Instituto Superior de Engenharia de Lisboa

Álgebra Linear e Geometria Analítica

Espaços vectoriais
Matrizes
Funções lineares
Determinantes

Volume 1

Licenciado em Engenharia Electrotécnica pelo IST

Versão 3.7, Março de 2009

Instituto Superior de Engenharia de Lisboa

ÁLGEBRA LINEAR E GEOMETRIA ANALÍTICA

Espaços vectoriais

Matrizes

Funções lineares

Determinantes

Volume 1

ISEL - DEETC

Por Engº Carlos M. Ribeiro

Licenciado em Engenharia Electrotécnica pelo IST

Versão 3.7, Março de 2009

Álgebra Linear e Geometria Analítica, Volume 1

Licenciado em Engenharia Electrotécnica pelo IST
E-mail: cmribeiro@deetc.isel.ipl.pt
Web: http://www.deetc.isel.ipl.pt/paginaspessoais/carlosribeiro
Versão 3.7, Março de 2009

Conteúdo

Lista de figuras

Capítulo 1

Capítulo 2

Capítulo 3

Capítulo 4

Simbologia

Lógica

$\neg p$ .... negação (not) da proposição p
$p \vee q$ .... disjunção (or) das proposições p e q
$p \dot{\vee} q$ .... disjunção exclusiva (xor) das proposições p e q
$p \wedge q$ .... conjunção (and) das proposições p e q
$p \Rightarrow q$ .... p implica q
$p \equiv q, p \Leftrightarrow q$ .... equivalência das proposições p e q
sse .... se e só se
$\forall$ .... quantificador universal (qualquer que seja...)
$\exists$ .... quantificador existencial (existe pelo menos um...)
$\exists^1$ .... quantificador existencial exclusivo (existe um e um só...)
q.e.d. .... quod erat demonstrandum

Conjuntos

$\{a, b, c, \dots\}$ .... conjunto formado por $a, b, c, \dots$
$\{x \in X \colon \mathrm{p}(x)\}$ .... conjunto dos elementos de X com a propriedade $\mathrm{p}(x)$
$\emptyset, \{\}$ .... conjunto vazio
$\mathscr{P}(X)$ .... conjunto das partes (subconjuntos) do conjunto X
$A^{\mathrm{c}}, \overline{A}$ .... complementar do conjunto A
$A \setminus B$ .... diferença entre os conjuntos A e B
$A \cup B$ .... reunião dos conjuntos A e B
$A \cap B$ .... intersecção dos conjuntos A e B
(a, b) .... par ordenado formado pelos objectos a e b
$A \times B$ .... produto cartesiano dos conjuntos A e B
$(x_1, x_2, \dots, x_n), (x_k)_{1 \le k \le n}$ .... lista de comprimento n de elementos de um conjunto X
$\emptyset, ()$ .... lista vazia
$\mathrm{pr}_k\, x$ .... kª projecção ou componente da lista x
X^n .... conjunto das sequências de n elementos de X
$(x_k)_{k \in I}$ .... família de elementos de um conjunto X indexada por I
$\mathrm{pr}_i\, x$ .... projecção ou componente de índice i da família x (x_i)
X^I .... conjunto das famílias de elementos de X indexadas por I
X/r .... conjunto quociente do conjunto X pela relação de equivalência r
$\bigcup_{i \in I} A_i$ .... reunião duma família de conjuntos
$\bigcap_{i \in L} A_i$ .... intersecção duma família de conjuntos

Relações binárias

$=$ .... igual
$\neq$ .... diferente
$\in$ .... pertence a, é elemento de
$\notin$ .... não pertence a, não é elemento de
$\subset$ .... está contido em, é uma parte de
$\supset$ .... contém, é sobreconjunto de
$\not\subset$ .... não está contido em, não é parte de

Funções

$f(x)$.. valor da função f no ponto x
Y^X.. conjunto das funções de X em Y
$f\colon X \to Y, X \xrightarrow{f} Y$........................ f é função de X em Y
$x \mapsto y, x \mapsto f(x), x \overset{f}{\mapsto} y$.............. x é aplicado por f em y ou $f(x)$
$\mathscr{I}_X$.. relação (função) identidade no conjunto X
$f(A)$.. imagem directa do conjunto A por f
$f^{-1}(A)$.. pré-imagem do conjunto A por f
$f^{-1}(y)$.. pré-imagem do conjunto singular $\{y\}$ ou traço de y por f
f^{-1}.. função inversa da função injectiva f
$g \circ f$.. composta das funções g e f (g após f)
$\lim\limits_{x \to a} f(x)$.. limite da função f no ponto a
$f'(a), f''(a), f^{(k)}(a)$........................ derivada de 1ª ordem, 2ª ordem, ordem k de f em a
$f', f'', f^{(k)}$.. função derivada de 1ª ordem, 2ª ordem, ordem k de f
$\ln x, \log_a x$.. logaritmo neperiano de $x > 0$, logaritmo de x na base a
$e^x, \exp x$.. função exponencial de base e
a^x.. função exponencial de base $a > 0$
sin, cos.. funções seno e coseno
arcsin, arccos.................................. funções arcseno e arccoseno
$\int_a^b f(x)\,\mathrm{d}x$.. integral definido da função f entre os pontos a e b
$\int_I f(x)\,\mathrm{d}x$.. integral definido da função f estendido ao intervalo I

Estruturas

$x + y$.. soma dos elementos x e y de um grupóide aditivo
$x \times y, x \cdot y, x\,y$.............................. produto dos elementos x e y de um grupóide multiplicativo
0.. elemento neutro (zero) de um grupóide aditivo
1.. elemento neutro (um, unidade ou identidade) de um grupóide multiplicativo
$-x$.. oposto (simétrico) do elemento regular x, num monóide aditivo
$x^{-1}, 1/x$.. oposto (inverso) do elemento regular x, num monóide multiplicativo
$\sum\limits_{k=1}^{n} x_k$.. soma da lista $(x_i)_{1 \le i \le n}$ de elementos de um semigrupo comutativo aditivo

$\prod_{k=1}^{n} x_k$ .... produto da lista $(x_i)_{1\leq i\leq n}$ de elementos de um semigrupo comutativo multiplicativo
$A \cong B$ .... A é isomorfo de B

Números Inteiros e Racionais

$\mathbb{N}_0$ .... conjunto dos números naturais com zero $\{0, 1, 2, \ldots, n, \ldots\}$
$\mathbb{N}, \mathbb{Z}^+$ .... conjunto dos números naturais $\{1, 2, \ldots, n, \ldots\}$
$[m, n]$ .... conjunto dos inteiros entre m e n inclusivé
$[1, n]$ .... intervalo $\{1, 2, \ldots, n\} \subset \mathbb{N}$
$\mathbb{Z}$ .... conjunto dos inteiros
$\mathbb{Z}^-$ .... conjunto dos inteiros < 0
$\mathbb{Z}^*$ .... conjunto dos inteiros $\neq 0$
$\mathbb{Z}_p$ .... anel dos inteiros módulo p (corpo, se p é primo)
$\mathbb{Q}$ .... conjunto dos números racionais
$\mathbb{Q}^+$ .... conjunto dos racionais > 0
$\mathbb{Q}_0^+$ .... conjunto dos racionais ≥ 0
$\mathbb{Q}^-$ .... conjunto dos racionais < 0
$\mathbb{Q}_0^-$ .... conjunto dos racionais ≤ 0
$\mathbb{Q}^*$ .... conjunto dos racionais $\neq 0$

Combinatória

$n!$ .... factorial de n
P_n, A_n^n .... permutações de n
$\binom{n}{p}$ .... combinações de n elementos tomados p a p (coeficientes binomiais)
A_p^n .... arranjos de n elementos tomados p a p

Números reais

$\mathbb{R}$ .... conjunto dos números reais
$\mathbb{R}^+$ .... conjunto dos números reais > 0
$\mathbb{R}_0^+$ .... conjunto dos números reais ≥ 0
$\mathbb{R}^-$ .... conjunto dos números reais < 0
$\mathbb{R}_0^-$ .... conjunto dos números reais ≤ 0
$\mathbb{R}^*$ .... conjunto dos números reais $\neq 0$
$\overline{\mathbb{R}}$ .... recta acabada
$\dot{\mathbb{R}}$ .... recta projectiva
π .... pi, razão entre o perímetro e o diâmetro de qualquer circunferência
e .... número de Neper, base da função exponencial $x \mapsto \exp(x)$
$x \leq y, y \geq x$ .... relação de ordem em $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ ou $\mathbb{R}$
$x < y, y > x$ .... relação de ordem estrita em $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ ou $\mathbb{R}$
$(a, b), [a, b], (a, b], [a, b)$ .... intervalos limitados em $\mathbb{R}$ ou num espaço ordenado

$(-\infty, b), (-\infty, b]$ .... intervalos não limitados inferiormente em $\mathbb{R}$ ou num espaço ordenado (secções inferiores)
$(a, +\infty), [a, +\infty)$ .... intervalos não limitados superiormente em $\mathbb{R}$ ou num espaço ordenado (secções superiores)
$a \approx b$ .... a aproximadamente igual a b
$|x|$ .... módulo do real x
$\max X$ .... máximo do conjunto X
$\min X$ .... mínimo do conjunto X
$\sup X$ .... supremo do conjunto X
$\inf X$ .... ínfimo do conjunto X
$-\infty, +\infty$ .... menos infinito, mais infinito em $\overline{\mathbb{R}}$

Números complexos

$\mathbb{C}$ .... conjunto dos números complexos
$\mathbb{C}^*$ .... conjunto dos números complexos $\neq 0$
$i\mathbb{R}$ .... conjunto dos imaginários puros
$x + iy$ .... forma algébrica de um complexo
$r\text{cis}(\theta)$ .... forma trigonométrica de um complexo
$re^{i\theta}$ .... forma exponencial de um complexo
$\overline{z}$ .... conjugado do complexo z
$A + B$ .... soma dos subconjuntos A e B de $\mathbb{C}$
zA .... produto do complexo z pelo subconjunto $A \subset \mathbb{C}$
$\Re(z), \Im(z)$ .... parte real do complexo z, parte imaginária do complexo z
$|z|$ .... módulo ou valor absoluto do complexo z
$\arg(z)$ .... argumento principal do complexo z
$\text{Arg}(z)$ .... conjunto de todos os argumentos do complexo z
$e^z, \exp(z)$ .... função exponencial complexa
$\exp^{-1}(z)$ .... conjunto de todos os logaritmos do complexo $z \neq 0$
$\ln(z)$ .... valor principal do logaritmo do complexo $z \neq 0$

Quaterniões

$\mathbb{H}$ .... conjunto dos quaterniões de Hamilton
$\mathbb{H}^*$ .... conjunto dos quaterniões não nulos
$\mathbb{H}_1$ .... conjunto dos quaterniões unitários
$\mathbb{H}_3$ .... conjunto dos quaterniões puros
$*$ .... multiplicação de quaterniões

Espaços vectoriais

$\alpha\vec{x}$ .... produto de escalar por vector
$\vec{x} + \vec{y}$ .... soma de vectores
$\vec{x}\,\vec{y}$ .... produto de vectores (numa álgebra linear)
$\vec{x}, \vec{0}, \vec{u}$ .... vector, vector nulo, vector unidade (numa álgebra linear com unidade)
$\sum_{i=1}^{n} \alpha_i \vec{x}_i$ .... combinação linear dos vectores da sequência $(\vec{x}_i)_{1 \leq i \leq n}$

$\sum_{i\in\emptyset}\alpha_i\vec{x}_i=\vec{o}$ .... combinação linear vazia
$\mathrm{pr}_k\,\vec{x}$ .... kª projecção ou componente do vector $\vec{x}\in\mathbb{K}^n$
$F\preceq E$ .... F é subespaço vectorial de E
$\prod_{i\in I}E_i$ .... produto cartesiano dos espaços vectoriais da família $(E_i)_{i\in I}$ de espaços vectoriais sobre um mesmo corpo $\mathbb{K}$
$\prod_{i=1}^{n}E_i$ ou $E_1\times E_2\times\cdots\times E_n$ .... produto cartesiano dos espaços vectoriais da sequência $(E_i)_{1\le i\le n}$ de espaços sobre um mesmo corpo $\mathbb{K}$
$\bigoplus_{i\in I}E_i$ .... soma directa externa da família $(E_i)_{i\in I}$ de espaços vectoriais sobre um mesmo corpo $\mathbb{K}$
$\bigoplus_{i\in I}E_i$ .... soma directa da família $(E_i)_{i\in I}$ de subespaços
$\sum_{i\in I}E_i$ .... soma da família $(E_i)_{i\in I}$ de subespaços
$A+B$ .... soma dos subespaços A e B
$A\oplus B$ .... soma directa dos subespaços A e B
$E^{(I)}$ .... conjunto das famílias $(\vec{x}_i)_{i\in I}$ de vectores de E indexadas por I e tais que $\vec{x}_i=\vec{0}$ no complementar de uma parte finita de I.
E/F .... espaço vectorial quociente de E pelo subespaço F
S^1,S^2,S^3 .... espaços de segmentos orientados com origem num ponto O
$\overrightarrow{AB}\sim\overrightarrow{CD}$ .... equipolência entre os segmentos orientados $\overrightarrow{AB}$ e $\overrightarrow{CD}$
$\widetilde{S}^1,\widetilde{S}^2,\widetilde{S}^3$ .... conjunto dos segmentos aplicados em qualquer ponto
$\widehat{S}^1,\widehat{S}^2,\widehat{S}^3$ .... espaços vectoriais dos vectores livres
$\mathfrak{C}(\mathbb{N},\mathbb{R})$ .... espaço das sucessões reais convergentes
$\mathfrak{C}_0(\mathbb{N},\mathbb{R})$ .... espaço das sucessões reais convergentes para 0
$l^2_{\mathbb{R}}$ .... espaço de Hilbert de sucessões reais
$\mathscr{B}(I,\mathbb{R})$ .... espaço das funções reais limitadas no conjunto I
$\mathscr{D}(I,\mathbb{R})$ .... espaço das funções reais diferenciáveis em I
$\mathscr{C}(I,\mathbb{R})$ .... espaço das funções reais contínuas no conjunto I
$\mathscr{C}^k(I,\mathbb{R})$ .... espaço das funções reais de classe C^k no conjunto I
$\mathscr{C}^\infty(I,\mathbb{R})$ .... espaço das funções reais de classe C^∞ no conjunto I
$\mathfrak{C}(\mathbb{N},\mathbb{C})$ .... espaço das sucessões complexas convergentes
$\mathfrak{C}_0(\mathbb{N},\mathbb{C})$ .... espaço das sucessões complexas convergentes para 0
$l^2_{\mathbb{C}}$ .... espaço de Hilbert de sucessões complexas
$\mathscr{B}(I,\mathbb{C})$ .... espaço das funções complexas limitadas no conjunto I
$\mathscr{D}(I,\mathbb{C})$ .... espaço das funções complexas diferenciáveis em I
$\mathscr{C}(I,\mathbb{C})$ .... espaço das funções complexas contínuas no conjunto I
$\mathscr{C}^k(I,\mathbb{C})$ .... espaço das funções complexas de classe C^k no conjunto I
$\mathscr{C}^\infty(I,\mathbb{C})$ .... espaço das funções complexas de classe C^∞ no conjunto I
$\mathscr{P}_n(\mathbb{K})$ .... espaço dos polinómios de grau $\le n$ e de coeficientes no corpo $\mathbb{K}$
$\mathscr{P}(\mathbb{K})$ .... álgebra dos polinómios de coeficientes no corpo $\mathbb{K}$

$L_{\mathbb{K}}(x), L_{\mathbb{K}}(\vec{x}_i)_{1\leq i\leq m}$........................ subespaço gerado pela lista $x = (\vec{x}_i)_{1\leq i\leq m}$ de vectores de um espaço vectorial E sobre o corpo $\mathbb{K}$
$L_{\mathbb{K}}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$........................ subespaço gerado pela lista $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ de vectores de um espaço vectorial E sobre o corpo $\mathbb{K}$
$< \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m >$ .... subespaço gerado pela lista $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ de vectores de um espaço vectorial E
$\dim_{\mathbb{K}} E, \dim E$................................ dimensão do espaço vectorial E sobre o corpo $\mathbb{K}$
$E \cong F$... E é isomorfo de F

Matrizes

$[a_{ij}]_{\substack{1\leq i\leq m\\1\leq j\leq n}}$.. matriz de tipo $m \times n$ de elementos a_{ij} de um conjunto X
$[a_{ij}]_{1\leq i,j\leq n}$...................................... matriz quadrada de ordem n de elementos a_{ij} de um conjunto X
$\mathbb{K}^{m,n}$.. espaço vectorial das matrizes do tipo $m \times n$ de elementos no corpo $\mathbb{K}$
$\mathbb{K}^{n,n}$.. álgebra das matrizes quadradas de ordem n de elementos no corpo $\mathbb{K}$
$O_{m,n}$... matriz nula do tipo $m \times n$
O_n.. matriz nula de ordem n
I_n.. matriz identidade de ordem n
$A \pm B$.. soma (diferença) das matrizes A e B
αA.. produto do escalar α pela matriz A
A^{T}, A^t, A'....................................... matriz transposta de A
$\overline{A}$.. matriz conjugada de A
A^*.. matriz transconjugada de A
AB... produto das matrizes A e B
$\mathrm{c}(A)$.. característica da matriz A
$\mathrm{tr}(A)$.. traço da matriz quadrada A
$\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$........................ matriz diagonal de elementos diagonais $\lambda_1, \lambda_2, \ldots, \lambda_n$
$T_{ee'}$.. matriz de mudança da base e para a base e'
$A[i_1, \ldots, i_p; j_1, \ldots, j_q]$.................... Submatriz de A, obtida seleccionando as linhas $i_1, \ldots, i_p$ e as colunas $j_1, \ldots, j_q$ de A
$A[\{i_1, \ldots, i_p\}; \{j_1, \ldots, j_q\}]$............ Submatriz de A, obtida seleccionando as linhas $i_1, \ldots, i_p$ e as colunas $j_1, \ldots, j_q$ de A
$A[I; J]$... Submatriz de A, obtida seleccionando as linhas cujos índices pertencem aos conjuntos I e J
$A(i_1, \ldots, i_p; j_1, \ldots, j_q)$.................. Submatriz de A, obtida eliminando as linhas $i_1, \ldots, i_p$ e as colunas $j_1, \ldots, j_q$ de A
$A[i_1, \ldots, i_p; j_1, \ldots, j_q)$.................. Submatriz de A, obtida seleccionando as linhas $i_1, \ldots, i_p$ e eliminando as colunas $j_1, \ldots, j_q$ de A
$A(i_1, \ldots, i_p; j_1, \ldots, j_q]$.................. Submatriz de A, obtida eliminando as linhas $i_1, \ldots, i_p$ e seleccionando as colunas $j_1, \ldots, j_q$ de A

Funções lineares

$\mathrm{Hom}(E,F), \mathscr{L}(E,F)$ .... espaço vectorial das funções lineares de E em F
$\mathrm{End}(E)$ .... álgebra dos endomorfismos de E
$\mathrm{Ker}(h)$, $\mathrm{Nuc}(h)$ .... núcleo da aplicação linear h
$\mathrm{Img}(h)$, $\mathrm{Im}(h)$, $h(E)$ .... imagem da aplicação linear h definida em E
$\mathrm{c}(h)$, c_h .... característica da aplicação linear h
$\mathrm{n}(h)$, n_h .... nulidade da aplicação linear h
$\mathscr{O}_{FE}$ .... função (linear) nula do espaço E no espaço F
$\mathscr{O}_E$ .... endomorfismo nulo no espaço E
$\det(h)$ .... determinante do endomorfismo h
$M_{fe}(h)$ .... representação matricial da função linear h nas bases f e e
$M_e(h)$ .... representação matricial do endomorfismo h na base e

Funções multilineares e determinantes

$\mathfrak{S}_n$ .... grupo simétrico de ordem n
$\mathscr{I}_n$ .... permutação identidade de ordem n
σ, τ .... permutação, transposição
σ^{-1} .... permutação inversa de σ
$\varepsilon(\sigma)$ .... sinal ou paridade da permutação σ.
$\det(A[i_1,\ldots,i_p;j_1,\ldots,j_p])$ .... menor de ordem p da matriz A
$\det(A(i_1,\ldots,i_p;j_1,\ldots,j_p))$ .... menor complementar do anterior
$\mathrm{cof}(\det(A[i_1,\ldots,i_p;j_1,\ldots,j_p]))$ .... cofactor do menor $\det(A[i_1,\ldots,i_p;j_1,\ldots,j_p])$
$\mathrm{cof}(a_{ij}), \mathrm{cof}a_{ij}$ .... cofactor de a_{ij}
$\widehat{A}$ .... matriz complementar de A (matriz dos cofactores dos elementos de A)
$\mathrm{adj}A$ .... matriz adjunta de A (transposta da anterior)
$\det_e(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n)$ .... determinante dos vectores $(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n)$ na base e
$\det(u)$ .... determinante do endomorfismo u
$\det(A), |A|$ .... determinante da matriz A
$\mathscr{F}(E^p,F)$ .... espaço de todas as funções de E^p em F
$\mathscr{M}_p(E,F)$ .... espaço das funções p-lineares sobre E com valores em F
$\mathscr{A}_p(E,F)$ .... espaço das funções p-lineares alternadas sobre E com valores em F
$\mathscr{S}_p(E,F)$ .... espaço das funções p-lineares simétricas sobre E com valores em F
$\mathscr{O}_{FE^p}$ .... função p-linear nula de E^p em F
$\mathscr{O}_{\mathbb{K}E^p}$ .... forma p-linear nula de E^p em $\mathbb{K}$

Valores e vectores próprios

$m_g(\lambda)$ .... multiplicidade geométrica do valor próprio λ
$m_a(\lambda)$ .... multiplicidade algébrica do valor próprio λ
$E_\lambda(h)$ .... Subespaço próprio do endomorfismo h associado ao valor próprio λ
$\mathscr{E}(h)$ .... espectro do endomorfismo h
$p_h(\lambda)$ (resp. $p_A(\lambda)$) .... polinómio característico do endomorfismo h (resp. da matriz A)

Espaços euclidianos

$\vec{x}\cdot\vec{y}, \langle\vec{x},\vec{y}\rangle, \vec{x}|\vec{y}$ .... produto interno dos vectores $\vec{x}$ e $\vec{y}$
$\|\vec{x}\|$ .... norma euclidiana ou hermitiana do vector $\vec{x}$
vers$\vec{x}$ .... versor do vector não nulo $\vec{x}$
proj $_{\vec{y}}\,\vec{x}$ .... projecção ortogonal de $\vec{x}$ sobre $\vec{y}\neq\vec{0}$
$\vec{x}\perp\vec{y}$ .... ortogonalidade entre vectores de um espaço com produto interno
$A^\perp$ .... complemento ortogonal de A
o.n. .... (base) *ortonormada*
o.n.d. .... (base) *ortonormada directa*
$G(x), G(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_m)$ .... determinante de Gram da lista $x=(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_m)$
G_x .... matriz de Gram da lista x
$\mathscr{B}(E)$ .... conjunto das bases de um espaço vectorial
$\mathscr{O}(E)$ .... conjunto das orientações de um espaço vectorial real
sgn .... sinal de uma base, num espaço euclidiano orientado
$[\![\vec{x}_1,\vec{x}_2,\dots,\vec{x}_n]\!]$ .... produto misto dos vectores $\vec{x}_1,\vec{x}_2,\dots,\vec{x}_n$
$\vec{x}_1\times\vec{x}_2\times\dots\times\vec{x}_{n-1}$ .... produto externo dos vectores $\vec{x}_1,\vec{x}_2,\dots,\vec{x}_{n-1}$
$\vec{x}_1\times\vec{x}_2$ .... produto externo dos vectores $\vec{x}_1$ e $\vec{x}_2$ num espaço euclidiano orientado de dimensão 3
$\mathscr{B}(E,F;\mathbb{K})$ .... espaço das formas bilineares sobre $E\times F$
$\mathscr{B}(E;\mathbb{K})$ .... espaço das formas bilineares sobre E
$\mathscr{A}(E;\mathbb{K})$ .... espaço das formas bilineares autoadjuntas sobre E
$\mathscr{Q}(E;\mathbb{K})$ .... espaço das formas quadráticas sobre E
$\mathscr{S}(E,F;\mathbb{C})$ .... espaço das formas sesquilineares sobre $E\times F$
$\mathscr{S}(E;\mathbb{C})$ .... espaço das formas sesquilineares sobre E
$\mathscr{A}_\mathbb{C}(E,\mathbb{C})$ .... espaço das formas sesquilineares hermitianas sobre E
$\mathscr{Q}_\mathbb{C}(E;\mathbb{R})$ .... espaço das formas quadráticas hermitianas sobre E
A^+ .... matriz pseudoinversa de A

Geometria analítica

$\mathscr{E},\mathscr{F}$ .... Espaço afim, subespaço afim
$\overrightarrow{AB}$, $B-A$ .... vector de origem A e extremidade B
dim$\mathscr{E}$ .... dimensão do espaço afim $\mathscr{E}$
$\mathscr{R}^n$ .... espaços afins euclidianos reais orientados de dimensão n

$(O; e)$, $(O; \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$	referencial num espaço afim
o.n.	abreviatura de *ortonormado* (para referenciais)
o.n.d.	abreviatura de *ortonormado directo* (para referenciais)
$P(x_1, x_2, \ldots, x_n)$	ponto P de um espaço afim com coordenadas $(x_1, x_2, \ldots, x_n)$ num dado referencial
$P \equiv (x_1, x_2, \ldots, x_n)$	ponto P de um espaço afim com coordenadas $(x_1, x_2, \ldots, x_n)$ num dado referencial
$\vec{v}(v_1, v_2, \ldots, v_n)$	vector de coordenadas $(v_1, v_2, \ldots, v_n)$ numa dada base
$\langle A; F\rangle$	Subespaço afim gerado pelo ponto A e pelo subespaço vectorial F

1

Espaços vectoriais

1.1 Introdução

Neste capítulo, faremos o estudo de uma importante estrutura algébrica muito utilizada nas ciências aplicadas e na engenharia: o *Espaço Vectorial*. Como pré-requisitos para o presente capítulo, mencionemos noções elementares sobre teoria dos conjuntos, lógica e estruturas algébricas (consultar o apêndice B); algum conhecimento dos números reais e complexos é também conveniente. Será útil a leitura prévia do apêndice B, no qual se definem as noções de *operação unária*, *binária* e n-*ária*; os conceitos de lei de *composição interna* e *externa*; as estruturas algébricas: *grupóides*, *semigrupos*, *monóides*, *grupos*, *anéis*, *anéis de integridade* e *corpos comutativos* (casos de $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$). Assume ainda relevo a noção de *homomorfismo* e, em particular, a de *isomorfismo*. Em anexo a este capítulo, apresenta-se informação relativa à forma como se pode utilizar o software MATHEMATICA© para operar com vectores.

1.2 Axiomática dos espaços vectoriais

Em tudo o que se segue neste manual, $\mathbb{K}$ designa um *corpo comutativo*, que designaremos simplesmente por *corpo*. Para a definição de Espaço Vectorial, seguiremos a via axiomática. As noções primitivas são as de *escalar* (elemento de um corpo comutativo $\mathbb{K}$) e de *vector* (elemento de um segundo conjunto E, eventualmente igual a $\mathbb{K}$). Para distinguir claramente os escalares dos vectores, nestes últimos usaremos uma seta superior (por exemplo, $\vec{x}$). Assim, poremos a seguinte

Definição 1.1 – Espaço vectorial – *Seja E um conjunto não vazio e $(\mathbb{K}, +, \times)$ um corpo comutativo (ver apêndice* B*). Diz-se que E (cujos elementos serão chamados* **vectores***) é um* **espaço vectorial** *(ou* **linear***) sobre o corpo $\mathbb{K}$ (cujos elementos serão chamados* **escalares***), sse E estiver munido de uma lei de composição interna $+: E^2 \to E; (\vec{x}, \vec{y}) \mapsto \vec{x} + \vec{y}$*[1] *(adição vectorial) e de uma lei de composição externa $\cdot: \mathbb{K} \times E \to E; (\alpha, \vec{x}) \mapsto \alpha\vec{x}$ (produto de escalar por vector ou multiplicação escalar, cujo resultado designaremos simplesmente por $\alpha\vec{x}$ em vez de $\alpha \cdot \vec{x}$) satisfazendo os seguintes oito axiomas:*

[A1] $\underset{\vec{x},\vec{y}\in E}{\forall}\ \vec{x} + \vec{y} = \vec{y} + \vec{x}$

[A2] $\underset{\vec{x},\vec{y},\vec{z}\in E}{\forall}\ (\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$

[A3] $\underset{\vec{o}\in E}{\exists}\ \underset{\vec{x}\in E}{\forall}\ \vec{x} + \vec{o} = \vec{o} + \vec{x} = \vec{x}$

[A4] $\underset{\vec{x}\in E}{\forall}\ \underset{-\vec{x}\in E}{\exists}\ \vec{x} + (-\vec{x}) = (-\vec{x}) + \vec{x} = \vec{o}$

[P1] $\underset{\substack{\alpha,\beta\in\mathbb{K} \\ \vec{x}\in E}}{\forall}\ (\alpha + \beta)\vec{x} = \alpha\vec{x} + \beta\vec{x}$

[P2] $\underset{\substack{\alpha\in\mathbb{K} \\ \vec{x},\vec{y}\in E}}{\forall}\ \alpha(\vec{x} + \vec{y}) = \alpha\vec{x} + \alpha\vec{y}$

---

1 Não existe ambiguidade no uso do sinal + para as adições em $\mathbb{K}$ e em E: o contexto algébrico em que estes símbolos ocorrem determina facilmente qual a operação em questão.

[P3] $\underset{\substack{\alpha,\beta\in\mathbb{K}\\ \vec{x}\in E}}{\forall}\ \alpha(\beta\vec{x}) = (\alpha\beta)\vec{x}$ [2]

[P4] $\underset{\vec{x}\in E}{\forall}\ 1\vec{x} = \vec{x}$

Os primeiros quatro axiomas dizem exclusivamente respeito à adição vectorial e são, por ordem, a *comutatividade*, a *associatividade*, a existência de *elemento neutro* $\vec{o}$ e a existência de *simétrico* para cada vector de E e significam que $(E, +)$ é um *Grupo Comutativo*. Como em qualquer grupo aditivo, o elemento neutro $\vec{o}$ (*vector nulo*, também designado por $\vec{o}_E$ se houver necessidade de salientar qual é o espaço vectorial) é único, o mesmo se dizendo do *simétrico* $-\vec{x}$ de cada vector $\vec{x} \in E$. Como em qualquer grupo comutativo aditivo $(E, +)$, definiremos a *subtracção vectorial* por

$$\vec{x} - \vec{y} = \vec{x} + (-\vec{y}) \tag{1.1}$$

Os axiomas P1-P4 caracterizam o comportamento do produto de escalar por vector relativamente à adição escalar, à adição vectorial e ao produto de escalares. Os dois primeiros são, por ordem, a *distributividade* em relação à adição escalar e a *distributividade* em relação à adição vectorial. O terceiro é a *associatividade mista*. O último exige que a *unidade* do corpo $\mathbb{K}$ seja elemento neutro à esquerda na multiplicação escalar.

Normalmente, chama-se espaço vectorial ao conjunto E mas faz-se notar que, formalmente, o espaço vectorial é a estrutura algébrica formada pelo conjunto E, o corpo $\mathbb{K}$ e as duas operações (adição vectorial e produto de escalar por vector) satisfazendo os axiomas A1-A4 e P1-P4, ou seja, o quaterno $(E, \mathbb{K}, +, \cdot)$. Dois espaços vectoriais $(E, \mathbb{K}, +, \cdot)$ e $(E', \mathbb{K}', \oplus, \odot)$ são, pois, iguais sse $E = E' \wedge \mathbb{K} = \mathbb{K}' \wedge + = \oplus \wedge \cdot = \odot$.

Adjectiva-se, por vezes, o espaço vectorial segundo o corpo subjacente: assim, diremos que E é um espaço vectorial *racional* se $\mathbb{K} = \mathbb{Q}$, um espaço vectorial *real* se $\mathbb{K} = \mathbb{R}$ e um espaço vectorial *complexo* se $\mathbb{K} = \mathbb{C}$.

A definição seguinte carateriza a noção de *Álgebra Linear*, juntando às operações definidas num espaço vectorial uma segunda lei de composição interna em E, chamada *multiplicação vectorial*.

Definição 1.2 – Álgebra linear – *Seja $(E, \mathbb{K}, +, \cdot)$ um espaço vectorial sobre $\mathbb{K}$, com operações $+$ e $\cdot$. Diz-se que E é uma **álgebra linear** sobre $\mathbb{K}$ se estiver definida uma segunda lei de composição interna $*\colon E^2 \to E; (\vec{x}, \vec{y}) \mapsto \vec{x}*\vec{y}$ (multiplicação vectorial) satisfazendo os axiomas abaixo indicados:*

[M1] $\underset{\vec{x},\vec{y},\vec{z}\in E}{\forall}\ (\vec{x}+\vec{y})*\vec{z} = \vec{x}*\vec{z} + \vec{y}*\vec{z}$

[M2] $\underset{\vec{x},\vec{y},\vec{z}\in E}{\forall}\ \vec{x}*(\vec{y}+\vec{z}) = \vec{x}*\vec{y} + \vec{x}*\vec{z}$

[M3] $\underset{\substack{\alpha\in\mathbb{K}\\ \vec{x},\vec{y}\in E}}{\forall}\ (\alpha\vec{x})*\vec{y} = \vec{x}*(\alpha\vec{y}) = \alpha(\vec{x}*\vec{y})$

---

2 Segundo o que é habitual, designamos por $\alpha\beta$ (e não $\alpha \times \beta$) o produto dos escalares α e β, no corpo $\mathbb{K}$.

A álgebra linear diz-se *associativa*, se a multiplicação $*$ vectorial o for:

[M4] $\underset{\vec{x},\vec{y},\vec{z}\in E}{\forall} (\vec{x}*\vec{y})*\vec{z} = \vec{x}*(\vec{y}*\vec{z})$,

M1/M2 são as distributividades à direita e à esquerda da multiplicação em relação à adição vectorial. M3 caracteriza o comportamento da multiplicação vectorial em relação à multiplicação de escalar por vector. M1, M2 e M4 significam que $(E, +, *)$ constitui um anel. Por satisfazer M1, M2 e M3, diz-se que a multiplicação vectorial é linear em cada um dos factores ou que é uma função bilinear sobre E com valores em E (ver capítulo 4). A álgebra diz-se *real* ou *complexa* conforme $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Se for $\vec{x}*\vec{y} = \vec{y}*\vec{x}$, para quaisquer $\vec{x}, \vec{y} \in E$, diremos que a álgebra (e o anel $(E, +, *)$) é *comutativa* e se $*$ tiver elemento neutro $\vec{u} \in E$ tal que $\vec{x}*\vec{u} = \vec{u}*\vec{x} = \vec{x}$, para qualquer $\vec{x} \in E$, diremos que a álgebra tem *unidade* ou *identidade*. Por exemplo, $\mathbb{C}$ é um espaço vectorial sobre $\mathbb{R}$ e uma álgebra linear comutativa com unidade 1 sobre $\mathbb{R}$, se considerarmos como multiplicação vectorial $*$ a multiplicação de complexos. O apêndice D (*quaterniões e rotações 3D*) fornece um exemplo de álgebra linear $\mathbb{H}$ não comutativa com elemento neutro e onde todos os elementos não nulos têm inverso (corpo não comutativo).

1.3 Consequências algébricas dos axiomas

Os axiomas enunciados atrás têm um certo número de consequências algébricas, de entre as quais se salientam as proposições a seguir demonstradas. Observe-se que estas proposições são verificadas em qualquer espaço vectorial, uma vez que elas só dependem dos referidos axiomas.

Proposição 1.1 *Em qualquer espaço vectorial E sobre um corpo $\mathbb{K}$, o produto de um escalar por um vector é nulo sse um dos factores for nulo*

$$\alpha\vec{x} = \vec{o} \Leftrightarrow \alpha = 0 \vee \vec{x} = \vec{o} \tag{1.2}$$

Demonstração:

Para qualquer $\alpha \in \mathbb{K}$ e qualquer $\vec{x} \in E$, tem-se, sucessivamente,

$$\vec{o} + \alpha\vec{x} = \alpha\vec{x} = (0 + \alpha)\vec{x} = 0\vec{x} + \alpha\vec{x}$$

Pela lei do corte, segue-se que $0\vec{x} = \vec{o}$. Do mesmo modo, para qualquer $\alpha \in \mathbb{K}$ e qualquer $\vec{x} \in E$, tem-se, sucessivamente,

$$\vec{o} + \alpha\vec{x} = \alpha\vec{x} = \alpha(\vec{o} + \vec{x}) = \alpha\vec{o} + \alpha\vec{x}$$

De novo, da lei do corte resulta $\alpha\vec{o} = \vec{o}$. Portanto, se $\alpha = 0$ ou $\vec{x} = \vec{o}$, teremos $\alpha\vec{x} = \vec{o}$.

A implicação recíproca $\alpha\vec{x} = \vec{o} \Rightarrow \alpha = 0 \vee \vec{x} = \vec{o}$ é equivalente a

$$\alpha\vec{x} = \vec{o} \wedge \alpha \neq 0 \Rightarrow \vec{x} = \vec{o}$$

Provemos a implicação anterior: sendo, $\alpha \neq 0$ existe $\alpha^{-1} \in \mathbb{K}$ tal que $\alpha^{-1}\alpha = 1$. Calculando, finalmente, $\vec{x}$ vem

$$\vec{x} = 1\vec{x} = \left(\alpha^{-1}\alpha\right)\vec{x} = \alpha^{-1}(\alpha\vec{x}) = \alpha^{-1}\vec{o} = \vec{o} \qquad \square$$

Proposição 1.2 *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$. Então, para qualquer escalar $\alpha \in \mathbb{K}$ e qualquer vector $\vec{x} \in E$, tem-se*

$$-(\alpha\vec{x}) = (-\alpha)\vec{x} = \alpha(-\vec{x}) \tag{1.3}$$

Demonstração:

Tem-se

$$\alpha\vec{x} + (-\alpha)\vec{x} = (\alpha + (-\alpha))\vec{x} = 0\vec{x} = \vec{o}$$

Da mesma forma, tem-se

$$\alpha\vec{x} + \alpha(-\vec{x}) = \alpha(\vec{x} + (-\vec{x})) = \alpha\vec{o} = \vec{o}$$

o que termina a demonstração. $\square$

Proposição 1.3 *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$. Então, para quaisquer escalares $\alpha, \beta \in \mathbb{K}$ e quaisquer vectores $\vec{x}, \vec{y} \in E$, tem-se*

$$(\alpha - \beta)\vec{x} = \alpha\vec{x} - \beta\vec{x} \tag{1.4.1}$$

$$\alpha(\vec{x} - \vec{y}) = \alpha\vec{x} - \alpha\vec{y} \tag{1.4.2}$$

Demonstração:

Tem-se, sucessivamente, por definição de subtracção e pela proposição anterior:

$$\begin{aligned}(\alpha - \beta)\vec{x} &= (\alpha + (-\beta))\vec{x} = \alpha\vec{x} + (-\beta)\vec{x} = \alpha\vec{x} + (-(\beta\vec{x})) = \alpha\vec{x} - \beta\vec{x}\\ \alpha(\vec{x} - \vec{y}) &= \alpha(\vec{x} + (-\vec{y})) = \alpha\vec{x} + \alpha(-\vec{y}) = \alpha\vec{x} + (-(\alpha\vec{y})) = \alpha\vec{x} - \alpha\vec{y}\end{aligned} \qquad \square$$

1.4 Exemplos de espaços vectoriais

Exemplo 1.1 O axioma A3 garante que $E \neq \emptyset$, visto que $\vec{o} \in E$. O presente exemplo vai mostrar que não é necessário mais qualquer vector em E para se obter um espaço vectorial. Sendo $\mathbb{K}$ um corpo e $E = \{\vec{o}\}$ um conjunto singular,[3] E constitui um espaço vectorial sobre $\mathbb{K}$, com as operações definidas, para qualquer $\alpha \in \mathbb{K}$, por

[A] $\quad \vec{o} + \vec{o} = \vec{o}$ (1.5)

[P] $\quad \alpha\vec{o} = \vec{o}$ (1.6)

---

3 De facto, não importa a natureza do único elemento de E. Estamos, simplesmente, a usar a notação $\vec{o}$ para designar esse elemento.

Exemplo 1.2 Fazendo $E = \mathbb{K}$, obtém-se com a adição e o produto de $\mathbb{K}$ um espaço vectorial sobre $\mathbb{K}$. O *vector nulo* é o zero 0 de $\mathbb{K}$ e o simétrico de $x \in \mathbb{K}$ é $-x$. Observe que, neste caso, os vectores confundem-se com os escalares.

Exemplo 1.3 Seja $\mathbb{K}$ um corpo, $n \in \mathbb{N}$ um inteiro positivo e $E = \mathbb{K}^n$. Então, o conjunto $\mathbb{K}^n$ de todas as sequências (listas) de comprimento n de escalares de $\mathbb{K}$ é um espaço vectorial sobre $\mathbb{K}$ (chamado *espaço cartesiano* de dimensão n sobre o corpo $\mathbb{K}$), com as seguintes operações, para quaisquer $\alpha \in \mathbb{K}$, $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ e $(y_1, y_2, \ldots, y_n) \in \mathbb{K}^n$

[A] $$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \tag{1.7}$$

[P] $$\alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n) \tag{1.8}$$

Os escalares $x_1, x_2, \ldots, x_n$ são chamados as *componentes* ou *projecções* do vector $(x_1, x_2, \ldots, x_n)$, escrevendo-se, às vezes, $x_k = \mathrm{pr}_k\, \vec{x} = \mathrm{pr}_k(x_1, x_2, \ldots, x_n)$. Observe-se que, nestas definições, usam-se (nos segundos membros) as operações pré-existentes em $\mathbb{K}$ para definir as operações novas no espaço vectorial $\mathbb{K}^n$. Observe-se ainda que o *vector nulo* de $\mathbb{K}^n$ é $\vec{o} = (0, 0, \ldots, 0)$ e que o *simétrico* de $(x_1, x_2, \ldots, x_n)$ é $(-x_1, -x_2, \ldots, -x_n)$ ou seja,

$$\vec{o} = (0, 0, \ldots, 0) \tag{1.9}$$

$$-(x_1, x_2, \ldots, x_n) = (-x_1, -x_2, \ldots, -x_n) \tag{1.10}$$

Observe-se a generalidade deste exemplo, que tem uma infinidade de casos particulares como sejam, por exemplo, $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{C}$, $\mathbb{C}^2$, $\mathbb{C}^n$, etc.

Exemplo 1.4 Seja $\mathbb{K}$ um corpo e $E = \mathbb{K}^{\mathbb{N}}$ o conjunto das *sucessões* de escalares de $\mathbb{K}$. Então $\mathbb{K}^{\mathbb{N}}$ é um espaço vectorial sobre $\mathbb{K}$, com a adição usual de sucessões e o produto de escalar por sucessão definidos, para quaisquer $\alpha \in \mathbb{K}$, $(x_n) \in \mathbb{K}^{\mathbb{N}}$ e $(y_n) \in \mathbb{K}^{\mathbb{N}}$, por:

[A] $$(x_n) + (y_n) = (x_n + y_n) \tag{1.11}$$

[P] $$\alpha(x_n) = (\alpha x_n) \tag{1.12}$$

Observe-se que, nestas definições, novamente se utiliza a estrutura algébrica pré-existente em $\mathbb{K}$ para definir as operações do espaço vectorial. Observe-se ainda que o *vector nulo* de $\mathbb{K}^{\mathbb{N}}$ é a sucessão constante nula $0 = (0)$ e que o *simétrico* de uma sucessão (x_n) é a sucessão $(-x_n)$ ou seja,

$$0 = (0) \tag{1.13}$$

$$-(x_n) = (-x_n) \tag{1.14}$$

Como caso particular, citemos o espaço vectorial real $\mathbb{R}^{\mathbb{N}}$ das sucessões de números reais e o espaço complexo $\mathbb{C}^{\mathbb{N}}$ das sucessões de números complexos. É óbvio que poderíamos também indexar as sucessões usando $\mathbb{N}_0$ e definindo de igual modo as operações, obtendo-se o espaço $\mathbb{K}^{\mathbb{N}_0}$.

Exemplo 1.5 Seja $\mathbb{K}$ um corpo e $E = \mathbb{K}^{(\mathbb{N}_0)}$ o conjunto das sucessões de escalares de $\mathbb{K}$, cujos termos são todos nulos, com excepção de um número finito deles, ou seja[4]

$$\mathbb{K}^{(\mathbb{N}_0)} = \big\{(x_n) \in \mathbb{K}^{\mathbb{N}_0} \colon \{n \colon x_n \neq 0\} \text{ é finito}\big\} \subset \mathbb{K}^{\mathbb{N}_0}$$

Com as operações definidas por (1.11) e (1.12) obtém-se um novo espaço vectorial. O *vector nulo* e o *simétrico* de (x_n) são, como anteriormente,

$$0 = (0) \tag{1.15}$$

$$-(x_n) = (-x_n) \tag{1.16}$$

Exemplo 1.6 Seja $\mathbb{K}$ um corpo, A um conjunto e $E = \mathbb{K}^A$ o conjunto das *funções escalares* definidas em A, $\mathbb{K}^A = \{f\colon\ f\colon A \to \mathbb{K}\}$. Então $\mathbb{K}^A$ é um espaço vectorial sobre $\mathbb{K}$, se definirmos a soma de funções $f, g \in \mathbb{K}^A$ e o produto do escalar $\alpha \in \mathbb{K}$ pela função $f \in \mathbb{K}^A$ como sendo as funções $f + g$ e αf dadas, para todo o $x \in A$, por

[A] $\qquad (f + g)(x) = f(x) + g(x) \qquad (1.17)$

[P] $\qquad (\alpha f)(x) = \alpha f(x) \qquad (1.18)$

Observe-se que, novamente, se utiliza a estrutura algébrica pré-existente em $\mathbb{K}$ para definir as operações necessárias ao espaço vectorial $\mathbb{K}^A$. Observe-se ainda que o *vector nulo* de $\mathbb{K}^A$ é a função constante nula 0 e que o *simétrico* da função f é a função $-f$, estando estas definidas, para todo o $x \in A$, por:

$$0(x) = 0 \tag{1.19}$$

$$(-f)(x) = -f(x) \tag{1.20}$$

Observe-se que os exemplos 1.3 e 1.4 são casos particulares do presente exemplo, com $A = \{1, 2, \ldots, n\}$ e $A = \mathbb{N}$ ou $A = \mathbb{N}_0$, respectivamente. O próprio Exemplo 1.1 pode também ser visto à luz do caso $\mathbb{K}^A$, com $A = \emptyset$ e em que, portanto, E é o conjunto singular constituído apenas pela função vazia (de $\emptyset$ em $\mathbb{K}$).

Segundo este exemplo, poderemos falar do espaço vectorial real $\mathbb{R}^I$ das funções reais definidas no intervalo $I = [0, 1]$ ou do espaço vectorial complexo $\mathbb{C}^I$ das funções complexas definidas no mesmo intervalo (ou noutro qualquer conjunto).

Exemplo 1.7 Seja $\mathbb{K}$ um corpo, $n \in \mathbb{N}_0$ um inteiro não negativo e $E = \mathscr{P}_n(\mathbb{K})$ o conjunto dos polinómios de grau menor ou igual a n e com coeficientes em $\mathbb{K}$.

$$\mathscr{P}_n(\mathbb{K}) = \left\{p\colon p(x) = \sum_{k=0}^{n} a_k x^k, \text{com } a_k \in \mathbb{K}\right\} \tag{1.21}$$

---

4 Observe-se que este conjunto $\mathbb{K}^{(\mathbb{N}_0)}$ é o mesmo que o conjunto das sucessões de escalares, cujos termos se anulam, a partir de alguma ordem.

$\mathcal{P}_n(\mathbb{K})$ é um espaço vectorial sobre $\mathbb{K}$, se definirmos a soma dos polinómios $p, q \in \mathcal{P}_n(\mathbb{K})$, onde $p(x) = \sum_{k=0}^{n} a_k x^k$ e $q(x) = \sum_{k=0}^{n} b_k x^k$, e o produto do escalar $\alpha \in \mathbb{K}$ pelo polinómio $p \in \mathcal{P}_n(\mathbb{K})$ como sendo os polinómios $p+q$ e αp dados por:

$$\text{[A]} \qquad (p+q)(x) = \sum_{k=0}^{n} (a_k + b_k) x^k \tag{1.22}$$

$$\text{[P]} \qquad (\alpha p)(x) = \sum_{k=0}^{n} (\alpha a_k) x^k \tag{1.23}$$

Observe-se que também aqui se utiliza a estrutura algébrica pré-existente em $\mathbb{K}$ para definir as operações do espaço vectorial. Observe-se ainda que o *vector nulo* de $\mathcal{P}_n(\mathbb{K})$ é o polinómio nulo 0 e que o *simétrico* do polinómio p tal que $p(x) = \sum_{k=0}^{n} a_k x^k$ é o polinómio $-p$, estando estes definidos para todo o $x \in \mathbb{K}$, por:

$$0(x) = \sum_{k=0}^{n} 0 x^k \tag{1.24}$$

$$(-p)(x) = \sum_{k=0}^{n} (-a_k) x^k \tag{1.25}$$

É claro que, se $0 \leq m \leq n$, então $\mathcal{P}_m(\mathbb{K}) \subset \mathcal{P}_n(\mathbb{K})$. Em face do exposto anteriormente, podemos afirmar, por exemplo, que o conjunto $\mathcal{P}_3(\mathbb{R})$ dos polinómios de coeficientes reais e grau menor ou igual a 3 é um espaço vectorial real.

Exemplo 1.8 Sejam $(a_0, a_1, \ldots, a_n, \ldots) \in \mathbb{K}^{(\mathbb{N}_0)}$ as sucessões de escalares de um corpo $\mathbb{K}$, nulas a partir de alguma ordem (para cada sucessão, existe uma ordem r tal que $a_i = 0$, para $i > r$) e considere-se o conjunto $E = \mathcal{P}(\mathbb{K})$ dos polinómios $\sum_{k=0}^{+\infty} a_k x^k$ cujos coeficientes são os termos daquelas sucessões; note-se que, de facto, este somatório tem apenas um número finito de parcelas ($\leq r+1$), sendo nulas as restantes. É claro que $\mathcal{P}(\mathbb{K}) = \bigcup_{n \in \mathbb{N}_0} \mathcal{P}_n(\mathbb{K})$.

$$\mathcal{P}(\mathbb{K}) = \left\{ p \colon p(x) = \sum_{k=0}^{+\infty} a_k x^k, \text{com } (a_k) \in \mathbb{K}^{(\mathbb{N}_0)} \right\} \tag{1.26}$$

$\mathcal{P}(\mathbb{K})$ é um espaço vectorial sobre $\mathbb{K}$, se definirmos a soma de polinómios e o produto de escalar por um polinómio através de:

$$\text{[A]} \qquad (p+q)(x) = \sum_{k=0}^{+\infty} (a_k + b_k) x^k \tag{1.22.1}$$

$$\text{[P]} \qquad (\alpha p)(x) = \sum_{k=0}^{+\infty} (\alpha a_k) x^k \tag{1.23.1}$$

Observe-se ainda que o *vector nulo* de $\mathcal{P}(\mathbb{K})$ é o polinómio nulo $0 = \sum_{k=0}^{+\infty} 0x^k$ e que o *simétrico* do polinómio p tal que $p(x) = \sum_{k=0}^{+\infty} a_k x^k$ é o polinómio definido por:

$$(-p)(x) = \sum_{k=0}^{+\infty} (-a_k)x^k \tag{1.25.1}$$

Tem-se, como é óbvio,

$$0 \le m \le n \Rightarrow \mathcal{P}_0(\mathbb{K}) \subset \mathcal{P}_m(\mathbb{K}) \subset \mathcal{P}_n(\mathbb{K}) \subset \mathcal{P}(\mathbb{K})$$

Consideremos, agora, dois polinómios p e q dados por

$$p(x) = \sum_{k=0}^{+\infty} a_k x^k$$
$$q(x) = \sum_{k=0}^{+\infty} b_k x^k$$

e o seu produto usual

$$(pq)(x) = p(x)q(x) = \left(\sum_{k=0}^{+\infty} a_k x^k\right)\left(\sum_{k=0}^{+\infty} b_k x^k\right) = \sum_{k=0}^{+\infty} c_k x^k$$

onde, para $k \in \mathbb{N}_0$,

$$c_k = \sum_{i+j=k} a_i b_j = \sum_{i=0}^{k} a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + \cdots + a_{k-1} b_1 + a_k b_0$$

Esta multiplicação verifica os axiomas M1 a M4, tratando-se de uma *multiplicação vectorial* e, portanto, $\mathcal{P}(\mathbb{K})$ é uma *Álgebra Linear* (esta álgebra é, ainda, comutativa e com unidade: o polinómio constante $u(x) = 1$) sobre $\mathbb{K}$.

Segundo o que acabámos de ver e a título de exemplo, os conjuntos $\mathcal{P}(\mathbb{R})$ e $\mathcal{P}(\mathbb{C})$ dos polinómios de coeficientes reais (respectivamente, complexos) constitui uma Álgebra Linear sobre o corpo dos reais (respectivamente, complexos).

Exemplo 1.9 Seja $\mathbb{K} = \mathbb{R}$ e $E = S^3$ o conjunto dos segmentos orientados (incluindo o segmento nulo, de comprimento 0) aplicados num ponto O do espaço tridimensional. Definindo a adição em S^3 através da regra do triângulo[(5)] (ver figura 1.1) e o produto de um número $\alpha \in \mathbb{R}$ por um segmento $\overrightarrow{OP}$ como sendo um novo segmento orientado aplicado em O, com a mesma direcção de $\overrightarrow{OP}$, o mesmo sentido ou o oposto conforme $\alpha > 0$ ou $\alpha < 0$ e um

---

5 A soma de dois segmentos $\overrightarrow{OA}$ e $\overrightarrow{OB}$ é o segmento $\overrightarrow{OC}$, em que C é a extremidade do segmento de origem A e *equipolente* a $\overrightarrow{OB}$ (ou a extremidade do segmento de origem B e *equipolente* a $\overrightarrow{OA}$). Por *equipolente* entende-se com a mesma direcção, sentido e comprimento (ver exemplo 1.10).

comprimento igual a $|\alpha|$ vezes o comprimento de $\overrightarrow{OP}$ (ver figura 1.2), obtém-se um espaço vectorial real.

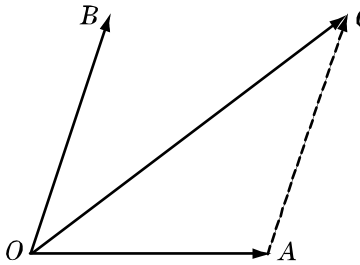


Fig. 1.1 – Regra do triângulo: $\overrightarrow{AC} \sim \overrightarrow{OB}$.

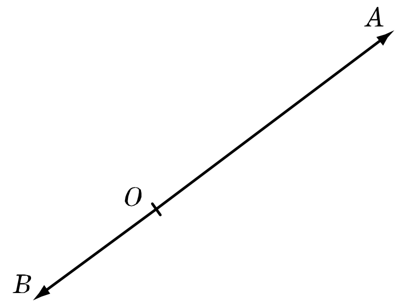


Fig. 1.2 – Multiplicação escalar: $\overrightarrow{OA} = -2\,\overrightarrow{OB}$.

É deste exemplo que provêm os termos *vector* e *espaço vectorial*. É ainda neste espaço que, a nível elementar, se pensa quando se fala em grandezas *escalares* (massa, temperatura, trabalho, energia, etc) e grandezas *vectoriais* (força, velocidade, campo eléctrico, etc), distinguindo-se estas das primeiras pelo facto de, além de *intensidade*, possuirem também *direcção* e *sentido*.

Sendo, ainda, $\mathbb{K} = \mathbb{R}$ e $E = S^2$ o conjunto dos segmentos orientados aplicados num ponto O de um *plano* qualquer e usando as regras anteriores para as operações de adição vectorial e multiplicação escalar, obtém-se um novo espaço vectorial real.

Por último, se $E = S^1$ for constituído pelos segmentos orientados existentes sobre uma *recta* e aplicados num ponto O desta e usando as regras anteriores para as operações, S^1 constituirá também um espaço vectorial real.

Exemplo 1.10 Seja $\mathbb{K} = \mathbb{R}$ e $E = \widetilde{S}^3$ o conjunto dos segmentos orientados (incluindo os segmentos nulos, de comprimento 0) aplicados em qualquer ponto do espaço tridimensional. Diremos que $\overrightarrow{AB}$ e $\overrightarrow{CD}$ são *equipolentes* (e escreve-se $\overrightarrow{AB} \sim \overrightarrow{CD}$) sse o ponto médio M de $\overline{AD}$ coincidir com o ponto médio de $\overline{BC}$ (ver figura 1.3).

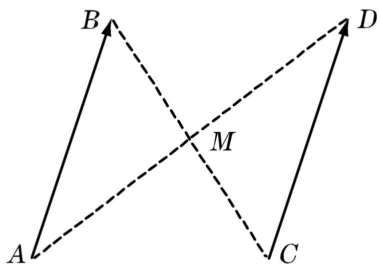


Fig. 1.3 – Segmentos equipolentes.

A relação $\sim$ de equipolência é uma relação de equivalência (ver secção A.3 do apêndice A), o que determina a partição de $\widetilde{S}^3$ em classes de equivalência formadas por segmentos orientados equipolentes entre si, as quais constituem o chamado *conjunto-quociente* de $\widetilde{S}^3$ por $\sim$, designado por $\widetilde{S}^3/\sim$. Cada classe de equivalência é chamada um *vector livre* (ou *deslizante*) e o referido conjunto quociente será designado por $\widehat{S}^3 = \widetilde{S}^3/\sim$. A classe (única) a que pertence um segmento $\overrightarrow{AB}$ é designada por $[\overrightarrow{AB}\,]$, dizendo-se que os segmentos pertencentes a $[\overrightarrow{AB}\,]$ são os *representantes* da classe. Para cada ponto O do espaço, existe um e um só representante de cada classe aplicado em O. Vamos, agora, ver que $\widehat{S}^3$ constitui um espaço vectorial real se definirmos as operações do seguinte modo:

[A] A soma $[\overrightarrow{AB}\,]+[\overrightarrow{CD}\,]$ dos vectores livres $[\overrightarrow{AB}\,]$ e $[\overrightarrow{CD}\,]$ é a classe (vector livre) a que pertence a soma (pela regra do triângulo) de dois representantes de $[\overrightarrow{AB}\,]$ e $[\overrightarrow{CD}\,]$ aplicados num mesmo ponto O qualquer do espaço tridimensional, isto é,

$$[\overrightarrow{AB}\,]+[\overrightarrow{CD}\,]=[\overrightarrow{OP}\,] \tag{1.27}$$

em que

$$\overrightarrow{OP} = \overrightarrow{OM} + \overrightarrow{ON}\text{, com } \overrightarrow{OM} \sim \overrightarrow{AB} \text{ e } \overrightarrow{ON} \sim \overrightarrow{CD}\,.$$

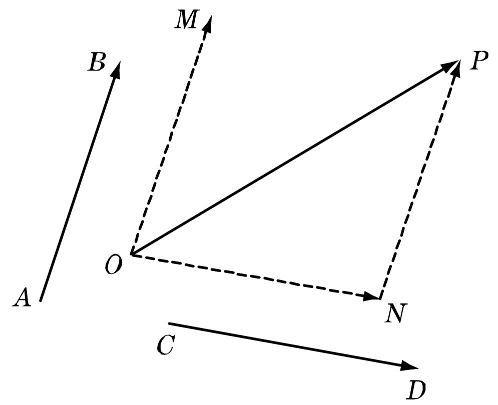


Fig. 1.4 – Adição vectorial em $\widehat{S}^3$.

[P] Quanto à multiplicação escalar, poremos, para qualquer $\alpha \in \mathbb{R}$,

$$\alpha[\overrightarrow{AB}\,]=[\alpha\overrightarrow{AB}\,] \tag{1.28}$$

onde $\alpha\overrightarrow{AB}$ é calculado de acordo com a regra do exemplo 1.9.

Na prática, em $\widehat{S}^3$ tudo se passa como se um vector livre fosse um segmento orientado que pode deslocar-se paralelamente a si próprio e estar, portanto, aplicado em qualquer ponto do espaço tridimensional (daí a designação de deslizante). De modo semelhante se podem definir os espaços vectoriais reais $\widehat{S}^2$ e $\widehat{S}^1$ dos vectores livres de um plano e de uma recta, respectivamente.

Exemplo 1.11 O conjunto $E=\mathbb{R}^+=(0,+\infty)$ constitui um espaço vectorial real pondo, para quaisquer $x,y\in\mathbb{R}^+$ e $\alpha\in\mathbb{R}$,

[A] $$x+y=xy \tag{1.29}$$

[P] $$\alpha x = x^\alpha = e^{\alpha \ln x} \tag{1.30}$$

O *vector nulo* deste espaço é 1 e o *simétrico* de x é $-x=1/x$.

Exemplo 1.12 Seja $(E_i)_{i\in I}$ uma família qualquer de espaços vectoriais sobre o mesmo corpo $\mathbb{K}$ e consideremos o produto cartesiano $E = \prod\limits_{i\in I} E_i$ dos espaços E_i

$$E = \prod_{i\in I} E_i = \left\{ (\vec{x}_i)_{i\in I} \colon \underset{i\in I}{\forall}\, \vec{x}_i \in E_i \right\}$$

Considerem-se, agora, as operações definidas por:

[A] $\quad (\vec{x}_i)_{i\in I} + (\vec{y}_i)_{i\in I} = (\vec{x}_i + \vec{y}_i)_{i\in I}$ (1.31.1)

[P] $\quad \alpha(\vec{x}_i)_{i\in I} = (\alpha\vec{x}_i)_{i\in I}$ (1.31.2)

Obtém-se, deste modo, um espaço vectorial sobre $\mathbb{K}$, chamado *espaço produto* dos E_i.

O *vector nulo* de E é constituído pelos vectores nulos dos E_i e o *simétrico* pelos simétricos (nos E_i) de cada vector componente:

$$\vec{o} = (\vec{o}_{E_i})_{i\in I} \tag{1.32.1}$$

$$-(\vec{x}_i)_{i\in I} = (-\vec{x}_i)_{i\in I} \tag{1.32.2}$$

Particularmente, podemos fazer os $(E_i)_{i\in I}$ todos iguais a um mesmo espaço vectorial E, obtendo-se então o espaço

$$E^I = \left\{ (\vec{x}_i)_{i\in I} \colon \underset{i\in I}{\forall}\, \vec{x}_i \in E \right\}$$

Exemplo 1.13 Se, no exemplo anterior, fizermos $I = \{1, 2, \ldots, n\}$, para $n > 0$, obtém-se o produto cartesiano

$$E = E_1 \times E_2 \times \cdots \times E_n = \prod_{i=1}^{n} E_i = \left\{ (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) \colon \underset{1\le i\le n}{\forall}\, \vec{x}_i \in E_i \right\}$$

As equações (1.31.1) e (1.31.2) podem, neste caso, escrever-se, para quaisquer $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n), (\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n) \in E$ e $\alpha \in \mathbb{K}$, na forma:

[A] $\quad (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) + (\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n) = (\vec{x}_1 + \vec{y}_1, \vec{x}_2 + \vec{y}_2, \ldots, \vec{x}_n + \vec{y}_n)$ (1.33.1)

[P] $\quad \alpha(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = (\alpha\vec{x}_1, \alpha\vec{x}_2, \ldots, \alpha\vec{x}_n)$ (1.33.2)

O *vector nulo* de E é constituído pelos vectores nulos dos E_i e o *simétrico* pelos simétricos (nos E_i) de cada vector componente :

$$\vec{o} = (\vec{o}_{E_1}, \vec{o}_{E_2}, \ldots, \vec{o}_{E_n}) \tag{1.34.1}$$

$$-(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = (-\vec{x}_1, -\vec{x}_2, \ldots, -\vec{x}_n) \tag{1.34.2}$$

Particularmente, podemos fazer os $(E_i)_{1\leq i\leq n}$ todos iguais a um mesmo espaço vectorial E, obtendo-se, então, o espaço

$$E^n = \left\{(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) \colon \underset{1\leq i\leq n}{\forall} \vec{x}_i \in E\right\}$$

Mais particularmente, se fizermos $E = \mathbb{K}$, teremos o espaço cartesiano $\mathbb{K}^n$ do exemplo 1.3.

Exemplo 1.14 Seja $(E, \mathbb{K}, +, \cdot)$ um espaço vectorial sobre um corpo $\mathbb{K}$ e $\mathbb{K}' \subset \mathbb{K}$ um subcorpo de $\mathbb{K}$. Se considerarmos a restrição $\odot$ do produto $\cdot$ ao conjunto $\mathbb{K}' \times E$ $\odot = \cdot|_{\mathbb{K}'\times E} \colon \mathbb{K}' \times E \to E$, facilmente se prova que $(E, \mathbb{K}', +, \odot)$ constitui igualmente um espaço vectorial, mas agora sobre $\mathbb{K}'$. Observe-se que este espaço é diferente do primeiro (embora com os mesmos vectores). Adiante (ver exemplo 1.56) veremos que a dimensão do novo espaço é, em geral, maior do que a do primeiro (pelo facto de existirem menos escalares). Em linguagem prática, enuncia-se por vezes esta propriedade dizendo que «Se E é um espaço vectorial sobre $\mathbb{K}$ e $\mathbb{K}'$ é subcorpo de $\mathbb{K}$, E é também espaço vectorial sobre $\mathbb{K}'$». O *vector nulo* e o *simétrico* de cada vector de E são evidentemente os mesmos nos dois espaços (visto que o grupo comutativo $(E, +)$ é o mesmo).

Este exemplo significa, pois, que todo o espaço vectorial *complexo* é automaticamente um espaço vectorial *real* e que este é, por sua vez, um espaço vectorial *racional*.

Exemplo 1.15 Neste exemplo vamos ver que, a partir de um espaço vectorial real, é sempre possível construir um espaço vectorial complexo. Seja, então, E um espaço vectorial real e façamos $E' = E^2 = \{(\vec{x}, \vec{y}) \colon \vec{x} \in E \wedge \vec{y} \in E\}$. O leitor pode, a título de exercício, verificar que E' é um espaço vectorial complexo, para as operações a seguir indicadas, onde $\alpha + i\beta \in \mathbb{C}$ (e, claro $\alpha, \beta \in \mathbb{R}$) e $(\vec{x}, \vec{y}), (\vec{x}', \vec{y}') \in E'$:

[A] $$(\vec{x}, \vec{y}) + (\vec{x}', \vec{y}') = (\vec{x} + \vec{x}', \vec{y} + \vec{y}') \tag{1.35}$$

[P] $$(\alpha + i\beta)(\vec{x}, \vec{y}) = (\alpha\vec{x} - \beta\vec{y}, \beta\vec{x} + \alpha\vec{y}) \tag{1.36}$$

Note-se que também aqui se utilizaram as operações pré-definidas no espaço E para definir as operações de E' e que o *vector nulo* de E' é $(\vec{o}, \vec{o})$ e $(-\vec{x}, -\vec{y})$ é o *simétrico* de $(\vec{x}, \vec{y})$.

Do acima exposto resulta, por exemplo, que $\mathbb{R}^4$ é espaço vectorial complexo, com as operações:

[A] $$(x_1, x_2, x_3, x_4) + (y_1, y_2, y_3, y_4) = (x_1 + y_1, x_2 + y_2, x_3 + y_3, x_4 + y_4)$$

[P] $$(\alpha + i\beta)(x_1, x_2, x_3, x_4) = (\alpha x_1 - \beta x_3, \alpha x_2 - \beta x_4, \beta x_1 + \alpha x_3, \beta x_2 + \alpha x_4)$$

1.5 Combinações lineares. Subespaços

Definição 1.3 – Combinação linear – *Seja $m > 0$ um número inteiro,*

$$x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$$

uma sequência de m vectores de um espaço vectorial E sobre um corpo $\mathbb{K}$ e $\vec{\alpha} = (\alpha_i)_{1\leq i\leq m}$ uma sequência de m escalares (ou seja, um vector do espaço $\mathbb{K}^m$). Chama-se ***combinação***

linear *dos $\vec{x}_i$ a uma soma (que resulta num vector de E) do tipo*

$$\sum_{i=1}^{m} \alpha_i \vec{x}_i \tag{1.37.1}$$

Se $m = 0$, põe-se, por definição:

$$\sum_{i \in \emptyset} \alpha_i \vec{x}_i = \vec{o} \tag{1.37.2}$$

Os exemplos seguintes ilustram o conceito de combinação linear:

Exemplo 1.16 Em $\mathbb{R}^3$, o vector $(-6, -11, 6)$ é combinação linear dos vectores da sequência $x = ((1, -2, 1), (-2, -1, 2), (1, 2, 1))$, visto que

$$(-6, -11, 6) = 2(1, -2, 1) + 3(-2, -1, 2) - 2(1, 2, 1)$$

Neste caso, esta é a única combinação linear dos vectores $(1, -2, 1), (-2, -1, 2)$ e $(1, 2, 1)$ que resulta no vector $(-6, -11, 6)$: de facto, a igualdade

$$(-6, -11, 6) = x(1, -2, 1) + y(-2, -1, 2) + z(1, 2, 1)$$

equivale ao sistema de equações

$$\begin{cases} x - 2y + z &= -6 \\ -2x - y + 2z &= -11 \\ x + 2y + z &= 6 \end{cases}$$

o qual é *determinado* e tem apenas a solução $x = 2, y = 3, z = -2$.

Exemplo 1.17 Em $\mathbb{R}^3$, o vector $(-2, -11, 10)$ é também combinação linear dos vectores da lista $u = ((-2, -1, 2), (-1, -8, 7), (1, -2, 1))$, visto que

$$\begin{aligned}(-2, -11, 10) &= 1(-2, -1, 2) + 1(-1, -8, 7) + 1(1, -2, 1) \\ &= -1(-2, -1, 2) + 2(-1, -8, 7) - 2(1, -2, 1)\end{aligned}$$

Observe que, neste caso, existem mesmo infinitas combinações lineares dos vectores de u que resultam no vector $(-2, -11, 10)$, duas das quais se indicam em cima: de facto, tem-se, para qualquer $a \in \mathbb{R}$,

$$(-2, -11, 10) = (3 - 2a)(-2, -1, 2) + a(-1, -8, 7) + (4 - 3a)(1, -2, 1)$$

Procedendo como anteriormente, a igualdade

$$(-2, -11, 10) = x(-2, -1, 2) + y(-1, -8, 7) + z(1, -2, 1)$$

equivale ao seguinte sistema de equações lineares:

$$\begin{cases} -2x - y + z &= -2 \\ -x - 8y - 2z &= -11 \\ 2x + 7y + z &= 10 \end{cases}$$

Este sistema é *indeterminado* e tem uma infinidade de soluções que são (a é real arbitrário):

$$x = 3 - 2a, y = a, z = 4 - 3a, \text{ onde } a \in \mathbb{R}$$

Exemplo 1.18 Em $\mathbb{R}^3$, o vector $(-2, 2, 1)$ não é combinação linear dos vectores da lista $v = ((-2, -1, 2), (-1, -8, 7), (1, -2, 1))$. Agora, a igualdade

$$(-2, 2, 1) = x(-2, -1, 2) + y(-1, -8, 7) + z(1, -2, 1)$$

é equivalente ao sistema de equações lineares

$$\begin{cases} -2x - y + z &= -2 \\ -x - 8y - 2z = 2 \\ 2x + 7y + z &= 1 \end{cases}$$

O sistema anterior não tem qualquer solução (sistema dito *impossível* ou *inconsistente*) não sendo, pois, possível exprimir $(-2, 2, 1)$ como combinação linear dos vectores de v.

Exemplo 1.19 No espaço vectorial real $\mathbb{R}^{\mathbb{R}}$ das funções de $\mathbb{R}$ em $\mathbb{R}$ e para qualquer $a \in \mathbb{R}$, as funções $x \mapsto \cos(x + a)$ e $x \mapsto \sin(x + a)$ são combinação linear das funções da lista $(\cos, \sin)$, visto que

$$\begin{cases} \cos(x + a) = \cos a \cos x - \sin a \sin x \\ \sin(x + a) \ = \sin a \cos x + \cos a \sin x \end{cases}$$

Estas combinações lineares são únicas: por exemplo e em relação à primeira, de

$$\underset{x \in \mathbb{R}}{\forall} (\cos(x + a) = u \cos x + v \sin x)$$

resulta, com $x = 0$ e $x = \pi/2$,

$$\begin{cases} u = \cos a \\ v = \cos(\pi/2 + a) = -\sin a \end{cases}$$

Exemplo 1.20 No espaço $\mathscr{P}_n(\mathbb{K})$ dos polinómios de grau menor ou igual a n, todo o polinómio $p \in \mathscr{P}_n(\mathbb{K})$ é combinação linear dos polinómios da lista $c = (1, x, x^2, \ldots, x^n)$, visto que

$$p(x) = \sum_{k=1}^{n} a_k x^k$$

Observe-se que, para cada polinómio p, a combinação linear anterior é única (método dos coeficientes indeterminados).

Em muitos casos, restringindo o número de vectores de um espaço vectorial E sobre um corpo $\mathbb{K}$ e usando as mesmas regras operatórias de E, obtém-se ainda um espaço vectorial sobre $\mathbb{K}$: é o caso dos chamados *subespaços* de E, cuja definição se apresenta a seguir:

Definição 1.4 – Subespaço – *Sendo E um espaço vectorial sobre um corpo $\mathbb{K}$ e $F \subset E$ um subconjunto de E, diz-se que F é um* **subespaço** *vectorial de E e escrevemos $F \preceq E$ sse*

[S1] F não é vazio, isto é, $F \neq \emptyset$.

[S2] F é *fechado* em relação à adição vectorial, ou seja,

$$\underset{\vec{x},\vec{y}}{\forall} (\vec{x} \in F \wedge \vec{y} \in F \Rightarrow \vec{x} + \vec{y} \in F)$$

[S3] F é *fechado* para a multiplicação escalar, isto é,

$$\underset{\alpha,\vec{x}}{\forall} (\alpha \in \mathbb{K} \wedge \vec{x} \in F \Rightarrow \alpha\vec{x} \in F)$$

É fácil reconhecer que a conjunção de S2 e S3 é equivalente a

[S4] F é fechado para as combinações lineares, ou seja,

$$\underset{\alpha,\beta,\vec{x},\vec{y}}{\forall} (\alpha \in \mathbb{K} \wedge \beta \in \mathbb{K} \wedge \vec{x} \in F \wedge \vec{y} \in F \Rightarrow \alpha\vec{x} + \beta\vec{y} \in F)$$

Facilmente se mostra que estas condições são as necessárias e suficientes para que F, munido da *restrição da adição vectorial em E a F^2* (S2 assegura que esta restrição é uma lei de composição interna *em* F) e da *restrição da multiplicação escalar em E a* $\mathbb{K} \times F$ (S3 garante que esta é uma aplicação *em* F), seja ainda um espaço vectorial sobre $\mathbb{K}$. A relação $\preceq$ entre subespaços de E é uma relação de ordem lata parcial (*reflexiva*, *anti-simétrica* e *transitiva* – ver apêndice A, definição A.6).

Observe-se que S3 implica, com $\alpha = -1$, que F é fechado em relação à *simetrização*, ou seja, $\vec{x} \in F \Rightarrow -\vec{x} \in F$. Os fechos de F em relação à *simetrização* e à *adição* levam ao fecho de F em relação à *subtracção* vectorial, definida por (1.1). Isto implica então, por S1, que existe pelo menos um vector $\vec{x} \in F$ e portanto $\vec{x} - \vec{x} = \vec{o} \in F$: logo, todo o subespaço $F \subset E$ contém, portanto, o *vector nulo* de E (que é também o *vector nulo* de F). Isto mostra que a condição S1 poderá ser substituída por:

[S0] $\vec{o} \in F$.

O conjunto singular formado apenas pelo vector nulo $\{\vec{o}\}$ é, como facilmente se reconhece, um subespaço de E e constitui, obviamente, o *menor* subespaço de E possível (o *mínimo* no sentido da relação de ordem $\preceq$), sendo o *maior* o próprio E (o *máximo* no sentido da referida relação $\preceq$). $\{\vec{o}\}$ e E são chamados os *subespaços triviais* de E, dizendo-se os restantes *não triviais* ou *próprios*.

S2 e S3 significam igualmente que, se F é subespaço de E, F conterá qualquer combinação linear de vectores seus, isto é, se $(\vec{x}_i)_{1\leq i\leq m}$ for uma lista de vectores de F (eventualmente vazia), então, para quaisquer escalares $(\alpha_i)_{1\leq i\leq m}$,

$$\sum_{i=1}^{m} \alpha_i\vec{x}_i \in F \tag{1.38}$$

Nas proposições seguintes, dá-se conta do comportamento dos subespaços em relação às operações com conjuntos.

Proposição 1.4 – Intersecção de subespaços – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e $(F_i)_{i\in I}$ uma família qualquer (eventualmente infinita) de subespaços de E. Então o conjunto $F = \bigcap\limits_{i\in I} F_i$ é ainda um subespaço de E.*

Demonstração:

Basta provar que S1, S2 e S3 se verificam em relação a F.

S1: Como os F_i são subespaços de E, tem-se $\vec{o} \in F_i$ para todo o $i \in I$ e portanto $\vec{o} \in F$. Então F é não vazio.

S2: Sejam $\vec{x}$ e $\vec{y}$ vectores de F. Então, $\vec{x}, \vec{y} \in F_i$ para todo o $i \in I$. Como os F_i são subespaços, $\vec{x} + \vec{y} \in F_i$, para todo o $i \in I$ e portanto $\vec{x} + \vec{y} \in F$.

S3: Seja $\vec{x}$ um vector de F e α um escalar de $\mathbb{K}$. Então $\vec{x} \in F_i$, para todo o $i \in I$. Sendo os F_i subespaços, ter-se-á $\alpha\vec{x} \in F_i$ para todo o $i \in I$ e, em consequência, $\alpha\vec{x} \in F$, o que termina a demonstração. ☐

Corolário 1.4.1 *Se F e G são subespaços de um espaço vectorial E, então $F \cap G$ é também um subespaço de E.*

Proposição 1.5 – Reunião de subespaços – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e F e G subespaços de E. Então, $F \cup G$ é um subespaço de E sse $F \subset G \vee G \subset F$.*

Demonstração:

Como $F \subset G \Rightarrow F \cup G = G$ e $G \subset F \Rightarrow F \cup G = F$, $F \cup G$ será, evidentemente, um subespaço caso se dê uma daquelas inclusões. Vejamos, agora, a recíproca que é equivalente a $F \cup G$ é subespaço $\wedge$ $F \not\subset G \Rightarrow G \subset F$: seja $\vec{u}$ um vector de F que não pertença a G e $\vec{v}$ um vector qualquer de G. Ambos pertencem a $F \cup G$ e, sendo este um subespaço, o mesmo sucederá com $\vec{u} + \vec{v}$. Sendo assim, tem-se $\vec{u} + \vec{v} \in F \vee \vec{u} + \vec{v} \in G$. Mas a hipótese $\vec{u} + \vec{v} \in G$ é absurda, visto que então, como G é subespaço, ter-se-ia $(\vec{u} + \vec{v}) - \vec{v} = \vec{u} \in G$, o que contraria a nossa escolha inicial de $\vec{u}$. Portanto, terá de ser $\vec{u} + \vec{v} \in F$, o que, sendo F um subespaço, implica $(\vec{u} + \vec{v}) - \vec{u} = \vec{v} \in F$. Isto mostra que $G \subset F$, provando deste modo a proposição. ☐

Proposição 1.6 – Subtracção de subespaços – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e F e G subespaços de E. Então, $F \setminus G$ não é um subespaço de E.*

Demonstração:

Basta notar que $F \setminus G$ não contém o vector nulo. ☐

De seguida, apresentam-se vários exemplos de subespaços vectoriais dalguns dos espaços anteriormente apresentados:

Exemplo 1.21 Seja $\vec{a} = (a_1, a_2, \ldots, a_n)$ um vector arbitrário do espaço cartesiano $\mathbb{K}^n$. Os subconjuntos $F_{\vec{a}}$ de $\mathbb{K}^n$ definidos por

$$F_{\vec{a}} = \left\{ (x_1, x_2, \ldots, x_n) \colon \sum_{j=1}^{n} a_j x_j = 0 \right\}$$

constituem subespaços de $\mathbb{K}^n$ que coincidem com $\mathbb{K}^n$, se $\vec{a} = \vec{o}$. Se $\vec{a} \neq \vec{o}$, estes subespaços são, por vezes, chamados *hiperplanos* de $\mathbb{K}^n$. Por exemplo, no caso de $\mathbb{R}^2$ (e interpretando geometricamente) e com $\vec{a} \neq \vec{o}$, os conjuntos $F_{\vec{a}}$ serão *rectas* passando pela origem; Em $\mathbb{R}^3$ e de novo com $\vec{a} \neq \vec{o}$, serão *planos* passando pela origem.

Exemplo 1.22 De acordo com a proposição 1.4, para qualquer sequência $(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m)$ de m vectores de $\mathbb{K}^n$, o conjunto

$$F = \bigcap_{i=1}^{m} F_{\vec{a}_i} \tag{1.39}$$

onde os $(F_{\vec{a}_i})_{1 \leq i \leq m}$ são espaços do tipo construído no exemplo anterior, constituirá igualmente um subespaço de $\mathbb{K}^n$. Sendo os $\vec{a}_i$ da forma $\vec{a}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, os vectores $(x_1, x_2, \ldots, x_n) \in F$ satisfarão o sistema de equações lineares homogéneas seguinte

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = 0 \\ \cdots & \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = 0 \end{cases}$$

Se algum dos vectores $\vec{a}_i$ for nulo, o subespaço $F_{\vec{a}_i}$ correspondente pode ser eliminado da expressão (1.39), visto que $F_{\vec{a}_i} = \mathbb{K}^n$ é elemento neutro na intersecção (portanto, naquela expressão podem figurar só os *hiperplanos* de $\mathbb{K}^n$).

Como casos particulares, aponte-se, por exemplo em $\mathbb{R}^3$, a intersecção de dois ($m = 2$ e $n = 3$) planos não coincidentes passando pela origem, que produz uma *recta* passando também pela origem. Portanto, as *rectas passando pela origem* são subespaços vectoriais de $\mathbb{R}^3$.

Exemplo 1.23 Sendo $\vec{x}$ um vector qualquer de um espaço vectorial E sobre um corpo $\mathbb{K}$, o conjunto

$$\mathbb{K}\vec{x} = \left\{ \vec{y}\colon \underset{\alpha \in \mathbb{K}}{\exists}\ \vec{y} = \alpha\vec{x} \right\}$$

dos múltiplos escalares de $\vec{x}$ é um subespaço de E, como facilmente se reconhece. Se $\vec{x} \neq \vec{o}$, este subespaço é, por vezes, designado por *recta* de E; se $\vec{x} = \vec{o}$, fica como é óbvio $\mathbb{K}\vec{x} = \{\vec{o}\}$.

Exemplo 1.24 Seja $(E_i)_{i \in I}$ uma família qualquer de espaços vectoriais sobre o mesmo corpo $\mathbb{K}$. O subconjunto de $\prod\limits_{i \in I} E_i$ formado pelas famílias $(\vec{x}_i)_{i \in I}$ de vectores dos E_i, com $\vec{x}_i \in E_i$ para todo o $i \in I$, e tais que $\vec{x}_i = \vec{o}_{E_i}$, para i no complementar de uma parte finita J de I, constitui também um subespaço de $\prod\limits_{i \in I} E_i$ chamado *soma directa externa* dos E_i e designado por $\bigoplus\limits_{i \in I} E_i \subset \prod\limits_{i \in I} E_i$. É claro que, se I é *finito*, tem-se $\bigoplus\limits_{i \in I} E_i = \prod\limits_{i \in I} E_i$. Se, em particular, os E_i forem todos iguais a um mesmo espaço vectorial E, a soma directa externa designa-se por $E^{(I)}$ (coincidente com E^I, se I for *finito*): trata-se, portanto, do conjunto das famílias $(\vec{x}_i)_{i \in I}$ de vectores de E tais que $\vec{x}_i = \vec{o}$ no complementar $I \setminus J$ de uma parte J finita do conjunto I de índices:

$$\bigoplus_{i\in I} E_i = \left\{ (\vec{x}_i)_{i\in I} \colon \underset{i\in I}{\forall} (\vec{x}_i \in E_i) \wedge \underset{\substack{J\subset I\\ J \text{ finito}}}{\exists} \ \underset{i\in I\setminus J}{\forall} \vec{x}_i = \vec{o}_{E_i} \right\}$$

$$E^{(I)} = \left\{ (\vec{x}_i)_{i\in I} \colon \underset{i\in I}{\forall} (\vec{x}_i \in E) \wedge \underset{\substack{J\subset I\\ J \text{ finito}}}{\exists} \ \underset{i\in I\setminus J}{\forall} \vec{x}_i = \vec{o} \right\}$$

Exemplo 1.25 O espaço $\mathbb{K}^{(\mathbb{N}_0)}$ do exemplo 1.5 é subespaço do espaço $\mathbb{K}^{\mathbb{N}_0}$ do exemplo 1.4.

Exemplo 1.26 Consideremos os espaços $\mathbb{C}^{\mathbb{N}}$ e $\mathbb{R}^{\mathbb{N}}$ das sucessões complexas (ou reais) e, nestes, os subconjuntos $\mathscr{B}(\mathbb{N},\mathbb{C})$ e $\mathscr{B}(\mathbb{N},\mathbb{R})$, obviamente não vazios, formados pelas sucessões *limitadas*

$$\mathscr{B}(\mathbb{N},\mathbb{C}) = \left\{ (z_n) \colon \underset{r\in\mathbb{R}^+}{\exists} \ \underset{n\in\mathbb{N}}{\forall} |z_n| < r \right\}$$

$$\mathscr{B}(\mathbb{N},\mathbb{R}) = \left\{ (x_n) \colon \underset{r\in\mathbb{R}^+}{\exists} \ \underset{n\in\mathbb{N}}{\forall} |x_n| < r \right\}$$

Como se sabe, a soma de sucessões limitadas é ainda uma sucessão limitada e o produto de um complexo (respectivamente, de um real) por uma sucessão complexa (respectivamente, real) limitada também o é; portanto, estamos de novo em presença de um subespaço de $\mathbb{C}^{\mathbb{N}}$ (respectivamente, de $\mathbb{R}^{\mathbb{N}}$).

Exemplo 1.27 Consideremos o espaço $\mathbb{C}^{\mathbb{N}}$ das sucessões complexas (ou o espaço $\mathbb{R}^{\mathbb{N}}$ das sucessões reais) e nestes os subconjuntos não vazios $\mathfrak{C}(\mathbb{N},\mathbb{C})$ e $\mathfrak{C}(\mathbb{N},\mathbb{R})$ formados pelas *sucessões convergentes*

$$\mathfrak{C}(\mathbb{N},\mathbb{C}) = \left\{ (z_n) \colon \underset{L\in\mathbb{C}}{\exists} \lim_{n\to\infty} z_n = L \right\}$$

$$\mathfrak{C}(\mathbb{N},\mathbb{R}) = \left\{ (x_n) \colon \underset{L\in\mathbb{R}}{\exists} \lim_{n\to\infty} x_n = L \right\}$$

Como se sabe, a soma de sucessões convergentes é ainda uma sucessão convergente e o produto de um complexo (ou de um real) por uma sucessão complexa (ou real) convergente também o é; portanto, estamos em presença de um subespaço de $\mathbb{C}^{\mathbb{N}}$ (ou de $\mathbb{R}^{\mathbb{N}}$). Mais ainda, como toda a sucessão convergente é limitada, $\mathfrak{C}(\mathbb{N},\mathbb{C})$ e $\mathfrak{C}(\mathbb{N},\mathbb{R})$ são subespaços de $\mathscr{B}(\mathbb{N},\mathbb{C})$ e de $\mathscr{B}(\mathbb{N},\mathbb{R})$, respectivamente.

Exemplo 1.28 Os conjuntos $\mathfrak{C}_0(\mathbb{N},\mathbb{C})$ e $\mathfrak{C}_0(\mathbb{N},\mathbb{R})$ das sucessões complexas ou reais convergentes para 0 (*infinitésimos*) são subespaços dos espaços $\mathfrak{C}(\mathbb{N},\mathbb{C})$ e $\mathfrak{C}(\mathbb{N},\mathbb{R})$ do exemplo anterior

$$\mathfrak{C}_0(\mathbb{N},\mathbb{C}) = \left\{ (z_n) \colon \lim_{n\to\infty} z_n = 0 \right\}$$

$$\mathfrak{C}_0(\mathbb{N},\mathbb{R}) = \left\{ (x_n) \colon \lim_{n\to\infty} x_n = 0 \right\}$$

Como se sabe, a sucessão nula é um infinitésimo, a soma de infinitésimos é ainda um infinitésimo e o produto de um número complexo (respectivamente, um número real) por um infinitésimo também o é; portanto, trata-se de subespaços de $\mathfrak{C}(\mathbb{N},\mathbb{C})$ e $\mathfrak{C}(\mathbb{N},\mathbb{R})$, respectivamente. Tem-se, afinal,

$$\mathfrak{C}_0(\mathbb{N},\mathbb{C}) \subset \mathfrak{C}(\mathbb{N},\mathbb{C}) \subset \mathscr{B}(\mathbb{N},\mathbb{C}) \subset \mathbb{C}^{\mathbb{N}}$$
$$\mathfrak{C}_0(\mathbb{N},\mathbb{R}) \subset \mathfrak{C}(\mathbb{N},\mathbb{R}) \subset \mathscr{B}(\mathbb{N},\mathbb{R}) \subset \mathbb{R}^{\mathbb{N}}$$

Exemplo 1.29 Consideremos o espaço $\mathbb{C}^{\mathbb{N}}$ (ou $\mathbb{R}^{\mathbb{N}}$) das sucessões complexas (respectivamente, reais) e neste o subconjunto não vazio $l^2_{\mathbb{C}}$ (respectivamente, $l^2_{\mathbb{R}}$) formado pelas sucessões (u_n) tais que a série de termo geral $|u_n|^2$ é convergente:

$$l^2_{\mathbb{C}} = \left\{ (z_n) \colon \underset{L\in\mathbb{R}}{\exists} \sum_{n=1}^{+\infty} |z_n|^2 = L \right\}$$
$$l^2_{\mathbb{R}} = \left\{ (x_n) \colon \underset{L\in\mathbb{R}}{\exists} \sum_{n=1}^{+\infty} x_n^2 = L \right\}$$

Não é difícil mostrar que estamos, de novo, em presença de subespaços de $\mathbb{C}^{\mathbb{N}}$ e $\mathbb{R}^{\mathbb{N}}$, respectivamente, e que são ainda subespaços dos espaços respectivos do exemplo anterior (para o fecho em relação à adição, observe que, para quaisquer sucessões (z_n) e (w_n) complexas ou reais, se tem: $|z_n + w_n|^2 \le 2|z_n|^2 + 2|w_n|^2$). Estes espaços chamam-se *espaços de Hilbert*[(6)] de sucessões.

Exemplo 1.30 No espaço vectorial complexo (respectivamente, real) $\mathbb{C}^I$ (respectivamente, $\mathbb{R}^I$) das funções escalares complexas (respectivamente, reais) definidas num intervalo $I \subset \mathbb{R}$, considere-se o conjunto $\mathscr{C}(I,\mathbb{C})$ (respectivamente, $\mathscr{C}(I,\mathbb{R})$) das funções *contínuas* em I

$$\mathscr{C}(I,\mathbb{C}) = \left\{ f \in \mathbb{C}^I \colon \underset{a\in I}{\forall} \lim_{x\to a} f(x) = f(a) \right\}$$
$$\mathscr{C}(I,\mathbb{R}) = \left\{ f \in \mathbb{R}^I \colon \underset{a\in I}{\forall} \lim_{x\to a} f(x) = f(a) \right\}$$

Como é sabido, a soma de funções contínuas no intervalo I é uma função contínua nesse intervalo, o mesmo se passando com o produto de um escalar por uma função contínua. Também se utilizam as notações $\mathscr{C}^0(I,\mathbb{C})$ e $\mathscr{C}^0(I,\mathbb{R})$. Em face do acima exposto, podemos afirmar que $\mathscr{C}(I,\mathbb{C})$ e $\mathscr{C}(I,\mathbb{R})$ são subespaços de $\mathbb{C}^I$ e $\mathbb{R}^I$, respectivamente.

Exemplo 1.31 No espaço vectorial complexo (respectivamente, real) $\mathbb{C}^I$ (respectivamente, $\mathbb{R}^I$) das funções escalares complexas (respectivamente, reais) definidas num intervalo $I \subset \mathbb{R}$,

---

6 *Hilbert, David*: matemático alemão (Königsberg 1862 – Göttingen 1943). Mais geralmente, um *espaço de Hilbert* é um espaço vectorial real ou complexo com produto interno e *completo* em relação à norma definida por este (ver capítulo 6).

considere-se o conjunto $\mathscr{B}(I,\mathbb{C})$ (respectivamente, $\mathscr{B}(I,\mathbb{R})$) das funções *limitadas* em I

$$\mathscr{B}(I,\mathbb{C}) = \left\{ f \in \mathbb{C}^I \colon \underset{L\in\mathbb{R}^+}{\exists} \ \underset{x\in I}{\forall} |f(x)| < L \right\}$$
$$\mathscr{B}(I,\mathbb{R}) = \left\{ f \in \mathbb{R}^I \colon \underset{L\in\mathbb{R}^+}{\exists} \ \underset{x\in I}{\forall} |f(x)| < L \right\}$$

Sabe-se que a soma de funções limitadas em I é também limitada e que o mesmo sucede com o produto de um escalar (real ou complexo) por uma função limitada: está-se, pois, mais uma vez na presença de um subespaço de $\mathbb{C}^I$ ou de $\mathbb{R}^I$, conforme o caso.

Exemplo 1.32 Seja $I \subset \mathbb{R}$ um intervalo aberto e $\mathscr{D}(I,\mathbb{C})$ e $\mathscr{D}(I,\mathbb{R})$ os conjuntos das funções de I em $\mathbb{C}$ ou de I em $\mathbb{R}$ *diferenciáveis* em I. Também aqui estamos na presença de subespaços de $\mathbb{C}^I$ ou de $\mathbb{R}^I$:

$$\mathscr{D}(I,\mathbb{C}) = \left\{ f \in \mathbb{C}^I \colon \underset{a\in I}{\forall} \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \in \mathbb{C} \right\}$$
$$\mathscr{D}(I,\mathbb{R}) = \left\{ f \in \mathbb{R}^I \colon \underset{a\in I}{\forall} \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \in \mathbb{R} \right\}$$

Exemplo 1.33 Seja $I \subset \mathbb{R}$ um intervalo aberto, $k \geq 0$ um inteiro e $\mathscr{C}^k(I,\mathbb{C})$ e $\mathscr{C}^k(I,\mathbb{R})$ os conjuntos das funções de I em $\mathbb{C}$ ou de I em $\mathbb{R}$ k *vezes diferenciáveis* em I e com derivada $f^{(k)}$ de ordem k *contínua* em I (entende-se por derivada de ordem 0 a própria função, isto é $\mathscr{C}^0(I,\mathbb{C}) = \mathscr{C}(I,\mathbb{C})$ e $\mathscr{C}^0(I,\mathbb{R}) = \mathscr{C}(I,\mathbb{R})$). Estas funções dizem-se de *classe* $\mathscr{C}^k$ em I. Também aqui estamos na presença de subespaços de $\mathbb{C}^I$ e de $\mathbb{R}^I$:

$$\mathscr{C}^k(I,\mathbb{C}) = \left\{ f \in \mathbb{C}^I \colon f^{(k)} \in \mathscr{C}(I,\mathbb{C}) \right\}$$
$$\mathscr{C}^k(I,\mathbb{R}) = \left\{ f \in \mathbb{R}^I \colon f^{(k)} \in \mathscr{C}(I,\mathbb{R}) \right\}$$

Tem-se, neste caso (ver figura 1.5),

$$k \geq k' \geq 0 \Rightarrow \mathscr{C}^k(I,\mathbb{C}) \subset \mathscr{C}^{k'}(I,\mathbb{C}) \subset \mathscr{C}^0(I,\mathbb{C})$$

Exemplo 1.34 Seja $I \subset \mathbb{R}$ um intervalo aberto e $\mathscr{C}^\infty(I,\mathbb{C})$ e $\mathscr{C}^\infty(I,\mathbb{R})$ os conjuntos das funções de I em $\mathbb{C}$ ou de I em $\mathbb{R}$ *infinitamente diferenciáveis* em I (funções de *classe* $\mathscr{C}^\infty$ em I). Pela proposição 1.4, também aqui estamos na presença de subespaços de $\mathbb{C}^I$ e de $\mathbb{R}^I$, que são a intersecção dos espaços do exemplo anterior:

$$\mathscr{C}^\infty(I,\mathbb{C}) = \left\{ f \in \mathbb{C}^I \colon \underset{k\in\mathbb{N}_0}{\forall} f^{(k)} \text{ existe} \right\} = \bigcap_{k\in\mathbb{N}_0} \mathscr{C}^k(I,\mathbb{C})$$
$$\mathscr{C}^\infty(I,\mathbb{R}) = \left\{ f \in \mathbb{R}^I \colon \underset{k\in\mathbb{N}_0}{\forall} f^{(k)} \text{ existe} \right\} = \bigcap_{k\in\mathbb{N}_0} \mathscr{C}^k(I,\mathbb{R})$$

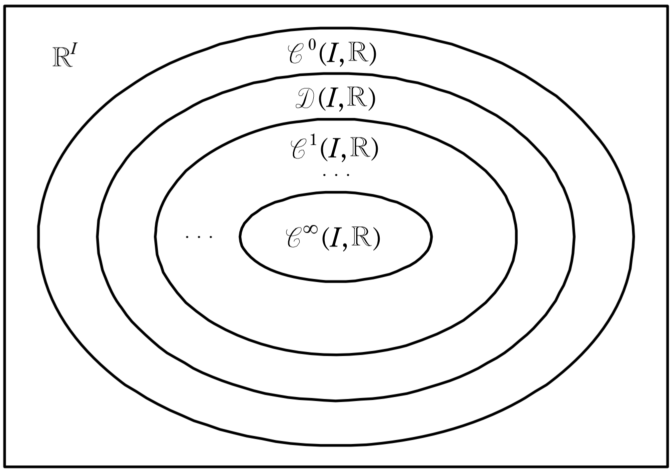


Fig. 1.5 – Os espaços de funções reais de classe $\mathscr{C}^k$ e de classe $\mathscr{C}^\infty$.

Exemplo 1.35 Os espaços $\mathscr{P}_n(\mathbb{K})$ dos polinómios de grau menor ou igual a um certo $n \geq 0$ (Exemplo 1.7) são subespaços uns dos outros: tem-se, de facto, $\mathscr{P}_m(\mathbb{K}) \subset \mathscr{P}_n(\mathbb{K}) \subset \mathscr{P}(\mathbb{K})$, se $0 \leq m \leq n$. A figura seguinte ilustra a situação:

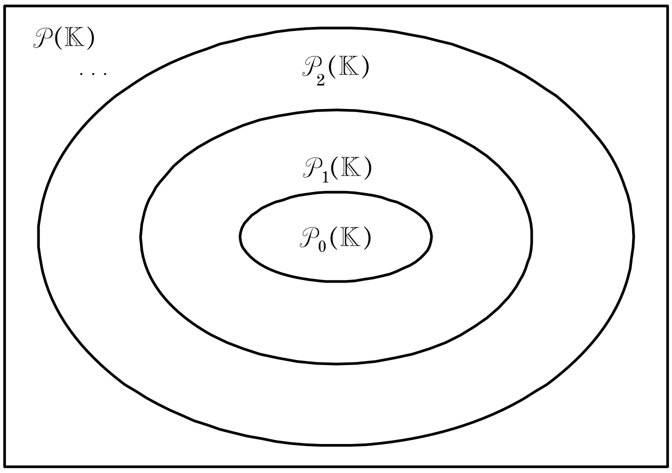


Fig. 1.6 – Os espaços de polinómios de coeficientes num corpo $\mathbb{K}$.

Exemplo 1.36 No espaço S^3 dos segmentos orientados aplicados num ponto O do espaço tridimensional, o conjunto P dos segmentos orientados aplicados em O e existentes sobre um qualquer *plano* passando por O constitui, como facilmente se verifica, um subespaço de S^3. O mesmo se pode dizer do conjunto R dos segmentos orientados aplicados em O e existentes sobre uma *recta* que contenha O.

Vamos, agora, provar que o conjunto de todas as combinações lineares de uma lista de vectores de um espaço vectorial E constituem sempre um *subespaço* de E e que este é o *menor* subespaço de E que contém simultaneamente todos os vectores da lista:

Proposição 1.7 *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ uma lista de $m \geq 0$ vectores de E. O conjunto A de todas as combinações lineares possíveis dos vectores de x é um* ***subespaço*** *de E, contendo todos os vectores de x e que está contido em* *qualquer subespaço de E que contenha igualmente os vectores de x: é, portanto, o* ***menor*** *subespaço de E contendo os vectores de x.*

Demonstração:

Se $m = 0$, tem-se $x = \emptyset$. Neste caso, por ser $\sum_{i\in\emptyset} \alpha_i \vec{x}_i = \vec{o}$, tem-se $A = \{\vec{o}\}$, que satisfaz todas as condições do enunciado. Passemos, então, ao caso $m > 0$: a lista x será

$$x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$$

É óbvio que $A \neq \emptyset$, visto que $\vec{o} = \sum_{i=1}^{m} 0\vec{x}_i \in A$. Por outro lado, A é fechado para a adição, pois que $\sum_{i=1}^{m} \alpha_i \vec{x}_i + \sum_{i=1}^{m} \beta_i \vec{x}_i = \sum_{i=1}^{m} (\alpha_i + \beta_i)\vec{x}_i$, e é igualmente fechado para a multiplicação escalar, visto que $\beta \sum_{i=1}^{m} \alpha_i \vec{x}_i = \sum_{i=1}^{m} (\beta\alpha_i)\vec{x}_i$. Isto significa que A é subespaço de E. O subespaço A contém, obviamente, todos os $\vec{x}_i$ já que $\vec{x}_i = \sum_{j=1}^{m} \delta_{ij}\vec{x}_j$ [7]. Por último, se F é um subespaço de E contendo os $\vec{x}_i$, então, por (1.38), F conterá todas as combinações lineares dos $\vec{x}_i$, o que significa que $A \subset F$, terminando assim demonstração. ☐

O subespaço A representa, assim, o *menor* subespaço de E contendo todos os vectores $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$ ou ainda a *intersecção* de todos os subespaços de E contendo aqueles vectores. Isto leva-nos a pôr a seguinte

Definição 1.5 – Subespaço gerado por uma lista de vectores – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ uma lista de $m \geq 0$ vectores de E. O subespaço A de E formado por todas as combinações lineares dos vectores de x é chamado* ***subespaço gerado*** *por x ou* ***envolvente linear*** *de x e designa-se por $L_{\mathbb{K}}(x)$, $L_{\mathbb{K}}(\vec{x}_i)_{1\leq i\leq m}$ ou por $L_{\mathbb{K}}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ ou ainda, quando se subentende o corpo, $L(x)$, $L(\vec{x}_i)_{1\leq i\leq m}$ ou por $L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$*[8]*:*

$$\begin{cases} L(\emptyset) = L() = \{\vec{o}\}, & \text{se } m = 0 \\ L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) = \left\{\vec{y} \in E\colon \underset{(\alpha_i)\in\mathbb{K}^m}{\exists} \vec{y} = \sum_{i=1}^{m} \alpha_i \vec{x}_i\right\}, & \text{se } m > 0 \end{cases} \tag{1.40}$$

Da definição anterior e da proposição 1.7 resultam, para qualquer subespaço F de E contendo os vectores $(\vec{x}_i)_{1\leq i\leq m}$, as relações:

---

7 δ_{ij} é o *símbolo de Kronecker*, definido por $\delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$

Kronecker, Leopold: matemático alemão (Liegnitz 1823 – Berlim 1891).

8 Também se utilizam as notações $< x >$ e $< \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m >$.

$$\begin{aligned} &\underset{1\le i\le m}{\forall}\ \vec{x}_i \in L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) \\ &L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) \subset F \end{aligned} \tag{1.41}$$

Pode facilmente ser verificado que $L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ é, afinal, a soma (no sentido da definição 1.12) dos subespaços $\mathbb{K}\vec{x}_i$, mencionados no exemplo 1.23 (os $\mathbb{K}\vec{x}_i$ são *rectas* de E, se $\vec{x}_i \neq \vec{o}$):

$$L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) = \sum_{i=1}^{m} (\mathbb{K}\vec{x}_i) \tag{1.42}$$

Quando $L(x)$ é o próprio espaço vectorial E, diz-se que x é uma sequência ou lista *geradora* (de E) ou ainda que x constitui um *sistema de geradores de* E:

Definição 1.6 – Sistema de geradores – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e x uma lista de $m \geq 0$ vectores de E. Se $L(x) = E$, diremos que x é uma lista* ***geradora*** *(de E) ou um* ***sistema de geradores*** *de E.*

$$x \text{ é um sistema de (vectores) geradores de } E \ \Leftrightarrow L(x) = E \tag{1.43}$$

A definição anterior equivale, obviamente, a dizer que todo o vector $\vec{y} \in E$ é combinação linear dos vectores $(\vec{x}_i)_{i\in I}$:

$$x \text{ é um sistema de geradores de } E \ \Leftrightarrow \underset{\vec{y}}{\forall} \left(\vec{y} \in E \Rightarrow \underset{(\alpha_i)\in\mathbb{K}^m}{\exists} \vec{y} = \sum_{i\in I} \alpha_i \vec{x}_i \right) \tag{1.44}$$

Caso exista alguma lista geradora de E (necessariamente de comprimento m *finito*), diremos que o espaço vectorial E é do *tipo finito* ou de *dimensão finita*). A seguir, formaliza-se esta definição:

Definição 1.7 – Espaço de dimensão finita – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$. Diz-se que o espaço E é do* ***tipo finito*** *(ou de* ***dimensão finita****) sse existe alguma sequência $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ de vectores de E geradora de E.*

Os espaços que não são gerados por qualquer lista $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ de vectores seus dizem-se de *dimensão infinita*.

Exemplo 1.37 As sequências $\emptyset$, $(\vec{o})$, $(\vec{o}, \vec{o})$, etc geram, todas elas, o espaço trivial $\{\vec{o}\}$.

Exemplo 1.38 A lista de n vectores de $\mathbb{K}^n$

$$c = ((1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1))$$

gera $\mathbb{K}^n$, visto que todo o vector $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ é combinação linear dos vectores de c:

$$(x_1, x_2, \ldots, x_n) = x_1(1, 0, \ldots, 0) + x_2(0, 1, \ldots, 0) + \cdots + x_n(0, 0, \ldots, 1)$$

Exemplo 1.39 No espaço $\mathcal{P}_n(\mathbb{K})$ dos polinómios de grau menor ou igual a n e coeficientes em $\mathbb{K}$, a sequência de $n + 1$ polinómios

$$c = \big(1, x, x^2, \ldots, x^n\big)$$

gera $\mathscr{P}_n(\mathbb{K})$, pois tem-se, para qualquer polinómio $p \in \mathscr{P}_n(\mathbb{K})$:

$$p(x) = \sum_{k=1}^{n} a_k x^k$$

Exemplo 1.40 O espaço $\mathscr{P}(\mathbb{K})$ de todos os polinómios de coeficientes em $\mathbb{K}$ não é gerado por qualquer lista de polinómios seus, visto que, se a lista for vazia, gera apenas o polinómio nulo e, se não o for, não pode, obviamente, gerar os polinómios cujo grau seja superior ao máximo dos graus dos polinómios presentes na lista. Isto significa que $\mathscr{P}(\mathbb{K})$ tem dimensão infinita.

Exemplo 1.41 A lista $u = ((1,-1,1),(2,1,1),(3,0,2))$ não é geradora de $\mathbb{R}^3$, visto que

$$\begin{aligned} L(u) &= \{(x,y,z)\colon (x,y,z) = a(1,-1,1) + b(2,1,1) + c(3,0,2);\ a,b,c \in \mathbb{R}\} \\ &= \{(x,y,z)\colon (x,y,z) = (a+2b+3c, -a+b, a+b+2c)\} \\ &= \{(x,y,z)\colon 2x - y - 3z = 0\} \neq \mathbb{R}^3 \end{aligned}$$

Geometricamente, $L(u)$ representa um plano passando pela origem.

Exemplo 1.42 Para qualquer $m \in \mathbb{N}$, a sequência $u_m = ((u_{1n}),(u_{2n}),\ldots,(u_{mn}))$ de sucessões (u_{kn}) de escalares de um corpo $\mathbb{K}$ definidas a seguir não pode gerar $\mathbb{K}^{\mathbb{N}}$, visto que nenhuma sucessão com algum termo não nulo e de ordem superior a m pode ser combinação linear das sucessões $(u_{kn})_{\substack{1\le k\le m\\ n\in\mathbb{N}}}$:

$$\begin{cases} (u_{1n}) = (1,0,\ldots,0,\ldots) \\ (u_{2n}) = (0,1,\ldots,0,\ldots) \\ \cdots \\ (u_{mn}) = (0,0,\ldots,1,\ldots) \end{cases}$$

Adiante veremos que o espaço $\mathbb{K}^{\mathbb{N}}$ tem também dimensão infinita.

Exemplo 1.43 No espaço vectorial real $\mathbb{R}^{\mathbb{R}}$ das funções de $\mathbb{R}$ em $\mathbb{R}$, as funções $x \mapsto (\cos x)^n$ pertencem ao subespaço de $\mathbb{R}^{\mathbb{R}}$ gerado por $(1, \cos x, \cos(2x), \ldots, \cos(nx))$. Isto resulta da fórmula de *Euler*[(9)] e do binómio de *Newton*[(10)], que permitem concluir que (ver igualdades C.57 e C.58 no apêndice C):

$$\begin{aligned} (\cos x)^{2p} &= \tfrac{1}{2^{2p}}\binom{2p}{p} + \tfrac{1}{2^{2p-1}}\sum_{k=0}^{p-1}\tbinom{2p}{k}\cos\big(2(p-k)x\big) \\ (\cos x)^{2p-1} &= \tfrac{1}{2^{2p-2}}\sum_{k=0}^{p-1}\tbinom{2p-1}{k}\cos\big((2(p-k)-1)x\big) \end{aligned}$$

Exemplo 1.44 Se E for um espaço vectorial sobre um corpo $\mathbb{K}$ e se $\mathbb{K}' \subset \mathbb{K}$ for subcorpo de $\mathbb{K}$, uma lista $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ de vectores de E gera dois subespaços: o subespaço $L_{\mathbb{K}}(x)$, para combinações lineares com escalares no corpo $\mathbb{K}$

---

9 *Euler, Leonhard*: matemático suíço (Bâle 1707 – São Petersburgo 1783).

10 *Newton, Sir Isaac*: físico, matemático e astrónomo inglês (Woolsthorpe 1642 – Kensington 1727).

$$L_{\mathbb{K}}(x) = \left\{ \vec{u} = \sum_{i=1}^{m} u_i \vec{x}_i \wedge u_i \in \mathbb{K} \right\} \subset E$$

e o subespaço $L_{\mathbb{K}'}(x)$, quando os escalares nas combinações lineares estão em $\mathbb{K}'$

$$L_{\mathbb{K}'}(x) = \left\{ \vec{u} = \sum_{i=1}^{m} u_i \vec{x}_i \wedge u_i \in \mathbb{K}' \right\} \subset E$$

Se $\mathbb{K}' \neq \mathbb{K}$, estes dois subespaços de E não são iguais, tendo-se obviamente

$$L_{\mathbb{K}'}(x) \subset L_{\mathbb{K}}(x)$$

É ainda evidente que $L_{\mathbb{K}'}(x)$ é espaço sobre $\mathbb{K}'$, mas não sobre $\mathbb{K}$, visto que não é fechado para o produto por escalar (de $\mathbb{K}$). Quanto a $L_{\mathbb{K}}(x)$, será espaço vectorial sobre $\mathbb{K}$ (e também sobre $\mathbb{K}'$).

1.6 Independência linear e bases. Dimensão

Na presente secção, vamos definir a importante noção de *independência linear*, bem como as noções de *base* e *dimensão* de um espaço vectorial. Comecemos pela primeira:

Definição 1.8 – Dependência/independência linear – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$, $m \in \mathbb{N}$ e $x = (\vec{x}_i)_{1\leq i\leq m}$ uma sequência de vectores de E. Diz-se que x é* ***linearmente dependente*** *(sobre $\mathbb{K}$) sse existe um vector $\vec{\alpha} = (\alpha_i)_{1\leq i\leq m} \in \mathbb{K}^m \setminus \{\vec{o}\}$ tal que $\sum_{i=1}^{m} \alpha_i \vec{x}_i = \vec{o}$:*

$$x = (\vec{x}_i)_{1\leq i\leq m} \text{ é linearmente dependente } \Leftrightarrow \underset{\vec{\alpha}\in\mathbb{K}^m}{\exists} \left(\vec{\alpha} \neq \vec{o} \wedge \sum_{i=1}^{m} \alpha_i \vec{x}_i = \vec{o} \right) \quad (1.45)$$

Se a lista x de vectores não for linearmente dependente, diremos que é ***linearmente independente*** *(sobre $\mathbb{K}$). Tem-se, portanto:*

$$x = (\vec{x}_i)_{1\leq i\leq m} \text{ é linearmente independente } \Leftrightarrow \underset{\vec{\alpha}\in\mathbb{K}^m}{\forall} \left(\sum_{i=1}^{m} \alpha_i \vec{x}_i = \vec{o} \Rightarrow \vec{\alpha} = \vec{o} \right) \quad (1.46)$$

Por definição, diremos que a lista $\emptyset$ vazia é linearmente independente.

Observações:

- Observe-se aqui que a dependência linear de uma lista de vectores *depende do corpo $\mathbb{K}$ de escalares* que se considere:

Se E é espaço vectorial sobre $\mathbb{K}$ e se $\mathbb{K}' \subset \mathbb{K}$ é subcorpo de $\mathbb{K}$, uma lista $x = (\vec{x}_i)_{1\leq i\leq m}$ de m vectores de E, x pode ser linearmente *dependente* sobre $\mathbb{K}$ e linearmente *independente* sobre $\mathbb{K}'$. O que se pode afirmar, com generalidade, é

$$\mathbb{K}' \subset \mathbb{K} \wedge x \text{ é linearmente dependente sobre } \mathbb{K}' \Rightarrow x \text{ é linearmente dependente sobre } \mathbb{K}.$$

e isto equivale a

$$\mathbb{K}' \subset \mathbb{K} \wedge x \text{ é linearmente independente sobre } \mathbb{K} \Rightarrow x \text{ é linearmente independente sobre } \mathbb{K}'.$$

Por exemplo, a lista $x = (\vec{x}_1, \vec{x}_2) = ((1+i, i), (1-i, 1))$ de vectores de $\mathbb{C}^2$ é linearmente dependente sobre $\mathbb{C}$ (observe que $\vec{x}_1 = i\vec{x}_2$), mas é linearmente independente sobre $\mathbb{R}$, pois se $a, b \in \mathbb{R}$,

$$a(1+i, i) + b(1-i, 1) = (0,0) \Rightarrow ((a+b) + i(a-b), b + ia) = (0,0) \Rightarrow a = b = 0$$

no entanto, qualquer lista $\mathbb{R}$-dependente é também $\mathbb{C}$-dependente e toda a lista $\mathbb{C}$-independente será igualmente $\mathbb{R}$-independente.

- Consideremos uma lista linearmente independente de p vectores de $\mathbb{K}^n$

$$x = ((x_{11}, x_{21}, \ldots, x_{n1}), (x_{12}, x_{22}, \ldots, x_{n2}), \ldots, (x_{1p}, x_{2p}, \ldots, x_{np}))$$

Nestas condições, quaisquer que sejam os escalares $x_{n+1,1}, x_{n+1,2}, \ldots, x_{n+1,p}$, a lista x' de p vectores de $\mathbb{K}^{n+1}$ obtida de x juntando cada um destes escalares aos seus vectores

$$x' = ((x_{11}, x_{21}, \ldots, x_{n1}, x_{n+1,1}), (x_{12}, x_{22}, \ldots, x_{n2}, x_{n+1,2}), \ldots, (x_{1p}, x_{2p}, \ldots, x_{np}, x_{n+1,p}))$$

é ainda linearmente independente, o que resulta de ser $\sum_{k=1}^{p} \alpha_k \vec{x}'_k = \vec{o}$ equivalente a $\sum_{k=1}^{p} \alpha_k \vec{x}_k = \vec{o}$ e $\sum_{k=1}^{p} \alpha_k x_{n+1,k} = 0$.

Se uma sequência x de vectores é linearmente independente, todo o vector de $L(x)$ é combinação linear dos vectores de x e esta combinação linear é *única*, o que se prova na seguinte

Proposição 1.8 – Unicidade da combinação linear – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e $x = (\vec{x}_i)_{1 \le i \le m}$ uma lista de m vectores de E. A sequência x é linearmente independente sse, para cada $\vec{y} \in L(x)$, existe um e um só $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbb{K}^m$ tal que*

$$\vec{y} = \sum_{i=1}^{m} \alpha_i \vec{x}_i$$

Observe-se que, se $\vec{y} = \vec{o}$, então $\vec{\alpha} = (0, 0, \ldots, 0) = \vec{o}$.

Demonstração:

Suponhamos que $\vec{y}$ é um vector de $L(x)$ tal que

$$\vec{y} = \sum_{i=1}^{m} \alpha_i \vec{x}_i = \sum_{i=1}^{m} \beta_i \vec{x}_i$$

daqui se segue que

$$\sum_{i=1}^{m}(\alpha_i-\beta_i)\vec{x}_i=\vec{o}$$

Sendo a lista $(\vec{x}_i)_{1\le i\le m}$ linearmente independente, a igualdade anterior implica $\alpha_i-\beta_i=0$, o que equivale a $\alpha_i=\beta_i$, para todo o $i\in\{1,2,\ldots,m\}$, sendo portanto única a combinação linear que exprime $\vec{y}$ nos $\vec{x}_i$. Reciprocamente, se x for linearmente dependente, existirá $\vec{\beta}=(\beta_i)\neq\vec{o}$ tal que

$$\sum_{i=1}^{m}\beta_i\vec{x}_i=\vec{o}$$

então, se $\vec{y}\in L(x)$, existirá $\vec{\alpha}=(\alpha_1,\alpha_2,\ldots,\alpha_m)\in\mathbb{K}^m$ tal que

$$\vec{y}=\sum_{i=1}^{m}\alpha_i\vec{x}_i=\sum_{i=1}^{m}\alpha_i\vec{x}_i+\vec{o}=\sum_{i=1}^{m}\alpha_i\vec{x}_i+\sum_{i=1}^{m}\beta_i\vec{x}_i=\sum_{i=1}^{m}(\alpha_i+\beta_i)\vec{x}_i$$

Basta, agora, observar que os escalares $\alpha_i+\beta_i$ não são todos iguais aos α_i respectivos, visto que $\vec{\beta}=(\beta_i)\neq\vec{o}$. Portanto, $\vec{y}$ exprime-se de duas formas diferentes como combinação linear dos $\vec{x}_i$, nomeadamente, $\sum\limits_{i=1}^{m}\alpha_i\vec{x}_i$ e $\sum\limits_{i=1}^{m}(\alpha_i+\beta_i)\vec{x}_i$. $\square$

A proposição anterior garante, portanto, que se $x=(\vec{x}_i)_{1\le i\le m}$ for uma lista *linearmente independente* de vectores de E, todo o vector $\vec{y}\in E$ se exprime *quando muito de uma forma* como combinação linear dos $\vec{x}_i$ (de uma e uma só forma quando $\vec{y}\in L(x)$ e de nenhuma, quando $\vec{y}\notin L(x)$). Para garantir que todo o vector de E se possa exprimir nos $\vec{x}_i$ será necessário que $L(x)=E$ e, portanto, que x seja também geradora E. Isto leva-nos à importante noção de *Base*:

Definição 1.9 – Base de um espaço de dimensão finita – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e $x=(\vec{x}_i)_{1\le i\le n}$ uma lista de n vectores de E . Diremos que x é uma* ***base*** *de E sse x gera E e x é linearmente independente:*

$$x \text{ é uma } \textbf{base} \text{ de } E \Leftrightarrow L(x)=E\wedge x \text{ é linearmente independente} \tag{1.47}$$

Em virtude da proposição 1.8, se $x=(\vec{x}_i)_{1\le i\le n}$ é uma base de E, todo o vector $\vec{y}$ de E se exprime de *uma e uma só* forma como combinação linear $\sum\limits_{i=1}^{n}\alpha_i\vec{x}_i$ dos vectores de x. Os escalares α_i únicos envolvidos em tal combinação linear são chamados as *coordenadas (contravariantes) de $\vec{y}$ em relação à base x* e o vector $\vec{\alpha}=(\alpha_1,\alpha_2,\ldots,\alpha_n)\in\mathbb{K}^n$ é o *vector das coordenadas* de $\vec{y}$ ou a *representação* de $\vec{y}$, relativa à base x.

De seguida, vamos enunciar e provar um certo número de proposições envolvendo os conceitos de *subespaço gerado*, *independência linear* e *base*. Uma delas permitir-nos-á definir rigorosamente a noção de *dimensão* de um espaço vectorial, como sendo o número de vectores de uma qualquer base desse espaço. Outras três de entre elas (as *operações elementares*) serão essenciais para fundamentar o Algoritmo de Condensação, de grande utilidade em vários temas da Álgebra Linear.

Proposição 1.9 *Uma sequência $x = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m)$ de vectores de um espaço vectorial E sobre um corpo $\mathbb{K}$ é linearmente dependente sse existir nela um vector $\vec{x}_k$ que seja combinação linear dos* ***restantes****:*

$$(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m) \text{ é linearmente dependente} \Leftrightarrow \underset{\substack{1\le k\le m \\ \alpha_i\in\mathbb{K}}}{\exists} \vec{x}_k = \sum_{i\neq k} \alpha_i \vec{x}_i \tag{1.48}$$

Demonstração:

A condição é suficiente, pois

$$\underset{\substack{1\le k\le m \\ \alpha_i\in\mathbb{K}}}{\exists} \vec{x}_k = \sum_{i\neq k} \alpha_i \vec{x}_i \Rightarrow \underset{\substack{1\le k\le m \\ \alpha_i\in\mathbb{K}}}{\exists} \sum_{i\neq k} \alpha_i \vec{x}_i - \vec{x}_k = \vec{o}$$

mas esta última igualdade mostra o anulamento de uma combinação linear dos vectores de x, na qual $\alpha_k = -1 \neq 0$, o que implica a dependência linear de x. Reciprocamente, de x for linearmente dependente, existe $\alpha_k \neq 0$ tal que

$$\sum_{i=1}^{m} \alpha_i \vec{x}_i = \sum_{i\neq k} \alpha_i \vec{x}_i + \alpha_k \vec{x}_k = \vec{o}$$

daqui se segue facilmente que

$$\vec{x}_k = \sum_{i\neq k} -\frac{\alpha_i}{\alpha_k} \vec{x}_i \qquad \square$$

Os seguintes corolários da proposição anterior são óbvios:

Corolário 1.9.1 *Uma sequência $(\vec{x}_1)$ formada por um só vector é linearmente dependente sse $\vec{x}_1 = \vec{o}$:*

$$(\vec{x}_1) \text{ é linearmente dependente} \Leftrightarrow \vec{x}_1 = \vec{o} \tag{1.48.1}$$

Corolário 1.9.2 *Uma sequência $(\vec{x}_1, \vec{x}_2)$ formada por dois vectores é linearmente dependente sse um desses vectores for múltiplo escalar do outro, isto é:*

$$(\vec{x}_1, \vec{x}_2) \text{ é linearmente dependente} \Leftrightarrow \underset{\alpha\in\mathbb{K}}{\exists} \vec{x}_1 = \alpha\vec{x}_2 \vee \underset{\beta\in\mathbb{K}}{\exists} \vec{x}_2 = \beta\vec{x}_1 \tag{1.48.2}$$

Corolário 1.9.3 *Se algum vector numa sequência for nulo, a sequência é linearmente dependente.*

Corolário 1.9.4 *Se, numa sequência, existirem dois vectores iguais, a sequência é linearmente dependente.*

Corolário 1.9.5 *Se, numa sequência, um dos vectores for múltiplo escalar de outro, a sequência é linearmente dependente.*

Corolário 1.9.6 *Se, numa sequência, existir uma subsequência linearmente dependente, então a sequência será linearmente dependente.*

Corolário 1.9.7 *Toda a subsequência de uma sequência linearmente independente será, também, linearmente independente.*

Proposição 1.10 *Uma sequência $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ de vectores de um espaço vectorial E sobre um corpo $\mathbb{K}$ é linearmente dependente sse existir nela um vector $\vec{x}_k$ que seja combinação linear dos* ***anteriores****:*

$$(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) \text{ é linearmente dependente} \Leftrightarrow \underset{\substack{1\le k\le m \\ \alpha_i\in\mathbb{K}}}{\exists} \vec{x}_k = \sum_{i<k} \alpha_i \vec{x}_i \tag{1.49}$$

Demonstração:

A condição é obviamente suficiente, se atendermos a que $\vec{x}_k = \sum\limits_{i<k} \alpha_i \vec{x}_i = \sum\limits_{i<k} \alpha_i \vec{x}_i + \sum\limits_{i>k} 0\vec{x}_i$ e à proposição 1.9 anterior. Reciprocamente, sendo x linearmente dependente, existe pelo menos um $k \in \{1, 2, \ldots, m\}$ tal que $\alpha_k \neq 0$ e $\sum\limits_{i=1}^{m} \alpha_i \vec{x}_i = \vec{o}$. Portanto, o conjunto

$$A = \{i \colon \alpha_i \neq 0\} \subset \{1, 2, \ldots, m\}$$

é finito e não vazio e, portanto, tem um máximo $k = \max A$. Devido à definição de k, tem-se $\alpha_i = 0$ para $i > k$ e $\alpha_k \neq 0$. Então, fica simplesmente

$$\sum_{i<k} \alpha_i \vec{x}_i + \alpha_k \vec{x}_k = \vec{o}$$

donde se conclui

$$\vec{x}_k = \sum_{i<k} -\frac{\alpha_i}{\alpha_k}\, \vec{x}_i$$

Observe-se que, se $k = 1$, fica $\alpha_1 \vec{x}_1 = \vec{o}$ e $\alpha_1 \neq 0$, donde $\vec{x}_1 = \vec{o}$: mesmo assim, poderemos dizer que $\vec{x}_1$ é combinação linear (vazia) dos vectores anteriores. $\square$

A proposição anterior é equivalente a dizer que uma sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ é *linearmente independente* sse nenhum dos seus vectores for combinação linear dos anteriores (isto é, $\vec{x}_1 \neq \vec{o}$ e nenhum dos $\vec{x}_k$, com $k > 1$, é combinação linear dos anteriores).

Exemplo 1.45 A sequência $((0,0,0,3), (0,0,1,-1), (0,-1,-4,0), (2,0,1,-3))$ de vectores de $\mathbb{R}^4$ é linearmente independente, visto que nenhum dos seus vectores é combinação linear dos anteriores (isto sucederá igualmente com qualquer lista de m vectores *não nulos* de $\mathbb{K}^n$, cujas primeiras componentes se anulem, mas diminuindo sempre o número de 0's nessas primeiras componentes de cada vector para o seguinte da lista).

Proposição 1.11 *Uma sequência* $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ *de vectores de um espaço vectorial* E *sobre um corpo* $\mathbb{K}$ *é linearmente dependente sse existir nela um vector* $\vec{x}_k$ *que seja combinação linear dos* ***seguintes****:*

$$(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) \text{ é linearmente dependente} \Leftrightarrow \underset{\substack{1\le k\le m\\ \alpha_i\in\mathbb{K}}}{\exists} \vec{x}_k = \sum_{i>k} \alpha_i \vec{x}_i \tag{1.50}$$

Demonstração:

A condição é obviamente suficiente, se atendermos a que $\vec{x}_k = \sum_{i>k} \alpha_i \vec{x}_i = \sum_{i<k} 0\vec{x}_i + \sum_{i>k} \alpha_i \vec{x}_i$ e à proposição 1.9. Reciprocamente, sendo x linearmente dependente, existe pelo menos um $k \in \{1, 2, \ldots, m\}$ tal que $\alpha_k \neq 0$ e $\sum_{i=1}^{m} \alpha_i \vec{x}_i = \vec{o}$. Portanto, o conjunto

$$A = \{i \colon \alpha_i \neq 0\} \subset \{1, 2, \ldots, m\}$$

é finito e não vazio e, portanto, tem um mínimo $k = \min A$. Devido à definição de k, tem-se $\alpha_i = 0$ para $i < k$ e $\alpha_k \neq 0$. Então, fica simplesmente

$$\alpha_k \vec{x}_k + \sum_{i>k} \alpha_i \vec{x}_i = \vec{o}$$

donde se conclui

$$\vec{x}_k = \sum_{i>k} -\frac{\alpha_i}{\alpha_k} \vec{x}_i$$

Observe-se que, se $k = m$, fica $\alpha_m \vec{x}_m = \vec{o}$ e $\alpha_m \neq 0$, donde $\vec{x}_m = \vec{o}$: mesmo assim, poderemos dizer que $\vec{x}_m$ é combinação linear (vazia) dos vectores seguintes. ☐

A proposição anterior é equivalente a dizer que uma sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ é *linearmente independente* sse nenhum dos seus vectores for combinação linear dos seguintes (isto é, $\vec{x}_m \neq \vec{o}$ e nenhum dos $\vec{x}_k$, com $k < m$, é combinação linear dos vectores seguintes).

Exemplo 1.46 A sequência $((2, 0, 1, -3), (0, 0, -4, 0), (0, 0, 0, -1))$ de vectores de $\mathbb{R}^4$ é linearmente independente, visto que nenhum dos seus vectores é combinação linear dos seguintes. Isto sucederá igualmente com qualquer lista de $m \leq n$ vectores não nulos de $\mathbb{K}^n$, cujas primeiras componentes se anulem, mas aumentando sempre o número de 0's dessas primeiras componentes de cada vector para o seguinte da lista: tais listas dizem-se *escalonadas*; segue-se a definição:

Definição 1.10 – Lista escalonada em $\mathbb{K}^n$ **–** Consideremos uma lista

$$x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$$

com

$$\begin{cases} \vec{x}_1 = (x_{11}, x_{12}, \dots, x_{1n}) \\ \vec{x}_2 = (x_{21}, x_{22}, \dots, x_{2n}) \\ \dots \\ \vec{x}_m = (x_{m1}, x_{m2}, \dots, x_{mn}) \end{cases}$$

de $m > 0$ vectores de $\mathbb{K}^n$. Diremos que a lista x está na forma **escalonada** sse se der uma das condições seguintes:

i) $\vec{x}_1 = \vec{x}_2 = \dots = \vec{x}_m = \vec{o}$

ii) *Existe um r, com $1 \le r \le m$, tal que*

$$\begin{cases} i \le r \Rightarrow \vec{x}_i \ne \vec{o} \\ i > r \Rightarrow \vec{x}_i = \vec{o} \end{cases}$$

e, além disso, a sequência k_i, definida para $i \in \{1, 2, \dots, r\}$, por

$$k_i = \min\{j : x_{ij} \ne 0\}$$

é estritamente crescente, isto é,

$$1 \le k_1 < k_2 < \dots < k_r \le n$$

Portanto, a lista do exemplo anterior é escalonada e são-no também, por exemplo, as listas

$((0, 0, 1, -3), (0, 0, 0, 2), (0, 0, 0, 0))$, em $\mathbb{R}^4$
$((0, 3, 1, -3, 5), (0, 0, -4, 0, 2), (0, 0, 0, -1, 1), (0, 0, 0, 0, -1))$, em $\mathbb{R}^5$
$((1 - 2i, 0, 1, -2i), (0, 2 + i, 0, -2 - i), (0, 0, 0, 3 - 2i), (0, 0, 0, 0), (0, 0, 0, 0))$, em $\mathbb{C}^4$

Da proposição 1.11, resulta imediatamente o seguinte

Corolário 1.11.1 *Toda a lista* ***escalonada*** *de vectores* ***não nulos*** *de $\mathbb{K}^n$ é linearmente independente.*

Demonstração:

Nas condições do enunciado, nenhum dos vectores da lista pode ser combinação linear dos seguintes e, portanto, a lista é linearmente independente. □

Proposição 1.12 *Seja $x_m = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m)$ uma lista linearmente independente de m vectores de um espaço vectorial E sobre um corpo $\mathbb{K}$ e $\vec{x}_{m+1}$ um vector de $E \setminus L(x_m)$ (portanto, x_m não é geradora de E, caso contrário, não existiria um tal vector $\vec{x}_{m+1}$). Nestas condições, a lista de $m + 1$ vectores $x_{m+1} = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m, \vec{x}_{m+1})$ é ainda linearmente independente e gera um subespaço $L(x_{m+1})$ de E contendo estritamente $L(x_m)$:*

$$\begin{gathered} x_m \text{ é linear/ independente} \wedge \vec{x}_{m+1} \in E \setminus L(x_m) \Rightarrow x_{m+1} \text{ é linear/ independente} \\ \vec{x}_{m+1} \in E \setminus L(x_m) \Rightarrow L(x_m) \subset L(x_{m+1}) \wedge L(x_m) \ne L(x_{m+1}) \end{gathered} \tag{1.51}$$

Demonstração:

A igualdade $\sum_{i=1}^{m} \alpha_i \vec{x}_i = \sum_{i=1}^{m} \alpha_i \vec{x}_i + 0\vec{x}_{m+1}$ mostra que toda a combinação linear de vectores de x_m é combinação linear dos vectores de x_{m+1}, logo: $L(x_m) \subset L(x_{m+1})$. A hipótese de que $\vec{x}_{m+1} \notin L(x_m)$ (mas obviamente que $\vec{x}_{m+1} \in L(x_{m+1})$) mostra, por outro lado, que $L(x_m) \neq L(x_{m+1})$.

Por fim, para provar que x_{m+1} é linearmente independente, atendamos à definição e ponhamos

$$\sum_{i=1}^{m} \alpha_i \vec{x}_i + \beta \vec{x}_{m+1} = \vec{o}$$

Esta igualdade implica que $\beta = 0$ (se assim não fosse, tinha-se $\vec{x}_{m+1} = \sum_{i=1}^{m} (-\alpha_i/\beta)\vec{x}_i$ ou seja, $\vec{x}_{m+1} \in L(x_m)$, contrariamente à hipótese!). Portanto, na igualdade anterior fica apenas

$$\sum_{i=1}^{m} \alpha_i \vec{x}_i = \vec{o}$$

e, agora, a independência linear de x_m mostra que terá que ser, igualmente, $\alpha_i = 0$, para todo o $i = 1, \ldots, m$. □

Como consequência da proposição anterior, temos o seguinte corolário que mostra que, num espaço de dimensão *infinita*, existem listas de vectores *linearmente independentes* com tantos vectores quantos queiramos:

Corolário 1.12.1 *Seja E um espaço de dimensão **infinita** sobre um corpo $\mathbb{K}$. Para todo o $n \in \mathbb{N}$ existe uma sequência de n vectores $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ linearmente independente.*

Demonstração:

Por indução em n: é claro que $E \neq \{\vec{o}\}$; portanto e para $n = 1$, é linearmente independente a sequência $x_1 = (\vec{x}_1)$, onde $\vec{x}_1 \neq \vec{o}$ é um vector não nulo arbitrário em E.

Se for $x_n = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ uma sequência linearmente independente com n vectores, como $E \setminus L(x_n) \neq \emptyset$ (visto que E tem dimensão infinita, x_n não pode gerar E), podemos escolher um vector $\vec{x}_{n+1} \in E \setminus L(x_n)$ e a proposição anterior permite concluir que $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n, \vec{x}_{n+1})$ é ainda linearmente independente e com $n + 1$ vectores. □

Mais adiante (proposição 1.17), veremos que a recíproca deste corolário é, igualmente, verdadeira. Vamos, agora, ver que, se o vector que se junta a uma lista de vectores pertencer ao subespaço que ela gera, a nova lista gera ainda esse subespaço e é *linearmente dependente*:

Lema 1.13.1 *Seja $x_m = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ uma lista de m vectores de um espaço vectorial E sobre um certo corpo $\mathbb{K}$ e $\vec{x}_{m+1} \in L(x_m)$ um vector qualquer do subespaço gerado por x_m. Então, a lista $x_{m+1} = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m, \vec{x}_{m+1})$ gera o mesmo subespaço que x_m gera e é* *linearmente dependente:*

$$\begin{aligned}&\vec{x}_{m+1} \in L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) \Rightarrow (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m, \vec{x}_{m+1}) \text{ é linearmente dependente}\\ &\vec{x}_{m+1} \in L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) \Rightarrow L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m, \vec{x}_{m+1}) = L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)\end{aligned} \tag{1.52}$$

Demonstração:

Visto ser $\vec{x}_{m+1} \in L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$, a proposição 1.9 mostra que x_{m+1} é linearmente dependente. Por outro lado, a igualdade $\sum_{i=1}^{m} \alpha_i \vec{x}_i = \sum_{i=1}^{m} \alpha_i \vec{x}_i + 0\vec{x}_{m+1}$ mostra que

$$L(x_m) \subset L(x_{m+1})$$

Para mostrar a inclusão inversa, basta atender a que $\vec{x}_{m+1} \in L(x_m)$:

$$\vec{z} \in L(x_{m+1}) \Rightarrow \vec{z} = \sum_{i=1}^{m+1} \alpha_i \vec{x}_i = \sum_{i=1}^{m} \alpha_i \vec{x}_i + \alpha_{m+1} \sum_{i=1}^{m} \beta_i \vec{x}_i = \sum_{i=1}^{m} (\alpha_i + \alpha_{m+1}\beta_i)\vec{x}_i$$

a última expressão mostra que $\vec{z} \in L(x_m)$ e que, portanto, $L(x_{m+1}) \subset L(x_m)$. □

Lema 1.13.2 *Seja $x_m = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ uma lista* ***linearmente dependente*** *de m vectores de um espaço vectorial E sobre um certo corpo $\mathbb{K}$. A proposição 1.10 garante a existência de um vector $\vec{x}_k$ da lista x_m que é combinação linear dos anteriores. A nova lista*

$$x_{m-1} = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{k-1}, \vec{x}_{k+1}, \ldots, \vec{x}_m)$$

de $m-1$ vectores obtida retirando $\vec{x}_k$ à lista x_m gera ainda o mesmo subespaço que era gerado por x_m:

$$\vec{x}_k = \sum_{i<k} \beta_i \vec{x}_i \Rightarrow L(\vec{x}_i)_{i \neq k} = L(\vec{x}_i)$$

Demonstração:

É óbvio que $L(x_{m-1}) \subset L(x_m)$. Quanto à inclusão inversa, tem-se, como no lema anterior:

$$\vec{z} \in L(x_m) \Rightarrow \vec{z} = \sum_{i=1}^{m} \alpha_i \vec{x}_i = \sum_{i \neq k} \alpha_i \vec{x}_i + \alpha_k \sum_{i<k} \beta_i \vec{x}_i = \sum_{i<k} (\alpha_i + \alpha_k \beta_i)\vec{x}_i + \sum_{i>k} \alpha_i \vec{x}_i$$

a última expressão mostra que $\vec{z} \in L(x_{m-1})$ e que, portanto, $L(x_m) \subset L(x_{m-1})$. □

É claro que, com demonstração análoga, é também possível retirar duma lista um vector seu que seja combinação linear dos *restantes* ou dos *seguintes* (proposições 1.9 e 1.11) sem que, com isso, se altere o respectivo subespaço gerado.

Proposição 1.13 *Seja E um espaço vectorial de dimensão finita e $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ uma lista geradora de E. Se $y = (\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m)$ for uma lista de vectores de E linearmente independente, então tem-se, necessariamente,*

$$m \leq n \tag{1.53.1}$$

Além disso, o espaço E é gerado por uma lista de n vectores que inclui todos os vectores de y:

$$E = L(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = L(\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m, \vec{x}_{i_1}, \vec{x}_{i_2}, \ldots, \vec{x}_{i_{n-m}}) \tag{1.53.2}$$

Demonstração:

Como x é geradora de E, o lema 1.13.1 permite construir a lista

$$y_1 = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n, \vec{y}_1)$$

de $n+1$ vectores que será ainda geradora, mas linearmente dependente. Recorrendo ao lema 1.13.2, conclui-se que existe em y_1 um vector que é combinação linear dos seguintes e que é possível retirá-lo desta lista sem que se altere o subespaço gerado; mas esse vector não pode ser $\vec{y}_1$, visto que ele não é nulo (Porquê?). Portanto, um dos vectores $\vec{x}_i$ pode ser retirado da lista y_1, obtendo-se uma nova sequência x_1 de n vectores geradora de E. Introduzamos, agora, nesta lista o vector $\vec{y}_2$. De acordo com o lema 1.13.1 a lista de $n+1$ vectores obtida

$$y_2 = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n, \vec{y}_1, \vec{y}_2)$$

é geradora de E e linearmente dependente. Portanto, existirá de novo um vector que é combinação linear dos seguintes e, tal como antes, esse vector não poderá ser nenhum dos $\vec{y}_i$, devido à independência linear destes. Logo, o vector a retirar (lema 1.13.2) terá que ser um dos $\vec{x}_i$, obtendo-se uma nova lista x_2 com n vectores geradora de E. Este processo de aplicação alternada dos lemas 1.13.1 e 1.13.2 leva à construção de listas x_i com n vectores, geradoras de E, e onde os vectores $\vec{x}_i$ vão sendo sucessivamente substituídos pelos $\vec{y}_k$. Se fosse $m > n$ e ao fim de n aplicações de cada um dos lemas referidos, obtinha-se uma lista x_n geradora de E formada pelos n primeiros vectores de y, ficando ainda de fora dessa lista os últimos $m - n > 0$ vectores de y. Sendo essa lista x_n geradora, o vector $\vec{y}_{n+1}$ seria combinação linear dos primeiros n vectores $\vec{y}_k$, o que significa que a lista y seria linearmente dependente, em contradição com a hipótese. Fica, assim, provado que deverá ser $m \leq n$. Portanto, e aplicando m vezes cada um dos lemas anteriores, somos conduzidos a uma lista x_m geradora de E, contendo todos os m vectores $\vec{y}_k$ e ainda $n - m \geq 0$ dos $\vec{x}_i$ (os únicos vectores de x que não foram retirados ao longo do processo), obtendo-se o resultado do enunciado: uma lista geradora de E do tipo $x_m = (\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m, \vec{x}_{i_1}, \vec{x}_{i_2}, \ldots, \vec{x}_{i_{n-m}})$. □

Com base na proposição anterior é fácil provar que, num espaço de dimensão finita, todas as bases têm o mesmo número de vectores, facto que é fundamental para a definição da importante noção de *dimensão* de um espaço vectorial:

Proposição 1.14 – Invariância do comprimento das bases – *Seja E um espaço de dimensão finita e considerem-se duas bases, $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ e $y = (\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m)$ desse espaço. Então, será $m = n$.*

Demonstração:

Basta aplicar a anterior proposição 1.13 duas vezes: x é linearmente independente e y é geradora de E, portanto

$$n \leq m$$

Por outro lado, y é linearmente independente e x é geradora de E, portanto

$$m \leq n$$

Pela propriedade anti-simétrica da relação de igualdade, será $m = n$. □

O resultado anterior garante que todas as bases de um mesmo espaço vectorial E de dimensão finita sobre certo corpo $\mathbb{K}$ são constituídas pelo mesmo número de vectores, o qual representa, assim, um atributo intrínseco do espaço E a que chamaremos a *dimensão* de E (sobre $\mathbb{K}$). Formaliza-se, agora, a definição:

Definição 1.11 – Dimensão de um espaço vectorial – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$. Chama-se* ***dimensão*** *de E (sobre $\mathbb{K}$) ao número de vectores de uma base arbitrária de E e usa-se a notação* $\dim_{\mathbb{K}} E$ *ou, mais simplesmente,* $\dim E$*, se subentendermos o corpo.*

Portanto, para calcular a dimensão de um espaço vectorial, teremos que procurar uma sua base qualquer e verificar o número de vectores que a constitui. Convém observar que a dimensão de um espaço vectorial depende do corpo $\mathbb{K}$, tanto como de E: como se viu, se $\mathbb{K}'$ for um subcorpo de $\mathbb{K}$, E será também espaço sobre $\mathbb{K}'$ mas a dimensão de E sobre $\mathbb{K}'$ será, em geral, maior que a dimensão sobre $\mathbb{K}$, como se ilustra por alguns dos exemplos que seguem:

Exemplo 1.47 O espaço $\{\vec{o}\}$ tem uma única base, a saber, a sequência vazia $\emptyset$. Portanto:

$$\dim\{\vec{o}\} = 0$$

Exemplo 1.48 Em $\mathbb{K}^n$, a sequência de n vectores

$$c = ((1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1))$$

é uma base, visto que, como facilmente se comprova (prove-o, como exercício), é linearmente independente e gera $\mathbb{K}^n$. Nesta base, as coordenadas de um vector $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ são os próprios escalares $x_1, x_2, \ldots, x_n$. É esta a única base de $\mathbb{K}^n$ onde tal sucede, daí a sua importância e o facto de ser designada por base *canónica* de $\mathbb{K}^n$. Podemos, pois, escrever:

$$\dim_{\mathbb{K}} \mathbb{K}^n = n$$

Exemplo 1.49 No espaço $\mathcal{P}_n(\mathbb{K})$ dos polinómios de coeficientes em $\mathbb{K}$ e grau inferior ou igual a n, a lista de $n + 1$ polinómios

$$c_n = \left(1, x, x^2, \ldots, x^n\right)$$

constitui uma base, como facilmente se prova (demonstre!). Neste caso, também é óbvio que as coordenadas de um polinómio são os respectivos coeficientes, o que só acontece com esta base. Por isso é chamada base *canónica* de $\mathcal{P}_n(\mathbb{K})$. Conclui-se, portanto, que:

$$\dim_{\mathbb{K}} \mathcal{P}_n(\mathbb{K}) = n + 1$$

Este exemplo mostra também que, no espaço $\mathcal{P}(\mathbb{K})$, existem listas linearmente independentes com tantos vectores quantos queiramos: por exemplo, as listas c_n com qualquer $n \in \mathbb{N}_0$.

Exemplo 1.50 Em S^1, uma base é constituída por qualquer segmento orientado *não nulo*, tendo-se, portanto:

$$\dim_{\mathbb{R}} S^1 = 1$$

Exemplo 1.51 Em S^2, uma base é constituída por quaisquer dois segmentos orientados *não colineares*. Por isso,

$$\dim_{\mathbb{R}} S^2 = 2$$

Exemplo 1.52 Em S^3, uma base é constituída por quaisquer três segmentos orientados *não coplanares*. Logo:

$$\dim_{\mathbb{R}} S^3 = 3$$

Exemplo 1.53 No espaço $\mathbb{R}^+$ do exemplo 1.11 e sendo $u \in \mathbb{R}^+ \setminus \{1\}$, (u) será uma base já que, se $x \in \mathbb{R}^+$, tem-se:

$$x = u^{\log_u x} = (\log_u x)u$$

o que mostra que u gera $\mathbb{R}^+$ e é, obviamente, linearmente independente pois que

$$\alpha u = u^\alpha = 1 \Rightarrow \alpha = \log_u 1 = 0$$

Neste caso, a *coordenada* de $x \in \mathbb{R}^+$ na base (u) é $\log_u x$. Tem-se, pois, $\dim_{\mathbb{R}} \mathbb{R}^+ = 1$.

Exemplo 1.54 Seja E um espaço vectorial de dimensão n sobre um corpo $\mathbb{K}$, com uma base $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$, e considere-se o espaço E^m (ver exemplo 1.13). A lista de mn vectores de E^m

$$c = \begin{pmatrix} (\vec{e}_1, \vec{o}, \dots, \vec{o}), & (\vec{e}_2, \vec{o}, \dots, \vec{o}), & \cdots & (\vec{e}_n, \vec{o}, \dots, \vec{o}), \\ (\vec{o}, \vec{e}_1, \dots, \vec{o}), & (\vec{o}, \vec{e}_2, \dots, \vec{o}), & \cdots & (\vec{o}, \vec{e}_n, \dots, \vec{o}), \\ \cdots & \cdots & \cdots & \cdots \\ (\vec{o}, \vec{o}, \dots, \vec{e}_1), & (\vec{o}, \vec{o}, \dots, \vec{e}_2), & \cdots & (\vec{o}, \vec{o}, \dots, \vec{e}_n) \end{pmatrix}$$

é uma base de E^m (demonstre!). Portanto:

$$\dim_{\mathbb{K}} E^m = mn = m \dim_{\mathbb{K}} E$$

Da mesma forma e mais geralmente, em relação ao espaço produto, tem-se:

$$\dim_{\mathbb{K}} \left(\prod_{i=1}^{m} E_i \right) = \sum_{i=1}^{m} \dim_{\mathbb{K}} E_i$$

Exemplo 1.55 $\mathbb{C}$ é um espaço vectorial sobre $\mathbb{R}$ e a lista

$$c = (1, i)$$

é uma base deste espaço visto que, para qualquer $z = a + ib \in \mathbb{C}$,

$$z = a \cdot 1 + b \cdot i, \text{ onde } a, b \in \mathbb{R}$$

e, portanto, qualquer vector $z \in \mathbb{C}$ é combinação linear (com escalares reais) dos vectores 1 e i;

logo, $\mathbb{C} = L(1, i)$. Por outro lado, c é linearmente independente, como resulta de

$$a \cdot 1 + b \cdot i = 0 \Rightarrow a = 0 \wedge b = 0$$

Deste modo, tem-se :

$$\dim_{\mathbb{R}} \mathbb{C} = 2$$

É claro que a dimensão de $\mathbb{C}$ sobre $\mathbb{C}$ é 1: $\dim_{\mathbb{C}} \mathbb{C} = 1$.

Exemplo 1.56 No exemplo 1.14, vimos que, se E é um espaço vectorial sobre um corpo $\mathbb{K}$ e $\mathbb{K}' \subset \mathbb{K}$ for um subcorpo de $\mathbb{K}$, E será igualmente espaço vectorial sobre $\mathbb{K}'$; em particular, $\mathbb{K}$ será um espaço vectorial sobre $\mathbb{K}'$. Vamos, agora, ver que, se as dimensões de E sobre $\mathbb{K}$ e de $\mathbb{K}$ sobre $\mathbb{K}'$ forem finitas, o mesmo sucede com a dimensão de E sobre $\mathbb{K}'$ e que esta é igual ao produto das dimensões de E sobre $\mathbb{K}$ e de $\mathbb{K}$ sobre $\mathbb{K}'$:

$$\dim_{\mathbb{K}'} E = \dim_{\mathbb{K}} E \cdot \dim_{\mathbb{K}'} \mathbb{K}$$

Suponha-se que $\dim_{\mathbb{K}} E = n$ e $\dim_{\mathbb{K}'} \mathbb{K} = m$ e sejam

$$e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) \text{ e } k = (k_1, k_2, \ldots, k_m)$$

bases de E sobre $\mathbb{K}$ e de $\mathbb{K}$ sobre $\mathbb{K}'$, respectivamente.

Basta provar que a lista $c = (k_i \vec{e}_j)_{\binom{1 \leq i \leq m}{1 \leq j \leq n}}$ de mn vectores de E é uma base de E sobre $\mathbb{K}'$: de facto, a lista c gera E

$$L(c) = E$$

visto que, para qualquer $\vec{x} \in E$ se tem $\vec{x} = \sum\limits_{j=1}^{n} x_j \vec{e}_j$, onde os $x_j \in \mathbb{K}$ são únicos. Porém, os x_j são "vectores" do espaço $\mathbb{K}$ e exprimir-se-ão na base k deste espaço, de forma única, como combinação linear dos k_i, vindo portanto: $x_j = \sum\limits_{i=1}^{m} x_{ij} k_i$, onde os $x_{ij} \in \mathbb{K}'$; substituindo esta expressão na anterior, vem

$$\vec{x} = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_{ij} k_i \right) \vec{e}_j = \sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij} (k_i \, \vec{e}_j), \text{ em que } x_{ij} \in \mathbb{K}'$$

A igualdade anterior prova que os mn vectores $k_i \, \vec{e}_j$ geram E (sobre $\mathbb{K}'$). Vejamos, agora, a independência linear de c:

$$\sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij} (k_i \, \vec{e}_j) = \vec{o} \Rightarrow \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_{ij} k_i \right) \vec{e}_j = \vec{o} \Rightarrow \underset{1 \leq j \leq n}{\forall} \sum_{i=1}^{m} x_{ij} k_i = 0 \Rightarrow \underset{1 \leq j \leq n}{\forall} \; \underset{1 \leq i \leq m}{\forall} x_{ij} = 0$$

Fica, assim, provado que c é uma base de E sobre $\mathbb{K}'$.

Assim e por exemplo, todo o espaço vectorial *complexo* E é também um espaço vectorial *real* e, no exemplo 1.55, vimos que $\dim_{\mathbb{R}} \mathbb{C} = 2$; estes resultados implicam que:

$$\dim_{\mathbb{R}} E = \dim_{\mathbb{C}} E \cdot \underbrace{\dim_{\mathbb{R}} \mathbb{C}}_{2} = 2 \dim_{\mathbb{C}} E$$

De facto, se $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ for uma base de E sobre $\mathbb{C}$ ($\dim_{\mathbb{C}} E = n$), então a lista de $2n$ vectores

$$c = (\vec{e}_1, i\vec{e}_1, \vec{e}_2, i\vec{e}_2, \ldots, \vec{e}_n, i\vec{e}_n)$$

será uma base de E sobre $\mathbb{R}$, visto que $(1, i)$ é base de $\mathbb{C}$ sobre $\mathbb{R}$: c gera E (através de escalares reais), visto que por hipótese, para qualquer $\vec{x} \in E$ existem complexos $z_k = \alpha_k + i\beta_k$ tais que:

$$\vec{x} = \sum_{k=1}^{n} (\alpha_k + i\beta_k)\vec{e}_k = \sum_{k=1}^{n} (\alpha_k \vec{e}_k + \beta_k(i\vec{e}_k)) = \sum_{k=1}^{n} \alpha_k \vec{e}_k + \sum_{k=1}^{n} \beta_k(i\vec{e}_k)$$

a última expressão dá $\vec{x}$ como combinação linear dos vectores de c, através dos escalares reais α_k e β_k, provando assim que c gera E (usando escalares reais). A sucessão de igualdades anteriores (em sentido inverso) mostra, por fim, que, se uma combinação dos vectores de c é nula, os escalares reais α_k e β_k terão de ser todos nulos, o que significa a independência linear de c.

Exemplo 1.57 Em relação ao espaço complexo E' do exemplo 1.15, se $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ for uma base de E, então a lista de n vectores

$$c = ((\vec{e}_1, \vec{o}), (\vec{e}_2, \vec{o}), \ldots, (\vec{e}_n, \vec{o}))$$

será uma base de $E' = E^2$. De facto, c gera E', visto que para qualquer $(\vec{x}, \vec{y}) \in E'$ se tem:

$$\begin{aligned}(\vec{x}, \vec{y}) &= \left(\sum_{k=1}^{n} x_k \vec{e}_k, \sum_{k=1}^{n} y_k \vec{e}_k\right) = \sum_{k=1}^{n} (x_k \vec{e}_k, y_k \vec{e}_k) \\ &= \sum_{k=1}^{n} (x_k + iy_k)(\vec{e}_k, \vec{o})\end{aligned}$$

onde os $x_k + iy_k$ são escalares complexos. Para provar que c é linearmente independente, recorra-se à definição respectiva: anulemos uma combinação linear (com escalares complexos $z_k = \alpha_k + i\beta_k$) dos $(\vec{e}_k, \vec{o})$ em E', resultando sucessivamente,

$$\begin{aligned}\sum_{k=1}^{n} z_k(\vec{e}_k, \vec{o}) = \sum_{k=1}^{n} (\alpha_k + i\beta_k)(\vec{e}_k, \vec{o}) &= (\vec{o}, \vec{o}) \\ \sum_{k=1}^{n} (\alpha_k \vec{e}_k, \beta_k \vec{e}_k) \qquad &= (\vec{o}, \vec{o}) \\ \left(\sum_{k=1}^{n} \alpha_k \vec{e}_k, \sum_{k=1}^{n} \beta_k \vec{e}_k\right) \qquad &= (\vec{o}, \vec{o})\end{aligned}$$

da última igualdade resulta, finalmente:

$$\begin{cases} \sum_{k=1}^{n} \alpha_k \vec{e}_k = \vec{o} \\ \sum_{k=1}^{n} \beta_k \vec{e}_k = \vec{o} \end{cases}$$

o que, sendo e linearmente independente, implica, para todo o $k \in \{1, 2, \ldots, n\}$,

$$\begin{cases} \alpha_k = 0 \\ \beta_k = 0 \end{cases}$$

donde se segue que, para todo o $k \in \{1, 2, \ldots, n\}$, se tem $\alpha_k = \beta_k = 0$ ou seja,

$$z_k = \alpha_k + i\beta_k = 0$$

Portanto, tem-se $\dim_{\mathbb{C}} E' = \dim_{\mathbb{R}} E$. Podemos ainda observar que as sequências

$$\begin{aligned} u &= ((\vec{o}, \vec{e}_1), (\vec{o}, \vec{e}_2), \ldots, (\vec{o}, \vec{e}_n)) \\ v &= ((\vec{e}_1, \vec{e}_1), (\vec{e}_2, \vec{e}_2), \ldots, (\vec{e}_n, \vec{e}_n)) \end{aligned}$$

serão também bases de E' sobre $\mathbb{C}$ (demonstre!).

Proposição 1.15 *Todo o espaço vectorial E de dimensão finita n sobre um corpo $\mathbb{K}$ é isomorfo*[11] *do espaço cartesiano $\mathbb{K}^n$.*

Demonstração:

Seja $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ uma base de E. Pela proposição 1.8, a aplicação $f_e\colon \mathbb{K}^n \to E$ definida por

$$f_e(x_1, x_2, \ldots, x_n) = \sum_{k=1}^{n} x_k \vec{e}_k$$

é *injectiva* (devido à *independência linear* de e) e *sobrejectiva* (devido a ser e uma lista *geradora* de E); por outro lado, facilmente se verifica que f_e transforma as operações do espaço cartesiano $\mathbb{K}^n$ nas respectivas operações do espaço vectorial E, isto é,

$$\begin{aligned} f_e((x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)) &= f_e(x_1, x_2, \ldots, x_n) + f_e(y_1, y_2, \ldots, y_n) \\ f_e(\alpha(x_1, x_2, \ldots, x_n)) &= \alpha f_e(x_1, x_2, \ldots, x_n) \end{aligned}$$

(no capítulo 3, as funções satisfazendo as duas condições anteriores serão chamadas *lineares* ou *homomorfismos* e, se forem bijectivas, *isomorfismos*). Portanto, f_e é um *isomorfismo* de $\mathbb{K}^n$ sobre E. ☐

Para exprimir o facto de dois espaços E e F sobre o mesmo corpo $\mathbb{K}$ serem isomorfos, escrevemos $E \cong F$; usando esta notação, podemos enunciar a proposição anterior escrevendo:

$$\dim_{\mathbb{K}} E = n \Rightarrow E \cong \mathbb{K}^n$$

A proposição anterior mostra que $\mathbb{K}^n$ constitui, por assim dizer, o *modelo* de espaço vectorial de dimensão n sobre $\mathbb{K}$ e que os restantes espaços vectoriais E sobre $\mathbb{K}$ com igual dimensão são diferentes de $\mathbb{K}^n$ apenas quanto à natureza dos seus vectores (e, claro, quanto à definição das operações). É, portanto, indiferente manipular algebricamente os vectores de E ou as suas *coordenadas* (que são vectores de $\mathbb{K}^n$) em relação a uma qualquer *base* de E.

A relação $\cong$ entre espaços vectoriais sobre o mesmo corpo é uma relação de *equivalência* (ou seja, é *reflexiva*, *simétrica* e *transitiva*), que divide o conjunto dos espaços vectoriais de dimensão finita sobre $\mathbb{K}$ em classes de espaços isomorfos, os quais são algebricamente idênticos. O corolário 1.15.1 mostra que dois espaços pertencem à mesma classe sse têm a mesma dimensão.

---

11 Dois espaços vectoriais E e F sobre o mesmo corpo $\mathbb{K}$ dizem-se *isomorfos* sse existir uma bijecção $f\colon E \to F$ tal que: $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$ e $f(\alpha\vec{x}) = \alpha f(\vec{x})$.

Do exposto, pode escrever-se, por exemplo, para todo o $n \geq 0$,

$$\mathscr{P}_n(\mathbb{K}) \cong \mathbb{K}^{n+1}$$

Do mesmo modo, facilmente se conclui (mostre-o a título de exercício!) que a aplicação

$$f\colon \mathbb{K}^{(\mathbb{N}_0)} \to \mathscr{P}(\mathbb{K});\ (a_0, a_1, \dots, a_r, 0, 0, \dots) \mapsto \sum_{k=0}^{+\infty} a_k x^k$$

é um isomorfismo de $\mathbb{K}^{(\mathbb{N}_0)}$ sobre $\mathscr{P}(\mathbb{K})$, que são, portanto, espaços vectoriais isomorfos:

$$\mathscr{P}(\mathbb{K}) \cong \mathbb{K}^{(\mathbb{N}_0)}$$

A proposição anterior mostra que, por cada base e que se fixe em E, fica definido um isomorfismo f_e de $\mathbb{K}^n$ sobre E, que estabelece uma correspondência biunívoca entre as *coordenadas* (em $\mathbb{K}^n$) de um vector $\vec{x} \in E$ em relação à base e e esse vector $\vec{x}$, de tal modo que é equivalente trabalhar com os vectores de E ou com as suas *coordenadas* em $\mathbb{K}^n$.

Quando dois espaços são isomorfos, isso significa que são algebricamente iguais (diz-se que são iguais, *a menos de um isomorfismo*) e, por vezes, esses espaços são identificados um com o outro: assim, por exemplo, S^3 pode identificar-se com $\mathbb{R}^3$ (e ainda com $\mathscr{P}_2(\mathbb{R})$) e diz-se que S^3 é a *interpretação geométrica* de $\mathbb{R}^3$, confundindo-se cada *terno ordenado* (x, y, z) de números reais com um *segmento orientado* de S^3. Analogamente, S^2 é a *interpretação geométrica* de $\mathbb{R}^2$ e S^1 é a *interpretação geométrica* de $\mathbb{R}^1$: é neste sentido que falamos, às vezes, em *rectas* de $\mathbb{R}^3$, *planos* de $\mathbb{R}^3$, etc (de facto, as rectas, planos, etc existem em S^3!).

Corolário 1.15.1 *Dois espaços vectoriais E e F de dimensão finita sobre o mesmo corpo $\mathbb{K}$ são isomorfos sse têm a mesma dimensão.*

Demonstração:

A condição é necessária: se $\dim E = \dim F = n$, então E e F são ambos isomorfos de $\mathbb{K}^n$ e, portanto, isomorfos um do outro.

Reciprocamente, suponha-se que E e F são isomorfos e que $\dim E = m$ e $\dim F = n$; então E é isomorfo de $\mathbb{K}^m$ e F é isomorfo de $\mathbb{K}^n$, donde, $\mathbb{K}^m \cong \mathbb{K}^n$, o que implica $m = n$. $\square$

Proposição 1.16 *Num espaço vectorial E de dimensão finita n sobre um corpo $\mathbb{K}$:*

i) *Qualquer lista linearmente independente tem, quando muito, n vectores.*

ii) *Qualquer lista geradora de E tem, pelo menos, n vectores.*

iii) *Qualquer lista linearmente independente e com n vectores é uma base de E.*

iv) *Qualquer lista geradora de E e com n vectores é uma base de E.*

v) *Toda a lista linearmente independente de vectores de E é prolongável a uma base.*

vi) *Toda a lista geradora de E contém uma base.*

vii) Se $F \subset E$ é subespaço de E, então $\dim F \leq \dim E = n$.

viii) Se $F \subset E$ é subespaço de E e $\dim F = \dim E$, *então $F = E$.*

Demonstração:

i) Resulta da proposição 1.13, visto que as bases de E têm n vectores e são geradoras de E.

ii) Resulta da proposição 1.13, visto que as bases de E têm n vectores e são independentes.

iii) Uma tal lista é, com certeza, geradora de E: se o não fosse, a proposição 1.12 permitia acrescentar um vector à lista, obtendo-se uma lista de $n+1 > n$ vectores linearmente independentes, o que é contraditório com a alínea *i)*.

iv) Uma tal lista é, com certeza, linearmente independente: se o não fosse, o lema 1.13.2 permitia retirar um vector à lista, mantendo-a geradora de E, mas apenas com $n-1 < n$ vectores, o que contraria a alínea *ii)*.

v) Como se viu, se $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ é linearmente independente será $m \leq n$. A proposição 1.12 permite, enquanto a lista não for geradora de E, ir acrescentando vectores (mantendo a independência linear) até que, quando tivermos acrescentado $n-m$ vectores, teremos uma lista de n vectores linearmente independentes e, portanto, uma base de E, contendo a lista dada.

Isto mostra, em particular, que todo o espaço vectorial de dimensão finita tem, pelo menos, uma base, a qual se pode obter pelo algoritmo que acabámos de utilizar nesta demonstração:

```
x = ∅
Enquanto L(x) ≠ E
  Acrescentar um vector arbitrário de E \ L(x) à lista x
Fim
```

O ciclo anterior termina ao fim de $n = \dim_{\mathbb{K}} E$ iterações e, no final, x será uma base de E. Observe que o algoritmo funciona, mesmo no caso limite $E = \{\vec{o}\}$, obtendo-se, neste caso, a base $x = \emptyset$. Cabe aqui notar que, num espaço de dimensão infinita, a iteração neste algoritmo não termina nunca (*Loop* infinito!).

vi) Como se viu anteriormente, se $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ é geradora de E, será $m \geq n$. O lema 1.13.2 permite retirar vectores à lista x (mantendo-a geradora), enquanto ela for linearmente dependente. Após retirar $m-n$ vectores, obtém-se uma lista geradora com n vectores a qual será, como vimos antes, linearmente independente, logo uma base de E contida em x. Portanto, pode obter-se uma base de E, a partir de uma sequência x geradora, mediante o algoritmo:

```
Enquanto x é linearmente dependente
  Retirar a x um vector que seja combinação linear dos restantes(12)
Fim
```

O ciclo anterior termina ao fim de $m-n$ iterações e, por fim, x será uma base de E. No caso limite $E = \{\vec{o}\}$, acabará por obter-se a sequência vazia $\emptyset$, que é linearmente independente e

---

12 Ou combinação linear dos vectores *anteriores* ou dos *seguintes* (proposições 1.10 e 1.11).

uma base de $E = L(\emptyset) = \{\vec{o}\}$. Neste caso, a lista inicial era, necessariamente:

$$x = (\vec{o}, \vec{o}, \ldots, \vec{o})$$

vii) Como uma base de F é, por definição, linearmente independente em F (e, portanto, também em E, visto que o corpo $\mathbb{K}$ é o mesmo), a alínea *i)* garante imediatamente a desigualdade desejada.

viii) Se $\dim F = \dim E$, isso significa que toda a base b de F é, igualmente, base de E. Portanto $F = L(b)$ e $E = L(b)$, donde $F = E$. ☐

As duas primeiras alíneas da proposição anterior permitem dar a seguinte interpretação da *dimensão* de um espaço vectorial:

- A dimensão de E representa o *número máximo de vectores* que uma lista *linearmente independente* pode ter em E.
- A dimensão de E representa o *número mínimo de vectores* que uma lista *geradora de* E deverá possuir.
- A dimensão de E é igual a n sse existe uma lista linearmente independente com n vectores e toda a lista com $n+1$ vectores é linearmente dependente.

A seguinte proposição dá uma condição necessária e suficiente para que um espaço vectorial tenha dimensão infinita:

Proposição 1.17 – Caracterização dos espaços de dimensão infinita – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$. Então, E tem dimensão* ***infinita*** *sse, para todo o $n \in \mathbb{N}$, existe uma lista* ***linearmente independente*** *com n vectores.*

Demonstração:

O corolário 1.12.1 mostra que a condição é necessária. Mas ela é também suficiente, em virtude da alínea *i)* da proposição anterior. ☐

Exemplo 1.58 O espaço vectorial $\mathbb{K}^{\mathbb{N}}$ das sucessões de escalares, citado no exemplo 1.42, tem dimensão infinita, visto que as listas u_m de sucessões aí apresentadas são linearmente independentes. Também são de dimensão infinita os espaços $\mathbb{K}^{(\mathbb{N})}$ e $\mathbb{K}^{(\mathbb{N}_0)}$.

Exemplo 1.59 O espaço $\mathscr{P}(\mathbb{K})$ tem dimensão infinita, visto que as listas de polinómios

$$c_n = \left(1, x, x^2, \ldots, x^n\right)$$

são linearmente independentes, para todo o $n \in \mathbb{N}$ (método dos coeficientes indeterminados).

Exemplo 1.60 Como as funções $x \mapsto 1$, $x \mapsto x$, $x \mapsto x^2, \ldots, x \mapsto x^n$ são linearmente independentes para todo o $n \in \mathbb{N}$, terão também dimensão infinita os espaços $\mathbb{R}^I$, $\mathscr{C}(I, \mathbb{R})$, $\mathscr{D}(I, \mathbb{R})$, $\mathscr{C}^k(I, \mathbb{R})$, etc, onde $I = (a, b) \subset \mathbb{R}$, com $-\infty \leq a < b \leq +\infty$, é um intervalo aberto.

Exemplo 1.61 O espaço $\mathbb{R}$ tem dimensão infinita sobre $\mathbb{Q}$: de facto, o número real π é *transcendente* (isto é, pode demonstrar-se que π não é solução de equação algébrica alguma com coeficientes racionais não todos nulos). Então, as listas

$$x_n = \left(1, \pi, \pi^2, \ldots, \pi^n\right)$$

são linearmente independentes, para todo o $n \in \mathbb{N}$, visto que $\sum_{k=0}^{n} a_k \pi^{\,k} = 0$ (onde os $a_k \in \mathbb{Q}$) só é possível se forem nulos todos os a_k.

Proposição 1.18 – Propriedades das operações elementares – *Cada uma das seguintes operações (ditas elementares), quando realizadas sobre uma lista $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ de $m > 0$ vectores de um espaço vectorial E sobre um corpo $\mathbb{K}$, não alteram a dependência ou independência linear nem o subespaço gerado por essa lista (nem, obviamente, a dimensão deste subespaço e o comprimento m da lista):*

i) *Troca entre si de dois elementos $\vec{x}_i$ e $\vec{x}_j$ da lista, com $i \neq j$ e, claro, $m \geq 2$:*

$$(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots, \vec{x}_m) \to (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_j, \ldots, \vec{x}_i, \ldots, \vec{x}_m) \tag{1.54}$$

ii) *Substituição de um vector $\vec{x}_i$ da lista pelo seu produto por um escalar $\alpha \neq 0$, com $m \geq 1$:*

$$(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_m) \to (\vec{x}_1, \vec{x}_2, \ldots, \alpha\vec{x}_i, \ldots, \vec{x}_m), \textit{ onde } \alpha \neq 0 \tag{1.55}$$

iii) *Sendo $m \geq 2$, substituição de um vector $\vec{x}_i$ da lista pela sua soma com o produto de outro vector $\vec{x}_j$, com $j \neq i$, multiplicado por um escalar β arbitrário:*

$$(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_m) \to (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i + \beta\vec{x}_j, \ldots, \vec{x}_m) \tag{1.56}$$

Demonstração:

i) Esta propriedade é consequência da comutatividade e associatividade da adição vectorial.

ii) Seja $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_m)$ e $x' = (\vec{x}_1, \vec{x}_2, \ldots, \alpha\vec{x}_i, \ldots, \vec{x}_m)$. A igualdade

$$\sum_{k=1}^{m} \beta_k \vec{x}_k = \sum_{k \neq i} \beta_k \vec{x}_k + \tfrac{\beta_i}{\alpha}(\alpha\vec{x}_i)$$

mostra que toda a combinação linear de vectores de x será combinação linear dos vectores de x', portanto $L(x) \subset L(x')$. A igualdade anterior mostra ainda que se x' é linearmente independente, também x o será, visto $\beta_i/\alpha = 0$ implica $\beta_i = 0$. Por outro lado, a igualdade

$$\sum_{k \neq i} \beta_k \vec{x}_k + \beta_i(\alpha\vec{x}_i) = \sum_{k \neq i} \beta_k \vec{x}_k + (\beta_i\alpha)\vec{x}_i$$

mostra que as combinações lineares dos vectores de x' são também combinação linear dos vectores de x, logo $L(x') \subset L(x)$. Dela se segue também que, se x for linearmente independente, x' também será, já que por ser $\alpha \neq 0$, $\beta_i\alpha = 0$ implicará $\beta_i = 0$.

iii) Como anteriormente, façamos

$$x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_m) \text{ e } x' = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i + \beta\vec{x}_j, \ldots, \vec{x}_m)$$

A igualdade

$$\sum_{k=1}^{m} \alpha_k \vec{x}_k = \sum_{k\neq i,j} \alpha_k \vec{x}_k + \alpha_i(\vec{x}_i + \beta\vec{x}_j) + (\alpha_j - \alpha_i\beta)\vec{x}_j$$

mostra que toda a combinação linear dos vectores de x é combinação linear dos vectores de x', ou seja, $L(x) \subset L(x')$. Ela mostra também que, se x' for linearmente independente, também o será x, pois que

$$\underset{k\neq i,j}{\forall} (\alpha_k = 0) \wedge \alpha_i = 0 \wedge \alpha_j - \alpha_i\beta = 0 \Rightarrow \underset{k\neq i,j}{\forall} (\alpha_k = 0) \wedge \alpha_i = 0 \wedge \alpha_j = 0$$

Por outro lado, a igualdade

$$\sum_{k\neq i} \alpha_k \vec{x}_k + \alpha_i(\vec{x}_i + \beta\vec{x}_j) = \sum_{k\neq j} \alpha_k \vec{x}_k + (\alpha_i\beta + \alpha_j)\vec{x}_j$$

mostra que uma combinação linear dos vectores de x' é forçosamente combinação linear dos vectores de x, logo $L(x') \subset L(x)$. Da igualdade anterior se segue também que, se x for linearmente independente, x' também será, já que

$$\underset{k\neq j}{\forall} (\alpha_k = 0) \wedge \alpha_i\beta + \alpha_j = 0 \Rightarrow \underset{k\neq j}{\forall} (\alpha_k = 0) \wedge \alpha_j = 0 \qquad \square$$

As operações das três alíneas da proposição anterior são chamadas *operações elementares de tipos 1, 2 e 3*, respectivamente, sobre a lista x de vectores. Por vezes, usam-se as seguintes notações para indicar aquelas operações:

$$\begin{aligned} &\vec{x}_i \overset{E^1}{\longleftrightarrow} \vec{x}_j \\ &\vec{x}_i \xrightarrow{E^2} \alpha\vec{x}_i, \qquad \text{onde } \alpha \neq 0 \\ &\vec{x}_i \xrightarrow{E^3} \vec{x}_i + \beta\vec{x}_j \end{aligned}$$

É óbvio que qualquer composição de operações daqueles três tipos não alterará também o subespaço gerado pela lista nem a sua dependência ou independência lineares. Em particular, tem-se o seguinte:

Corolário 1.18.1 *Cada uma das seguintes operações, quando realizadas sobre uma lista $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ de $m > 0$ vectores de um espaço vectorial E sobre um corpo $\mathbb{K}$, não alteram a dependência ou independência linear nem o subespaço gerado por essa lista (nem, obviamente, o seu comprimento m):*

i) Sendo $m \geq 2$, substituição de um vector $\vec{x}_i$ da lista pela soma do produto de um escalar $\alpha \neq 0$ por esse vector com o produto de um escalar β por outro vector $\vec{x}_j$, com $j \neq i$:

$$(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_m) \to (\vec{x}_1, \vec{x}_2, \ldots, \alpha\vec{x}_i + \beta\vec{x}_j, \ldots, \vec{x}_m), \textit{ onde } \alpha \neq 0 \qquad (1.57)$$

ii) *Sendo $m \geq 1$, substituição de um vector $\vec{x}_i$ por uma combinação linear $\sum_{k=1}^{m} \alpha_k \vec{x}_k$ dos vectores da lista x, desde que $\alpha_i \neq 0$:*

$$(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_m) \to \left(\vec{x}_1, \vec{x}_2, \ldots, \sum_{k=1}^{m} \alpha_k \vec{x}_k, \ldots, \vec{x}_m\right), \textit{ onde } \alpha_i \neq 0 \tag{1.58}$$

Demonstração:

i) Esta operação é a composição $E^3 \circ E^2$ de uma operação do tipo 3 com outra do tipo 2:

$$\begin{aligned} &\vec{x}_i \xrightarrow{E^2} \alpha \vec{x}_i \\ &\alpha \vec{x}_i \xrightarrow{E^3} \alpha \vec{x}_i + \beta \vec{x}_j \end{aligned}$$

ii) Esta operação é a composição $E^3_{m-1} \circ \cdots \circ E^3_2 \circ E^3_1 \circ E^2$ de uma operação do tipo 2 seguida por $m-1$ operações de tipo 3:

$$\begin{aligned} &\vec{x}_i \xrightarrow{E^2} \alpha_i \vec{x}_i, \text{ onde } \alpha_i \neq 0 \\ &\alpha_i \vec{x}_i \xrightarrow{E^3_1} \alpha_i \vec{x}_i + \alpha_{k_1} \vec{x}_{k_1}, \text{ onde } k_1 \neq i \\ &\alpha_i \vec{x}_i + \alpha_1 \vec{x}_1 \xrightarrow{E^3_2} \alpha_i \vec{x}_i + \alpha_{k_1} \vec{x}_{k_1} + \alpha_{k_2} \vec{x}_{k_2}, \text{ onde } k_1 \neq k_2 \neq i \\ &\ldots \\ &\alpha_i \vec{x}_i + \sum_{r=1}^{m-2} \alpha_{k_r} \vec{x}_{k_r} \xrightarrow{E^3_{m-1}} \sum_{k=1}^{m} \alpha_k \vec{x}_k \end{aligned}$$

□

A primeira operação deste corolário é, por vezes, chamada *operação de tipo 4*. Analogamente às operações dos tipos 1, 2 e 3, pode usar-se a notação:

$$\vec{x}_i \xrightarrow{E^4} \alpha \vec{x}_i + \beta \vec{x}_j, \text{ onde } \alpha \neq 0$$

1.7 Soma de subespaços. Soma directa

Nesta secção, vamos analisar de mais perto um processo de gerar novos subespaços de um espaço vectorial E dado, a partir de outros subespaços de E. Comecemos por provar a seguinte proposição:

Proposição 1.19 *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e $(E_k)_{1 \leq k \leq m}$ uma lista de m subespaços de E. O conjunto $A \subset E$ formado por todas as somas $\sum_{k=1}^{m} \vec{x}_k$ de vectores $\vec{x}_k \in E_k$ é ainda um subespaço de E.*

Demonstração:

$\vec{o} \in A$, pelo que S1 é obviamente válida, visto que $\sum_{k=1}^{m} \vec{o} = \vec{o}$. Quanto a S2 tem-se, se $\vec{x}$ e $\vec{y} \in A$,

$$\vec{x} + \vec{y} = \sum_{k=1}^{m} \vec{x}_k + \sum_{k=1}^{m} \vec{y}_k = \sum_{k=1}^{m} (\vec{x}_k + \vec{y}_k)$$

e, como $\vec{x}_k + \vec{y}_k \in E_k$, será $\vec{x} + \vec{y} \in A$. Por fim, S3 resulta de $\alpha \sum_{k=1}^{m} \vec{x}_k = \sum_{k=1}^{m} (\alpha \vec{x}_k)$ e de ser $\alpha \vec{x}_k \in E_k$. □

O subespaço A mencionado na proposição anterior é chamado *soma* dos subespaços E_k e designado pela notação $\sum_{k=1}^{m} E_k$, como se formaliza na seguinte:

Definição 1.12 – Soma de subespaços – *Seja E um espaço vectorial e $(E_k)_{1 \leq k \leq m}$ uma sequência de m subespaços de E. O subespaço formado por todas as somas de vectores $\vec{x}_k \in E_k$ é chamado* ***soma*** *dos subespaços E_k e designado por $\sum_{k=1}^{m} E_k$:*

$$\sum_{k=1}^{m} E_k = \left\{ \vec{y} \in E \colon \underset{\vec{x}_k \in E_k}{\exists} \vec{y} = \sum_{k=1}^{m} \vec{x}_k \right\} \tag{1.59.1}$$

em particular, se $m = 2$:

$$E_1 + E_2 = \left\{ \vec{y} \in E \colon \underset{\substack{\vec{x}_1 \in E_1 \\ \vec{x}_2 \in E_2}}{\exists} \vec{y} = \vec{x}_1 + \vec{x}_2 \right\} \tag{1.59.2}$$

A proposição que se segue dá conta das principais propriedades da soma de subespaços.

Proposição 1.20 – Propriedades da soma de subespaços – *Seja $(E_k)_{1 \leq k \leq m}$ uma lista de subespaços de um espaço vectorial E. Então:*

i) *Para todo o $i \in \{1, 2, \ldots, m\}$,*

$$E_i \subset \sum_{k=1}^{m} E_k$$

ii)

$$\bigcup_{k=1}^{m} E_k \subset \sum_{k=1}^{m} E_k$$

iii) *Se F é um subespaço de E contendo todos os E_k (logo, F também conterá a reunião dos E_k), então:*

$$\sum_{k=1}^{m} E_k \subset F$$

iv) *Se σ é uma permutação de $\{1, 2, \ldots, m\}$, então:*

$$\sum_{k=1}^{m} E_{\sigma_k} = \sum_{k=1}^{m} E_k$$

v) *Para todo o $i \in \{1, 2, \ldots, m\}$,*

$$\sum_{k=1}^{m} E_k = E_i \Leftrightarrow \sum_{k \neq i} E_k \subset E_i$$

vi) *Tem-se:*

$$\dim\left(\sum_{k=1}^{m} E_k\right) \leq \sum_{k=1}^{m} \dim E_k$$

Demonstração:

i) A primeira inclusão resulta de que, para qualquer $i \in \{1, 2, \ldots, m\}$, se tem:

$$\vec{x} \in E_i \Rightarrow \vec{x} = \vec{x} + \sum_{k \neq i} \vec{o} \Rightarrow \vec{x} \in \sum_{k=1}^{m} E_k$$

ii) A segunda inclusão obtém-se reunindo as m inclusões em *i*)

$$\underset{1 \leq i \leq m}{\forall}\left(E_i \subset \sum_{k=1}^{m} E_k\right) \Rightarrow \bigcup_{i=1}^{m} E_i \subset \bigcup_{i=1}^{m}\left(\sum_{k=1}^{m} E_k\right) = \sum_{k=1}^{m} E_k$$

iii) Quanto à terceira proposição, se $\vec{x} \in \sum_{k=1}^{m} E_k$, $\vec{x}$ será soma de vectores $\vec{x}_k \in E_k$ os quais pertencerão também a F (em virtude de ser $E_k \subset F$). Como F é subespaço, $\sum_{k=1}^{m} \vec{x}_k$ pertencerá a F, o que significa que $\sum_{k=1}^{m} E_k \subset F$.

iv) Resulta da comutatividade e associatividade da adição vectorial.

v) Como $\sum_{k \neq i} \vec{x}_k = \sum_{k \neq i} \vec{x}_k + \vec{o}$ e o vector nulo pertence a E_i, segue-se que

$$\sum_{k \neq i} E_k \subset \sum_{k=1}^{m} E_k$$

Então, se $\sum_{k=1}^{m} E_k = E_i$, vem imediatamente $\sum_{k\neq i} E_k \subset E_i$.

Reciprocamente, suponha-se que $\sum_{k\neq i} E_k \subset E_i$ e seja $\vec{x} \in \sum_{k=1}^{m} E_k$: então será

$$\vec{x} = \sum_{k=1}^{m} \vec{x}_k = \vec{x}_i + \underbrace{\sum_{k\neq i} \vec{x}_k}_{\in \sum_{k\neq i} E_k \subset E_i} \in E_i$$

o que mostra que $\sum_{k=1}^{m} E_k \subset E_i$ e isto juntamente com *i)* termina a demonstração.

vi) Supondo dim$E_k = n_k$ e que $e_k = (\vec{e}_{k1}, \vec{e}_{k2}, \ldots, \vec{e}_{kn_k})$ é base de E_k, mostra-se facilmente que a junção $e = e_1 \vee e_2 \vee \cdots \vee e_m = (\vec{e}_{kj_k})_{\substack{1\leq k\leq m \\ 1\leq j_k\leq n_k}}$ das bases e_k constitui um sistema de geradores de $\sum_{k=1}^{m} E_k$: de facto, tem-se, sucessivamente:

$$\vec{x} \in \sum_{k=1}^{m} E_k \Rightarrow \vec{x} = \sum_{k=1}^{m} \vec{x}_k = \sum_{k=1}^{m} \left(\sum_{j_k=1}^{n_k} \alpha_{kj_k} \vec{e}_{kj_k} \right) = \sum_{k=1}^{m} \sum_{j_k=1}^{n_k} (\alpha_{kj_k} \vec{e}_{kj_k})$$

Esta última expressão mostra que $\vec{x}$ é combinação linear dos $(\vec{e}_{kj_k})_{\substack{1\leq k\leq m \\ 1\leq j_k\leq n_k}}$, o que termina a demonstração. $\square$

A proposição anterior significa que a soma dos subespaços E_k é o *menor* subespaço de E contendo cada um dos E_k e, portanto, contendo $\bigcup_{k=1}^{m} E_k$ (note que esta reunião não é, em geral, um subespaço de E).

Corolário 1.20.1 *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e A, B, C subespaços de E. Tem-se:*

i) $A \subset A + B \wedge B \subset A + B$

ii) $A \cup B \subset A + B$

iii) Se F é um subespaço de E contendo A e B, tem-se: $A + B \subset F$

iv) $A + B = B + A$

v) $A + B = A \Leftrightarrow B \subset A$

vi) dim$(A + B) \leq$ dim$A +$ dim$B$

vii) dim$(A + B) =$ dim$A +$ dim$B -$ dim$(A \cap B)$

viii) $A + \{\vec{o}\} = \{\vec{o}\} + A = A$

ix) $A + A = A$

Demonstração:

As primeiras seis alíneas resultam da proposição anterior, fazendo $m = 2$.

As alíneas *viii)* e *ix)* resultam de *v)*, fazendo $B = \{\vec{o}\}$ e $B = A$, respectivamente.

Quanto a *vii)*, seja $s = \dim A \cap B$, $m = \dim A$ e $n = \dim B$. Como $A \cap B$ é um subespaço de A e de B, será $s \leq m$ e $s \leq n$. Tome-se, então, uma base $c = (\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_s)$ de $A \cap B$ e prologuemo-la (proposição 1.16.*v*) a uma base $a = (\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_s, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{m-s})$ de A e também a uma base $b = \left(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_s, \vec{b}_1, \vec{b}_2, \ldots, \vec{b}_{n-s}\right)$ de B.

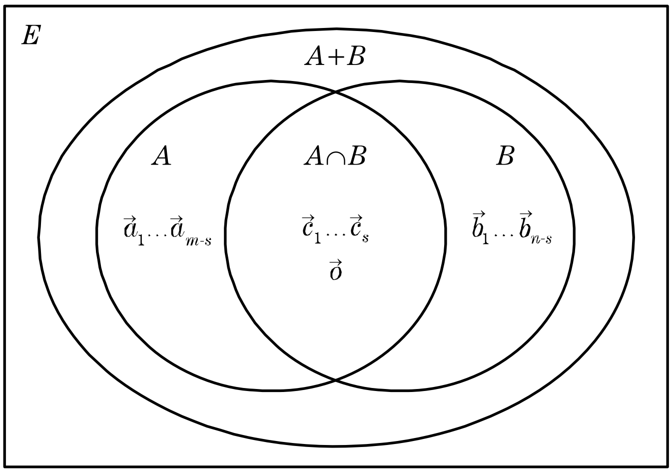


Fig. 1.7 – Os espaços A, B, $A \cap B$ e as suas bases.

Para provar a igualdade desejada, basta provar que a lista

$$r = \left(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_s, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{m-s}, \vec{b}_1, \vec{b}_2, \ldots, \vec{b}_{n-s}\right)$$

de $m + n - s$ vectores de E é uma base de $A + B$: comecemos por ver que:

$$L(r) = A + B$$

De facto,

$$\begin{aligned} \vec{z} \in A + B \Rightarrow \vec{z} &= \vec{a} + \vec{b}, \text{ onde } \vec{a} \in A \text{ e } \vec{b} \in B \\ &= \left(\sum_{k=1}^{s} \alpha_k \vec{c}_k + \sum_{k=1}^{m-s} \beta_k \vec{a}_k \right) + \left(\sum_{k=1}^{s} \gamma_k \vec{c}_k + \sum_{k=1}^{n-s} \eta_k \vec{b}_k \right) \\ &= \sum_{k=1}^{s} (\alpha_k + \gamma_k) \vec{c}_k + \sum_{k=1}^{m-s} \beta_k \vec{a}_k + \sum_{k=1}^{n-s} \eta_k \vec{b}_k \end{aligned}$$

a última expressão mostra que $\vec{z}$ é combinação linear dos vectores de r e, portanto, $A + B \subset L(r)$. A inclusão inversa é óbvia, já que

$$\begin{aligned} \vec{z} \in L(r) \Rightarrow \vec{z} = \underbrace{\sum_{k=1}^{s} \alpha_k \vec{c}_k + \sum_{k=1}^{m-s} \beta_k \vec{a}_k}_{\in A} + \underbrace{\sum_{k=1}^{n-s} \gamma_k \vec{b}_k}_{\in B} \\ = \vec{a} + \vec{b} \in A + B \end{aligned}$$

Vejamos, agora, que r é linearmente independente, recorrendo à definição:

$$\sum_{k=1}^{s} \alpha_k \vec{c}_k + \sum_{k=1}^{m-s} \beta_k \vec{a}_k + \sum_{k=1}^{n-s} \gamma_k \vec{b}_k = \vec{o} \tag{1.60}$$

o que equivale a

$$\underbrace{\sum_{k=1}^{s} \alpha_k \vec{c}_k + \sum_{k=1}^{m-s} \beta_k \vec{a}_k}_{\in A} = \underbrace{\sum_{k=1}^{n-s} (-\gamma_k) \vec{b}_k}_{\in B}$$

A igualdade anterior mostra que $\sum_{k=1}^{n-s} (-\gamma_k) \vec{b}_k \in A \cap B$, donde, existem $\alpha_k \in \mathbb{K}$ tais que:

$$\sum_{k=1}^{n-s} (-\gamma_k) \vec{b}_k = \sum_{k=1}^{s} \alpha_k \vec{c}_k$$

o que equivale a

$$\sum_{k=1}^{s} \alpha_k \vec{c}_k + \sum_{k=1}^{n-s} \gamma_k \vec{b}_k = \vec{o}$$

atendendo a que b é uma base de B, tem-se, forçosamente, $\gamma_k = 0$, para $k = 1, \ldots, n - s$. Substituindo os valores de γ_k em (1.60) e atendendo a que a é uma base, resulta o anulamento dos α_k ($k = 1, \ldots, s$) e dos β_k ($k = 1, \ldots, m - s$), terminando a demonstração. $\square$

Dependendo dos subespaços $(E_i)_{1\leq i\leq m}$, a soma em que se decompõem os vectores de $\sum_{k=1}^{m} E_k$ pode ou não ser *única*; quando for única, diz-se que a soma é *directa*:

Definição 1.13 – Soma directa de subespaços – *Uma soma de subespaços $(E_k)_{1\leq k\leq m}$ de um espaço vectorial E diz-se* ***directa*** *sse, para cada vector $\vec{u} \in F = \sum_{k=1}^{m} E_k$, existe uma e uma só sequência $x = (\vec{x}_i)_{1\leq i\leq m}$ de vectores, em que $\vec{x}_i \in E_i$, tal que*

$$\vec{u} = \sum_{k=1}^{m} \vec{x}_k \tag{1.61}$$

Neste caso, a soma (directa) dos E_k designa-se por $\bigoplus_{k=1}^{m} E_k$, em vez de $\sum_{k=1}^{m} E_k$. (em particular, note-se que, para o vector nulo, a soma referida será : $\vec{o} = \sum_{k=1}^{m} \vec{o}$).

Na proposição seguinte, vamos ver que uma soma de subespaços é directa sse a intersecção de cada subespaço com a soma dos restantes se reduzir ao vector nulo.

Proposição 1.21 – Caracterização da soma directa – *A soma $F = \sum_{k=1}^{m} E_k$ de uma sequência $(E_k)_{1\leq k\leq m}$ de subespaços de um espaço vectorial E é* ***directa*** *sse, para cada*

$i \in \{1, 2, \ldots, m\}$, *a intersecção* $E_i \cap \sum\limits_{k \neq i} E_k$ *se reduz a* $\{\vec{o}\}$.

$$F = \bigoplus_{k=1}^{m} E_k \Leftrightarrow F = \sum_{k=1}^{m} E_k \wedge \underset{1 \leq i \leq m}{\forall} E_i \cap \sum_{k \neq i} E_k = \{\vec{o}\} \tag{1.62}$$

Demonstração:

1. A condição é suficiente:

Suponhamos que $\vec{u} \in \sum\limits_{k=1}^{m} E_k$ e que existem duas decomposições de $\vec{u}$ em vectores dos E_k:

$$\vec{u} = \sum_{k=1}^{m} \vec{x}_k = \sum_{k=1}^{m} \vec{y}_k \Rightarrow \sum_{k=1}^{m} (\vec{x}_k - \vec{y}_k) = \vec{o}$$

Para cada $i \in \{1, 2, \ldots, m\}$, pode pois escrever-se

$$(\vec{x}_i - \vec{y}_i) + \sum_{k \neq i} (\vec{x}_k - \vec{y}_k) = \vec{o}$$

Portanto, para qualquer $i \in \{1, 2, \ldots, m\}$,

$$\underbrace{(\vec{x}_i - \vec{y}_i)}_{\in E_i} = \underbrace{\sum_{k \neq i} (\vec{y}_k - \vec{x}_k)}_{\in \sum\limits_{k \neq i} E_k}$$

Mas, como os $(E_k)_{1 \leq k \leq m}$ são subespaços, $\vec{x}_i - \vec{y}_i \in E_i$ e $\sum\limits_{k \neq i} (\vec{y}_k - \vec{x}_k) \in \sum\limits_{k \neq i} E_k$ e, portanto, para qualquer $k \in \{1, 2, \ldots, m\}$:

$$\vec{x}_i - \vec{y}_i \in E_i \cap \sum_{k \neq i} E_k = \{\vec{o}\}$$

daqui se conclui que

$$\underset{1 \leq i \leq m}{\forall} \vec{x}_i = \vec{y}_i$$

o que significa que a decomposição de $\vec{u}$ é única e a soma é directa.

2. A condição é necessária:

Suponha-se que existia um $i \in \{1, 2, \ldots, m\}$ tal que $G_i = E_i \cap \sum\limits_{k \neq i} E_k \neq \{\vec{o}\}$ e seja $\vec{a}_i \neq \vec{o}$ um vector de G_i (observe que $\vec{a}_i \in E_i$ e $\vec{a}_i \in \sum\limits_{k \neq i} E_k$). Como $\vec{a}_i \in \sum\limits_{k \neq i} E_k$, $\vec{a}_i$ será decomponível numa soma de vectores $(\vec{a}_k)_{1 \leq k \leq m}$, tais que $\vec{a}_k \in E_k$ para todo o $k \neq i$, satisfazendo

$$\vec{a}_i = \sum_{k \neq i} \vec{a}_k$$

Se for $\vec{u}$ um vector de $\sum\limits_{k=1}^{m} E_k$ será, com $\vec{x}_k \in E_k$:

$$\begin{aligned}\vec{u} &= \sum_{k=1}^{m} \vec{x}_k = \vec{x}_i + \sum_{k\neq i} \vec{x}_k \\ &= (\vec{x}_i + \vec{a}_i) + \sum_{k\neq i} \vec{x}_k - \vec{a}_i \\ &= (\vec{x}_i + \vec{a}_i) + \sum_{k\neq i} \vec{x}_k - \sum_{k\neq i} \vec{a}_k \\ &= (\vec{x}_i + \vec{a}_i) + \sum_{k\neq i} (\vec{x}_k - \vec{a}_k)\end{aligned}$$

como, para qualquer $k \in \{1, 2, \ldots, m\}$, $\vec{x}_k - \vec{a}_k \in E_k$ e $\vec{x}_i + \vec{a}_i \neq \vec{x}_i$, as expressões

$$\vec{u} = \sum_{k=1}^{m} \vec{x}_k \quad \text{e} \quad \vec{u} = (\vec{x}_i + \vec{a}_i) + \sum_{k\neq i} (\vec{x}_k - \vec{a}_k)$$

são duas decomposições de $\vec{u}$ diferentes numa soma de vectores dos E_k. Daqui se segue que a soma $\sum\limits_{k=1}^{m} E_k$ não é directa. □

A proposição anterior tem o seguinte corolário imediato, fazendo $m = 2$:

Corolário 1.21.1 *Se A e B são subespaços de um espaço vectorial E, então:*

$$F = A \oplus B \Leftrightarrow F = A + B \wedge A \cap B = \{\vec{o}\}$$

Exemplo 1.62 Os conjuntos

$$\begin{aligned}P &= \Big\{g\colon \underset{z\in\mathbb{C}}{\forall}\, g(-z) = g(z)\Big\} \subset \mathbb{C}^{\mathbb{C}} \\ I &= \Big\{h\colon \underset{z\in\mathbb{C}}{\forall}\, p(-z) = -h(z)\Big\} \subset \mathbb{C}^{\mathbb{C}}\end{aligned}$$

formados pelas funções *pares* e pelas funções *ímpares*, respectivamente, são subespaços de $\mathbb{C}^{\mathbb{C}}$ (demonstre!). Como a única função simultaneamente par e ímpar é a função nula ($P \cap I = \{o\}$) e, para toda a função $f \in \mathbb{C}^{\mathbb{C}}$, se tem

$$f(z) = \frac{1}{2}(f(z) + f(-z)) + \frac{1}{2}(f(z) - f(-z))$$

onde a primeira parcela é função par e a segunda é função ímpar (verifique!), será $\mathbb{C}^{\mathbb{C}} = P \oplus I$.

Vamos, agora ver que a dimensão de uma soma directa de subespaços é igual à soma das dimensões destes:

Proposição 1.22 – Dimensão da soma directa – *Sejam $E_1, E_2, \ldots, E_p$ subespaços de dimensões finitas $n_1, n_2, \ldots, n_p$ de um espaço vectorial E. Então:*

$$\dim\Big(\bigoplus_{k=1}^{p} E_k\Big) = \sum_{k=1}^{p} \dim(E_k) \tag{1.63}$$

Demonstração:

Tome-se, para cada subespaço E_k, uma base

$$e_k = (\vec{e}_{k1}, \vec{e}_{k2}, \ldots, \vec{e}_{kn_k});\ k = 1, 2, \ldots, p$$

e mostremos que a lista de $n = \sum_{k=1}^{p} n_k$ vectores de E

$$\begin{aligned} e &= e_1 \vee e_2 \vee \ldots \vee e_p \\ &= \left(\vec{e}_{11}, \vec{e}_{12}, \ldots, \vec{e}_{1n_1}, \vec{e}_{21}, \vec{e}_{22}, \ldots, \vec{e}_{2n_2}, \ldots, \vec{e}_{p1}, \vec{e}_{p2}, \ldots, \vec{e}_{pn_p}\right) \end{aligned}$$

é uma *base* de $F = \bigoplus_{k=1}^{p} E_k$:

- A lista e *gera* F. De facto, se $\vec{x} \in F$, então $\vec{x}$ é soma única de vectores dos E_k (por definição de soma directa)

$$\vec{x} = \sum_{k=1}^{p} \vec{x}_k,\ \text{onde } \vec{x}_k \in E_k \tag{1.64}$$

Mas, sendo $\vec{x}_k \in E_k$, o vector $\vec{x}_k$ será combinação linear única dos vectores da base e_k:

$$\vec{x}_k = \sum_{i=1}^{n_k} \alpha_{ki}\vec{e}_{ki},\ k = 1, 2, \ldots, p$$

Substituindo esta última expressão em (1.64), obtém-se

$$\vec{x} = \sum_{k=1}^{p} \sum_{i=1}^{n_k} \alpha_{ki}\vec{e}_{ki}$$

o que mostra que $\vec{x} \in L(e)$.

- A lista e é *linearmente independente*. De

$$\sum_{k=1}^{p} \sum_{i=1}^{n_k} \alpha_{ki}\vec{e}_{ki} = \vec{o} = \vec{o} + \vec{o} + \cdots + \vec{o}$$

vem, por ser $\sum_{i=1}^{n_k} \alpha_{ki}\vec{e}_{ki} \in E_k$ e F ser uma soma directa dos E_k,

$$\sum_{i=1}^{n_k} \alpha_{ki}\vec{e}_{ki} = \vec{o},\ k = 1, 2, \ldots, p$$

Como os vectores $\vec{e}_{ki}$ são linearmente independentes (e_k é uma base de E_k) segue-se que

$$\alpha_{ki} = 0,\ \begin{cases} k = 1, 2, \ldots, p \\ i\ = 1, 2, \ldots, n_k \end{cases}$$

□

1.8 Anexos: vectores e o MATHEMATICA©

MATHEMATICA©

O software MATHEMATICA© permite a manipulação de vectores através da noção de lista, que é mais geral que a de vector; um vector não contém listas como elementos seus. Existem várias funções para manipulação de listas e, de entre estas, salientamos:

- `List[]` ou `{}`

Devolve a lista formada pelos argumentos.

- `Plus[]` ou `+`

Adiciona as listas elemento a elemento.

- `Times[]` ou `*` ou espaço

Multiplica escalares entre si e também escalares por listas.

- `Part[]` ou `[[]]`

Permite extrair elementos e sublistas de uma lista.

- `Take[]` e `Select[]`

Permitem extrair sublistas, segundo vários critérios.

- `Sort[]`, `Join[]` e `Union[]`

Permitem ordenar, concatenar e reunir listas.

- `Length[]` e `Dimensions[][[1]]`

Devolvem o comprimento de uma lista.

- `VectorQ[]`

Função booleana que devolve `True` sse o argumento é um vector.

- `Table[]`, `Range[]` e `Array[]`

Permitem construir listas.

- `Sum[]` e `Product[]`

Calculam somas e produtos dos elementos de uma lista.

Nas páginas seguintes, apresenta-se um notebook ilustrando o uso de algumas das funções aqui mencionadas:

- **Uma função** F **pode utilizar-se com 3 sintaxes possíveis:** F[x]**, ou** F@x **ou ainda** x//F

```
Sin[Pi / 4]
```

$$\frac{1}{\sqrt{2}}$$

```
Sin@(Pi / 4)
```

$$\frac{1}{\sqrt{2}}$$

```
(Pi / 4) // Sin
```

$$\frac{1}{\sqrt{2}}$$

Vectores

1.8.1 O que são?

No MATHEMATICA, um vector é uma lista de escalares racionais, reais ou complexos

```
v = {-2, 3, 1, 2, 4}
```

```
{-2, 3, 1, 2, 4}
```

```
z = {3 / 4, -5 / 8, 2 / 7, -2 / 9}
```

$$\left\{\frac{3}{4}, -\frac{5}{8}, \frac{2}{7}, -\frac{2}{9}\right\}$$

As componentes podem ser números irracionais ou transcendentes...

```
t = {Sin[Pi / 4], Cos[3 * Pi / 5], -Sqrt[5], ArcSin[Sqrt[3] / 2]}
```

$$\left\{\frac{1}{\sqrt{2}}, \frac{1}{4}\left(1-\sqrt{5}\right), -\sqrt{5}, \frac{\pi}{3}\right\}$$

E podem ser complexos...

```
w = {1 + I, -3 + 2 * I, -1 + 2 * I, I, (4 + I) ^2}
```

$$\{1+i, -3+2i, -1+2i, i, 15+8i\}$$

Somando com + ou `Plus[]`...

```
v + w
```

$$\{-1+i, 2i, 2i, 2+i, 19+8i\}$$

```
Plus[v, w]
```

$$\{-1+i, 2i, 2i, 2+i, 19+8i\}$$

```
z - t
```

$$\left\{\frac{3}{4}-\frac{1}{\sqrt{2}}, -\frac{5}{8}+\frac{1}{4}\left(-1+\sqrt{5}\right), \frac{2}{7}+\sqrt{5}, -\frac{2}{9}-\frac{\pi}{3}\right\}$$

1.8.2. Acesso às componentes de um vector

- **Podemos extrair componentes de um vector com** `w[[]]` **ou** `Part[]` **e** `Take[]`

```
w[[3]]
```

$$-1+2i$$

```
w[[{3}]]
```

```
{-1 + 2 i}
```

```
w[[{1, 3, 3}]]
```

```
{1 + i, -1 + 2 i, -1 + 2 i}
```

```
Part[w, {1, 3, 3}]
```

```
{1 + i, -1 + 2 i, -1 + 2 i}
```

Extrair as 2 primeiras componentes

```
Take[w, 2]
```

```
{1 + i, -3 + 2 i}
```

Ou as 2 últimas

```
Take[w, -2]
```

```
{i, 15 + 8 i}
```

Extrair da 2ª até à 5ª componentes

```
Take[w, {2, 5}]
```

```
{-3 + 2 i, -1 + 2 i, i, 15 + 8 i}
```

Ou da 1ª à 5ª de 2 em 2

```
Take[w, {1, 5, 2}]
```

```
{1 + i, -1 + 2 i, 15 + 8 i}
```

- Extrair elementos de acordo com condições: `Select[]`

```
w = {10, 7, 13, 4, 11, 16, 9, 20, 7, 13};
```

```
Select[w, OddQ]
```

```
{7, 13, 11, 9, 7, 13}
```

```
Select[w, EvenQ]
```

```
{10, 4, 16, 20}
```

```
Select[w, # < 9 || # > 15 &]
```

```
{7, 4, 16, 20, 7}
```

- Ordenar

```
Sort[w]
```

```
{4, 7, 7, 9, 10, 11, 13, 13, 16, 20}
```

- Juntar listas (concatenação e reunião)

```
u = {16, 3, 12, 8, 3, 21, 15, 20, 13};
```

```
Join[u, w]
```

```
{16, 3, 12, 8, 3, 21, 15, 20, 13, 10, 7, 13, 4, 11, 16, 9, 20, 7, 13}
```

```
Union[u, w]
```

```
{3, 4, 7, 8, 9, 10, 11, 12, 13, 15, 16, 20, 21}
```

1.8.3. Dimensão (comprimento) de um vector

- **Pode obter-se o comprimento do vector com** `Length[]` **ou** `Dimensions[]`**...**

```
Length[v]
```

```
5
```

```
Dimensions[v][[1]]
```

```
5
```

1.8.4. Trata-se mesmo de um vector?

```
VectorQ[v]
```

```
True
```

```
VectorQ[Pi]
```

```
False
```

```
VectorQ[{{2, 3}, 5}]
```

```
False
```

1.8.5. Geração de vectores

- **Podemos gerar vectores, através das funções** `Table[]` **e** `Range[]`

```
Table[i^2 / (1 + i), {i, 4}]
```

$\left\{\frac{1}{2}, \frac{4}{3}, \frac{9}{4}, \frac{16}{5}\right\}$

Lista dos valores de 2i, para i entre 4 e 16 de 3 em 3

```
Table[2 * i, {i, 4, 16, 3}]
```

```
{8, 14, 20, 26, 32}
```

Lista dos complexos 3k^2+2kI, para k entre 1 e 3

```
Table[3 * k^2 + 2 * k * I, {k, 3}]
```

```
{3 + 2 i, 12 + 4 i, 27 + 6 i}
```

Lista de 5 inteiros aleatórios entre -9 e 9 divididos por 2

$\texttt{Table}\left[\frac{1}{2}\,\texttt{RandomInteger}[\{-9, 9\}], \{5\}\right]$

$\left\{-1, \frac{3}{2}, -\frac{9}{2}, 2, 4\right\}$

Vector de 6 uns e zeros alternados. OddQ[i] dá True sse i é ímpar e EvenQ[i] dá True sse i é par

```
Table[If[OddQ[i], 1, 0], {i, 6}]
```

```
{1, 0, 1, 0, 1, 0}
```

Vector dos primeiros 10 números primos...

```
Table[Prime[i], {i, 10}]
```

```
{2, 3, 5, 7, 11, 13, 17, 19, 23, 29}
```

Lista de 5 inteiros entre 1 e 5

```
Range[5]
```

```
{1, 2, 3, 4, 5}
```

Lista dos inteiros entre 4 e 8

```
Range[4, 8]
```

```
{4, 5, 6, 7, 8}
```

Lista dos racionais entre 3/4 e 13/2, de 3/4 em 3/4

```
Range[3 / 4, 13 / 2, 3 / 4]
```

$$\left\{\frac{3}{4}, \frac{3}{2}, \frac{9}{4}, 3, \frac{15}{4}, \frac{9}{2}, \frac{21}{4}, 6\right\}$$

```
Remove["Global`*"]
```

- **Podemos gerar vectores simbólicos com a função** `Array[]`

```
v = Array[y, {6}]
```

```
{y[1], y[2], y[3], y[4], y[5], y[6]}
```

1.8.6. Operações vectoriais

- **Podemos somar, subtrair e multiplicar os elementos de vectores**

Soma dos elementos de um vector

```
Sum[v[[i]], {i, Length[v]}]
```

```
y[1] + y[2] + y[3] + y[4] + y[5] + y[6]
```

Produto dos elementos de um vector

```
Product[v[[i]], {i, Length[v]}]
```

```
y[1] y[2] y[3] y[4] y[5] y[6]
```

Multiplicação de escalar por vector

```
u = 3 * v
```

```
{3 y[1], 3 y[2], 3 y[3], 3 y[4], 3 y[5], 3 y[6]}
```

Soma e subtração de vectores

```
{u + v, u - v}
```

```
{{4 y[1], 4 y[2], 4 y[3], 4 y[4], 4 y[5], 4 y[6]}, {2 y[1], 2 y[2], 2 y[3], 2 y[4], 2 y[5], 2 y[6]}}
```

2

Matrizes

2.1 Introdução

No presente capítulo abordaremos a noção de matriz de um ponto de vista autónomo, isto é, sem relevar a sua correspondência com os homomorfismos de espaços vectoriais de dimensão finita, tema que será tratado no capítulo 3. Serão estudados os espaços vectoriais de matrizes, bem como a álgebra linear (e anel) das matrizes quadradas.

A interpretação das matrizes como sequências de vectores de um espaço cartesiano conduzir-nos-á à noção de característica de uma matriz e a proposição 1.18 vai permitir a fundamentação do importante algoritmo – dito de condensação – para a sua determinação. O paralelismo entre a proposição 1.18 e os princípios de equivalência dos sistemas de equações lineares levam ao método de eliminação de Gauss-Jordan[(1)] para a resolução daquele importante tipo de sistemas. A aplicação de métodos matriciais aos sistemas de equações lineares são um bom exemplo prático do poder unificador da ferramenta matricial. Por último, os resultados obtidos sobre sistemas de equações vão permitir-nos o estudo do problema da inversão de uma matriz (não necessariamente quadrada), a obtenção de um algoritmo de condensação para a determinação da(s) inversa(s) e ainda estudar o problema da divisão matricial esquerda e direita. Abordamos também a teoria das matrizes elementares e a sua relação com a inversão, a característica e a equivalência de matrizes. Por último, trataremos das mudanças de base nos espaços vectoriais de dimensão finita.

2.2 Noção de matriz sobre um corpo. Alguns tipos de matrizes

Começaremos por definir a noção de matriz:

Definição 2.1 – Matriz sobre um corpo – *Sejam $m, n \in \mathbb{N}$ números naturais, $\mathbb{K}$ um corpo e considerem-se os intervalos $[1, m]$ e $[1, n]$ de $\mathbb{N}$. Chamaremos* ***matriz do tipo** $m \times n$ sobre o corpo $\mathbb{K}$ a uma família $A = [a_{ij}]_{(i,j)\in[1,m]\times[1,n]}$ de elementos de $\mathbb{K}$ indexada por $[1, m] \times [1, n]$, isto é a uma função $A\colon [1, m] \times [1, n] \to \mathbb{K}$; $(i, j) \mapsto a_{ij}$ [(2)]. O escalar de $\mathbb{K}$ que é o valor de A no ponto $(i, j) \in [1, m] \times [1, n]$ é designado por a_{ij} (ou A_{ij}) e a matriz A é também designada por $[a_{ij}]_{\substack{1\le i\le m\\ 1\le j\le n}}$ ou, mais simplesmente, $[a_{ij}]$[(3)], quando não houver interesse em especificar o tipo $m \times n$ da matriz.*

Uma matriz fica completamente determinada pelos seus valores nos $m \times n$ pontos do domínio $[1, m] \times [1, n]$ e, portanto, pode ser representada por uma tabela de duas entradas constituída por m linhas e n colunas de escalares a_{ij} do corpo $\mathbb{K}$ colocados entre parêntesis rectos[(3)]. Como o primeiro índice i é constante em cada linha da tabela, diz-se que ele é o *índice das linhas* da matriz; do mesmo modo e sendo o segundo índice j constante em cada coluna, j é

---

1 *Gauss, Carl Friedrich*: matemático, físico e astrónomo alemão (Brunswick 1777 – Göttingen 1855).
Jordan, Marie Ennemond Camille: matemático francês (Lyon 1838 – Paris 1922).

2 Mais geralmente, podem considerar-se matrizes de tipo $m \times n$ sobre qualquer conjunto X, como sendo as aplicações de $[1, m] \times [1, n]$ em X, sendo o seu conjunto designado por $X^{m,n}$.

3 Também se usam as notações $(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$, $\|a_{ij}\|_{\substack{1\le i\le m\\ 1\le j\le n}}$ e ainda

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \text{ ou } \left\| \begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{matrix} \right\|.$$

chamado o *índice das colunas* da matriz.

$$A = [a_{ij}]_{\substack{1\le i\le m\\1\le j\le n}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{2.1}$$

A matriz diz-se *racional*, *real* ou *complexa* consoante $\mathbb{K} = \mathbb{Q}$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$ e designaremos o conjunto de todas as matrizes do tipo $m \times n$ sobre um corpo $\mathbb{K}$ por $\mathbb{K}^{m,n}$.

É óbvio que, se o corpo $\mathbb{K}$ for finito (por exemplo, $\mathbb{Z}_p$ – os inteiros módulo p – onde p é um número primo)[4] o mesmo acontecerá com $\mathbb{K}^{m,n}$, tendo-se

$$\#(\mathbb{K}^{m,n}) = (\#\mathbb{K})^{mn} \tag{2.2}$$

Se o corpo $\mathbb{K}$ for infinito (casos de $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$), o mesmo sucederá com $\mathbb{K}^{m,n}$.

Da definição de matriz, resulta que duas matrizes são *iguais* sse têm o mesmo tipo e coincidem em todos os pontos $(i,j) \in [1,m] \times [1,n]$, isto é,

$$[a_{ij}]_{\substack{1\le i\le m\\1\le j\le n}} = [b_{ij}]_{\substack{1\le i\le m'\\1\le j\le n'}} \Leftrightarrow m = m' \wedge n = n' \wedge \underset{\substack{1\le i\le m\\1\le j\le n}}{\forall} a_{ij} = b_{ij} \tag{2.3}$$

Para cada $1 \le i \le m$, a sequência $(a_{ij})_{1\le j\le n}$ é chamada a *linha* i de A (ou iª linha de A) e, para cada $1 \le j \le n$, a sequência $(a_{ij})_{1\le i\le m}$ é chamada a *coluna* j (ou jª coluna de A). O vector de $\mathbb{K}^n$ cujas componentes constituem a linha i de A é chamado o iº *vector-linha* de A e o vector de $\mathbb{K}^m$ cujas componentes constituem a coluna j de A é o jº *vector-coluna* de A. Deste modo, pode a matriz A ser considerada como uma lista de m vectores de $\mathbb{K}^n$ (as linhas de A) ou como uma lista de n vectores de $\mathbb{K}^m$ (as colunas de A). Da mesma maneira que as *filas* (termo que designa, indistintamente, as linhas ou colunas) de uma matriz do tipo $m \times n$ se identificam com vectores de $\mathbb{K}^n$ ou de $\mathbb{K}^m$, uma matriz do tipo 1×1 é também identificada com o escalar único a_{11} que a constitui. As observações anteriores mostram que se pode definir a noção de matriz com base na noção de lista: trata-se de uma lista de m listas (as linhas) cada uma das quais com igual número n de escalares de $\mathbb{K}$. Pode aplicar-se recursivamente esta ideia e constituirem-se listas de q matrizes $[a_{ijk}]$ (ou seja, listas de listas de listas de escalares) e assim sucessivamente, obtendo-se a noção recursiva de *tensor* de ordem p[5], como uma lista de tensores de ordem $p-1$; a ordem é, afinal, o número de índices presentes em $\left[a_{i_1 i_2 \cdots i_p}\right]$; uma matriz $[a_{ij}]$ é, portanto, um *tensor de 2ª ordem*, um vector (v_i) é um *tensor de 1ª ordem*, considerando-se um escalar como um *tensor de ordem 0*.

Se $m = 1$, a matriz A constitui uma *matriz-linha* ou *vector-linha* e consiste no vector $(a_{11}, a_{12}, \ldots, a_{1n}) \in \mathbb{K}^n$ matricialmente, continuaremos a escrever

$$A = [\, a_{11} \;\; a_{12} \;\; \cdots \;\; a_{1n} \,] \in \mathbb{K}^{1,n}$$

---

4 Trata-se dos corpos $(\mathbb{Z}_p, \oplus, \otimes)$ onde $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$, p é um número primo e as operações são definidas por: $x \oplus y = (x+y) \bmod p$ e $x \otimes y = (xy) \bmod p$. Nestas igualdades, $a \bmod b$ designa o resto da divisão inteira de a por b.

5 Corresponde, nas linguagens de programação, à noção de “array” de dimensão p, sendo que as matrizes são os “arrays” de dimensão 2.

Se $n = 1$, tem-se uma *matriz-coluna* ou *vector-coluna* $(a_{11}, a_{21}, \ldots, a_{m1}) \in \mathbb{K}^m$ que, matricialmente, se escreve

$$A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \in \mathbb{K}^{m,1}$$

Adiante referiremos a operação de *transposição* matricial que nos permitirá escrever a matriz-coluna de uma forma tipograficamente mais conveniente, do modo seguinte

$$A = [\, a_{11} \;\; a_{21} \;\; \cdots \;\; a_{m1} \,]^{\mathrm{T}} \in \mathbb{K}^{m,1}$$

Se $m = n$, a matriz diz-se *quadrada de ordem* n e o conjunto destas matrizes é designado por $\mathbb{K}^{n,n}$; se $m \neq n$, dizemos que a matriz é *rectangular*.

Numa matriz $[a_{ij}]$ quadrada de ordem n, os elementos a_{ii} constituem a *diagonal principal*, 1ª diagonal ou diagonal descendente; os elementos a_{ij}, com $i + j = n + 1$, constituem a *diagonal secundária*, 2ª diagonal ou diagonal ascendente. Numa matriz rectangular, chama-se por vezes, *diagonal* à lista dos elementos da forma a_{ii} ($1 \leq i \leq \min\{m, n\}$).

Definem-se vários tipos de matrizes, conforme a configuração dos escalares presentes em (2.1). Seguem-se alguns exemplos:

Exemplo 2.1 Se $m = n$ e $a_{ij} = 0$, para $i > j$, a matriz diz-se *triangular superior*; a título de exemplo, apresenta-se uma matriz complexa triangular superior:

$$A = \begin{bmatrix} i & 1+i & -3 \\ 0 & 2 & -i \\ 0 & 0 & 0 \end{bmatrix}$$

Exemplo 2.2 Se $m = n$ e $a_{ij} = 0$, para $i < j$, a matriz diz-se *triangular inferior*; por exemplo, a seguinte matriz real é triangular inferior:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & -1 & 0 \\ 1 & 0 & 4 \end{bmatrix}$$

Exemplo 2.3 Se uma matriz é simultaneamente triangular superior e inferior, isto é, $m = n$ e $a_{ij} = 0$, para $i \neq j$, a matriz diz-se *matriz diagonal* e, sendo a_{ii} os elementos da diagonal principal, escreve-se diag$(a_{11}, a_{22}, \ldots, a_{nn})$; por exemplo, a seguinte matriz real é diagonal:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix} = \mathrm{diag}(1, -1, 4)$$

Uma matriz rectangular de tipo $m \times n$, chama-se *matriz diagonal* sse for $a_{ij} = 0$, para $i \neq j$, isto é, são nulos os elementos de A que não estejam na diagonal, por exemplo,

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Exemplo 2.4 Uma matriz diagonal quadrada cujos elementos diagonais são iguais entre si chama-se *matriz escalar* (adiante justificaremos esta denominação), isto é, $m = n$ e $a_{ij} = 0$, para $i \neq j$, e $a_{ii} = k$; a seguinte matriz real A é matriz escalar:

$$A = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \text{diag}(-2, -2, -2)$$

Exemplo 2.5 Entre as matrizes escalares refiram-se as matrizes *identidade*, em que $k = 1$ (mais tarde, veremos a razão deste nome). A matriz identidade de ordem n é designada por I_n ou simplesmente I, quando a ordem n for subentendida. O respectivo elemento genérico é o símbolo ou *delta de Kronecker* δ_{ij} (ver nota de rodapé 7 no capítulo 1), definido a seguir

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{2.4}$$

A título de exemplo, mostram-se as matrizes identidade de 2ª e 3ª ordens

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Exemplo 2.6 As matrizes $m \times n$ cujos elementos são todos nulos são chamadas *matrizes nulas* e designam-se por $O_{m,n}$ ou, sendo quadradas de ordem n, simplesmente por O_n. Caso não haja necessidade de indicar o tipo, usa-se apenas O. A seguir, mostram-se exemplos:

$$O_{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad O_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Exemplo 2.7 As matrizes de tipo $m \times n$, cujos elementos são $\left[h_{ij} = \frac{1}{i+j-1}\right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ são denominadas *matrizes de Hilbert* e designam-se por $H_{m,n}$. Usa-se a notação H_n para a matriz de Hilbert quadrada de ordem n. A seguir, mostra-se um exemplo:

$$H_3 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$$

Exemplo 2.8 Uma matriz quadrada de ordem n cujos elementos são tais que

$$a_{ij} = 0, \text{ se } i > j + 1 \vee j > i + 1$$

é chamada *matriz tridiagonal* (note que qualquer matriz de ordem $n \leq 2$ é tridiagonal). A seguir, mostra-se um exemplo de uma matriz real tridiagonal de 4ª ordem:

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ -3 & -1 & 3 & 0 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Exemplo 2.9 Diremos que uma matriz $E \in \mathbb{K}^{m,n}$ é *escalonada* sse as suas linhas formam uma lista escalonada de vectores de $\mathbb{K}^n$ (ver definição 1.10). Assim, são escalonadas as matrizes

$$A = \begin{bmatrix} 2 & 1 & -3 \\ 0 & -1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 2 & -1 & 0 & 1 \\ 0 & 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} -3 & 1 & 0 & 2 & 2 \\ 0 & 2 & -1 & 1 & 0 \\ 0 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Na secção 2.16, ilustra-se a construção de alguns dos tipos anteriores de matrizes, usando o software MATHEMATICA©.

2.3 Espaço linear das matrizes

Nesta secção, iremos definir um certo número de operações algébricas sobre matrizes que nos levarão à construção dos espaços vectoriais $\mathbb{K}^{m,n}$ e das álgebras $\mathbb{K}^{n,n}$ de matrizes quadradas de ordem ordem n. Comecemos por definir a adição matricial:

Definição 2.2 – Adição de matrizes – *Dadas duas matrizes $A = [a_{ij}]$ e $B = [b_{ij}]$ sobre um corpo $\mathbb{K}$ e do mesmo tipo $m \times n$ (ou seja, $A, B \in \mathbb{K}^{m,n}$), chamaremos soma de A e B à nova matriz $A + B \in \mathbb{K}^{m,n}$ cujos elementos c_{ij} estão definidos por:*

$$c_{ij} = a_{ij} + b_{ij};\ i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n \tag{2.5}$$

ou ainda

$$(A + B)_{ij} = A_{ij} + B_{ij};\ i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n \tag{2.6}$$

Portanto, as matrizes somam-se, somando (em $\mathbb{K}$) os elementos homólogos das matrizes A e B e só faz sentido adicionar matrizes do mesmo tipo. As propriedades formais da adição em $\mathbb{K}$, à custa da qual se definiu a adição matricial, implicam imediatamente que esta adição é:

- *Associativa*:

$$\underset{A,B,C \in \mathbb{K}^{m,n}}{\forall} (A + B) + C = A + (B + C)$$

- *Comutativa*:

$$\underset{A,B \in \mathbb{K}^{m,n}}{\forall} A + B = B + A$$

- *Tem elemento neutro, que é a matriz nula* $O_{m,n}$:

$$\underset{O_{m,n}\in\mathbb{K}^{m,n}}{\exists}\ \underset{A\in\mathbb{K}^{m,n}}{\forall} A + O_{m,n} = O_{m,n} + A = A$$

- *Toda a matriz* A *tem simétrica (única)* $-A$:

$$\underset{A\in\mathbb{K}^{m,n}}{\forall}\ \underset{-A\in\mathbb{K}^{m,n}}{\exists} A + (-A) = (-A) + A = O_{m,n}$$

As propriedades anteriores mostram que $(\mathbb{K}^{m,n}, +)$ constitui um *grupo comutativo aditivo*. Como em qualquer grupo aditivo, define-se a *subtracção* por

$$A - B = A + (-B) \tag{2.7}$$

o que é equivalente a

$$(A - B)_{ij} = A_{ij} - B_{ij};\ i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n \tag{2.8}$$

Com vista a obter um espaço vectorial, vamos a seguir definir a operação de *produto* de um escalar de $\mathbb{K}$ por uma matriz do tipo $m \times n$ sobre $\mathbb{K}$. Trata-se de uma *lei de composição externa* definida em $\mathbb{K} \times \mathbb{K}^{m,n}$ e com valores em $\mathbb{K}^{m,n}$:

Definição 2.3 – Produto de escalar por matriz – *Dado um escalar* $\alpha \in \mathbb{K}$ *e uma matriz* $A = [a_{ij}]$ *sobre um corpo* $\mathbb{K}$ *do tipo* $m \times n$ *(ou seja,* $A \in \mathbb{K}^{m,n}$*), chamaremos produto de* α *por* A *à nova matriz* $\alpha A \in \mathbb{K}^{m,n}$ *cujos elementos* b_{ij} *estão definidos por:*

$$b_{ij} = \alpha\, a_{ij};\ i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n \tag{2.9}$$

ou ainda

$$(\alpha A)_{ij} = \alpha A_{ij};\ i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n \tag{2.10}$$

Portanto, multiplica-se um escalar por uma matriz, multiplicando (em $\mathbb{K}$) esse escalar por cada um dos elementos da matriz. Das propriedades do produto no corpo $\mathbb{K}$, resultam imediatamente as seguintes propriedades para o produto definido por (2.9):

- *Distributividade em relação à adição escalar*:

$$\underset{\substack{\alpha,\beta\in\mathbb{K}\\ A\in\mathbb{K}^{m,n}}}{\forall} (\alpha + \beta)A = \alpha A + \beta A$$

- *Distributividade em relação à adição matricial*:

$$\underset{\substack{\alpha\in\mathbb{K}\\ A,B\in\mathbb{K}^{m,n}}}{\forall} \alpha(A + B) = \alpha A + \alpha B$$

- *Associatividade mista*:

$$\underset{\substack{\alpha,\beta\in\mathbb{K}\\ A\in\mathbb{K}^{m,n}}}{\forall} \alpha(\beta A) = (\alpha\beta)A$$

- A unidade do corpo é *elemento neutro à esquerda* no produto de escalar por matriz:

$$\underset{A\in\mathbb{K}^{m,n}}{\forall} 1A = A$$

As propriedades anteriores constituem os axiomas A1-A4 e P1-P4 do capítulo 1 e significam, portanto, que $\mathbb{K}^{m,n}$ é, com as operações definidas atrás, um *espaço vectorial* sobre o corpo $\mathbb{K}$ em que os "vectores" são as matrizes do tipo $m \times n$ de elementos em $\mathbb{K}$. A família de mn matrizes $(C_{\alpha\beta})_{\substack{1\le\alpha\le m\\1\le\beta\le n}}$ definidas por

$$(C_{\alpha\beta})_{ij} = \delta_{i\alpha}\delta_{j\beta} \tag{2.11}$$

formam uma base de $\mathbb{K}^{m,n}$, na qual as *coordenadas* de uma matriz $X = [x_{ij}] \in \mathbb{K}^{m,n}$ são os próprios escalares x_{ij} que a constituem:

$$X = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}C_{ij}$$

Por esta razão, a base mencionada é chamada *base canónica* de $\mathbb{K}^{m,n}$. O que acabou de se expor mostra que

$$\dim_{\mathbb{K}}\mathbb{K}^{m,n} = mn \tag{2.12}$$

Por exemplo, a base canónica de $\mathbb{C}^{2,2}$ (sobre $\mathbb{C}$) será

$$c = \left(\begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}0&1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$$

e a base canónica de $\mathbb{C}^{2,2}$ (sobre $\mathbb{R}$) é

$$c = \left(\begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}i&0\\0&0\end{bmatrix}, \begin{bmatrix}0&1\\0&0\end{bmatrix}, \begin{bmatrix}0&i\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&0\end{bmatrix}, \begin{bmatrix}0&0\\i&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}, \begin{bmatrix}0&0\\0&i\end{bmatrix}\right)$$

Pela proposição 1.15 e atendendo a (2.12), podemos afirmar que o espaço $\mathbb{K}^{m,n}$ é *isomorfo* do espaço *cartesiano* $\mathbb{K}^{mn}$ e é óbvio que, por exemplo, a aplicação

$$[a_{ij}]_{\substack{1\le i\le m\\1\le j\le n}} \mapsto (a_{11}, a_{12}, \dots, a_{1n}, a_{21}, a_{22}, \dots, a_{2n}, \dots, a_{m1}, a_{m2}, \dots, a_{mn})$$

é um *isomorfismo* de $\mathbb{K}^{m,n}$ sobre $\mathbb{K}^{mn}$.

Sendo $\mathbb{K}^{m,n}$ um espaço vectorial sobre $\mathbb{K}$, podemos recorrer às proposições 1.1, 1.2 e 1.3 para concluir imediatamente que:

$$\underset{\substack{\alpha\in\mathbb{K}\\A\in\mathbb{K}^{m,n}}}{\forall} \big(\alpha A = O_{m,n} \Leftrightarrow \alpha = 0 \vee A = O_{m,n}\big)$$

$$\underset{\substack{\alpha\in\mathbb{K}\\A\in\mathbb{K}^{m,n}}}{\forall} \big(-(\alpha A) = (-\alpha)A = \alpha(-A)\big)$$

$$\underset{\substack{\alpha,\beta\in\mathbb{K}\\A\in\mathbb{K}^{m,n}}}{\forall} \big((\alpha-\beta)A = \alpha A - \beta A\big)$$

$$\underset{\substack{\alpha\in\mathbb{K}\\A,B\in\mathbb{K}^{m,n}}}{\forall} \big(\alpha(A-B) = \alpha A - \alpha B\big)$$

Observe-se, por último, que uma matriz escalar de ordem n será sempre da forma kI_n, onde $k \in \mathbb{K}$ é um escalar

$$\text{diag}(k, k, \dots, k) = kI_n$$

2.4 Álgebra e Anel das matrizes quadradas

Vamos, agora, definir uma terceira operação matricial: o produto de matrizes.

Definição 2.4 – Produto matricial – *Sejam $m, n, p \in \mathbb{N}$ e $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{n,p}$ matrizes dos tipos $m \times n$ e $n \times p$ sobre um corpo $\mathbb{K}$. O produto das matrizes A e B (por esta ordem) é a matriz $AB \in \mathbb{K}^{m,p}$ do tipo $m \times p$, cujos elementos c_{ij} se definem por:*

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj};\ i = 1, 2, \dots, m;\ j = 1, 2, \dots, p \tag{2.13}$$

ou seja

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj};\ i = 1, 2, \dots, m;\ j = 1, 2, \dots, p \tag{2.13}$$

As mp igualdades (2.13) mostram que o elemento $(AB)_{ij}$ da linha i e coluna j de AB se calcula mediante a soma dos produtos dos n elementos da linha i de A pelos n elementos respectivos da coluna j de B e que o número de linhas de AB é o número de linhas de A, sendo o seu número de colunas igual ao número de colunas de B. Note-se que a condição prévia para que o produto AB, por esta ordem, esteja definido é que o número de colunas de A seja igual ao número de linhas de B (igual a n). Daqui resulta imediatamente que para que uma matriz A se possa multiplicar por si mesma ($A = B$), é necessário que ela seja *quadrada* o que significa que a *potência* de base A e expoente natural só estará definida quando A for quadrada. Convém ainda observar que o produto matricial será uma *lei de composição interna* (em $\mathbb{K}^n$), apenas quando $m = n = p$.

Nas condições da definição anterior e se $m \neq p$, o produto BA não está definido. Se $m = p \neq n$, então AB e BA estão definidos mas $AB \in \mathbb{K}^{m,m}$ enquanto que $BA \in \mathbb{K}^{n,n}$, o que significa que será, certamente, $AB \neq BA$. Finalmente, se tivermos $m = n = p$, AB e BA serão ambas matrizes quadradas de ordem n, mas, ainda assim, as igualdades (2.13) implicam que, em geral, será $AB \neq BA$, como mostra o exemplo seguinte:

Exemplo 2.10 Considerem-se as matrizes

$$A = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} \text{ e } B = \begin{bmatrix} 1 & -3 \\ 2 & -1 \end{bmatrix}$$

ter-se-á, neste caso,

$$AB = \begin{bmatrix} 4 & -7 \\ -3 & -1 \end{bmatrix} \text{ e } BA = \begin{bmatrix} -1 & 7 \\ 3 & 4 \end{bmatrix} \neq AB$$

Do exposto, podemos concluir que, para ser $AB = BA$, é necessário (mas não suficiente) que A e B sejam matrizes *quadradas e da mesma ordem*. Existem, de facto, casos de matrizes quadradas em que $AB = BA$ e, quando tal acontecer, diremos que A e B são matrizes *permutáveis* ou *comutáveis*. É o que acontece, entre outros, nos casos a seguir indicados:

Exemplo 2.11 Caso de ser nulo um dos factores:

$$\underset{A\in\mathbb{K}^{n,n}}{\forall}\big(AO_n = O_nA = O_n\big)$$

Para quaisquer $i, j = 1, 2, \ldots, n$, as igualdades (2.13) implicam imediatamente

$$\begin{cases} (AO_n)_{ij} = \displaystyle\sum_{k=1}^{n} A_{ik}(O_n)_{kj} = \sum_{k=1}^{n} A_{ik}0 = \sum_{k=1}^{n} 0 = 0 = (O_n)_{ij} \\ (O_nA)_{ij} = \displaystyle\sum_{k=1}^{n} (O_n)_{ik}A_{kj} = \sum_{k=1}^{n} 0A_{kj} = \sum_{k=1}^{n} 0 = 0 = (O_n)_{ij} \end{cases}$$

Exemplo 2.12 Caso de um dos factores ser a matriz identidade I_n:

$$\underset{A\in\mathbb{K}^{n,n}}{\forall}\big(AI_n = I_nA = A\big)$$

As igualdades (2.13) levam directamente, para $i, j = 1, 2, \ldots, n$ e onde δ_{ij} é o símbolo de Kronecker, às igualdades

$$\begin{cases} (AI_n)_{ij} = \displaystyle\sum_{k=1}^{n} A_{ik}(I_n)_{kj} = \sum_{k=1}^{n} A_{ik}\delta_{kj} = A_{ij} \\ (I_nA)_{ij} = \displaystyle\sum_{k=1}^{n} (I_n)_{ik}A_{kj} = \sum_{k=1}^{n} \delta_{ik}A_{kj} = A_{ij} \end{cases}$$

Exemplo 2.13 Sejam A e B as matrizes seguintes

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & -2 \\ 2 & 1 & -2 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & -1 & 1 \\ -2 & -6 & 5 \\ -1 & -4 & 3 \end{bmatrix}$$

Verifica-se, neste caso, que

$$AB = BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

De tudo o que se viu anteriormente, pode concluir-se que o produto matricial *não é comutativo* e a discussão que fizémos pode ser condensada do seguinte modo:

$$\left.\begin{matrix} A \in \mathbb{K}^{m,n} \\ B \in \mathbb{K}^{n,p} \end{matrix}\right\} \Rightarrow \begin{cases} m \neq p \Rightarrow \text{Existe } AB, \text{ mas não existe } BA. \\ m = p \Rightarrow \begin{cases} m \neq n \Rightarrow AB \in \mathbb{K}^{m,m} \wedge BA \in \mathbb{K}^{n,n} \Rightarrow AB \neq BA \\ m = n \Rightarrow \begin{cases} AB, BA \in \mathbb{K}^{n,n} \Rightarrow \text{Em geral, será } AB \neq BA, \\ \text{mas existem matrizes permutáveis.} \end{cases} \end{cases} \end{cases}$$

A seguinte proposição resume algumas das propriedades do produto matricial:

Proposição 2.1 – Propriedades do produto matricial – *Sejam $m, n, p, q \in \mathbb{N}$ e $\mathbb{K}$ um corpo. O produto de matrizes sobre $\mathbb{K}$ satisfaz as seguintes propriedades:*

i) *Associatividade: para quaisquer matrizes $A \in \mathbb{K}^{m,n}$, $B \in \mathbb{K}^{n,p}$ e $C \in \mathbb{K}^{p,q}$, tem-se*

$$(AB)C = A(BC)$$

ii) *Distributividade à direita em relação à adição (e subtracção) matriciais: sendo $A, B \in \mathbb{K}^{m,n}$ e $C \in \mathbb{K}^{n,p}$, tem-se:*

$$(A + B)C = AC + BC$$

iii) *Distributividade à esquerda em relação à adição (e subtracção) matriciais: sendo $A \in \mathbb{K}^{m,n}$ e $B, C \in \mathbb{K}^{n,p}$, tem-se:*

$$A(B + C) = AB + AC$$

iv) *Para quaisquer matrizes $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{n,p}$ e qualquer escalar $\alpha \in \mathbb{K}$, verificam-se as igualdades seguintes*

$$(\alpha A)B = A(\alpha B) = \alpha(AB)$$

v) *Existência de elementos neutros (as matrizes identidade) à esquerda e à direita: para qualquer matriz $A \in \mathbb{K}^{m,n}$, tem-se:*

$$I_m A = A I_n = A$$

vi) *Anulamento do produto: para quaisquer matrizes $A \in \mathbb{K}^{m,n}$, $B \in \mathbb{K}^{n,p}$, tem-se*

$$A = O_{m,n} \vee B = O_{n,p} \Rightarrow AB = O_{m,p}$$

Observe-se que, ao contrário do que acontece com o produto de escalares, a recíproca desta implicação não é válida: AB pode ser a matriz nula, sem que nem A nem B o sejam.

Demonstração:

i) Para quaisquer $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, q$:

$$\begin{aligned}
((AB)C)_{ij} &= \sum_{r=1}^{p} (AB)_{ir} C_{rj} = \sum_{r=1}^{p} \left(\sum_{k=1}^{n} A_{ik} B_{kr} \right) C_{rj} = \sum_{r=1}^{p} \sum_{k=1}^{n} A_{ik} B_{kr} C_{rj} \\
&= \sum_{k=1}^{n} \sum_{r=1}^{p} A_{ik} B_{kr} C_{rj} = \sum_{k=1}^{n} \left(A_{ik} \sum_{r=1}^{p} B_{kr} C_{rj} \right) = \sum_{k=1}^{n} A_{ik} \left(\sum_{r=1}^{p} B_{kr} C_{rj} \right) \\
&= \sum_{k=1}^{n} A_{ik} (BC)_{kj} = (A(BC))_{ij}
\end{aligned}$$

ii) Neste caso, a igualdade resulta de ser, para $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, p$:

$$\begin{aligned}((A+B)C)_{ij} &= \sum_{k=1}^{n}(A+B)_{ik}C_{kj} = \sum_{k=1}^{n}(A_{ik}+B_{ik})C_{kj} = \sum_{k=1}^{n}(A_{ik}C_{kj}+B_{ik}C_{kj})\\ &= \sum_{k=1}^{n}A_{ik}C_{kj} + \sum_{k=1}^{n}B_{ik}C_{kj} = (AC)_{ij} + (BC)_{ij} = (AC+BC)_{ij}\end{aligned}$$

iii) Tal como anteriormente, teremos, para $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, p$:

$$\begin{aligned}(A(B+C))_{ij} &= \sum_{k=1}^{n}A_{ik}(B+C)_{kj} = \sum_{k=1}^{n}A_{ik}(B_{kj}+C_{kj}) = \sum_{k=1}^{n}(A_{ik}B_{kj}+A_{ik}C_{kj})\\ &= \sum_{k=1}^{n}A_{ik}B_{kj} + \sum_{k=1}^{n}A_{ik}C_{kj} = (AB)_{ij} + (AC)_{ij} = (AB+AC)_{ij}\end{aligned}$$

iv) Resulta imediatamente das igualdades seguintes, para $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, p$:

$$\begin{cases}((\alpha A)B)_{ij} = \displaystyle\sum_{k=1}^{n}(\alpha A)_{ik}B_{kj} = \sum_{k=1}^{n}(\alpha A_{ik})B_{kj}\\ \qquad\qquad\quad = \displaystyle\sum_{k=1}^{n}\alpha(A_{ik}B_{kj}) = \alpha\sum_{k=1}^{n}A_{ik}B_{kj} = \alpha(AB)_{ij}\\ (A(\alpha B))_{ij} = \displaystyle\sum_{k=1}^{n}A_{ik}(\alpha B)_{kj} = \sum_{k=1}^{n}A_{ik}(\alpha B_{kj})\\ \qquad\qquad\quad = \displaystyle\sum_{k=1}^{n}\alpha(A_{ik}B_{kj}) = \alpha\sum_{k=1}^{n}A_{ik}B_{kj} = \alpha(AB)_{ij}\end{cases}$$

v) Tem-se, para $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$ e sendo δ_{ij} o símbolo de Kronecker:

$$\begin{cases}(I_mA)_{ij} = \displaystyle\sum_{k=1}^{m}(I_m)_{ik}A_{kj} = \sum_{k=1}^{m}\delta_{ik}A_{kj} = A_{ij}\\ (AI_n)_{ij} = \displaystyle\sum_{k=1}^{n}A_{ik}(I_n)_{kj} = \sum_{k=1}^{n}A_{ik}\delta_{kj} = A_{ij}\end{cases}$$

vi) Tem-se, para $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, p$:

$$\begin{cases}(O_{m,n}B)_{ij} = \displaystyle\sum_{k=1}^{n}(O_{m,n})_{ik}B_{kj} = \sum_{k=1}^{n}0\,B_{kj} = \sum_{k=1}^{n}0 = 0 = (O_{m,p})_{ij}\\ (AO_{n,p})_{ij} = \displaystyle\sum_{k=1}^{n}A_{ik}(O_{n,p})_{kj} = \sum_{k=1}^{n}A_{ik}\,0 = \sum_{k=1}^{n}0 = 0 = (O_{m,p})_{ij}\end{cases}$$

□

Exemplo 2.14 Um produto matricial pode ser nulo, sem que nenhum dos factores o seja, como mostra o exemplo seguinte

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3 & -2 \end{bmatrix} \wedge B = \begin{bmatrix} 1 & -1 \\ 2 & -2 \\ 3 & -3 \end{bmatrix} \Rightarrow AB = O_2$$

Exemplo 2.15 Note-se, de passagem, que as “leis do corte” à esquerda e à direita não são válidas (na mesma forma em que são válidas para os escalares). Por exemplo a lei do corte à esquerda

$$\underset{\substack{A\in\mathbb{K}^{m,n} \\ B,C\in\mathbb{K}^{n,p}}}{\forall} (A \neq O_{m,n} \wedge AB = AC \Rightarrow B = C)$$

não é válida, como mostra o exemplo seguinte

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3 & -2 \end{bmatrix}, B = \begin{bmatrix} 3 & 1 \\ 1 & 0 \\ 0 & -2 \end{bmatrix}, C = \begin{bmatrix} 2 & 2 \\ -1 & 2 \\ -3 & 1 \end{bmatrix}$$

O leitor pode verificar que, neste caso, será

$$A \neq O_{2,3} \wedge AB = AC = \begin{bmatrix} 5 & 2 \\ 3 & 4 \end{bmatrix} \wedge B \neq C$$

Outra situação que exige cuidado é a da multiplicação de ambos os membros de uma igualdade matricial por uma mesma matriz: há que fazer a multiplicação à esquerda em ambos os membros ou à direita em ambos os membros; por exemplo, não são, em geral, válidas as implicações

$$A = B \Rightarrow CA = BC \text{ e } A = B \Rightarrow AD = DB$$

Se $m = n$, o produto matricial constitui uma *lei de composição interna* em $\mathbb{K}^{n,n}$ e as quatro primeiras alíneas da proposição anterior são precisamente os axiomas M1 a M4 do capítulo 1. O produto matricial é, pois, uma função bilinear associativa em $\mathbb{K}^{n,n}$ e este é igualmente um espaço vectorial (de dimensão n^2) sobre $\mathbb{K}$: conclui-se que o conjunto $\mathbb{K}^{n,n}$ das matrizes quadradas de ordem n é (com a adição matricial (2.5), o produto por escalar (2.9) e com o produto matricial (2.13)) uma *Álgebra Linear* não comutativa sobre $\mathbb{K}$ e a penúltima alínea da proposição anterior mostra que esta álgebra tem *unidade* ou *identidade*, precisamente a matriz identidade de ordem n, I_n. É claro, igualmente, que $(\mathbb{K}^{n,n}, +, \times)$ constitui um *Anel* não comutativo com identidade, mas não um *Corpo* (se $n > 1$, o produto matricial não é comutativo); quando estudarmos a inversão matricial, veremos que existe uma razão adicional para que $(\mathbb{K}^{n,n}, +, \times)$ não seja um corpo, pois há em $\mathbb{K}^{n,n}$ elementos *singulares* não nulos.

Se $E = kI_n$ for uma *matriz escalar*, então, para qualquer matriz $A \in \mathbb{K}^{m,n}$,

$$AE = A(kI_n) = k(AI_n) = kA$$

ou seja, multiplicar A por uma matriz escalar E à direita equivale a multiplicar A pelo escalar k que forma a diagonal de E. Do mesmo modo, se $E = kI_m$ for uma matriz escalar, então, para qualquer matriz $A \in \mathbb{K}^{m,n}$

$$EA = (kI_m)A = k(I_mA) = kA$$

ou seja, multiplicar A por uma matriz escalar E à esquerda equivale a multiplicar A pelo escalar k que forma a diagonal de E: a isto se deve a designação de *matriz escalar* para as matrizes da forma kI_n, visto que elas se comportam na multiplicação por outra matriz como se de escalares se tratasse.

Para as matrizes quadradas (e só para estas) é possível definir, por recorrência, a potenciação com base matricial e expoente natural (para algumas matrizes quadradas, passaremos, mais tarde, a expoente inteiro negativo) , pondo:

Definição 2.5 – Potência de matriz e expoente natural – *Seja $A \in \mathbb{K}^{n,n}$ uma matriz quadrada de ordem n sobre um corpo $\mathbb{K}$. A potência de base A e expoente $k \in \mathbb{N}_0$ define-se pondo, por recorrência,*

$$\begin{cases} A^0 = I_n \\ A^k = A^{k-1}A, \; k > 0 \end{cases} \tag{2.14}$$

É claro que a aplicação repetida de (2.14) dá, para $k > 0$,

$$A^k = \underbrace{A\,A \cdots A}_{k \text{ factores}}$$

A potenciação goza de algumas das propriedades habituais, mas nem todas, como consta da

Proposição 2.2 – Propriedades da potência – *Para quaisquer matrizes $A, B \in \mathbb{K}^{n,n}$ e quaisquer $p, q \in \mathbb{N}_0$, tem-se:*

i) $A^pA^q = A^{p+q}$

ii) $(A^p)^q = A^{p\,q}$

iii) Se A e B são permutáveis, então

$$AB^p = B^pA$$

iv) Se A e B são permutáveis, então

$$A^pB^p = (AB)^p$$

(Se $p \in \{0, 1\}$, a igualdade é válida, mesmo que A e B não sejam permutáveis).

Demonstração:

i) Por indução em q: se $q = 0$, tem-se:

$$A^pA^0 = A^pI_n = A^p = A^{p+0}$$

Se $A^pA^q = A^{p+q}$, então:

$$A^pA^{q+1} = A^p(A^qA) = (A^pA^q)A = (A^{p+q})A = A^{(p+q)+1} = A^{p+(q+1)}$$

ii) Novamente por indução em q. Se $q = 0$, tem-se:

$$\big(A^p\big)^0 = I_n = A^0 = A^{p0}$$

Se $(A^p)^q = A^{pq}$, então:

$$\big(A^p\big)^{q+1} = \big(A^p\big)^q A^p = A^{pq}A^p = A^{pq+p} = A^{p(q+1)}$$

iii) Suponha-se $AB = BA$ e mostremos a igualdade $AB^p = B^pA$ por indução em p:

Se $p = 0$, será

$$AB^0 = AI_n = A = I_nA = B^0A$$

e, supondo que $AB = BA$ e $AB^p = B^pA$, então:

$$AB^{p+1} = A(B^pB) = (AB^p)B = (B^pA)B = B^p(AB) = B^p(BA) = (B^pB)A = B^{p+1}A$$

iv) De novo por indução em p:

Se $p = 0$, será

$$A^0B^0 = I_nI_n = B^0A^0$$

Por outro lado, supondo que $AB = BA$ e $A^pB^p = (AB)^p$, tem-se:

$$\begin{aligned} A^{p+1}B^{p+1} &= (A^pA)(B^pB) = A^p(AB^p)B = A^p(B^pA)B \\ &= (A^pB^p)(AB) = (AB)^p(AB) = (AB)^{p+1} \end{aligned}$$

□

As igualdades (2.5), (2.9) e (2.14) dão significado às expressões

$$p(A) = \sum_{k=0}^{m} c_kA^k = c_0I_n + c_1A + c_2A^2 + \cdots + c_mA^m \tag{2.15}$$

onde $m \geq 0$ é um inteiro, os $(c_k)_{0\leq k\leq m}$ são escalares de um corpo $\mathbb{K}$ e $A \in \mathbb{K}^{n,n}$ é matriz quadrada de ordem n: são os *polinómios matriciais* na matriz quadrada A, de coeficientes $c_k \in \mathbb{K}$ e de grau $\leq m$ e é óbvio que $p(A)$ é ainda uma matriz quadrada de ordem n.

Vejamos alguns exemplos:

Exemplo 2.16 Se A e B são matrizes permutáveis e se $p(B) = \sum\limits_{k=0}^{m} c_kB^k$ for um polinómio arbitrário em B, vamos ver, graças a algumas das propriedades da álgebra matricial e à 3ª alínea da proposição 2.2, que A e $p(B)$ são também permutáveis:

$$\begin{aligned} Ap(B) &= A\sum_{k=0}^{m} c_k B^k = \sum_{k=0}^{m} A\big(c_k B^k\big) = \sum_{k=0}^{m} c_k\big(AB^k\big) \\ &= \sum_{k=0}^{m} c_k\big(B^k A\big) = \sum_{k=0}^{m}\big(c_k B^k\big)A = \left(\sum_{k=0}^{m} c_k B^k\right)A = p(B)A \end{aligned}$$

O que acabámos de mostrar no exemplo anterior implica em particular que, para qualquer matriz quadrada A e qualquer polinómio p, as matrizes A *e* $p(A)$ *são sempre permutáveis*

$$Ap(A) = p(A)A \tag{2.16}$$

Exemplo 2.17 A não comutatividade do produto matricial acarreta a não validade de algumas manipulações algébricas que estamos habituados a fazer sobre escalares. Seguem-se alguns exemplos destas diferenças entre a manipulação de escalares e de matrizes:

- Em geral, a fórmula do *binómio de Newton* não é válida, isto é, para $m \geq 2$ inteiro e $A, B \in \mathbb{K}^{n,n}$, tem-se, geralmente:

$$(A+B)^m \neq \sum_{k=0}^{m} \binom{m}{k} A^{m-k} B^k \tag{2.17}$$

Por exemplo, para

$$A = \begin{bmatrix} 2 & 1 & -1 \\ -2 & 0 & 1 \\ 3 & 1 & -3 \end{bmatrix} \text{ e } B = \begin{bmatrix} -3 & -1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & 1 \end{bmatrix}$$

vem

$$\begin{cases} (A+B)^2 = \begin{bmatrix} 4 & 0 & -3 \\ 11 & 4 & -1 \\ -9 & 0 & 7 \end{bmatrix} \\ A^2 + 2AB + B^2 = \begin{bmatrix} -3 & 2 & 7 \\ 10 & 1 & 4 \\ -22 & 1 & 17 \end{bmatrix} \neq (A+B)^2 \end{cases}$$

Mas, para quaisquer matrizes $A, B \in \mathbb{K}^{n,n}$ será sempre, por exemplo,

$$\begin{aligned} (A+B)^2 &= (A+B)(A+B) = A^2 + AB + BA + B^2 \\ (A+B)^3 &= \big(A^2 + AB + BA + B^2\big)(A+B) \\ &= A^3 + A^2B + ABA + AB^2 + BA^2 + BAB + B^2A + B^3 \end{aligned} \tag{2.18}$$

■ Se $A, B \in \mathbb{K}^{n,n}$ são matrizes *permutáveis* e $m \in \mathbb{N}_0$ é um inteiro, então será válida a fórmula do *binómio de Newton* (ver secção 11 do apêndice B):

$$A \text{ e } B \text{ são } \textit{permutáveis} \Rightarrow (A+B)^m = \sum_{k=0}^{m} \binom{m}{k} A^{m-k} B^k \tag{2.19}$$

Por exemplo, seja $A \in \mathbb{R}^{3,3}$ a matriz já anteriormente referida e C o seguinte polinómio do 3º grau na matriz A

$$C = p(A) = A^3 + 3A^2 + A - 2I_3 = \begin{bmatrix} -1 & 5 & 1 \\ -8 & -7 & 1 \\ -1 & 3 & 0 \end{bmatrix}$$

Devido a (2.16), A e C são permutáveis (verifique!) e, portanto, o binómio de Newton é aplicável a A e C; por exemplo, para expoente $m = 3$, tem-se

$$\begin{cases} (A+C)^3 = \begin{bmatrix} 325 & -54 & -108 \\ 54 & 325 & 54 \\ 270 & 0 & -107 \end{bmatrix} \\ A^3 + 3A^2C + 3AC^2 + C^3 = \begin{bmatrix} 325 & -54 & -108 \\ 54 & 325 & 54 \\ 270 & 0 & -107 \end{bmatrix} = (A+C)^3 \end{cases}$$

■ A fórmula da *diferença de quadrados* não é válida, isto é, para $A, B \in \mathbb{K}^{n,n}$, tem-se, geralmente:

$$A^2 - B^2 \neq (A+B)(A-B) \tag{2.20}$$

Por exemplo, e de novo para as matrizes A e B anteriores, vem

$$\begin{cases} A^2 - B^2 = \begin{bmatrix} -9 & 2 & 9 \\ 0 & -1 & -12 \\ -4 & 3 & 9 \end{bmatrix} \\ (A+B)(A-B) = \begin{bmatrix} -2 & 0 & -1 \\ 1 & 2 & -17 \\ 9 & 2 & -1 \end{bmatrix} \neq A^2 - B^2 \end{cases}$$

Aqui, a fórmula que será sempre verificada é

$$A^2 - AB + BA - B^2 = (A+B)(A-B) \tag{2.21}$$

■ Se $A, B \in \mathbb{K}^{n,n}$ são matrizes *permutáveis*, será então válida a fórmula da *diferença de quadrados*

$$A \text{ e } B \text{ são } \textit{permutáveis} \Rightarrow A^2 - B^2 = (A+B)(A-B) \tag{2.22}$$

O leitor pode verificar que, para as matrizes permutáveis A e C já referidas, virá

$$\begin{cases} A^2 - C^2 = \begin{bmatrix} 39 & 38 & -2 \\ -64 & -13 & 14 \\ 18 & 26 & 5 \end{bmatrix} \\ (A+C)(A-C) = \begin{bmatrix} 39 & 38 & -2 \\ -64 & -13 & 14 \\ 18 & 26 & 5 \end{bmatrix} = A^2 - C^2 \end{cases}$$

2.5 Transposição e transconjugação

Vamos tratar agora de uma nova operação matricial que, ao contrário das anteriormente estudadas, é uma operação unária que transforma matrizes em novas matrizes e a que se chama *transposição*:

Definição 2.6 – Transposição – *Seja $\mathbb{K}$ um corpo e $m, n \in \mathbb{N}$. A* ***Transposta*** *de uma matriz $A \in \mathbb{K}^{m,n}$ é a matriz $A^{\mathrm{T}} \in \mathbb{K}^{n,m}$ (também designada por A^t ou A'), cujas linhas são as colunas de A (ou vice-versa, o que é, obviamente, equivalente). Sendo $\left(A^{\mathrm{T}}\right)_{ji}$ o elemento da linha j e da coluna i de A^{T} e A_{ij} o elemento da linha i e da coluna j de A, tem-se, portanto,*

$$\left(A^{\mathrm{T}}\right)_{ji} = A_{ij},\ i = 1, 2, \ldots, m \text{ e } j = 1, 2, \ldots, n \tag{2.23}$$

A transposição satisfaz um certo número de propriedades muito simples que se resumem a seguir:

Proposição 2.3 – Propriedades da transposição – *Seja $\mathbb{K}$ um corpo e $m, n, p \in \mathbb{N}$. Para quaisquer matrizes $A, B \in \mathbb{K}^{m,n}$, $C \in \mathbb{K}^{n,p}$ e qualquer escalar α, tem-se:*

i) $\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A$

ii) $(A \pm B)^{\mathrm{T}} = A^{\mathrm{T}} \pm B^{\mathrm{T}}$

iii) $(\alpha A)^{\mathrm{T}} = \alpha A^{\mathrm{T}}$

iv) $(AC)^{\mathrm{T}} = C^{\mathrm{T}} A^{\mathrm{T}}$

Demonstração:

i) Para qualquer $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$, virá:

$$\left(\left(A^{\mathrm{T}}\right)^{\mathrm{T}}\right)_{ij} = \left(A^{\mathrm{T}}\right)_{ji} = A_{ij}$$

ii) Tem-se, para $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$:

$$\left((A \pm B)^{\mathrm{T}}\right)_{ji} = (A \pm B)_{ij} = A_{ij} \pm B_{ij} = \left(A^{\mathrm{T}}\right)_{ji} \pm \left(B^{\mathrm{T}}\right)_{ji} = \left(A^{\mathrm{T}} \pm B^{\mathrm{T}}\right)_{ji}$$

iii) Neste caso e para $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$, ficará

$$\left((\alpha A)^{\mathrm{T}}\right)_{ji} = (\alpha A)_{ij} = \alpha A_{ij} = \alpha\left(A^{\mathrm{T}}\right)_{ji} = \left(\alpha A^{\mathrm{T}}\right)_{ji}$$

iv) Para $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, p$, tem-se:

$$\left((AC)^{\mathrm{T}}\right)_{ji} = (AC)_{ij} = \sum_{k=1}^{n} A_{ik}C_{kj} = \sum_{k=1}^{n}\left(A^{\mathrm{T}}\right)_{ki}\left(C^{\mathrm{T}}\right)_{jk} = \sum_{k=1}^{n}\left(C^{\mathrm{T}}\right)_{jk}\left(A^{\mathrm{T}}\right)_{ki} = \left(C^{\mathrm{T}}A^{\mathrm{T}}\right)_{ji}$$

□

Para qualquer matriz complexa $A \in \mathbb{C}^{m,n}$, a *matriz conjugada* de A é designada por $\overline{A} \in \mathbb{C}^{m,n}$ e é formada pelos complexos conjugados dos elementos de A:

$$\left(\overline{A}\right)_{ij} = \overline{A_{ij}};\ i = 1, 2, \ldots, m \text{ e } j = 1, 2, \ldots, n \tag{2.24}$$

A seguinte proposição é consequência das propriedades do operador conjugação em $\mathbb{C}$ e tem demonstração imediata (deixada ao cuidado do leitor):

Proposição 2.4 – Propriedades da conjugação – *Sejam $m, n, p \in \mathbb{N}$. Para quaisquer matrizes $A, B \in \mathbb{C}^{m,n}$, $C \in \mathbb{C}^{n,p}$ e qualquer número complexo α, tem-se:*

i) $\overline{\overline{A}} = A$

ii) $\overline{A \pm B} = \overline{A} \pm \overline{B}$

iii) $\overline{\alpha A} = \overline{\alpha}\,\overline{A}$. *Se $\alpha \in \mathbb{R}$, tem-se $\overline{\alpha A} = \alpha\overline{A}$.*

iv) Para qualquer matriz $A \in \mathbb{C}^{m,n}$, tem-se $A \in \mathbb{R}^{m,n} \Leftrightarrow \overline{A} = A$.

v) $\overline{AC} = \overline{A}\,\overline{C}$

A matriz transposta de $\overline{A}$ é chamada a *transconjugada* de A e designada por A^*.

Definição 2.7 – Transconjugação – *Sejam $m, n \in \mathbb{N}$ e $A \in \mathbb{C}^{m,n}$ uma matriz complexa. A* ***transconjugada*** *de A é a matriz A^* definida por*

$$A^* = \left(\overline{A}\right)^{\mathrm{T}} \tag{2.25}$$

As proposições 2.3 e 2.4 implicam imediatamente as seguintes propriedades para o operador transconjugação, cuja demonstração é óbvia (faça-a, a título de exercício!):

Proposição 2.5 – Propriedades da transconjugação – *Sejam $m, n, p \in \mathbb{N}$. Para quaisquer matrizes $A, B \in \mathbb{C}^{m,n}$, $C \in \mathbb{C}^{n,p}$ e qualquer número complexo α, tem-se:*

i) $(A^*)^* = A$

ii) $(A \pm B)^* = A^* \pm B^*$

iii) $(\alpha A)^* = \overline{\alpha} A^*$. *Se* $\alpha \in \mathbb{R}$, *tem-se* $(\alpha A)^* = \alpha A^*$.

iv) Para qualquer matriz $A \in \mathbb{C}^{m,n}$, *tem-se* $A \in \mathbb{R}^{m,n} \Leftrightarrow A^* = A^{\mathrm{T}}$.

v) $(AC)^* = C^* A^*$

De acordo com o seu comportamento perante os operadores transposição e transconjugação, definem-se vários tipos importantes de matrizes: as matrizes *simétricas*, as matrizes *hemi-simétricas* (ou *anti-simétricas*), as matrizes *hermitianas* (ou *hermíticas*) e as matrizes *hemi-* (ou *anti-*) *hermitianas* (ou *hermíticas*). Seguem-se as definições:

Definição 2.8 – Matriz simétrica – *Seja* $n \in \mathbb{N}$ *e* $A = [\, a_{ij} \,] \in \mathbb{K}^{n,n}$ *uma matriz quadrada de ordem* n *sobre um corpo* $\mathbb{K}$. *A matriz* A *diz-se* ***simétrica*** *sse for invariante por transposição, isto é:*

$$A \text{ é simétrica} \Leftrightarrow A^{\mathrm{T}} = A \tag{2.26}$$

(Observe que a igualdade anterior implica que a matriz A *seja quadrada). Em termos dos elementos* a_{ij} *de* A*, será:*

$$a_{ij} = a_{ji},\ i, j = 1, 2, \ldots, n \tag{2.27}$$

Os escalares a_{ij} e a_{ji} figuram na matriz A em posições simétricas relativamente à diagonal principal respectiva. O seguinte exemplo mostra uma matriz simétrica complexa:

Exemplo 2.18 A seguinte matriz $A \in \mathbb{C}^{3,3}$ é simétrica:

$$A = \begin{bmatrix} 0 & 1-i & 3i \\ 1-i & -2i & 4 \\ 3i & 4 & 1+i \end{bmatrix}$$

Definição 2.9 – Matriz anti-simétrica – *Seja* $n \in \mathbb{N}$ *e* $A = [\, a_{ij} \,] \in \mathbb{K}^{n,n}$ *uma matriz quadrada de ordem* n *sobre um corpo* $\mathbb{K}$. *A matriz* A *diz-se* ***anti-simétrica*** *(ou* ***hemi-simétrica****) sse a matriz transposta de* A *for igual à matriz* $-A$*:*

$$A \text{ é anti-simétrica} \Leftrightarrow A^{\mathrm{T}} = -A \tag{2.28}$$

(Observe que a igualdade anterior implica, uma vez mais, que a matriz A *seja quadrada). Em termos dos elementos* a_{ij} *de* A*, será:*

$$a_{ij} = -a_{ji},\ i, j = 1, 2, \ldots, n \tag{2.29}$$

As igualdades anteriores implicam, em particular para os elementos da diagonal principal, que se tem $a_{ii} = -a_{ii}$. Se o corpo $\mathbb{K}$ tiver característica diferente de 2, isto conduz a que

$a_{ii} = 0$. É o que acontece com os corpos $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}^{(6)}$. O seguinte exemplo mostra uma matriz real hemi-simétrica:

Exemplo 2.19 A seguinte matriz $A \in \mathbb{R}^{3,3}$ é hemi-simétrica:

$$A = \begin{bmatrix} 0 & 1 & 3 \\ -1 & 0 & -4 \\ -3 & 4 & 0 \end{bmatrix}$$

Exemplo 2.20 Os conjuntos $S = \{A: A^{\mathrm{T}} = A\} \subset \mathbb{K}^{n,n}$ das matrizes *simétricas* de ordem n e $H = \{A: A^{\mathrm{T}} = -A\} \subset \mathbb{K}^{n,n}$ das matrizes *hemi-simétricas* da mesma ordem constituem subespaços de $\mathbb{K}^{n,n}$ (isto é, são satisfeitos os requesitos S1-S3 do capítulo 1, facto cuja demonstração deixamos ao cuidado do leitor, como exercício).

Além disso, tem-se, para qualquer matriz $A \in \mathbb{K}^{n,n}$,

$$A = \tfrac{1}{2}\left(A + A^{\mathrm{T}}\right) + \tfrac{1}{2}\left(A - A^{\mathrm{T}}\right)$$

sendo que $\frac{1}{2}(A + A^{\mathrm{T}})$ é uma matriz *simétrica* e $\frac{1}{2}(A - A^{\mathrm{T}})$ é *hemi-simétrica,* como se mostra a seguir:

$$\begin{cases} \left(\tfrac{1}{2}\left(A + A^{\mathrm{T}}\right)\right)^{\mathrm{T}} = \tfrac{1}{2}\left(A + A^{\mathrm{T}}\right)^{\mathrm{T}} = \tfrac{1}{2}\left(A^{\mathrm{T}} + \left(A^{\mathrm{T}}\right)^{\mathrm{T}}\right) = \tfrac{1}{2}(A^{\mathrm{T}} + A) = \tfrac{1}{2}\left(A + A^{\mathrm{T}}\right) \\ \left(\tfrac{1}{2}\left(A - A^{\mathrm{T}}\right)\right)^{\mathrm{T}} = \tfrac{1}{2}\left(A - A^{\mathrm{T}}\right)^{\mathrm{T}} = \tfrac{1}{2}\left(A^{\mathrm{T}} - \left(A^{\mathrm{T}}\right)^{\mathrm{T}}\right) = \tfrac{1}{2}(A^{\mathrm{T}} - A) = -\left(\tfrac{1}{2}\left(A - A^{\mathrm{T}}\right)\right) \end{cases}$$

A igualdade anterior significa que o espaço $\mathbb{K}^{n,n}$ é *soma* dos seus subespaços S e H (ver definição 1.12):

$$\mathbb{K}^{n,n} = S + H$$

Além disso, se $A \in S \cap H$, então $A^{\mathrm{T}} = A = -A$, donde $A = -A$. Conclui-se que, *se o corpo* $\mathbb{K}$ *tiver característica diferente de 2*, será $A = O_n$ e, obviamente, $S \cap H = \{O_n\}$ e o corolário 1.21.1 permite deduzir que $\mathbb{K}^{n,n}$ será *soma directa* de S e H:

$$\mathbb{K}^{n,n} = S \oplus H \tag{2.30}$$

Tem-se ainda, pela proposição 1.22 e atendendo a que $\dim(\mathbb{K}^{n,n}) = n^2$,

$$\begin{cases} \dim H = \frac{n^2-n}{2} = \frac{1}{2}n(n-1) \\ \dim S \ = n^2 - \frac{n^2-n}{2} = \frac{n^2+n}{2} = \frac{1}{2}n(n+1) \end{cases} \tag{2.31}$$

Definição 2.10 – Matriz hermitiana – *Seja* $n \in \mathbb{N}$ *e* $A = [\,a_{ij}\,] \in \mathbb{C}^{n,n}$ *uma matriz complexa quadrada de ordem* n*. A matriz* A *diz-se* ***hermitiana*** *(ou* ***hermítica****) sse for invariante por transconjugação, isto é:*

$$A \text{ é hermitiana } \Leftrightarrow A^* = A \tag{2.32}$$

---

6 Mas não com $\mathbb{Z}_2$, por exemplo. Aqui tem-se $0+0=0$ e também $1+1=0$, o que mostra que $-0=0$ e $-1=1$: Portanto, 0 *não é* o único escalar simétrico de si mesmo.

(Observe que a igualdade anterior implica que a matriz A seja quadrada). Em termos dos elementos de A, será:

$$\overline{a_{ij}} = a_{ji},\ i, j = 1, 2, \ldots, n \tag{2.33}$$

As igualdades anteriores implicam, para os elementos da diagonal principal, que se tem $\overline{a_{ii}} = a_{ii}$, o que é equivalente a $a_{ii} \in \mathbb{R}$. O seguinte exemplo mostra uma matriz hermitiana:

Exemplo 2.21 A seguinte matriz $A \in \mathbb{C}^{3,3}$ é hermitiana:

$$A = \begin{bmatrix} 2 & 1+i & -3i \\ 1-i & -3 & -4 \\ 3i & -4 & 0 \end{bmatrix}$$

Definição 2.11 – Matriz anti-hermitiana – *Seja $n \in \mathbb{N}$ e $A = [\,a_{ij}\,] \in \mathbb{C}^{n,n}$ uma matriz complexa quadrada de ordem n. A matriz A diz-se* ***anti-hermitiana*** *(ou* ***anti-hermítica****)*[7] *sse a sua transconjugada coincidir com a simétrica:*

$$A \text{ é anti-hermitiana } \Leftrightarrow A^* = -A \tag{2.34}$$

(Observe que a igualdade anterior implica que a matriz seja quadrada). Em termos dos elementos de A, será:

$$\overline{a_{ij}} = -a_{ji},\ i, j = 1, 2, \ldots, n \tag{2.35}$$

As igualdades anteriores implicam, para os elementos da diagonal principal, que se tem $\overline{a_{ii}} = -a_{ii}$, o que é equivalente a $a_{ii} \in i\mathbb{R}$ (conjunto dos imaginários puros). A seguir, exemplifica-se uma matriz anti-hermitiana:

Exemplo 2.22 A seguinte matriz $A \in \mathbb{C}^{3,3}$ é anti-hermitiana:

$$A = \begin{bmatrix} -2i & 1+i & -3i \\ -1+i & 3i & 4 \\ -3i & -4 & 0 \end{bmatrix}$$

Exemplo 2.23 $\mathbb{C}^{n,n}$ é, como se viu, um espaço vectorial de dimensão n^2 sobre $\mathbb{C}$ e, portanto, terá dimensão $2n^2$ sobre $\mathbb{R}$ (ver exemplo 1.56).

Sejam ainda $H = \{A\colon A^* = A\} \subset \mathbb{C}^{n,n}$ e $H' = \{A\colon A^* = -A\} \subset \mathbb{C}^{n,n}$ respectivamente os conjuntos das matrizes *hermitianas* e *hemi-hermitianas* de ordem n. Estes conjuntos *não constituem* subespaços de $\mathbb{C}^{n,n}$, como espaço vectorial *complexo*, visto que a condição S2 – fecho em relação ao produto por escalar complexo – não se verifica, atendendo à 3ª alínea da proposição 2.5; no entanto, considerando $\mathbb{C}^{n,n}$ um espaço vectorial *real* e atendendo àquela alínea, os conjuntos H e H' já são subespaços de $\mathbb{C}^{n,n}$ (isto é, são satisfeitos os requesitos S1-S3 do capítulo 1, facto cuja demonstração deixamos ao cuidado do leitor, como exercício). Além disso, tem-se, para qualquer matriz $A \in \mathbb{C}^{n,n}$,

$$A = \tfrac{1}{2}(A + A^*) + \tfrac{1}{2}(A - A^*)$$

sendo que $\frac{1}{2}(A + A^*)$ é uma matriz *hermitiana* e $\frac{1}{2}(A - A^*)$ é *hemi-hermitiana*, o que se prova

---

7 Usam-se também os termos *hemi-hermitiana* ou *hemi-hermítica*.

a seguir:

$$\begin{cases} \left(\frac{1}{2}\left(A+A^*\right)\right)^* = \frac{1}{2}\left(A+A^*\right)^* = \frac{1}{2}\left(A^*+\left(A^*\right)^*\right) = \frac{1}{2}(A^*+A) = \frac{1}{2}\left(A+A^*\right) \\ \left(\frac{1}{2}\left(A-A^*\right)\right)^* = \frac{1}{2}\left(A-A^*\right)^* = \frac{1}{2}\left(A^*-\left(A^*\right)^*\right) = \frac{1}{2}(A^*-A) = -\left(\frac{1}{2}\left(A-A^*\right)\right) \end{cases}$$

A igualdade anterior significa que o espaço vectorial *real* $\mathbb{C}^{n,n}$ é *soma* dos subespaços H e H' (ver definição 1.12):

$$\mathbb{C}^{n,n} = H + H'$$

Além disso, se $A \in H \cap H'$, então $A^* = A = -A$, donde $A = -A$, ou seja, $A = O_n$ e, obviamente, $H \cap H' = \{O_n\}$ e o corolário 1.21.1 permite deduzir que $\mathbb{C}^{n,n}$ é *soma directa* de H e H':

$$\mathbb{C}^{n,n} = H \oplus H' \tag{2.36}$$

Quanto às dimensões de H e H', tem-se (como exercício, encontre uma base de cada subespaço),

$$\text{dim}_{\mathbb{R}} H = \text{dim}_{\mathbb{R}} H' = n + 2\,\frac{n^2-n}{2} = n^2 \tag{2.37}$$

Observe-se que estes resultados estão de acordo com a proposição 1.22:

$$2n^2 = \text{dim}_{\mathbb{R}} \mathbb{C}^{n,n} = \text{dim}_{\mathbb{R}} H + \text{dim}_{\mathbb{R}} H' = n^2 + n^2$$

Para terminar esta secção, faremos referência a um operador matricial, chamado *traço*, que actua sobre matrizes quadradas e que passamos a definir,

Definição 2.12 – Traço de uma matriz – *Dada uma matriz $A \in \mathbb{K}^{n,n}$ quadrada de ordem n, chama-se* ***traço*** *de A à soma dos elementos da diagonal principal de A*

$$\text{tr}(A) = \sum_{i=1}^{n} a_{ii} \in \mathbb{K} \tag{2.38}$$

As principais propriedades do operador traço resumem-se na

Proposição 2.6 – Propriedades do traço – *As principais propriedades do operador* ***traço*** *são:*

i) *Para quaisquer matrizes $A, B \in \mathbb{K}^{n,n}$, tem-se:*

$$\text{tr}(A \pm B) = \text{tr}(A) \pm \text{tr}(B)$$

ii) *Para qualquer escalar $k \in \mathbb{K}$ e qualquer matriz $A \in \mathbb{K}^{n,n}$, tem-se:*

$$\text{tr}(kA) = k\,\text{tr}(A)$$

iii) *Para qualquer matriz* $A \in \mathbb{K}^{n,n}$, *vem:*

$$\operatorname{tr}(A^{\mathrm{T}}) = \operatorname{tr}(A)$$

iv) *Para quaisquer matrizes* $A \in \mathbb{K}^{m,n}$ *e* $B \in \mathbb{K}^{n,m}$ *e mesmo quando for* $AB \neq BA$, *tem-se sempre:*

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

Demonstração:

i) Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$. Então

$$\operatorname{tr}(A \pm B) = \sum_{i=1}^{n} (a_{ii} \pm b_{ii}) = \sum_{i=1}^{n} a_{ii} \pm \sum_{i=1}^{n} b_{ii} = \operatorname{tr}(A) \pm \operatorname{tr}(B)$$

ii) Sendo $A = [a_{ij}]$, obtém-se

$$\operatorname{tr}(kA) = \sum_{i=1}^{n} (ka_{ii}) = k \sum_{i=1}^{n} a_{ii} = k \operatorname{tr}(A)$$

iii) Basta observar que A e A^{T} têm diagonais principais iguais.

iv) Sejam $A = [a_{ij}]$ e $B = [b_{jk}]$. O resultado pretendido decorre das igualdades sucessivas

$$\operatorname{tr}(AB) = \sum_{i=1}^{m} (AB)_{ii} = \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} b_{ki} = \sum_{k=1}^{n} \sum_{i=1}^{m} b_{ki} a_{ik} = \sum_{k=1}^{n} (BA)_{kk} = \operatorname{tr}(BA) \qquad \square$$

Na secção 2.16, mostramos o uso do software MATHEMATICA© para levar a cabo as operações matriciais aqui estudadas.

2.6 Submatrizes. Matrizes de blocos. Operações por blocos

Considere-se uma matriz $A \in \mathbb{K}^{m,n}$ e sejam $p \in [1, m]$ e $q \in [1, n]$ dois inteiros. Sejam

$$\begin{aligned} I &= \{i_1, i_2, \ldots, i_p\} \subset [1, m] \\ J &= \{j_1, j_2, \ldots, j_q\} \subset [1, n] \end{aligned} \tag{2.39}$$

duas partes não vazias de $[1, m]$ e de $[1, n]$ respectivamente (supomos $i_1 < i_2 < \cdots < i_p$ e $j_1 < j_2 < \cdots < j_q$). A matriz de tipo $p \times q$ sobre $\mathbb{K}$ obtida de A tomando apenas os elementos das linhas com índices pertencentes a I e das colunas com índices pertencentes a J é chamada uma *submatriz* extraída de A de tipo $p \times q$ e será designada por uma das notações

$$A[I; J] = A[\{i_1, i_2, \ldots, i_p\};\{j_1, j_2, \ldots, j_q\}] \in \mathbb{K}^{p,q} \tag{2.40}$$

ou mais abreviadamente por

$$A[i_1, i_2, \ldots, i_p; j_1, j_2, \ldots, j_q] \in \mathbb{K}^{p,q} \tag{2.41}$$

Nalgumas situações (como acontecerá no capítulo 4) é útil indicar os índices das filas *eliminadas*, ao invés de indicarmos as filas *seleccionadas*, como começámos por fazer. Neste sentido, se fizermos

$$\begin{aligned} I' &= [1,m] \setminus I = \{r_1, r_2, \ldots, r_{m-p}\} \\ J' &= [1,n] \setminus J = \{s_1, s_2, \ldots, s_{n-q}\} \end{aligned}$$

designaremos a submatriz extraída de A indicando as filas *eliminadas* por

$$A(I;J) = A(i_1, i_2, \ldots, i_p; j_1, j_2, \ldots, j_q) = A[I';J'] \in \mathbb{K}^{m-p,n-q} \tag{2.42}$$

A convenção que faremos é, portanto, a de que:

- Para indicarmos as filas *seleccionadas*, usamos parêntesis *rectos*.
- Para indicarmos as filas *eliminadas*, usamos parêntesis *curvos*.

A sintaxe anterior pode aplicar-se *separadamente* às linhas e às colunas de A e, sendo I e J partes não vazias de $[1,m]$ e $[1,n]$ respectivamente, temos as seguintes quatro alternativas de notação:

$A[I;J] = A[i_1, \ldots, i_p; j_1, \ldots, j_q] \in \mathbb{K}^{p,q}$	Indicação de linhas e colunas seleccionadas.
$A(I;J) = A(i_1, \ldots, i_p; j_1, \ldots, j_q) \in \mathbb{K}^{m-p,n-q}$	Indicação de linhas e colunas eliminadas.
$A[I;J) = [i_1, \ldots, i_p; j_1, \ldots, j_q) \in \mathbb{K}^{p,n-q}$	Indicação de linhas seleccionadas e colunas eliminadas.
$A(I;J] = A(i_1, \ldots, i_p; j_1, \ldots, j_q] \in \mathbb{K}^{m-p,q}$	Indicação de linhas eliminadas e colunas seleccionadas.

É claro que, para $p = q = 1$, se terá $A[i;j] = [a_{ij}]$. O número de sequências nas condições (2.39) é de $\binom{m}{p}$ e $\binom{n}{q}$, respectivamente. Portanto o número de submatrizes de tipo $p \times q$ existentes numa matriz de tipo $m \times n$ é de $\binom{m}{p} \times \binom{n}{q}$.

Se os índices em (2.39) forem *consecutivos* (isto é, I e J são *subintervalos* de $[1,m]$ e de $[1,n]$ respectivamente), as submatrizes respectivas chamam-se *blocos* de tipo $p \times q$. Portanto, um bloco de A é uma submatriz formada por filas *consecutivas* de A (mas não necessariamente obtida eliminando filas consecutivas); o número de blocos de tipo $p \times q$ existente numa matriz $m \times n$ é de $(m-p+1) \times (n-q+1)$.

Exemplo 2.24 Sendo

$$A = \begin{bmatrix} 2 & -3 & -4 & 0 \\ -1 & 3 & -2 & 4 \\ -5 & 1 & 5 & -3 \end{bmatrix} \in \mathbb{R}^{3,4}$$

tem-se, por exemplo,

$$\begin{aligned} A[1,3;2,4] = A(2;1,3) = A[1,3;1,3) = A(2;2,4] &= \begin{bmatrix} -3 & 0 \\ 1 & -3 \end{bmatrix} \in \mathbb{R}^{2,2} \\ A[2,3;1,2,3] = A(1;4) = A[2,3;4) = A(1;1,2,3] &= \begin{bmatrix} -1 & 3 & -2 \\ -5 & 1 & 5 \end{bmatrix} \in \mathbb{R}^{2,3} \end{aligned}$$

A primeira submatriz (do tipo 2×2) não é um bloco de A, pois as filas lá representadas não são consecutivas; a 2ª destas submatrizes já constitui um bloco de A do tipo 2×3, constituído pelas linhas consecutivas 2 e 3 e pelas colunas consecutivas $1, 2$ e 3.

Sendo $1 \leq p \leq m$ e $1 \leq q \leq n$, é possível dividir uma matriz de tipo $m \times n$ em $p \times q$ blocos, mediante a construção de uma *fragmentação* da matriz: uma fragmentação de A obtém-se construindo *partições* dos conjuntos dos índices $\{1, 2, \ldots, m\}$ das linhas e dos índices $\{1, 2, \ldots, n\}$ das colunas em p e q subconjuntos respectivamente, através de sequências crescentes de índices de linhas

$$2 \leq i_1 < i_2 < \cdots < i_{p-1} \leq m$$

e de índices de colunas

$$2 \leq j_1 < j_2 < \cdots < j_{q-1} \leq n$$

$$\begin{cases} \{1, 2, \ldots, m\} = \{1, \ldots, i_1 - 1\} \cup \{i_1, \ldots, i_2 - 1\} \cup \cdots \cup \{i_{p-1}, \ldots, m\} \\ \{1, 2, \ldots, n\} \ = \{1, \ldots, j_1 - 1\} \cup \{j_1, \ldots, j_2 - 1\} \cup \cdots \cup \{j_{q-1}, \ldots, n\} \end{cases} \tag{2.43}$$

As partições acima determinam $p \times q$ blocos $(A_{rs})_{\substack{1 \leq r \leq p \\ 1 \leq s \leq q}}$ e a matriz A de blocos resultante escreve-se na forma

$$A = \left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1q} \\ \hline A_{21} & A_{22} & \cdots & A_{2q} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{p1} & A_{p2} & \cdots & A_{pq} \end{array}\right] \tag{2.44}$$

Assim, a matriz A é do tipo $m \times n$, como matriz de escalares, mas é do tipo $p \times q$, como matriz de blocos.

Exemplo 2.25 A seguinte matriz real está fragmentada em 2×3 blocos

$$\left[\begin{array}{cc|c|cc} 2 & -1 & 0 & -3 & 1 \\ 1 & 3 & 2 & 4 & 0 \\ \hline 3 & -2 & 1 & 5 & -1 \\ 0 & -1 & 2 & 0 & 3 \\ -1 & 2 & -3 & -4 & 2 \end{array}\right] = \left[\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \end{array}\right]$$

em que

$$\begin{cases} A_{11} = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}, \ A_{12} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \ A_{13} = \begin{bmatrix} -3 & 1 \\ 4 & 0 \end{bmatrix} \\ A_{21} = \begin{bmatrix} 3 & -2 \\ 0 & -1 \\ -1 & 2 \end{bmatrix}, \ A_{22} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}, \ A_{23} = \begin{bmatrix} 5 & -1 \\ 0 & 3 \\ -4 & 2 \end{bmatrix} \end{cases}$$

Neste caso, a fragmentação apresentada corresponde às partições (2.43) seguintes:

$$\begin{cases} \{1,2,3,4,5\} = \{1,2\} \cup \{3,4,5\} & \text{, para as linhas} \\ \{1,2,3,4,5\} = \{1,2\} \cup \{3\} \cup \{4,5\}, & \text{para as colunas} \end{cases}$$

A matriz é quadrada de 5ª ordem, como matriz real e é do tipo 2×3, como matriz de blocos.

Reciprocamente, sejam $p, q \in \mathbb{N}$ e $(A_{rs})_{\substack{1\le r\le p\\1\le s\le q}}$ uma família de $p \times q$ matrizes sobre um corpo $\mathbb{K}$ tais que, para cada $1 \le r \le p$, todas as matrizes $(A_{rs})_{1\le s\le q}$ têm o mesmo número m_r de linhas e, para cada $1 \le s \le q$, todas as matrizes $(A_{rs})_{1\le r\le p}$ têm o mesmo número n_s de colunas. Pode, então, construir-se uma matriz de blocos do tipo $p \times q$ que será designada por $[A_{rs}]_{\substack{1\le r\le p\\1\le s\le q}}$ e que, como matriz de escalares, será do tipo $m \times n$, onde $m = \sum\limits_{r=1}^{p} m_r$ e $n = \sum\limits_{s=1}^{q} n_s$.

As operações matriciais podem ser realizadas *por blocos*. Como regra geral, podemos dizer que se aplicam às matrizes de blocos os mesmos procedimentos válidos para as matrizes de escalares. Tudo se passa como se as operações matriciais que atrás definimos fossem, afinal, aplicadas a blocos do tipo 1×1.

Vejamos, uma a uma, as operações matriciais:

- *Adição por blocos*: sejam $p, q \in \mathbb{N}$ e $(A_{rs})_{\substack{1\le r\le p\\1\le s\le q}}$ e $(B_{rs})_{\substack{1\le r\le p\\1\le s\le q}}$ duas famílias de matrizes sobre um corpo $\mathbb{K}$ tais que, para cada $1 \le r \le p$, todas as matrizes $(A_{rs})_{1\le s\le q}$ e $(B_{rs})_{1\le s\le q}$ têm o mesmo número m_r de linhas e, para cada $1 \le s \le q$, todas as matrizes $(A_{rs})_{1\le r\le p}$ e $(B_{rs})_{1\le r\le p}$ têm o mesmo número n_s de colunas. A soma (respectivamente, a diferença) das matrizes de blocos $A = [A_{rs}]_{\substack{1\le r\le p\\1\le s\le q}}$ e $B = [B_{rs}]_{\substack{1\le r\le p\\1\le s\le q}}$ é a nova matriz de blocos $A \pm B$ que se obtém somando (ou subtraindo) os blocos homólogos de cada uma delas:

$$\left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1q} \\ \hline A_{21} & A_{22} & \cdots & A_{2q} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{p1} & A_{p2} & \cdots & A_{pq} \end{array}\right] \pm \left[\begin{array}{c|c|c|c} B_{11} & B_{12} & \cdots & B_{1q} \\ \hline B_{21} & B_{22} & \cdots & B_{2q} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline B_{p1} & B_{p2} & \cdots & B_{pq} \end{array}\right] = \left[\begin{array}{c|c|c|c} A_{11}\pm B_{11} & A_{12}\pm B_{12} & \cdots & A_{1q}\pm B_{1q} \\ \hline A_{21}\pm B_{21} & A_{22}\pm B_{22} & \cdots & A_{2q}\pm B_{2q} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{p1}\pm B_{p1} & A_{p2}\pm B_{p2} & \cdots & A_{pq}\pm B_{pq} \end{array}\right] \tag{2.45}$$

- *Multiplicação de escalar por matriz de blocos*: sejam $p, q \in \mathbb{N}$ e $(A_{rs})_{\substack{1\le r\le p\\1\le s\le q}}$ uma família de matrizes sobre um corpo $\mathbb{K}$ tal que, para cada $1 \le r \le p$, todas as matrizes $(A_{rs})_{1\le s\le q}$ têm o mesmo número m_r de linhas e, para cada $1 \le s \le q$, todas as matrizes $(A_{rs})_{1\le r\le p}$ têm o mesmo número n_s de colunas. Sendo $\alpha \in \mathbb{K}$, o produto de α pela matriz A de blocos $A = [A_{rs}]_{\substack{1\le r\le p\\1\le s\le q}}$, é a matriz obtida multiplicando por α cada um dos blocos A_{rs}

$$\alpha \left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1q} \\ \hline A_{21} & A_{22} & \cdots & A_{2q} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{p1} & A_{p2} & \cdots & A_{pq} \end{array}\right] = \left[\begin{array}{c|c|c|c} \alpha A_{11} & \alpha A_{12} & \cdots & \alpha A_{1q} \\ \hline \alpha A_{21} & \alpha A_{22} & \cdots & \alpha A_{2q} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \alpha A_{p1} & \alpha A_{p2} & \cdots & \alpha A_{pq} \end{array}\right] \tag{2.46}$$

- *Transposição de matrizes por blocos*: sejam $p, q \in \mathbb{N}$ e $(A_{rs})_{\substack{1\le r\le p\\ 1\le s\le q}}$ uma família de matrizes sobre um corpo $\mathbb{K}$ tal que, para cada $1 \le r \le p$, todas as matrizes $(A_{rs})_{1\le s\le q}$ têm o mesmo número m_r de linhas e, para cada $1 \le s \le q$, todas as matrizes $(A_{rs})_{1\le r\le p}$ têm o mesmo número n_s de colunas. A transposta da matriz $A = [A_{rs}]_{\substack{1\le r\le p\\ 1\le s\le q}}$ de blocos é a matriz que se obtém transpondo cada bloco A_{rs} e ainda transpondo A como se de uma matriz de escalares se tratasse

$$\left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1q} \\ \hline A_{21} & A_{22} & \cdots & A_{2q} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{p1} & A_{p2} & \cdots & A_{pq} \end{array}\right]^{\mathrm{T}} = \left[\begin{array}{c|c|c|c} A_{11}^{\mathrm{T}} & A_{21}^{\mathrm{T}} & \cdots & A_{p1}^{\mathrm{T}} \\ \hline A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}} & \cdots & A_{p2}^{\mathrm{T}} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{1q}^{\mathrm{T}} & A_{2q}^{\mathrm{T}} & \cdots & A_{pq}^{\mathrm{T}} \end{array}\right] \tag{2.47}$$

- *Produto de matrizes por blocos*: sejam $p, q, u \in \mathbb{N}$ e $(A_{rs})_{\substack{1\le r\le p\\ 1\le s\le q}}$ e $(B_{st})_{\substack{1\le s\le q\\ 1\le t\le u}}$ duas famílias de matrizes sobre um corpo $\mathbb{K}$ tais que, para cada $1 \le r \le p$, todas as matrizes $(A_{rs})_{1\le s\le q}$ têm o mesmo número m_r de linhas e, para cada $1 \le s \le q$, todas as matrizes $(A_{rs})_{1\le r\le p}$ têm o mesmo número n_s de colunas. Suponha-se ainda que, para cada $1 \le s \le q$, todas as matrizes $(B_{st})_{1\le t\le u}$ têm n_s linhas e que, para cada $1 \le t \le u$, todas as matrizes $(B_{st})_{1\le s\le q}$ têm o mesmo número x_t de colunas. Portanto, as matrizes A_{rs} são do tipo $m_r \times n_s$ e as matrizes B_{st} são do tipo $n_s \times x_t$. O produto da matriz de blocos $A = [A_{rs}]_{\substack{1\le r\le p\\ 1\le s\le q}}$ pela matriz de blocos $B = [B_{st}]_{\substack{1\le s\le q\\ 1\le t\le u}}$ é a nova matriz de $p \times u$ blocos $C = [C_{rt}]_{\substack{1\le r\le p\\ 1\le t\le u}}$, em que

$$C_{rt} = \sum_{s=1}^{q} A_{rs}B_{st} \text{ e onde } C_{rt} \text{ é matriz } m_r \times x_t \tag{2.48}$$

As igualdades matriciais anteriores podem escrever-se na forma

$$\left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1q} \\ \hline A_{21} & A_{22} & \cdots & A_{2q} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{p1} & A_{p2} & \cdots & A_{pq} \end{array}\right] \left[\begin{array}{c|c|c|c} B_{11} & B_{12} & \cdots & B_{1u} \\ \hline B_{21} & B_{22} & \cdots & B_{2u} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline B_{q1} & B_{q2} & \cdots & B_{qu} \end{array}\right] = \left[\begin{array}{c|c|c|c} \sum\limits_{s=1}^{q} A_{1s}B_{s1} & \sum\limits_{s=1}^{q} A_{1s}B_{s2} & \cdots & \sum\limits_{s=1}^{q} A_{1s}B_{su} \\ \hline \sum\limits_{s=1}^{q} A_{2s}B_{s1} & \sum\limits_{s=1}^{q} A_{2s}B_{s2} & \cdots & \sum\limits_{s=1}^{q} A_{2s}B_{su} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \sum\limits_{s=1}^{q} A_{ps}B_{s1} & \sum\limits_{s=1}^{q} A_{ps}B_{s2} & \cdots & \sum\limits_{s=1}^{q} A_{ps}B_{su} \end{array}\right] \tag{2.49}$$

Na secção 2.16, refere-se a utilização do software MATHEMATICA© para tratar matrizes por blocos, através do *package* **`ALGA`Matrizes``**.

2.7 Característica de uma matriz

Nesta secção, vamos ver que, em qualquer matriz de tipo $m \times n$ (eventualmente rectangular) sobre um corpo $\mathbb{K}$, o subespaço de $\mathbb{K}^n$ gerado pelas linhas e o subespaço de $\mathbb{K}^m$ gerado pelas colunas têm a mesma dimensão. Essa dimensão comum é um atributo importante da matriz a que chamaremos a respectiva *característica*.

Proposição 2.7 *Seja* $A = [\,a_{ij}\,] \in \mathbb{K}^{m,n}$ *uma matriz do tipo* $m \times n$ *sobre um corpo* $\mathbb{K}$*, com linhas* $(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m)$ *e colunas* $(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n)$*. O subespaço* $L(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m) \subset \mathbb{K}^n$ *e o subespaço* $L(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n) \subset \mathbb{K}^m$ *têm dimensões iguais.*

Demonstração:

Seja

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

e sejam

$$\vec{r}_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \ldots, \vec{r}_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$$

as m linhas de A. Suponha-se que $\dim L(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m) = r$ e que a lista $s = (\vec{s}_i)_{1 \le i \le r}$ formada pelos r vectores de $\mathbb{K}^n$ seguintes formam uma *base* de $L(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m)$:

$$\vec{s}_1 = (s_{11}, s_{12}, \ldots, s_{1n}), \ldots, \vec{s}_2 = (s_{21}, s_{22}, \ldots, s_{2n}), \ldots, \vec{s}_r = (s_{r1}, s_{r2}, \ldots, s_{rn})$$

Então, cada uma das linhas $(\vec{r}_i)_{1 \le i \le m}$ de A é combinação linear (única) dos $(\vec{s}_k)_{1 \le k \le r}$

$$\begin{cases} \vec{r}_1 &= u_{11}\vec{s}_1 + u_{12}\vec{s}_2 + \cdots + u_{1r}\vec{s}_r = \sum\limits_{k=1}^{r} u_{1k}\vec{s}_k \\ \vec{r}_2 &= u_{21}\vec{s}_1 + u_{22}\vec{s}_2 + \cdots + u_{2r}\vec{s}_r = \sum\limits_{k=1}^{r} u_{2k}\vec{s}_k \\ \quad\cdots & \\ \vec{r}_m &= u_{m1}\vec{s}_1 + u_{m2}\vec{s}_2 + \cdots + u_{mr}\vec{s}_r = \sum\limits_{k=1}^{r} u_{mk}\vec{s}_k \end{cases}$$

onde os u_{ik} são escalares únicos. As m igualdades vectoriais em $\mathbb{K}^n$ anteriores equivalem às $m \times n$ igualdades escalares, válidas para cada $j = 1, 2, \ldots, n$,

$$\begin{cases} a_{1j} &= u_{11}s_{1j} + u_{12}s_{2j} + \cdots + u_{1r}s_{rj} = \sum\limits_{k=1}^{r} u_{1k}s_{kj} \\ a_{2j} &= u_{21}s_{1j} + u_{22}s_{2j} + \cdots + u_{2r}s_{rj} = \sum\limits_{k=1}^{r} u_{2k}s_{kj} \\ \quad\cdots & \\ a_{mj} &= u_{m1}s_{1j} + u_{m2}s_{2j} + \cdots + u_{mr}s_{rj} = \sum\limits_{k=1}^{r} u_{mk}s_{kj} \end{cases}$$

Mas as igualdades escalares anteriores equivalem também às seguintes n igualdades vectoriais (no espaço $\mathbb{K}^m$), para $j = 1, 2, \ldots, n$,

$$\vec{c}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = s_{1j} \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{m1} \end{bmatrix} + s_{2j} \begin{bmatrix} u_{12} \\ u_{22} \\ \vdots \\ u_{m2} \end{bmatrix} + \cdots + s_{rj} \begin{bmatrix} u_{1r} \\ u_{2r} \\ \vdots \\ u_{mr} \end{bmatrix}$$

as quais mostram que as colunas $\vec{c}_j$ de A são combinação linear dos r vectores de $\mathbb{K}^m$

$$\vec{u}_1 = \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{m1} \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} u_{12} \\ u_{22} \\ \vdots \\ u_{m2} \end{bmatrix}, \ldots, \vec{u}_r = \begin{bmatrix} u_{1r} \\ u_{2r} \\ \vdots \\ u_{mr} \end{bmatrix}$$

Portanto, pode concluir-se que

$$L(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n) \subset L(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r)$$

A inclusão anterior implica

$$\mathrm{dim}L(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n) \leq \mathrm{dim}L(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r) \leq r = \mathrm{dim}L(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m)$$

ficando, então

$$\mathrm{dim}L(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n) \leq \mathrm{dim}L(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m)$$

Repetindo o argumento anterior em relação a A^{T} e atendendo a que as linhas de A^{T} são as colunas de A e vice-versa, conclui-se que será também

$$\mathrm{dim}L(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m) \leq \mathrm{dim}L(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n)$$

As duas últimas desigualdades equivalem à igualdade pretendida

$$\mathrm{dim}L(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m) = \mathrm{dim}L(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n) \qquad \square$$

A proposição anterior permite-nos pôr a seguinte

Definição 2.13 – Característica de uma matriz – *Sejam $m, n \in \mathbb{N}$ e $A \in \mathbb{K}^{m,n}$ uma matriz de tipo $m \times n$ sobre um corpo $\mathbb{K}$. Chama-se* ***característica*** *de A à dimensão do subespaço gerado pelas linhas (ou colunas) de A a qual será designada por* $\mathrm{c}(A)$.

Da definição anterior resultam imediatamente algumas propriedades que se resumem na

Proposição 2.8 – Propriedades da característica – *Seja $A \in \mathbb{K}^{m,n}$ uma matriz de tipo $m \times n$ sobre um corpo $\mathbb{K}$. Então:*

i) $\mathrm{c}(O_{m,n}) = 0$ e $\mathrm{c}(I_n) = n$

ii) $\mathrm{c}(A^{\mathrm{T}}) = \mathrm{c}(A)$.

iii) Se $\alpha \in \mathbb{K} \setminus \{0\}$, então $\mathrm{c}(\alpha A) = \mathrm{c}(A)$.

iv) $\mathrm{c}(A) \leq \min\{m, n\}$.

v) Se $B \in \mathbb{K}^{n,p}$ for uma matriz de tipo $n \times p$ sobre $\mathbb{K}$, tem-se

$$\mathrm{c}(AB) \leq \min\{\mathrm{c}(A), \mathrm{c}(B)\}$$

vi) *A característica de A é invariante, por operações elementares executadas sobre as filas de A.*

vii) *Se A é uma matriz* ***escalonada***, *a sua característica é igual ao número de linhas não nulas nela existentes.*

Demonstração:

i) O subespaço gerado pela linhas é $\{(0, 0, \dots, 0)\} = \{\vec{o}\}$. Portanto,

$$\mathrm{c}(O_{m,n}) = \dim\{\vec{o}\} = 0$$

Por outro lado, as linhas de I_n constituem a *base canónica* de $\mathbb{K}^n$ e o subespaço por elas gerado é o próprio $\mathbb{K}^n$: portanto,

$$\mathrm{c}(I_n) = \dim(\mathbb{K}^n) = n$$

ii) É consequência imediata da proposição 2.7.

iii) Basta observar que as linhas de αA se obtêm das linhas de A por operações elementares do tipo 2 (com $\alpha \neq 0$); deste modo, a proposição 1.18 garante que os subespaços de $\mathbb{K}^n$ gerados pelas linhas de αA e de A são iguais, o mesmo acontecendo, portanto, às características respectivas (que são as dimensões daqueles subespaços).

iv) Como A tem m linhas $(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_m)$, será $\mathrm{c}(A) = \dim L(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_m) \leq m$. Porém, tem-se $L(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_m) \subset \mathbb{K}^n$ e então fica também $\mathrm{c}(A) = \dim L(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_m) \leq n$. Daqui resulta $\mathrm{c}(A) \leq \min\{m, n\}$, como se pretendia.

v) Fragmentando A nas suas n colunas $[\vec{c}_j]_{1 \leq j \leq n}$ e B nos seus $n \times p$ elementos $[b_{jk}]_{\substack{1 \leq j \leq n \\ 1 \leq k \leq p}}$ e calculando o produto AB por blocos, resulta a matriz $U = AB = [\vec{u}_k]_{1 \leq k \leq p}$ fragmentada também nas suas p colunas $\vec{u}_k$ onde

$$\vec{u}_k = \sum_{j=1}^{n} b_{jk}\vec{c}_j\,;\ k = 1, \dots, p$$

As p igualdades vectoriais anteriores (em $\mathbb{K}^m$) significam que as colunas de AB são combinação linear das colunas de A e que, portanto,

$$L(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p) \subset L(\vec{c}_1, \vec{c}_2, \dots, \vec{c}_n) \Rightarrow \mathrm{c}(AB) \leq \mathrm{c}(A)$$

Raciocinando em termos das transpostas e recorrendo a *ii)*, à proposição 2.3.*iv* e à desigualdade anterior (aplicada a $B^{\mathrm{T}}A^{\mathrm{T}}$), tem-se

$$\mathrm{c}(AB) = \mathrm{c}\big((AB)^{\mathrm{T}}\big) = \mathrm{c}\big(B^{\mathrm{T}}A^{\mathrm{T}}\big) \leq \mathrm{c}\big(B^{\mathrm{T}}\big) = \mathrm{c}(B)$$

Portanto, é simultaneamente $\mathrm{c}(AB) \leq \mathrm{c}(A)$ e $\mathrm{c}(AB) \leq \mathrm{c}(B)$, o que implica a desigualdade desejada: $\mathrm{c}(AB) \leq \min\{\mathrm{c}(A), \mathrm{c}(B)\}$.

vi) Observe-se que, pela proposição 1.18, uma operação elementar executada sobre as linhas de uma matriz $A \in \mathbb{K}^{m,n}$ não altera o subespaço vectorial $V \subset \mathbb{K}^n$ gerado por elas, nem a dependência ou independência linear das referidas linhas; portanto, a dimensão de V, que é a característica de A, também não se altera. Obviamente que idêntico raciocínio vale também para as colunas de A. Portanto:

A característica de $A \in \mathbb{K}^{m,n}$ é invariante por operações elementares sobre as filas de A

vii) Se A é escalonada, as linhas nulas que eventualmente ocorram (na parte inferior de A) podem ser retiradas da lista das linhas, que isso não alterará o subespaço $V \subset \mathbb{K}^n$ gerado pela referida lista (as linhas nulas são combinação linear das anteriores – lema 1.13.2) e o corolário 1.11.1 garante que as linhas não nulas restantes são linearmente independentes: conclui-se daqui que a característica de uma matriz escalonada é igual ao número de linhas não nulas nela existentes:

$$A \in \mathbb{K}^{m,n} \text{ é escalonada} \Rightarrow \mathrm{c}(A) \text{ é o número de linhas de } A \text{ não nulas.}$$ □

2.8 Algoritmo de condensação vertical

O problema de calcular a característica de uma matriz A dada é muito importante em álgebra linear e ele pode ser resolvido através de um algoritmo, baseado nas operações elementares sobre listas de vectores (ver proposição 1.18) os quais, no caso vertente, são as linhas (ou as colunas) da matriz A e conhecido por algoritmo de *condensação vertical* (se as operações incidem sobre linhas) ou algoritmo de *condensação horizontal* (se as operações incidem sobre as colunas de A). O algoritmo tem, como veremos, as mais variadas aplicações, para além do cálculo da característica de uma matriz: resolução de sistemas de equações lineares, inversão de matrizes, cálculo de determinantes, etc; ele resolve também o problema vectorial típico de calcular uma *base* e a *dimensão* do subespaço gerado por uma lista de vectores dada, determinando ainda se esta é ou não *linearmente independente*.

O algoritmo de condensação vertical baseia-se nas alíneas *vi)* e *vii)* da proposição 2.8 e visa, por meio de uma sequência adequada de operações elementares sobre as linhas, transformar qualquer matriz $A \in \mathbb{K}^{m,n}$ numa matriz E escalonada, sendo que este processo não altera a característica de A

$$\mathrm{c}(A) = \mathrm{c}(E) = \text{número de linhas não nulas de } E$$

O referido algoritmo de condensação vertical pode ser informalmente delineado nos termos seguintes:

1. *Procurar um elemento não nulo na matriz A, fazendo esta busca por colunas da esquerda para a direita e de cima para baixo. Se não foi encontrado nenhum elemento diferente de zero, tem-se $A = O_{m,n}$ e o algoritmo terminou com característica igual a 0. Se foi encontrado um elemento não nulo, ele pertencerá a uma certa linha i e a uma coluna k, ou seja, será $a_{ik} \neq 0$ e a característica será, pelo menos, 1.*

2. *Trocar a linha i onde ocorreu o elemento a_{ik} anteriormente referido com a primeira linha (operação elementar do tipo 1).*

3. *Observe-se que, neste ponto, será $a_{1k} \neq 0$; multiplique-se a primeira linha por a_{1k}^{-1}, obtendo-se $a_{1k} = 1$ (operação elementar do tipo 2); este elemento é chamado o **1º redutor** ou **1º pivot** do algoritmo.*

4. *Para todas as linhas L_i a partir da segunda (inclusivé), substituam-se as linhas pela diferença entre cada uma delas e o produto de a_{ik} pela primeira linha (operação do*

tipo 3): isto é, L_i será substituída por $L_i - a_{ik}L_1$. Atendendo a que $a_{1k} = 1$, isto fará com que o elemento a_{ik} seja substituído por $a_{ik} - a_{ik} \cdot a_{1k} = a_{ik} - a_{ik} \cdot 1 = 0$.

5. *Repetir os todos os passos anteriores, para o bloco $A(1; 1, 2, \ldots, k) \in \mathbb{K}^{m-1,n-k}$ obtido de A eliminando a primeira linha e as primeiras k colunas, se este não for vazio; caso contrário, o algoritmo terminou.*

Apresentamos, a seguir, o algoritmo de condensação vertical de um modo mais formal, para transformar uma matriz $A = [a_{ij}] \in \mathbb{K}^{m,n}$ noutra matriz E escalonada e, em simultâneo, calcular a característica r de A: os dados de entrada são a matriz $A = [a_{ij}]$. As saídas são a característica r de A e a matriz escalonada B. As convenções de notação feitas na versão em *pseudo-código* abaixo mostrada são:

=.................................*Afectação entre escalares (ou vectores de $\mathbb{K}^n$).*

`aik`.............................. *Escalar: elemento da linha `i` e coluna `k` da matriz A.*

`Linhai`...................... *Vector de $\mathbb{K}^n$ constituído pela linha `i` da matriz A.*

`v` $\leftrightarrow$ `w`......................... *Troca dos vectores `v` e `w` de $\mathbb{K}^n$ entre si.*

$\wedge$ ....*Conjunção.*

`% Texto`....................*Comentário.*

Segue-se o algoritmo:

```
m=#Linhas(a); n=#Colunas(a);
r=0; k=0;
b=a;
Enquanto r < m ∧ k < n
    % Busca do Pivot:
    Fazer
        k=k+1;
        i=r;
        Fazer
          i=i+1
        Enquanto bik = 0 ∧ i < m
    Enquanto bik = 0 ∧ k < n;
    Se bik ≠ 0 então                    % bik é pivot?
        % Redução da coluna k de b:
        r=r+1;
        Se i ≠ r então Linhai ↔ Linhar;      % Op. de tipo 1
        Linhar=Linhar/brk;                   % Op. de tipo 2
        Para i=r+1 até m
          Linhai=Linhai-bik*Linhar           % Op. de tipo 3
        FimPara
    FimSe
FimEnquanto;
Saída: {r, b}
```

Na secção 2.16, apresenta-se uma implementação do algoritmo anterior utilizando o software MATHEMATICA©: trata-se da função **`Caracteristica`** definida no *package* chamado **`` ALGA`Matrizes` ``**.

Quando se faz manualmente a condensação de uma matriz, raramente se segue rigorosamente o algoritmo anterior. Algumas modificações do algoritmo incluem:

- *No passo 3 do algoritmo esboçado na pág. 97 deste capítulo, não se divide a primeira linha pelo pivot (forçando o pivot a ser igual a* 1*). Deste modo, se a referida linha for constituída por inteiros, não aparecerão fracções nos outros elementos da linha, o que complicará os cálculos subsequentes.*

- *No passo 4, em vez de fazermos operações do tipo 3 para anular os elementos por baixo do pivot, fazemos operações do tipo 4: para* $i = 2, \ldots, m$*, substitui-se* L_i *por* $a_{1k}L_i - a_{ik}L_1$ *o que conduz a substituir* a_{ik} *por* $a_{1k}a_{ik} - a_{ik}a_{1k} = 0$*, como se pretendia.*

- *No passo 1, escolher, sempre que possível, um elemento igual a* 1 *(e não apenas diferente de zero), pois isso possibilitará a utilização de operações do tipo 3 (e não do tipo 4) para reduzir a zero os elementos debaixo do pivot.*

- *Observe-se que, se o objectivo é apenas o cálculo da característica de* A*, podemos fazer, ao longo do processo, operações sobre linhas e também sobre colunas.*

- *Ao programar o algoritmo e com vista a minimizar os erros de arredondamento típicos do uso da representação de vírgula flutuante para implementar os números reais, usa-se para pivot o elemento de maior valor absoluto da primeira coluna da matriz na qual este elemento se não anule.*

Vamos agora apresentar alguns exemplos de utilização do algoritmo de condensação vertical:

Exemplo 2.26 Calcule-se a característica da seguinte matriz $A \in \mathbb{R}^{5,4}$, usando condensação:

$$A = \begin{bmatrix} 2 & -1 & 3 & 0 \\ -3 & 1 & 2 & -1 \\ 5 & -2 & 3 & -2 \\ -2 & 0 & 8 & 1 \\ 1 & -1 & 10 & -4 \end{bmatrix}$$

Tem-se, executando as operações elementares indicadas e indicando dentro de rectângulos os pivots utilizados,

$$\begin{bmatrix} 2 & -1 & 3 & 0 \\ -3 & 1 & 2 & -1 \\ 5 & -2 & 3 & -2 \\ -2 & 0 & 8 & 1 \\ 1 & -1 & 10 & -4 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_5} \begin{bmatrix} \boxed{1} & -1 & 10 & -4 \\ -3 & 1 & 2 & -1 \\ 5 & -2 & 3 & -2 \\ -2 & 0 & 8 & 1 \\ 2 & -1 & 3 & 0 \end{bmatrix} \xrightarrow[\substack{L_4 \to L_4 + 2L_1 \\ L_5 \to L_5 - 2L_1}]{\substack{L_2 \to L_2 + 3L_1 \\ L_3 \to L_3 - 5L_1}} \begin{bmatrix} 1 & -1 & 10 & -4 \\ 0 & -2 & 32 & -13 \\ 0 & 3 & -47 & 18 \\ 0 & -2 & 28 & -7 \\ 0 & 1 & -17 & 8 \end{bmatrix}$$

$$\xrightarrow{L_2 \leftrightarrow L_5} \begin{bmatrix} 1 & -1 & 10 & -4 \\ 0 & \boxed{1} & -17 & 8 \\ 0 & 3 & -47 & 18 \\ 0 & -2 & 28 & -7 \\ 0 & -2 & 32 & -13 \end{bmatrix} \xrightarrow[L_5 \to L_5 + 2L_2]{\substack{L_3 \to L_3 - 3L_2 \\ L_4 \to L_4 + 2L_2}} \begin{bmatrix} 1 & -1 & 10 & -4 \\ 0 & 1 & -17 & 8 \\ 0 & 0 & 4 & -6 \\ 0 & 0 & -6 & 9 \\ 0 & 0 & -2 & 3 \end{bmatrix}$$

$$\xrightarrow[L_5 \to -L_5]{\substack{L_3 \to \frac{1}{2}L_3 \\ L_4 \to -\frac{1}{3}L_4}} \begin{bmatrix} 1 & -1 & 10 & -4 \\ 0 & 1 & -17 & 8 \\ 0 & 0 & \boxed{2} & -3 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 2 & -3 \end{bmatrix} \xrightarrow[L_5 \to L_5 - L_3]{L_4 \to L_4 - L_3} \left[\begin{array}{cccc} 1 & -1 & 10 & -4 \\ 0 & 1 & -17 & 8 \\ 0 & 0 & 2 & -3 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Como existem 3 linhas não nulas na matriz escalonada final, a *característica* da matriz A (tal como a de todas as matrizes que ocorreram ao longo do algoritmo) é 3, visto ser esta a dimensão do subespaço de $\mathbb{R}^4$ gerado pelas suas 5 linhas (as quais são *linearmente dependentes* o que era, aliás, óbvio visto tratar-se de 5 vectores do espaço cartesiano $\mathbb{R}^4$ cuja dimensão é 4); uma *base* deste subespaço é formada pelas 3 primeiras linhas da matriz escalonada obtida.

2.9 Sistemas de equações lineares. Princípios de equivalência

Nesta secção, aborda-se uma importante aplicação da teoria das matrizes e do algoritmo de condensação: a discussão e resolução de sistemas de equações lineares num corpo $\mathbb{K}$ qualquer. Consideraremos sistemas com qualquer número $m \in \mathbb{N}$ de equações e qualquer número $n \in \mathbb{N}$ de incógnitas e com coeficientes e incógnitas pertencentes a um certo corpo $\mathbb{K}$. A chamada *forma canónica* de um tal sistema será

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots & \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{cases} \text{, em que } \begin{cases} a_{ij}, x_j, b_i \in \mathbb{K} \\ i = 1, \ldots, m;\ j = 1, \ldots, n \end{cases} \tag{2.50}$$

ou ainda,

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \text{ , onde } i = 1, \ldots, m \tag{2.50}$$

A forma canónica caracteriza-se, pois, por:

- Estão no 1º membro todos os termos contendo as incógnitas (os termos $a_{ij}x_j$).
- Os termos independentes de incógnitas estão no 2º membro (os termos b_i).
- A ordem das incógnitas é a mesma em todas as equações, embora seja arbitrária.

Se os segundos membros das equações (2.50) forem todos *nulos*, o sistema diz-se *homogéneo*, dizendo-se também homogénea cada equação do sistema.

Poremos, agora, duas definições importantes

Definição 2.14 – Solução de um sistema – *Chama-se* ***solução*** *do sistema* (2.50) *a qualquer vector*

$$\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$$

tal que se verifiquem em simultâneo as m *igualdades* (2.50)*. O conjunto* S *das soluções do sistema é o subconjunto de* $\mathbb{K}^n$ *que resulta da intersecção dos subconjuntos de* $\mathbb{K}^n$ *constituídos pelas soluções de cada equação isoladamente; sendo* S_i *o conjunto das soluções da* i*-ésima equação* (2.50)*, será portanto,*

$$\begin{aligned} S &= \bigcap_{i=1}^{m} S_i \\ S_i &= \left\{ (x_1, x_2, \ldots, x_n) \colon \sum_{j=1}^{n} a_{ij} x_j = b_i \right\} \subset \mathbb{K}^n \end{aligned} \tag{2.51}$$

Definição 2.15 – Sistemas equivalentes – *Diremos que dois sistemas de equações lineares são* ***equivalentes****, sse o conjunto* S *das soluções é o mesmo para os dois sistemas.*

Isto implica imediatamente que o número n de incógnitas tem de ser igual em ambos sistemas, mas não necessariamente o número m de equações, como se ilustra facilmente com o exemplo seguinte: se uma das equações (a equação de ordem k), tiver $a_{kj} = b_k = 0$ (ou seja, a k-ésima equação é $0x_1 + 0x_2 + \cdots + 0x_n = 0$), então será $S_k = \mathbb{K}^n$. Como $\mathbb{K}^n$ é o elemento neutro da intersecção (2.51), podemos eliminar $S_k = \mathbb{K}^n$ daquela intersecção sem alterar o conjunto S, isto é, podemos eliminar a k-ésima equação do sistema, obtendo-se, assim, um novo sistema (de $m-1$ equações) equivalente ao primeiro.

Note-se ainda que, se a equação k do sistema for da forma

$$0x_1 + 0x_2 + \cdots + 0x_n = b_k \neq 0,$$

será $S_k = \emptyset$ e, como $\emptyset$ é elemento absorvente para a intersecção (2.51), virá $S = \emptyset$: o sistema não terá qualquer solução (dizendo-se que é *impossível*, incoerente ou incompatível).

Existem 3 tipos de transformações que podem ser feitas sobre um sistema de equações lineares e que não alteram o conjunto S das suas soluções: essas transformações constam dos chamados *princípios de equivalência* dos sistemas:

Proposição 2.9 – Princípios de equivalência – *Cada uma das seguintes operações, quando executadas sobre um sistema de equações* (2.50)*, não altera o conjunto* S *das soluções definido por* (2.51)*, resultando, portanto, num sistema equivalente ao primeiro:*

i) *Troca entre si de duas equações do sistema.*

ii) *Multiplicação de ambos os membros de uma das equações por um escalar não nulo.*

iii) *Substituição de uma das equações pela que dela se obtém, somando-a membro a membro com outra das equações multiplicada por qualquer escalar.*

Demonstração:

Considerem-se o sistema 1 dado por (2.50) e o sistema 2 obtido após execução de cada uma das operações referidas e designem-se por S_i^1 e S_i^2 os conjuntos de soluções da i-ésima equação daqueles dois sistemas e por S^1 e S^2 os conjuntos das soluções dos mesmos:

i) Resulta imediatamente da comutatividade e associatividade da intersecção (2.51).

ii) Tem-se, neste caso e sendo $1 \leq k \leq m$ e $\alpha \in \mathbb{K} \setminus \{0\}$,

$$S_k^1 = \left\{(x_1, x_2, \ldots, x_n) \colon \sum_{j=1}^{n} a_{kj} x_j = b_k\right\}$$
$$S_k^2 = \left\{(x_1, x_2, \ldots, x_n) \colon \sum_{j=1}^{n} (\alpha a_{kj}) x_j = \alpha b_k\right\}$$

As implicações

$$(x_1, x_2, \ldots, x_n) \in S_k^1 \Rightarrow \sum_{j=1}^{n} a_{kj} x_j = b_k \Rightarrow \sum_{j=1}^{n} (\alpha a_{kj}) x_j = \alpha b_k \Rightarrow (x_1, x_2, \ldots, x_n) \in S_k^2$$

mostram que $S_k^1 \subset S_k^2$. Reciprocamente, se $\alpha \neq 0$, existe α^{-1} e tem-se:

$$(x_1, x_2, \ldots, x_n) \in S_k^2 \Rightarrow \sum_{j=1}^{n} (\alpha a_{kj}) x_j = \alpha b_k \Rightarrow$$
$$\alpha^{-1} \sum_{j=1}^{n} (\alpha a_{kj}) x_j = \alpha^{-1} \alpha b_k \Rightarrow \sum_{j=1}^{n} a_{kj} x_j = b_k \Rightarrow (x_1, x_2, \ldots, x_n) \in S_k^1$$

o que mostra que $S_k^2 \subset S_k^1$ e que, portanto, $S_k^1 = S_k^2$. S é, portanto, igual em ambos os sistemas.

iii) Sejam $1 \leq i \neq j \leq m$ e considere-se o sistema 2 que resulta de (2.50), substituindo a equação i pela sua soma com o produto de um escalar β pela equação j. Comecemos por observar que os conjuntos das soluções de ambos os sistemas se pode decompor da forma seguinte

$$S^1 = \left(S_i^1 \cap S_j^1\right) \cap \left(\bigcap_{k \neq i,j} S_k^1\right)$$
$$S^2 = \left(S_i^2 \cap S_j^2\right) \cap \left(\bigcap_{k \neq i,j} S_k^2\right)$$

Como só a iª equação será distinta em ambos sistemas, será $S_k^1 = S_k^2$, para $k \neq i$ e, portanto $\bigcap_{k \neq i,j} S_k^1 = \bigcap_{k \neq i,j} S_k^2$.

Bastará, agora, provar que $S_i^1 \cap S_j^1 = S_i^2 \cap S_j^2$; tem-se, para o sistema 1(2.50),

$$\begin{cases} S_i^1 = \left\{(x_1, x_2, \ldots, x_n) \colon \sum_{k=1}^{n} a_{ik} x_k = b_i\right\} \subset \mathbb{K}^n \\ S_j^1 = \left\{(x_1, x_2, \ldots, x_n) \colon \sum_{k=1}^{n} a_{jk} x_k = b_j\right\} \subset \mathbb{K}^n \end{cases}$$

e, para o segundo sistema, vem:

$$\begin{cases} S_i^2 = \left\{(x_1, x_2, \ldots, x_n)\colon \sum\limits_{k=1}^{n} (a_{ik} + \beta a_{jk})x_k = b_i + \beta b_j\right\} \subset \mathbb{K}^n \\ S_j^2 = S_j^1 = \left\{(x_1, x_2, \ldots, x_n)\colon \sum\limits_{k=1}^{n} a_{jk}x_k = b_j\right\} \subset \mathbb{K}^n \end{cases}$$

As implicações

$$\begin{aligned} &(x_1, x_2, \ldots, x_n) \in S_i^1 \cap S_j^1 \Rightarrow \sum_{k=1}^{n} a_{ik}x_k = b_i \wedge \sum_{k=1}^{n} a_{jk}x_k = b_j \Rightarrow \\ &\sum_{k=1}^{n} a_{ik}x_k = b_i \wedge \sum_{k=1}^{n} a_{jk}x_k = b_j \wedge \sum_{k=1}^{n} (\beta a_{jk})x_k = \beta b_j \Rightarrow \\ &\sum_{k=1}^{n} (a_{ik} + \beta a_{jk})x_k = b_i + \beta b_j \wedge \sum_{k=1}^{n} a_{jk}x_k = b_j \Rightarrow (x_1, x_2, \ldots, x_n) \in S_i^2 \cap S_j^2 \end{aligned}$$

mostram que $S_i^1 \cap S_j^1 \subset S_i^2 \cap S_j^2$. Por outro lado, usando o mesmo raciocínio, mas com $-\beta$ em vez de β, conclui-se que $S_i^2 \cap S_j^2 \subset S_i^1 \cap S_j^1$ e que, portanto, $S_i^1 \cap S_j^1 = S_i^2 \cap S_j^2$. □

Com base no que vimos anteriormente, podemos ainda observar que, a estes 3 princípios de equivalência, podemos juntar uma 4ª operação que conduz a um sistema equivalente (mas com $m - 1$ equações): a que nos permite retirar do sistema qualquer equação do tipo $0x_1 + 0x_2 + \cdots + 0x_n = 0$.

É costume classificarem-se os sistemas (2.50), consoante a cardinalidade do conjunto $S \subset \mathbb{K}^n$ das suas soluções, do modo seguinte

$$\text{Sistema}\begin{cases} \text{Possível } (S \neq \emptyset)\begin{cases} \text{Determinado } (\#S = 1) \\ \text{Indeterminado } (\#S > 1) \end{cases} \\ \text{Impossível } (S = \emptyset) \end{cases}$$

2.10 Formas matricial e vectorial de um sistema

Vamos, agora, ver que as m equações simultâneas (2.50) equivalem a uma única igualdade matricial e ainda a uma igualdade vectorial; estas igualdades constituem a *forma matricial* do sistema e a sua *forma vectorial*. Para tal, introduzam-se a *matriz dos coeficientes* ou *matriz simples* do sistema $A = [\,a_{ij}\,] \in \mathbb{K}^{m,n}$, a *matriz-coluna ou vector-coluna das incógnitas* $X = [\,x_1 \quad x_2 \quad \cdots \quad x_n\,]^{\mathrm{T}}$ e a *matriz-coluna ou vector-coluna dos termos independentes* $B = [\,b_1 \quad b_2 \quad \cdots \quad b_m\,]^{\mathrm{T}}$. Usando estas matrizes, facilmente se reconhece que as m igualdades simultâneas (2.50) equivalem à igualdade matricial única

$$AX = B \tag{2.52}$$

A igualdade anterior constitui a *forma matricial* do sistema de equações, que tem (2.50) por forma canónica. Nesta forma, é evidente que cada X satisfazendo a igualdade anterior constitui uma *solução* do sistema. É ainda útil introduzir-se a *matriz completa* ou *ampliada* do sistema $A' = \big[A \,\big|\, B\big] \in \mathbb{K}^{m,n+1}$.

Dado um sistema $AX = B$, diz-se que o sistema $AX = O$ (com idêntica matriz dos coeficientes A) é o *sistema homogéneo associado a* $AX = B$; este sistema homogéneo tem, portanto, os mesmos coeficientes a_{ij} das incógnitas, mas termos independentes b_i nulos.

Fragmentando a matriz A nas suas n colunas $A_1, A_2, \ldots, A_n$ e a matriz X nas suas n linhas (que são, afinal, as incógnitas x_j), obtém-se

$$x_1 A_1 + x_2 A_2 + \cdots + x_n A_n = B \tag{2.53.1}$$

ou, mais abreviadamente,

$$\sum_{j=1}^{n} x_j A_j = B \tag{2.53.2}$$

É claro que poderíamos usar a notação vectorial, designando-se as colunas de A por $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ e B por $\vec{b}$, vindo então

$$\sum_{j=1}^{n} x_j \vec{a}_j = \vec{b} \tag{2.53.3}$$

Qualquer das igualdades (2.53) é uma igualdade vectorial em $\mathbb{K}^m$ que traduz o facto de ser B *combinação linear* das colunas $(A_j)_{1 \le j \le n}$ da matriz A. Qualquer das igualdades vectoriais (2.53) é designada por *forma vectorial* do sistema (2.50) e esta forma é a chave para a resolução das duas principais questões que se põem relativamente ao sistema:

- Qual a condição necessária e suficiente para que o sistema seja possível?
- Verificando-se a condição anterior, em que condições será o sistema determinado?

A seguinte proposição responde às duas questões anteriores:

Proposição 2.10 – Teorema da existência e da unicidade – *Considere-se um sistema de m equações lineares a n incógnitas, $AX = B$, com matriz completa A' e sejam $r = \mathrm{c}(A)$ e $s = \mathrm{c}(A')$. Então:*

i) *O sistema é possível sse $s = r$.*

ii) *O sistema é determinado sse, além de $s = r$, é ainda $r = n$.*

Demonstração:

i) Comecemos por observar que, como A' tem apenas mais uma coluna do que A, a sua característica s só poderá ser igual a r ou a $r + 1$. A questão da existência de solução para o sistema pode ter uma formulação vectorial, graças a (2.53): o sistema é *possível* sse o vector B pertencer ao subespaço de $\mathbb{K}^m$ gerado pela lista $A = (A_1, A_2, \ldots, A_n)$ de vectores de $\mathbb{K}^m$ e, pelo que vimos no capítulo 1, isso acontece sse $L(A_1, A_2, \ldots, A_n) = L(A_1, A_2, \ldots, A_n, B)$ ou seja, sse $s = r$.

ii) Supondo que $s = r$, ter-se-á $B \in L(A_1, A_2, \ldots, A_n)$ e o sistema será *determinado* sse os escalares x_j em (2.53) forem únicos. Para que isso aconteça, a proposição 1.8 mostra que deverá a lista $(A_1, A_2, \ldots, A_n)$ das colunas de A ser *linearmente independente* (ou seja, deverá esta lista constituir uma *base* de $L(A_1, A_2, \ldots, A_n)$), pelo que a característica r de A deverá ser igual a n. $\square$

Observações:

- A proposição 2.8.*iv* mostra que é sempre $r \leq n$ e, quando um sistema for possível ($s = r$), a diferença $n - r \geq 0$ é chamada o *grau de indeterminação* do sistema. Se este for igual a 1, diremos que o sistema é *simplesmente indeterminado*; quando for igual a 2, diremos que *é duplamente indeterminado* e assim sucessivamente. Um sistema é, pois, *determinado* sse $s = r$, ou seja, sse o seu grau de indeterminação $n - r$ for igual a 0.

- Podemos, agora, precisar melhor a classificação dos sistemas feita anteriormente

$$\text{Sistema}\begin{cases} s = r \Rightarrow \text{Possível } (S \neq \emptyset)\begin{cases} r = n \Rightarrow \text{Determinado } (\#S = 1). \\ r < n \Rightarrow \text{Indeterminado de grau } n - r > 0 \ (\#S > 1). \end{cases} \\ s = r + 1 \Rightarrow \text{Impossível } (S = \emptyset). \end{cases}$$

- Suponha-se que a característica r de A é igual ao número de equações m; neste caso, é necessariamente $s = m$, visto que A' tem também m linhas e não pode ter característica $m + 1$: o sistema é, portanto, *possível*, para todo o $B \in \mathbb{K}^m$. Outra forma de obter este resultado é observar que, quando $r = m$, fica $L(A_1, A_2, \ldots, A_n) = \mathbb{K}^m$, o que leva a que todo o $B \in \mathbb{K}^m$ seja combinação linear dos $(A_1, A_2, \ldots, A_n)$.

- Pelo que acabámos de ver, $r < m$ é condição necessária (mas não suficiente) para que um sistema seja *impossível*.

- Note-se ainda que se for $m < n$ (número de incógnitas superior ao de equações), ter-se-á concerteza $r < n$ e isto significa que o sistema não pode ser determinado (só poderá ser *impossível* ou *indeterminado*).

- Tudo o que acima foi dito pode ser usado para obter uma formulação alternativa da discussão de um sistema:

$$\text{Sistema}\begin{cases} r = m \Rightarrow s = m \Rightarrow \text{Possível } (S \neq \emptyset)\begin{cases} r = n \Rightarrow \text{Determinado } (\#S = 1). \\ r < n \Rightarrow \text{Indeterminado de grau } n - r > 0 \ (\#S > 1). \end{cases} \\ r < m \Rightarrow \begin{cases} s = r \Rightarrow \text{Possível } (S \neq \emptyset)\begin{cases} r = n \Rightarrow \text{Determinado } (\#S = 1). \\ r < n \Rightarrow \text{Indeterminado de grau } n - r > 0 \ (\#S > 1). \end{cases} \\ s = r + 1 \Rightarrow \text{Impossível } (S = \emptyset). \end{cases} \end{cases}$$

2.11 Algoritmo de Gauss-Jordan

Observando que existe um perfeito paralelismo entre os princípios de equivalência dos sistemas e as 3 operações elementares referidas na proposição 1.18, vemos que, aplicando estas às linhas (e só às linhas) da matriz completa A', obtém-se a matriz completa de um novo sistema que será equivalente ao primeiro. Observe-se que a única operação admitida sobre colunas é a troca de colunas da matriz A, a qual corresponde à *reordenação* das incógnitas nas equações (2.50), devendo, no final, ter-se essa reordenação em conta. Se usarmos o algoritmo de condensação anteriormente referido, poderemos discutir e posteriormente resolver qualquer sistema de equações lineares dado; tal é a base do *algoritmo de Gauss-Jordan* para a resolução de sistemas de equações lineares:

- Quando a matriz completa A' estiver na forma escalonada, o mesmo se passa com a matriz A dos coeficientes e ficaremos a conhecer as características s e r, assim como se o sistema é ou não possível.

- Sendo o sistema possível e comparando r com n, saberemos se a solução é única.

- Após levar a matriz completa A' à forma escalonada, usam-se os r pivots já determinados para, com operações do tipo 3 ou 4, anularmos os elementos da matriz A situados nas r colunas dos pivots e acima destes, após o que poderemos trocar as colunas de A por forma a fazer com que as colunas contendo os pivots ocupem as primeiras r posições em A (isso equivale a alterar a ordem das incógnitas, o que deverá ser levado em conta, aquando da obtenção das soluções). Dividindo, se necessário, cada linha i (entre 1 e r) por a_{ii}, obtém-se a matriz completa A' de um sistema equivalente ao sistema dado e contendo, nas suas primeiras r filas, a *matriz identidade* de ordem r, como se indica a seguir

$$A' = \left[\begin{array}{c|c|c} I_r & C_{r,n-r} & B_{r,1} \\ \hline O_{m-r,r} & O_{m-r,n-r} & D_{m-r,1} \end{array}\right]$$

As primeiras r incógnitas (correspondentes às colunas de I_r) são chamadas *incógnitas principais* e, se o sistema for possível e determinado, todas as incógnitas são principais; as restantes $n-r$ incógnitas (correspondentes às colunas de $C_{r,n-r}$) são chamadas *incógnitas secundárias* ou *livres* (em número igual ao *grau de indeterminação* do sistema). Note-se que não existem, quando o sistema for determinado ($n-r=0$).

Podem, então, acontecer 3 casos:

1. $D_{m-r,1} \neq O_{m-r,1}$: neste caso, será $s=r+1$ e o sistema será *impossível*.

2. $D_{m-r,1} = O_{m-r,1}$: neste caso, será $s=r$ e o sistema será *possível*. Poderão agora acontecer duas situações:

2.1. $r=n$ (*sistema determinado*): neste caso, as matrizes $C_{r,n-r}$ e $O_{m-r,n-r}$ não existem, não havendo incógnitas secundárias e ficando apenas

$$A' = \left[\begin{array}{c|c} I_n & B_{n,1} \\ \hline O_{m-n,n} & O_{m-n,1} \end{array}\right]$$

A solução do sistema será, obviamente, $X = B_{n,1}$.

2.2. $r<n$ (*sistema indeterminado* de grau $n-r>0$): neste caso, existem r incógnitas principais e $n-r$ incógnitas secundárias, cujos valores ficam arbitrários. A matriz completa é

$$A' = \left[\begin{array}{c|c|c} I_r & C_{r,n-r} & B_{r,1} \\ \hline O_{m-r,r} & O_{m-r,n-r} & O_{m-r,1} \end{array}\right]$$

Fragmentando X em $\left[\begin{array}{c} X_r \\ \hline X_{n-r} \end{array}\right]$, onde X_r é matriz-coluna contendo as r incógnitas principais e X_{n-r} é outra matriz-coluna contendo as $n-r$ incógnitas secundárias, o produto AX realizado por blocos dá

$$AX = \left[\begin{array}{c|c} I_r & C_{r,n-r} \\ \hline O_{m-r,r} & O_{m-r,n-r} \end{array}\right] \left[\begin{array}{c} X_r \\ \hline X_{n-r} \end{array}\right] = \left[\begin{array}{c} X_r + C_{r,n-r}X_{n-r} \\ \hline O_{m-r,1} \end{array}\right] = \left[\begin{array}{c} B_{r,1} \\ \hline O_{m-r,1} \end{array}\right]$$

Da igualdade anterior, segue-se que

$$X_r + C_{r,n-r}X_{n-r} = B_{r,1}$$

donde,

$$X_r = B_{r,1} - C_{r,n-r}X_{n-r}$$

Por fim, obtêm-se as soluções do sistema de equações lineares $AX = B$, em função das $n - r$ incógnitas secundárias, contidas em X_{n-r}.

$$\begin{aligned} X &= \left[\begin{array}{c} X_r \\ \hline X_{n-r} \end{array}\right] = \left[\begin{array}{c} B_{r,1} - C_{r,n-r}X_{n-r} \\ \hline X_{n-r} \end{array}\right] \\ &= \left[\begin{array}{c} B_{r,1} - C_{r,n-r}X_{n-r} \\ \hline O_{n-r,1} + I_{n-r}X_{n-r} \end{array}\right] = \left[\begin{array}{c} B_{r,1} \\ \hline O_{n-r,1} \end{array}\right] + \left[\begin{array}{c} -C_{r,n-r} \\ \hline I_{n-r} \end{array}\right] X_{n-r} \end{aligned} \tag{2.54}$$

As $n - r$ incógnitas secundárias presentes em X_{n-r} ficam arbitrárias, tomando, cada uma delas, qualquer valor do corpo $\mathbb{K}$:

- Se $\mathbb{K}$ for *infinito*, haverá uma infinidade de soluções para o sistema (o conjunto S será infinito); se $n - r = 1$, diz-se que há uma *infinidade simples* de soluções; se $n - r = 2$, diz-se que há uma *dupla infinidade* de soluções e assim por diante.

- Se o corpo $\mathbb{K}$ for *finito*, então será também finito o número de soluções do sistema, tendo-se

$$\#S = (\#\mathbb{K})^{n-r} \tag{2.55}$$

Na última das igualdades (2.54), vemos que a *solução geral* X de $AX = B$ é soma de

$$X_0 = \left[\begin{array}{c} B_{r,1} \\ \hline O_{n-r,1} \end{array}\right]$$

(que é uma *solução particular* do sistema – exactamente aquela que se obtém, fazendo $X_{n-r} = O$) com

$$Y = \left[\begin{array}{c} -C_{r,n-r} \\ \hline I_{n-r} \end{array}\right] X_{n-r}$$

que é a *solução geral* do sistema homogéneo associado $AX = O$ e onde X_{n-r} é vector-coluna contendo as $n - r$ incógnitas secundárias $(x_{i_1}, x_{i_2}, \ldots, x_{i_{n-r}})$. É claro que, se o desejarmos, podemos substituir as $n - r$ incógnitas secundárias por outros tantos parâmetros escalares arbitrários $(t_k)_{1 \leq k \leq n-r}$.

Fragmentando a matriz $\left[\begin{array}{c} -C_{r,n-r} \\ \hline I_{n-r} \end{array}\right] \in \mathbb{K}^{n,n-r}$ nas suas $n - r$ colunas $Y_1, Y_2, \ldots, Y_{n-r}$,

$$\left[\begin{array}{c} -C_{r,n-r} \\ \hline I_{n-r} \end{array}\right] = \left[\begin{array}{c|c|c|c} Y_1 & Y_2 & \cdots & Y_{n-r} \end{array}\right]$$

e fragmentando também X_{n-r} nas suas $n-r$ linhas (os parâmetros $t_k = x_{i_k}$), obtém-se

$$X = X_0 + \sum_{k=1}^{n-r} t_k Y_k \tag{2.56}$$

Vemos, agora, que as soluções de $AX = O$ formam o conjunto

$$S_0 = \left\{ Y \colon Y = \sum_{k=1}^{n-r} t_k Y_k \wedge t_k \in \mathbb{K} \right\} = L(Y_1, Y_2, \ldots, Y_{n-r}) \subset \mathbb{K}^n$$

S_0 é, portanto, o subespaço de $\mathbb{K}^n$ gerado por $(Y_1, Y_2, \ldots, Y_{n-r})$ e estes são linearmente independentes formando, pois, uma *base* de S_0; então:

$$\mathrm{dim} S_0 = n - r \tag{2.57}$$

O facto de que o conjunto S_0 das soluções de um sistema homogéneo $AX = O$ constitui sempre (ao contrário do que acontece com o conjunto das soluções de um sistema não homogéneo) um *subespaço* vectorial de $\mathbb{K}^n$ pode ser provado directamente, atendendo às condições S1, S2 e S3 mencionadas no capítulo 1:

- Condição S1:

É claro que um sistema homogéneo tem sempre, pelo menos, a solução $X = O$, portanto

$$O \in S_0$$

- Condição S2:

$$\begin{aligned} X, X' \in S_0 &\Rightarrow AX = O \wedge AX' = O \Rightarrow AX + AX' = O + O \\ &\Rightarrow A(X + X') = O \Rightarrow X + X' \in S_0 \end{aligned}$$

- Condição S3:

$$\alpha \in \mathbb{K} \wedge X \in S_0 \Rightarrow \alpha \in \mathbb{K} \wedge AX = O \Rightarrow \alpha(AX) = \alpha O \Rightarrow A(\alpha X) = O \Rightarrow \alpha X \in S_0$$

Também a decomposição (2.56) pode ser deduzida de uma forma directa, como mostra a proposição seguinte que estabelece a relação entre as soluções de $AX = B$ e de $AX = O$:

Proposição 2.11 – Relação entre as soluções de $AX = B$ **e** $AX = O$ **–** *Sejam $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,1}$ matrizes dadas e considere-se o sistema de m equações e n incógnitas $AX = B$, com X matriz das incógnitas e o respectivo sistema homogéneo associado $AX = O$. Sempre que o sistema $AX = B$ seja possível, a sua solução geral X obtém-se de uma solução particular X_0, somando-lhe a solução geral Y do sistema homogéneo associado $AX = O$.*

Demonstração:

Sejam S e $S_0^{(8)}$ os conjuntos de soluções de $AX = B$ e de $AX = O$, respectivamente, e fixemos uma solução $X_0 \in S$ qualquer de $AX = B$ (solução particular). Comecemos por provar que a soma de X_0 com qualquer solução $Y \in S_0$ de $AX = O$ dá-nos ainda uma solução de $AX = B$:

$$\begin{aligned} X_0 \in S \wedge Y \in S_0 &\Rightarrow AX_0 = B \wedge AY = O \Rightarrow AX_0 + AY = B + O \\ &\Rightarrow A(X_0 + Y) = B \Rightarrow X_0 + Y \in S \end{aligned}$$

Porém, todas as soluções de $AX = B$ são desta forma, visto que, sendo X qualquer solução de $AX = B$, se tem

$$\begin{aligned} X \in S \wedge X_0 \in S &\Rightarrow AX = B \wedge AX_0 = B \Rightarrow AX - AX_0 = B - B \\ &\Rightarrow A(X - X_0) = O \Rightarrow X - X_0 \in S_0 \end{aligned}$$

Mas, como $X = X_0 + \underbrace{(X - X_0)}_{\in S_0}$, conclui-se que a solução X é, de facto, soma da solução particular X_0 com uma solução (que é $X - X_0 = Y$) de $AX = O$. $\square$

Da proposição anterior, segue-se que, para qualquer $X_0 \in S$, se tem:

$$S = \left\{ X \colon \underset{Y}{\exists} (X = X_0 + Y \wedge Y \in S_0) \right\} \tag{2.58}$$

Podemos, pois, afirmar que o conjunto S resulta de fazer uma *translação* de S_0, definida pelo vector X_0 (ver esquema na figura 2.1).

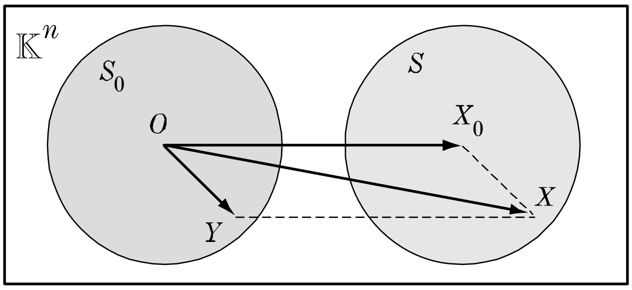


Fig. 2.1 – Relação entre as soluções dos sistemas $AX = B$ e $AX = O$.

Se o sistema $AX = B$ é real, com 2 incógnitas e possível, então:

- Se é *determinado*, $S_0 = \{(0,0)\}$ e $S = \{X_0\}$ são conjuntos singulares.
- Se é *simplesmente indeterminado*, S_0 e S são duas *rectas* paralelas em que S_0 passa pela origem (ver figura 2.2).
- Se é *duplamente indeterminado*, então $S_0 = S = \mathbb{R}^2$ (situação só possível com $A = O$ e $B = O$).

---

8 S_0 e S são respectivamente um *subespaço* de $\mathbb{K}^n$ e uma *variedade linear* (ver capítulo 7) de dimensão $n - r$.

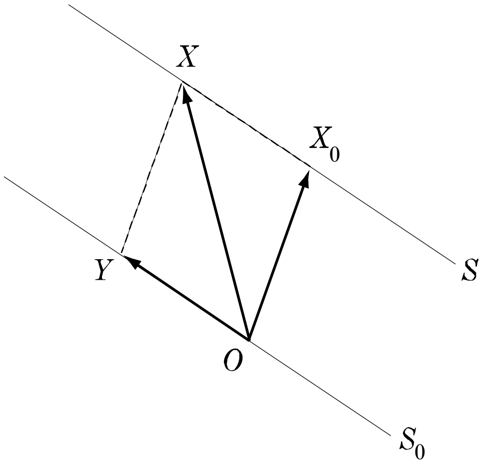


Fig. 2.2 – Caso de um sistema simplesmente indeterminado.

Se o sistema $AX = B$ é real, com 3 incógnitas e possível, então:

- Se é *determinado*, $S_0 = \{(0,0,0)\}$ e $S = \{X_0\}$ são conjuntos singulares.

- Se é *simplesmente indeterminado*, S_0 e S são duas *rectas* paralelas em que S_0 passa pela origem (ver figura 2.2).

- Se é *duplamente indeterminado*, S_0 e S são dois *planos* paralelos em que S_0 passa pela origem (ver figura 2.3).

- Se é *triplamente indeterminado*, então $S_0 = S = \mathbb{R}^3$ (situação só possível com $A = O$ e $B = O$).

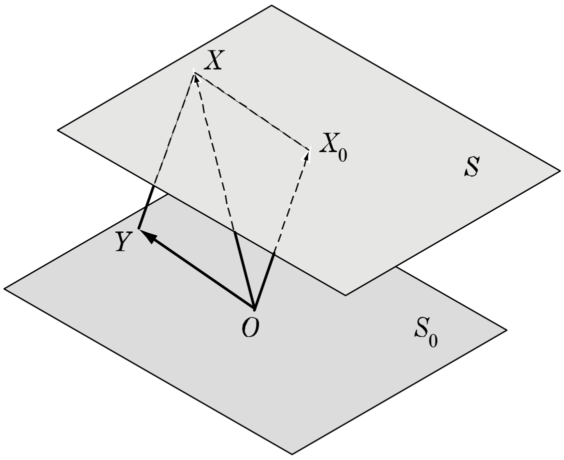


Fig. 2.3 – Caso de um sistema duplamente indeterminado.

A seguir, apresenta-se formalmente, o algoritmo de Gauss-Jordan, para a resolução de um sistema de m equações a n incógnitas, numa versão em *pseudo-código* e, na secção 2.16, apresenta-se a implementação na linguagem do software MATHEMATICA© constituindo a função `GaussJordan`. O algoritmo tem como entradas a matriz a_{ij} dos coeficientes e o vector b_i dos termos independentes e, como saída, uma solução particular $(sp_j)_{1\leq j\leq n}$ do sistema e os vectores $(sh_{kj})_{\substack{1\leq k\leq n-r\\ 1\leq j\leq n}}$ de uma base do espaço das soluções do sistema homogéneo associado; valem as convenções tipográficas feitas na pág. 98 do presente capítulo:

```
m=#Linhas(a); n=#Colunas(a);
s=0; r=0; k=0;
c=[a,b];                % Matriz completa
Para k=1 até n
  piv_k=False           % piv_k=True sse k^a incógnita é principal
FimPara;
Enquanto s < m ∧ k < n+1
    % Busca do Pivot:
    Fazer
        k=k+1;
        i=s;
        Fazer
          i=i+1
        Enquanto c_ik = 0 ∧ i < m
    Enquanto c_ik = 0 ∧ k < n+1;
    Se c_ik ≠ 0 então                          % c_ik é pivot?
        % Redução da coluna k de c:
        s=s+1;
        Se k ≤ n então
          r=r+1;
          piv_k=True
        FimSe;
        Se i ≠ s então Linha_i ↔ Linha_s;         % Op. de tipo 1
        Linha_s=Linha_s/c_sk;                     % Op. de tipo 2
        Para i=1 até m
          Se i ≠ s então Linha_i=Linha_i-c_ik*Linha_s   % Op. de tipo 3
        FimPara
    FimSe
FimEnquanto;

% Discussão da solução e saída de resultados:
Se s ≠ r então
  Saída: Sistema impossível
Senão
  % Solução particular sp:
  Para k=1 até n
    sp_k=0
  FimPara;
  i=0;
  Para k=1 até n
    Se piv_k = True então
      i=i+1;
      sp_k = c_i,n+1
    FimSe
  FimPara;

  % Solução sh do sistema homogéneo associado:
  Para i=1 até n-r
    Para k=1 até n
      sh_ik=0
    FimPara
  FimPara;
  t=0;
  Para s=1 até n
    Se piv_s = False então
      t=t+1;
      i=0;
      Para k=1 até n
        Se piv_k = True então
          i=i+1;
          sh_tk=-c_is
        Senão
          sh_ts=1
        FimSe
```

```
      FimPara
    FimSe
  FimPara;
  Saída: {sp, sh}
FimSe
```

Observações:

- O algoritmo de eliminação Gauss-Jordan conduz a um grande número de operações de divisão/multiplicação e de subtracção. Para obtermos uma estimativa do número de operações necessárias à resolução de um sistema, consideremos a aplicação do método a um sistema de n equações e n incógnitas e convencionando considerar a multiplicação e subtração incluídas nas operações de tipo 3 como uma operação única (o equivale a contar apenas as operações de multiplicação/divisão); Para a condensação da matriz completa $\left[A \,|\, B\right]$, teremos:

– Considerando o 1º pivot $a_{11} \neq 0$, para anular o elemento a_{i1} da linha i e 1ª coluna teremos que fazer a operação $L_i \to L_i - \frac{a_{i1}}{a_{11}} L_1$ de tipo 3 e isto traduz-se por 1 divisão ($\frac{a_{i1}}{a_{11}}$) e ainda n operações de multiplicação/subtracção (correspondentes aos $n-1$ elementos a_{ij} da linha i, com $1 < j \leq n$ e ao termo independente b_i), o que perfaz um total de $1 + (n-1) + 1 = n + 1$ operações. Para anular todos os elementos a_{i1} para $i > 1$, faremos portanto $(n-1)(n+1) = n^2 - 1$ operações. Para os pivots seguintes, o número de operações irá sempre diminuindo, de modo que, para se obter uma matriz em escada, será necessário um total T_1 de operações dado por

$$T_1 = \sum_{k=1}^{n} \left(k^2 - 1\right) = \sum_{k=1}^{n} k^2 - \sum_{k=1}^{n} 1 = \frac{1}{6} n(n+1)(2n+1) - n = \frac{1}{6} n\big(n-1\big)(2n+5) \approx \frac{1}{3} n^3$$

Para valores de n grandes, $n^3/3$ será uma boa estimativa do número de operações necessárias à obtenção da matriz escalonada (dizemos que o número de operações é da ordem de $n^3/3$).

– A fase de retro-substituição requer uma operação (divisão) para a incógnita x_n, duas para x_{n-1}, três para x_{n-3} e assim por diante. O total T_2 de operações nesta fase será de

$$T_2 = \sum_{k=1}^{n} k = \frac{1}{2} n(n+1)$$

– Do que vimos nos pontos anteriores, o total geral T de operações conducentes à resolução de um sistema de n equações e n incógnitas é portanto de

$$T = T_1 + T_2 = \frac{1}{6} n\big(n-1\big)(2n+5) + \frac{1}{2} n(n+1) = \frac{1}{3} n\big(n^2 + 3n - 1\big) \approx \frac{1}{3} n^3$$

Para valores elevados de n, o número de operações necessárias à resolução de um sistema de n equações lineares e n incógnitas é, pois, da ordem de $n^3/3$.

- O algoritmo é muito sensível à propagação de erros de truncatura nos cálculos, podendo facilmente conduzir a soluções completamente erradas, para determinadas matrizes “mal-condicionadas”. A situação de instabilidade nos cálculos ocorre em particular quando os pivots têm um valor pequeno (em módulo); para evitar pivots pequenos, uma possível estratégia chamada *escolha parcial de pivots*[9] consiste em escolher o elemento de maior valor absoluto na coluna relevante para pivot (em vez do primeiro não nulo dessa coluna, como fizémos) à custa de uma troca de linhas (operação de tipo 1); um aperfeiçoamento adicional deste método é a chamada *escolha total de pivots*[10], na qual se escolhe para pivot o elemento de maior valor absoluto considerando também as colunas de A à direita da coluna relevante e fazendo uma troca de colunas – e de linhas – para levar o pivot à posição corrente em consideração (reordenação das incógnitas e operação de tipo 1).

Exemplo 2.27 Considere-se o sistema de 4 equações lineares a 3 incógnitas reais

$$\begin{cases} x+2y+2z &= 2 \\ 3x-2y-z &= 5 \\ 2x-5y+3z &= -4 \\ x+4y+6z &= 0 \end{cases}$$

Condensemos a matriz completa do sistema:

$$A' = \left[\begin{array}{ccc|c} 1 & 2 & 2 & 2 \\ 3 & -2 & -1 & 5 \\ 2 & -5 & 3 & -4 \\ 1 & 4 & 6 & 0 \end{array}\right] \to \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & -8 & -7 & -1 \\ 0 & -9 & -1 & -8 \\ 0 & 2 & 4 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & -8 & -7 & -1 \\ 0 & -9 & -1 & -8 \\ 0 & 1 & 2 & -1 \end{bmatrix} \to$$
$$\begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & -9 & -1 & -8 \\ 0 & -8 & -7 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 17 & -17 \\ 0 & 0 & 9 & -9 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \to \left[\begin{array}{ccc|c} 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Neste caso, é $s = r = n = 3$ e, portanto, o sistema é *possível* e *determinado*. A última linha nula pode ser retirada da matriz, obtendo-se a matriz completa de um sistema *equivalente* ao dado e, para determinar a única solução existente, basta transformar a matriz dos coeficientes (quadrada de 3ª ordem) na matriz identidade

$$\begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

O conjunto das soluções do sistema é, pois, $S = \{(2, 1, -1)\}$.

Exemplo 2.28 Vejamos, agora, o sistema em $\mathbb{R}$

$$\begin{cases} x+5y+4z-13w = 3 \\ 3x-y+2z+5w &= 2 \\ 2x+2y+3z-4w = 1 \end{cases}$$

[9] Em inglês, *partial pivoting*.

[10] Em inglês, *total pivoting*.

Procedendo como anteriormente, obtém-se

$$\left[\begin{array}{cccc|c}1 & 5 & 4 & -13 & 3\\ 3 & -1 & 2 & 5 & 2\\ 2 & 2 & 3 & -4 & 1\end{array}\right] \rightarrow \left[\begin{array}{ccccc}1 & 5 & 4 & -13 & 3\\ 0 & -16 & -10 & 44 & -7\\ 0 & -8 & -5 & 22 & -5\end{array}\right] \rightarrow \left[\begin{array}{cccc|c}1 & 5 & 4 & -13 & 3\\ 0 & -16 & -10 & 44 & -7\\ 0 & 0 & 0 & 0 & 3\end{array}\right]$$

Tem-se, neste caso, $s = 3$ e $r = 2$ e, então, o sistema não tem qualquer solução (sistema *impossível*), sendo $S = \emptyset$.

Exemplo 2.29 Seja, agora, o sistema real

$$\begin{cases} x + 2y - 3z + 2w = 2\\ 2x + 5y - 8z + 6w = 5\\ 3x + 4y - 5z + 2w = 4\end{cases}$$

Partindo, de novo, da matriz completa e condensando, vem

$$\left[\begin{array}{cccc|c}1 & 2 & -3 & 2 & 2\\ 2 & 5 & -8 & 6 & 5\\ 3 & 4 & -5 & 2 & 4\end{array}\right] \rightarrow \left[\begin{array}{ccccc}1 & 2 & -3 & 2 & 2\\ 0 & 1 & -2 & 2 & 1\\ 0 & -2 & 4 & -4 & -2\end{array}\right] \rightarrow \left[\begin{array}{cccc|c}1 & 2 & -3 & 2 & 2\\ 0 & 1 & -2 & 2 & 1\\ 0 & 0 & 0 & 0 & 0\end{array}\right]$$

Agora, será $s = r = 2$ e $n = 4$; consequentemente, o sistema é *possível* e *duplamente indeterminado* ($n - r = 4 - 2 = 2$). As incógnitas principais são as que correspondem às colunas onde ocorreram os pivots, ou seja, são x e y; as incógnitas secundárias são as restantes z e w. De novo, a última linha pode ser eliminada, obtendo-se a matriz completa de um sistema *equivalente* ao sistema inicial e, para obter a *solução geral* do sistema (que já sabemos ter uma *dupla infinidade* de soluções), basta transformar a matriz de 2ª ordem formada pelas duas colunas das incógnitas principais na matriz identidade I_2:

$$\left[\begin{array}{ccccc}1 & 2 & -3 & 2 & 2\\ 0 & 1 & -2 & 2 & 1\end{array}\right] \rightarrow \left[\begin{array}{cc|cc|c}1 & 0 & 1 & -2 & 0\\ 0 & 1 & -2 & 2 & 1\end{array}\right]$$

Tem-se, portanto

$$\begin{cases} x = -z + 2w\\ y = 1 + 2z - 2w\\ z = z\\ w = w\end{cases}, \text{ em que } z, w \in \mathbb{R}$$

O conjunto S das soluções do sistema dado é, portanto,

$$\begin{aligned} S &= \{(-z + 2w, 1 + 2z - 2w, z, w) \colon z, w \in \mathbb{R}\}\\ &= \{(0, 1, 0, 0) + z(-1, 2, 1, 0) + w(2, -2, 0, 1) \colon z, w \in \mathbb{R}\} \subset \mathbb{R}^4\end{aligned}$$

Exemplo 2.30 Considere-se o corpo $(\mathbb{Z}_5, +, \times)$ dos inteiros módulo 5, onde as operações estão definidas por

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\times$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Tendo em atenção estas tabelas, podemos construir a tabela dos *simétricos* e *inversos*:

x	$-x$	x^{-1}
0	0	$-$
1	4	1
2	3	3
3	2	2
4	1	4

Determinemos as soluções do sistema de equações lineares nas incógnitas $x, y, z \in \mathbb{Z}_5$:

$$\begin{cases} x - 2y + 2z &= 3 \\ 3y - 4z &= 1 \\ 3x + 3y - 2z &= 3 \end{cases}$$

Procedendo como nos exemplos anteriores e atendendo às tabelas anteriores para as operações em $\mathbb{Z}_5$, vem

$$\left[\begin{array}{ccc|c} 1 & -2 & 2 & 3 \\ 0 & 3 & -4 & 1 \\ 3 & 3 & -2 & 3 \end{array}\right] \to \begin{bmatrix} 1 & 3 & 2 & 3 \\ 0 & 3 & 1 & 1 \\ 3 & 3 & 3 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 2 & 3 \\ 0 & 3 & 1 & 1 \\ 0 & 4 & 2 & 4 \end{bmatrix} \to \left[\begin{array}{ccc|c} 1 & 3 & 2 & 3 \\ 0 & 3 & 1 & 1 \\ 0 & 0 & 2 & 3 \end{array}\right]$$

Aqui, é $s = r = n = 3$ e o sistema é *possível* e *determinado*. Vamos, agora, obter a solução do sistema, observando que $2^{-1} = 3$ e $4^{-1} = 4$,

$$\begin{bmatrix} 1 & 3 & 2 & 3 \\ 0 & 3 & 1 & 1 \\ 0 & 0 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 2 & 3 \end{bmatrix} \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

A solução do sistema é, portanto, $(3, 4, 4)$.

Exemplo 2.31 Considere, de novo em $\mathbb{Z}_5$, o sistema

$$\begin{cases} -4x + 3y - 3z = -1 \\ 2x - 4y + 3z &= 0 \\ -x + 2y - 3z &= 3 \end{cases}$$

Tem-se, agora,

$$\left[\begin{array}{ccc|c} -4 & 3 & -3 & -1 \\ 2 & -4 & 3 & 0 \\ -1 & 2 & -3 & 3 \end{array}\right] \to \begin{bmatrix} 1 & 3 & 2 & 4 \\ 2 & 1 & 3 & 0 \\ 4 & 2 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 2 & 4 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 4 & 2 \end{bmatrix} \to \left[\begin{array}{ccc|c} 1 & 3 & 2 & 4 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

O sistema tem $s = r = 2$ e $n = 3$, sendo *possível* e *simplesmente indeterminado*. Para obter as soluções, continuamos a condensação, eliminando a última linha e trocando entre si a segunda e a terceira coluna (nova ordem para as incógnitas: x, z, y),

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} 3 & 0 & 4 & 4 \\ 0 & 4 & 0 & 2 \end{bmatrix} \to \left[\begin{array}{cc|c|c} 1 & 0 & 3 & 3 \\ 0 & 1 & 0 & 3 \end{array}\right]$$

As incógnitas principais são x e z e a incógnita secundária é y. A matriz completa obtida permite escrever a solução geral do sistema:

$$\begin{cases} x = 3 - 3y = 3 + 2y \\ y = y \\ z = 3 \end{cases}, \text{ onde } y \in \mathbb{Z}_5$$

O conjunto S das soluções do sistema tem, portanto, 5 elementos (ver equação (2.55)), obtidos da solução geral anterior fazendo y tomar todos os valores de $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$:

$$S = \{(3, 0, 3), (0, 1, 3), (2, 2, 3), (4, 3, 3), (1, 4, 3)\}$$

Exemplo 2.32 Consideremos o sistema real seguinte, onde a e b são parâmetros reais

$$\begin{cases} x + ay + az = 1 \\ x + (a + b)y + z = a \\ bx + aby + bz = a^2 b \end{cases}$$

Como os coeficientes do sistema dependem dos parâmetros $a, b \in \mathbb{R}$, destes dependerá igualmente a natureza do sistema. Analisemos a situação:

$$\left[\begin{array}{ccc|c} 1 & a & a & 1 \\ 1 & a+b & 1 & a \\ b & ab & b & a^2 b \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & a & a & 1 \\ 0 & b & 1-a & a-1 \\ 0 & 0 & (1-a)b & (a^2-1)b \end{array}\right]$$

O seguinte esquema resume os casos possíveis e mostra que o sistema nunca é impossível:

$$\text{Sistema} \begin{cases} b = 0 \begin{cases} a = 1 \Rightarrow s = r = 1 \wedge n = 3 \Rightarrow \text{sistema } \textit{duplamente indeterminado.} \\ a \neq 1 \Rightarrow s = r = 2 \wedge n = 3 \Rightarrow \text{sistema } \textit{simplesmente indeterminado.} \end{cases} \\ b \neq 0 \begin{cases} a = 1 \Rightarrow s = r = 2 \wedge n = 3 \Rightarrow \text{sistema } \textit{simplesmente indeterminado.} \\ a \neq 1 \Rightarrow s = r = 3 \wedge n = 3 \Rightarrow \text{sistema } \textit{determinado.} \end{cases} \end{cases}$$

Vamos resolver o sistema nos quatro casos anteriores:

- Para $a = 1 \wedge b = 0$, vem, tomando x para incógnita principal,

$$\begin{cases} x = 1 - y - z \\ y = y \\ z = z \end{cases}, \text{ em que } y, z \in \mathbb{R}$$

Isto permite obter a *solução geral* na forma vectorial:

$$(x, y, z) = (1 - y - z, y, z) = (1, 0, 0) + y(-1, 1, 0) + z(-1, 0, 1), \text{ em que } y, z \in \mathbb{R}$$

$(1, 0, 0)$ é uma *solução particular* do sistema dado (é a que se obtém, fazendo $y = z = 0$) e $y(-1, 1, 0) + z(-1, 0, 1)$ é a *solução geral* do sistema homogéneo associado ao sistema inicial, a qual constitui o *subespaço* de $\mathbb{R}^3$ gerado por $(-1, 1, 0)$ e $(-1, 0, 1)$ os quais formam uma *base* deste subespaço, que terá dimensão 2 (igual ao grau de indeterminação).

- Para $a \neq 1 \wedge b = 0$, eliminando a terceira linha (nula) e trocando entre si a segunda e terceira coluna (ordem das incógnitas: x, z, y), vem

$$\begin{bmatrix} 1 & a & a & 1 \\ 0 & 0 & 1-a & a-1 \end{bmatrix} \to \begin{bmatrix} 1 & a & a & 1 \\ 0 & 1-a & 0 & a-1 \end{bmatrix} \to$$
$$\to \begin{bmatrix} 1 & a & a & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix} \to \left[\begin{array}{cc|c|c} 1 & 0 & a & a+1 \\ 0 & 1 & 0 & -1 \end{array}\right]$$

As variáveis principais são x e z e a solução geral é

$$\begin{cases} x = a + 1 - ay \\ y = y \\ z = -1 \end{cases} \text{, em que } y \in \mathbb{R}$$

A solução geral na forma vectorial é

$$(x, y, z) = (a + 1 - ay, y, -1) = (a + 1, 0, -1) + y(-a, 1, 0)\text{, em que } y \in \mathbb{R}$$

$(a + 1, 0, -1)$ é uma *solução particular* do sistema dado (é a que se obtém, fazendo $y = 0$) e $y(-a, 1, 0)$ é a *solução geral* do sistema homogéneo associado ao sistema dado, a qual constitui o *subespaço* de $\mathbb{R}^3$ gerado por $(-a, 1, 0)$ sendo este vector uma *base* deste subespaço, que terá dimensão 1 (igual ao grau de indeterminação).

- Para $a = 1 \wedge b \neq 0$, eliminando a última linha (nula), vem

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & b & 0 & 0 \end{bmatrix} \to \begin{bmatrix} -b & 0 & -b & -b \\ 0 & b & 0 & 0 \end{bmatrix} \to \left[\begin{array}{cc|c|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{array}\right]$$

As variáveis principais são x e y e a solução geral é

$$\begin{cases} x = 1 - z \\ y = 0 \\ z = z \end{cases} \text{, em que } z \in \mathbb{R}$$

A solução geral na forma vectorial é

$$(x, y, z) = (1 - z, 0, z) = (1, 0, 0) + z(-1, 0, 1)\text{, em que } z \in \mathbb{R}$$

$(1, 0, 0)$ é uma *solução particular* do sistema dado (é a que se obtém, fazendo $z = 0$) e $z(-1, 0, 1)$ é a *solução geral* do sistema homogéneo associado ao sistema dado, a qual constitui o *subespaço* de $\mathbb{R}^3$ gerado por $(-1, 0, 1)$ sendo este vector uma *base* deste subespaço, que terá dimensão 1 (igual ao grau de indeterminação).

- Finalmente, para $a \neq 1 \wedge b \neq 0$, vem sucessivamente,

$$\begin{bmatrix} 1 & a & a & 1 \\ 0 & b & 1-a & a-1 \\ 0 & 0 & (1-a)b & (a^2-1)b \end{bmatrix} \to \begin{bmatrix} 1 & a & a & 1 \\ 0 & b & 1-a & a-1 \\ 0 & 0 & 1 & -(a+1) \end{bmatrix} \to$$
$$\begin{bmatrix} 1 & a & 0 & a(a+1)+1 \\ 0 & b & 0 & a(1-a) \\ 0 & 0 & 1 & -(a+1) \end{bmatrix} \to \begin{bmatrix} b & 0 & 0 & b(a(a+1)+1)-a^2(1-a) \\ 0 & b & 0 & a(1-a) \\ 0 & 0 & 1 & -(a+1) \end{bmatrix} \to$$
$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{b}(a^3+a^2(b-1)+ab+b) \\ 0 & 1 & 0 & \frac{a}{b}(1-a) \\ 0 & 0 & 1 & -(a+1) \end{array}\right]$$

Pelo que a única solução é

$$\begin{cases} x = \frac{1}{b}(a^3+a^2(b-1)+ab+b) \\ y = \frac{a}{b}(1-a) \\ z = -(a+1) \end{cases}$$

Exemplo2.33 Vejamos, agora, o sistema real de 3 equações e 3 incógnitas seguinte, em que a e b são parâmetros reais

$$\begin{cases} ax+y+az & = 2b+bz \\ 2x-y-az & = a-2b+ax-2bz \\ ax+y+(a-b)z & = a+3b+(b^2-a^2)z \end{cases}$$

Primeiro, há que levar o sistema nas incógnitas reais (x, y, z) à forma canónica:

$$\begin{cases} ax+y+(a-b)z & = 2b \\ (2-a)x-y+(2b-a)z & = a-2b \\ ax+y+(a-b)(1+a+b)z & = a+3b \end{cases}$$

Condensemos a matriz ampliada $\left[A \middle| B\right]$ do sistema, por meio de operações elementares e começando por trocar as 2 primeiras colunas, o que significa que a ordem das incógnitas passou a ser (y, x, z):

$$\begin{bmatrix} a & 1 & a-b & 2b \\ 2-a & -1 & 2b-a & a-2b \\ a & 1 & (a-b)(1+a+b) & a+3b \end{bmatrix} \xrightarrow{C_1 \leftrightarrow C_2} \begin{bmatrix} 1 & a & a-b & 2b \\ -1 & 2-a & 2b-a & a-2b \\ 1 & a & (a-b)(1+a+b) & a+3b \end{bmatrix}$$
$$\xrightarrow[L_3 \to L_3 - L_1]{L_2 \to L_2 + L_1} \begin{bmatrix} 1 & a & a-b & 2b \\ 0 & 2 & b & a \\ 0 & 0 & (a-b)(a+b) & a+b \end{bmatrix}$$

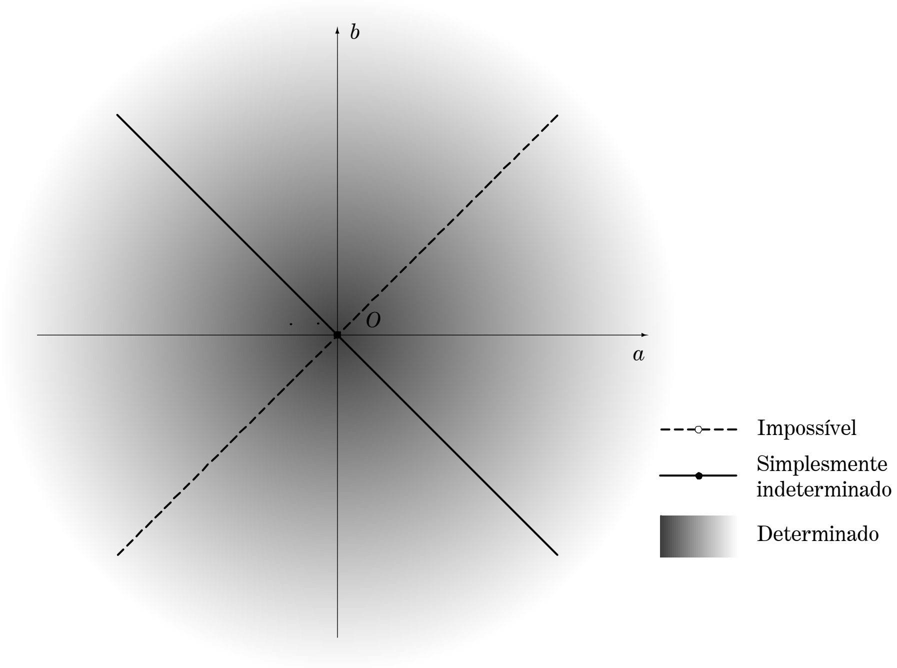


Fig. 2.4 – Discussão do sistema no plano dos parâmetros (a, b).

É, agora, possível discutir as soluções do sistema, em função dos valores reais dos parâmetros a e b e onde $r = \text{c}(A)$ e $s = \text{c}\big(\big[A \,|\, B\big]\big)$:

$$\text{Sistema}\begin{cases} a+b=0 \Rightarrow r=s=2 \Rightarrow \text{possível e } n-r=3-2=1 \text{ (simplesmente indet}^{\text{o}}\text{)} \\ a+b\neq 0 \Rightarrow s=3 \begin{cases} a-b=0 \Rightarrow r=2 \Rightarrow \text{impossível} \\ a-b\neq 0 \Rightarrow r=3 \Rightarrow \text{possível e } n-r=0 \text{ (determinado)} \end{cases} \end{cases}$$

Considerando o espaço $\mathbb{R}^2$ dos parâmetros (a, b), podemos ilustrar toda a discussão anterior na figura 2.4.

Vamos, agora, resolver o sistema nos casos de possibilidade:

- Se for $a + b = 0$, a matriz completa do sistema será, após eliminar a 3ª linha (nula) e substituir b por $-a$:

$$\begin{bmatrix} 1 & a & 2a & -2a \\ 0 & 2 & -a & a \end{bmatrix} \xrightarrow{L_1 \to 2L_1 - aL_2} \begin{bmatrix} 2 & 0 & 4a+a^2 & -4a-a^2 \\ 0 & 2 & -a & a \end{bmatrix} \xrightarrow[L_2 \to \frac{1}{2}L_2]{L_1 \to \frac{1}{2}L_1} \begin{bmatrix} 1 & 0 & \frac{a(4+a)}{2} & -\frac{a(4+a)}{2} \\ 0 & 1 & -\frac{a}{2} & \frac{a}{2} \end{bmatrix}$$

Portanto, as soluções são, para quaisquer $a, z \in \mathbb{R}$, (relembre a ordem (y, x, z) das incógnitas!)

$$\begin{cases} x = \dfrac{a}{2} + \dfrac{a}{2}z \\ y = -\dfrac{a(4+a)}{2} - \dfrac{a(4+a)}{2}z \\ z = z \end{cases}$$

Vectorialmente, fica

$$\begin{aligned}(x,y,z) &= \left(\frac{a}{2}+\frac{a}{2}z, -\frac{a(4+a)}{2}-\frac{a(4+a)}{2}z, z\right)\\ &= \left(\frac{a}{2}, -\frac{a(4+a)}{2}, 0\right)+z\left(\frac{a}{2}, -\frac{a(4+a)}{2}, 1\right)\end{aligned}$$

O vector $\left(\frac{a}{2}, -\frac{a(4+a)}{2}, 0\right)$ é uma solução particular do sistema inicial (obtida quando $z=0$) e a parcela $z\left(\frac{a}{2}, -\frac{a(4+a)}{2}, 1\right)$ é a solução geral do sistema homogéneo associado àquele.

- Se for $a+b\neq 0 \wedge a-b\neq 0$, a matriz completa do sistema será, após dividir a 3ª linha por $(a-b)(a+b)$,

$$\begin{bmatrix}1 & a & a-b & 2b\\ 0 & 2 & b & a\\ 0 & 0 & 1 & \frac{1}{a-b}\end{bmatrix} \xrightarrow[L_1\to L_1+(b-a)L_3]{L_2\to L_2-bL_3} \begin{bmatrix}1 & a & 0 & 2b-1\\ 0 & 2 & 0 & a-\frac{b}{a-b}\\ 0 & 0 & 1 & \frac{1}{a-b}\end{bmatrix}$$
$$\xrightarrow[L_2\to\frac{1}{2}L_2]{L_1\to L_1-\frac{a}{2}L_2} \begin{bmatrix}1 & 0 & 0 & 2b-1-\frac{a}{2}\left(a-\frac{b}{a-b}\right)\\ 0 & 1 & 0 & \frac{a}{2}-\frac{b}{2(a-b)}\\ 0 & 0 & 1 & \frac{1}{a-b}\end{bmatrix}$$

A solução (única) do sistema é, em função de a e b (observe-se de novo a ordem (y,x,z) das incógnitas!):

$$(x,y,z) = \left(\frac{a}{2}-\frac{b}{2(a-b)}, 2b-1-\frac{a}{2}\left(a-\frac{b}{a-b}\right), \frac{1}{a-b}\right)$$

Exemplo 2.34 – Métodos de Jacobi e de Gauss-Seidel – Vamos agora mencionar dois métodos iterativos para resolver sistemas $\sum_{j=1}^{n} a_{ij}\,x_j = b_i$ $(i=1,2,\ldots,n)$ de n equações lineares e n incógnitas determinados ($s=r=n$) e ainda com $a_{ii}\neq 0$. Resolvendo cada equação em ordem a x_i, obtemos

$$x_i = \frac{1}{a_{ii}}\Big(b_i - \sum_{j\neq i} a_{ij}\,x_j\Big)$$

Definamos, por recorrência, uma sucessão de pontos $\vec{x}_k = (x_{k,1}, x_{k,2},\ldots,x_{k,n})\in\mathbb{K}^n$, do modo seguinte:

$$\begin{cases}\vec{x}_0 &= (x_{0,1}, x_{0,2},\ldots,x_{0,n})\in\mathbb{K}^n\\ x_{k,i} &= \frac{1}{a_{ii}}\Big(b_i-\sum\limits_{j\neq i} a_{ij}\,x_{k-1,j}\Big); k>0 \text{ e } i=1,2,\ldots,n\end{cases}$$

Prova-se que, se esta sucessão for convergente, ela converge para a solução do sistema. Este método é conhecido por *método de Jacobi*$^{(11)}$.

---

11 *Jacobi, Carl Gustav Jacob*: matemático alemão (Potsdam 1804 – Berlim 1851).

Uma variante do algoritmo anterior que, normalmente, converge mais rapidamente para a solução é o *método de Gauss-Seidel*[12], no qual se utilizam imediatamente as coordenadas do vector $\vec{x}_k$ assim que se determinam, isto é, põe-se

$$\begin{cases} \vec{x}_0 &= (x_{0,1}, x_{0,2}, \ldots, x_{0,n}) \in \mathbb{K}^n \\ x_{k,i} &= \frac{1}{a_{ii}}\Big(b_i - \sum\limits_{j<i} a_{ij}\, x_{k,j} - \sum\limits_{j>i} a_{ij}\, x_{k-1,j}\Big); k > 0 \text{ e } i = 1, 2, \ldots, n \end{cases}$$

Uma condição suficiente (mas não necessária) para que ambos os métodos convirjam é que, para cada i, os a_{ii} sejam superiores em módulo à soma dos módulos dos restantes a_{ij}, isto é,

$$\underset{1\le i\le n}{\forall} |a_{ii}| > \sum_{j\neq i} |a_{ij}|$$

A seguir apresenta-se a implementação do método de Gauss-Seidel no MATHEMATICA©, na qual a iteração termina quando $\|\vec{x}_k - \vec{x}_{k-1}\| < \varepsilon$ (ε especificado) ou quando se atinge um certo número especificado de pontos calculados. Nesta implementação, é devolvida a solução e o número de pontos calculados:

```
GaussSeidel[a_?MatrixQ, b_?VectorQ, x0_?VectorQ, ε_, maxiter_?IntegerQ] :=
  Module[{y0 = x0, y1 = x0, n = Length[x], i, j, cont = True, m = 0},
    While[cont,
      For[i = 1, i <= n, i++,
        y1[[i]] = 1/a[[i, i]]*(b[[i]] - Sum[a[[i, j]]*y1[[j]], {j, 1, i - 1}] -
        Sum[a[[i, j]]*y0[[j]], {j, i + 1, n}]);
        m++
      ];
      cont = Norm[y1 - y0] >= ε && m <= maxiter;
      y0 = y1
    ];
    {y1, m}
  ]
```

A função seguinte devolve a lista de pontos calculados (incluindo $\vec{x}_0$):

```
GaussSeidelPoints[a_?MatrixQ, b_?VectorQ, x0_?VectorQ, ε_, maxiter_?IntegerQ] :=
  Module[{y0 = x0, y1 = x0, n = Length[x], i, j, cont = True, lista, m = 0},
    lista = {y1};
    While[cont,
      For[i = 1, i <= n, i++,
        y1[[i]] = 1/a[[i, i]]*(b[[i]] - Sum[a[[i, j]]*y1[[j]], {j, 1, i - 1}] -
        Sum[a[[i, j]]*y0[[j]], {j, i + 1, n}]);
        lista = Join[lista, {y1}];
        m++
      ];
      cont = Norm[y1 - y0] >= ε && m <= maxiter;
      y0 = y1
    ];
    lista
  ]
```

Na secção 2.16, mostramos o uso que se pode fazer do software MATHEMATICA©, na resolução e discussão de sistemas de equações lineares, incluindo exemplos de uso das funções anteriores e da sua interpretação geométrica em $\mathbb{R}^2$ (sistemas de 2 equações e 2 incógnitas sobre o corpo $\mathbb{R}$).

---

12 *Seidel, Philipp Ludwig von*: matemático alemão (Zweibrücken 1821 – München 1896).

2.12 Inversão matricial

Como vimos anteriormente, a matriz identidade é elemento neutro para o produto matricial e, para matrizes rectangulares $m \times n$, existem elementos neutros distintos à esquerda e à direita, respectivamente I_m e I_n. Para analisarmos a questão da inversão de uma matriz rectangular A e porque o produto matricial não é comutativo, temos que considerar "inversas direitas" (quando a "inversa" é o 2º factor) e "inversas esquerdas" (quando a "inversa" constitui o 1º factor).

Definição 2.16 – Inversas laterais – *Dada uma matriz $A \in \mathbb{K}^{m,n}$ de tipo $m \times n$ sobre um corpo $\mathbb{K}$, chama-se **inversa direita** de A a qualquer matriz $X_d \in \mathbb{K}^{n,m}$ tal que*

$$AX_d = I_m \tag{2.59}$$

*e **inversa esquerda** de A a toda a matriz $X_e \in \mathbb{K}^{n,m}$ tal que*

$$X_eA = I_n \tag{2.60}$$

Observe-se que as matrizes X_d, X_e, I_m e I_n só poderiam ser dos tipos indicados, por ser $A \in \mathbb{K}^{m,n}$. Podemos ver de imediato que as matrizes X_d e X_e, se existirem simultaneamente, serão necessariamente iguais e únicas e que, nesse caso, a matriz A tem que ser quadrada: a igualdade daquelas duas matrizes resulta de ser

$$X_d = I_nX_d = (X_eA)X_d = X_e(AX_d) = X_eI_m = X_e \Rightarrow X_d = X_e$$

Por outro lado, se X_{d_1} e X_{d_2} são inversas direitas de A e X_e for uma inversa esquerda, procedendo como anteriormente, virá

$$X_{d_1} = X_e \wedge X_{d_2} = X_e \Rightarrow X_{d_1} = X_{d_2}$$

o que mostra a unicidade da inversa direita (caso exista também inversa esquerda). Do mesmo modo se pode provar a unicidade da inversa esquerda (caso exista inversa direita). Por outro lado, se ambas as inversas existem, das alíneas *iv)* e *v)* da proposição 2.8 resulta

$$\left.\begin{array}{l} \mathrm{c}(I_m) \leq \mathrm{c}(A) \leq m \Rightarrow m \leq \mathrm{c}(A) \leq m \Rightarrow \mathrm{c}(A) = m \\ \mathrm{c}(I_n) \leq \mathrm{c}(A) \leq n \Rightarrow n \leq \mathrm{c}(A) \leq n \Rightarrow \mathrm{c}(A) = n \end{array}\right\} \Rightarrow m = n$$

e, portanto, a matriz A será quadrada[(13)]; isto equivale a dizer que, se A for rectangular, seguramente não existirá pelo menos uma das inversas.

É ainda fácil reduzir o problema da inversa esquerda ao problema da inversa direita, por transposição; de facto

$$X_eA = I_n \Leftrightarrow (X_eA)^{\mathrm{T}} = I_n^{\mathrm{T}} \Leftrightarrow A^{\mathrm{T}}X_e^{\mathrm{T}} = I_n \tag{2.61}$$

e estas equivalências mostram que X_e é uma inversa esquerda de A sse X_e^{T} é uma inversa direita de A^{T} e, como $X_e = (X_e^{\mathrm{T}})^{\mathrm{T}}$, poderemos concluir que se obtém uma inversa *esquerda* de A, mediante a transposição de uma inversa *direita* de A^{T}. Graças a (2.61) reduz-se, portanto, a questão da inversa esquerda de A ao problema da inversa direita (de A^{T}).

---

13 Esta conclusão pode também ser tirada recorrendo às propriedades do *traço*: Sendo $X \in \mathbb{K}^{n,m}$ uma inversa (bilateral) de $A \in \mathbb{K}^{m,n}$, será $AX = I_m$ e $XA = I_n$. O resultado decorre de ser $m = \mathrm{tr}(AX) = \mathrm{tr}(XA) = n$ e, portanto, $m = n$.

Vamos apresentar, de seguida, a condição necessária e suficiente para que uma matriz $A \in \mathbb{K}^{m,n}$ tenha inversa direita e também a condição para existir inversa esquerda:

Proposição 2.12 – Existência de inversas laterais – *Seja $A \in \mathbb{K}^{m,n}$ uma matriz de tipo $m \times n$ sobre um corpo $\mathbb{K}$ e $r = \mathrm{c}(A)$. então:*

i) *A tem inversa(s) direita(s) sse $r = m$.*

ii) *A tem inversa(s) esquerda(s) sse $r = n$.*

Demonstração:

i) Atente-se em (2.59) e fragmentemos as matrizes X_d e I_m nas suas m colunas

$$X_d = \left[X_1 \,\middle|\, X_2 \,\middle|\, \cdots \,\middle|\, X_m\right] \quad I_m = \left[B_1 \,\middle|\, B_2 \,\middle|\, \cdots \,\middle|\, B_m\right]$$

Observe-se, desde já, que os vectores B_i de $\mathbb{K}^m$ são linearmente independentes e constituem a base canónica de $\mathbb{K}^m$, donde

$$L(B_1, B_2, \ldots, B_m) = \mathbb{K}^m$$

Fazendo o produto AX_d por blocos, verifica-se facilmente que a igualdade (2.59) equivale às m igualdades

$$AX_i = B_i, \text{ para } i = 1, 2, \ldots, m \tag{2.62}$$

e cada uma destas igualdades é a forma matricial de um sistema de m equações e n incógnitas, sendo que estes m sistemas têm em comum a matriz A dos coeficientes.

É, agora, óbvio que a matriz X_d satisfazendo (2.59) existe sse forem simultaneamente possíveis todos estes sistemas, ou seja, sse os vectores $B_i \in \mathbb{K}^m$ pertencerem ao subespaço de $\mathbb{K}^m$ gerado pelas colunas $(A_j)_{1 \le j \le n}$ da matriz A (ver demonstração da proposição 2.10), o que equivale evidentemente a

$$L(B_1, B_2, \ldots, B_m) \subset L(A_1, A_2, \ldots, A_n)$$

Porém tem-se, como se viu, $L(B_1, B_2, \ldots, B_m) = \mathbb{K}^m$ e a inclusão anterior transforma-se em

$$\mathbb{K}^m \subset L(A_1, A_2, \ldots, A_n)$$

o que, juntamente com a inclusão óbvia $L(A_1, A_2, \ldots, A_n) \subset \mathbb{K}^m$, acaba por dar a equivalência que pretendíamos demonstrar

$$\text{Existe inversa direita de } A \Leftrightarrow L(A_1, A_2, \ldots, A_n) = \mathbb{K}^m \Leftrightarrow r = m$$

ii) Resulta imediatamente do resultado anterior, por transposição,

$$\text{Existe inversa esquerda de } A \Leftrightarrow \text{Existe inversa direita de } A^{\mathrm{T}} \Leftrightarrow \mathrm{c}\big(A^{\mathrm{T}}\big) = n \Leftrightarrow \mathrm{c}(A) = r = n$$

□

Em face dos resultados provados na proposição anterior, podemos fazer a discussão do que se pode passar em relação à inversão de uma matriz $A \in \mathbb{K}^{m,n}$

$$A \in \mathbb{K}^{m,n}\begin{cases} m < n \Rightarrow r \leq m < n \Rightarrow r < n \;\Rightarrow \begin{cases}\text{Não existe inversa esquerda de } A\\ \text{e existirá inversa direita, sse } r = m.\end{cases}\\ n < m \Rightarrow r \leq n < m \Rightarrow r < m \Rightarrow \begin{cases}\text{Não existe inversa direita de } A\\ \text{e existirá inversa esquerda, sse } r = n.\end{cases}\\ m = n \qquad\qquad\qquad\qquad\qquad \Rightarrow \begin{cases} r = n \Rightarrow \begin{cases}\text{Existem ambas as inversas de } A\\ \text{e são iguais e únicas, como provámos.}\end{cases}\\ r < n \Rightarrow \text{Não existe qualquer das inversas de } A.\end{cases}\end{cases}$$

Vamos, de seguida, analisar cada um dos casos mencionados no quadro anterior:

- Se for $r < m < n$, então não existem inversas.

- Quando $r = m < n$ A não tem inversa esquerda, mas tem mais do que uma inversa direita, visto que os sistemas (2.62) são todos possíveis e indeterminados de grau $n - m$. Para calcularmos as matrizes inversas direitas de A, apenas teremos que resolver os m sistemas (2.62), todos com grau de indeterminação $n - m$, o que pode ser feito em simultâneo, visto terem todos a mesma matriz A de coeficientes; obtém-se, assim, o algoritmo

$$\left[A \mid B_1 \mid B_2 \mid \cdots \mid B_m\right] = \left[A \mid I_m\right] \rightarrow \left[I_m \mid C_{m,n-m} \mid \underset{m,m}{B}\right] \tag{2.63}$$

e daqui se deduzem as colunas X_i que formarão as matrizes inversas direitas

$$X_d = \left[X_1 \mid X_2 \mid \cdots \mid X_m\right]$$

Se o corpo $\mathbb{K}$ for infinito ($\mathbb{Q}, \mathbb{R}, \mathbb{C}$), haverá uma infinidade de inversas direitas; se $\mathbb{K}$ for finito, então existirão $n - m$ parâmetros arbitrários por cada sistema (as incógnitas secundárias desses sistemas) e, sendo estes em número de m, haverá $m(n - m)$ parâmetros; no total, o número de inversas direitas é de $(\#\mathbb{K})^{m(n-m)}$.

O exemplo seguinte ilustra o cálculo das inversas direitas:

Exemplo 2.35 A seguinte matriz A real não terá inversa esquerda ($3 = m < n = 5$).

$$A = \begin{bmatrix} 1 & -2 & 1 & 0 & 3\\ 0 & 1 & 3 & 4 & 2\\ -1 & 1 & -1 & -1 & -2 \end{bmatrix}$$

Para calcularmos as possíveis inversas direitas, condensemos a matriz

$$\left[\begin{array}{ccccc|ccc} 1 & -2 & 1 & 0 & 3 & 1 & 0 & 0\\ 0 & 1 & 3 & 4 & 2 & 0 & 1 & 0\\ -1 & 1 & -1 & -1 & -2 & 0 & 0 & 1 \end{array}\right] \rightarrow \begin{bmatrix} 1 & -2 & 1 & 0 & 3 & 1 & 0 & 0\\ 0 & 1 & 3 & 4 & 2 & 0 & 1 & 0\\ 0 & -1 & 0 & -1 & 1 & 1 & 0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 & 7 & 8 & 7 & 1 & 2 & 0\\ 0 & 1 & 3 & 4 & 2 & 0 & 1 & 0\\ 0 & 0 & 3 & 3 & 3 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 0 & 0 & 3 & 0 & -4 & -1 & -7\\ 0 & 1 & 0 & 1 & -1 & -1 & 0 & -1\\ 0 & 0 & 3 & 3 & 3 & 1 & 1 & 1 \end{bmatrix}$$

$$\rightarrow \left[\begin{array}{ccc|cc|ccc} 1 & 0 & 0 & 1 & 0 & -\frac{4}{3} & -\frac{1}{3} & -\frac{7}{3} \\ 0 & 1 & 0 & 1 & -1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}\right]$$

A característica de A é $r = 3 = m$ e, portanto, existem inversas direitas X_d, cujas colunas X_1, X_2 e X_3 serão as soluções gerais dos 3 sistemas, de 3 equações e 5 incógnitas, duplamente indeterminados que acabámos de resolver em cima. As soluções gerais são

$$X_1 = \begin{bmatrix} -\frac{4}{3} - c_{12} \\ -1 + c_{11} - c_{12} \\ \frac{1}{3} - c_{11} - c_{12} \\ c_{12} \\ c_{11} \end{bmatrix}$$

$$X_2 = \begin{bmatrix} -\frac{1}{3} - c_{22} \\ c_{21} - c_{22} \\ \frac{1}{3} - c_{21} - c_{22} \\ c_{22} \\ c_{21} \end{bmatrix}$$

$$X_3 = \begin{bmatrix} -\frac{7}{3} - c_{32} \\ -1 + c_{31} - c_{32} \\ \frac{1}{3} - c_{31} - c_{32} \\ c_{32} \\ c_{31} \end{bmatrix}$$

onde os c_{ij} são 6 parâmetros reais arbitrários.

Existe, pois, uma infinidade de inversas direitas de A, com expressão geral

$$X_d = \begin{bmatrix} -\frac{4}{3} - c_{12} & -\frac{1}{3} - c_{22} & -\frac{7}{3} - c_{32} \\ -1 + c_{11} - c_{12} & c_{21} - c_{22} & -1 + c_{31} - c_{32} \\ \frac{1}{3} - c_{11} - c_{12} & \frac{1}{3} - c_{21} - c_{22} & \frac{1}{3} - c_{31} - c_{32} \\ c_{12} & c_{22} & c_{32} \\ c_{11} & c_{21} & c_{31} \end{bmatrix}, \text{ onde } c_{ij} \in \mathbb{R}$$

Pode verificar-se o resultado:

$$AX_d =$$

$$= \begin{bmatrix} 1 & -2 & 1 & 0 & 3 \\ 0 & 1 & 3 & 4 & 2 \\ -1 & 1 & -1 & -1 & -2 \end{bmatrix} \begin{bmatrix} -\frac{4}{3} - c_{12} & -\frac{1}{3} - c_{22} & -\frac{7}{3} - c_{32} \\ -1 + c_{11} - c_{12} & c_{21} - c_{22} & -1 + c_{31} - c_{32} \\ \frac{1}{3} - c_{11} - c_{12} & \frac{1}{3} - c_{21} - c_{22} & \frac{1}{3} - c_{31} - c_{32} \\ c_{12} & c_{22} & c_{32} \\ c_{11} & c_{21} & c_{31} \end{bmatrix}$$

$$= I_3$$

- Se for $r < n < m$, não existem inversas.

- Quando $r = n < m$, A não tem inversa direita, mas terá várias inversas esquerdas. Para calcular as inversas esquerdas de A, aplica-se (2.63) a A^{T} e, após se obter as inversas direitas de A^{T}, transpõem-se estas. Neste caso, os n sistemas (2.61) vão ser todos indeterminados de grau $m - n$.

Se o corpo $\mathbb{K}$ for infinito (casos de $\mathbb{Q}, \mathbb{R}, \mathbb{C}$), haverá uma infinidade de inversas esquerdas; se $\mathbb{K}$ for finito, então existirão $m - n$ parâmetros arbitrários por cada sistema (as incógnitas secundárias desses sistemas) e, sendo n o número de sistemas, haverá $n(m-n)$ parâmetros pertencentes a $\mathbb{K}$ e, ao todo, existirão $(\#\mathbb{K})^{n(m-n)}$ inversas esquerdas. Ilustra-se, no seguinte exemplo, o cálculo das inversas esquerdas:

Exemplo 2.36 Considere-se a matriz A seguinte

$$A = \begin{bmatrix} 1 & 0 & -1 \\ -2 & 1 & 1 \\ 1 & 3 & -1 \\ 0 & 4 & -1 \\ 3 & 2 & -2 \end{bmatrix}$$

a qual não tem inversas direitas, visto que se tem $3 = n < m = 5.$ Calculemos as suas inversas esquerdas, se existirem: para isso, determinamos as inversas direitas de A^{T}, o que já fizemos no exemplo anterior; agora, para se obter as inversas esquerdas procuradas, basta transpor o resultado obtido nesse exemplo:

$$X_e = \begin{bmatrix} -\frac{4}{3} - c_{12} & -1 + c_{11} - c_{12} & \frac{1}{3} - c_{11} - c_{12} & c_{12} & c_{11} \\ -\frac{1}{3} - c_{22} & c_{21} - c_{22} & \frac{1}{3} - c_{21} - c_{22} & c_{22} & c_{21} \\ -\frac{7}{3} - c_{32} & -1 + c_{31} - c_{32} & \frac{1}{3} - c_{31} - c_{32} & c_{32} & c_{31} \end{bmatrix}, \text{ onde } c_{ij} \in \mathbb{R}$$

Segue-se a verificação do resultado

$$X_e A = \begin{bmatrix} -\frac{4}{3} - c_{12} & -1 + c_{11} - c_{12} & \frac{1}{3} - c_{11} - c_{12} & c_{12} & c_{11} \\ -\frac{1}{3} - c_{22} & c_{21} - c_{22} & \frac{1}{3} - c_{21} - c_{22} & c_{22} & c_{21} \\ -\frac{7}{3} - c_{32} & -1 + c_{31} - c_{32} & \frac{1}{3} - c_{31} - c_{32} & c_{32} & c_{31} \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ -2 & 1 & 1 \\ 1 & 3 & -1 \\ 0 & 4 & -1 \\ 3 & 2 & -2 \end{bmatrix} = I_3$$

- Se for $r < m = n$, não existirá inversa e diz-se que a matriz quadrada A é *não invertível*.

- Para $r = m = n$, existem os dois tipos de inversas *em simultâneo*, sendo então estas iguais e únicas, isto é, $X_d = X_e$ e a matriz quadrada A diz-se *invertível*, sendo a inversa (bilateral) designada por $A^{-1} = X_d = X_e$. Portanto, tem-se

$$A \in \mathbb{K}^{n,n} \text{ é } \textit{invertível} \Leftrightarrow \mathrm{c}(A) = n \tag{2.64}$$

e podemos ainda escrever

$$A \in \mathbb{K}^{n,n} \text{ é } \textit{invertível} \Leftrightarrow \underset{A^{-1} \in \mathbb{K}^{n,n}}{\exists} AA^{-1} = A^{-1}A = I_n \tag{2.65}$$

As igualdades anteriores mostram que A e A^{-1} são sempre *permutáveis* e o que acabamos de ver significa que, no *anel* não comutativo com identidade $(\mathbb{K}^{n,n}, +, \times)$ das matrizes quadradas de ordem n, os elementos *invertíveis* são as matrizes cuja característica coincide com a ordem (matrizes ditas *regulares*) e os elementos *não invertíveis* são as matrizes com caraterística inferior à ordem (matrizes ditas *singulares*, sendo que por entre estas está, é claro, a matriz nula O_n, com $r = 0 < n$). Há, pois, elementos não invertíveis e não nulos, (por exemplo, em $\mathbb{R}^{2,2}$, a matriz $\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$ não é invertível, com $r = 1 < 2$). Isto significa que, para $n > 1$, $(\mathbb{K}^{n,n}, +, \times)$ não é um *corpo*, porque o produto matricial não é comutativo e ainda porque existem matrizes não nulas e não invertíveis. No anel $(\mathbb{K}^{n,n}, +, \times)$ os termos *invertível* (existência de inversa) e *regular* (característica igual à ordem) são, pois, sinónimos, o mesmo acontecendo com *não invertível* e *singular*.

Agora, os n sistemas (2.61) são todos possíveis e determinados (o que implica a existência unicidade da inversa bilateral) e (2.63) simplifica-se, obtendo-se o algoritmo de condensação seguinte

$$\left[A \mid B_1 \mid B_2 \mid \cdots \mid B_n\right] \to \left[A \mid X_1 \mid X_2 \mid \cdots \mid X_n\right]$$

ou seja,

$$\left[A \,|\, I_n\right] \to \left[I_n \,|\, A^{-1}\right] \tag{2.66}$$

Exemplo 2.37 Vamos usar o algoritmo anterior para inverter a matriz quadrada

$$A = \begin{bmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{bmatrix}$$

Condensemos, então, a matriz $\left[A \,|\, I_3\right]$:

$$\left[\begin{array}{ccc|ccc} -1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 4 & -2 & 5 & 0 & 0 & 1 \end{array}\right] \to \begin{bmatrix} -1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 5 & -6 & 2 & 1 & 0 \\ 0 & 6 & -7 & 4 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 5 & 0 & 3 & -1 & 2 & 0 \\ 0 & 5 & -6 & 2 & 1 & 0 \\ 0 & 0 & 1 & 8 & -6 & 5 \end{bmatrix} \to$$

$$\begin{bmatrix} 5 & 0 & 0 & -25 & 20 & -15 \\ 0 & 5 & 0 & 50 & -35 & 30 \\ 0 & 0 & 1 & 8 & -6 & 5 \end{bmatrix} \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -5 & 4 & -3 \\ 0 & 1 & 0 & 10 & -7 & 6 \\ 0 & 0 & 1 & 8 & -6 & 5 \end{array}\right]$$

Daqui resulta que a inversa é:

$$A^{-1} = \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix}$$

Como verificação, podem calcular-se os produtos

$$AA^{-1} = \begin{bmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{bmatrix} \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

$$A^{-1}A = \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix} \begin{bmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

No *package* **`ALGA`Matrizes``**, implementado em linguagem MATHEMATICA© e descrito na secção 2.16, definem-se as funções

`InversaDireita[A,c]` e **`InversaEsquerda[A,c]`**

que calculam as inversas laterais de `A` tendo como parâmetros os **`c[i]`**. Naquela secção, apresentam-se exemplos de uso dessas funções.

No referido *package*, figura também a implementação do algoritmo (2.66), constituindo a função **`Inversa[A]`**; trata-se aqui de mero exercício demonstrativo, pois o software já possui uma função para esse efeito chamada **`Inverse[A]`**.

A seguinte proposição dá conta de algumas das propriedades da inversão, para matrizes quadradas

Proposição 2.13 – Propriedades da inversa – *Sejam $A, B \in \mathbb{K}^{n,n}$ matrizes quadradas, $\alpha \in \mathbb{K}$ e $k \in \mathbb{N}_0$, então:*

i) *A é invertível sse A^{-1} é invertível e tem-se*

$$\left(A^{-1}\right)^{-1} = A$$

ii) *A é invertível sse A^{T} é invertível e*

$$\left(A^{\mathrm{T}}\right)^{-1} = \left(A^{-1}\right)^{\mathrm{T}}$$

iii) *Se $\alpha \neq 0$ (ou seja, se α é invertível no corpo $\mathbb{K}$) e A é invertível, então αA é invertível e tem-se*

$$(\alpha A)^{-1} = \alpha^{-1} A^{-1} = \frac{1}{\alpha}\, A^{-1}$$

iv) *A é invertível sse o sistema homogéneo $AX = O$ é determinado, tendo apenas a solução trivial $X = O$.*

v) *A e B são invertíveis sse AB é invertível e*

$$(AB)^{-1} = B^{-1} A^{-1}$$

Mais geralmente, para todo o $p \in \mathbb{N}$, as matrizes de uma sequência $(A_k)_{1 \leq k \leq p}$ são invertíveis sse o produto $\prod_{k=1}^{p} A_k$ o for e tem-se

$$(A_1 A_2 \cdots A_p)^{-1} = A_p^{-1} \cdots A_2^{-1} A_1^{-1} \tag{2.67}$$

vi) *A é invertível sse A^k é invertível e tem-se*

$$\left(A^k\right)^{-1} = \left(A^{-1}\right)^k$$

vii) *Se A é invertível, então*

$$A^{-1}B = BA^{-1} \Leftrightarrow AB = BA$$

isto é, A^{-1} e B permutam sse A e B o fazem.

Demonstração:

i) É consequência de ser $AA^{-1} = A^{-1}A = I_n$.

ii) Resulta de $A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = (A^{-1}A)^{\mathrm{T}} = I_n^{\mathrm{T}} = I_n$.

iii) É consequência imediata de ser $(\alpha A)\big(\frac{1}{\alpha}\,A^{-1}\big) = \big(\alpha\frac{1}{\alpha}\big)(AA^{-1}) = 1I_n = I_n$.

iv) Resulta das equivalências

$$A \text{ é invertível } \Leftrightarrow r = \mathrm{c}(A) = n \Leftrightarrow AX = O \text{ é sistema determinado}$$

v) Suponha-se que A e B são invertíveis; então existem A^{-1} e B^{-1} e as igualdades

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= ((AB)B^{-1})A^{-1} = (A(BB^{-1}))A^{-1}\\ &= (AI_n)A^{-1} = AA^{-1} = I_n\end{aligned}$$

mostram que $B^{-1}A^{-1}$ é a inversa (necessariamente única) de AB e que, portanto, AB é invertível.

Provemos, agora, a implicação contrária: se uma das matrizes A ou B não é invertível, será

$$\min(\mathrm{c}(A), \mathrm{c}(B)) < n$$

e a proposição 2.8.*v* implica que

$$\mathrm{c}(AB) \leq \min(\mathrm{c}(A), \mathrm{c}(B)) < n \Rightarrow \mathrm{c}(AB) < n$$

o que mostra que AB não é invertível.

Quanto à segunda afirmação, ela demonstra-se facilmente por indução.

vi) Por indução em k: para $k = 0$, fica

$$(A^0)^{-1} = I_n^{-1} = I_n = (A^{-1})^0$$

Admitindo que, para $k \geq 0$, é $\big(A^k\big)^{-1} = (A^{-1})^k$, então, para $k+1$ vem

$$\big(A^{k+1}\big)^{-1} = \big(A^kA\big)^{-1} = A^{-1}\big(A^k\big)^{-1} = (A^{-1})(A^{-1})^k = (A^{-1})^k(A^{-1}) = (A^{-1})^{k+1}$$

vii) Multiplicando ambos os membros de $AB = BA$ à esquerda e à direita por A^{-1}, obtém-se

$$AB = BA \Rightarrow A^{-1}(AB)A^{-1} = A^{-1}(BA)A^{-1} \Rightarrow BA^{-1} = A^{-1}B$$

Por outro lado, multiplicando ambos os membros de $A^{-1}B = BA^{-1}$ à esquerda e à direita por A, obtém-se

$$A^{-1}B = BA^{-1} \Rightarrow A(A^{-1}B)A = A(BA^{-1})A \Rightarrow BA = AB$$ □

Observações:

- Sejam $A \in \mathbb{K}^{m,m}$ e $B \in \mathbb{K}^{n,n}$ matrizes sobre um corpo $\mathbb{K}$ com ordens m e n respectivamente e invertíveis (isto é, c$(A) = m$ e c$(B) = n$). Então, para qualquer matriz $C \in \mathbb{K}^{m,n}$, a matriz de blocos

$$U = \begin{bmatrix} A & C \\ O_{n,m} & B \end{bmatrix} \in \mathbb{K}^{m+n,m+n}$$

é também invertível e tem-se:

$$U^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}CB^{-1} \\ O_{n,m} & B^{-1} \end{bmatrix}$$

, o que pode ser facilmente verificado multiplicando U e U^{-1} por blocos.

Do mesmo modo, para qualquer matriz $D \in \mathbb{K}^{n,m}$, a matriz de blocos

$$V = \begin{bmatrix} A & O_{m,n} \\ D & B \end{bmatrix} \in \mathbb{K}^{m+n,m+n}$$

é igualmente invertível, sendo:

$$V^{-1} = \begin{bmatrix} A^{-1} & O_{m,n} \\ -B^{-1}DA^{-1} & B^{-1} \end{bmatrix}$$

, o que pode também ser facilmente verificado multiplicando V e V^{-1} por blocos.

Podemos, agora, usar a alínea *vi)* da proposição anterior para definir a potência de expoente k inteiro negativo de uma matriz quadrada A regular, pondo

$$A^{-k} = \left(A^k\right)^{-1} = \left(A^{-1}\right)^k, \text{ com } k \in \mathbb{N} \tag{2.68}$$

Deste modo e para matrizes A regulares, faz sentido a potência A^k, com expoente $k \in \mathbb{Z}$.

As inversas podem ser utilizadas para resolver sistemas de equações lineares, podendo ter-se em consideração os 3 casos seguintes, onde $A \in \mathbb{K}^{m,n}$:

- Matriz A com característica $r = m < n$.

Neste caso, existem inversas direitas X_d e qualquer sistema $AX = B$ é possível, sendo $X = X_dB$ solução do sistema, qualquer que seja a inversa direita X_d usada, porque

$$AX = A(X_dB) = (AX_d)B = I_mB = B$$

- Matriz A com característica $r = n < m$.

Neste caso, existem inversas esquerdas X_e e, se o sistema $AX = B$ for possível, ele será determinado; podemos, então, multiplicar ambos os membros de $AX = B$ à esquerda por X_e, obtendo-se a solução X_eB para o sistema, a qual não dependerá da inversa esquerda X_e utilizada:

$$AX = B \Rightarrow X_e(AX) = X_eB \Rightarrow (X_eA)X = X_eB \Rightarrow I_nX = X_eB \Rightarrow X = X_eB$$

- Matriz A quadrada com característica $r = m = n$.

É um caso particular dos anteriores, obtendo-se a solução $A^{-1}B$ para o sistema $AX = B$:

$$AX = B \Rightarrow A^{-1}(AX) = A^{-1}B \Rightarrow \big(A^{-1}A\big)X = A^{-1}B \Rightarrow I_nX = A^{-1}B \Rightarrow X = A^{-1}B$$

Observe-se que os métodos anteriores para resolver sistemas não têm grande interesse prático, visto que é maior o trabalho de condensação para inverter A do que o correspondente trabalho para resolver $AX = B$ pelo algoritmo de Gauss-Jordan.

2.13 Matrizes elementares

Definição 2.17 – Matriz elementar – *Chama-se matriz elementar de ordem n a uma matriz obtida de I_n por uma única operação elementar sobre linhas ou sobre colunas; essa matriz elementar diz-se associada à operação elementar que lhe deu origem.*

Observações:

- Da definição anterior resulta a existência de 3 tipos de matrizes elementares, correspondentes aos 3 tipos de operações elementares:

$$E_1 = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix},\ E_2 = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix},\ E_3 = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & \beta & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

- Trocando entre si as linhas i e j de I_n ou as colunas com os mesmos índices obtém-se a mesma matriz elementar E_1 (operação de tipo 1). Esta matriz é simétrica e, portanto, igual à transposta.

- Multiplicando uma linha i de I_n por um escalar $\alpha \neq 0$ obtém-se a mesma matriz elementar E_2 obtida ao multiplicar a coluna i por α. Esta matriz é simétrica e, portanto, igual à transposta.

- As matrizes elementares obtidas pela mesma operação elementar de tipo 3 sobre linhas e sobre colunas de I_n são transpostas.

- Em face das observações anteriores, podemos afirmar que a matriz elementar obtida de I_n por uma operação elementar (de qualquer dos 3 tipos) sobre colunas é a transposta da matriz elementar obtida pela mesma operação elementar sobre linhas de I_n.

Seja, agora, $A \in \mathbb{K}^{m,n}$ uma matriz de tipo $m \times n$ sobre um corpo $\mathbb{K}$ e $(L_1, L_2, \ldots, L_m)$ as suas linhas. As linhas de I_m constituem a base canónica $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m)$ de $\mathbb{K}^m$ e facilmente se reconhece que multiplicar à esquerda o vector-linha $\vec{e}_i$ por A resulta na linha i de A, para todo o

$i = 1, 2, \ldots, m$.

$$\vec{e}_i A = [\,0\ \ 0\ \cdots\ 1\ \cdots\ 0\,] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [\,a_{i1}\ \ a_{i2}\ \cdots\ a_{in}\,] = L_i$$

Do mesmo modo, sejam $(C_1, C_2, \ldots, C_n)$ as colunas da mesma matriz $A \in \mathbb{K}^{m,n}$. As colunas de I_n constituem a base canónica $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de $\mathbb{K}^n$ e é fácil reconhecer que a multiplicação à direita do vector-coluna $\vec{e}_j$ por A resulta na coluna j de A, para todo o $j = 1, 2, \ldots, n$.

$$A\vec{e}_j = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = C_j$$

Com os resultados anteriores, podemos agora mostrar que a matriz $e(A)$ resultante da execução de uma qualquer operação elementar e sobre as linhas de A é igual ao produto à esquerda da matriz elementar associada àquela operação elementar por $A \in \mathbb{K}^{m,n}$, como se prova na seguinte

Proposição 2.14 *Seja $A \in \mathbb{K}^{m,n}$ uma matriz de tipo $m \times n$ sobre um corpo $\mathbb{K}$ e e uma operação elementar (de qualquer dos 3 tipos) sobre linhas de A. Então:*

$$e(A) = \underbrace{e(I_m)}_{\substack{\text{matriz} \\ \text{elementar } E}} A = EA$$

em que E é a matriz elementar obtida de I_m pela mesma operação elementar e sobre linhas.

Demonstração:

Vamos ver separadamente os 3 tipos de operações elementares: para isso, designemos por $e(A)$ a matriz resultante de aplicar uma operação elementar e à matriz A e por $E = e(I)$ um dos 3 tipos de matriz elementar:

- Operação de tipo 1: sejam $1 \leq i < j \leq m$ e E a matriz elementar resultante de trocar as linhas i e j de I_m; fazendo o produto EA por blocos, vem:

$$EA = e(I_m)A = \begin{bmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_j \\ \vdots \\ \vec{e}_i \\ \vdots \\ \vec{e}_m \end{bmatrix} A = \begin{bmatrix} \vec{e}_1 A \\ \vdots \\ \vec{e}_j A \\ \vdots \\ \vec{e}_i A \\ \vdots \\ \vec{e}_m A \end{bmatrix} = \begin{bmatrix} L_1 \\ \vdots \\ L_j \\ \vdots \\ L_i \\ \vdots \\ L_m \end{bmatrix} = e(A)$$

- Operação de tipo 2: seja $1 \leq i \leq m$ e E a matriz elementar resultante de multiplicar a linha i de I_m pelo escalar $\alpha \neq 0$; calculando o produto EA por blocos, vem:

$$EA = e(I_m)A = \begin{bmatrix} \vec{e}_1 \\ \vdots \\ \alpha\vec{e}_i \\ \vdots \\ \vec{e}_m \end{bmatrix} A = \begin{bmatrix} \vec{e}_1 A \\ \vdots \\ \alpha(\vec{e}_i A) \\ \vdots \\ \vec{e}_m A \end{bmatrix} = \begin{bmatrix} L_1 \\ \vdots \\ \alpha L_i \\ \vdots \\ L_m \end{bmatrix} = e(A)$$

- Operação de tipo 3: sejam $1 \leq i < j \leq m$ e E a matriz elementar resultante de somar à linha i o produto de $\beta \in \mathbb{K}$ pela linha j de I_m; fazendo o produto EA por blocos, vem:

$$EA = e(I_m)A = \begin{bmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_i + \beta\vec{e}_j \\ \vdots \\ \vec{e}_j \\ \vdots \\ \vec{e}_m \end{bmatrix} A = \begin{bmatrix} \vec{e}_1 A \\ \vdots \\ (\vec{e}_i + \beta\vec{e}_j)A \\ \vdots \\ \vec{e}_j A \\ \vdots \\ \vec{e}_m A \end{bmatrix} = \begin{bmatrix} \vec{e}_1 A \\ \vdots \\ \vec{e}_i A + \beta(\vec{e}_j A) \\ \vdots \\ \vec{e}_j A \\ \vdots \\ \vec{e}_m A \end{bmatrix} = \begin{bmatrix} L_1 \\ \vdots \\ L_i + \beta L_j \\ \vdots \\ L_j \\ \vdots \\ L_m \end{bmatrix} = e(A)$$

Portanto, executar uma operação elementar e sobre linhas de A é multiplicar à esquerda a matriz A pela matriz elementar associada àquela operação e. □

Resultado análogo vale também para as operações sobre colunas de A, mas o produto pela matriz elementar deve agora ser efectuado à direita:

Proposição 2.15 *Seja $A \in \mathbb{K}^{m,n}$ uma matriz de tipo $m \times n$ sobre um corpo $\mathbb{K}$ e e uma operação elementar (de qualquer dos 3 tipos) sobre colunas de A. Então:*

$$e(A) = A\,\underbrace{e(I_n)}_{\substack{\text{matriz} \\ \text{elementar } E}} = AE$$

em que E é a matriz elementar obtida de I_n pela mesma operação elementar e sobre colunas.

Demonstração:

A demonstração é análoga à anterior e é deixada ao cuidado do leitor. □

Da definição resulta imediatamente que todas as matrizes elementares são *regulares*, visto que elas resultam da execução de uma operação elementar sobre uma matriz regular (a identidade). Na secção anterior, vimos que uma matriz quadrada A de ordem n é regular sse a sua característica é igual a n: por outras palavras, $A \in \mathbb{K}^{n,n}$ é regular sse A pode ser transformada na matriz identidade I_n por meio de um número finito de operações elementares sobre linhas. Pelo que acima foi dito, esta condição equivale à existência de uma sequência de matrizes elementares $(E_1, E_2, \ldots, E_p)$ de ordem n tais que

$$E_p \cdots E_2 E_1 A = I_n$$

A igualdade anterior mostra que o produto de matrizes elementares $E_p \cdots E_2 E_1$ é a matriz inversa de A,

$$A^{-1} = E_p \cdots E_2 E_1 \tag{2.69}$$

e invertendo ambos os membros e atendendo a (2.67), obtém-se:

$$A = E_1^{-1} E_2^{-1} \ldots E_p^{-1}$$

Podemos, portanto, enunciar a seguinte

Proposição 2.16 *Uma matriz quadrada é regular sse ela é produto de matrizes elementares.*

Se atendermos a que I_n é elemento neutro do produto matricial, a igualdade (2.69) pode escrever-se:

$$A^{-1} = E_p \cdots E_2 E_1 = E_p \cdots E_2 E_1 I_n \tag{2.70}$$

Mas esta igualdade significa que A^{-1} se obtém de I_n pela mesma sequência de operações elementares sobre linhas que transformaram A em I_n: obtém-se assim de novo o algoritmo expresso em (2.66):

$$\left[A \,\middle|\, I_n\right] \to \left[I_n \,\middle|\, A^{-1}\right]$$

Graças às matrizes elementares, podemos mesmo generalizar este algoritmo, permitindo que a condensação envolva também operações elementares sobre colunas:

Proposição 2.17 *Seja $A \in \mathbb{K}^{n,n}$ uma matriz quadrada regular de ordem n sobre um corpo $\mathbb{K}$. Fazendo operações elementares sobre linhas e sobre colunas da matriz (de ordem $2n$)*

$$\left[\begin{array}{c|c} A & I_n \\ \hline I_n & O_n \end{array}\right]$$

de modo a levá-la à forma

$$\left[\begin{array}{c|c} I_n & C \\ \hline B & O_n \end{array}\right]$$

tem-se

$$A^{-1} = BC$$

Demonstração:

Como I_n se obtém de A por operações elementares sobre linhas e colunas, existem sequências de matrizes elementares $(E_1, E_2, \ldots, E_p)$ e $\left(E'_1, E'_2, \ldots, E'_q\right)$ tais que

$$E_p \cdots E_2 E_1 A E'_1 E'_2 \ldots E'_q = I_n$$

Sejam $C = E_p \cdots E_2 E_1$ e $B = E'_1 E'_2 \ldots E'_q$. As matrizes B e C são regulares (por serem produto de matrizes regulares); então

$$CAB = I_n \Rightarrow C^{-1}CAB = C^{-1} \Rightarrow AB = C^{-1} \Rightarrow ABC = C^{-1}C \Rightarrow ABC = I_n$$

A última igualdade mostra que $A^{-1} = BC$, que é o resultado pretendido. $\square$

Observações:

- Em geral, começamos por transformar A numa matriz T_s triangular superior, por meio de operações elementares sobre as linhas de $\left[A \,|\, I_n\right]$, o que nos conduz à matriz C.

- Neste ponto, conhecemos a característica de A e saberemos se existe inversa.

- Se $\mathrm{c}(A) = n$, transformamos a matriz triangular T_s obtida anteriormente na matriz identidade I_n, por meio de operações elementares sobre as colunas de $\left[\begin{array}{c} T_s \\ \hline I_n \end{array}\right]$. Este processo fornece-nos a matriz B.

- Por fim, calcula-se o produto $BC = A^{-1}$.

Exemplo 2.38 Usemos o algoritmo anterior para calcular a inversa de

$$A = \begin{bmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{bmatrix}$$

Temos:

$$\left[\begin{array}{ccc|ccc} -1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 4 & -2 & 5 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & -2 & 3 & -1 & 0 & 0 \\ 0 & 5 & -6 & 2 & 1 & 0 \\ 0 & 6 & -7 & 4 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & -2 & 3 & -1 & 0 & 0 \\ 0 & 5 & -6 & 2 & 1 & 0 \\ 0 & 0 & 1 & 8 & -6 & 5 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 5 & -6 & 2 & 1 & 0 \\ 0 & 0 & 1 & 8 & -6 & 5 \\ \hline 1 & 2 & -3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 5 & 0 & 2 & 1 & 0 \\ 0 & 0 & 5 & 8 & -6 & 5 \\ \hline 1 & 2 & -3 & 0 & 0 & 0 \\ 0 & 1 & 6 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 8 & -6 & 5 \\ \hline 1 & \frac{2}{5} & -\frac{3}{5} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{6}{5} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right]$$

Portanto, tem-se:

$$B = \begin{bmatrix} 1 & \frac{2}{5} & -\frac{3}{5} \\ 0 & \frac{1}{5} & \frac{6}{5} \\ 0 & 0 & 1 \end{bmatrix} \text{ e } C = \begin{bmatrix} -1 & 0 & 0 \\ 2 & 1 & 0 \\ 8 & -6 & 5 \end{bmatrix}$$

A inversa é então dada por

$$A^{-1} = BC = \begin{bmatrix} 1 & \frac{2}{5} & -\frac{3}{5} \\ 0 & \frac{1}{5} & \frac{6}{5} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 2 & 1 & 0 \\ 8 & -6 & 5 \end{bmatrix} = \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix}$$

Vamos agora definir uma relação de equivalência no conjunto $\mathbb{K}^{m,n}$ das matrizes de tipo $m \times n$ sobre um corpo $\mathbb{K}$. É a noção de equivalência de matrizes:

Definição 2.18 – Matrizes equivalentes – *Diz-se que duas matrizes* $A, B \in \mathbb{K}^{m,n}$ *são* ***equivalentes*** *sse existem duas matrizes quadradas regulares* $P \in \mathbb{K}^{m,m}$ *e* $Q \in \mathbb{K}^{n,n}$ *tais que*

$$B = PAQ \tag{2.71}$$

A relação binária anterior é de *equivalência* (ver Apêndice A, definição A.5), visto que $A = I_m A I_n$, o que mostra que a relação é *reflexiva*; por outro lado, a implicação

$$B = PAQ \Rightarrow A = P^{-1}BQ^{-1}$$

mostra que a relação é *simétrica*. Por fim, a implicação

$$B = PAQ \wedge C = P'BQ' \Rightarrow C = P'(PAQ)Q' = (P'P)A(QQ')$$

prova que a relação é *transitiva*.

Como P e Q são regulares sse são produto de matrizes elementares, podemos dizer que A e B são equivalentes sse B se obtém de A por meio de operações elementares sobre linhas e colunas:

$$B = \underbrace{(E_p \cdots E_2 E_1)}_{P} A \underbrace{\left(E_1' E_2' \ldots E_q'\right)}_{Q}$$

Como as operações elementares sobre as filas de A não alteram a sua característica, isto significa que, se A e B são equivalentes, as suas características são necessariamente iguais. No capítulo 3, veremos que a recíproca desta implicação é também verdadeira (corolário 3.12.1).

2.14 Divisão matricial

Devido à não comutatividade do produto matricial, o problema da divisão matricial tem que ser posto de duas formas diferentes: a *divisão esquerda* (divisor à esquerda) e a *divisão direita* (divisor à direita). Assim, dadas uma matriz $A \in \mathbb{K}^{m,n}$ (o divisor) e uma matriz $B \in \mathbb{K}^{m,p}$ (o dividendo) com o mesmo número m de linhas que a matriz A, chamaremos quociente da *divisão esquerda* de B por A a qualquer matriz $X \in \mathbb{K}^{n,p}$ tal que

$$AX = B \tag{2.72}$$

De forma semelhante, se $B \in \mathbb{K}^{p,n}$ (o dividendo) for uma matriz com o mesmo número n de colunas que a matriz A, o quociente da *divisão direita* de B por A é qualquer matriz $X \in \mathbb{K}^{p,m}$ tal que

$$XA = B \tag{2.73}$$

Podemos fixar a nossa atenção exclusivamente na divisão esquerda, porque a questão da divisão direita se pode reduzir facilmente àquela, por transposição

$$XA = B \Leftrightarrow A^{\mathrm{T}} X^{\mathrm{T}} = B^{\mathrm{T}} \tag{2.74}$$

A equivalência anterior mostra que dividir B por A à direita equivale a dividir B^{T} por A^{T} à esquerda.

O problema de calcular X em (2.72) é equivalente à resolução de p sistemas de equações lineares com m equações e n incógnitas cada um, para o que basta fragmentar as matrizes X e B nas suas p colunas

$$\begin{aligned} X &= \left[X_1 \mid X_2 \mid \cdots \mid X_p\right] \\ B &= \left[B_1 \mid B_2 \mid \cdots \mid B_p\right] \end{aligned}$$

Fazendo o produto AX por blocos, a equação (2.72) transforma-se em

$$\left[AX_1 \mid AX_2 \mid \cdots \mid AX_p\right] = \left[B_1 \mid B_2 \mid \cdots \mid B_p\right]$$

o que equivale aos p sistemas com a mesma matriz A de coeficientes

$$AX_k = B_k \text{ , com } k = 1, 2, \ldots, p \tag{2.75}$$

e é, agora, óbvio que o problema da divisão esquerda tem solução sse estes sistemas são simultaneamente possíveis, o que acontece sse os vectores $B_k \in \mathbb{K}^m$ forem todos combinação linear das colunas da matriz A (ver proposição 2.10), ou seja sse $\mathrm{c}(A) = \mathrm{c}\big(\left[A \mid B\right]\big)$. Se esta condição se der, podemos calcular as soluções resolvendo simultaneamente os sistemas anteriores, mediante a condensação da matriz $\left[A \mid B\right] \in \mathbb{K}^{m,n+p}$. Se a característica r de A for igual a n, todos os sistemas serão determinados e há uma só solução X para o problema da divisão esquerda; se for $r < n$, então os p sistemas serão todos indeterminados com o mesmo grau de indeterminação $n - r$ e existirá mais que uma solução. Neste caso, as soluções X serão função de $p(n - r)$ parâmetros escalares arbitrários (as $n - r$ incógnitas secundárias de cada sistema), o que implica a existência de $(\#\mathbb{K})^{p(n-r)}$ soluções, se o corpo $\mathbb{K}$ for finito. Podemos, pois, apresentar a seguinte

Proposição 2.18 – Divisão esquerda – *Sejam $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,p}$ matrizes com o mesmo número de linhas. O problema da divisão esquerda de B por A é solúvel sse $c(A) = c\big(\left[A \mid B\right]\big)$ e, sendo $r = c(A)$, existe uma e uma só solução sse $r = n$. Quando $r < n$, existem várias soluções (uma infinidade, se $\mathbb{K}$ for infinito e em número de $(\#\mathbb{K})^{p(n-r)}$, se $\mathbb{K}$ for finito).*

Em virtude de (2.74) podemos enunciar proposição idêntica para a divisão direita:

Proposição 2.19 – Divisão direita – *Sejam $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{p,n}$ matrizes com o mesmo número de colunas. O problema da divisão direita de B por A é solúvel sse $c(A) = c\left(\begin{bmatrix} A \\ \hline B \end{bmatrix}\right)$ e, sendo $r = c(A)$, existe uma e uma só solução sse $r = m$. Quando $r < m$, existem várias soluções (uma infinidade, se $\mathbb{K}$ for infinito e em número de $(\#\mathbb{K})^{p(m-r)}$, se $\mathbb{K}$ for finito).*

O problema da resolução de um sistema $AX = B$ na forma matricial é, afinal, o problema da divisão esquerda de uma matriz-coluna B pela matriz A.

Identicamente, o problema da inversa direita de A é equivalente à divisão esquerda da matriz identidade I_m por $A \in \mathbb{K}^{m,n}$, visto que se trata de resolver a equação matricial

$$AX_d = I_m$$

Exemplo 2.39 Para dividir B por A à esquerda,

$$A = \begin{bmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 1 & 3 & -3 \end{bmatrix} \text{ e } B = \begin{bmatrix} 3 & -6 \\ 5 & -1 \\ 8 & -7 \end{bmatrix}$$

condensemos a matriz $\left[A \,|\, B\right]$

$$\left[\begin{array}{ccc|cc} -1 & 2 & -3 & 3 & -6 \\ 2 & 1 & 0 & 5 & -1 \\ 1 & 3 & -3 & 8 & -7 \end{array}\right] \to \begin{bmatrix} -1 & 2 & -3 & 3 & -6 \\ 0 & 5 & -6 & 11 & -13 \\ 0 & 5 & -6 & 11 & -13 \end{bmatrix} \to \left[\begin{array}{ccc|cc} 5 & 0 & 3 & 7 & 4 \\ 0 & 5 & -6 & 11 & -13 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Temos, neste caso, $\mathrm{c}(A) = \mathrm{c}\big(\left[A \,|\, B\right]\big) = 2$ e, portanto, a divisão esquerda é possível; os dois sistemas de equações lineares que acabámos de resolver são simplesmente indeterminados e as suas soluções são

$$X_1 = \begin{bmatrix} \frac{7}{5} - \frac{3}{5}\alpha_{11} & \frac{11}{5} + \frac{6}{5}\alpha_{11} & \alpha_{11} \end{bmatrix}^{\mathrm{T}} \text{ e } X_2 = \begin{bmatrix} \frac{4}{5} - \frac{3}{5}\alpha_{21} & -\frac{13}{5} + \frac{6}{5}\alpha_{21} & \alpha_{21} \end{bmatrix}^{\mathrm{T}}$$

O resultado da divisão é

$$X = \left[X_1 \;|\; X_2\right] = \begin{bmatrix} \frac{7}{5} - \frac{3}{5}\alpha_{11} & \frac{4}{5} - \frac{3}{5}\alpha_{21} \\ \frac{11}{5} + \frac{6}{5}\alpha_{11} & -\frac{13}{5} + \frac{6}{5}\alpha_{21} \\ \alpha_{11} & \alpha_{21} \end{bmatrix} = \begin{bmatrix} \frac{7}{5} - 3c_{11} & \frac{4}{5} - 3c_{21} \\ \frac{11}{5} + 6c_{11} & -\frac{13}{5} + 6c_{21} \\ 5c_{11} & 5c_{21} \end{bmatrix}$$

O leitor pode verificar a correcção do resultado, calculando $AX - B$ e verificando que é igual a $O_{3,2}$. Existe, pois, uma dupla infinidade de soluções para este problema de divisão esquerda.

Tem particular interesse o caso em que $A \in \mathbb{K}^{n,n}$ é matriz quadrada de ordem n *invertível*. Neste caso, as proposições anteriores mostram que os problemas de divisão esquerda e direita têm uma e uma só solução; para a divisão esquerda, fica, com $B \in \mathbb{K}^{n,p}$,

$$AX = B \Rightarrow A^{-1}(AX) = A^{-1}B \Rightarrow \big(A^{-1}A\big)X = A^{-1}B \Rightarrow I_nX = A^{-1}B \Rightarrow X = A^{-1}B$$

e o quociente único $A^{-1}B \in \mathbb{K}^{n,p}$ é designado pela notação $B\backslash A$:

$$B\backslash A = A^{-1}B$$

Para a divisão direita, vem, com $B \in \mathbb{K}^{p,n}$,

$$XA = B \Rightarrow (XA)A^{-1} = BA^{-1} \Rightarrow X\big(AA^{-1}\big) = BA^{-1} \Rightarrow XI_n = BA^{-1} \Rightarrow X = BA^{-1}$$

e para a matriz $BA^{-1} \in \mathbb{K}^{p,n}$ usa-se a notação B/A:

$$B/A = BA^{-1}$$

Em qualquer dos casos, o quociente tem o mesmo tipo que o dividendo B. Para poder ser $B\backslash A = B/A$, terá que ser B quadrada de ordem n e, mesmo nesse caso, ainda pode ser $B\backslash A \neq B/A$. Porém, a proposição 2.13.*vii* mostra que $A^{-1}B = BA^{-1}$ sse $AB = BA$. Portanto, se as matrizes A e B forem permutáveis, então a divisão esquerda e a divisão direita têm a mesma solução (quociente):

$$AB = BA \Rightarrow B\backslash A = B/A, \text{ em que } A, B \in \mathbb{K}^{n,n} \wedge A \text{ é invertível}$$

Exemplo 2.40 Para as matrizes não permutáveis

$$A = \begin{bmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 2 & -3 \\ 2 & -1 & -3 \end{bmatrix}$$

calculemos A^{-1}:

$$\left[\begin{array}{ccc|ccc} -1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 4 & -2 & 5 & 0 & 0 & 1 \end{array}\right] \to \begin{bmatrix} -1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 5 & -6 & 2 & 1 & 0 \\ 0 & 6 & -7 & 4 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 5 & 0 & 3 & -1 & 2 & 0 \\ 0 & 5 & -6 & 2 & 1 & 0 \\ 0 & 0 & 1 & 8 & -6 & 5 \end{bmatrix} \to$$
$$\begin{bmatrix} 5 & 0 & 0 & -25 & 20 & -15 \\ 0 & 5 & 0 & 50 & -35 & 30 \\ 0 & 0 & 1 & 8 & -6 & 5 \end{bmatrix} \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -5 & 4 & -3 \\ 0 & 1 & 0 & 10 & -7 & 6 \\ 0 & 0 & 1 & 8 & -6 & 5 \end{array}\right]$$

donde resulta que a inversa de A é:

$$A^{-1} = \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix}$$

Os quocientes esquerdo e direito de B por A são diferentes, porque A e B não são permutáveis:

$$B\backslash A = A^{-1}B = \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix} \begin{bmatrix} 0 & 1 & -2 \\ -1 & 2 & -3 \\ 2 & -1 & -3 \end{bmatrix} = \begin{bmatrix} -10 & 6 & 7 \\ 19 & -10 & -17 \\ 16 & -9 & -13 \end{bmatrix}$$
$$B/A = BA^{-1} = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 2 & -3 \\ 2 & -1 & -3 \end{bmatrix} \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix} = \begin{bmatrix} -6 & 5 & -4 \\ 1 & 0 & 0 \\ -44 & 33 & -27 \end{bmatrix}$$

Exemplo 2.41 Para as matrizes permutáveis (verifique!)

$$A = \begin{bmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{bmatrix} \text{ e } B = \begin{bmatrix} -58 & 32 & -75 \\ -22 & 28 & -54 \\ 64 & -14 & 56 \end{bmatrix}$$

tem-se, como se viu no exemplo anterior,

$$A^{-1} = \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix}$$

Os quocientes esquerdo e direito de B por A são, agora, iguais:

$$B\backslash A = A^{-1}B = \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix} \begin{bmatrix} -58 & 32 & -75 \\ -22 & 28 & -54 \\ 64 & -14 & 56 \end{bmatrix} = \begin{bmatrix} 10 & -6 & -9 \\ -42 & 40 & -36 \\ -12 & 18 & 4 \end{bmatrix}$$
$$B/A = BA^{-1} = \begin{bmatrix} -58 & 32 & -75 \\ -22 & 28 & -54 \\ 64 & -14 & 56 \end{bmatrix} \begin{bmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{bmatrix} = \begin{bmatrix} 10 & -6 & -9 \\ -42 & 40 & -36 \\ -12 & 18 & 4 \end{bmatrix}$$

O *package* **`ALGA`Matrizes``**, descrito na secção 2.16, implementa as funções **`DivisaoEsquerda[A,B,c]`** e **`DivisaoDireita[A,B,c]`** que resolvem, quando possível, o problema da divisão matricial de **`B`** por **`A`**, em função de parâmetros **`c`** arbitrários, apresentando-se ainda exemplos do seu uso.

2.15 Mudança de base

Vamos começar por introduzir notações matriciais para escrever as combinações lineares.

Num espaço vectorial E de dimensão finita n sobre um corpo $\mathbb{K}$, considere-se uma base

$$e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$$

Sabemos que, dado um vector $\vec{x} \in E$, são únicas as suas coordenadas $(x_i)_{1\leq i\leq n}$ em relação à base e e será

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i$$

Introduzindo a matriz-linha simbólica[14] $[\,\vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n\,]$ e a matriz-coluna $X_e = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ contendo as coordenadas referidas e identificando o seu produto (de tipo 1×1) com o vector $\vec{x} \in E$, podemos escrever

$$\vec{x} = [\,\vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n\,] X_e$$

No mesmo espaço vectorial E de dimensão finita n sobre um corpo $\mathbb{K}$, considerem-se, agora, duas bases

$$\begin{aligned} e &= (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) \\ e' &= \left(\vec{e}\,'_1, \vec{e}\,'_2, \ldots, \vec{e}\,'_n\right) \end{aligned}$$

---

14 Observe que esta matriz-linha é constituída por vectores: Não se trata, pois, de uma matriz sobre o corpo $\mathbb{K}$, mas sim sobre o espaço E.

Cada vector $\vec{e}\,'_j$ há-de exprimir-se nos $\vec{e}_i$ de uma e uma só forma, através das respectivas *coordenadas* $t_{ij} \in \mathbb{K}$ relativas à base e

$$\vec{e}\,'_j = \sum_{i=1}^{n} t_{ij}\, \vec{e}_i \tag{2.76}$$

As n igualdades vectoriais anteriores podem escrever-se matricialmente

$$[\,\vec{e}\,'_1 \;\; \vec{e}\,'_2 \;\; \cdots \;\; \vec{e}\,'_n\,] = [\,\vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n\,]\, T_{ee'} \tag{2.77}$$

A matriz $T_{ee'} = [\, t_{ij}\,]$, cuja coluna j contém as coordenadas de $\vec{e}\,'_j$ na base e, é chamada *matriz de mudança* ou *de passagem* da base e para a base e' e o facto de e' ser uma base de E implica que as suas colunas são *linearmente independentes* e que, portanto, $T_{ee'}$ é *regular*. Multiplicando ambos os membros de (2.77) por $T_{ee'}^{-1}$ à direita, obtém-se

$$[\,\vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n\,] = [\,\vec{e}\,'_1 \;\; \vec{e}\,'_2 \;\; \cdots \;\; \vec{e}\,'_n\,] T_{ee'}^{-1}$$

o que permite concluir que

$$T_{e'e} = T_{ee'}^{-1} \tag{2.78}$$

Além da relação anterior, é óbvio que

$$T_{ee} = I_n \tag{2.79}$$

Se for, agora, $e'' = \left(\vec{e}\,''_1, \vec{e}\,''_2, \ldots, \vec{e}\,''_n\right)$ uma terceira base de E, ter-se-á

$$\begin{cases} [\,\vec{e}\,''_1 \;\; \vec{e}\,''_2 \;\; \cdots \;\; \vec{e}\,''_n\,] = [\,\vec{e}\,'_1 \;\; \vec{e}\,'_2 \;\; \cdots \;\; \vec{e}\,'_n\,] T_{e'e''} \\ [\,\vec{e}\,'_1 \;\; \vec{e}\,'_2 \;\; \cdots \;\; \vec{e}\,'_n\,] = [\,\vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n\,]\, T_{ee'} \end{cases}$$

Substituindo a 2ª destas igualdades na 1ª, obtém-se

$$[\,\vec{e}\,''_1 \;\; \vec{e}\,''_2 \;\; \cdots \;\; \vec{e}\,''_n\,] = [\,\vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n\,]\,(T_{ee'} T_{e'e''})$$

concluindo-se que

$$T_{ee''} = T_{ee'} T_{e'e''} \tag{2.80}$$

Como se transformam as coordenadas de um vector $\vec{x}$ quando se muda de base? tem-se

$$\vec{x} = [\,\vec{e}\,'_1 \;\; \vec{e}\,'_2 \;\; \cdots \;\; \vec{e}\,'_n\,] X_{e'} = [\,\vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n\,] \underbrace{T_{ee'} X_{e'}}_{X_e} = [\,\vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n\,] X_e$$

Como as coordenadas de $\vec{x}$ na base e são únicas, deverá ser

$$X_e = T_{ee'} X_{e'} \tag{2.81}$$

Observe-se que as coordenadas do vector ocupam nesta igualdade posições *contrárias* às que são ocupadas pelos vectores das bases: daqui a designação de *coordenadas contra-variantes* que também lhes é dada.

Exemplo 2.42 Considerem-se, em $\mathbb{R}^3$, as bases $e = ((1,0,1),(-1,1,1,),(0,1,-1))$ e $e' = ((0,1,1),(1,2,-1),(-1,0,1))$. Para calcular a matriz $T_{ee'}$ de mudança da base e para e', teremos de exprimir os $\vec{e}'_j$ nos $\vec{e}_i$. Este processo conduz-nos a três sistemas de equações lineares com 3 equações e 3 incógnitas cada um, mas todos com a mesma matriz dos coeficientes: podemos, portanto, resolvê-los de uma só vez condensando a matriz seguinte

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 & 2 & 0 \\ 1 & 1 & -1 & 1 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 & 2 & 0 \\ 0 & 2 & -1 & 1 & -2 & 2 \end{bmatrix} \to$$

$$\to \begin{bmatrix} 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & -3 & -1 & -6 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 3 & 0 & 2 & 0 & 2 \\ 0 & 0 & -3 & -1 & -6 & 2 \end{bmatrix} \to$$
$$\to \begin{bmatrix} 3 & 0 & 0 & 2 & 3 & -1 \\ 0 & 3 & 0 & 2 & 0 & 2 \\ 0 & 0 & -3 & -1 & -6 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 2/3 & 1 & -1/3 \\ 0 & 1 & 0 & 2/3 & 0 & 2/3 \\ 0 & 0 & 1 & 1/3 & 2 & -2/3 \end{bmatrix}$$

Daqui se conclui que

$$T_{ee'} = \frac{1}{3}\begin{bmatrix} 2 & 3 & -1 \\ 2 & 0 & 2 \\ 1 & 6 & -2 \end{bmatrix}$$

Por exemplo, o vector $\vec{x} = 3\vec{e}'_1 - 2\vec{e}'_2 + 4\vec{e}'_3$ exprime-se na base e por

$$\vec{x} = [\,\vec{e}_1\ \ \vec{e}_2\ \ \vec{e}_3\,]\frac{1}{3}\begin{bmatrix} 2 & 3 & -1 \\ 2 & 0 & 2 \\ 1 & 6 & -2 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix} = \frac{1}{3}(-4\vec{e}_1 + 14\vec{e}_2 - 17\vec{e}_3)$$

Exemplo 2.43 Em relação ao exemplo anterior, as matrizes de mudança da base canónica c para as bases e e e' são, respectivamente,

$$T_{ce} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$
$$T_{ce'} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$

A expressão (2.80) permite confirmar a matriz obtida no exemplo anterior, já que

$$T_{ee'} = T_{ec}T_{ce'} = T_{ce}^{-1}T_{ce'} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & 2 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \\ 1 & -1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & 3 & -1 \\ 2 & 0 & 2 \\ 1 & 6 & -2 \end{bmatrix}$$

em que a inversa de T_{ce} se pode obter pelo algoritmo de condensação.

As coordenadas do vector $\vec{x}$ do exemplo anterior na base canónica serão

$$X_c = T_{ce'} X_{e'} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} -6 \\ -1 \\ 9 \end{bmatrix}$$

Portanto, $\vec{x} = (-6, -1, 9)$; este resultado pode ser confirmado por substituição directa dos $\vec{e}_i'$ em $\vec{x} = 3\vec{e}_1' - 2\vec{e}_2' + 4\vec{e}_3'$.

Exemplo 2.44 Podemos generalizar o procedimento usado nos exemplos anteriores, para obter a matriz $T_{ee'}$ de passagem de uma base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de $\mathbb{K}^n$ para uma segunda base $e' = (\vec{e}_1', \vec{e}_2', \ldots, \vec{e}_n')$:

$$[\,\vec{e}_1' \;\; \vec{e}_2' \;\; \cdots \;\; \vec{e}_n'\,] = [\,\vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n\,]\, T_{ee'}$$

Trata-se de calcular as coordenadas dos vectores $\vec{e}_i'$ em relação à base e e, para tal, será necessário resolver os n sistemas de equações lineares nas incógnitas $t_{ji} \in \mathbb{K}$ resultantes das n igualdades vectoriais

$$\vec{e}_i' = \sum_{j=1}^{n} t_{ji} \vec{e}_j; \; i = 1, 2, \ldots, n$$

Todos estes sistemas serão determinados e terão a mesma matriz E dos coeficientes (cujas colunas são constituídas pelos vectores da base e); as matrizes dos termos independentes destes sistemas são os n vectores de e'. Podemos, pois, resolvê-los de uma só vez fazendo a condensação vertical da matriz $\left[E \,\middle|\, E'\right]$ de tipo $n \times 2n$ (em que E' tem por colunas os vectores de e'), por forma a obter a matriz $\left[I_n \,\middle|\, T_{ee'}\right]$ do mesmo tipo e cuja metade direita será a matriz de mudança de base pretendida:

$$\left[E \,\middle|\, E'\right] \to \left[I_n \,\middle|\, T_{ee'}\right]$$

Se e for a base canónica de $\mathbb{K}^n$, ficará $E = I_n$ e, então, não será necessária a condensação, resultando imediatamente $T_{ee'} = E'$; se e' for a base canónica, então será $E' = I_n$ e a condensação $\left[E \,\middle|\, I_n\right] \to \left[I_n \,\middle|\, T_{ee'}\right]$ mostra, de acordo com (2.66), que será $T_{ee'} = E^{-1}$.

2.16 Anexos: matrizes e o MATHEMATICA©

MATHEMATICA©

O software MATHEMATICA© dispõe de funções capazes de resolver grande parte dos problemas por nós analisados neste capítulo; de entre as referidas funções salientamos:

- **`MatrixQ[A]`**

Verifica se o argumento **`A`** é uma matriz, devolvendo **`True`** ou **`False`**.

- **`IdentityMatrix[n]`**

Devolve a matriz *identidade* de ordem n.

- **`DiagonalMatrix[vector]`**

Devolve a matriz *diagonal* com as componentes do **`vector`** na diagonal.

- **`RowReduce[A]`**

Reduz a matriz **`A`** à forma *escalonada*, com pivots iguais a 1 e anulando todos os elementos nas colunas dos pivots à excepção destes. Facilmente se obtém a *característica* de **`A`**, por simples observação do resultado (número de linhas não nulas).

Se aplicada à matriz completa dum sistema, permite fàcilmente resolvê-lo. Para obter a matriz completa **`ab`** a partir das matrizes **`a`** e **`b`**, pode usar-se

```
ab = Transpose[Join[Transpose[a], {b}]]
```

- **`MatrixRank[A, ZeroTest->`***`test`***`]`**

Calcula a característica da matriz **`A`**, usando a função *`test`* para testar se os elementos de **`A`** são nulos.

- **`LinearSolve[A, B]`**

Calcula um vector **`sp`** *solução particular* do sistema **`A.X==B`**, se este for possível ou dá uma mensagem de erro, quando o sistema for impossível. Se o sistema for determinado, será esta a sua única solução; se for indeterminado, obtemos a solução geral mediante soma com uma combinação linear arbitrária dos vectores fornecidos pela função **`NullSpace[A]`**.

```
sp = LinearSolve[A,B]
```

- **`NullSpace[A]`**

Calcula uma base **`sh`** do espaço vectorial das *soluções do sistema homogéneo* **`A.X==0`**. A base é devolvida como uma lista **`sh`** de vectores (uma matriz do tipo $(n-r)\times n$).

```
sh = NullSpace[A]
```

Quando indeterminado, a solução geral **`sg`** de **`A.X==B`** obtém-se por

```
sg = sp + Array[c, {Length[sh]}].sh
```

onde os **`c[1]`**, **`c[2]`**, $\ldots$, **`c[n-r]`** são parâmetros arbitrários.

- **`Solve[A.X == B, X]`**

Resolve a *equação* **`A.X == B,`** em ordem às variáveis presentes na lista **`X`**. No caso de sistemas indeterminados, podemos escolher as r variáveis principais e colocá-las na lista **`X`**.

- **`Reduce[A.X == B, X]`**

Resolve a *equação* **`A.X == B,`** em ordem às variáveis mencionadas na lista **`X`**. Esta função é particularmente potente, permitindo mesmo fazer a discussão de sistemas com parâmetros (estes não são obviamente parte da lista **`X`**). A solução vem como uma expressão lógica envolvendo os operadores booleanos **`&&`** e || (conjunção e disjunção).

- **`Inverse[A]`**

Calcula a matriz *inversa* da matriz quadrada regular **`A`** ou dá uma mensagem de erro, se a matriz for singular.

A título de ilustração dos algoritmos descritos neste capítulo, apresenta-se nas próximas páginas, um *package* chamado **`ALGA`Matrizes``** implementado em MATHEMATICA©, onde se utilizam os algoritmos de condensação mencionados. Os módulos implementados são:

- **`Caracteristica[A,PivotUm -> opcao]`**

Devolve a lista formada pela *característica* e pela forma *escalonada* da matriz, com opção para que os pivots sejam 1 ou não (por defeito, não são).

- **`GaussJordan[A,B]`**

Usa o algoritmo de *Gauss-Jordan* na resolução do sistema de equações lineares **`A.X == B`**.

- **`Inversa[A]`**

Determina a *inversa* da matriz quadrada **`A`**, por condensação.

- **`InversaDireita[A,c]`**

Calcula as *inversas direitas* de **`A`**, se existirem, com parâmetros **`c`**.

- **`InversaEsquerda[A,c]`**

Calcula as *inversas esquerdas* de **`A`**, se existirem, com parâmetros **`c`**.

- **`DivisaoEsquerda[A,B,c]`**

Determina a divisão esquerda de **`B por A`**, se existir, com parâmetros **`c`**.

- **`DivisaoDireita[A,B,c]`**

Determina a divisão direita de **`B por A`**, se existir, com parâmetros **`c`**.

- **`GaussSeidel[A,B,X0,tol,max]`** determina a solução do sistema **`AX == B`**, a partir da estimativa inicial **`X0`** com a tolerancia **`tol`** e com um máximo de pontos **`max`** e devolve a lista formada pela solução aproximada e pelo número de pontos calculados (excluindo **`X0`**).

- **`GaussSeidelPoints[A,B,X0,tol,max]`** determina a lista de um máximo de **`max+1`** soluções aproximadas (incluindo **`X0`**) do sistema **`AX == B`**, a partir da estimativa inicial **`X0`** com a tolerancia **`tol`**.

- **`Passagem[e1,e2]`**

Determina a matriz de mudança da base **`e1`** para a base **`e2`**; mais geralmente, determina a matriz cujas colunas são as coordenadas dos vectores da lista **`e2`** em relação à base **`e1`**.

- **`Coordenadas[e1,e2,xe1]`**

Transforma as coordenadas **`xe1`** de um vector da base **`e1`** para a base **`e2`**.

Segue-se a listagem do *package* **``ALGA`Matrizes` ``**:

```
(* Package by Carlos Ribeiro, Março 2009 *)
(* Contexto : ALGA`Matrizes` *)
(* Versão : 3.7 *)
(* Versão do Mathematica : 7.0 *)

BeginPackage["ALGA`Matrizes`"]
```

```
(* ------------------------------ HELP ON-LINE ------------------------------ *)

AppendRows::usage = "AppendRows[mat1, mat2, ...] dá uma nova matriz composta pelas submatrizes \
mat1, mat2, ..., juntando-lhes as linhas. As submatrizes deverão ter o mesmo número de linhas."

AppendColumns::usage = "AppendColumns[mat1, mat2, ...] dá uma nova matriz composta pelas \
submatrizes mat1, mat2, ..., juntando-lhes as colunas. As submatrizes deverão ter o mesmo \
número de colunas."

ZeroMatrix::usage = "ZeroMatrix[m] devolve a matriz nula quadrada de ordem m. ZeroMatrix[m, n] \
devolve a matriz nula de tipo m x n."

SquareMatrixQ::usage = "SquareMatrixQ[mat] testa se a matriz mat é quadrada."

Op1::usage = "Op1[A,i,j] Matriz que resulta de trocar as linhas i e j na matriz A."

Op2::usage = "Op2[A,i,x] Matriz que resulta de multiplicar a linha i da matriz A pelo \
escalar x não-nulo."

Op3::usage = "Op3[A,i,j,y] matriz obtida de A, substituindo a linha i pela sua soma com \
o produto do escalar y pela linha j."

Op4::usage = "Op4[A,i,j,x,y] matriz obtida de A, substituindo a linha i \
pela soma do produto com o produto do escalar y pela linha j."

PivotUm::usage = "PivotUm é uma opção da função Caracteristica[] que indica que os pivots \
deverão ser 1s, quando PivotUm for igual a True. Por defeito é PivotUm -> False."

Caracteristica::usage="Caracteristica[A, PivotUm -> Opção] Devolve a lista composta pela \
característica de A, seguida da matriz levada à forma escalonada. PivotUm é uma opção que \
indica que os pivots deverão ser 1s, quando PivotUm for igual a True. \
Por defeito é PivotUm -> False."

GaussJordan::usage="GaussJordan[A,B] Resolve o sistema A.X = B, devolvendo uma lista formada \
por uma solução particular e pela lista contendo os vectores duma base do espaço das soluções \
do sistema homogéneo A.X = O."

Inversa::usage="Inversa[A] calcula a matriz inversa da matriz quadrada A."

InversaDireita::usage="InversaDireita[A,c] calcula a(s) matriz(es) inversa(s) direita(s) da \
matriz A e usando c como parâmetro."

InversaEsquerda::usage="InversaEsquerda[A,c] calcula a(s) matriz(es) inversa(s) esquerda(s) \
da matriz A e usando c como parâmetro."

DivisaoEsquerda::usage="DivisaoEsquerda[A,B,c] calcula o quociente de B por A, com o divisor \
A à esquerda e usando c como parâmetro."

DivisaoDireita::usage="DivisaoDireita[A, B, c] calcula o quociente de B por A, com o divisor \
A à direita e usando c como parâmetro."

GaussSeidel::usage="GaussSeidel[a, b, x0, err, iter] tenta resolver o sistema ax=b pelo método \
de Gauss-Seidel, a partir de um ponto inicial x0 com um erro inferior a err e realizando um \
máximo de iter iterações."

GaussSeidelPoints::usage="GaussSeidelPoints[a, b, x0, err, iter] devolve a lista de pontos \
obtida pelo algoritmo de Gauss-Seidel na resolução do sistema ax=b, a partir de um ponto \
inicial x0 com um erro inferior a err e realizando um máximo de iter iterações."

Passagem::usage="Passagem[e1, e2] Devolve a matriz cujas colunas são as coordenadas dos \
vectores da lista e2, na base formada pelos vectores da lista e1; se e2 for outra base, \
calcula a matriz de mudança de e1 para e2. A lista e1 deve ser uma base e os vectores de e2 \
devem pertencer ao espaço gerado por e1."
```

```
Coordenadas::usage="Coordenadas[e1, e2, xe1] Devolve as coordenadas em relação à base e2 de \
um vector, dadas as suas coordenadas xe1 em relação à base e1."

Begin["`Private`"]
(* -------------------------------- DEFAULTS -------------------------------- *)

Options[Caracteristica] = {PivotUm -> False};

(* --------------------------- MENSAGENS DE ERRO ---------------------------- *)

Op2::escalarnulo3 = "Operação do tipo 2 inválida. O 3° argumento não pode ser nulo.";

Op4::escalarnulo4 = "Operação do tipo 4 inválida. O 4° argumento não pode ser nulo.";

GaussJordan::errdimensao = "GaussJordan[A,B]: N° de linhas da matriz A diferente do \
comprimento do vector B.";
GaussJordan::impossivel = "Sistema AX = B impossível.";

Caracteristica::matrizvazia = "Matriz vazia."

Inversa::inverrdimensao = "A matriz a inverter não é quadrada ou é vazia.";
Inversa::singular = "A matriz é singular.";

InversaDireita::naoexiste = "A matriz não tem inversa(s) direita(s).";

InversaEsquerda::naoexiste = "A matriz não tem inversa(s) esquerda(s).";

DivisaoEsquerda::naoexiste = "A divisão esquerda não é possível.";
DivisaoEsquerda::dimensao = "O número de linhas das matrizes a dividir não é igual";

DivisaoDireita::naoexiste = "A divisão direita não é possível.";
DivisaoDireita::dimensao = "O número de colunas das matrizes a dividir não é igual";

GaussSeidel::errdim = GaussSeidelPoints::errdim = "O comprimento do vector inicial é \
diferente da ordem da matriz simples do sistema.";
GaussSeidel::errindep = GaussSeidelPoints::errindep = "O comprimento do vector dos termos \
independentes é diferente da ordem da matriz simples do sistema.";
GaussSeidel::errdiag=GaussSeidelPoints::errdiag="Existe(m) elemento(s) nulo(s) na diagonal \
da matriz simples.";
GaussSeidel::dimensao = GaussSeidelPoints::dimensao = "A matriz simples do sistema \
não é quadrada.";

Passagem::naobase = "A primeira lista não é uma base.";
Passagem::errdim = "Os vectores da segunda lista não pertencem ao espaço gerado pelos \
vectores da primeira lista.";

Coordenadas::naobase = "A segunda lista não é uma base.";
Coordenadas::errdim = "As bases indicadas nos dois primeiros argumentos não são do mesmo \
espaço ou o vector tem a dimensão errada.";

(* ------------------------------ IMPLEMENTAÇÃO ------------------------------ *)

(* Implementação da função ZeroMatrix *)
ZeroMatrix[0, ___] := {}
ZeroMatrix[m_Integer, 0] := Table[{}, {m}]

ZeroMatrix[m_Integer, n_Integer] := Normal[SparseArray[{},{m, n}]] /; \
m >= 0 && n>=0

ZeroMatrix[m_Integer] := ZeroMatrix[m, m] /; m >= 0

(* Implementação da função SquareMatrixQ *)
```

```
SquareMatrixQ[a_?MatrixQ] := SameQ @@ Dimensions[a]

(* Implementação da função AppendRows *)
SameRowSize[l_List] := (SameQ @@ (Dimensions[#][[1]]& /@ l) )
AppendRows[l__?MatrixQ] := Transpose[Join @@ Transpose /@ {l}] /; \
SameRowSize[{l}]

(* Implementação da função AppendColumns *)
SameColumnSize[l_List] := (SameQ @@ (Dimensions[#][[2]]& /@ l) )
AppendColumns[l__?MatrixQ] := Join[l] /; SameColumnSize[{l}]

(* Implementação das operações elementares *)
Op1[a_?MatrixQ, i_Integer, j_Integer] :=
  Module[{temp = a[[i]], b = a}, b[[i]] = b[[j]]; b[[j]] = temp; b]

Op2[a_?MatrixQ, i_Integer, x_] :=
  Module[{b = a}, If[x == 0, Message[Op2::escalarnulo3],
    b[[i]] = x b[[i]]; b]]

Op3[a_?MatrixQ, i_Integer, j_Integer, x_] :=
  Module[{b = a}, b[[i]] = b[[i]] + x b[[j]]; b]

Op4[a_?MatrixQ, i_Integer, j_Integer, x_, y_] :=
  Module[{b = a}, If[x == 0, Message[Op4::escalarnulo4],
    b[[i]] = x b[[i]] + y b[[j]]; b]]

(* Implementação do módulo Caracteristica *)
Caracteristica[a_?MatrixQ, opt___?OptionQ] :=
  Module[{m = Length[a], n = Dimensions[a][[2]], b = a, r = 0, k = 0, i, piv1},
    {piv1} = {PivotUm}/.Flatten[{opt, Options[Caracteristica]}];
    If[m == 0 || n== 0, Message[Caracteristica::matrizvazia];
      Print["Caracteristica[", a, ", PivotUm -> ", piv1, "]"],
      While[r < m && k < n,
        While[k++; i = r; While[i++; b[[i, k]] == 0 && i < m, Null];
          b[[i, k]] == 0 && k < n, Null
        ];
        If[b[[i, k]] != 0,
          r++;
          b = Op1[b, r, i];
          If[piv1,
            b = Op2[b, r, 1/b[[r, k]]];
            For[i = r + 1, i <= m, i++, b = Op3[b, i, r, -b[[i, k]]]],
            For[i = r + 1, i <= m, i++, If[b[[i,k]] != 0,
              b = Op4[b, i, r, b[[r, k]], -b[[i,k]]]]]
          ]
        ]
      ]; {r,b}
    ]
  ]

(* Implementação do módulo GaussJordan *)
GaussJordan[a_?MatrixQ, b_?VectorQ] :=
  Module[{m = Length[a], n = Dimensions[a][[2]], s = 0, r = 0,
    k = 0, c, i, piv, sp, sh},
    If[m != Length[b],
      Message[GaussJordan::errdimensao],
      (* Condensação da matriz completa *)
      c = Transpose[Join[Transpose[a],{b}]];
      piv = Table[False,{n}];
      While[s < m && k < n+1,
        While[k++; i = s; While[i++; c[[i, k]] == 0 && i < m, Null];
          c[[i, k]] == 0 && k < n+1, Null
```

```
        ];
        If[c[[i, k]] != 0,
          s++;
          If[k <= n, r++; piv[[k]] = True];
          If[i != s, c = Op1[c, s, i]];
          c = Op2[c, s, 1/c[[s, k]]];
          For[i = 1, i <= m, i++,
            If[i != s, c = Op3[c, i, s, -c[[i, k]]]]
          ]
        ]
      ];
      (* Discussão da solução e preparação da saída *)
      If[s != r,
        (* Sistema impossível *)
        Message[GaussJordan::impossivel];
        Print["GaussJordan[", a, ", ", b, "]"],
        (* Sistema possível *)
        (* sp: solução particular *)
        sp = Table[0,{n}];
        For[k = 1; i = 0, k <= n, k++,
          If[piv[[k]], i++; sp[[k]] = c[[i, n+1]]]
        ];
        (* sh: base do espaço das soluções do sistema homogéneo associado *)
        sh = Table[0,{n-r},{n}];
        For[s = 1; t = 0, s <= n, s++,
          If[Not[piv[[s]]],
            t++;
            For[k = 1; i = 0, k <= n, k++,
              If[piv[[k]], i++; sh[[t, k]] = -c[[i, s]], sh[[t, s]] = 1]
            ]
          ]
        ];
        {sp, sh}
      ]
    ]
  ]

(* Implementação do módulo Inversa *)
Inversa[a_?MatrixQ] :=
  Module[{n = Length[a], cont = True, r = 0, k = 0, b, i},
    If[Not[SquareMatrixQ[a]],
      Message[Inversa::inverrdimensao]; Print["Inversa[", a, "]"],
      (* Inversão da matriz *)
      b = Transpose[Join[Transpose[a],IdentityMatrix[n]]];
      While[r < n && k < n && cont,
        k++; i = r; While[i++; b[[i, k]] == 0 && i < n, Null];
        If[b[[i, k]] == 0,
          cont = False,
          r++;
          If[i != r, b = Op1[b, r, i]];
          b = Op2[b, r, 1/b[[r, k]]];
          For[i = 1, i <= n, i++,
            If[i != r, b = Op3[b, i, r, -b[[i, k]]]]
          ]
        ]
      ];
      If[cont, Transpose[Take[Transpose[b],-n]],
        Message[Inversa::singular];Print["Inversa[", a, "]"]
      ]
    ]
  ]
```

```
(* Implementação do módulo InversaDireita *)
InversaDireita[a_?MatrixQ, c_]:=
  Module[{sh = NullSpace[a], i},
    If[Caracteristica[a][[1]] == Length[a],
      Transpose[Table[LinearSolve[a,IdentityMatrix[Length[a]][[i]]]+
      Array[c[i],{Length[sh]}].sh,{i,Length[a]}]],
      Message[InversaDireita::naoexiste];
      Print["InversaDireita[", a, ", ", c, "]"]
    ]
  ]

(* Implementação do módulo InversaEsquerda *)
InversaEsquerda[a_?MatrixQ, c_] :=
  If[Caracteristica[a][[1]] == Dimensions[a][[2]],
    Transpose[InversaDireita[Transpose[a], c]],
    Message[InversaEsquerda::naoexiste];
    Print["InversaEsquerda[", a, ", ", c, "]"]
  ]

(* Implementação do módulo DivisaoEsquerda *)
DivisaoEsquerda[a_?MatrixQ,b_?MatrixQ, c_]:=
  Module[{sh=NullSpace[a]},
    If[Length[a]==Length[b],
      If[Caracteristica[a][[1]]== \
        Caracteristica[Join[Transpose[a],Transpose[b]]][[1]],
        Transpose[Table[LinearSolve[a,b[[All,i]]]+Array[c[i],{Length[sh]}].sh, \
          {i,Dimensions[b][[2]]}]],
        Message[DivisaoEsquerda::naoexiste];
        Print["DivisaoEsquerda[",a,", ",b,"]"]
      ],
      Message[DivisaoEsquerda::dimensao];Print["DivisaoEsquerda[",a,",",b,"]"]
    ]
  ]

(* Implementação do módulo DivisaoDireita *)
DivisaoDireita[a_?MatrixQ,b_?MatrixQ, c_]:=
  Module[{sh=NullSpace[a]},
    If[Dimensions[a][[2]]==Dimensions[b][[2]],
      If[Caracteristica[a][[1]]==Caracteristica[Join[a,b]][[1]],
        Transpose[DivisaoEsquerda[Transpose[a],Transpose[b],c]],
        Message[DivisaoDireita::naoexiste]; Print["DivisaoDireita[",a,", ",b,"]"]
      ],
      Message[DivisaoDireita::dimensao];Print["DivisaoDireita[",a,", ",b,"]"]
    ]
  ]

(* Implementação do módulo GaussSeidel *)
GaussSeidel[a_?MatrixQ,b_?VectorQ,x0_?VectorQ,ε_,maxiter_?IntegerQ]:=
  Module[{y0=x0,y1=x0,n=Length[x0],i,j,cont=True,m=0,ok=True},
    If[SquareMatrixQ[a],
      If[Length[a]==Length[b],
        If[Length[a]==Length[x0],
          For[i=1,i<=n,i++,ok=ok && a[[i,i]]!=0];
          If[ok,
            While[cont,
              For[i=1,i<=n,i++,
                y1[[i]]=1/a[[i,i]]*(b[[i]]-Sum[a[[i,j]]*y1[[j]],{j,1,i-1}]-\
                        Sum[a[[i,j]]*y0[[j]],{j,i+1,n}]);
                m++
              ];
              cont=Norm[y1-y0]>=ε && m<=maxiter;
              y0=y1
```

```
            ];
            {y1,m},
            Message[GaussSeidel::errdiag]
          ],
          Message[GaussSeidel::errdim]
        ],
        Message[GaussSeidel::errindep]
      ],
      Message[GaussSeidel::dimensao]
    ]
  ]

(* Implementação do módulo GaussSeidelPoints *)
GaussSeidelPoints[a_?MatrixQ,b_?VectorQ,x0_?VectorQ,ε_,maxiter_?IntegerQ]:=
  Module[{y0=x0,y1=x0,n=Length[x0],i,j,cont=True,lista,m=0,ok=True},
    If[SquareMatrixQ[a],
      If[Length[a]==Length[b],
        If[Length[a]==Length[x0],
          For[i=1,i<=n,i++,ok=ok && a[[i,i]]!=0];
            If[ok,
              lista={y1};
              While[cont,
                For[i=1,i<=n,i++,
                  y1[[i]]=1/a[[i,i]]*(b[[i]]-Sum[a[[i,j]]*y1[[j]],{j,1,i-1}]-\
                          Sum[a[[i,j]]*y0[[j]],{j,i+1,n}]);
                  lista=Join[lista,{y1}];
                  m++
                ];
                cont=Norm[y1-y0]>=ε && m<=maxiter;
                y0=y1
              ];
            lista,
            Message[GaussSeidelPoints::errdiag]
          ],
          Message[GaussSeidelPoints::errdim]
        ],
        Message[GaussSeidelPoints::errindep]
      ],
      Message[GaussSeidelPoints::dimensao]
    ]
  ]

(* Implementação do módulo Passagem *)
Passagem[e1_?MatrixQ,e2_?MatrixQ]:=
  Module[{n=Length[e1]},
    Which[Not[SquareMatrixQ[e1]], Message[Passagem::naobase],
      MatrixRank[e1]!=n, Message[Passagem::naobase],
      Dimensions[e2][[2]]!=n, Message[Passagem::errdim],
      True, Transpose[e2.Inverse[e1]]
    ]
  ]

(* Implementação do módulo Coordenadas *)
Coordenadas[e1_?MatrixQ,e2_?MatrixQ,xe1_?VectorQ]:=
  Module[{n=Length[e1]},
    Which[Not[SquareMatrixQ[e1]], Message[Coordenadas::naobase],
      MatrixRank[e1]!=n, Message[Coordenadas::naobase],
      Length[e2]!=n || Length[xe1]!=n, Message[Coordenadas::errdim],
      True, Passagem[e2,e1].xe1
    ]
  ]
```

```
End[]
EndPackage[]
```

O *package* **ALGA`Matrizes`** deverá ser previamente carregado por meio de um dos comandos

```
<<ALGA`Matrizes`
Get["ALGA`Matrizes`"]
Needs["ALGA`Matrizes`"]
DeclarePackage["ALGA`Matrizes`"]
```

Nas páginas seguintes, apresenta-se um *notebook* onde se ilustra o uso do *package* anterior e das funções próprias do software MATHEMATICA© para resolver vários problemas matriciais.

Matrizes

2.16.1. O que são?

- **No MATHEMATICA, uma matriz é uma lista de listas do mesmo comprimento**

```
a = {{2, 2, 1}, {3, -2, 1}, {1, -5, 1}, {3, 2, -1}};
```

```
b = {{2, -5}, {3, 4, 1}};
```

```
c = {3, -2, 5};
```

- **Pode formatar-se uma matriz com a função** `MatrixForm[]`

```
a // MatrixForm
```

$$\begin{pmatrix} 2 & 2 & 1 \\ 3 & -2 & 1 \\ 1 & -5 & 1 \\ 3 & 2 & -1 \end{pmatrix}$$

2.16.2. Acesso aos elementos de uma matriz

- **Podemos extrair partes de uma matriz**

Obtém-se o elemento na posição `(3,2)` da matriz por meio de `a[[3,2]]` ou `Part[a,3,2]`

```
{a[[3, 2]], Part[a, 3, 2]}
```

```
{-5, -5}
```

A linha 3

```
a[[3]]
```

```
{1, -5, 1}
```

A coluna 2

```
a[[All, 2]]
```

```
{2, -2, -5, 2}
```

Podemos extrair as 2 primeiras linhas com `Take[a,2]`

```
Take[a, 2] // MatrixForm
```

$$\begin{pmatrix} 2 & 2 & 1 \\ 3 & -2 & 1 \end{pmatrix}$$

Ou as 2 últimas linhas

```
Take[a, -2] // MatrixForm
```

$$\begin{pmatrix} 1 & -5 & 1 \\ 3 & 2 & -1 \end{pmatrix}$$

Podemos extrair as 2 primeiras colunas

```
Take[a, All, 2] // MatrixForm
```

$$\begin{pmatrix} 2 & 2 \\ 3 & -2 \\ 1 & -5 \\ 3 & 2 \end{pmatrix}$$

Podemos extrair as 2 últimas colunas

```
Take[a, All, -2] // MatrixForm
```

$$\begin{pmatrix} 2 & 1 \\ -2 & 1 \\ -5 & 1 \\ 2 & -1 \end{pmatrix}$$

Ou as linhas de ordem par

```
Take[a, {2, 4, 2}] // MatrixForm
```

$$\begin{pmatrix} 3 & -2 & 1 \\ 3 & 2 & -1 \end{pmatrix}$$

Ou as colunas 2 e 3

```
Take[a, All, {2, 3}] // MatrixForm
```

$$\begin{pmatrix} 2 & 1 \\ -2 & 1 \\ -5 & 1 \\ 2 & -1 \end{pmatrix}$$

2.16.3. Dimensões (tipo) de uma matriz

Podemos obter o vector contendo o número de linhas e de colunas (tipo da matriz)

```
Dimensions[a]
```

```
{4, 3}
```

E extrair daí o número de linhas

```
Dimensions[a][[1]]
```

```
4
```

```
Length[a]
```

```
4
```

E o número de colunas

```
Dimensions[a][[2]]
```

```
3
```

2.16.4. Trata-se mesmo de uma matriz?

A função `MatrixQ` verifica se um objecto é uma matriz

```
MatrixQ[a]
```

```
True
```

```
MatrixQ[b]
```

```
False
```

```
MatrixQ[c]
```

```
False
```

2.16.5. Geração de matrizes

Podemos gerar matrizes identidade de qualquer ordem

```
IdentityMatrix[4] // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Podemos gerar matrizes diagonais

```
DiagonalMatrix[{2, -3, 0, 1}] // MatrixForm
```

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Matriz diagonal complexa com elementos diagonais aleatórios

```
DiagonalMatrix[Table[RandomInteger[{-10, 10}] +
    i RandomInteger[{-10, 10}], {5}]] // MatrixForm
```

$$\begin{pmatrix} -9+10\,i & 0 & 0 & 0 & 0 \\ 0 & -6+9\,i & 0 & 0 & 0 \\ 0 & 0 & -5-2\,i & 0 & 0 \\ 0 & 0 & 0 & 3-8\,i & 0 \\ 0 & 0 & 0 & 0 & -10-10\,i \end{pmatrix}$$

Podem gerar-se outras matrizes com `Table[]` ou `Array[]`

```
(a = Table[1 / (i + j - 1), {i, 3}, {j, 4}]) // MatrixForm
```

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \end{pmatrix}$$

```
(b = Array[x, {3, 4}]) // MatrixForm
```

$$\begin{pmatrix} x[1,1] & x[1,2] & x[1,3] & x[1,4] \\ x[2,1] & x[2,2] & x[2,3] & x[2,4] \\ x[3,1] & x[3,2] & x[3,3] & x[3,4] \end{pmatrix}$$

```
(c = RandomInteger[{-9, 9}, {4, 3}]) // MatrixForm
```

$$\begin{pmatrix} 0 & -1 & 9 \\ 6 & -9 & 4 \\ 5 & 3 & -9 \\ 3 & 7 & 8 \end{pmatrix}$$

Podemos gerar uma matriz tridiagonal de 8ª ordem com elementos aleatórios inteiros entre `-10` e `10` nas 3 diagonais

```
Table[If[Abs[i - j] ≤ 1, RandomInteger[{-10, 10}], 0],
  {i, 8}, {j, 8}] // MatrixForm
```

$$\begin{pmatrix} -2 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ -9 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & -8 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & -9 & 6 & -6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 9 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & -5 \end{pmatrix}$$

Matriz 5x6 formada por números primos escolhidos aleatoriamente de entre os primeiros 2000

```
Table[Prime[RandomInteger[{1, 2000}]], {i, 5}, {j, 6}] // MatrixForm
```

$$\begin{pmatrix} 8273 & 9377 & 3271 & 16\,927 & 11\,681 & 15\,889 \\ 773 & 1483 & 8501 & 5639 & 4759 & 10\,657 \\ 17\,093 & 2711 & 3617 & 14\,983 & 12\,763 & 14\,699 \\ 1579 & 4591 & 11\,159 & 9601 & 1559 & 2207 \\ 2417 & 13\,691 & 1451 & 479 & 3863 & 4793 \end{pmatrix}$$

2.16.6. Operações matriciais

Podem somar-se matrizes com + ou `Plus[]`

```
(c = a + b) // MatrixForm
```

$$\begin{pmatrix} 1 + x[1,1] & \frac{1}{2} + x[1,2] & \frac{1}{3} + x[1,3] & \frac{1}{4} + x[1,4] \\ \frac{1}{2} + x[2,1] & \frac{1}{3} + x[2,2] & \frac{1}{4} + x[2,3] & \frac{1}{5} + x[2,4] \\ \frac{1}{3} + x[3,1] & \frac{1}{4} + x[3,2] & \frac{1}{5} + x[3,3] & \frac{1}{6} + x[3,4] \end{pmatrix}$$

Multiplicar por escalar

```
(d = -Pi * a) // MatrixForm
```

$$\begin{pmatrix} -\pi & -\frac{\pi}{2} & -\frac{\pi}{3} & -\frac{\pi}{4} \\ -\frac{\pi}{2} & -\frac{\pi}{3} & -\frac{\pi}{4} & -\frac{\pi}{5} \\ -\frac{\pi}{3} & -\frac{\pi}{4} & -\frac{\pi}{5} & -\frac{\pi}{6} \end{pmatrix}$$

Podem transpor-se matrizes

```
(e = Transpose[d]) // MatrixForm
```

$$\begin{pmatrix} -\pi & -\frac{\pi}{2} & -\frac{\pi}{3} \\ -\frac{\pi}{2} & -\frac{\pi}{3} & -\frac{\pi}{4} \\ -\frac{\pi}{3} & -\frac{\pi}{4} & -\frac{\pi}{5} \\ -\frac{\pi}{4} & -\frac{\pi}{5} & -\frac{\pi}{6} \end{pmatrix}$$

Podem multiplicar-se matrizes com . ou `Dot[]` - atenção à não comutatividade

```
(d.e) // MatrixForm
```

$$\begin{pmatrix} \frac{205\,\pi^2}{144} & \frac{4\,\pi^2}{5} & \frac{17\,\pi^2}{30} \\ \frac{4\,\pi^2}{5} & \frac{1669\,\pi^2}{3600} & \frac{\pi^2}{3} \\ \frac{17\,\pi^2}{30} & \frac{\pi^2}{3} & \frac{869\,\pi^2}{3600} \end{pmatrix}$$

```
(e.d) // MatrixForm
```

$$\begin{pmatrix} \frac{49\,\pi^2}{36} & \frac{3\,\pi^2}{4} & \frac{21\,\pi^2}{40} & \frac{73\,\pi^2}{180} \\ \frac{3\,\pi^2}{4} & \frac{61\,\pi^2}{144} & \frac{3\,\pi^2}{10} & \frac{7\,\pi^2}{30} \\ \frac{21\,\pi^2}{40} & \frac{3\,\pi^2}{10} & \frac{769\,\pi^2}{3600} & \frac{\pi^2}{6} \\ \frac{73\,\pi^2}{180} & \frac{7\,\pi^2}{30} & \frac{\pi^2}{6} & \frac{469\,\pi^2}{3600} \end{pmatrix}$$

As matrizes podem ser complexas

```
(a = Table[2 * i / (j - 2 * I), {i, 4}, {j, 3}]) // MatrixForm
```

$$\begin{pmatrix} \frac{2}{5}+\frac{4\,i}{5} & \frac{1}{2}+\frac{i}{2} & \frac{6}{13}+\frac{4\,i}{13} \\ \frac{4}{5}+\frac{8\,i}{5} & 1+i & \frac{12}{13}+\frac{8\,i}{13} \\ \frac{6}{5}+\frac{12\,i}{5} & \frac{3}{2}+\frac{3\,i}{2} & \frac{18}{13}+\frac{12\,i}{13} \\ \frac{8}{5}+\frac{16\,i}{5} & 2+2\,i & \frac{24}{13}+\frac{16\,i}{13} \end{pmatrix}$$

```
(b = Table[RandomInteger[{-20, 20}] + RandomInteger[{-10, 10}] i,
    {i, 4}, {j, 3}]) // MatrixForm
```

$$\begin{pmatrix} -20-7\,i & 7+2\,i & -20-9\,i \\ 18-2\,i & 9-7\,i & 8-2\,i \\ -11-2\,i & -1+5\,i & 2+i \\ -13+8\,i & -5+i & 0 \end{pmatrix}$$

E podemos subtrair

```
(c = a - b) // MatrixForm
```

$$\begin{pmatrix} \frac{102}{5}+\frac{39\,i}{5} & -\frac{13}{2}-\frac{3\,i}{2} & \frac{266}{13}+\frac{121\,i}{13} \\ -\frac{86}{5}+\frac{18\,i}{5} & -8+8\,i & -\frac{92}{13}+\frac{34\,i}{13} \\ \frac{61}{5}+\frac{22\,i}{5} & \frac{5}{2}-\frac{7\,i}{2} & -\frac{8}{13}-\frac{i}{13} \\ \frac{73}{5}-\frac{24\,i}{5} & 7+i & \frac{24}{13}+\frac{16\,i}{13} \end{pmatrix}$$

E transpor

```
Transpose[c] // MatrixForm
```

$$\begin{pmatrix} \frac{102}{5}+\frac{39\,i}{5} & -\frac{86}{5}+\frac{18\,i}{5} & \frac{61}{5}+\frac{22\,i}{5} & \frac{73}{5}-\frac{24\,i}{5} \\ -\frac{13}{2}-\frac{3\,i}{2} & -8+8\,i & \frac{5}{2}-\frac{7\,i}{2} & 7+i \\ \frac{266}{13}+\frac{121\,i}{13} & -\frac{92}{13}+\frac{34\,i}{13} & -\frac{8}{13}-\frac{i}{13} & \frac{24}{13}+\frac{16\,i}{13} \end{pmatrix}$$

E conjugar

```
(d = Conjugate[c]) // MatrixForm
```

$$\begin{pmatrix} \frac{102}{5}-\frac{39\,i}{5} & -\frac{13}{2}+\frac{3\,i}{2} & \frac{266}{13}-\frac{121\,i}{13} \\ -\frac{86}{5}-\frac{18\,i}{5} & -8-8\,i & -\frac{92}{13}-\frac{34\,i}{13} \\ \frac{61}{5}-\frac{22\,i}{5} & \frac{5}{2}+\frac{7\,i}{2} & -\frac{8}{13}+\frac{i}{13} \\ \frac{73}{5}+\frac{24\,i}{5} & 7-i & \frac{24}{13}-\frac{16\,i}{13} \end{pmatrix}$$

Podemos definir a função transconjugada

```
TransConjugada[a_?MatrixQ] := Transpose[Conjugate[a]]
```

Podemos, agora, usá-la

```
TransConjugada[d] // MatrixForm
```

$$\begin{pmatrix} \frac{102}{5}+\frac{39\,i}{5} & -\frac{86}{5}+\frac{18\,i}{5} & \frac{61}{5}+\frac{22\,i}{5} & \frac{73}{5}-\frac{24\,i}{5} \\ -\frac{13}{2}-\frac{3\,i}{2} & -8+8\,i & \frac{5}{2}-\frac{7\,i}{2} & 7+i \\ \frac{266}{13}+\frac{121\,i}{13} & -\frac{92}{13}+\frac{34\,i}{13} & -\frac{8}{13}-\frac{i}{13} & \frac{24}{13}+\frac{16\,i}{13} \end{pmatrix}$$

```
TransConjugada[TransConjugada[d]] == d
```

```
True
```

Podemos calcular potências de matrizes

```
(a = Table[(i + j) / (1 + Abs[i - j]), {i, 4}, {j, 4}]) // MatrixForm
```

$$\begin{pmatrix} 2 & \frac{3}{2} & \frac{4}{3} & \frac{5}{4} \\ \frac{3}{2} & 4 & \frac{5}{2} & 2 \\ \frac{4}{3} & \frac{5}{2} & 6 & \frac{7}{2} \\ \frac{5}{4} & 2 & \frac{7}{2} & 8 \end{pmatrix}$$

```
MatrixPower[a, 3] // MatrixForm
```

$$\begin{pmatrix} \frac{3301}{36} & \frac{5153}{32} & \frac{12\,593}{54} & \frac{154\,805}{576} \\ \frac{5153}{32} & \frac{581}{2} & \frac{59\,935}{144} & \frac{11\,317}{24} \\ \frac{12\,593}{54} & \frac{59\,935}{144} & \frac{5783}{9} & \frac{211\,333}{288} \\ \frac{154\,805}{576} & \frac{11\,317}{24} & \frac{211\,333}{288} & \frac{22\,651}{24} \end{pmatrix}$$

```
MatrixPower[a, -2] // MatrixForm
```

$$\begin{pmatrix} \frac{3\,956\,854\,176}{6\,902\,452\,561} & -\frac{1\,709\,058\,624}{6\,902\,452\,561} & -\frac{120\,503\,304}{6\,902\,452\,561} & -\frac{170\,052\,096}{6\,902\,452\,561} \\ -\frac{1\,709\,058\,624}{6\,902\,452\,561} & \frac{1\,609\,475\,444}{6\,902\,452\,561} & -\frac{442\,309\,104}{6\,902\,452\,561} & \frac{41\,067\,096}{6\,902\,452\,561} \\ -\frac{120\,503\,304}{6\,902\,452\,561} & -\frac{442\,309\,104}{6\,902\,452\,561} & \frac{659\,269\,044}{6\,902\,452\,561} & -\frac{231\,678\,624}{6\,902\,452\,561} \\ -\frac{170\,052\,096}{6\,902\,452\,561} & \frac{41\,067\,096}{6\,902\,452\,561} & -\frac{231\,678\,624}{6\,902\,452\,561} & \frac{266\,544\,416}{6\,902\,452\,561} \end{pmatrix}$$

```
MatrixPower[a, 0] // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2.16.7. Matrizes permutáveis

- **Definir função polinómio** `PolAleatorio` **na matriz mat com coeficientes inteiros aleatórios entre** -9 **e** 9 **e grau aleatório até** `maxgrau` **- devolve uma lista contendo o grau gerado, os coeficientes gerados e a matriz obtida**

```
PolAleatorio[mat_?MatrixQ, maxgrau_?(IntegerQ[#1] && #1 ≥ 0 &)] :=
 Module[{grau = RandomInteger[{0, maxgrau}], coefs},
  coefs = RandomInteger[{-9, 9}, grau + 1];
  {grau, coefs, Sum[coefs[[k + 1]] MatrixPower[mat, k], {k, 0, grau}]}]
```

- **A função** `MatAleatoria` **gera uma matriz quadrada de ordem aleatória entre** 2 **e** 5 **com elementos inteiros aleatórios entre** -9 **e** 9

```
MatAleatoria[] :=
 Module[{n = RandomInteger[{2, 5}]}, RandomInteger[{-9, 9}, {n, n}]]
```

a é uma matriz quadrada aleatória

```
(a = MatAleatoria[]) // MatrixForm
```

$$\begin{pmatrix} -1 & 9 \\ 3 & -2 \end{pmatrix}$$

b é a matriz polinómio em a com grau `gr` até 4 e coeficientes `cf`

```
{gr, cf, b} = PolAleatorio[a, 4];
```

O grau, os coeficientes e a matriz polinómio em a gerada anteriormente

```
{gr, cf, b // MatrixForm}
```

$$\left\{2, \{4, -1, -3\}, \begin{pmatrix} -79 & 72 \\ 24 & -87 \end{pmatrix}\right\}$$

Verificar que a e b são permutáveis

```
(p = a.b) == (q = b.a)
```

```
True
```

```
Map[MatrixForm, {p, q}]
```

$$\left\{\begin{pmatrix} 295 & -855 \\ -285 & 390 \end{pmatrix}, \begin{pmatrix} 295 & -855 \\ -285 & 390 \end{pmatrix}\right\}$$

Verificação de que $(a.b)^4 = a^4.b^4$

```
(p = MatrixPower[a.b, 4]) == (q = MatrixPower[a, 4].MatrixPower[b, 4])
```

```
True
```

```
p // MatrixForm
```

$$\begin{pmatrix} 47\,152\,726\,420 & -71\,836\,014\,105 & -32\,930\,742\,250 \\ 6\,191\,942\,216 & 11\,031\,472\,641 & 28\,865\,807\,720 \\ 46\,472\,670\,946 & -37\,496\,710\,654 & 27\,669\,372\,445 \end{pmatrix}$$

```
q // MatrixForm
```

$$\begin{pmatrix} 47152726420 & -71836014105 & -32930742250 \\ 6191942216 & 11031472641 & 28865807720 \\ 46472670946 & -37496710654 & 27669372445 \end{pmatrix}$$

Binómio de Newton: $(a + b)^3 = a^3 + 3a^2 b + 3ab^2 + b^3$

```
(p = MatrixPower[a + b, 3]) ==
 (q = MatrixPower[a, 3] + 3 MatrixPower[a, 2].b +
    3 a.MatrixPower[b, 2] + MatrixPower[b, 3])
```

```
True
```

```
p // MatrixForm
```

$$\begin{pmatrix} 109719 & -147917 & -116298 \\ 34776 & 32712 & 55328 \\ 116074 & -126686 & 49876 \end{pmatrix}$$

```
q // MatrixForm
```

$$\begin{pmatrix} 109719 & -147917 & -116298 \\ 34776 & 32712 & 55328 \\ 116074 & -126686 & 49876 \end{pmatrix}$$

Verificação de que $a^2 - b^2 = (a + b).(a - b)$

```
(p = MatrixPower[a, 2] - MatrixPower[b, 2]) == (q = (a + b).(a - b))
```

```
True
```

```
p // MatrixForm
```

$$\begin{pmatrix} 31 & -3513 & -3022 \\ 944 & -1812 & 1292 \\ 2866 & -3274 & -1496 \end{pmatrix}$$

```
q // MatrixForm
```

$$\begin{pmatrix} 31 & -3513 & -3022 \\ 944 & -1812 & 1292 \\ 2866 & -3274 & -1496 \end{pmatrix}$$

2.16.8. Matrizes não permutáveis

Gerar nova matriz b aleatória e da mesma ordem que a matriz a

```
(b = RandomInteger[{-9, 9}, {Length[a], Length[a]}]) // MatrixForm
```

$$\begin{pmatrix} -6 & 2 & 8 \\ -7 & 3 & 9 \\ 2 & -9 & -8 \end{pmatrix}$$

Verificar que a e b não são permutáveis

```
(p = a.b) == (q = b.a)
```

```
False
```

```
p // MatrixForm
```

$$\begin{pmatrix} 25 & -77 & -75 \\ 21 & -23 & -39 \\ -34 & -25 & 16 \end{pmatrix}$$

```
q // MatrixForm
```

$$\begin{pmatrix} -36 & 84 & -4 \\ -44 & 96 & -5 \\ 56 & -79 & -42 \end{pmatrix}$$

Verificação de que $(a.b)^4 \neq a^4.b^4$

```
(p = MatrixPower[a.b, 4]) == (q = MatrixPower[a, 4].MatrixPower[b, 4])
```

```
False
```

```
p // MatrixForm
```

$$\begin{pmatrix} 4\,919\,860 & 2\,285\,749 & -2\,625\,150 \\ 4\,475\,298 & -1\,269\,389 & -4\,874\,232 \\ -6\,424\,717 & 6\,942\,125 & 10\,797\,643 \end{pmatrix}$$

```
q // MatrixForm
```

$$\begin{pmatrix} -7\,198\,325 & -863\,315 & 6\,766\,995 \\ -2\,205\,775 & 13\,419\,573 & 11\,763\,441 \\ -1\,075\,200 & 20\,772\,113 & 15\,810\,796 \end{pmatrix}$$

Binómio de Newton não se verifica: $(a + b)^3 \neq a^3 + 3a^2b + 3ab^2 + b^3$

```
(p = MatrixPower[a + b, 3]) ==
 (q = MatrixPower[a, 3] + 3 MatrixPower[a, 2].b +
    3 a.MatrixPower[b, 2] + MatrixPower[b, 3])
```

```
False
```

```
p // MatrixForm
```

$$\begin{pmatrix} 68 & 32 & 625 \\ -176 & 229 & 671 \\ -33 & -13 & 171 \end{pmatrix}$$

```
q // MatrixForm
```

$$\begin{pmatrix} -328 & 1770 & 1320 \\ -512 & 2323 & 1874 \\ 164 & -708 & -1527 \end{pmatrix}$$

Verificação de que $a^2 - b^2 \neq (a + b).(a - b)$

```
(p = MatrixPower[a, 2] - MatrixPower[b, 2]) == (q = (a + b).(a - b))
```

```
False
```

```
p // MatrixForm
```

$$\begin{pmatrix} -38 & 133 & 144 \\ -55 & 115 & 81 \\ -81 & 5 & 26 \end{pmatrix}$$

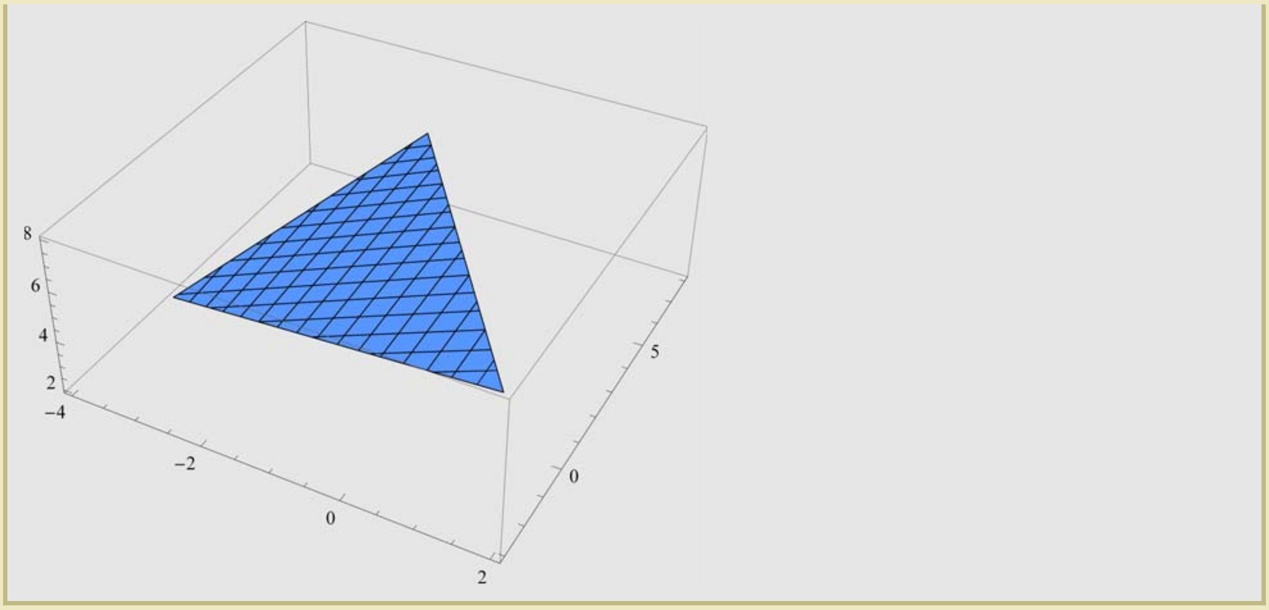
```
q // MatrixForm
```

$$\begin{pmatrix} -99 & 294 & 215 \\ -120 & 234 & 115 \\ 9 & -49 & -32 \end{pmatrix}$$

2.16.9. Representação gráfica

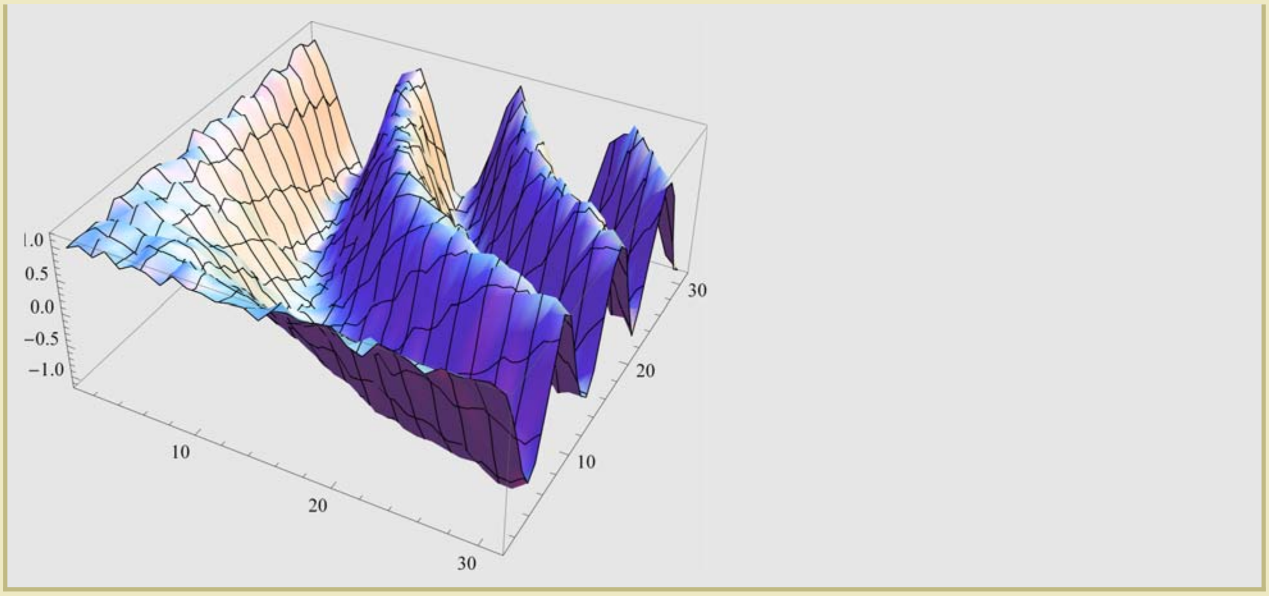
Podemos representar matrizes geometricamente em 3D

```
ListPlot3D[a]
```



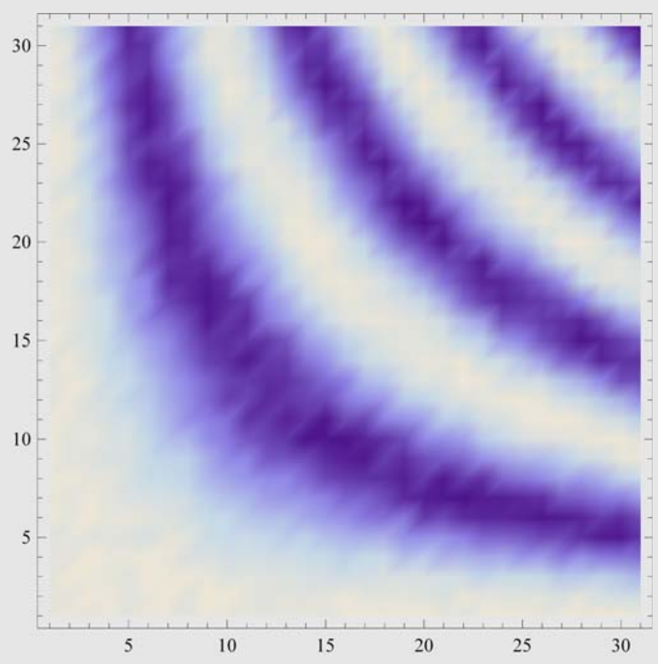

```
b = Table[Cos[x y] + RandomReal[{-0.15`, 0.15`}],
    {x, 0, 3π/2, π/20}, {y, 0, 3π/2, π/20}];
```

```
ListPlot3D[b]
```



E podemos representar matrizes geometricamente em 2D

```
ListDensityPlot[b, Mesh → False]
```



2.16.10. Característica

```
Needs["ALGA`Matrizes`"]
```

```
(a = {{1, -2, 1, 0, 3}, {0, 1, 3, 4, 2}, {-1, 1, -1, -1, -2}}) //
 MatrixForm
```

$$\begin{pmatrix} 1 & -2 & 1 & 0 & 3 \\ 0 & 1 & 3 & 4 & 2 \\ -1 & 1 & -1 & -1 & -2 \end{pmatrix}$$

Cálculo da característica

```
?Caracteristica
```

Caracteristica[A, PivotUm –> Opção] Devolve a lista composta pela característica de A, seguida da matriz levada à forma escalonada. PivotUm é uma opção que indica que os pivots deverão ser 1s, quando PivotUm for igual a True. Por defeito é PivotUm –> False.

```
{MatrixRank[a], Caracteristica[a][[1]]}
```

```
3
```

```
Remove["Global`*"]
```

2.16.11. Sistemas

Sistema determinado

```
(a = {{1, 2, 2}, {3, -2, -1}, {2, -5, 3}, {1, 4, 6}}) // MatrixForm
```

$$\begin{pmatrix} 1 & 2 & 2 \\ 3 & -2 & -1 \\ 2 & -5 & 3 \\ 1 & 4 & 6 \end{pmatrix}$$

```
b = {2, 5, -4, 0};
```

Uso de RowReduce[]

```
(ab = Transpose[Join[Transpose[a], {b}]]) // MatrixForm
```

$$\begin{pmatrix} 1 & 2 & 2 & 2 \\ 3 & -2 & -1 & 5 \\ 2 & -5 & 3 & -4 \\ 1 & 4 & 6 & 0 \end{pmatrix}$$

```
RowReduce[ab] // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Uso de `GaussJordan[]`

```
?GaussJordan
```

GaussJordan[A,B] Resolve o sistema A.X = B, devolvendo uma lista formada por uma solução particular e pela lista contendo os vectores duma base do espaço das soluções do sistema homogéneo A.X = O.

```
solu = GaussJordan[a, b]
```

```
{{2, 1, -1}, {}}
```

```
(sp = solu[[1]]) // MatrixForm
```

$$\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

Verificação do resultado

```
a.sp == b
```

```
True
```

Uso de `LinearSolve[]` e `NullSpace[]`

```
(sp = LinearSolve[a, b]) // MatrixForm
```

$$\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

```
NullSpace[a]
```

```
{}
```

Uso de Solve[]

```
X = {x, y, z};
```

```
sistema = a.X == b
```

```
{x + 2 y + 2 z, 3 x - 2 y - z, 2 x - 5 y + 3 z, x + 4 y + 6 z} == {2, 5, -4, 0}
```

```
Solve[sistema, X]
```

```
{{x -> 2, y -> 1, z -> -1}}
```

Uso de Reduce[]

```
Reduce[sistema, X]
```

```
x == 2 && y == 1 && z == -1
```

```
Remove["Global`*"]
```

Sistema impossível

```
(a = {{1, 5, 4, -13}, {3, -1, 2, 5}, {2, 2, 3, -4}}) // MatrixForm
```

$$\begin{pmatrix} 1 & 5 & 4 & -13 \\ 3 & -1 & 2 & 5 \\ 2 & 2 & 3 & -4 \end{pmatrix}$$

```
b = {3, 2, 1};
```

Uso de RowReduce[]

```
(ab = Transpose[Join[Transpose[a], {b}]]) // MatrixForm
```

$$\begin{pmatrix} 1 & 5 & 4 & -13 & 3 \\ 3 & -1 & 2 & 5 & 2 \\ 2 & 2 & 3 & -4 & 1 \end{pmatrix}$$

```
RowReduce[ab] // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & \frac{7}{8} & \frac{3}{4} & 0 \\ 0 & 1 & \frac{5}{8} & -\frac{11}{4} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Uso de GaussJordan[]

```
solu = GaussJordan[a, b]
```

GaussJordan::impossivel : Sistema A.X = B impossível.

GaussJordan[{{1, 5, 4, -13}, {3, -1, 2, 5}, {2, 2, 3, -4}}, {3, 2, 1}]

Uso de LinearSolve[]

```
(sp = LinearSolve[a, b]) // MatrixForm
```

LinearSolve::nosol : Linear equation encountered that has no solution. >>

LinearSolve[{{1, 5, 4, -13}, {3, -1, 2, 5}, {2, 2, 3, -4}}, {3, 2, 1}]

Uso de Solve[]

```
X = {x, y, z, w};
```

```
sistema = a.X == b
```

```
{-13 w + x + 5 y + 4 z, 5 w + 3 x - y + 2 z, -4 w + 2 x + 2 y + 3 z} == {3, 2, 1}
```

```
Solve[sistema, X]
```

```
{}
```

Uso de Reduce[]

```
Reduce[sistema, X]
```

```
False
```

```
Remove["Global`*"]
```

- **Sistema duplamente indeterminado**

```
(a = {{1, 2, -3, 2}, {2, 5, -8, 6}, {3, 4, -5, 2}}) // MatrixForm
```

$$\begin{pmatrix} 1 & 2 & -3 & 2 \\ 2 & 5 & -8 & 6 \\ 3 & 4 & -5 & 2 \end{pmatrix}$$

```
b = {2, 5, 4};
```

Uso de RowReduce[]

```
(ab = Transpose[Join[Transpose[a], {b}]]) // MatrixForm
```

$$\begin{pmatrix} 1 & 2 & -3 & 2 & 2 \\ 2 & 5 & -8 & 6 & 5 \\ 3 & 4 & -5 & 2 & 4 \end{pmatrix}$$

```
RowReduce[ab] // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 1 & -2 & 0 \\ 0 & 1 & -2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Uso de `GaussJordan[]`

```
solu = GaussJordan[a, b]
```

```
{{0, 1, 0, 0}, {{-1, 2, 1, 0}, {2, -2, 0, 1}}}
```

```
(sp = solu[[1]]) // MatrixForm
```

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

```
(sh = solu[[2]]) // MatrixForm
```

$$\begin{pmatrix} -1 & 2 & 1 & 0 \\ 2 & -2 & 0 & 1 \end{pmatrix}$$

```
Remove[c]
```

```
(sg = sp + Array[c, {Length[sh]}].sh) // MatrixForm
```

```
( -c[1] + 2 c[2]         )
( 1 + 2 c[1] - 2 c[2]    )
( c[1]                   )
( c[2]                   )
```

Verificação do resultado

```
Simplify[a.sg - b]
```

```
{0, 0, 0}
```

4 soluções particulares do sistema, para valores aleatórios inteiros dos parâmetros `c[i]` entre `-6` e `6`

```
Table[sg /. Table[c[i] → RandomInteger[{-6, 6}], {i, 2}], {4}]
```

```
{{6, -3, 2, 4}, {11, -13, -3, 4}, {-4, 11, 6, 1}, {-6, 11, 4, -1}}
```

Uso de `LinearSolve[]` e `NullSpace[]`

```
(sp = LinearSolve[a, b]) // MatrixForm
```

```
( 0 )
( 1 )
( 0 )
( 0 )
```

```
sh = NullSpace[a]
```

```
{{2, -2, 0, 1}, {-1, 2, 1, 0}}
```

```
(sg = sp + Array[c, {Length[sh]}].sh) // MatrixForm
```

$$\begin{pmatrix} 2\,c[1] - c[2] \\ 1 - 2\,c[1] + 2\,c[2] \\ c[2] \\ c[1] \end{pmatrix}$$

Uso de `Solve[]`

```
X = {x, y, z, w};
```

```
sistema = a.X == b
```

```
{2 w + x + 2 y - 3 z, 6 w + 2 x + 5 y - 8 z, 2 w + 3 x + 4 y - 5 z} == {2, 5, 4}
```

Resolver em ordem a x e y

```
Solve[sistema, X[[{1, 2}]]]
```

```
{{x → 2 w - z, y → 1 - 2 w + 2 z}}
```

Uso de `Reduce[]`

```
Reduce[sistema, X]
```

$z == -1 + x + y \;\&\&\; w == -\frac{1}{2} + x + \frac{y}{2}$

```
Remove["Global`*"]
```

- **Discussão de um sistema: caso 1**

```
(a = {{1, 1, 1, 3}, {2, 2, 1, 1}, {1, 1, 3, 14}, {0, 0, 2, u}}) //
 MatrixForm
```

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 3 & 14 \\ 0 & 0 & 2 & u \end{pmatrix}$$

```
b = {1, 2, 4, v};
```

```
X = {x, y, z, w};
```

```
sistema = a.X == b
```

```
{3 w + x + y + z, w + 2 x + 2 y + z, 14 w + x + y + 3 z, u w + 2 z} == {1, 2, 4, v}
```

```
Reduce[sistema, X]
```

$$u == \frac{30 + v}{3} \ \&\& \ y == 7 - x \ \&\& \ z == -15 \ \&\& \ w == 3$$

```
Remove["Global`*"]
```

- **Discussão de um sistema: caso 2**

```
(a = {{u^2, 1, u}, {u^2 - v^2, 1, u + 1}, {3 u^2 - 2 v^2, 2, 2 u + 1}}) //
 MatrixForm
```

$$\begin{pmatrix} u^2 & 1 & u \\ u^2 - v^2 & 1 & 1 + u \\ 3 u^2 - 2 v^2 & 2 & 1 + 2 u \end{pmatrix}$$

```
b = {v - u, v, 3 v};
```

```
X = {x, y, z};
```

```
sistema = a.X == b
```

$\{u^2 x + y + u z, (u^2 - v^2) x + y + (1 + u) z, (3 u^2 - 2 v^2) x + 2 y + (1 + 2 u) z\} == \{-u + v, v, 3 v\}$

```
Reduce[sistema, X]
```

$(u == -v \&\& y == 2 v - v^2 - v^2 x + v^3 x \&\& z == -v + v^2 x) \mid\mid$
$\left((u - v) (u + v) \neq 0 \&\& x == \frac{1}{u - v} \&\& y == -2 u - u^2 - v^2 x - u v^2 x \&\& z == u + v^2 x\right)$

```
Remove["Global`*"]
```

2.16.12. Método de Gauss-Seidel

- `GaussSeidel[a, b, x0, tol,max]` **determina a solução do sistema** `ax = b`**, a partir da estimativa inicial** `x0` **com a tolerancia** `tol` **e calculando um máximo de** `max` **pontos e devolve a lista formada pela solução aproximada e pelo número de pontos calculados.**
- `GaussSeidelPoints[a, b, x0, tol,max]` **determina a lista das soluções aproximadas do sistema** `ax = b`**, a partir da estimativa inicial** `x0` **com a tolerancia** `tol` **e com um máximo de** `max+1` **pontos.**

Primeiro exemplo: uso da função `GaussSeidel`.

```
?GaussSeidel
```

GaussSeidel[a, b, x0, err, iter] tenta resolver o sistema ax=b pelo método de Gauss–Seidel, a partir de um ponto inicial x0 com um erro inferior a err e realizando um máximo de iter iterações.

```
(a = {{6, -2, 3}, {2, 4, -1}, {4, 2, 7}}) // MatrixForm
```

$$\begin{pmatrix} 6 & -2 & 3 \\ 2 & 4 & -1 \\ 4 & 2 & 7 \end{pmatrix}$$

```
b = {10, 4, 20}
```

```
{10, 4, 20}
```

```
LinearSolve[a, b]
```

```
{1, 1, 2}
```

```
x0 = {0, 0, 0}; ϵ = 10^(-6);
```

```
GaussSeidel[a, b, x0, ϵ, 100]
```

$$\left\{\left\{\frac{1\,076\,052\,546\,519\,130\,475}{1\,076\,052\,481\,665\,702\,912}, \frac{4\,304\,209\,768\,276\,963\,961}{4\,304\,209\,926\,662\,811\,648}, \frac{10\,043\,156\,375\,399\,369\,573}{5\,021\,578\,247\,773\,280\,256}\right\}, 36\right\}$$

```
N[%[[1]], 6]
```

```
{1.00000, 1.00000, 2.00000}
```

Segundo exemplo: uso das funções `GaussSeidel` e `GaussSeidelPoints` e interpretação geométrica.

```
?GaussSeidelPoints
```

GaussSeidelPoints[a, b, x0, err, iter] devolve a lista de pontos obtida pelo algoritmo de Gauss–Seidel na resolução do sistema ax=b, a partir de um ponto inicial x0 com um erro inferior a err e realizando um máximo de iter iterações.

```
(a = {{3, -2}, {3, 4}}) // MatrixForm
```

$$\begin{pmatrix} 3 & -2 \\ 3 & 4 \end{pmatrix}$$

```
b = {-1, 11}
```

```
{-1, 11}
```

```
LinearSolve[a, b]
```

```
{1, 2}
```

```
x0 = {0, 1}; ϵ = 10^(-4);
```

```
GaussSeidel[a, b, x0, ϵ, 100]
```

$$\left\{\left\{\frac{49152}{49152}, \frac{131071}{65536}\right\}, 32\right\}$$

```
N[%[[1]], 6]
```

```
{1.00002, 1.99998}
```

```
pontos = GaussSeidelPoints[a, b, x0, ϵ, 100]
```

$$\left\{\{0,1\},\left\{\frac{1}{3},1\right\},\left\{\frac{1}{3},\frac{5}{2}\right\},\left\{\frac{4}{3},\frac{5}{2}\right\},\left\{\frac{4}{3},\frac{7}{4}\right\},\left\{\frac{5}{6},\frac{7}{4}\right\},\left\{\frac{5}{6},\frac{17}{8}\right\},\left\{\frac{13}{12},\frac{17}{8}\right\},\left\{\frac{13}{12},\frac{31}{16}\right\},\left\{\frac{23}{24},\frac{31}{16}\right\},\right.$$
$$\left\{\frac{23}{24},\frac{65}{32}\right\},\left\{\frac{49}{48},\frac{65}{32}\right\},\left\{\frac{49}{48},\frac{127}{64}\right\},\left\{\frac{95}{96},\frac{127}{64}\right\},\left\{\frac{95}{96},\frac{257}{128}\right\},\left\{\frac{193}{192},\frac{257}{128}\right\},\left\{\frac{193}{192},\frac{511}{256}\right\},$$
$$\left\{\frac{383}{384},\frac{511}{256}\right\},\left\{\frac{383}{384},\frac{1025}{512}\right\},\left\{\frac{769}{768},\frac{1025}{512}\right\},\left\{\frac{769}{768},\frac{2047}{1024}\right\},\left\{\frac{1535}{1536},\frac{2047}{1024}\right\},\left\{\frac{1535}{1536},\frac{4097}{2048}\right\},$$
$$\left\{\frac{3073}{3072},\frac{4097}{2048}\right\},\left\{\frac{3073}{3072},\frac{8191}{4096}\right\},\left\{\frac{6143}{6144},\frac{8191}{4096}\right\},\left\{\frac{6143}{6144},\frac{16\,385}{8192}\right\},\left\{\frac{12\,289}{12\,288},\frac{16\,385}{8192}\right\},$$
$$\left.\left\{\frac{12\,289}{12\,288},\frac{32\,767}{16\,384}\right\},\left\{\frac{24\,575}{24\,576},\frac{32\,767}{16\,384}\right\},\left\{\frac{24\,575}{24\,576},\frac{65\,537}{32\,768}\right\},\left\{\frac{49\,153}{49\,152},\frac{65\,537}{32\,768}\right\},\left\{\frac{49\,153}{49\,152},\frac{131\,071}{65\,536}\right\}\right\}$$

```
Length[pontos]
```

```
33
```

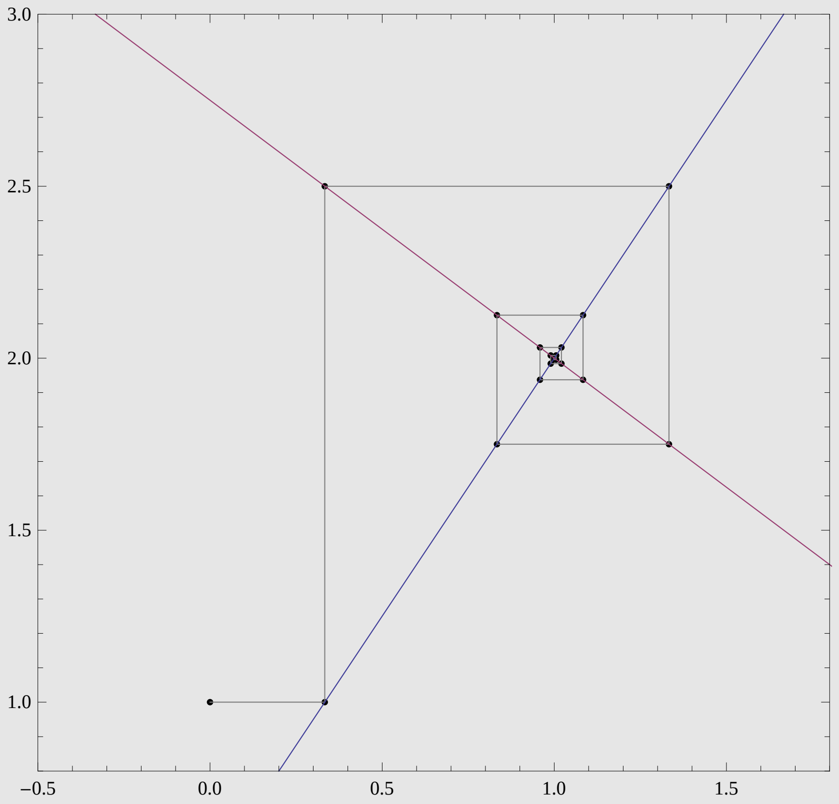
```
equacoestraj = Simplify[Table[
    (1 - t) * pontos[[i]] + t * pontos[[i + 1]], {i, Length[pontos] - 1}]];
```

```
sopontos = Graphics[ListPlot[pontos,
    PlotRange → {{-1 / 2, 9 / 5}, {4 / 5, 3}}, AspectRatio → Automatic,
    Frame -> True, PlotStyle → Black, Axes → False]];
```

```
trajectoria = Graphics[ParametricPlot[equacoestraj, {t, 0, 1},
    PlotRange → {{-1 / 2, 9 / 5}, {4 / 5, 3}}, AspectRatio → Automatic,
    Frame -> True, PlotStyle → Gray, Axes → False]];
```

```
rectas = Graphics[Plot[{(3 x + 1) / 2, (11 - 3 x) / 4},
    {x, -1 / 2, 2}, PlotRange → {{-1 / 2, 9 / 5}, {4 / 5, 3}},
    AspectRatio → Automatic, Frame -> True, Axes → False]];
```

```
Show[sopontos, trajectoria, rectas]
```



Terceiro exemplo: uso das funções `GaussSeidel` e `GaussSeidelPoints` e interpretação geométrica.

```
(a = {{-3, 2}, {1, -2}}) // MatrixForm
```

$$\begin{pmatrix} -3 & 2 \\ 1 & -2 \end{pmatrix}$$

```
b = {-1, -1}
```

```
{-1, -1}
```

```
LinearSolve[a, b]
```

```
{1, 1}
```

```
x0 = {1, 0}; ϵ = 10^(-6);
```

```
GaussSeidel[a, b, x0, ϵ, 100]
```

$$\left\{\left\{\frac{4\,782\,967}{4\,782\,969}, \frac{4\,782\,968}{4\,782\,969}\right\}, 28\right\}$$

```
N[%[[1]], 6]
```

```
{1.00000, 1.00000}
```

```
pontos = GaussSeidelPoints[a, b, x0, ϵ, 100]
```

$$\Big\{\{1, 0\}, \left\{\frac{1}{3}, 0\right\}, \left\{\frac{1}{3}, \frac{2}{3}\right\}, \left\{\frac{7}{9}, \frac{2}{3}\right\}, \left\{\frac{7}{9}, \frac{8}{9}\right\}, \left\{\frac{25}{27}, \frac{8}{9}\right\}, \left\{\frac{25}{27}, \frac{26}{27}\right\}, \left\{\frac{79}{81}, \frac{26}{27}\right\}, \left\{\frac{79}{81}, \frac{80}{81}\right\}, \left\{\frac{241}{243}, \frac{80}{81}\right\},$$
$$\left\{\frac{241}{243}, \frac{242}{243}\right\}, \left\{\frac{727}{729}, \frac{242}{243}\right\}, \left\{\frac{727}{729}, \frac{728}{729}\right\}, \left\{\frac{2185}{2187}, \frac{728}{729}\right\}, \left\{\frac{2185}{2187}, \frac{2186}{2187}\right\}, \left\{\frac{6559}{6561}, \frac{2186}{2187}\right\},$$
$$\left\{\frac{6559}{6561}, \frac{6560}{6561}\right\}, \left\{\frac{19\,681}{19\,683}, \frac{6560}{6561}\right\}, \left\{\frac{19\,681}{19\,683}, \frac{19\,682}{19\,683}\right\}, \left\{\frac{59\,047}{59\,049}, \frac{19\,682}{19\,683}\right\}, \left\{\frac{59\,047}{59\,049}, \frac{59\,048}{59\,049}\right\},$$
$$\left\{\frac{177\,145}{177\,147}, \frac{59\,048}{59\,049}\right\}, \left\{\frac{177\,145}{177\,147}, \frac{177\,146}{177\,147}\right\}, \left\{\frac{531\,439}{531\,441}, \frac{177\,146}{177\,147}\right\}, \left\{\frac{531\,439}{531\,441}, \frac{531\,440}{531\,441}\right\},$$
$$\left\{\frac{1\,594\,321}{1\,594\,323}, \frac{531\,440}{531\,441}\right\}, \left\{\frac{1\,594\,321}{1\,594\,323}, \frac{1\,594\,322}{1\,594\,323}\right\}, \left\{\frac{4\,782\,967}{4\,782\,969}, \frac{1\,594\,322}{1\,594\,323}\right\}, \left\{\frac{4\,782\,967}{4\,782\,969}, \frac{4\,782\,968}{4\,782\,969}\right\}\Big\}$$

```
Length[pontos]
```

```
29
```

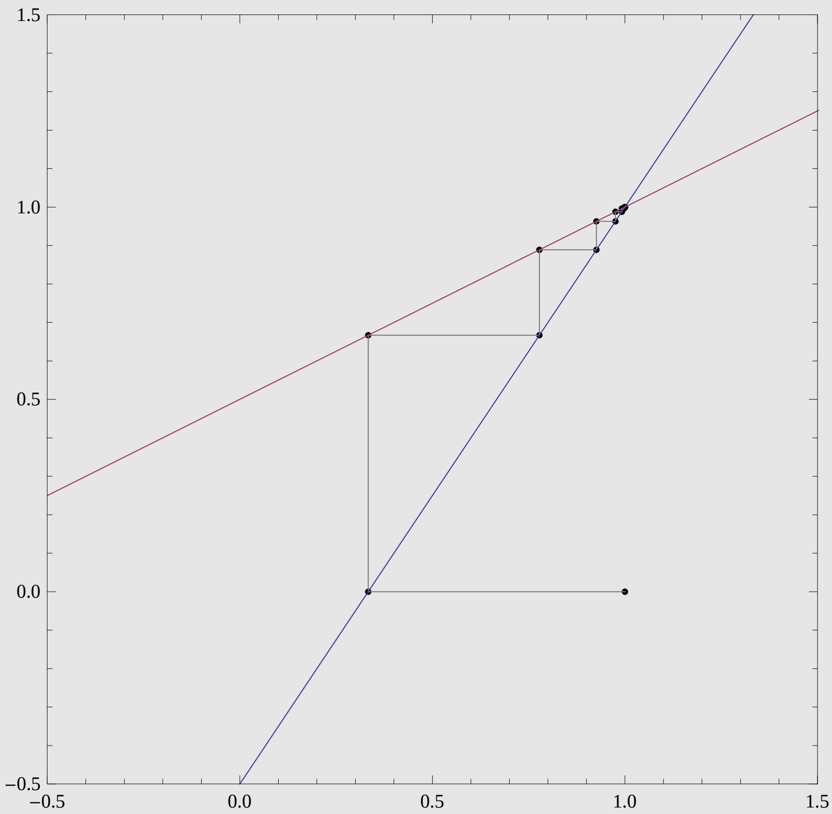
```
equacoestraj = Simplify[Table[
    (1 - t) * pontos[[i]] + t * pontos[[i + 1]], {i, Length[pontos] - 1}]];
```

```
sopontos =
  Graphics[ListPlot[pontos, PlotRange → {{-1 / 2, 3 / 2}, {-1 / 2, 3 / 2}},
    AspectRatio → Automatic, Frame -> True,
    PlotStyle → Black, Axes → False]];
```

```
trajectoria =
  Graphics[ParametricPlot[equacoestraj, {t, 0, 1}, PlotRange →
     {{-1 / 2, 3 / 2}, {-1 / 2, 3 / 2}}, AspectRatio → Automatic,
    Frame -> True, PlotStyle → Gray, Axes → False]];
```

```
rectas = Graphics[Plot[{(3 x - 1) / 2, (x + 1) / 2},
    {x, -1 / 2, 2}, PlotRange → {{-1 / 2, 3 / 2}, {-1 / 2, 3 / 2}},
    AspectRatio → Automatic, Frame -> True, Axes → False]];
```

```
Show[sopontos, trajectoria, rectas]
```



2.16.13. Inversão de matrizes rectangulares

Inversas direitas

```
?InversaDireita
```

InversaDireita[A,c] calcula a(s) matriz(es) inversa(s) direita(s) da matriz A e usando c como parâmetro.

```
(a = {{1, -2, 1, 0, 3}, {0, 1, 3, 4, 2}, {-1, 1, -1, -1, -2}}) //
 MatrixForm
```

$$\begin{pmatrix} 1 & -2 & 1 & 0 & 3 \\ 0 & 1 & 3 & 4 & 2 \\ -1 & 1 & -1 & -1 & -2 \end{pmatrix}$$

Cálculo das inversas direitas de a

```
(x = InversaDireita[a, c]) // MatrixForm
```

$$\begin{pmatrix} -\frac{4}{3} - c[1][2] & -\frac{1}{3} - c[2][2] & -\frac{7}{3} - c[3][2] \\ -1 + c[1][1] - c[1][2] & c[2][1] - c[2][2] & -1 + c[3][1] - c[3][2] \\ \frac{1}{3} - c[1][1] - c[1][2] & \frac{1}{3} - c[2][1] - c[2][2] & \frac{1}{3} - c[3][1] - c[3][2] \\ c[1][2] & c[2][2] & c[3][2] \\ c[1][1] & c[2][1] & c[3][1] \end{pmatrix}$$

Verificação do cálculo anterior

```
Simplify[a.x] // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Exemplo de 4 matrizes inversas direitas para valores aleatórios (inteiros entre -9 e 9) dos parâmetros c[i][j]

```
Table[x /. Flatten[Table[c[i][j] → RandomInteger[{-9, 9}],
    {i, 3}, {j, 2}]] // MatrixForm, {4}]
```

$$\left\{\begin{pmatrix} -\frac{25}{3} & -\frac{4}{3} & -\frac{13}{3} \\ -7 & 8 & -12 \\ -\frac{23}{3} & -\frac{29}{3} & \frac{22}{3} \\ 7 & 1 & 2 \\ 1 & 9 & -9 \end{pmatrix}, \begin{pmatrix} -\frac{4}{3} & -\frac{4}{3} & \frac{2}{3} \\ 5 & -5 & -3 \\ -\frac{17}{3} & \frac{10}{3} & \frac{25}{3} \\ 0 & 1 & -3 \\ 6 & -4 & -5 \end{pmatrix}, \begin{pmatrix} -\frac{13}{3} & -\frac{1}{3} & -\frac{22}{3} \\ -13 & -5 & -14 \\ \frac{19}{3} & \frac{16}{3} & \frac{10}{3} \\ 3 & 0 & 5 \\ -9 & -5 & -8 \end{pmatrix}, \begin{pmatrix} -\frac{16}{3} & -\frac{19}{3} & \frac{20}{3} \\ -11 & -2 & 0 \\ \frac{7}{3} & -\frac{29}{3} & \frac{52}{3} \\ 4 & 6 & -9 \\ -6 & 4 & -8 \end{pmatrix}\right\}$$

Cálculo das inversas esquerdas de a (não existem)

```
? InversaEsquerda
```

InversaEsquerda[A,c] calcula a(s) matriz(es) inversa(s) esquerda(s) da matriz A e usando c como parâmetro.

```
InversaEsquerda[a, c]
```

InversaEsquerda::naoexiste : A matriz não tem inversa(s) esquerda(s).

InversaEsquerda[{{1, -2, 1, 0, 3}, {0, 1, 3, 4, 2}, {-1, 1, -1, -1, -2}}, c]

Soluções de um sistema a.x == b (possível e indeterminado), usando as inversas direitas de a

```
b = a.{2, -1, 3, 2, -1}
```

{4, 14, -6}

```
Simplify[x.b] // MatrixForm
```

$$\begin{pmatrix} 4 - 4\,c[1][2] - 14\,c[2][2] + 6\,c[3][2] \\ 2\,(1 + 2\,c[1][1] - 2\,c[1][2] + 7\,c[2][1] - 7\,c[2][2] - 3\,c[3][1] + 3\,c[3][2]) \\ -2\,(-2 + 2\,c[1][1] + 2\,c[1][2] + 7\,c[2][1] + 7\,c[2][2] - 3\,c[3][1] - 3\,c[3][2]) \\ 4\,c[1][2] + 14\,c[2][2] - 6\,c[3][2] \\ 4\,c[1][1] + 14\,c[2][1] - 6\,c[3][1] \end{pmatrix}$$

Uso de `GaussJordan[]` para o mesmo efeito

```
Remove[k]
```

```
m = GaussJordan[a, b];
```

```
m[[1]] + Array[k, {Length[m[[2]]]}].m[[2]] // MatrixForm
```

$$\begin{pmatrix} 4 - k[1] \\ 2 - k[1] + k[2] \\ 4 - k[1] - k[2] \\ k[1] \\ k[2] \end{pmatrix}$$

- **Inversas esquerdas**

```
(a = {{1, 0, -1}, {-2, 1, 1}, {1, 3, -1}, {0, 4, -1}, {3, 2, -2}}) //
 MatrixForm
```

$$\begin{pmatrix} 1 & 0 & -1 \\ -2 & 1 & 1 \\ 1 & 3 & -1 \\ 0 & 4 & -1 \\ 3 & 2 & -2 \end{pmatrix}$$

Cálculo da característica

```
Caracteristica[a][[1]]
```

```
3
```

Cálculo das inversas esquerdas y de a

```
(y = InversaEsquerda[a, c]) // MatrixForm
```

$$\begin{pmatrix} -\frac{4}{3}-c[1][2] & -1+c[1][1]-c[1][2] & \frac{1}{3}-c[1][1]-c[1][2] & c[1][2] & c[1][1] \\ -\frac{1}{3}-c[2][2] & c[2][1]-c[2][2] & \frac{1}{3}-c[2][1]-c[2][2] & c[2][2] & c[2][1] \\ -\frac{7}{3}-c[3][2] & -1+c[3][1]-c[3][2] & \frac{1}{3}-c[3][1]-c[3][2] & c[3][2] & c[3][1] \end{pmatrix}$$

Verificação do cálculo anterior

```
Simplify[y.a] // MatrixForm  (* verificação do cálculo anterior *)
```

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Exemplo de 3 matrizes inversas esquerdas para valores aleatórios (inteiros entre -9 e 9) dos parâmetros c[i][j]

```
Table[y /. Flatten[Table[c[i][j] → RandomInteger[{-9, 9}],
      {i, 3}, {j, 2}]] // MatrixForm, {3}]
```

$$\left\{\begin{pmatrix} -\frac{16}{3} & -9 & \frac{1}{3} & 4 & -4 \\ -\frac{28}{3} & -10 & -\frac{23}{3} & 9 & -1 \\ -\frac{25}{3} & -15 & \frac{7}{3} & 6 & -8 \end{pmatrix}, \begin{pmatrix} \frac{14}{3} & 1 & \frac{31}{3} & -6 & -4 \\ -\frac{28}{3} & -11 & -\frac{20}{3} & 9 & -2 \\ -\frac{16}{3} & -9 & \frac{7}{3} & 3 & -5 \end{pmatrix}, \begin{pmatrix} -\frac{1}{3} & -4 & \frac{16}{3} & -1 & -4 \\ \frac{8}{3} & 10 & -\frac{11}{3} & -3 & 7 \\ \frac{8}{3} & 3 & \frac{19}{3} & -5 & -1 \end{pmatrix}\right\}$$

Cálculo das inversas direitas de a (não existem)

```
InversaDireita[a, c]
```

InversaDireita::naoexiste : A matriz não tem inversa(s) direita(s).

InversaDireita[{{1, 0, -1}, {-2, 1, 1}, {1, 3, -1}, {0, 4, -1}, {3, 2, -2}}, c]

Sistema a.x == b possível e determinado com solução imposta {2,-3,1}

```
b = a.{2, -3, 1}
```

```
{1, -6, -8, -13, -2}
```

Solução do sistema a.x == b (possível e determinado), usando as inversas esquerdas y de a.
A solução não depende da inversa esquerda usada.

```
x = Simplify[y.b]
```

```
{2, -3, 1}
```

Sistema a.x == b impossível

```
b = {1, 1, 1, 1, 1};
```

```
m = GaussJordan[a, b]
```

```
GaussJordan::impossivel : Sistema A.X = B impossível.
```

```
GaussJordan[{{1, 0, -1}, {-2, 1, 1}, {1, 3, -1}, {0, 4, -1}, {3, 2, -2}}, {1, 1, 1, 1, 1}]
```

Neste caso, x = y.b não é solução de a.x == b

```
x = Simplify[y.b]
```

```
{c[1][1] - 2 (1 + c[1][2]), c[2][1] - 2 c[2][2], -3 + c[3][1] - 2 c[3][2]}
```

2.16.14. Inversão de matrizes quadradas

```
? Inversa
```

Inversa[A] calcula a matriz inversa da matriz quadrada A.

```
(a = {{-1, 2, -3}, {2, 1, 0}, {4, -2, 5}}) // MatrixForm
```

$$\begin{pmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{pmatrix}$$

Cálculo da inversa (bilateral)

```
(x = Inversa[a]) // MatrixForm
```

$$\begin{pmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{pmatrix}$$

Verificação do cálculo anterior

```
Simplify[a.x] // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

```
Simplify[x.a] // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Verificação das funções do package `` ALGA`Matrizes` ``, recorrendo à função `Inverse[]`

```
(y = Inverse[a]) // MatrixForm
```

$$\begin{pmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{pmatrix}$$

```
Inversa[a] == y
```

```
True
```

```
InversaDireita[a, c] == y
```

```
True
```

```
InversaEsquerda[a, c] == y
```

```
True
```

2.16.15. Divisão matricial

- **Divisão esquerda**

```
?DivisaoEsquerda
```

DivisaoEsquerda[A,B,c] calcula o quociente de B por A, com o divisor A à esquerda e usando c como parâmetro.

```
(a = {{-1, 2, -3}, {2, 1, 0}, {1, 3, -3}}) // MatrixForm
```

$$\begin{pmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 1 & 3 & -3 \end{pmatrix}$$

```
(b = a.{{2, -1}, {1, 1}, {-1, 3}}) // MatrixForm
```

$$\begin{pmatrix} 3 & -6 \\ 5 & -1 \\ 8 & -7 \end{pmatrix}$$

Divisão esquerda de b por a

```
(x = DivisaoEsquerda[a, b, c]) // MatrixForm
```

$$\begin{pmatrix} \frac{7}{5} - 3\,c[1][1] & \frac{4}{5} - 3\,c[2][1] \\ \frac{11}{5} + 6\,c[1][1] & -\frac{13}{5} + 6\,c[2][1] \\ 5\,c[1][1] & 5\,c[2][1] \end{pmatrix}$$

Verificação do resultado

```
Simplify[a.x == b]
```

```
True
```

4 soluções da divisão esquerda para valores aleatórios (inteiros entre -9 e 9) dos parâmetros c[i][1]

```
Table[x /. Table[c[i][1] → RandomInteger[{-9, 9}], {i, 2}] //
  MatrixForm, {4}]
```

$$\left\{ \begin{pmatrix} \frac{22}{5} & \frac{109}{5} \\ -\frac{19}{5} & -\frac{223}{5} \\ -5 & -35 \end{pmatrix}, \begin{pmatrix} -\frac{53}{5} & \frac{139}{5} \\ \frac{131}{5} & -\frac{283}{5} \\ 20 & -45 \end{pmatrix}, \begin{pmatrix} -\frac{38}{5} & -\frac{41}{5} \\ \frac{101}{5} & \frac{77}{5} \\ 15 & 15 \end{pmatrix}, \begin{pmatrix} \frac{112}{5} & -\frac{41}{5} \\ -\frac{199}{5} & \frac{77}{5} \\ -35 & 15 \end{pmatrix} \right\}$$

- **Divisão direita**

```
?DivisaoDireita
```

DivisaoDireita[A, B, c] calcula o quociente de B por A, com o divisor A à direita e usando c como parâmetro.

```
(a = {{-1, 2, -3, 1}, {2, 1, 0, -1}, {1, 3, -3, 0}}) // MatrixForm
```

$$\begin{pmatrix} -1 & 2 & -3 & 1 \\ 2 & 1 & 0 & -1 \\ 1 & 3 & -3 & 0 \end{pmatrix}$$

```
(b = {{2, -1, 1}, {1, 1, 3}}.a) // MatrixForm
```

$$\begin{pmatrix} -3 & 6 & -9 & 3 \\ 4 & 12 & -12 & 0 \end{pmatrix}$$

Divisão direita de b por a

```
(x = DivisaoDireita[a, b, c]) // MatrixForm
```

$$\begin{pmatrix} 3 - c[1][1] & -c[1][1] & c[1][1] \\ 4 - c[2][1] & 4 - c[2][1] & c[2][1] \end{pmatrix}$$

Verificação do resultado

```
Simplify[x.a == b]
```

```
True
```

Divisões com divisor a invertível e dividendo b não permutável com a

```
(a = {{-1, 2, -3}, {2, 1, 0}, {4, -2, 5}}) // MatrixForm
```

$$\begin{pmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{pmatrix}$$

```
(b = {{0, 1, -2}, {-1, 2, -3}, {2, -1, -3}}) // MatrixForm
```

$$\begin{pmatrix} 0 & 1 & -2 \\ -1 & 2 & -3 \\ 2 & -1 & -3 \end{pmatrix}$$

```
a.b == b.a
```

```
False
```

```
(x = DivisaoEsquerda[a, b, c]) // MatrixForm
```

$$\begin{pmatrix} -10 & 6 & 7 \\ 19 & -10 & -17 \\ 16 & -9 & -13 \end{pmatrix}$$

Verificação do resultado

```
a.x == b
```

```
True
```

```
(y = DivisaoDireita[a, b, c]) // MatrixForm
```

$$\begin{pmatrix} -6 & 5 & -4 \\ 1 & 0 & 0 \\ -44 & 33 & -27 \end{pmatrix}$$

Verificação do resultado

```
y.a == b
```

```
True
```

As divisões dão resultados diferentes

```
x == y
```

```
False
```

Divisões com divisor a invertível e dividendo b permutável com a

A matriz b é um polinómio em a; portanto, é permutável com a

```
(b = 2 * MatrixPower[a, 3] - MatrixPower[a, 2] +
    3 * MatrixPower[a, 1] - 4 MatrixPower[a, 0]) // MatrixForm
```

$$\begin{pmatrix} -58 & 32 & -75 \\ -22 & 28 & -54 \\ 64 & -14 & 56 \end{pmatrix}$$

Verificação de que a e b são permutáveis

```
a.b == b.a
```

```
True
```

As divisões esquerda e direita dão o mesmo resultado

```
(x = DivisaoEsquerda[a, b, c]) // MatrixForm
```

$$\begin{pmatrix} 10 & -6 & -9 \\ -42 & 40 & -36 \\ -12 & 18 & 4 \end{pmatrix}$$

```
(y = DivisaoDireita[a, b, c]) // MatrixForm
```

$$\begin{pmatrix} 10 & -6 & -9 \\ -42 & 40 & -36 \\ -12 & 18 & 4 \end{pmatrix}$$

As divisões dão resultados iguais

```
x == y
```

```
True
```

2.16.16. Mudanças de base

```
? Passagem
```

Passagem[e1, e2] Devolve a matriz cujas colunas são as coordenadas dos vectores da lista e2, na base formada pelos vectores da lista e1; se e2 for outra base, calcula a matriz de mudança de e1 para e2. A lista e1 deve ser uma base e os vectores de e2 devem pertencer ao espaço gerado por e1.

Definam-se duas bases de $\mathbb{R}^3$:

```
e1 = {{1, -2, 1}, {3, 0, -1}, {-2, 2, 3}}
```

```
{{1, -2, 1}, {3, 0, -1}, {-2, 2, 3}}
```

```
e2 = {{3, 0, 1}, {1, 1, 1}, {1, -3, 1}}
```

```
{{3, 0, 1}, {1, 1, 1}, {1, -3, 1}}
```

A matriz de mudança de e1 para e2 é:

```
(t12 = Passagem[e1, e2]) // MatrixForm
```

$$\begin{pmatrix} \frac{6}{11} & \frac{1}{22} & \frac{29}{22} \\ \frac{13}{11} & \frac{15}{22} & -\frac{5}{22} \\ \frac{6}{11} & \frac{6}{11} & -\frac{2}{11} \end{pmatrix}$$

A matriz de mudança de e2 para e1 é:

```
(t21 = Passagem[e2, e1]) // MatrixForm
```

$$\begin{pmatrix} 0 & 2 & -\frac{5}{2} \\ \frac{1}{4} & -\frac{9}{4} & \frac{37}{8} \\ \frac{3}{4} & -\frac{3}{4} & \frac{7}{8} \end{pmatrix}$$

Verificação dos resultados:

```
t12.t21 == IdentityMatrix[Length[e1]]
```

```
True
```

Defina-se um vector através das suas coordenadas na base e1:

```
xe1 = {4, -2, 5}
```

```
{4, -2, 5}
```

As suas coordenadas na base e2 são:

```
? Coordenadas
```

Coordenadas[e1, e2, xe1] Devolve as coordenadas em relação
à base e2 de um vector, dadas as suas coordenadas xe1 em relação à base e1.

```
Coordenadas[e1, e2, xe1]
```

$$\left\{-\frac{33}{2}, \frac{229}{8}, \frac{71}{8}\right\}$$

Definam-se duas bases de $\mathbb{C}^2$:

```
e1 = {{1 + I, -2 + 2 I}, {3 - 2 I, 5 I}}
```

$\{\{1+i, -2+2i\}, \{3-2i, 5i\}\}$

```
e2 = {{3 - I, 1 + I}, {1 - 2 I, -3}}
```

$\{\{3-i, 1+i\}, \{1-2i, -3\}\}$

A matriz de mudança de e1 para e2 é:

```
(t12 = Passagem[e1, e2]) // MatrixForm
```

$$\begin{pmatrix} -\frac{35}{17}-\frac{21 i}{17} & -\frac{26}{17}+\frac{49 i}{17} \\ \frac{9}{17}+\frac{19 i}{17} & \frac{30}{17}+\frac{i}{17} \end{pmatrix}$$

A matriz de mudança de e2 para e1 é:

```
(t21 = Passagem[e2, e1]) // MatrixForm
```

$$\begin{pmatrix} \frac{3}{20}+\frac{4 i}{5} & \frac{29}{20}+\frac{2 i}{5} \\ \frac{9}{20}-\frac{7 i}{20} & \frac{7}{20}-\frac{21 i}{20} \end{pmatrix}$$

Verificação dos resultados:

```
t12.t21 == IdentityMatrix[Length[e1]]
```

```
True
```

Defina-se um vector através das suas coordenadas na base `e2`:

```
xe2 = {4 - 2 I, -2 + 5 I}
```

$\{4-2\,i,\ -2+5\,i\}$

As suas coordenadas na base `e1` são:

```
Coordenadas[e2, e1, xe2]
```

$$\left\{-\frac{375}{17}-\frac{242\,i}{17},\ \frac{9}{17}+\frac{206\,i}{17}\right\}$$

3

Funções lineares

3.1 Introdução

Até ao presente capítulo, temos estudado os espaços vectoriais e as suas principais propriedades algébricas. Vamos agora estudar as funções que respeitam a estrutura algébrica daqueles espaços, ou seja, os *homomorfismos* dos espaços vectoriais. Veremos que estas funções constituem, por sua vez, um espaço vectorial e que este estudo permitirá lançar nova luz sobre a álgebra das matrizes, a qual veremos ser isomorfa da álgebra das funções lineares entre espaços com dimensão finita. As funções lineares, para além do seu interesse intrínseco, desempenham também um importante papel no cálculo diferencial das funções de mais de uma variável e que permitirá aproximar localmente as funções não lineares diferenciáveis por funções lineares, como o leitor poderá constatar quando estudar aquele tema nas cadeiras de Análise Matemática.

3.2 Funções lineares. Núcleo e imagem

Comecemos por definir a noção de função linear:

Definição 3.1 – Função linear – *Sejam E e F espaços vectoriais sobre um mesmo corpo $\mathbb{K}$ e $h\colon E \to F$ uma função. Diz-se que h é uma* ***função linear*** *ou um* ***homomorfismo*** *de E em F sse forem verificadas as condições seguintes, para quaisquer vectores $\vec{x}, \vec{x}\,' \in E$ e qualquer escalar $\alpha \in \mathbb{K}$:*

[L1] $h\big(\vec{x} + \vec{x}\,'\big) = h(\vec{x}) + h\big(\vec{x}\,'\big)$

[L2] $h(\alpha\vec{x}) = \alpha h(\vec{x})$

Portanto, a função h é linear sse transforma as operações características do espaço vectorial E nas respectivas operações em F. Convém observar que E e F devem ser espaços sobre o mesmo corpo $\mathbb{K}$, sem o que a condição L2 não faria sentido. A conjunção das duas condições anteriores é equivalente à condição única seguinte, onde α e β são escalares e $\vec{x}, \vec{x}\,' \in E$:

[L3] $h\big(\alpha\vec{x} + \beta\vec{x}\,'\big) = \alpha h(\vec{x}) + \beta h\big(\vec{x}\,'\big)$

Fazendo, em particular, $\alpha = 0$ e $\alpha = -1$ em L2, obtém-se

$$\begin{aligned} h(\vec{o}_E) &= \vec{o}_F \\ h(-\vec{x}) &= -h(\vec{x}) \end{aligned} \tag{3.1}$$

Da igualdade (1.1) que define a subtracção vectorial e da segunda das igualdades anteriores, obtém-se, para quaisquer vectores $\vec{x}, \vec{x}\,' \in E$,

$$h\big(\vec{x} - \vec{x}\,'\big) = h(\vec{x}) - h\big(\vec{x}\,'\big) \tag{3.2}$$

Portanto, as funções lineares respeitam a subtracção. Por outro lado, L1 e L2 implicam que, para qualquer combinação linear $\sum\limits_{j=1}^{m} \alpha_j\,\vec{e}_j$ de vectores de E, se tem:

$$h\left(\sum_{j=1}^{m} \alpha_j\,\vec{e}_j\right) = \sum_{j=1}^{m} \alpha_j h(\vec{e}_j) \tag{3.3}$$

Observe, ainda, que a primeira das igualdades (3.1) mostra que a igualdade (3.3) é válida, mesmo quando $m = 0$.

A proposição seguinte descreve o comportamento das funções lineares no que respeita aos subespaços de E e de F:

Proposição 3.1 *Sejam E e F espaços vectoriais sobre um mesmo corpo $\mathbb{K}$ e $h\colon E \to F$ uma função linear:*

i) *Se $A \subset E$ é um subespaço de E, então $h(A)$ é um subespaço de F. Em particular, o contradomínio $h(E)$ de h é um subespaço de F.*

ii) *Se $A \subset E$ é um subespaço de E e se $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m)$ for um sistema de vectores geradores de A, a sequência $h(e) = (h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_m))$ será um sistema de geradores de $h(A)$, isto é,*

$$h(A) = L(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_m))$$

Em particular, se $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ for uma base de E, então $(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_n))$ será um sistema de geradores do contradomínio $h(E)$.

iii) *Se $A \subset F$ é um subespaço de F, então $h^{-1}(A)$ é um subespaço de E. Em particular, o traço de h segundo $\vec{o}_F$ é um subespaço de E.*

Demonstração:

i) Mostremos que $h(A)$ satisfaz as condições S1, S2 e S3 do capítulo 1 (definição de subespaço):

Quanto a S1, a primeira das igualdades (3.1) mostra que

$$A \text{ é subespaço de } E \Rightarrow \vec{o}_E \in A \Rightarrow h(\vec{o}_E) \in h(A) \Rightarrow \vec{o}_F \in h(A) \Rightarrow h(A) \neq \emptyset$$

Para o fecho em relação à adição S2, tem-se, sucessivamente,

$$\begin{aligned} \vec{y}, \vec{y}\,' \in h(A) \Rightarrow \underset{\vec{x}, \vec{x}\,' \in A}{\exists} \vec{y} = h(\vec{x}) \wedge \vec{y}\,' = h\big(\vec{x}\,'\big) \Rightarrow \\ \Rightarrow \vec{y} + \vec{y}\,' = h(\vec{x}) + h\big(\vec{x}\,'\big) = h\big(\vec{x} + \vec{x}\,'\big) \Rightarrow \vec{y} + \vec{y}\,' \in h(A) \end{aligned}$$

A última implicação deve-se ao facto de se ter $\vec{x} + \vec{x}\,' \in A$, por ser A um subespaço de E e, consequentemente, $\vec{y} + \vec{y}\,' \in h(A)$.

Por último e para S3 (fecho em relação ao produto por escalar), seja $\vec{y} \in h(A)$ e $\alpha \in \mathbb{K}$: existe $\vec{x} \in A$ tal que $\vec{y} = h(\vec{x})$ e então

$$\alpha\vec{y} = \alpha h(\vec{x}) = h(\alpha\vec{x})$$

Porém, como A é subespaço de E, será $\alpha\vec{x} \in A$ e, por consequência, $\alpha\vec{y} \in h(A)$.

Quanto à segunda parte, sendo E um subespaço de E, a sua imagem $h(E)$ – que é o contradomínio de h – será um subespaço de F.

ii) Como $\vec{e}_j \in A$, tem-se (para $j = 1, \ldots, m$) $h(\vec{e}_j) \in h(A)$ e, por ser $h(A)$ um subespaço vectorial de F:

$$L(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_m)) \subset h(A)$$

Reciprocamente, ter-se-á:

$$\begin{aligned}\vec{y} \in h(A) \Rightarrow \underset{\vec{x}\in A}{\exists}\, \vec{y} = h(\vec{x}) \Rightarrow \vec{y} = h\left(\sum_{j=1}^{m} x_j\, \vec{e}_j\right) = \sum_{j=1}^{m} x_j h(\vec{e}_j) \Rightarrow \\ \Rightarrow \vec{y} \in L(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_m)) \Rightarrow \\ \Rightarrow h(A) \subset L(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_m))\end{aligned}$$

Por fim, as duas inclusões anteriores equivalem a

$$h(A) = L(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_m))$$

Quanto à segunda afirmação, basta atender a que uma base $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de E gera E e, portanto, $(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_n))$ gera $h(E)$, que é o contradomínio de h.

iii) Basta mostrar, de novo, que $h^{-1}(A)$ satisfaz as condições S1, S2 e S3:

Como $\vec{o}_F \in A$, a primeira igualdade (3.1) mostra que o vector $\vec{o}_E$ (pelo menos) pertence a $h^{-1}(A)$ que, portanto, não será vazio (condição S1).

Quanto a S2, tem-se:

$$\begin{aligned}\vec{x}, \vec{x}\,' \in h^{-1}(A) \Rightarrow h(\vec{x}), h\big(\vec{x}\,'\big) \in A \Rightarrow h(\vec{x}) + h\big(\vec{x}\,'\big) \in A \Rightarrow \\ \Rightarrow h\big(\vec{x} + \vec{x}\,'\big) \in A \Rightarrow \vec{x} + \vec{x}\,' \in h^{-1}(A)\end{aligned}$$

Por fim, para S3, tem-se:

$$\alpha \in \mathbb{K} \wedge \vec{x} \in h^{-1}(A) \Rightarrow \alpha \in \mathbb{K} \wedge h(\vec{x}) \in A \Rightarrow \alpha h(\vec{x}) = h(\alpha\vec{x}) \in A \Rightarrow \alpha\vec{x} \in h^{-1}(A)$$

Quanto à segunda parte, basta considerar $A = \{\vec{o}_F\}$. □

Da proposição anterior, resulta que o contradomínio $h(E)$ de uma função linear $h\colon E \to F$ é sempre um *subespaço* de F. De igual modo, a pré-imagem do subespaço trivial $\{\vec{o}_F\} \subset F$ por h é sempre um *subespaço* de E: estes subespaços têm, respectivamente, o nome de *Imagem* e *Núcleo* (ou *Kernel*) da aplicação linear h, o que se formaliza na seguinte

Definição 3.2 – Imagem, núcleo, característica e nulidade – *Sejam E e F espaços vectoriais sobre um mesmo corpo $\mathbb{K}$ e $h\colon E \to F$ uma função linear. O contradomínio de h é um subespaço de F chamado* ***Imagem*** *de h e a pré-imagem de $\{\vec{o}_F\} \subset F$ por h é um subespaço de E chamado o* ***Núcleo****,* ***Kernel*** *ou* ***Espaço nulo***[1] *de h e designados respectivamente por* $\mathrm{Img}(h)$ *e* $\mathrm{Nuc}(h)$ *ou* $\mathrm{Ker}(h)$

---

1 Os autores anglo-saxónicos, usam o termo *nullspace*.

$$\begin{aligned}\mathrm{Img}(h) = \mathrm{Im}(h) = h(E) = \left\{\vec{y} \in F\colon \underset{\vec{x}\in E}{\exists}\, \vec{y} = h(\vec{x})\right\} \subset F \\ \mathrm{Ker}(h) = \mathrm{Nuc}(h) = \left\{\vec{x} \in E\colon h(\vec{x}) = \vec{o}_F\right\} \subset E\end{aligned} \tag{3.4}$$

A dimensão do espaço $\mathrm{Img}(h)$ *é chamada a* ***característica*** *da função* h *– designada por* $\mathrm{c}(h)$ *ou* c_h *– e a dimensão do núcleo é chamada a* ***nulidade*** *de* h*, que designaremos por* $\mathrm{n}(h)$ *ou* n_h*:*

$$\begin{aligned}\mathrm{c}(h) = c_h = \dim(\mathrm{Img}(h)) \\ \mathrm{n}(h) = n_h = \dim(\mathrm{Ker}(h))\end{aligned} \tag{3.5}$$

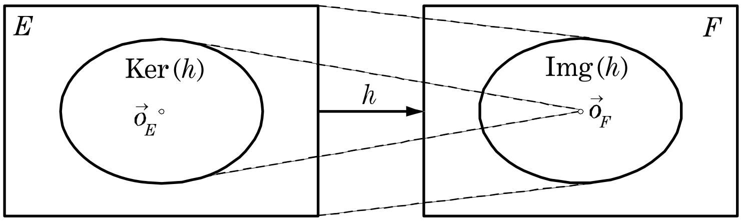


Fig. 3.1 – Esquema representativo do Núcleo e

Imagem de uma função linear $h\colon E \to F$.

O conjunto de todas as funções lineares de E em F designa-se por $\mathscr{L}(E,F)$ ou $\mathrm{Hom}(E,F)$.

Se o espaço F de chegada for o espaço $\mathbb{K}$, a função linear (de E em $\mathbb{K}$ e, portanto, assumindo valores escalares) toma o nome de *forma* linear.

Se $h\colon E \to F$ é injectiva, diz-se que h é um *monomorfismo* de E em F e se h for sobrejectiva, diz-se que h é um *epimorfismo* de E sobre F.

Se $h\colon E \to F$ é bijectiva, diz-se que h é um *isomorfismo* de E sobre F. Neste caso, os espaços vectoriais E e F são algebricamente idênticos, dizendo-se *isomorfos* (ou *iguais*, a menos de um isomorfismo) o que se descreve através da notação $E \cong F$ (leia-se "*E é isomorfo de F*").

Se $F = E$, a função linear toma o nome de *endomorfismo* de E e designaremos por $\mathrm{End}(E)$ o conjunto dos endomorfismos de E (ou seja, o conjunto $\mathscr{L}(E,E)$).

Se h é um endomorfismo bijectivo de E denominamo-lo por *automorfismo* de E, dizendo-se ainda que h é *invertível* ou *regular*.

Uma das propriedades interessantes do núcleo de um operador linear é a que o relaciona com a *injectividade* deste:

Proposição 3.2 – Caracterização das funções lineares injectivas – *Sejam E e F espaços vectoriais sobre um mesmo corpo $\mathbb{K}$ e $h\colon E \to F$ uma função linear. A função h é* ***injectiva*** *sse* $\mathrm{Ker}(h) = \{\vec{o}_E\}$.

Além disso, se $\vec{y}_0 \in F$, tem-se:

$$h^{-1}(\vec{y}_0) = \begin{cases} \emptyset & \textit{se } \vec{y}_0 \in F \setminus \mathrm{Img}(h) \\ \vec{x}_0 + \mathrm{Ker}(h) & \textit{se } \vec{y}_0 \in \mathrm{Img}(h) \textit{ e em que } h(\vec{x}_0) = \vec{y}_0 \end{cases} \tag{3.6.1}$$

Em particular, se h for injectiva e $\vec{y}_0 \in \mathrm{Img}(h)$, então:

$$h^{-1}(\vec{y}_0) = \{\vec{x}_0\}, \text{ onde } h(\vec{x}_0) = \vec{y}_0 \tag{3.6.2}$$

Demonstração:

A condição é necessária: como $\mathrm{Ker}(h)$ é subespaço de E, tem-se

$$\{\vec{o}_E\} \subset \mathrm{Ker}(h)$$

Por outro lado, se h é injectiva, então:

$$\vec{x} \in \mathrm{Ker}(h) \Rightarrow h(\vec{x}) = \vec{o}_F = h(\vec{o}_E) \Rightarrow \vec{x} = \vec{o}_E \Rightarrow \mathrm{Ker}(h) \subset \{\vec{o}_E\}$$

As inclusões anteriores equivalem à igualdade

$$\mathrm{Ker}(h) = \{\vec{o}_E\}$$

A condição é suficiente: de facto, se $\mathrm{Ker}(h) = \{\vec{o}_E\}$, então:

$$\begin{aligned} h(\vec{x}) = h\big(\vec{x}\,'\big) \Rightarrow h(\vec{x}) - h\big(\vec{x}\,'\big) = \vec{o}_F \Rightarrow h\big(\vec{x} - \vec{x}\,'\big) = \vec{o}_F \Rightarrow \\ \Rightarrow \vec{x} - \vec{x}\,' \in \mathrm{Ker}(h) = \{\vec{o}_E\} \Rightarrow \vec{x} - \vec{x}\,' = \vec{o}_E \Rightarrow \vec{x} = \vec{x}\,' \end{aligned}$$

o que mostra ser h injectiva.

Quanto à segunda parte da proposição, se $\vec{y}_0 \in F \setminus \mathrm{Img}(h)$, não existe qualquer vector $x \in \vec{E}$ tal que $h(\vec{x}) = \vec{y}_0$, o que significa que $h^{-1}(\vec{y}_0) = \emptyset$.

Se $\vec{y}_0 \in \mathrm{Img}(h)$, então será $h^{-1}(\vec{y}_0) \neq \emptyset$. Escolha-se em $h^{-1}(\vec{y}_0)$ um certo vector $\vec{x}_0$; por definição, tem-se $h(\vec{x}_0) = \vec{y}_0$ e mostremos que

$$h^{-1}(\vec{y}_0) = \vec{x}_0 + \mathrm{Ker}(h)$$

De facto,

$$\begin{aligned} \vec{x} \in \vec{x}_0 + \mathrm{Ker}(h) \Rightarrow \vec{x} = \vec{x}_0 + \vec{x}_1 \wedge \vec{x}_1 \in \mathrm{Ker}(h) \Rightarrow \\ \Rightarrow h(\vec{x}) = h(\vec{x}_0) + h(\vec{x}_1) = \vec{y}_0 + \vec{o}_E = \vec{y}_0 \Rightarrow \vec{x} \in h^{-1}(\vec{y}_0) \end{aligned}$$

o que mostra que $\vec{x}_0 + \mathrm{Ker}(h) \subset h^{-1}(\vec{y}_0)$.

Por outro lado, tem-se

$$\begin{aligned} \vec{x} \in h^{-1}(\vec{y}_0) \Rightarrow h(\vec{x}) = \vec{y}_0 \Rightarrow h(\vec{x}) - h(\vec{x}_0) = \vec{y}_0 - \vec{y}_0 = \vec{o}_E \Rightarrow \\ \Rightarrow h(\vec{x} - \vec{x}_0) = \vec{o}_E \Rightarrow \vec{x} - \vec{x}_0 \in \mathrm{Ker}(h) \Rightarrow \vec{x} = \vec{x}_0 + (\vec{x} - \vec{x}_0) \Rightarrow \vec{x} \in \vec{x}_0 + \mathrm{Ker}(h) \end{aligned}$$

as implicações anteriores mostram que também $h^{-1}(\vec{y}_0) \subset \vec{x}_0 + \mathrm{Ker}(h)$ e, portanto,

$$h^{-1}(\vec{y}_0) = \vec{x}_0 + \mathrm{Ker}(h)$$

□

A proposição seguinte reflecte o comportamento das aplicações lineares no que respeita à independência linear de vectores dos espaços envolvidos:

Proposição 3.3 *Sejam E e F espaços vectoriais sobre um mesmo corpo $\mathbb{K}$, $h\colon E \to F$ uma função linear e $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m)$ uma lista de vectores de E.*

i) *Se a lista $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m)$ for linearmente dependente, o mesmo sucede com a lista $h(e) = (h(\vec{e}_1), h(\vec{e}_2), \dots, h(\vec{e}_m))$.*

ii) *Se a lista (de vectores de* $\operatorname{Img}(h)$*)*

$$h(e) = (h(\vec{e}_1), h(\vec{e}_2), \dots, h(\vec{e}_m))$$

for linearmente independente, o mesmo sucede com a lista $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m)$.

iii) *Se h for injectiva e a lista $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m)$ for linearmente independente, o mesmo sucede com a lista $h(e) = (h(\vec{e}_1), h(\vec{e}_2), \dots, h(\vec{e}_m))$ de vectores de* $\operatorname{Img}(h)$.

iv) *Se h não é injectiva, existe uma sequência $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m)$ linearmente independente de vectores de E, tal que a sequência $h(e) = (h(\vec{e}_1), h(\vec{e}_2), \dots, h(\vec{e}_m))$ das imagens é linearmente dependente.*

Demonstração:

i) Suponhamos que e é linearmente dependente. Então, existe $\alpha_k \neq 0$ tal que $\sum_{j=1}^{m} \alpha_j \vec{e}_j = \vec{o}_E$ e, aplicando h a ambos os membros, obtém-se:

$$h\left(\sum_{j=1}^{m} \alpha_j \vec{e}_j\right) = h(\vec{o}_E) \Rightarrow \sum_{j=1}^{m} \alpha_j h(\vec{e}_j) = \vec{o}_F$$

Ora, a expressão obtida é uma combinação linear dos vectores $h(\vec{e}_j)$ na qual existe $1 \leq k \leq m$ tal que $\alpha_k \neq 0$, o que mostra que $h(e)$ é linearmente dependente.

ii) É a contra-recíproca da implicação anterior e, portanto, equivalente a esta.

iii) Mostremos que $h(e)$ é linearmente independente, se h for injectiva (o que, pela proposição 3.2, equivale a $\operatorname{Ker}(h) = \{\vec{o}_E\}$) e e for também linearmente independente

$$\sum_{j=1}^{m} \alpha_j h(\vec{e}_j) = \vec{o}_F \Rightarrow h\left(\sum_{j=1}^{m} \alpha_j \vec{e}_j\right) = \vec{o}_F \Rightarrow \sum_{j=1}^{m} \alpha_j \vec{e}_j \in \operatorname{Ker}(h) = \{\vec{o}_E\} \Rightarrow$$
$$\Rightarrow \sum_{j=1}^{m} \alpha_j \vec{e}_j = \vec{o}_E \Rightarrow \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$$

iv) Se h não é injectiva a proposição 3.2 garante que $\operatorname{Ker}(h) \neq \{\vec{o}_E\}$. Basta agora considerar uma base $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m)$ do núcleo, que será linearmente independente, e para a qual a sequência das imagens $h(e) = (h(\vec{e}_1), h(\vec{e}_2), \dots, h(\vec{e}_m)) = (\vec{o}_F, \vec{o}_F, \dots, \vec{o}_F)$ será, claro está, linearmente dependente. □

Vamos, a seguir, apresentar alguns exemplos de funções lineares, definidas em vários espaços vectoriais:

Exemplo 3.1 Sendo E e F espaços vectoriais sobre um corpo $\mathbb{K}$, a função *nula* $\mathbb{O}_{FE}\colon E \to F;\ \vec{x} \mapsto \vec{o}_F$ é, claramente, linear. Neste caso, tem-se

$$\begin{cases} \text{Ker}(\mathbb{O}_{FE}) = E \\ \text{Img}(\mathbb{O}_{FE}) = \{\vec{o}_F\} \end{cases}$$

Se $F = E$, usaremos a notação $\mathbb{O}_E$ para a função nula $\mathbb{O}_E\colon E \to E$.

Exemplo 3.2 Sendo $E \subset F$ um subespaço de um espaço vectorial F sobre um corpo $\mathbb{K}$, a *injecção canónica* $j\colon E \to F;\ \vec{x} \mapsto \vec{x}$ é linear, visto que, para quaisquer $\vec{x}, \vec{x}\,' \in E$ e $\alpha, \beta \in \mathbb{K}$,

$$j\big(\alpha\vec{x} + \beta\vec{x}\,'\big) = \alpha\vec{x} + \beta\vec{x}\,' = \alpha j(\vec{x}) + \beta j\big(\vec{x}\,'\big)$$

Para esta função, é

$$\begin{cases} \text{Ker}(j) = \{\vec{o}\} \\ \text{Img}(j) = E \end{cases}$$

Exemplo 3.3 Sendo E um espaço vectorial, a função *identidade em* E, $\mathscr{I}_E\colon E \to E;\ \vec{x} \mapsto \vec{x}$ é um automorfismo de E. Com efeito tem-se, como em cima,

$$\mathscr{I}_E\big(\alpha\vec{x} + \beta\vec{x}\,'\big) = \alpha\vec{x} + \beta\vec{x}\,' = \alpha\mathscr{I}_E(\vec{x}) + \beta\mathscr{I}_E\big(\vec{x}\,'\big)$$

Tal como no exemplo anterior, tem-se

$$\begin{cases} \text{Ker}(\mathscr{I}_E) = \{\vec{o}\} \\ \text{Img}(\mathscr{I}_E) = E \end{cases}$$

Exemplo 3.4 Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$, $h\colon E \to F$ uma função linear e $E' \subset E$ um *subespaço* de E. A *restrição* $h|_{E'}\colon E' \to F$ de h a E' é também uma função linear (demonstre!), que fica definida por

$$h|_{E'}(\vec{x}) = h(\vec{x}), \text{ em que } \vec{x} \in E'$$

Tem-se, neste caso,

$$\begin{cases} \text{Ker}(h|_{E'}) = \text{Ker}(h) \cap E' \\ \text{Img}(h|_{E'}) = h(E') \end{cases}$$

Exemplo 3.5 Considere-se um espaço vectorial E sobre um corpo $\mathbb{K}$ e o endomorfismo de E, $h_k\colon E \to E$ definido por:

$$h_k(\vec{x}) = k\vec{x}$$

onde $k \in \mathbb{K}$ é um escalar arbitrário. Como

$$h_k\big(\alpha\vec{x} + \beta\vec{x}\,'\big) = k\big(\alpha\vec{x} + \beta\vec{x}\,'\big) = k(\alpha\vec{x}) + k\big(\beta\vec{x}\,'\big) = \alpha(k\vec{x}) + \beta\big(k\vec{x}\,'\big) = \alpha h_k(\vec{x}) + \beta h_k\big(\vec{x}\,'\big)$$

segue-se que h_k é linear. O endomorfismo h_k é chamado uma *homotetia* de razão k. Se $k \neq 0$, tem-se:

$$\begin{cases} \text{Ker}(h_k) = \{\vec{o}\} \\ \text{Img}(h_k) = E \end{cases}$$

Se $k = 0$, então,

$$\begin{cases} \text{Ker}(h_k) = E \\ \text{Img}(h_k) = \{\vec{o}\} \end{cases}$$

Note que, se $k = 1$, fica $h_k = \mathscr{I}_E$ e, se $k = 0$, é $h_k = \mathscr{O}_E$ (função nula em E); por outro lado, se $k \neq 0$, h_k será um automorfismo de E, cujo inverso é $\vec{x} \mapsto \frac{1}{k}\vec{x}$.

Exemplo 3.6 Considerem-se os espaços $\mathbb{K}^{m,n}$ e $\mathbb{K}^{n,m}$ das matrizes de tipos $m \times n$ e $n \times m$ sobre o corpo $\mathbb{K}$. O operador *transposição* $^{\text{T}}\colon \mathbb{K}^{m,n} \to \mathbb{K}^{n,m}$ é linear. O núcleo é formado apenas pela matriz nula $O_{m,n}$ e a imagem é $\mathbb{K}^{n,m}$. Trata-se, portanto, de um isomorfismo de $\mathbb{K}^{m,n}$ sobre $\mathbb{K}^{n,m}$.

Exemplo 3.7 Considerem-se os espaços $\mathbb{C}^{m,n}$ e $\mathbb{C}^{n,m}$ das matrizes complexas de tipos $m \times n$ e $n \times m$ e o operador *transconjugação* $*\colon \mathbb{C}^{m,n} \to \mathbb{C}^{n,m}$ que transforma cada matriz complexa na sua transconjugada. Se os espaços forem considerados como espaços vectoriais *reais*, o operador é linear, já que, se $\alpha, \beta \in \mathbb{R}$,

$$(\alpha A + \beta B)^* = \alpha A^* + \beta B^*$$

Porém, se os espaços se considerarem como *complexos*, tem-se

$$(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$$

e o operador transconjugação *não será linear*.

Exemplo 3.8 Considere o espaço $\mathfrak{C}(\mathbb{N}, \mathbb{R})$ das sucessões reais convergentes e a função

$$\lim\colon \mathfrak{C}(\mathbb{N}, \mathbb{R}) \to \mathbb{R}; u_n \mapsto \lim u_n$$

que, a cada sucessão convergente, associa o seu limite. Esta função é linear, pois, como se sabe da análise, para quaisquer sucessões reais convergentes $(u_n), (v_n)$ e quaisquer escalares $\alpha, \beta \in \mathbb{R}$,

$$\lim(\alpha u_n + \beta v_n) = \alpha \lim u_n + \beta \lim v_n$$

O núcleo da função lim é o conjunto $\mathfrak{C}_0(\mathbb{N}, \mathbb{R})$ dos infinitésimos e a imagem é $\mathbb{R}$ (prove!):

$$\begin{cases} \text{Ker}(\lim) = \mathfrak{C}_0(\mathbb{N}, \mathbb{R}) \\ \text{Img}(\lim) = \mathbb{R} \end{cases}$$

Exemplo 3.9 Considere o espaço $\mathfrak{C}(\mathbb{N}, \mathbb{C})$ das sucessões complexas convergentes e a função

$$\lim\colon \mathfrak{C}(\mathbb{N}, \mathbb{C}) \to \mathbb{C}; u_n \mapsto \lim u_n$$

que, a cada sucessão convergente, associa o seu limite. Esta função é linear, pois, como se sabe da análise, para quaisquer sucessões complexas convergentes $(u_n), (v_n)$ e quaisquer escalares

$\alpha, \beta \in \mathbb{C}$

$$\lim(\alpha u_n + \beta v_n) = \alpha \lim u_n + \beta \lim v_n$$

O núcleo da função lim é o conjunto $\mathfrak{C}_0(\mathbb{N}, \mathbb{C})$ dos infinitésimos e a imagem é $\mathbb{C}$:

Exemplo 3.10 Sendo $\mathbb{K}$ um corpo e $\vec{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{K}^n$, as funções $h_{\vec{a}}\colon \mathbb{K}^n \to \mathbb{K}$ definidas, para todo o $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$, por

$$h_{\vec{a}}(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} a_j\, x_j \tag{3.7}$$

são lineares, pois, se $\vec{x} = (x_1, x_2, \ldots, x_n)$, $\vec{y} = (y_1, y_2, \ldots, y_n)$ e $\alpha,\beta \in \mathbb{K}$:

$$\begin{aligned} h_{\vec{a}}(\alpha\vec{x} + \beta\vec{y}) &= h_{\vec{a}}(\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, \ldots, \alpha x_n + \beta y_n) \\ &= \sum_{j=1}^{n} a_j(\alpha x_j + \beta y_j) = \sum_{j=1}^{n} a_j \alpha x_j + \sum_{j=1}^{n} a_j \beta y_j = \alpha \sum_{j=1}^{n} a_j x_j + \beta \sum_{j=1}^{n} a_j y_j \\ &= \alpha h_{\vec{a}}(x_1, x_2, \ldots, x_n) + \beta h_{\vec{a}}(y_1, y_2, \ldots, y_n) \end{aligned}$$

Observe-se que $h_{\vec{o}}$ é a função nula de $\mathbb{K}^n$ em $\mathbb{K}$ e que $h_{-\vec{a}} = -h_{\vec{a}}$.

Neste caso, tem-se

$$\begin{cases} \vec{a} = \vec{o} \Rightarrow \begin{cases} \operatorname{Ker}(h_{\vec{a}}) = \mathbb{K}^n \\ \operatorname{Img}(h_{\vec{a}}) = \{0\} \end{cases} \\ \vec{a} \neq \vec{o} \Rightarrow \begin{cases} \operatorname{Ker}(h_{\vec{a}}) = \left\{(x_1, x_2, \ldots, x_n) \colon \sum_{j=1}^{n} a_j\, x_j = 0\right\} \text{ é um } \textit{hiperplano} \text{ de } \mathbb{K}^n \\ \operatorname{Img}(h_{\vec{a}}) = \mathbb{K} \end{cases} \end{cases}$$

Reciprocamente, se $h\colon \mathbb{K}^n \to \mathbb{K}$ é uma forma linear e $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ é a base canónica de $\mathbb{K}^n$, tem-se

$$h(\vec{x}) = h(x_1, x_2, \ldots, x_n) = h\left(\sum_{j=1}^{n} x_j\, \vec{e}_j\right) = \sum_{j=1}^{n} x_j h(\vec{e}_j) = \sum_{j=1}^{n} a_j\, x_j$$

onde $a_j = h(\vec{e}_j) \in \mathbb{K}$. Portanto, as únicas formas lineares de $\mathbb{K}^n$ em $\mathbb{K}$ são as que se exprimem por meio da igualdade (3.7).

Exemplo 3.11 Sejam E e $(F_i)_{1 \leq i \leq m}$ espaços vectoriais sobre um corpo $\mathbb{K}$ e $h_i\colon E \to F_i$ uma sequência de m funções lineares. A função $h\colon E \to \prod_{i=1}^{m} F_i$ definida por

$$h(\vec{x}) = (h_1(\vec{x}), h_2(\vec{x}), \ldots, h_m(\vec{x}))$$

é linear, como se mostra a seguir, onde $\alpha, \beta \in \mathbb{K}$ e $\vec{x}, \vec{y} \in E$:

$$\begin{aligned} h(\alpha\vec{x} + \beta\vec{y}) &= (h_1(\alpha\vec{x} + \beta\vec{y}), h_2(\alpha\vec{x} + \beta\vec{y}), \ldots, h_m(\alpha\vec{x} + \beta\vec{y})) \\ &= (\alpha h_1(\vec{x}) + \beta h_1(\vec{y}), \alpha h_2(\vec{x}) + \beta h_2(\vec{y}), \ldots, \alpha h_m(\vec{x}) + \beta h_m(\vec{y})) \\ &= (\alpha h_1(\vec{x}), \alpha h_2(\vec{x}), \ldots, \alpha h_m(\vec{x})) + (\beta h_1(\vec{y}), \beta h_2(\vec{y}), \ldots, \beta h_m(\vec{y})) \\ &= \alpha(h_1(\vec{x}), h_2(\vec{x}), \ldots, h_m(\vec{x})) + \beta(h_1(\vec{y}), h_2(\vec{y}), \ldots, h_m(\vec{y})) \\ &= \alpha h(\vec{x}) + \beta h(\vec{y}) \end{aligned}$$

Quanto ao núcleo e imagem de h, tem-se:

$$\begin{cases} \mathrm{Ker}(h) = \bigcap_{i=1}^{m} \mathrm{Ker}(h_i) \\ \mathrm{Img}(h) \subset \prod_{i=1}^{m} \mathrm{Img}(h_i) \end{cases}$$

Como exercício, justifique o facto de a última relação não ser uma igualdade.

Exemplo 3.12 Dos dois exemplos anteriores, resulta que as funções lineares $h\colon \mathbb{K}^n \to \mathbb{K}^m$ são da forma

$$\vec{y} = h(\vec{x}) = h(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_m)$$

em que

$$y_i = \sum_{j=1}^{n} a_{ij}\, x_j;\ i = 1, 2, \ldots, m$$

onde os a_{ij} são escalares. As m igualdades escalares anteriores equivalem a uma igualdade matricial única:

$$Y = AX$$

onde $X = [\, x_1 \ \ x_2 \ \ \cdots \ \ x_n \,]^{\mathrm{T}}$, $Y = [\, y_1 \ \ y_2 \ \ \cdots \ \ y_m \,]^{\mathrm{T}}$ e $A = [a_{ij}] \in \mathbb{K}^{m,n}$.

Exemplo 3.13 Seja $(F_i)_{1 \le i \le m}$ uma lista de m espaços vectoriais sobre um corpo $\mathbb{K}$ e $F = \prod_{i=1}^{m} F_i$ o espaço vectorial produto. Considere as *projecções* $\mathrm{pr}_i\colon F \to F_i$ definidas por:

$$\mathrm{pr}_i(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) = \vec{x}_i;\ \ i = 1, \ldots, m$$

As m funções pr_i acima definidas são lineares, visto que, para quaisquer $\alpha, \beta \in \mathbb{K}$ e $\vec{x} = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m), \vec{y} = (\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m) \in F$, se tem:

$$\begin{aligned} \mathrm{pr}_i(\alpha\vec{x} + \beta\vec{y}) &= \mathrm{pr}_i(\alpha(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) + \beta(\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m)) = \\ &= \mathrm{pr}_i(\alpha\vec{x}_1 + \beta\vec{y}_1, \alpha\vec{x}_2 + \beta\vec{y}_2, \ldots, \alpha\vec{x}_m + \beta\vec{y}_m) \\ &= \alpha\vec{x}_i + \beta\vec{y}_i \\ &= \alpha\mathrm{pr}_i(\vec{x}) + \beta\mathrm{pr}_i(\vec{y}) \end{aligned}$$

Pode ainda o leitor mostrar que

$$\begin{cases} \mathrm{Ker}(\mathrm{pr}_i) = \prod_{k<i} F_k \times \{\vec{o}_{F_i}\} \times \prod_{k>i} F_k \\ \mathrm{Img}(\mathrm{pr}_i) = F_i \end{cases}$$

Exemplo 3.14 Sejam $(E_i)_{1 \le i \le m}$ e F espaços vectoriais sobre um corpo $\mathbb{K}$ e $h_i\colon E_i \to F$ funções lineares. Sendo $E = \prod_{i=1}^{m} E_i$ o espaço produto, consideremos a função $h\colon E \to F$ definida, para todo o $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) \in E$, por:

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) = \sum_{i=1}^{m} h_i(\vec{x}_i)$$

A função h anteriormente definida é também linear, pois tem-se, para quaisquer $\alpha, \beta \in \mathbb{K}$ e $\vec{x} = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m), \vec{y} = (\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m) \in E$:

$$\begin{aligned}
h(\alpha\vec{x} + \beta\vec{y}) &= h(\alpha(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m) + \beta(\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m)) \\
&= h(\alpha\vec{x}_1 + \beta\vec{y}_1, \alpha\vec{x}_2 + \beta\vec{y}_2, \ldots, \alpha\vec{x}_m + \beta\vec{y}_m) \\
&= \sum_{i=1}^{m} h_i(\alpha\vec{x}_i + \beta\vec{y}_i) = \sum_{i=1}^{m}(\alpha h_i(\vec{x}_i) + \beta h_i(\vec{y}_i)) \\
&= \alpha \sum_{i=1}^{m} h_i(\vec{x}_i) + \beta \sum_{i=1}^{m} h_i(\vec{y}_i) \\
&= \alpha h(\vec{x}) + \beta h(\vec{y})
\end{aligned}$$

Quanto ao núcleo e à imagem de h, tem-se (demonstre!):

$$\begin{cases} \mathrm{Ker}(h) \supset \prod_{i=1}^{m} \mathrm{Ker}(h_i) \\ \mathrm{Img}(h) = \sum_{i=1}^{m} \mathrm{Img}(h_i) \end{cases}$$

Exemplo 3.15 Seja $n \geq 0$ um inteiro. A função $D\colon \mathscr{P}_n(\mathbb{R}) \to \mathscr{P}_n(\mathbb{R})$ que transforma cada polinómio de grau inferior ou igual a n na sua derivada é um endomorfismo de $\mathscr{P}_n(\mathbb{R})$, cujo núcleo é $\mathscr{P}_0(\mathbb{R})$ e cuja imagem é $\mathscr{P}_{n-1}(\mathbb{R})$:

$$D\left(\sum_{k=0}^{n} a_k x^k\right) = \sum_{k=0}^{n-1}(k+1)a_{k+1}x^k$$

$$\begin{cases} \mathrm{Ker}(D) = \mathscr{P}_0(\mathbb{R}) \\ \mathrm{Img}(D) = \mathscr{P}_{n-1}(\mathbb{R}) \end{cases}$$

Daqui se segue que D não é *injectivo* nem *sobrejectivo*.

Exemplo 3.16 Seja $I \subset \mathbb{R}$ um intervalo aberto. O operador derivação $D\colon \mathscr{D}(I, \mathbb{R}) \to \mathbb{R}^I$ que transforma cada função f diferenciável em I na sua função derivada Df é linear, atendendo aos teoremas sobre diferenciabilidade (regras de derivação)

$$D(\alpha f + \beta g) = \alpha Df + \beta Dg$$

O *núcleo* de D é formado pelas *funções constantes* em I e a *imagem* de D é o conjunto das *funções primitiváveis* em I. O operador derivação D aqui definido não é *injectivo* nem *sobrejectivo*.

Exemplo 3.17 Seja $I \subset \mathbb{R}$ um intervalo aberto e $a \in I$. A função

$$\mathrm{Int}_a\colon \mathscr{C}(I, \mathbb{R}) \to \mathscr{C}(I, \mathbb{R});\ f \mapsto F_a$$

chamada *integral indefinido com origem* a, que transforma cada função f contínua em I na função F_a (de classe $\mathscr{C}^1$ e, portanto, contínua em I e cuja derivada é f, devido ao teorema fundamental do cálculo integral) definida, para todo o $x \in I$ por

$$F_a(x) = \int_a^x f(t)\mathrm{d}t;\ x \in I$$

é linear, visto que, como se sabe do cálculo integral, para quaisquer funções f e g contínuas em I e quaisquer reais α e β, tem-se

$$\int_a^x \alpha f(t) + \beta g(t)\mathrm{d}t = \alpha \int_a^x f(t)\mathrm{d}t + \beta \int_a^x g(t)\mathrm{d}t$$

Observe-se que a função F_a se anula, necessariamente, em a. O *núcleo* de Int_a é formado apenas pela função nula e a *imagem* de Int_a é formada pelas funções de $\mathscr{C}^1(I, \mathbb{R})$ que se anulam em a:

$$\mathrm{Img}(\mathrm{Int}_a) = \mathscr{C}^1(I, \mathbb{R}) \cap \{F \colon F(a) = 0\}$$

O operador Int_a é, pois, *injectivo* mas não *sobrejectivo*.

Exemplo 3.18 Seja E um espaço vectorial sobre um corpo $\mathbb{K}$, $m > 0$ um inteiro e

$$x = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m)$$

uma lista de m vectores de E. Considere-se, ainda, a função $h_x \colon \mathbb{K}^m \to E$, definida, para todo o vector $\vec{k} = (k_1, k_2, \dots, k_m) \in \mathbb{K}^m$, por

$$h_x(k_1, k_2, \dots, k_m) = \sum_{i=1}^{m} k_i \vec{x}_i$$

A função h_x é linear, como se mostra a seguir, onde α, β são escalares e

$$\vec{u} = (u_1, u_2, \dots, u_m) \text{ e } \vec{v} = (v_1, v_2, \dots, v_m)$$

são vectores de $\mathbb{K}^m$:

$$\begin{aligned} h_x(\alpha\vec{u} + \beta\,\vec{v}) &= h_x(\alpha u_1 + \beta\, v_1, \alpha u_2 + \beta\, v_2, \dots, \alpha u_m + \beta\, v_m) = \sum_{i=1}^{m} (\alpha u_i + \beta\, v_i)\vec{x}_i \\ &= \sum_{i=1}^{m} \alpha u_i \vec{x}_i + \sum_{i=1}^{m} \beta\, v_i \vec{x}_i = \alpha \sum_{i=1}^{m} u_i \vec{x}_i + \beta \sum_{i=1}^{m} v_i \vec{x}_i \\ &= \alpha h_x(\vec{u}) + \beta h_x(\vec{v}) \end{aligned}$$

É fácil reconhecer que o *subespaço gerado por* x é a *imagem* de h_x, que x é *geradora* de E sse h_x for *sobrejectiva*, que x é *linearmente independente* sse h_x for *injectiva* e que x é uma *base* de E sse h_x for um *isomorfismo* de $\mathbb{K}^m$ sobre E:

$$\begin{cases} L(x) = \mathrm{Img}(h_x) \\ x \text{ é } \textit{geradora} \text{ de } E \Leftrightarrow h_x \text{ é } \textit{sobrejectiva} \\ x \text{ é } \textit{linearmente independente} \Leftrightarrow h_x \text{ é } \textit{injectiva} \\ x \text{ é uma } \textit{base} \text{ de } E \Leftrightarrow h_x \text{ é um } \textit{isomorfismo} \text{ de } \mathbb{K}^m \text{ sobre } E \end{cases}$$

Daqui se pode, de novo, obter o resultado (já referido no capítulo 1 – proposição 1.15) de que *todo o espaço vectorial* E *de dimensão* n *sobre um corpo* $\mathbb{K}$ *é isomorfo do espaço cartesiano* $\mathbb{K}^n$.

3.3 Álgebra das funções lineares

Nesta secção, vamos introduzir em $\mathscr{L}(E, F)$ uma estrutura de *espaço vectorial* sobre o corpo $\mathbb{K}$ em relação ao qual estão definidos os espaços vectoriais E e F. Para tal, considerem-se duas funções lineares $h, k: E \to F$ e defina-se $h + k: E \to F$ como a função dada por:

$$(h+k)(\vec{x}) = h(\vec{x}) + k(\vec{x}); \vec{x} \in E \tag{3.8}$$

O leitor pode verificar facilmente que $h + k$ é ainda uma função *linear* (verifique as condições L1 e L2 para $h + k$!) e que $(\mathscr{L}(E, F), +)$ é um *grupo comutativo aditivo* (axiomas A1-A4 dos espaços vectoriais – capítulo 1), cujo *zero* é a função $\mathscr{O}_{FE}: E \to F$ identicamente nula em E, tal que, para todo o vector $\vec{x} \in E$,

$$\mathscr{O}_{FE}(\vec{x}) = \vec{o}_F; \vec{x} \in E \tag{3.9}$$

Por outro lado, a função *simétrica* de uma função $h: E \to F$ é a função $-h$ tal que, para todo o vector $\vec{x} \in E$:

$$(-h)(\vec{x}) = -(h(\vec{x})); \vec{x} \in E \tag{3.10}$$

Como em qualquer grupo comutativo, põe-se

$$h - k = h + (-k)$$

donde, para qualquer $\vec{x} \in E$,

$$(h-k)(\vec{x}) = h(\vec{x}) - k(\vec{x}); \vec{x} \in E \tag{3.11}$$

Sendo $\alpha \in \mathbb{K}$ e $h: E \to F$ uma função linear, define-se o produto de α por h como a nova função $\alpha h: E \to F$ obtida pondo, para todo o vector $\vec{x} \in E$,

$$(\alpha h)(\vec{x}) = \alpha(h(\vec{x})); \vec{x} \in E \tag{3.12}$$

A função αh é *linear* (Mostre!) e a operação anterior satisfaz os axiomas P1-P4 dos espaços vectoriais (capítulo 1). Em resultado disto, $\mathscr{L}(E, F)$ constitui um *espaço vectorial* sobre o corpo $\mathbb{K}$.

Vamos, agora, definir uma terceira operação binária entre funções lineares: o *produto* ou *composição* de funções lineares.

Proposição 3.4 – Composição de funções lineares – *Sejam* G, E *e* F *espaços vectoriais sobre um mesmo corpo* $\mathbb{K}$ *e* $k: G \to E$ *e* $h: E \to F$ *funções lineares. A função (designada por* $h \circ k$ *ou, simplesmente,* hk*)* $h \circ k: G \to F$ *definida, para todo o* $\vec{x} \in G$*, por*

$$(h \circ k)(\vec{x}) = h(k(\vec{x})); \vec{x} \in G \tag{3.13}$$

*é **linear** e é designada por **composta** ou **produto** de h e k (ver figura 3.2).*

Demonstração:

Para provar que $h \circ k$ é linear, mostremos que L1 e L2 se verificam. Quanto a L1:

$$\begin{aligned}(h \circ k)\big(\vec{x} + \vec{x}'\big) &= h\big(k\big(\vec{x} + \vec{x}'\big)\big) = h\big(k(\vec{x}) + k\big(\vec{x}'\big)\big) \\ &= h(k(\vec{x})) + h\big(k\big(\vec{x}'\big)\big) = (h \circ k)(\vec{x}) + (h \circ k)\big(\vec{x}'\big)\end{aligned}$$

Em relação a L2, tem-se:

$$(h \circ k)(\alpha\vec{x}) = h(k(\alpha\vec{x})) = h(\alpha k(\vec{x})) = \alpha(h(k(\vec{x}))) = \alpha((h \circ k)(\vec{x})) \qquad \square$$

$$\begin{tikzcd} G \arrow[r, "k"] \arrow[dr, "hk = h \circ k"'] & E \arrow[d, "h"] \\ & F \end{tikzcd}$$

Fig. 3.2 – Composição de funções lineares.

Observe-se que, nas condições da proposição anterior, existe $h \circ k$ e, se $G \neq F$, $k \circ h$ não existe.

Se $G = F \neq E$, então $h \circ k$ é um *endomorfismo* de F, enquanto que $k \circ h$ é um *endomorfismo* de E, sendo, portanto, $h \circ k \neq k \circ h$.

Por fim, se for $G = E = F$, então $h \circ k$ e $k \circ h$ são ambos endomorfismos de E. Mas, mesmo assim tem-se, geralmente,

$$h \circ k \neq k \circ h$$

No entanto, existem como veremos endomorfismos h, k para os quais é $h \circ k = k \circ h$; destes endomorfismos diremos serem *permutáveis* ou *comutáveis*. De tudo o que acaba de ser dito pode, portanto, afirmar-se que a composição de funções lineares *não é comutativa*.

A operação de composição anteriormente definida goza de certas propriedades em relação às operações dos espaços vectoriais $\mathscr{L}(G, E)$ e $\mathscr{L}(E, F)$ e que se resumem na seguinte

Proposição 3.5 – Propriedades da composição – *Sejam E, F, G e H espaços vectoriais sobre um corpo $\mathbb{K}$. A composição de funções lineares satisfaz as seguintes propriedades algébricas, para quaisquer $\alpha \in \mathbb{K}$, $k, k' \in \mathscr{L}(G, E)$, $h, h' \in \mathscr{L}(E, F)$, $u \in \mathscr{L}(F, H)$:*

i) *Associatividade:*

$$(u \circ h) \circ k = u \circ (h \circ k)$$

ii) *Comportamento da composição em relação ao produto por escalar:*

$$(\alpha h) \circ k = \alpha(h \circ k)$$
$$h \circ (\alpha k) = \alpha(h \circ k)$$

iii) *Distributividades à direita e à esquerda em relação à adição:*

$$h \circ (k + k') = h \circ k + h \circ k'$$
$$(h + h') \circ k = h \circ k + h' \circ k$$

iv) *Existência de elementos neutros à esquerda e à direita:*

$$\mathscr{I}_F \circ h = h$$
$$h \circ \mathscr{I}_E = h$$

v) *Propriedade das funções nulas em relação à composição:*

$$\mathscr{O}_{FE} \circ k = \mathscr{O}_{FG}$$
$$h \circ \mathscr{O}_{EG} = \mathscr{O}_{FG}$$

Demonstração:

i) Atente-se no esquema de composição seguinte

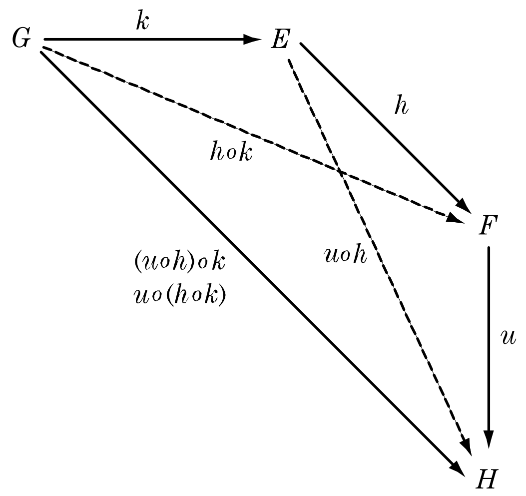


Fig. 3.3 – A associatividade da Composição de funções lineares.

De acordo com a figura, tem-se, sucessivamente, para qualquer $\vec{x} \in G$,

$$((u \circ h) \circ k)(\vec{x}) = (u \circ h)(k(\vec{x})) = u(h(k(\vec{x}))) = u((h \circ k)(\vec{x})) = (u \circ (h \circ k))(\vec{x})$$

ii) Observe-se o esquema de composição

$$\begin{tikzcd} G \arrow[r, "{k,\alpha k}"] \arrow[dr, swap, "{\substack{h\circ k \\ \alpha(h\circ k) \\ (\alpha h)\circ k \\ h\circ(\alpha k)}}"] & E \arrow[d, "{\substack{h \\ \alpha h}}"] \\ & F \end{tikzcd}$$

Fig. 3.4 – Composição de funções lineares e o produto de escalar por função linear.

Então, tem-se, para qualquer $\vec{x} \in G$,

$$((\alpha h) \circ k)(\vec{x}) = (\alpha h)(k(\vec{x})) = \alpha(h(k(\vec{x}))) = \alpha(h \circ k)(\vec{x}) = (\alpha(h \circ k))(\vec{x})$$

Identicamente, ter-se-á, atendendo à linearidade de h,

$$(h \circ (\alpha k))(\vec{x}) = h((\alpha k)(\vec{x})) = h(\alpha k(\vec{x})) = \alpha h(k(\vec{x})) = \alpha(h \circ k)(\vec{x}) = (\alpha(h \circ k))(\vec{x})$$

iii) De acordo com os esquemas de composição que se seguem, poderemos escrever, para qualquer $\vec{x} \in G$,

$$\begin{aligned} &((h + h') \circ k)(\vec{x}) = (h + h')(k(\vec{x})) = h(k(\vec{x})) + h'(k(\vec{x})) \\ &= (h \circ k)(\vec{x}) + (h' \circ k)(\vec{x}) = (h \circ k + h' \circ k)(\vec{x}) \end{aligned}$$

$$\begin{tikzcd} G \arrow[r, "{k,\, k',\, k+k'}"] \arrow[dr, swap, "{\substack{h\circ k \\ h\circ k' \\ h\circ(k+k')}}"] & E \arrow[d, "h"] \\ & F \end{tikzcd} \qquad \begin{tikzcd} G \arrow[r, "k"] \arrow[dr, swap, "{\substack{h\circ k \\ h'\circ k \\ (h+h')\circ k}}"] & E \arrow[d, "{\substack{h \\ h' \\ h+h'}}"] \\ & F \end{tikzcd}$$

Fig. 3.5 – Distributividades à esquerda e à direita em relação à adição.

Do mesmo modo, tem-se:

$$\begin{aligned} (h \circ (k + k'))(\vec{x}) &= h((k + k')(\vec{x})) = h(k(\vec{x}) + k'(\vec{x})) = h(k(\vec{x})) + h(k'(\vec{x})) \\ &= (h \circ k)(\vec{x}) + (h \circ k')(\vec{x}) = (h \circ k + h \circ k')(\vec{x}) \end{aligned}$$

iv) Atendendo às figuras seguintes,

$$\begin{tikzcd} E \arrow[r, "h"] \arrow[dr, swap, "\substack{h \\ \mathscr{J}_F \circ h}"] & F \arrow[d, "\mathscr{J}_F"] \\ & F \end{tikzcd} \qquad \begin{tikzcd} E \arrow[r, "\mathscr{J}_E"] \arrow[dr, swap, "\substack{h \\ h \circ \mathscr{J}_E}"] & E \arrow[d, "h"] \\ & F \end{tikzcd}$$

Fig. 3.6 – Funções identidade $\mathscr{J}_F$ e $\mathscr{J}_E$: elementos neutros à esquerda e à direita.

poderemos escrever, para qualquer $\vec{x} \in E$,

$$(\mathscr{J}_F \circ h)(\vec{x}) = \mathscr{J}_F(h(\vec{x})) = h(\vec{x})$$

De modo análogo e para qualquer $\vec{x} \in E$,

$$(h \circ \mathscr{J}_E)(\vec{x}) = h(\mathscr{J}_E(\vec{x})) = h(\vec{x})$$

v)

$$\begin{tikzcd} G \arrow[r, "k"] \arrow[dr, swap, "\substack{\mathscr{O}_{FG} \\ \mathscr{O}_{FE} \circ k}"] & E \arrow[d, "\mathscr{O}_{FE}"] \\ & F \end{tikzcd} \qquad \begin{tikzcd} G \arrow[r, "\mathscr{O}_{EG}"] \arrow[dr, swap, "\substack{\mathscr{O}_{FG} \\ h \circ \mathscr{O}_{EG}}"] & E \arrow[d, "h"] \\ & F \end{tikzcd}$$

Fig. 3.7 – As funções nulas $\mathscr{O}_{FE}$, $\mathscr{O}_{EG}$ e $\mathscr{O}_{FG}$ e a Composição de funções lineares.

Em relação à figura da esquerda, tem-se, para qualquer $\vec{x} \in G$,

$$(\mathscr{O}_{FE} \circ k)(\vec{x}) = \mathscr{O}_{FE}(k(\vec{x})) = \vec{o}_F = \mathscr{O}_{FG}(\vec{x})$$

Para a segunda igualdade, atente-se na figura anterior à direita, com $\vec{x} \in G$:

$$(h \circ \mathscr{O}_{EG})(\vec{x}) = h(\mathscr{O}_{EG}(\vec{x})) = h(\vec{o}_E) = \vec{o}_F = \mathscr{O}_{FG}(\vec{x}) \qquad \square$$

Observações:

- Observemos que a operação $\circ$ de composição é, em geral, uma lei de composição *externa*

$$\circ : \mathscr{L}(E, F) \times \mathscr{L}(G, E) \to \mathscr{L}(G, F);\ (h, k) \mapsto h \circ k$$

Porém, se $G = E = F$, h e k são endomorfismos de E e a composição é, então, uma lei de composição *interna* em $\mathrm{End}(E)$.

- As alíneas *i), iii)* e *iv)* da proposição anterior significam que $(\mathrm{End}(E), +, \circ)$ constitui um *anel não comutativo com unidade*, que é a função identidade em E, $\mathscr{I}_E$.

- Se atendermos também a *ii)*, conclui-se ser $\mathrm{End}(E)$ uma *álgebra linear* sobre o corpo $\mathbb{K}$ (segundo os axiomas M1-M4 do capítulo 1).

Vamos ver, de seguida, que é possível definir a potenciação (com expoente inteiro positivo) para qualquer endomorfismo de um espaço vectorial, pondo

Definição 3.3 – Potência de um endomorfismo com expoente inteiro – *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ e $h\colon E \to E$ um endomorfismo de E. Para todo o inteiro $m \geq 0$, poremos por recorrência,*

$$\begin{cases} h^0 = \mathscr{I}_E \\ h^m = h \circ h^{m-1}, \text{ se } m > 0 \end{cases}$$

Se h for um automorfismo de E, poderemos definir a potência de expoente inteiro negativo $-m \in \mathbb{Z}^-$, pondo:

$$h^{-m} = \left(h^{-1}\right)^m;\ m > 0$$

onde h^{-1} é o automorfismo inverso de h.

A definição anterior dá, portanto, para $m > 0$,

$$h^m = \underbrace{h \circ h \circ \cdots \circ h}_{m \text{ vezes}}$$

A associatividade implica que, para qualquer endomorfismo $h\colon E \to E$ e $m, n \in \mathbb{N}_0$, se tem

$$h^{m+n} = h^m \circ h^n$$

e a igualdade anterior vale para $m, n \in \mathbb{Z}$, se h for um automorfismo de E.

Sendo $m \geq 0$ e $(\alpha_k)_{0 \leq k \leq m}$ uma sequência de $m+1$ escalares de $\mathbb{K}$ pode, portanto, definir-se o *polinómio* no endomorfismo h, de grau inferior ou igual a m e coeficientes α_k, por

$$\sum_{k=0}^{m} \alpha_k h^k = \alpha_0 \mathscr{I}_E + \alpha_1 h + \alpha_2 h^2 + \cdots + \alpha_m h^m$$

3.4 Funções lineares em espaços de dimensão finita

Vamos, agora, ver algumas propriedades das funções lineares definidas num espaço com dimensão finita, entre as quais se inclui a que mostra que uma função linear definida num espaço E de dimensão finita fica completamente determinada pelos seus valores nos vectores de uma base de E:

Proposição 3.6 *Seja E um espaço vectorial sobre um corpo $\mathbb{K}$ com dimensão finita n e F um outro espaço vectorial sobre o mesmo corpo. Seja ainda $h: E \to F$ uma função linear. Então:*

i) *Se $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m)$ for uma sequência* ***geradora*** *de E, então*

$$h(e) = (h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_m))$$

será uma sequência ***geradora*** *de* $\text{Img}(h)$.

ii) *Se h é* ***injectiva*** *e $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ é uma* ***base*** *de E, então*

$$h(e) = (h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_n))$$

será uma ***base*** *de* $\text{Img}(h)$*, sendo então:* $\dim(\text{Img}(h)) = \dim E = n$.

iii) $\text{Ker}(h)$ *e* $\text{Img}(h)$ *têm dimensão* ***finita****, tendo-se*

$$\dim(\text{Ker}(h)) + \dim(\text{Img}(h)) = \dim E \tag{3.14.1}$$

ou seja,

$$\text{n}(h) + \text{c}(h) = \dim E \tag{3.14.2}$$

iv) *Se h é nula em cada um dos vectores de uma base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de E, então h é identicamente nula ($h = \mathbb{O}_{FE}$), isto é, para todo o $\vec{x} \in E$,*

$$h(\vec{x}) = \vec{o}_F;\ \vec{x} \in E$$

v) *Sendo $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ uma base de E e $(\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n)$ uma lista de n vectores de F, existe uma e uma só função linear $\varphi: E \to F$ tal que $\varphi(\vec{e}_j) = \vec{y}_j$, para $j = 1, 2, \ldots, n$.*

Demonstração:

i) Resulta imediatamente da proposição 3.1.*ii*, fazendo aí $A = E$ e notando que

$$h(E) = \text{Img}(h)$$

ii) Como e é uma base de E, e é *geradora* de E. Pela alínea anterior, a sequência $(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_n))$ gera $\text{Img}(h)$; atendendo, por outro lado, à proposição 3.3.*iii*, $h(e) = (h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_n))$ é *linearmente independente* (visto ser e linearmente independente e h injectiva); daqui se segue que $h(e)$ é uma *base* de $\text{Img}(h)$, donde resulta a igualdade das dimensões de E e $\text{Img}(h)$:

$$\dim(\text{Img}(h)) = n = \dim E$$

iii) A relação enunciada e que agora provaremos é, por vezes, denominada *teorema fundamental* das funções lineares. Como $\text{Ker}(h)$ é subespaço de E, tem-se, obviamente, $0 \leq \dim(\text{Ker}(h)) \leq n$. Seja então s a dimensão de $\text{Ker}(h)$ e

$$e_{\text{Ker}(h)} = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s)$$

uma *base* do núcleo de h. Observe-se que

$$\vec{e}_j \in \text{Ker}(h) \Rightarrow h(\vec{e}_j) = \vec{o}_F;\ j = 1, 2, \ldots, s$$

Como $e_{\text{Ker}(h)}$ é linearmente independente, podemos prolongar esta lista a uma *base* e_E de E (ver proposição 1.16.*v*), mediante a inserção de $n-s$ vectores $(\vec{e}_{s+1}, \vec{e}_{s+2}, \ldots, \vec{e}_n)$ de $E \setminus \text{Ker}(h)$:

$$e_E = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s, \vec{e}_{s+1}, \vec{e}_{s+2}, \ldots, \vec{e}_n)$$

A proposição fica demonstrada, se provarmos que a lista u de $n-s$ vectores de $\text{Img}(h)$ definida por

$$u = (h(\vec{e}_{s+1}), h(\vec{e}_{s+2}), \ldots, h(\vec{e}_n))$$

é uma *base* de $\text{Img}(h)$ e que, portanto, $\dim(\text{Img}(h)) = n - s$:

Como, para $s < j \leq n$, se tem $h(\vec{e}_j) \in \text{Img}(h)$, vem, por ser $\text{Img}(h)$ subespaço de F:

$$L(u) \subset \text{Img}(h) \tag{3.15}$$

Por outro lado, tem-se (justifique cada passagem!):

$$\begin{aligned}
&\vec{y} \in \text{Img}(h) \Rightarrow \underset{\vec{x}\in E}{\exists}\ \vec{y} = h(\vec{x}) \Rightarrow \\
&\qquad \Rightarrow \vec{y} = h\left(\sum_{j=1}^{n} x_j\, \vec{e}_j\right) = \sum_{j=1}^{n} x_j h(\vec{e}_j) = \underbrace{\sum_{j=1}^{s} x_j \overbrace{h(\vec{e}_j)}^{\vec{o}_F}}_{\vec{o}_F} + \sum_{j=s+1}^{n} x_j h(\vec{e}_j) \Rightarrow \\
&\qquad \Rightarrow \vec{y} = \sum_{j=s+1}^{n} x_j h(\vec{e}_j) \Rightarrow \vec{y} \in L(u) \Rightarrow \\
&\qquad \Rightarrow \text{Img}(h) \subset L(u)
\end{aligned} \tag{3.16}$$

As relações (3.15) e (3.16) equivalem a $L(u) = \text{Img}(h)$ e, portanto, u ***gera*** $\text{Img}(h)$.

Vejamos, por fim, que u é *linearmente independente*, o que resulta das implicações seguintes (justifique cada passagem!):

$$\begin{aligned}
&\sum_{j=s+1}^{n} x_j h(\vec{e}_j) = \vec{o}_F \Rightarrow h\left(\sum_{j=s+1}^{n} x_j\, \vec{e}_j\right) = \vec{o}_F \Rightarrow \sum_{j=s+1}^{n} x_j\, \vec{e}_j \in \text{Ker}(h) \Rightarrow \\
&\qquad \Rightarrow \sum_{j=s+1}^{n} x_j\, \vec{e}_j = \sum_{j=1}^{s} x_j\, \vec{e}_j \Rightarrow \sum_{j=1}^{s} -x_j\, \vec{e}_j + \sum_{j=s+1}^{n} x_j\, \vec{e}_j = \vec{o}_E \Rightarrow \\
&\qquad \Rightarrow x_1 = x_2 = \cdots = x_n = 0 \Rightarrow x_{s+1} = x_{s+2} = \cdots = x_n = 0
\end{aligned}$$

o que prova a independência linear de u.

iv) De facto, para qualquer vector $\vec{x} \in E$ e sendo x_j as coordenadas de $\vec{x}$ em relação à base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$, tem-se:

$$h(\vec{x}) = h\left(\sum_{j=1}^{n} x_j\, \vec{e}_j\right) = \sum_{j=1}^{n} x_j h(\vec{e}_j) = \sum_{j=1}^{n} x_j\, \vec{o}_F = \vec{o}_F$$

v) Considere-se a função $\varphi: E \to F$ definida por

$$\varphi(\vec{x}) = \sum_{j=1}^{n} x_j \, \vec{y}_j$$

onde os escalares x_j são as coordenadas (únicas) de $\vec{x}$ em relação à base e. É óbvio que a função φ (mostre que φ é linear!) satisfaz as condições do enunciado, visto que, para qualquer $i = 1, 2, \ldots, n$, se tem:

$$\varphi(\vec{e}_i) = \sum_{j=1}^{n} \delta_{ij} \, \vec{y}_j = \vec{y}_i$$

Fica, pois, demonstrada a *existência* de uma função linear φ nas condições do enunciado. Vejamos, agora, que φ é *única*: para isso, consideremos uma segunda função linear $\varphi': E \to F$ tal que

$$\varphi(\vec{e}_i) = \varphi'(\vec{e}_i) = \vec{y}_i; \; i = 1, 2, \ldots, n$$

As igualdades anteriores equivalem a

$$\varphi(\vec{e}_i) - \varphi'(\vec{e}_i) = (\varphi - \varphi')(\vec{e}_i) = \vec{o}_F; \; i = 1, 2, \ldots, n$$

Portanto, $\varphi - \varphi'$ anula-se em todos os vectores da base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$; pela alínea anterior, será $\varphi - \varphi' = \mathbb{O}_{FE}$ e, portanto, $\varphi = \varphi'$. ☐

Podemos, em consequência da proposição anterior, tirar importantes conclusões sobre funções lineares entre espaços de dimensão finita, como se apresenta no seguinte

Corolário 3.6.1 *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$ com dimensões finitas n e m, respectivamente e $h : E \to F$ uma função linear com característica r e nulidade s. Então:*

i) *h é injectiva sse $r = n$.*

ii) *h é sobrejectiva sse $r = m$.*

iii) *Se h é um isomorfismo de E sobre F, então $m = n$. Isto significa que, se $E \cong F$, tem-se necessariamente* $\dim E = \dim F$.

iv) $0 \leq r \leq \min\{m, n\}$.

v) $\max\{0, n - m\} \leq s \leq n$.

vi) *h é sobrejectiva $\Rightarrow m \leq n$.*

vii) *h é injectiva $\Rightarrow n \leq m$.*

viii) *$m = n \Rightarrow (h$ é injectiva $\Leftrightarrow h$ é sobrejectiva $\Leftrightarrow h$ é um isomorfismo de E sobre $F)$.*

ix) *Um endomorfismo de um espaço com dimensão finita é injectivo sse for sobrejectivo.*

x) *$h: E \to E$ é um automorfismo de $E \Leftrightarrow r = n$. (h diz-se também invertível ou regular)*

Demonstração:

Sejam $s = \text{dimKer}(h) \geq 0$ e $r = \text{dimImg}(h) \geq 0$, respectivamente, a *nulidade* e a *característica* de h e mostremos as relações anteriores:

i) h é injectiva $\Leftrightarrow \text{Ker}(h) = \{\vec{o}_E\} \Leftrightarrow s = 0 = n - r \Leftrightarrow r = n$.

ii) h é sobrejectiva $\Leftrightarrow \text{Img}(h) = F \Leftrightarrow r = m$.

iii) h é um isomorfismo de E sobre $F \Rightarrow h$ é bijectiva $\Rightarrow r = n \wedge r = m \Rightarrow m = n$.

iv) Como $\text{Img}(h) \subset F$, será $r \leq m$; por outro lado, $s + r = n$ implica $r \leq n$ e, portanto,

$$r \leq \min\{m, n\}$$

v) O núcleo é um subespaço de E e, portanto, $s = \text{dimKer}(h) \leq n$. Por outro lado,

$$r \leq m \Rightarrow -m \leq -r \Rightarrow n - m \leq n - r = s \Rightarrow n - m \leq s$$

A última desigualdade, juntamente com $0 \leq s$, implica finalmente

$$\max\{0, n - m\} \leq s \leq n$$

vi) Se h é sobrejectiva, então $r = m$. Porém, vimos em *iv)* que $r \leq n$, ou seja, $m \leq n$.

vii) Se h é injectiva, será $r = n$ e, como é sempre $r \leq m$, resulta $n \leq m$.

viii) Suponha-se que $m = n$: então, temos

$$h \text{ é injectiva} \Leftrightarrow r = n \Leftrightarrow r = m \Leftrightarrow h \text{ é sobrejectiva}$$

Portanto, as condições anteriores significam que quando $m = n$, h é bijectiva: logo, h será um isomorfismo de E sobre F.

ix) É consequência da alínea anterior, visto que a condição $m = n$ se dá necessariamente.

x) Como h é um endomorfismo de E, a condição $m = n$ é automaticamente satisfeita e, portanto, o resultado é consequência imediata da alínea *ix)*. □

Observação:

- O exemplo 3.17 mostra que a conclusão do corolário 3.6.1.*ix* não tem lugar para um endomorfismo de um espaço E com dimensão *infinita*: o operador

$$\text{Int}_a : \mathscr{C}(I, \mathbb{R}) \to \mathscr{C}(I, \mathbb{R});\ f \mapsto F_a$$

é, como se viu então, *injectivo* mas a sua imagem é

$$\text{Img}(\text{Int}_a) = \mathscr{C}^1(I, \mathbb{R}) \cap \{F : F(a) = 0\} \neq \mathscr{C}(I, \mathbb{R})$$

o que mostra que Int_a não é *sobrejectivo*!

A proposição seguinte, mostra que, quando o domínio de uma função linear tem dimensão infinita, o núcleo ou a imagem terão também, necessariamente, dimensão infinita.

Proposição 3.7 *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$ e $h : E \to F$ uma função linear. Se E tem dimensão infinita, então pelo menos um dos espaços* $\mathrm{Ker}(h)$ *ou* $\mathrm{Img}(h)$ *tem dimensão infinita.*

Demonstração:

Suponha-se, por absurdo, que E tem dimensão *infinita* e que $\mathrm{Ker}(h)$ e $\mathrm{Img}(h)$ têm dimensões *finitas* $s = \dim(\mathrm{Ker}(h))$ e $r = \dim(\mathrm{Img}(h))$. Considere-se, então, uma base de $\mathrm{Ker}(h)$

$$e_{\,\mathrm{Ker}(h)} = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s)$$

Como E tem dimensão infinita, podemos encontrar em E listas linearmente independentes com quantos vectores quantos queiramos (ver proposição 1.17): consequentemente, ampliemos a base anterior, acrescentando-lhe um número p de vectores superior a r, para obter a sequência *linearmente independente* de $s + p$ vectores de E:

$$e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s, \vec{e}_{s+1}, \vec{e}_{s+2}, \ldots, \vec{e}_{s+p}); \text{ onde } p > r$$

(claro que $h(\vec{e}_j) = \vec{o}_F$, para $j = 1, 2, \ldots, s$ e $h(\vec{e}_j) \neq \vec{o}_F$, para $j = s+1, s+2, \ldots, s+p$).

Consideremos, agora, a sequência u de p vectores de $\mathrm{Img}(h)$, formada pelas imagens dos $(\vec{e}_{s+j})_{1 \leq j \leq p}$:

$$u = (h(\vec{e}_{s+1}), h(\vec{e}_{s+2}), \ldots, h(\vec{e}_{s+p}))$$

Como $p > r = \dim(\mathrm{Img}(h))$, a sequência u é *linearmente dependente*. Portanto existem escalares $(\alpha_j)_{1 \leq j \leq p}$ não todos nulos, tais que

$$\sum_{j=1}^{p} \alpha_j h(\vec{e}_{s+j}) = \vec{o}_F \text{ e em que existe } k \in \{1, 2, \ldots, p\} \text{ tal que } \alpha_k \neq 0$$

Para finalizar a demonstração, atente-se, agora, na sequência de implicações (justifique cada passagem!)

$$\begin{aligned}
\sum_{j=1}^{p} \alpha_j h(\vec{e}_{s+j}) = \vec{o}_F &\Rightarrow h\left(\sum_{j=1}^{p} \alpha_j \vec{e}_{s+j}\right) = \vec{o}_F \\
&\Rightarrow \sum_{j=1}^{p} \alpha_j \vec{e}_{s+j} \in \mathrm{Ker}(h) \\
&\Rightarrow \sum_{j=1}^{p} \alpha_j \vec{e}_{s+j} = \sum_{j=1}^{s} \beta_j \, \vec{e}_j \\
&\Rightarrow \sum_{j=1}^{s} (-\beta_j) \vec{e}_j + \sum_{j=1}^{p} \alpha_j \vec{e}_{s+j} = \vec{o}_E \\
&\Rightarrow \text{A lista } e \text{ será } \textit{linearmente dependente}\text{, o que é absurdo.}
\end{aligned}$$

□

A proposição 3.2 permite interpretar o problema dos sistemas de equações lineares à luz da teoria das funções lineares: de facto, seja $AX = B$ a forma matricial de um sistema de m equações nas n incógnitas $X = [\,x_1 \;\, x_2 \;\, \cdots \;\, x_n\,]^{\mathrm{T}} \in \mathbb{K}^n$ e onde $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^m$. A função

$$h_A\colon \mathbb{K}^n \to \mathbb{K}^m;\ X \mapsto AX = \sum_{j=1}^{n} x_j A_j,\ \text{onde os } A_j \text{ são as colunas de } A,$$

é linear, como imediatamente se reconhece, e o conjunto S das soluções do sistema é formado pelos vectores de $\mathbb{K}^n$ cuja imagem por h_A é igual ao vector $B \in \mathbb{K}^m$ dos termos independentes do sistema. Portanto, trata-se afinal de calcular a pré-imagem de B por meio de h_A:

$$S = h_A^{-1}(B) = \{X\colon AX = B\} \subset \mathbb{K}^n$$

Seja $r = \mathrm{c}(A) = \dim(L(A_1, A_2, \ldots, A_n))$. A definição de h_A mostra que:

- A imagem de h_A é o subespaço gerado por $(A_1, A_2, \ldots, A_n)$

  $$\mathrm{Img}(h_A) = L(A_1, A_2, \ldots, A_n)$$

  portanto, tem-se:

  $$\mathrm{c}(h_A) = \dim(\mathrm{Img}(h_A)) = \dim(L(A_1, A_2, \ldots, A_n)) = r$$

As igualdades 3.6.1, por seu lado, mostram que:

- O sistema tem solução sse $B \in \mathrm{Img}(h_A) = L(A_1, A_2, \ldots, A_n)$ ou seja,

  $$L(A_1, A_2, \ldots, A_n) = L(A_1, A_2, \ldots, A_n, B)$$

  o que equivale ainda a:

  $$r = \mathrm{c}\big(\big[A \,\big|\, B\big]\big)$$

- O núcleo de h_A é o conjunto S_0 das soluções do sistema homogéneo associado $AX = O$:

  $$S_0 = \mathrm{Ker}(h_A) = \{X\colon AX = O\} \subset \mathbb{K}^n$$

  e a dimensão deste subespaço de $\mathbb{K}^n$ é o *grau de indeterminação* do sistema

  $$\dim(S_0) = \dim(\mathrm{Ker}(h_A)) = \dim(\mathbb{K}^n) - \dim(\mathrm{Img}(h_A)) = n - \mathrm{c}(A) = n - r$$

- Se $B \in \mathrm{Img}(h_A)$, as soluções do sistema $AX = B$ obtêm-se de uma sua solução particular X_0 somando-lhe o núcleo de h_A, ou seja, as soluções Y do sistema homogéneo associado $AX = O$:

  $$X = X_0 + Y$$

- O sistema é *determinado* sse $B \in \mathrm{Img}(h_A)$ e h_A for *injectiva* (o que equivale a $\mathrm{Ker}(h_A) = \{\vec{o}\}$ ou $r = n$), ou seja:

  $$AX = B \text{ é } \textit{determinado} \Leftrightarrow r = \mathrm{c}\big(\big[A \,\big|\, B\big]\big) \wedge r = n$$

- Se $r = m$, então h_A é *sobrejectiva* e o sistema $AX = B$ é *possível*, qualquer que seja o vector $B \in \mathbb{K}^m$.

- Se $m < n$, então é $r < n$ e h_A não será injectiva, ou seja, o sistema não é determinado.

Exemplo 3.19 Considere, na figura 3.8, a aplicação linear $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ que transforma o quadrado da esquerda no da direita, sendo cada triângulo levado por h no correspondente triângulo da mesma côr.

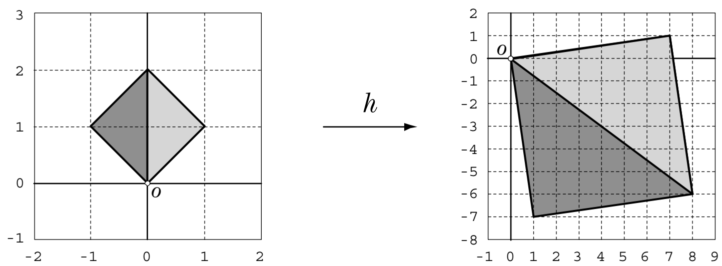


Fig. 3.8 – Acção da transformação h sobre um quadrado.

Determinemos $h(x,y)$, para todo o vector $(x,y) \in \mathbb{R}^2$. A imagem anterior mostra-nos que[2]

$$\begin{cases} h(1,1) &= (7,1) \\ h(-1,1) &= (1,-7) \end{cases}$$

Basta agora exprimir (x,y) na base $\big((1,1),(-1,1)\big)$, através das suas coordenadas (a,b) nesta base:

$$(x,y) = a(1,1) + b(-1,1) = (a-b, a+b)$$

Daqui resulta o sistema de equações lineares

$$\begin{cases} a - b = x \\ a + b = y \end{cases}$$

cuja solução é $a = \frac{1}{2}(x+y), b = \frac{1}{2}(y-x)$; portanto, tem-se

$$(x,y) = \frac{1}{2}(x+y)(1,1) + \frac{1}{2}(y-x)(-1,1)$$

Basta, agora, aplicar h a ambos os membros da igualdade vectorial precedente e atender à linearidade de h, para obtermos

$$\begin{aligned} h(x,y) &= \frac{1}{2}(x+y)h(1,1) + \frac{1}{2}(y-x)h(-1,1) \\ &= \frac{1}{2}(x+y)(7,1) + \frac{1}{2}(y-x)(1,-7) \\ &= (3x+4y, 4x-3y) \end{aligned}$$

3.5 Representação matricial de uma função linear

Na presente secção, vamos estabelecer um dos resultados mais importantes sobre funções lineares: o da correspondência entre as funções lineares (entre dois espaços de dimensão finita) e as matrizes. Após fixarmos bases no domínio (com dimensão n) e no espaço de chegada (com

---

2 É claro que poderíamos substituir qualquer destas igualdades por $h(0,2) = (8,-6)$.

dimensão m) de uma função linear h, fica univocamente determinada uma matriz de tipo $m \times n$, que se dirá a *representação matricial de* h em relação às referidas bases. Esta matriz depende não só de h, mas também das duas bases escolhidas.

Proposição 3.8 – Matriz de um operador linear – *Seja E um espaço vectorial, com dimensão n sobre um corpo $\mathbb{K}$, e $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ uma sua base. Seja, por outro lado, F um segundo espaço vectorial sobre o mesmo corpo, com dimensão m, e seleccione-se uma base $f = \left(\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\right)$ de F. Se $h\colon E \to F$ é uma função linear, existe uma e uma só matriz $M_{fe}(h) = [h_{ij}] \in \mathbb{K}^{m,n}$ tal que:*

i) *Entre as imagens dos $\vec{e}_j$ por h e os $\vec{f}_i$ verificam-se as relações*

$$h(\vec{e}_j) = \sum_{i=1}^{m} h_{ij}\vec{f}_i;\ j = 1, 2, \ldots, n \tag{3.17}$$

ou seja, matricialmente,

$$\left[\, h(\vec{e}_1)\ \ h(\vec{e}_2)\ \ \cdots\ \ h(\vec{e}_n) \,\right] = \left[\, \vec{f}_1\ \ \vec{f}_2\ \ \cdots\ \ \vec{f}_m \,\right] M_{fe}(h) \tag{3.18}$$

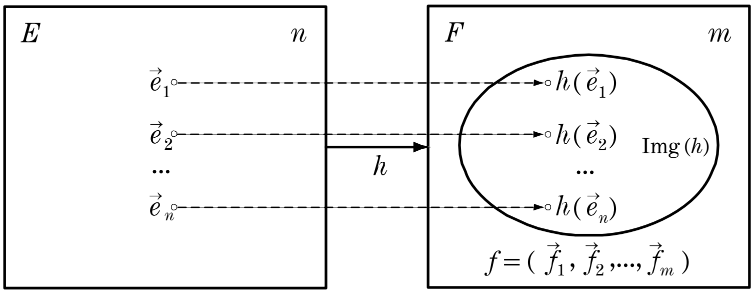


Fig. 3.9 – Diagrama auxiliar para a proposição 3.8.

ii) *Se $(x_1, x_2, \ldots, x_n)$ forem as coordenadas de um vector $\vec{x} \in E$ em relação à base e e $(y_1, y_2, \ldots, y_m)$ as coordenadas da sua imagem $\vec{y} = h(\vec{x})$ em relação à base f, tem-se:*

$$y_i = \sum_{j=1}^{n} h_{ij}x_j;\ i = 1, 2, \ldots, m \tag{3.19}$$

matricialmente, fica

$$Y_f = M_{fe}(h)X_e \tag{3.20}$$

onde

$$\begin{cases} X_e = \left[\, x_1\ \ x_2\ \ \cdots\ \ x_n \,\right]^{\mathrm{T}} \\ Y_f = \left[\, y_1\ \ y_2\ \ \cdots\ \ y_m \,\right]^{\mathrm{T}} \end{cases} \tag{3.21}$$

iii) *Tem-se sempre:*

$$\mathrm{c}(h) = \mathrm{c}(M_{fe}(h)) \tag{3.22}$$

Demonstração:

i) Os vectores $h(\vec{e}_j)$ pertencem a F e, como f é uma base de F, cada um deles exprime-se de *uma e uma só* forma como combinação linear dos $\vec{f}_i$ através das suas m *coordenadas* $(h_{ij})_{1\leq i\leq m}$ em relação à base f: portanto, a matriz $M_{fe}(h) = [h_{ij}]$, do tipo $m \times n$ tal que as relações (3.17) ou (3.18) se verificam é *única* e determinada *exclusivamente por* h *e pelo par de bases* (e, f).

ii) Por outro lado, tem-se, para qualquer vector $\vec{x} \in E$:

$$\begin{aligned}\vec{y} = h(\vec{x}) &= h\left(\sum_{j=1}^{n} x_j\vec{e}_j\right) \\ &= \sum_{j=1}^{n} x_j h(\vec{e}_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} h_{ij}\vec{f}_i \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} h_{ij}x_j\right)\vec{f}_i = \sum_{i=1}^{m} y_i\vec{f}_i\end{aligned}$$

Como as coordenadas de $\vec{y}$ na base f são únicas, deverá ter-se:

$$y_i = \sum_{j=1}^{n} h_{ij}x_j;\ i = 1, 2, \ldots, m$$

o que equivale à expressão matricial (3.20), onde as matrizes-coluna X_e e Y_f contêm, respectivamente, as coordenadas de $\vec{x}$ e da sua imagem $\vec{y}$ por h em relação às bases e e f, (segundo as expressões (3.21)) e $M_{fe}(h) = [h_{ij}]$.

iii) Comecemos por recordar que a característica de h é a *dimensão* do subespaço Img(h) e que a característica de $M_{fe}(h)$ é a *dimensão* do subespaço gerado pelas suas colunas. Mas, pela proposição 3.6.*i*, tem-se Img$(h) = L(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_n))$ e o resultado é, agora, uma consequência imediata do facto de as colunas de $M_{fe}(h)$ conterem precisamente as coordenadas dos $h(\vec{e}_j)$ em relação à base f (o espaço Img(h) é isomorfo do subespaço gerado pelas colunas de $M_{fe}(h)$). □

A matriz $M_{fe}(h)$ é chamada a *representação matricial de* h, em relação ao par de bases (e, f).

Observações:

- A matriz $M_{fe}(h)$ tem um número de linhas m igual à dimensão do espaço F de chegada e um número de colunas n igual à dimensão do espaço E domínio de h.

- A coluna j da matriz $M_{fe}(h)$ contém as *coordenadas* do vector $h(\vec{e}_j)$, relativamente à base f.

- No caso de h ser um *endomorfismo* de E, a mesma base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ serve para o domínio E e para o codomínio (embora não seja obrigatoriamente assim) e usaremos a notação $M_e(h)$ (em vez de $M_{ee}(h)$) para a representação matricial (quadrada de ordem n) de h em relação à base e comum.

Seguem-se alguns exemplos:

Exemplo 3.20 Seja $\mathscr{I}_E\colon E \to E$ a função identidade em E. Consideremos agora duas bases $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ e $e' = \bigl(\vec{e}\,'_1, \vec{e}\,'_2, \dots, \vec{e}\,'_n\bigr)$ de E e determinemos a representação matricial de $\mathscr{I}_E$ em relação às bases citadas (e' no domínio e e no codomínio): tem-se $\mathscr{I}_E\bigl(\vec{e}\,'_j\bigr) = \vec{e}\,'_j$, para todo o $j = 1, 2, \dots, n$, donde

$$\bigl[\,\mathscr{I}_E\bigl(\vec{e}\,'_1\bigr)\ \mathscr{I}_E\bigl(\vec{e}\,'_2\bigr)\ \cdots\ \mathscr{I}_E\bigl(\vec{e}\,'_n\bigr)\,\bigr] = [\,\vec{e}\,'_1\ \vec{e}\,'_2\ \cdots\ \vec{e}\,'_n\,] = [\,\vec{e}_1\ \vec{e}_2\ \cdots\ \vec{e}_n\,]M_{ee'}(\mathscr{I}_E)$$

Comparando com (2.77), obtemos

$$M_{ee'}(\mathscr{I}_E) = T_{ee'} \tag{3.23}$$

Vemos assim que a matriz que representa $\mathscr{I}_E$ nas bases e e e' é a *matriz de mudança* da base e para a base e' (ver secção 2.15). Em particular, quando $e' = e$, (2.79) mostra-nos que

$$M_{ee}(\mathscr{I}_E) = T_{ee} \Rightarrow M_e(\mathscr{I}_E) = I_n \tag{3.24}$$

o que mostra que a representação matricial da função identidade não depende da base e (adoptada para o domínio e para o codomínio).

Exemplo 3.21 Seja $E \subset F$ um subespaço com dimensão n de um espaço vectorial F de dimensão $m \geq n$ sobre um corpo $\mathbb{K}$ e considere-se a *injecção canónica* $j\colon E \to F;\ \vec{x} \mapsto \vec{x}$. Seja, agora, $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ uma base de E e $f = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n, \vec{e}_{n+1}, \vec{e}_{n+2}, \dots, \vec{e}_m)$ um prolongamento de e a uma base f de F (que existe, pela proposição 1.16.*v*). Tem-se, claro está, $j(\vec{e}_k) = \vec{e}_k$, para $k = 1, 2, \dots, n$. Portanto, a representação matricial de j, em relação ao par de bases (e, f) será:

$$M_{fe}(h) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} I_n \\ O_{m-n,n} \end{bmatrix}$$

Exemplo 3.22 Sejam E e F espaços vectoriais de dimensões n e m respectivamente e $\mathscr{O}_{FE}\colon E \to F;\ \vec{x} \mapsto \vec{o}_F$ a função nula. Para qualquer par de bases $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ e $f = \bigl(\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\bigr)$, tem-se $\mathscr{O}_{FE}(\vec{e}_k) = \vec{o}_F = \sum\limits_{i=1}^{m} 0\vec{f}_i$ e, portanto,

$$M_{fe}(\mathscr{O}_{FE}) = O_{m,n}$$

Observe-se que, também aqui, a matriz não depende da escolha das bases.

Exemplo 3.23 Seja E um espaço vectorial de dimensão finita n sobre um corpo $\mathbb{K}$ e $h: E \to E$ o endomorfismo de E mencionado no exemplo 3.5

$$h_k(\vec{x}) = k\vec{x} = k\mathscr{I}_E(\vec{x})$$

Para qualquer base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$, tem-se $h_k(\vec{e}_i) = k\vec{e}_i$; logo, a matriz de h_k nessa base e será:

$$M_e(h_k) = \begin{bmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k \end{bmatrix} = \text{diag}(k, k, \ldots, k) = kI_n = kM_e(\mathscr{I}_E)$$

Exemplo 3.24 Seja $h: E \to E$ um endomorfismo de um espaço vectorial E de dimensão n sobre um corpo $\mathbb{K}$ e suponha-se a existência de uma base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de E e de escalares $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ tais que, para $j = 1, 2, \ldots, n$,

$$h(\vec{e}_j) = \lambda_j \vec{e}_j$$

A representação matricial de h, em relação à base e será a matriz diagonal de ordem n

$$M_e(h) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

Exemplo 3.25 Sendo $\mathbb{K}$ um corpo e $\vec{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{K}^n$, considere-se a função linear $h_{\vec{a}}: \mathbb{K}^n \to \mathbb{K}$ definida, para todo o $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$, por (ver exemplo 3.10)

$$h_{\vec{a}}(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} a_j\, x_j$$

Daqui resulta que

$$h_{\vec{a}}(\vec{e}_i) = \sum_{j=1}^{n} a_j\, \delta_{ij} = a_i$$

onde $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ é a base canónica de $\mathbb{K}^n$. Portanto, sendo $f = (1)$ a base canónica de $\mathbb{K}$, tem-se:

$$M_{fe}(h_{\vec{a}}) = [\, a_1 \;\; a_2 \;\; \cdots \;\; a_n \,]$$

Exemplo 3.26 Em relação ao exemplo 3.12, a representação matricial de $h: \mathbb{K}^n \to \mathbb{K}^m$ em relação às bases canónicas de $\mathbb{K}^n$ e $\mathbb{K}^m$ é matriz $A = [a_{ij}]$ aí definida.

Exemplo 3.27 A função $D: \mathscr{P}_n(\mathbb{R}) \to \mathscr{P}_n(\mathbb{R})$ que transforma cada polinómio de grau inferior ou igual a n na sua derivada é, como se viu no exemplo 3.15, um endomorfismo de

$\mathscr{P}_n(\mathbb{R})$:

$$D\left(\sum_{k=0}^{n} a_k x^k\right) = \sum_{k=0}^{n-1} (k+1) a_{k+1} x^k$$

Considere-se, agora, a base canónica de $\mathscr{P}_n(\mathbb{R})$, $c = (1, x, x^2, \ldots, x^n)$. Ora,

$$\begin{cases} D(1) &= 0 \\ D(x) &= 1 \\ D(x^2) &= 2x \\ \cdots &= \cdots \\ D(x^n) &= nx^{n-1} \end{cases}$$

Daqui se segue que a representação matricial de D, em relação à base c será a matriz de ordem $n+1$:

$$M_c(D) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Exemplo 3.28 Considere-se a função linear (ver exemplo 3.17) Int: $\mathscr{P}_{n-1}(\mathbb{R}) \to \mathscr{P}_n(\mathbb{R})$, que transforma cada polinómio $p \in \mathscr{P}_{n-1}(\mathbb{R})$ de coeficientes reais e de grau menor ou igual a $n-1$ no seu integral indefinido $P \in \mathscr{P}_n(\mathbb{R})$ com origem 0:

$$\begin{cases} \text{Int}(p) = P, \quad \text{onde } P(x) = \displaystyle\int_0^x p(t)\mathrm{d}t \\ \text{Int}\left(\displaystyle\sum_{k=0}^{n-1} a_k x^k\right) = \displaystyle\sum_{k=0}^{n-1} \tfrac{a_k}{k+1} x^{k+1} \end{cases}$$

Daqui resulta:

$$\begin{cases} \text{Int}(1) &= x \\ \text{Int}(x) &= x^2/2 \\ \text{Int}(x^2) &= x^3/3 \\ \cdots & \cdots \\ \text{Int}(x^{n-1}) &= x^n/n \end{cases}$$

e, portanto, a representação de Int em relação às bases canónicas $c = (1, x, x^2, \ldots, x^{n-1})$ e

$c' = (1, x, x^2, \ldots, x^n)$ de $\mathcal{P}_{n-1}(\mathbb{R})$ e de $\mathcal{P}_n(\mathbb{R})$ será a matriz de tipo $(n+1) \times n$:

$$M_{c'c}(\text{Int}) = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \frac{1}{n} \end{bmatrix}$$

Exemplo 3.29 Considere, em $\mathbb{R}^2$, a base $e = ((1,1),(1,-1))$ e a função linear $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ definida por:

$$h(x,y) = (2x - y, x + y)$$

Para calcular a representação matricial de h, em relação à base e, temos de calcular as coordenadas de $h(1,1) = (1,2)$ e de $h(1,-1) = (3,0)$ em relação àquela base, para o que basta resolver os dois sistemas de equações lineares seguintes (que sabemos, de antemão, serem determinados – porquê?)

$$\begin{cases} \alpha + \beta = 1 \\ \alpha - \beta = 2 \end{cases} \qquad \begin{cases} \alpha + \beta = 3 \\ \alpha - \beta = 0 \end{cases}$$

Como a matriz dos coeficientes dos dois sistemas é a mesma, podemos resolvê-los em simultâneo, condensando a matriz seguinte

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & -1 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & 1 & -3 \end{bmatrix} \to \begin{bmatrix} 2 & 0 & 3 & 3 \\ 0 & -2 & 1 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 3/2 & 3/2 \\ 0 & 1 & -1/2 & 3/2 \end{bmatrix}$$

Atendendo às soluções dos sistemas, temos:

$$\begin{cases} h(1,1) \quad = (1,2) = \frac{3}{2}(1,1) - \frac{1}{2}(1,-1) \\ h(1,-1) = (3,0) = \frac{3}{2}(1,1) + \frac{3}{2}(1,-1) \end{cases}$$

A matriz de h é, pois:

$$M_e(h) = \frac{1}{2}\begin{bmatrix} 3 & 3 \\ -1 & 3 \end{bmatrix}$$

A matriz de h em relação à base canónica de $\mathbb{R}^2$, $c = ((1,0),(0,1))$, seria, como facilmente se conclui:

$$M_c(h) = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$$

Exemplo 3.30 Consideremos, em S^2, uma base $e = (\vec{e}_1, \vec{e}_2)$ formada por dois segmentos unitários e ortogonais (ver figura 3.10) aplicados num ponto O e a aplicação $\rho_\theta\colon S^2 \to S^2$ que roda cada segmento aplicado em O de um certo ângulo θ em torno de O.

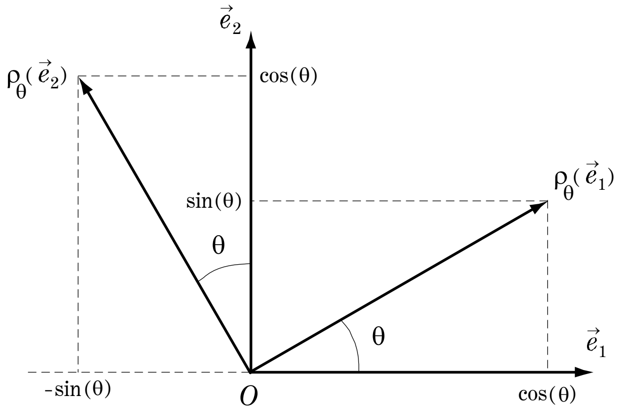


Fig. 3.10 – Rotação em S^2.

Da figura conclui-se facilmente que

$$\begin{cases} \rho_\theta(\vec{e}_1) = \cos(\theta)\vec{e}_1 + \sin(\theta)\vec{e}_2 \\ \rho_\theta(\vec{e}_2) = -\sin(\theta)\vec{e}_1 + \cos(\theta)\vec{e}_2 \end{cases}$$

Matricialmente, fica

$$[\,\rho_\theta(\vec{e}_1)\ \ \rho_\theta(\vec{e}_2)\,] = [\,\vec{e}_1\ \ \vec{e}_2\,]\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

o que mostra que a matriz de ρ_θ em relação à base e será:

$$M_e(\rho_\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Exemplo 3.31 Seja $e = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ a base canónica de $\mathbb{R}^3$ e considere uma função linear $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ que transforma o cubo dos vectores $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$, onde $x_1, x_2, x_3 \in [0, 1]$, na região plana sombreada da figura abaixo.

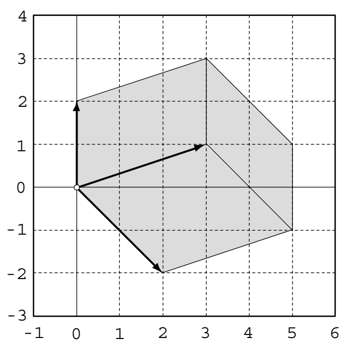


Fig. 3.11 – Imagem do cubo $[0, 1]^3 \subset \mathbb{R}^3$ por meio de f.

A matriz que representa f em relação às bases canónicas de $\mathbb{R}^3$ e de $\mathbb{R}^2$ pode ser determinada, se soubermos que $(3, -2, 4)$ é um vector do núcleo de f. A figura mostra que as imagens dos vectores da base canónica $e = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ de $\mathbb{R}^3$ são os vectores $(2, -2), (3, 1)$ e $(0, 2)$. O que a imagem não nos mostra é qual destes vectores é $f(\vec{e}_1)$, $f(\vec{e}_2)$ e $f(\vec{e}_3)$ respectivamente. Existem $3! = 6$ formas diferentes de ordenar os vectores $(2, -2), (3, 1)$ e $(0, 2)$. Mas o facto de sabermos que $(3, -2, 4)$ é um vector do núcleo de f e que f é linear permitirá esclarecer a questão, visto que terá que ser

$$f(3,-2,4) = f(3\vec{e}_1 - 2\vec{e}_2 + 4\vec{e}_3) = 3f(\vec{e}_1) - 2f(\vec{e}_2) + 4f(\vec{e}_3) = (0,0)$$

As seis correspondências possíveis encontram-se na tabela seguinte:

$f(\vec{e}_1)$	$f(\vec{e}_2)$	$f(\vec{e}_3)$	$3f(\vec{e}_1) - 2f(\vec{e}_2) + 4f(\vec{e}_3)$
$(3,1)$	$(0,2)$	$(2,-2)$	$(17,-9)$
$(3,1)$	$(2,-2)$	$(0,2)$	$(5,15)$
$(2,-2)$	$(3,1)$	$(0,2)$	$(0,0)$
$(2,-2)$	$(0,2)$	$(3,1)$	$(18,-6)$
$(0,2)$	$(2,-2)$	$(3,1)$	$(8,14)$
$(0,2)$	$(3,1)$	$(2,-2)$	$(2,-4)$

Dos casos possíveis, apenas um (o terceiro caso, a sombreado) torna nulo o vector $3f(\vec{e}_1) - 2f(\vec{e}_2) + 4f(\vec{e}_3)$, pelo que se conclui que as imagens dos vectores da base canónica de $\mathbb{R}^3$ por f são necessariamente as que se seguem:

$$\begin{cases} f(\vec{e}_1) = (2,-2) \\ f(\vec{e}_2) = (3,1) \\ f(\vec{e}_3) = (0,2) \end{cases}$$

Como as coordenadas dos vectores $f(\vec{e}_i)$ em relação à base canónica c de $\mathbb{R}^2$ coincidem com as suas próprias componentes, a matriz de f em relação às duas bases canónicas será, de imediato,

$$M_{ce}(f) = \begin{bmatrix} 2 & 3 & 0 \\ -2 & 1 & 2 \end{bmatrix}$$

Podemos, agora, determinar a função linear f:

$$\begin{aligned} f(x,y,z) &= f(x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3) = xf(\vec{e}_1) + yf(\vec{e}_2) + zf(\vec{e}_3) \\ &= x(2,-2) + y(3,1) + z(0,2) = (2x+3y, -2x+y+2z) \end{aligned}$$

3.6 Isomorfismo entre $\mathscr{L}(E,F)$ e $\mathbb{K}^{m,n}$

Veremos neste parágrafo que a correspondência entre as funções lineares e as matrizes é um *isomorfismo* de espaços lineares e que, no caso dos *endomorfismos*, é um isomorfismo de álgebras lineares:

Proposição 3.9 *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$ (com dimensões n e m, respectivamente) e $h, k\colon E \to F$ duas funções lineares. Sejam ainda $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ uma base de E e $f = \left(\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\right)$ uma base de F. Então:*

i) $M_{fe}(h+k) = M_{fe}(h) + M_{fe}(k)$ (3.25)

ii) $M_{fe}(\alpha h) = \alpha M_{fe}(h)$*, onde* $\alpha \in \mathbb{K}$*.* (3.26)

Demonstração:

i) Seja $u = h + k$. Devido a (3.17), tem-se:

$$\begin{cases} h(\vec{e}_j) = \sum_{i=1}^{m} h_{ij}\vec{f}_i \\ k(\vec{e}_j) = \sum_{i=1}^{m} k_{ij}\vec{f}_i \\ u(\vec{e}_j) = \sum_{i=1}^{m} u_{ij}\vec{f}_i \end{cases} \qquad \text{onde } j = 1, 2, \ldots, n$$

e onde

$$\begin{cases} M_{fe}(h) = [h_{ij}] \in \mathbb{K}^{m,n} \\ M_{fe}(k) = [k_{ij}] \in \mathbb{K}^{m,n} \\ M_{fe}(u) = [u_{ij}] \in \mathbb{K}^{m,n} \end{cases}$$

Para calcularmos a representação matricial de $u = h + k$, calculemos as imagens dos vectores da base e por meio de $h + k$ e atendamos à expressão (3.17), aplicada a h e a k, e à definição de adição de funções lineares:

$$\begin{cases} u(\vec{e}_j) = (h + k)(\vec{e}_j) = h(\vec{e}_j) + k(\vec{e}_j) \\ \qquad\;\; = \sum_{i=1}^{m} h_{ij}\vec{f}_i + \sum_{i=1}^{m} k_{ij}\vec{f}_i \\ \qquad\;\; = \sum_{i=1}^{m} \left(h_{ij}\vec{f}_i + k_{ij}\vec{f}_i\right) \\ \qquad\;\; = \sum_{i=1}^{m} (h_{ij} + k_{ij})\vec{f}_i \end{cases} \qquad \text{onde } j = 1, 2, \ldots, n$$

Porém, as coordenadas dos vectores $u(\vec{e}_j)$ na base f são únicas, pelo que terá de ser:

$$u_{ij} = h_{ij} + k_{ij} \;\text{ onde } \begin{cases} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{cases}$$

e as $m \times n$ igualdades anteriores equivalem à igualdade matricial desejada:

$$M_{fe}(h + k) = M_{fe}(h) + M_{fe}(k)$$

ii) Seja $u = \alpha h$. As igualdades (3.17), escrevem-se

$$\begin{cases} h(\vec{e}_j) = \sum_{i=1}^{m} h_{ij}\vec{f}_i \\ u(\vec{e}_j) = \sum_{i=1}^{m} u_{ij}\vec{f}_i \end{cases} \qquad \text{em que } j = 1, 2, \ldots, n$$

e

$$\begin{cases} M_{fe}(h) = [h_{ij}] \in \mathbb{K}^{m,n} \\ M_{fe}(u) = [u_{ij}] \in \mathbb{K}^{m,n} \end{cases}$$

Para calcularmos a representação matricial de $u = \alpha h$, calculemos as imagens dos vectores da base e por meio de αh e atendamos à expressão (3.17), aplicada a h, e à definição de produto de escalar por função linear:

$$\begin{cases} u(\vec{e}_j) = (\alpha h)(\vec{e}_j) = \alpha h(\vec{e}_j) \\ \qquad\quad = \alpha \sum\limits_{i=1}^{m} h_{ij}\vec{f}_i \\ \qquad\quad = \sum\limits_{i=1}^{m} (\alpha h_{ij})\vec{f}_i \end{cases} \qquad \text{onde } j = 1, 2, \ldots, n$$

Porém, as coordenadas dos vectores $u(\vec{e}_j)$ na base f são únicas, pelo que terá de ser

$$u_{ij} = \alpha h_{ij} \ \text{ onde} \begin{cases} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{cases}$$

e as $m \times n$ igualdades anteriores equivalem à igualdade matricial do enunciado:

$$M_{fe}(\alpha h) = \alpha M_{fe}(h)$$ □

A proposição anterior mostra que a função $M_{fe}: \mathscr{L}(E,F) \to \mathbb{K}^{m,n};\ h \mapsto M_{fe}(h)$ que transforma cada função linear na respectiva matriz (do tipo $m \times n$) em relação ao par de bases (e, f) satifaz as condições L1 e L2 e é, portanto, *linear*; além disso, a equação (3.20) mostra que, se $M_{fe}(h) = O_{m,n}$, então $h = \mathscr{O}_{FE}$ o que significa que Ker$(M_{fe}) = \{\mathscr{O}_{FE}\}$ e que, em virtude da proposição 3.2, M_{fe} é *injectiva* e, obviamente, também *sobrejectiva* (porquê?), ou seja, um *isomorfismo* de $\mathscr{L}(E,F)$ sobre $\mathbb{K}^{m,n}$: podemos então escrever

$$\begin{aligned} &\mathscr{L}(E,F) \cong \mathbb{K}^{m,n} \\ &\text{End}(E) \cong \mathbb{K}^{n,n} \end{aligned} \tag{3.27}$$

O isomorfismo anterior permite concluir que:

$$\begin{aligned} &\dim(\mathscr{L}(E,F)) = mn \\ &\dim(\text{End}(E)) = n^2 \end{aligned} \tag{3.28}$$

De tudo o que acima se disse resulta que a escolha de um par de bases (e, f) em E e F determina um isomorfismo M_{fe} de $\mathscr{L}(E,F)$ sobre $\mathbb{K}^{m,n}$ e que existem tantos isomorfismos quantos os pares (e, f) de bases de E e F.

De seguida, veremos que também o *produto* de matrizes corresponde à *composição* de funções lineares:

Proposição 3.10 *Sejam E, F e G espaços vectoriais sobre um corpo $\mathbb{K}$ (com dimensões n e m e p, respectivamente) e $k: G \to E$ e $h: E \to F$ duas funções lineares. Sejam ainda*

$$\begin{aligned} e &= (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) \\ f &= (\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m) \\ g &= (\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_p) \end{aligned}$$

bases de E, F e G, respectivamente. Nestas condições, tem-se:

$$M_{fg}(h \circ k) = M_{fe}(h)M_{eg}(k) \tag{3.29}$$

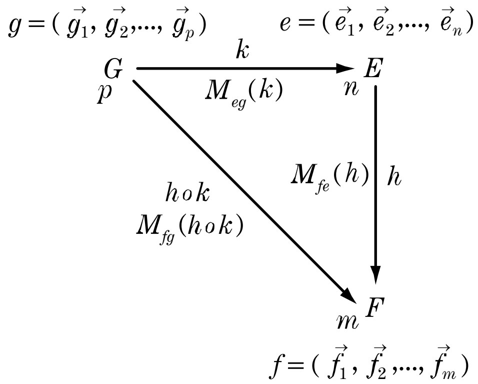


Fig. 3.12 – Matriz da aplicação composta.

Demonstração:

Seja $u = h \circ k$. Devido a (3.17), tem-se:

$$\begin{cases} h(\vec{e}_j) = \sum\limits_{i=1}^{m} h_{ij}\vec{f}_i \ \text{ onde } j = 1, 2, \ldots, n \\ k(\vec{g}_r) = \sum\limits_{j=1}^{n} k_{jr}\vec{e}_j \ \text{ onde } r = 1, 2, \ldots, p \\ u(\vec{g}_r) = \sum\limits_{i=1}^{m} u_{ir}\vec{f}_i \ \text{ onde } r = 1, 2, \ldots, p \end{cases}$$

e onde

$$\begin{cases} M_{fe}(h) = [h_{ij}] \in \mathbb{K}^{m,n} \\ M_{eg}(k) \ = [k_{jr}] \in \mathbb{K}^{n,p} \\ M_{fg}(u) = [u_{ir}] \in \mathbb{K}^{m,p} \end{cases}$$

Para calcularmos a representação matricial de $u = h \circ k$, calculemos as imagens dos vectores da base g por meio de $h \circ k$, de acordo com a expressão (3.17) e com a definição de composição:

$$\begin{aligned} u(\vec{g}_r) &= (h \circ k)(\vec{g}_r) = h(k(\vec{g}_r)) \\ &= h\left(\sum_{j=1}^{n} k_{jr}\vec{e}_j\right) = \sum_{j=1}^{n} k_{jr} h(\vec{e}_j) \\ &= \sum_{j=1}^{n} k_{jr} \sum_{i=1}^{m} h_{ij}\vec{f}_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} h_{ij}k_{jr}\right)\vec{f}_i \end{aligned} \qquad \text{onde } r = 1, 2, \ldots, p$$

Porém, as coordenadas dos vectores $u(\vec{g}_r)$ na base f são únicas, pelo que terá de ser:

$$u_{ir} = \sum_{j=1}^{n} h_{ij}k_{jr} \ \text{ onde } \begin{cases} i \ = 1, 2, \ldots, m \\ r = 1, 2, \ldots, p \end{cases}$$

e as $m \times p$ igualdades escalares anteriores equivalem à igualdade matricial (3.29):

$$M_{fg}(h \circ k) = M_{fe}(h) M_{eg}(k) \qquad \square$$

As equações (3.25), (3.26) e (3.29) significam que a função $M_e\colon \mathrm{End}(E) \to \mathbb{K}^{n,n}$ é um *isomorfismo* dos *anéis* (não comutativos e com unidade) $(\mathrm{End}(E), +, \circ)$ e $(\mathbb{K}^{n,n}, +, \times)$ e também das correspondentes *álgebras lineares*.

Se $E = F = G$ e h for um automorfismo do espaço n-dimensional E (o que equivale a ser $\mathrm{c}(h) = n$) e h^{-1} o automorfismo inverso, então (3.29) dá-nos, para qualquer base de E,

$$\begin{aligned} I_n &= M_e(\mathscr{I}_E) = M_e\big(h \circ h^{-1}\big) = M_e(h) M_e\big(h^{-1}\big) \\ I_n &= M_e(\mathscr{I}_E) = M_e\big(h^{-1} \circ h\big) = M_e\big(h^{-1}\big) M_e(h) \end{aligned}$$

Isto que mostra que

$$M_e\big(h^{-1}\big) = \big(M_e(h)\big)^{-1} \tag{3.30}$$

e é claro que h *é invertível se e só se o mesmo acontece com a matriz* $M_e(h)$, ou seja, se e só se ambos têm característica igual a n, ou seja ambos são *regulares*.

3.7 Alteração da representação matricial nas mudanças de base

Vamos ver, nesta secção, como se comporta a representação matricial de uma função linear h, perante mudança(s) de base nos dois espaços vectoriais envolvidos:

Proposição 3.11 – Mudança de bases – *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$, com dimensões finitas n e m respectivamente, e $h\colon E \to F$ uma função linear. Considerem-se, ainda, duas bases (e, e') de E e duas bases (f, f') de F. Se as bases se relacionam entre si por*

$$\begin{cases} [\,\vec{e}\,'_1 \quad \vec{e}\,'_2 \quad \cdots \quad \vec{e}\,'_n\,] = [\,\vec{e}_1 \quad \vec{e}_2 \quad \cdots \quad \vec{e}_n\,]\, T_{ee'} \\ [\,\vec{f}\,'_1 \quad \vec{f}\,'_2 \quad \cdots \quad \vec{f}\,'_m\,] = [\,\vec{f}_1 \quad \vec{f}_2 \quad \cdots \quad \vec{f}_m\,]\, T_{ff'} \end{cases} \tag{3.31}$$

então, as representações matriciais de h nos pares de bases (e, f) e (e', f') relacionam-se por

$$M_{f'e'}(h) = T_{f'f} M_{fe}(h)\, T_{ee'} = T_{ff'}^{-1} M_{fe}(h)\, T_{ee'} \tag{3.32}$$

Demonstração:

Atentemos no seguinte esquema de composição de funções

$$\begin{tikzcd}[column sep=huge, row sep=huge]
E \ (e) \arrow[r, "M_{fe}(h)", "h"'] & F \ (f) \arrow[d, "M_{f'f}(\mathscr{I}_F)", "\mathscr{I}_F"'] \\
E \ (e') \arrow[u, "M_{ee'}(\mathscr{I}_E)", "\mathscr{I}_E"'] \arrow[r, "h", "M_{f'e'}(h)"'] & F \ (f')
\end{tikzcd}$$

Fig. 3.13 – Matrizes da mesma função linear em diferentes pares de bases.

Temos, atendendo à proposição 3.5.*iv*,

$$h = \mathscr{I}_F \circ h \circ \mathscr{I}_E$$

A equação (3.29) implica imediatamente

$$M_{f'e'}(h) = M_{f'f}(\mathscr{I}_F) M_{fe}(h) M_{ee'}(\mathscr{I}_E)$$

Atendendo a (3.23) tem-se $M_{ee'}(\mathscr{I}_E) = T_{ee'}$ e $M_{f'f}(\mathscr{I}_F) = T_{f'f}$, pelo que a relação entre as representações de h nos dois pares bases é

$$M_{f'e'}(h) = T_{f'f} M_{fe}(h)\, T_{ee'}$$

Mas, (2.78) permite ainda escrever

$$M_{f'e'}(h) = T_{ff'}^{-1} M_{fe}(h)\, T_{ee'} \qquad \square$$

A relação (3.32) mostra-nos que as representações matriciais da mesma função linear h são matrizes *equivalentes* (ver definição 2.18). Reciprocamente, se $A, B \in \mathbb{K}^{m,n}$ são matrizes equivalentes, existem matrizes regulares $P \in \mathbb{K}^{m,m}$ e $Q \in \mathbb{K}^{n,n}$ tais que $B = PAQ$. Seja $h\colon E \to F$ a função linear que tem, em relação a um par de bases (e, f), a representação matricial A; em relação ao par de bases (e', f') de E e F tais que

$$\begin{aligned}
[\,\vec{e}_1' \quad \vec{e}_2' \quad \cdots \quad \vec{e}_n'\,] &= [\,\vec{e}_1 \quad \vec{e}_2 \quad \cdots \quad \vec{e}_n\,]\, Q \\
[\,\vec{f}_1' \quad \vec{f}_2' \quad \cdots \quad \vec{f}_m'\,] &= [\,\vec{f}_1 \quad \vec{f}_2 \quad \cdots \quad \vec{f}_m\,]\, P^{-1}
\end{aligned}$$

a representação de h será, como se viu, $B = PAQ$ e, portanto, A e B são representações da mesma função linear h, em bases diferentes: podemos, pois, enunciar o

Corolário 3.11.1 *Duas matrizes $A, B \in \mathbb{K}^{m,n}$ são equivalentes sse elas representam a mesma função linear de um espaço E de dimensão n para um espaço F de dimensão m (sobre $\mathbb{K}$).*

Como se viu no capítulo 2, a relação anterior é uma relação de equivalência que divide $\mathbb{K}^{m,n}$ em classes de equivalência de matrizes equivalentes; cada função linear $h\colon E \to F$ corresponde a uma destas classes e qualquer propriedade matricial p comum a todas as matrizes equivalentes (por exemplo, a sua característica) pode ser considerada uma propriedade da própria função linear h, pondo, por definição

$$p(h) = p(M_{fe}(h)),$$

onde $p(M_{fe}(h))$ não depende da escolha das bases (e, f). A este propósito, recorde-se a proposição 3.8.*iii* respeitante à característica de h.

Consideremos, agora, o caso especial dos endomorfismos de E em que, tomando as mesmas bases $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ e $e' = \left(\vec{e}_1', \vec{e}_2', \ldots, \vec{e}_n'\right)$ no domínio e no codomínio, as igualdades (3.32) se transformam em

$$M_{e'}(h) = T_{e'e} M_e(h)\, T_{ee'} = T_{ee'}^{-1} M_e(h)\, T_{ee'} \tag{3.33}$$

A igualdade (3.33) sugere a definição de uma nova relação (apenas entre matrizes quadradas) do seguinte modo:

Definição 3.4 – Matrizes semelhantes – *Dizemos que duas matrizes* $A, B \in \mathbb{K}^{n,n}$ *são* ***semelhantes*** *sse existe uma matriz regular* $T \in \mathbb{K}^{n,n}$ *tal que*

$$B = T^{-1} A T \tag{3.34}$$

Observações:

- A relação de semelhança definida em $\mathbb{K}^{n,n}$ é uma relação de *equivalência* (tal como a equivalência de matrizes): a *reflexividade* resulta de ser $B = I_n^{-1} A I_n$; a *simetria* é consequência da implicação

$$B = T^{-1} A T \Rightarrow A = T B T^{-1}$$

e a *transitividade* resulta da implicação

$$B = T^{-1} A T \wedge C = P^{-1} B P \Rightarrow C = P^{-1}\big(T^{-1} A T\big) P = \big(P^{-1} T^{-1}\big) A \big(TP\big) = (TP)^{-1} A (TP)$$

- Pela definição anterior, poderemos dizer que as representações matriciais $M_{e'}(h)$ e $M_e(h)$ do mesmo endomorfismo de um espaço de dimensão finita são matrizes *semelhantes*.

- Mas a recíproca também é verdadeira: se $A, B \in \mathbb{K}^{n,n}$ são semelhantes, existe uma matriz regular $T \in \mathbb{K}^{n,n}$ tal que $B = T^{-1} A T$. Se for $h\colon E \to E$ o endomorfismo de E representado por A numa certa base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de E, esse mesmo endomorfismo será, como vimos, representado por $B = T^{-1} A T$ na base $e' = \left(\vec{e}_1', \vec{e}_2', \ldots, \vec{e}_n'\right)$ tal que

$$[\, \vec{e}_1' \;\; \vec{e}_2' \;\; \cdots \;\; \vec{e}_n' \,] = [\, \vec{e}_1 \;\; \vec{e}_2 \;\; \cdots \;\; \vec{e}_n \,]\, T$$

- Observemos que as relações de *equivalência* e de *semelhança* de matrizes são diferentes entre si: em primeiro lugar, a *equivalência* define-se entre matrizes possivelmente rectangulares enquanto a *semelhança* diz respeito, exclusivamente, a matrizes quadradas; em segundo lugar, para as matrizes quadradas (caso em que ambas as relações estão definidas), a semelhança das matrizes implica obviamente a sua equivalência, mas a recíproca não é verdadeira: por exemplo, qualquer matriz regular $A \in \mathbb{K}^{n,n}$ é *equivalente* à matriz identidade I_n (isto é, A representa um automorfismo de um espaço de dimensão n), mas só a matriz identidade é *semelhante* à matriz identidade.

Do acima exposto, podemos concluir o seguinte

Corolário 3.11.2 *Duas matrizes $A, B \in \mathbb{K}^{n,n}$ são semelhantes sse elas representam o mesmo endomorfismo de um espaço E de dimensão n sobre $\mathbb{K}$.*

Observações:

- Como acabámos de ver, a relação de semelhança definida em $\mathbb{K}^{n,n}$ é uma relação de equivalência, ficando, pois, $\mathbb{K}^{n,n}$ dividido em classes de equivalência de matrizes semelhantes. A cada endomorfismo h de um espaço E de dimensão n sobre $\mathbb{K}$ corresponde uma destas classes e qualquer propriedade matricial p que seja comum a todas as matrizes semelhantes (isto é, invariante por uma transformação (3.33)) poderá ser atribuída à própria função h, pondo

$$p(h) = p(M_e(h)) \tag{3.35}$$

e onde $p(M_e(h))$ não depende da escolha da base e. Isto acontece, entre outras propriedades, com a *característica* de um endomorfismo h, que é, portanto a característica de qualquer matriz que o represente (ver proposição 3.8.*iii*).

- No capítulo 5, veremos outras propriedades matriciais invariantes pela transformação $A \mapsto T^{-1}AT$, como por exemplo, o *polinómio característico* de A, os *valores próprios* de A, o *determinante* de A, o *traço* de A, o *traço da adjunta* de A e, mais geralmente, a *soma dos menores principais* de uma dada ordem de A; todos estes invariantes podem, segundo (3.35), ser atribuídos a qualquer endomorfismo que seja representado matricialmente por A.

Exemplo 3.32 Para ilustrar o uso das expressões (3.33), podemos considerar a função $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ do exemplo 3.29 e calcular a matriz $M_{e'}(h)$ de h na base $e' = ((2,-1),(1,2))$ a partir da matriz $M_e(h)$ de h na base $e = ((1,1),(1,-1))$ aí considerada. Comecemos por calcular $T_{ee'}$, para o que deveremos exprimir os vectores $(2,-1)$ e $(1,2)$ como combinação linear dos vectores $(1,1)$ e $(1,-1)$. Para resolver os dois sistemas de equações obtidos, bastará condensar a matriz

$$\begin{bmatrix} 1 & 1 & 2 & 1 \\ 1 & -1 & -1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & -2 & -3 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -2 & -3 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1/2 & 3/2 \\ 0 & 1 & 3/2 & -1/2 \end{bmatrix}$$

Das soluções obtidas para os sistemas anteriores, vem

$$T_{ee'} = \frac{1}{2}\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix}$$

e, invertendo a matriz anterior, tem-se

$$T_{e'e} = T_{ee'}^{-1} = \frac{1}{5}\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix}$$

De (3.33) resulta então,

$$M_{e'}(h) = T_{e'e} M_e(h)\, T_{ee'} = \frac{1}{5}\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix} \frac{1}{2}\begin{bmatrix} 3 & 3 \\ -1 & 3 \end{bmatrix} \frac{1}{2}\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 9 & -3 \\ 7 & 6 \end{bmatrix}$$

Poderíamos também ter partido da matriz $M_c(h)$ de h na base canónica c de $\mathbb{R}^2$, que obtivemos também no exemplo 3.29. Agora devemos calcular $T_{ce'}$, para o que é necessário exprimir os vectores de e' na base canónica, o que é imediato:

$$T_{ce'} = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$$

Invertendo a matriz anterior, virá

$$T_{e'c} = T_{ce'}^{-1} = \frac{1}{5}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$$

A expressão (3.33) dá, por fim,

$$M_{e'}(h) = T_{e'c}M_c(h)\,T_{ce'} = \frac{1}{5}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 9 & -3 \\ 7 & 6 \end{bmatrix}$$

o que confirma o resultado anteriormente obtido para $M_{e'}(h)$.

Como vimos, a matriz de uma função linear depende das bases que se usem para a determinar. É, pois, natural que se procure escolher bases que conduzam a uma matriz o mais *simples* possível: a este respeito, a proposição seguinte mostra que, para toda a função linear h entre espaços de dimensão finita, existe um par de bases em relação às quais a matriz $[h_{ij}]$ da função é toda constituída por zeros, excepto os r elementos $(h_{ii})_{1\leq i\leq r}$ que valem 1 e onde r é a *característica* de h.

Proposição 3.12 *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$, com dimensões n e m respectivamente, e $h\colon E \to F$ uma função linear com característica r. Então, existe um par de bases (e, f) em relação às quais a matriz de h assume a forma:*

$$M_{fe}(h) = \begin{bmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{bmatrix} = J_r \tag{3.36}$$

Demonstração:

Sejam s e r, respectivamente, a nulidade a característica r. Pelo teorema fundamental (proposição 3.6.*iii)*, temos $r = n - s$.

Tomemos agora uma base $e' = (\vec{e}_{r+1}, \vec{e}_{r+2}, \ldots, \vec{e}_{r+s} = \vec{e}_n)$ de Ker(h) e prologuemo-la a uma base e de E mediante a junção de $r = n - s$ vectores $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_r)$ de E

$$e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_r, \vec{e}_{r+1}, \vec{e}_{r+2}, \ldots, \vec{e}_n)$$

Por construção, temos $h(\vec{e}_j) = \vec{o}$, para $r < j \leq n$. Quaisquer que sejam os vectores $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_r)$ que usámos para estender a base e' do núcleo, as suas imagens $f' = \big(h(\vec{e}_1), h(\vec{e}_2), \ldots, h(\vec{e}_r)\big)$ são linearmente independentes, como se mostra a seguir

$$\sum_{i=1}^{r} k_i h(\vec{e}_i) = \vec{o} \Rightarrow h\Big(\sum_{i=1}^{r} k_i\vec{e}_i\Big) = \vec{o} \Rightarrow \sum_{i=1}^{r} k_i\vec{e}_i \in \text{Ker}(h) \Rightarrow$$
$$\sum_{i=1}^{r} k_i\vec{e}_i = \sum_{i=r+1}^{n} k_i\vec{e}_i \Rightarrow \sum_{i=1}^{r} k_i\vec{e}_i - \sum_{i=r+1}^{n} k_i\vec{e}_i = \vec{o}$$

Obtivemos, assim, uma combinação linear nula dos vectores da base e de E, pelo que os k_i terão de ser todos nulos. Isto mostra que f' é uma base da imagem (contradomínio) de h visto tratar-se de vectores linearmente independentes em número de $r = \dim\bigl(\mathrm{Img}(h)\bigr)$.

Prolonguemos agora f' a uma base f de F, mediante a junção de $m-r$ vectores $\bigl(\vec{f}_{r+1}, \vec{f}_{r+2}, \dots, \vec{f}_m\bigr)$ de F:

$$f = \Bigl(\vec{f}_1 = h(\vec{e}_1), \vec{f}_2 = h(\vec{e}_2), \dots, \vec{f}_r = h(\vec{e}_r), \vec{f}_{r+1}, \vec{f}_{r+2}, \dots, \vec{f}_m\Bigr)$$

Por construção das bases e e f, temos

$$\begin{cases} h(\vec{e}_j) = \vec{f}_j & ; 1 \leq j \leq r \\ h(\vec{e}_j) = \vec{o} & ; r < j \leq n \end{cases}$$

Matricialmente, a relação entre os $h(\vec{e}_j)$ e os $\bigl(\vec{f}_i\bigr)$ fica, portanto:

$$\begin{aligned} &[\, h(\vec{e}_1)\ h(\vec{e}_2)\ \cdots\ h(\vec{e}_r)\ h(\vec{e}_{r+1})\ \cdots\ h(\vec{e}_n)\,] = \\ &= [\, \vec{f}_1\ \vec{f}_2\ \cdots\ \vec{f}_r\ \vec{f}_{r+1}\ \cdots\ \vec{f}_m\,] \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}}_{M_{fe}(h)} \\ &= [\, \vec{f}_1\ \vec{f}_2\ \cdots\ \vec{f}_r\ \vec{f}_{r+1}\ \cdots\ \vec{f}_m\,] \begin{bmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{bmatrix} \end{aligned}$$

o que termina a demonstração. De passagem, observe que $\mathrm{c}(M_{fe}(h)) = r$ e que este resultado está de acordo com a proposição 3.8.*iii*. □

Observações:

- A representação matricial J_r anterior é chamada forma *normal* ou *canónica* de h (ou de qualquer matriz A que represente h).

- Se uma matriz A tem característica r, ela é equivalente a uma matriz da forma (3.36); quer isto dizer que será possível levá-la à referida forma por operações elementares sobre linhas e colunas.

Podemos agora provar que se duas matrizes $A, B \in \mathbb{K}^{m,n}$ têm a mesma característica, elas são necessariamente equivalentes:

Corolário 3.12.1 *Sejam $A, B \in \mathbb{K}^{m,n}$ duas matrizes de tipo $m \times n$ sobre um corpo $\mathbb{K}$. As matrizes A e B são equivalentes sse* $\mathrm{c}(A) = \mathrm{c}(B)$.

Demonstração:

Como se viu na secção 2.13 sobre matrizes elementares, se A e B são equivalentes, as suas características são iguais.

Reciprocamente, seja $r = \mathrm{c}(A) = \mathrm{c}(B)$ e $h_A, h_B \colon \mathbb{K}^n \to \mathbb{K}^m$ funções lineares representadas em certo par de bases (e, f) por A e por B, respectivamente. As características de h_A e de h_B são iguais a r e, consequentemente, as formas canónicas de h_A e h_B são ambas iguais a (3.36); daqui conclui-se que A e B são ambas equivalentes a J_r e a transitividade da relação de equivalência mostra que A e B serão equivalentes entre si. $\square$

3.8 Anexos: funções lineares e o MATHEMATICA©

MATHEMATICA©

O MATHEMATICA© pode também ser um bom auxiliar no que respeita às funções lineares. Dispõe de funções capazes de resolver boa parte dos problemas por nós analisados neste capítulo; de entre as referidas funções salientamos:

- **`NullSpace[A]`**

Devolve uma base do *núcleo* do operador representado por **`A`**.

A expressão **`Length[NullSpace[A]]`** devolve a *nulidade* do operador representado por **`A`**.

- **`MatrixRank[A]`**

  `Caracteristica[A][[1]]` (do package **`` ALGA`Matrizes` ``**)

Devolve a *característica* do operador representado por **`A`**.

O Package **`` ALGA`FuncoesLineares` ``** implementa algumas funções adicionais que resolvem alguns problemas sobre funções lineares e consta das seguintes funções:

- **`MatrizRepNova[f1,e1,f2,e2,A1]`**

Devolve a representação em relação ao par de bases (**`f2,e2`**) da função linear que é representada por **`A1`**, em relação ao par de bases (**`f1,e1`**). Nestes pares, a primeira base é do espaço de chegada e a segunda é do espaço de partida.

- **`MatrizRep[f2,e2,A1]`**

Determina a representação matricial nas bases (**`f2,e2`**) do operador que é representado por **`A1`** em relação às bases canónicas. No par de bases indicadas, a primeira base é do espaço de chegada e a segunda é do espaço de partida.

- **`Imagem[A]`**

Devolve uma base da *imagem* do operador linear representado por **`A`**.

A expressão **`Length[Imagem[A]]`** devolve também a *característica* do operador representado por **`A`**.

- **`InjectivaQ[A]`**

Devolve **`True`** se e só se a função representada por **`A`** é *injectiva*.

- **`SobrejectivaQ[A]`**

Devolve **`True`** se e só se a função representada por **`A`** é *sobrejectiva*.

Apresentamos a seguir a implementação do *package* **`` ALGA`FuncoesLineares` ``**:

```
(* Package by Carlos Ribeiro, Março 2009 *)
(* Contexto : ALGA`FuncoesLineares` *)
(* Versão : 3.7 *)
(* Versão do Mathematica : 7.0 *)

BeginPackage["ALGA`FuncoesLineares`", {"ALGA`Matrizes`"}]

(* ------------------------------ HELP ON-LINE ------------------------------ *)

MatrizRepNova::usage = "MatrizRepNova[f1,e1,f2,e2,a1] devolve a representação em \
relação ao par de bases f2,e2 da função linear que é representada por a1,em relação \
ao par de bases f1,e1.Nestes pares,a primeira base é do espaço de chegada e a segunda \
é do espaço de partida."

MatrizRep::usage = "MatrizRep[f2,e2,a1] determina a representação matricial nas \
bases f2,e2 do operador que é representado por a1 em relação às bases canónicas. \
No par de bases indicadas, a primeira base é do espaço de chegada e a segunda \
é do espaço de partida."

Imagem::usage = "Imagem[a] devolve uma base da imagem (contradomínio) do operador \
representado por a."

InjectivaQ::usage = "InjectivaQ[a] devolve True se o operador representado por a é injectivo; \
False, no caso contrário."

SobrejectivaQ::usage =  "SobrejectivaQ[a] devolve True se o operador representado \
por a é sobrejectivo; False, no caso contrário."

Begin["`Private`"]

(* ---------------------------- MENSAGENS DE ERRO ----------------------------- *)

MatrizRepNova::naobase = "Uma das listas indicadas não é uma base";
MatrizRepNova::errdim = "As dimensões da matriz não são compatíveis com as bases indicadas";

(* ------------------------------ IMPLEMENTAÇÃO ------------------------------ *)

(* Implementação do módulo MatrizRepNova *)
MatrizRepNova[f1_?MatrixQ, e1_?MatrixQ, f2_?MatrixQ, e2_?MatrixQ, a1_?MatrixQ] :=
  Module[{t1 = Passagem[e1, e2], t2 = Passagem[f2, f1]},
    Which[
      Not[SquareMatrixQ[t1] && SquareMatrixQ[t2]],
      Message[MatrizRepNova::naobase],
      MatrixRank[f1] != Length[f1], Message[MatrizRepNova::naobase],
      MatrixRank[e2] != Length[e2], Message[MatrizRepNova::naobase],
      Length[t2] != Length[a1] || Length[t1] != Dimensions[a1][[2]],
      Message[MatrizRepNova::errdim],
      True, t2.a1.t1
    ]
  ]

(* Implementação do módulo MatrizRep *)
MatrizRep[f2_?MatrixQ, e2_?MatrixQ, a1_?MatrixQ] :=
  MatrizRepNova[IdentityMatrix[Length[f2]],IdentityMatrix[Length[e2]],f2,e2,a1]

(* Implementação do módulo Imagem *)
Imagem[a_?MatrixQ] :=
  Module[{m = Caracteristica[Transpose[a]]},
    Take[m[[2]],m[[1]]]
  ]

(* Implementação do módulo InjectivaQ *)
```

```
InjectivaQ[a_?MatrixQ] := MatrixRank[a]==Dimensions[a][[2]]

(* Implementação do módulo SobrejectivaQ *)
SobrejectivaQ[a_?MatrixQ] := MatrixRank[a]==Length[a]

End[]
EndPackage[]
```

O *package* **ALGA`FuncoesLineares`** deverá ser previamente carregado por meio de um dos comandos

```
<<ALGA`FuncoesLineares`
Get["ALGA`FuncoesLineares`"]
Needs["ALGA`FuncoesLineares`"]
DeclarePackage["ALGA`FuncoesLineares`"]
```

Nas páginas seguintes apresenta-se um *notebook* que ilustra o uso do MATHEMATICA© e destas funções para resolver alguns problemas sobre funções lineares.

Funções lineares

3.8.1. Núcleo, Imagem, nulidade e característica; injectividade e sobrejectividade

- **Com o MATHEMATICA, podemos estudar as funções lineares**

```
<< ALGA`FuncoesLineares`
```

Se a representação matricial (em certo par de bases) for

```
(a = {{2, 2, 1}, {3, -2, 1}, {1, -4, 0}, {1, 6, 1}}) // MatrixForm
```

$$\begin{pmatrix} 2 & 2 & 1 \\ 3 & -2 & 1 \\ 1 & -4 & 0 \\ 1 & 6 & 1 \end{pmatrix}$$

Uma base do núcleo será

```
ker = NullSpace[a]
```

```
{{-4, -1, 10}}
```

E a nulidade é

```
s = Length[ker]
```

```
1
```

Uma base da imagem (contradomínio) será

```
? Imagem
```

Imagem[a] devolve uma base da imagem (contradomínio) do operador representado por a.

```
img = Imagem[a]
```

```
{{2, 3, 1, 1}, {0, -10, -10, 10}}
```

A característica é

```
r = Length[img]
```

```
2
```

Ou ainda

```
r = MatrixRank[a]
```

```
2
```

O teorema fundamental sobre as dimensões

```
r + s == Dimensions[a][[2]]
```

```
True
```

A função é injectiva?

```
?InjectivaQ
```

InjectivaQ[a] devolve True se o operador representado por a é injectivo; False, no caso contrário.

```
InjectivaQ[a]
```

```
False
```

E sobrejectiva?

```
?SobrejectivaQ
```

SobrejectivaQ[a] devolve True se o operador representado por a é Sobrejectivo; False, no caso contrário.

```
SobrejectivaQ[a]
```

```
False
```

3.8.2. Mudança de bases e a representação matricial

- **As funções `MatrizRep[]` e `MatrizRepNova[]` permitem manipular a representação matricial, quando se mudam as bases**
  Exemplo 2.1

```
?MatrizRep
```

MatrizRep[f2,e2,a1] determina a representação matricial nas bases f2,e2 do
operador que é representado por a1 em relação às bases canónicas. No par de bases
indicadas, a primeira base é do espaço de chegada e a segunda é do espaço de partida.

```
?MatrizRepNova
```

MatrizRepNova[f1,e1,f2,e2,a1] devolve a representação em relação ao par de bases
f2,e2 da função linear que é representada por a1,em relação ao par de bases f1,e1.Nestes
pares,a primeira base é do espaço de chegada e a segunda é do espaço de partida.

Definam-se um par de bases em $\mathbb{R}^3$e $\mathbb{R}^4$ por

```
f2 = {{1, -1, 2}, {1, 1, 0}, {2, 1, -1}};
```

```
e2 = {{-1, 2, 1, 1}, {0, 2, 1, -1}, {2, 3, -2, 1}, {1, 1, 1, 1}};
```

Defina-se um segundo par de bases de $\mathbb{R}^3$e de $\mathbb{R}^4$ por

```
f3 = {{0, 1, -2}, {2, -1, 1}, {3, -1, 0}};
```

```
e3 = {{2, 1, 2, -2}, {1, 1, -1, 2}, {-1, 0, 1, 3}, {2, -1, 0, -3}};
```

A função `f`:$\mathbb{R}^4 \to \mathbb{R}^3$;`(x,y,z,w)`$\mapsto$`(2x+3y-z+2w,x-2y+4z-w,x+5y-5z+3w)` é representada em relação às bases canónicas `f1,e1` de $\mathbb{R}^3$e de $\mathbb{R}^4$ pela matriz 3×4 dos coeficientes da última expressão

```
(a1 = {{2, 3, -1, 2}, {1, -2, 4, -1}, {1, 5, -5, 3}}) // MatrixForm
```

$$\begin{pmatrix} 2 & 3 & -1 & 2 \\ 1 & -2 & 4 & -1 \\ 1 & 5 & -5 & 3 \end{pmatrix}$$

A função `f` é representada em relação às bases `f2` e `e2` de $\mathbb{R}^3$e de $\mathbb{R}^4$ por

```
(a2 = MatrizRep[f2, e2, a1]) // MatrixForm
```

$$\begin{pmatrix} \frac{7}{2} & 1 & 15 & 2 \\ \frac{3}{2} & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

A mesma função `f` é representada em relação às bases `f3` e `e3` de $\mathbb{R}^3$e de $\mathbb{R}^4$ pela matriz obtida por

```
(a3 = MatrizRepNova[f2, e2, f3, e3, a2]) // MatrixForm
```

$$\begin{pmatrix} 22 & 6 & 6 & 4 \\ 35 & 29 & 15 & -4 \\ -23 & -16 & -9 & 1 \end{pmatrix}$$

Ou ainda por

```
(a3 = MatrizRepNova[IdentityMatrix[Length[f3]],
    IdentityMatrix[Length[e3]], f3, e3, a1]) // MatrixForm
```

$$\begin{pmatrix} 22 & 6 & 6 & 4 \\ 35 & 29 & 15 & -4 \\ -23 & -16 & -9 & 1 \end{pmatrix}$$

Uma base do núcleo de `f` será

```
ker = NullSpace[a1]
```

```
{{-1, -4, 0, 7}, {-10, 9, 7, 0}}
```

E a nulidade de `f` é

```
s = Length[ker]
```

```
2
```

Uma base da imagem (contradomínio) de `f` será

```
img = Imagem[a1]
```

```
{{2, 1, 1}, {0, -7, 7}}
```

A característica de `f` é

```
r = Length[img]
```

```
2
```

Ou ainda

```
r = MatrixRank[a1]
```

```
2
```

O teorema fundamental sobre as dimensões

```
r + s == Dimensions[a1][[2]]
```

```
True
```

A função `f` é injectiva?

```
InjectivaQ[a1]
```

```
False
```

E sobrejectiva?

```
SobrejectivaQ[a1]
```

```
False
```

■ Exemplo 2.2

A função $f:\mathbb{R}^4 \to \mathbb{R}^3; (x,y,z,w) \mapsto (2x+3y-z+2w, x-2y+4z-w, 2x-3y+2z)$ é representada em relação às bases canónicas `f1,e1` de $\mathbb{R}^3$ e de $\mathbb{R}^4$ pela matriz 3×4 dos coeficientes da última expressão

```
(a1 = {{2, 3, -1, 2}, {1, -2, 4, -1}, {2, -3, 2, 0}}) // MatrixForm
```

$$\begin{pmatrix} 2 & 3 & -1 & 2 \\ 1 & -2 & 4 & -1 \\ 2 & -3 & 2 & 0 \end{pmatrix}$$

A função `f` é representada em relação às bases `f2` e `e2` de $\mathbb{R}^3$ e de $\mathbb{R}^4$ por

```
(a2 = MatrizRep[f2, e2, a1]) // MatrixForm
```

$$\begin{pmatrix} \frac{1}{4} & -\frac{1}{2} & \frac{21}{4} & \frac{5}{4} \\ -\frac{33}{4} & -\frac{5}{2} & -\frac{109}{4} & \frac{7}{4} \\ \frac{13}{2} & 3 & \frac{39}{2} & \frac{3}{2} \end{pmatrix}$$

A mesma função `f` é representada em relação às bases
`f3` e `e3` de $\mathbb{R}^3$ e de $\mathbb{R}^4$ pela matriz obtida por

```
(a3 = MatrizRepNova[f2, e2, f3, e3, a2]) // MatrixForm
```

$$\begin{pmatrix} 36 & -14 & 3 & 23 \\ 77 & -31 & 6 & 53 \\ -51 & 24 & -3 & -37 \end{pmatrix}$$

Ou ainda por

```
(a3 = MatrizRepNova[IdentityMatrix[Length[f3]],
   IdentityMatrix[Length[e3]], f3, e3, a1]) // MatrixForm
```

$$\begin{pmatrix} 36 & -14 & 3 & 23 \\ 77 & -31 & 6 & 53 \\ -51 & 24 & -3 & -37 \end{pmatrix}$$

Uma base do núcleo de `f` será

```
ker = NullSpace[a1]
```

```
{{-19, -6, 10, 33}}
```

E a nulidade de `f` é

```
s = Length[ker]
```

```
1
```

Uma base da imagem (contradomínio) de `f` será

```
img = Imagem[a1]
```

```
{{2, 1, 2}, {0, -7, -12}, {0, 0, 66}}
```

A característica de `f` é

```
r = Length[img]
```

```
3
```

Ou ainda

```
r = MatrixRank[a1]
```

```
3
```

O teorema fundamental sobre as dimensões

```
r + s == Dimensions[a1][[2]]
```

```
True
```

A função `f` é injectiva?

```
InjectivaQ[a1]
```

```
False
```

E sobrejectiva?

```
SobrejectivaQ[a1]
```

```
True
```

Exemplo 2.3

Definam-se um par de bases em $\mathbb{C}^3$e $\mathbb{C}^2$ por

```
f2 = {{1 + I, 2 I, 2 - I}, {1, I, 1 + I}, {2 - 5 I, 2 + I, 3 I}};
```

```
e2 = {{2 I, 1 - I}, {3 - 2 I, -2 + I}};
```

Defina-se um segundo par de bases de $\mathbb{C}^3$e de $\mathbb{C}^2$ por

```
f3 = {{2 - 2 I, 1 - I, -1 + I}, {2 - I, 0, 3 - I}, {I, 2 - I, 2 + I}};
```

```
e3 = {{2 - I, 3 + I}, {1 + 3 I, -1 + 2 I}};
```

A função $f:\mathbb{C}^2 \to \mathbb{C}^3; (z,w) \mapsto (2z+(1-i)w, (2-3i)z-iw, (2+i)z+(1+2i)w)$ é representada em relação às bases canónicas f1, e1 de $\mathbb{C}^3$ e de $\mathbb{C}^2$ pela matriz 3×2 dos coeficientes da última expressão

```
(a1 = {{2, 1 - I}, {2 - 3 I, -I}, {2 + I, 1 + 2 I}}) // MatrixForm
```

$$\begin{pmatrix} 2 & 1-i \\ 2-3i & -i \\ 2+i & 1+2i \end{pmatrix}$$

A função f é representada em relação às bases f2 e e2 de $\mathbb{C}^3$ e de $\mathbb{C}^2$ por

```
(a2 = MatrizRep[f2, e2, a1]) // MatrixForm
```

$$\begin{pmatrix} -\frac{1}{4}-\frac{31 i}{4} & -\frac{27}{2}+\frac{11 i}{2} \\ \frac{223}{20}+\frac{239 i}{20} & \frac{123}{10}-\frac{261 i}{10} \\ -\frac{19}{20}-\frac{67 i}{20} & -\frac{49}{10}+\frac{43 i}{10} \end{pmatrix}$$

A mesma função f é representada em relação às bases f3 e e3 de $\mathbb{C}^3$ e de $\mathbb{C}^2$ pela matriz obtida por

```
(a3 = MatrizRepNova[f2, e2, f3, e3, a2]) // MatrixForm
```

$$\begin{pmatrix} \frac{107}{34}-\frac{3 i}{2} & \frac{19}{17}+\frac{57 i}{17} \\ \frac{2}{17}+2 i & -\frac{29}{17}-\frac{2 i}{17} \\ \frac{24}{17}-\frac{42 i}{17} & \frac{52}{17}+\frac{41 i}{17} \end{pmatrix}$$

Ou ainda por

```
(a3 = MatrizRepNova[IdentityMatrix[Length[f3]],
    IdentityMatrix[Length[e3]], f3, e3, a1]) // MatrixForm
```

$$\begin{pmatrix} \frac{107}{34}-\frac{3 i}{2} & \frac{19}{17}+\frac{57 i}{17} \\ \frac{2}{17}+2 i & -\frac{29}{17}-\frac{2 i}{17} \\ \frac{24}{17}-\frac{42 i}{17} & \frac{52}{17}+\frac{41 i}{17} \end{pmatrix}$$

Uma base do núcleo de `f` será

```
ker = NullSpace[a1]
```

```
{}
```

E a nulidade de `f` é

```
s = Length[ker]
```

```
0
```

Uma base da imagem (contradomínio) de `f` será

```
img = Imagem[a1]
```

```
{{2, 2 - 3 i, 2 + i}, {0, 1 + 3 i, -1 + 5 i}}
```

A característica de `f` é

```
r = Length[img]
```

```
2
```

Ou ainda

```
r = MatrixRank[a1]
```

```
2
```

O teorema fundamental sobre as dimensões

```
r + s == Dimensions[a1][[2]]
```

```
True
```

A função `f` é injectiva?

```
InjectivaQ[a1]
```

```
True
```

E sobrejectiva?

```
SobrejectivaQ[a1]
```

```
False
```

4

Determinantes

4.1 Introdução

Em muitas aplicações da Álgebra Linear à Geometria e à Análise, o conceito de determinante desempenha um importante papel. Começaremos por um breve estudo do grupo simétrico, após o que abordaremos com generalidade a teoria das funções multilineares alternadas e as suas propriedades. Estas são, na essência, as propriedades dos determinantes. Definiremos depois as funções determinante numa base, determinante de uma matriz e determinante de um endomorfismo. Apresentaremos adiante três métodos gerais para o cálculo de determinantes de ordem n: o método de condensação, um método recursivo baseado no Teorema de Laplace restrito e um método misto, resultante do uso alternado de condensação e do Teorema de Laplace. Estudaremos também a relação entre a teoria dos determinantes e a característica de uma matriz, bem como a resolução de sistemas de equações lineares à custa de determinantes (regra de Cramer e Teorema de Rouché). O teorema de Laplace restrito conduzir-nos-á à definição da noção de matriz adjunta e à sua utilização na inversão de matrizes quadradas regulares. Na parte final, faremos referência à fórmula de Cauchy. No capítulo 6, aquando da abordagem do produto misto, veremos uma "interpretação geométrica" do determinante: o módulo do determinante de uma lista de vectores numa base ortonormada representa o volume do paralelepípedo definido por esses vectores.

4.2 Permutações. O grupo simétrico

Definição 4.1 – Permutação – *Seja $n \geq 1$ um inteiro. Uma* ***permutação***[1] *de ordem n é uma aplicação bijectiva $\sigma\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Designando por σ_k a imagem de k por meio de σ, cada permutação será definida pelas n imagens $\sigma_1, \sigma_2, \ldots, \sigma_n$ em que, devido à injectividade de σ, será $1 \leq \sigma_1 \neq \sigma_2 \neq \ldots \neq \sigma_n \leq n$. Utilizaremos a notação*

$$\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n) \tag{4.1}$$

para designar a permutação σ.

A função identidade

$$\mathscr{I}_n\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}; k \mapsto k$$

é uma permutação dita permutação identidade de ordem n e, usando a notação introduzida acima, poderemos escrever

$$\mathscr{I}_n = (1, 2, \ldots, n) \tag{4.2}$$

O número de permutações de ordem n é de $n(n-1)\cdots 2 \cdot 1 = n!$ e designaremos por $\mathfrak{S}_n$ o conjunto de todas as permutações de ordem n:

$$\begin{aligned} &\mathfrak{S}_n = \{\sigma\colon \sigma \text{ é uma permutação de ordem } n\} \\ &\#\mathfrak{S}_n = n! \end{aligned} \tag{4.3}$$

---

1 Ver o exemplo B.35 no apêndice B.

Definição 4.2 – Transposição – *Seja $n \geq 2$ um inteiro. Chama-se* ***transposição*** *a uma permutação $\tau \in \mathfrak{S}_n$, para a qual existem inteiros $i, j \in \{1, 2, \ldots, n\}$ com $i < j$, tais que*

$$\begin{cases} \tau_i = j \\ \tau_j = i \\ \tau_k = k; \quad k \neq i, j \end{cases} \tag{4.4}$$

portanto, será $\tau = (1, 2, \ldots, i-1, j, i+1, \ldots j-1, i, j+1, \ldots, n)$.

Sendo $n \geq 2$, o número de transposições de ordem n é de $\binom{n}{2} = \frac{1}{2}n(n-1)$ (tantas quantos os subconjuntos $\{i, j\}$ de $\{1, 2, \ldots, n\}$). Para $n = 1$, não existem transposições. De entre as transposições, podemos ainda salientar as chamadas *transposições elementares*, definidas como segue:

Definição 4.3 – Transposição elementar – *Seja $n \geq 2$ um inteiro. Chama-se* ***transposição elementar*** *a uma permutação $\tau \in \mathfrak{S}_n$, para a qual existe um inteiro $1 \leq i < n$, tal que*

$$\begin{cases} \tau_i = i+1 \\ \tau_{i+1} = i \\ \tau_k = k; \qquad k \neq i, i+1 \end{cases} \tag{4.5}$$

O número de transposições elementares é de $n - 1$, visto ser este o número de formas distintas de seleccionar o inteiro $1 \leq i < n$.

Exemplo 4.1 Existe apenas uma permutação de ordem 1, que é a identidade

$$\mathfrak{S}_1 = \{(1)\}$$

Não existem transposições, quando $n = 1$.

Exemplo 4.2 Consideremos $n = 2$. Existem $2! = 2$ permutações de 2ª ordem e, entre estas, $\binom{2}{2} = 1$ transposição (elementar). Tem-se

$$\mathfrak{S}_2 = \{(1, 2), (2, 1)\}$$

Neste caso, existe apenas a permutação identidade $\mathscr{I}_2 = (1, 2)$ e uma transposição elementar $\tau = (2, 1)$.

Exemplo 4.3 Consideremos $n = 3$. Existem $3! = 6$ permutações de 3ª ordem e, de entre estas, $\binom{3}{2} = 3$ são transposições. Tem-se

$$\mathfrak{S}_3 = \{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}$$

Neste caso, além da permutação identidade $\mathscr{I}_3 = (1, 2, 3)$, existem três transposições que são $(1, 3, 2), (2, 1, 3)$ e $(3, 2, 1)$. Destas, as duas primeiras são elementares.

Exemplo 4.4 Consideremos $n = 4$. Existem $4! = 24$ permutações de 4ª ordem e, de entre estas, $\binom{4}{2} = 6$ são transposições. Tem-se (ver figura 4.1)

$$\begin{aligned}\mathfrak{S}_4 = &\{(1,2,3,4),(1,2,4,3),(1,3,2,4),(1,3,4,2),(1,4,2,3),(1,4,3,2)\} \cup \\ &\{(2,1,3,4),(2,1,4,3),(2,3,1,4),(2,3,4,1),(2,4,1,3),(2,4,3,1)\} \cup \\ &\{(3,1,2,4),(3,1,4,2),(3,2,1,4),(3,2,4,1),(3,4,1,2),(3,4,2,1)\} \cup \\ &\{(4,1,2,3),(4,1,3,2),(4,2,1,3),(4,2,3,1),(4,3,1,2),(4,3,2,1)\}\end{aligned}$$

Aqui, além da permutação identidade $\mathscr{I}_4 = (1,2,3,4)$, existem $\binom{4}{2} = 6$ transposições que são $(1,2,4,3),(1,3,2,4),(1,4,3,2),(2,1,3,4),(3,2,1,4),(4,2,3,1)$. De entre estas, as duas primeiras e a quarta são elementares.

Para obter as 24 permutações de 4ª ordem, podemos recorrer ao esquema da figura 4.1, que pode ser usado também para outros valores de n. Este processo de "desdobramento" consiste em colocar nos nós-raízes (de nível 0) das n árvores os números 1 a n; cada um destes nós tem $n-1$ descendentes de nível 1 (diferentes da raiz); estes por sua vez têm $n-2$ descendentes (diferentes dos nós dos dois níveis anteriores) e assim por diante, até que, no nível $n-2$, existe apenas um descendente (o único dos números 1 a n que não foi usado antes).

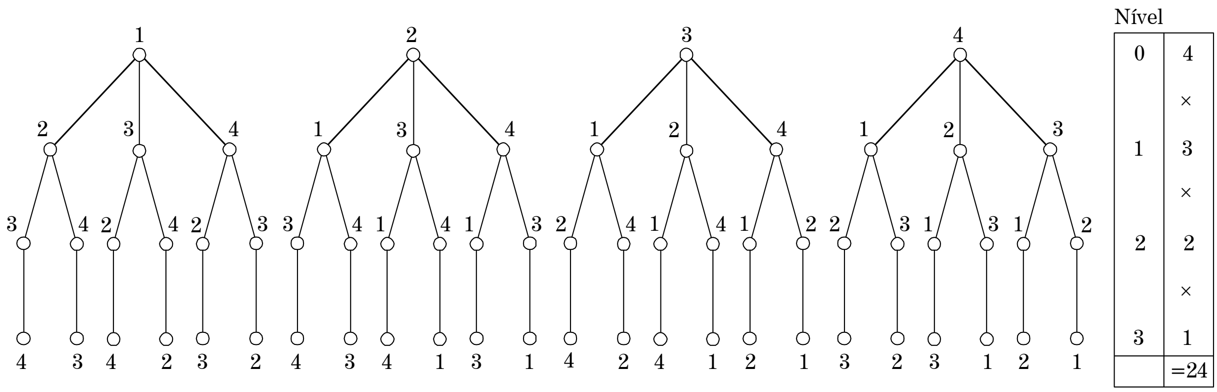


Fig. 4.1 – Esquema para obter as $4! = 24$ permutações de 4ª ordem.

Sejam σ_1 e σ_2 duas permutações de ordem n e consideremos, agora, a *composição* (também chamada *produto*) de σ_1 com σ_2:

$$\begin{tikzcd} \{1,2,\ldots,n\} \arrow[r, "\sigma_2"] \arrow[dr, "\sigma_1 o \sigma_2"'] & \{1,2,\ldots,n\} \arrow[d, "\sigma_1"] \\ & \{1,2,\ldots,n\} \end{tikzcd}$$

Fig. 4.2 – Composição (produto) de permutações.

Como a composta de duas bijecções é uma bijecção, segue-se que a composição $\circ$ é uma *lei de composição interna* em $\mathfrak{S}_n$. Portanto, $(\mathfrak{S}_n, \circ)$ constitui um grupóide e, para $n > 2$, este grupóide não é comutativo. Todavia, a composição é *associativa*, tem *elemento neutro* (a permutação identidade $\mathscr{I}_n$) e, sendo uma função bijectiva, cada permutação σ tem uma *inversa* σ^{-1} tal que $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathscr{I}_n$: trata-se, portanto, de um *grupo* (não comutativo para $n > 2$) a que chamaremos o *grupo simétrico de ordem* n.

Portanto, tem-se:

- $n > 2 \Rightarrow \underset{\sigma_1,\sigma_2 \in \mathfrak{S}_n}{\exists} \sigma_1 \circ \sigma_2 \neq \sigma_2 \circ \sigma_1$
- $\underset{\sigma_1,\sigma_2,\sigma_3 \in \mathfrak{S}_n}{\forall} (\sigma_1 \circ \sigma_2) \circ \sigma_3 = \sigma_1 \circ (\sigma_2 \circ \sigma_3)$
- $\underset{\sigma \in \mathfrak{S}_n}{\forall} \mathscr{I}_n \circ \sigma = \sigma \circ \mathscr{I}_n = \sigma$
- $\underset{\sigma \in \mathfrak{S}_n}{\forall} \underset{\sigma^{-1} \in \mathfrak{S}_n}{\exists} \sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathscr{I}_n$

Exemplo 4.5 Por exemplo, em $\mathfrak{S}_3$, calculemos $(1,3,2) \circ (3,2,1)$: tem-se, neste caso,

$$\begin{aligned} 1 &\mapsto 3 \mapsto 2 \\ 2 &\mapsto 2 \mapsto 3 \\ 3 &\mapsto 1 \mapsto 1 \end{aligned}$$

e, portanto, $(1,3,2) \circ (3,2,1) = (2,3,1)$. Do mesmo modo, para $(3,2,1) \circ (1,3,2)$, vem:

$$\begin{aligned} 1 &\mapsto 1 \mapsto 3 \\ 2 &\mapsto 3 \mapsto 1 \\ 3 &\mapsto 2 \mapsto 2 \end{aligned}$$

Portanto, será $(3,2,1) \circ (1,3,2) = (3,1,2) \neq (1,3,2) \circ (3,2,1) = (2,3,1)$. Seguem-se outros exemplos:

$$\begin{aligned} &(2,1,3) \circ (2,3,1) = (1,3,2) \neq (2,3,1) \circ (2,1,3) = (3,2,1) \\ &(3,1,2) \circ (2,3,1) = (1,2,3) = (2,3,1) \circ (3,1,2) = (1,2,3) = \mathscr{I}_3 \end{aligned}$$

Exemplo 4.6 O exemplo anterior mostra que $(2,3,1) = (3,1,2)^{-1}$. Tem-se ainda, a título de exemplo, algumas composições e inversas em $\mathfrak{S}_4$ e $\mathfrak{S}_5$ (verifique as igualdades!),

$$\begin{aligned} &(4,3,1,2) \circ (2,4,3,1) = (3,2,1,4) \\ &(3,5,2,1,4) \circ (4,2,1,5,3) = (1,5,3,4,2) \\ &(3,4,1,2)^{-1} = (3,4,1,2) \\ &(3,5,2,1,4)^{-1} = (4,3,1,5,2) \\ &(2,1,5,3,4)^{-1} = (2,1,4,5,3) \end{aligned}$$

Exemplo 4.7 Sendo τ uma transposição de ordem n, tem-se

$$\tau \circ \tau = \mathscr{I}_n \tag{4.6}$$

o que mostra que $\tau^{-1} = \tau$, para qualquer transposição τ.

Exemplo 4.8 Seja $1 \leq i < j \leq n$ e considere-se a transposição τ de i e j

$$\begin{cases} \tau_i = j \\ \tau_j = i \\ \tau_k = k, \quad k \neq i, j \end{cases}$$

e uma permutação $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_i, \ldots, \sigma_j, \ldots, \sigma_n)$. Calculemos $\sigma \circ \tau$, para o que basta atentar no esquema

$$\begin{cases} i \mapsto j \mapsto \sigma_j \\ j \mapsto i \mapsto \sigma_i \\ k \mapsto k \mapsto \sigma_k, \quad k \neq i, j \end{cases}$$

Daqui se conclui ser

$$\sigma \circ \tau = (\sigma_1, \sigma_2, \ldots, \sigma_j, \ldots, \sigma_i, \ldots, \sigma_n)$$

Portanto, *a troca de dois elementos σ_i e σ_j entre si numa permutação σ conduz a uma nova permutação que tem a forma $\sigma \circ \tau$, onde τ é uma transposição.* Em particular, *a troca de dois elementos* ***consecutivos*** *σ_i e σ_{i+1} entre si na permutação σ leva-nos a uma permutação da forma $\sigma \circ \tau$ em que τ é transposição elementar.*

Com base neste último exemplo, podemos provar a

Proposição 4.1 – Decomposição de uma permutação – *Seja $\sigma \in \mathfrak{S}_n$ uma permutação de ordem $n \geq 1$. Então, existem transposições $\tau_1, \tau_2, \ldots, \tau_p$, com $p < n$, tais que*

$$\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_p \tag{4.7}$$

Demonstração:

A demonstração consiste em transformar a permutação identidade $\mathscr{I}_n$ na permutação σ dada, por sucessivas trocas dos seus elementos: seja

$$\mathscr{I}_n = (1, 2, \ldots, n)$$

Se $\sigma_1 \neq 1$, então troquemos 1 com σ_1 (se for $\sigma_1 = 1$, passamos para o passo seguinte); após isto, obtém-se a permutação

$$\mathscr{I}_n \circ \tau_1 = (\sigma_1, 2, \ldots, n)$$

Se $\sigma_2 \neq 2$, então troquemos 2 com σ_2 (se for $\sigma_2 = 2$, saltamos para o passo seguinte); após esta operação, obtemos a permutação

$$\mathscr{I}_n \circ \tau_1 \circ \tau_2 = (\sigma_1, \sigma_2, \ldots, n)$$

Continuando com este procedimento, ao fim de um máximo de p trocas ($p \leq n-1$, sendo que $p = n-1$ só será atingido, se for sempre $\sigma_k \neq k$, $k = 1, 2, \ldots, n-1$), chega-se a

$$\mathscr{I}_n \circ \tau_1 \circ \tau_2 \circ \cdots \circ \tau_p = (\sigma_1, \sigma_2, \ldots, \sigma_n) = \sigma$$

e a proposição está demonstrada, visto que $\mathscr{I}_n \circ \tau_1 \circ \tau_2 \circ \cdots \circ \tau_p = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_p$. □

Vejamos, agora, que toda a transposição é um produto de transposições elementares:

Proposição 4.2 *Seja* $\tau \in \mathfrak{S}_n$ *uma transposição de ordem* $n \geq 2$. *Então, existem transposições elementares* $\xi_1, \xi_2, \ldots, \xi_k$ *tais que*

$$\tau = \xi_1 \circ \xi_2 \circ \cdots \circ \xi_k \tag{4.8}$$

Demonstração:

Considere-se uma transposição τ de ordem n dada por (4.4) e onde $i < j$. Se $j = i + 1$, o resultado é imediato, visto que τ é ela própria elementar. Se $i + 1 < j$, consideremos a permutação identidade

$$\mathscr{I}_n = (1, 2, \ldots, i, i + 1, \ldots, j - 1, j, \ldots, n)$$

e façamos as $r = j - i$ transposições elementares $\xi_1 \circ \xi_2 \circ \cdots \circ \xi_r$ que trocam sucessivamente os elementos nas posições consecutivas $(i, i + 1), (i + 1, i + 2), \ldots, (j - 1, j)$, o que levará o elemento i a ocupar a posição j (o elemento i avança desde a posição i até à posição j, através de r trocas entre elementos ocupando posições *consecutivas*)

$$\mathscr{I}_n \circ \xi_1 \circ \xi_2 \circ \cdots \circ \xi_r = (1, 2, \ldots, i + 1, i + 2, \ldots, j, i, \ldots, n)$$

De seguida, executemos as $s = j - i - 1$ transposições elementares $\xi_{r+1} \circ \cdots \circ \xi_{r+s}$ que trocam os elementos nas posições $(j - 1, j - 2), (j - 2, j - 3), \ldots, (i + 1, i)$, o que leva o elemento j da permutação anterior a ocupar a posição i (o elemento j recua até à posição i, exclusivamente através de trocas entre elementos ocupando posições *consecutivas*), obtendo-se finalmente, com $k = r + s = 2(j - i) - 1$ (sempre ímpar):

$$\xi_1 \circ \xi_2 \circ \cdots \circ \xi_k = (1, 2, \ldots, j, i + 1, \ldots, j - 1, i, \ldots, n) = \tau$$

o que termina a demonstração. □

Com as duas proposições anteriores, é óbvio que podemos garantir que toda a permutação de ordem n é decomponível num produto de transposições *elementares*:

Proposição 4.3 *Seja* $\sigma \in \mathfrak{S}_n$ *uma permutação de ordem* $n \geq 1$. *Então, existem transposições elementares* $\xi_1, \xi_2, \ldots, \xi_p$ *tais que*

$$\sigma = \xi_1 \circ \xi_2 \circ \cdots \circ \xi_p \tag{4.9}$$

Demonstração:

Basta substituir em (4.7) cada transposição pela sua decomposição (4.8) em transposições elementares. □

O leitor pode a título de exercício provar a última proposição directamente, sem usar (4.8). As proposições anteriores mostram que *toda a permutação é um produto de transposições* (que poderão também ser exclusivamente *elementares*) e diremos que σ é uma permutação *par* ou *ímpar*, consoante for par ou ímpar o número p de transposições nas decomposições (4.7) ou (4.9). Do mesmo modo, chamaremos *sinal* ou *paridade* de σ ao inteiro $\varepsilon(\sigma)$ definido por

$$\varepsilon(\sigma) = (-1)^p \tag{4.10}$$

Portanto, o sinal de σ é $+1$ ou -1, conforme a permutação σ (e p) for par ou ímpar. Da definição resulta imediatamente que todas as transposições (elementares ou não) são ímpares e que a permutação identidade $\mathscr{I}_n$ é par:

$$\begin{aligned} \varepsilon(\tau) &= -1 \\ \varepsilon(\mathscr{I}_n) &= +1 \end{aligned} \tag{4.11}$$

A demonstração da proposição 4.1 fornece-nos o algoritmo para determinar a paridade de uma permutação qualquer[(2)], como se mostra nos exemplos seguintes:

Exemplo 4.9 A única permutação de 1ª ordem $\mathscr{I}_1 = (1)$ é par

$$\varepsilon(\mathscr{I}_1) = +1$$

Exemplo 4.10 Determinemos as paridades das 2 permutações de 2ª ordem:

$$\begin{cases} \varepsilon(1,2) = +1 \\ \varepsilon(2,1) = -1 \end{cases}$$

Exemplo 4.11 Vejamos, agora, as paridades das 6 permutações de 3ª ordem:

$$\begin{cases} \varepsilon(1,2,3) = +1 \\ \varepsilon(1,3,2) = -1 \\ \varepsilon(2,1,3) = -1 \\ \varepsilon(2,3,1) = +1 \\ \varepsilon(3,1,2) = +1 \\ \varepsilon(3,2,1) = -1 \end{cases}$$

Exemplo 4.12 Determinemos as paridades de seis das 24 permutações de 4ª ordem:

$$\begin{cases} \varepsilon(1,2,3,4) = +1 \\ \varepsilon(1,3,2,4) = -1 \\ \varepsilon(2,4,3,1) = +1 \\ \varepsilon(4,2,3,1) = -1 \\ \varepsilon(3,1,4,2) = -1 \\ \varepsilon(3,4,2,1) = +1 \end{cases}$$

---

2 A paridade de $\sigma \in \mathfrak{S}_n$ pode também ser calculada mediante a contagem de todos os pares (σ_i, σ_j) para os quais se tenha $i < j \wedge \sigma_i > \sigma_j$ (chamados *inversões* de σ). A permutação σ será par ou ímpar, consoante for par ou ímpar o número resultante da contagem referida.

Eis como determinar as 5 últimas paridades anteriores, partindo da identidade $(1, 2, 3, 4)$:

$$\begin{cases} (1,2,3,4) \to (1,3,2,4) \\ (1,2,3,4) \to (2,1,3,4) \to (2,4,3,1) \\ (1,2,3,4) \to (4,2,3,1) \\ (1,2,3,4) \to (3,2,1,4) \to (3,1,2,4) \to (3,1,4,2) \\ (1,2,3,4) \to (3,2,1,4) \to (3,4,1,2) \end{cases}$$

Vamos, agora, demonstrar a proposição que relaciona o sinal da composta de duas permutações com os sinais de cada uma dessas permutações:

Proposição 4.4 – Paridade da composta – *Seja $n \geq 1$ um inteiro e $\sigma, \sigma' \in \mathfrak{S}_n$. Então*

$$\varepsilon(\sigma \circ \sigma') = \varepsilon(\sigma)\varepsilon(\sigma') \tag{4.12}$$

Demonstração:

Em virtude de (4.7), podemos escrever

$$\begin{cases} \sigma\ = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_p \Rightarrow \varepsilon(\sigma) = (-1)^p \\ \sigma' = \tau_1' \circ \tau_2' \circ \cdots \circ \tau_{p'}' \Rightarrow \varepsilon(\sigma') = (-1)^{p'} \end{cases}$$

e daqui conclui-se que

$$\sigma \circ \sigma' = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_p \circ \tau_1' \circ \tau_2' \circ \cdots \circ \tau_{p'}'$$

Deste modo, a paridade de $\sigma \circ \sigma'$ é definida por $p + p'$ e fica, por fim,

$$\varepsilon(\sigma \circ \sigma') = (-1)^{p+p'} = (-1)^p \times (-1)^{p'} = \varepsilon(\sigma)\varepsilon(\sigma') \qquad \square$$

A igualdade (4.12) significa que, se $n \geq 2$, a função

$$\varepsilon \colon \mathfrak{S}_n \to \{-1, +1\}$$

é um homomorfismo do grupo simétrico de ordem n $(\mathfrak{S}_n, \circ)$ sobre o grupo multiplicativo comutativo $(\{-1, +1\}, \times)$.

Desta proposição resulta imediatamente o seguinte

Corolário 4.4.1 (Teorema de Bézout)[3] *Seja $n \geq 2$ um inteiro e $\sigma \in \mathfrak{S}_n$. Então, para qualquer* ***transposição*** *$\tau \in \mathfrak{S}_n$, tem-se sempre*

$$\varepsilon(\sigma \circ \tau) = -\varepsilon(\sigma) \tag{4.13}$$

Demonstração:

Basta atender a que $\varepsilon(\tau) = -1$ e fazer $\sigma' = \tau$ em (4.12). $\square$

---

3 *Bézout, Etienne*: matemático francês (Nemours 1730 – Les Basses-Loges 1783).

Este corolário significa que *toda a permutação muda de paridade, quando sobre ela se executa uma transposição (Teorema de Bézout)*, o que permite olhar para o algoritmo que usámos para determinação da paridade de uma permutação a uma nova luz: partimos de uma permutação (a identidade $\mathscr{I}_n$) cuja paridade conhecemos (par) e, por cada transposição executada, muda essa paridade até obtermos a permutação cuja paridade se pretende determinar:

$$\underbrace{(1,2,3,4)}_{\text{par}} \to \underbrace{(1,3,2,4)}_{\text{ímpar}}$$
$$\underbrace{(1,2,3,4)}_{\text{par}} \to \underbrace{(2,1,3,4)}_{\text{ímpar}} \to \underbrace{(2,4,3,1)}_{\text{par}}$$
$$\underbrace{(1,2,3,4)}_{\text{par}} \to \underbrace{(4,2,3,1)}_{\text{ímpar}}$$
$$\underbrace{(1,2,3,4)}_{\text{par}} \to \underbrace{(3,2,1,4)}_{\text{ímpar}} \to \underbrace{(3,1,2,4)}_{\text{par}} \to \underbrace{(3,1,4,2)}_{\text{ímpar}}$$

$$\underbrace{(1,2,3,4)}_{\text{par}} \to \underbrace{(3,2,1,4)}_{\text{ímpar}} \to \underbrace{(3,4,1,2)}_{\text{par}}$$

Vamos, de seguida, tratar de mais uma consequência de (4.12): *uma permutação e a sua inversa têm sempre a mesma paridade.*

Corolário 4.4.2 *Seja $n \geq 1$ um inteiro e $\sigma \in \mathfrak{S}_n$. Então*

$$\varepsilon\big(\sigma^{-1}\big) = \varepsilon(\sigma) \tag{4.14}$$

Demonstração:

Observando que $\varepsilon(\mathscr{I}_n) = +1$ e que, por definição de inversa, se tem

$$\sigma \circ \sigma^{-1} = \mathscr{I}_n$$

bastará agora usar (4.12) para obter

$$\varepsilon\big(\sigma \circ \sigma^{-1}\big) = \varepsilon(\sigma)\varepsilon\big(\sigma^{-1}\big) = +1$$

Se σ *e* σ^{-1} tivessem paridades diferentes, seria $\varepsilon(\sigma)\varepsilon(\sigma^{-1}) = -1$, contrariando o resultado anterior. ☐

O teorema de Bézout tem também uma implicação interessante: quando $n \geq 2$, *das* $n!$ *permutações de ordem* n*, metade são permutações pares e a outra metade são ímpares*.

Corolário 4.4.3 *No grupo simétrico $\mathfrak{S}_n$ de ordem $n \geq 2$ existem sempre $\frac{n!}{2}$ permutações pares e $\frac{n!}{2}$ permutações ímpares.*

Demonstração:

Seja $\tau \in \mathfrak{S}_n$ uma transposição de ordem n (uma tal transposição existe, visto que $n \geq 2$) e $f\colon \mathfrak{S}_n \to \mathfrak{S}_n$ a função $\sigma \mapsto \sigma \circ \tau$. A função f é uma bijecção de $\mathfrak{S}_n$ sobre si próprio (demonstre!) que transforma as permutações pares em ímpares e vice-versa (Teorema de Bézout). Deste modo, se for p o número de permutações pares e i o número de permutações ímpares, será $p + i = n!$ e

$$p \geq i \wedge i \geq p$$

As desigualdades anteriores implicam imediatamente $p = i = \frac{n!}{2}$. □

Na secção 4.16, usamos o MATHEMATICA© para implementar várias funções úteis para trabalhar com permutações (ver *package* **`ALGA`Determinantes``**).

4.3 Funções multilineares

Definição 4.4 – Funções p-lineares. – *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$ e $p \in \mathbb{N}$ um inteiro positivo. Chama-se* ***função multilinear*** *de ordem p (ou ainda p-linear) sobre E com valores em F a uma função $h\colon E^p \to F$; $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) \mapsto h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p)$, tal que são lineares as p "funções parciais" $h_k\colon E \to F$; $\vec{x}_k \mapsto h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k, \dots, \vec{x}_p)$, isto é, para qualquer $k = 1, 2, \dots, p$ tem-se:*

[ML1] $h(\vec{x}_1, \dots, \vec{x}_k + \vec{y}_k, \dots, \vec{x}_p) = h(\vec{x}_1, \dots, \vec{x}_k, \dots, \vec{x}_p) + h(\vec{x}_1, \dots, \vec{y}_k, \dots, \vec{x}_p)$

[ML2] $h(\vec{x}_1, \dots, \alpha\vec{x}_k, \dots, \vec{x}_p) = \alpha h(\vec{x}_1, \dots, \vec{x}_k, \dots, \vec{x}_p)$

Quando h for uma função escalar ($F = \mathbb{K}$), diremos que se trata de uma ***forma multilinear****. O conjunto das aplicações p-lineares sobre E com valores em F constitui um subespaço vectorial do espaço vectorial $\mathscr{F}(E^p, F)$ de todas as funções de E^p em F*[4]*. Este subespaço vectorial designa-se por $\mathscr{M}_p(E, F)$. Quando $p = 1$, $\mathscr{M}_p(E, F)$ coincide com o espaço vectorial $\mathscr{L}(E, F)$ das funções lineares de E em F.*

A função h diz-se *linear*, *bilinear*, *trilinear*, etc, conforme for $p = 1$, $p = 2$, $p = 3$, etc.

Da condição [ML2] resulta (com $\alpha = 0$)

$$h(\vec{x}_1, \dots, \vec{o}_E, \dots, \vec{x}_p) = \vec{o}_F$$

portanto, *uma função multilinear anula-se se for nulo algum dos seus argumentos*; por outro lado, da aplicação repetida de ML1 e ML2 resulta a igualdade seguinte, para qualquer

4 O espaço $\mathscr{F}(E^p, F)$ está definido através das operações

$$(f + g)(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = f(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) + g(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p)$$
$$(\alpha f)(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = \alpha f(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p)$$

A função $\mathscr{O}_{FE^p}$ (identicamente) nula definida por

$$\mathscr{O}_{FE^p}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = \vec{o}_F$$

é o *vector nulo* deste espaço e o *simétrico* de f será a função $-f$ definida por

$$(-f)(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = -f(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p)$$

combinação linear $\vec{x} = \sum\limits_{i=1}^{m} x_i\vec{e}_i$ de m vectores $\vec{e}_i$ de E,

$$h\Big(\vec{x}_1, \dots, \sum_{i=1}^{m} x_i\vec{e}_i, \dots, \vec{x}_p\Big) = \sum_{i=1}^{m} x_i h(\vec{x}_1, \dots, \vec{e}_i, \dots, \vec{x}_p)$$

Sendo E de dimensão finita n e $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ uma base de E, cada um dos argumentos $\vec{x}_i$ de h será combinação linear dos vectores de e, através das suas coordenadas $(x_{ki})_{\substack{1\le k\le n\\ 1\le i\le p}}$ relativas à base e

$$\vec{x}_i = \sum_{k_i=1}^{n} x_{k_i i}\vec{e}_{k_i};\ i = 1, 2, \dots, p$$

Portanto e devido à multilinearidade de h, será

$$\begin{aligned} h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) &= h\left(\sum_{k_1=1}^{n} x_{k_1 1}\vec{e}_{k_1}, \sum_{k_2=1}^{n} x_{k_2 2}\vec{e}_{k_2}, \dots, \sum_{k_p=1}^{n} x_{k_p p}\vec{e}_{k_p}\right) \\ &= \sum_{k_1, k_2, \dots, k_p=1}^{n} x_{k_1 1}x_{k_2 2}\cdots x_{k_p p}\, h\big(\vec{e}_{k_1}, \vec{e}_{k_2}, \dots, \vec{e}_{k_p}\big) \end{aligned} \tag{4.15}$$

A última expressão mostra que uma função p-linear definida num espaço E com dimensão n fica inteiramente determinada pelos seus $\underbrace{n \times n \times \cdots \times n}_{p \text{ factores}} = n^p$ valores

$$h\big(\vec{e}_{k_1}, \vec{e}_{k_2}, \dots, \vec{e}_{k_p}\big) \in F$$

em todas as combinações $\big(\vec{e}_{k_1}, \vec{e}_{k_2}, \dots, \vec{e}_{k_p}\big)_{1\le k_1, k_2, \dots, k_p \le n}$ de p vectores da base e.

Exemplo 4.13 A função $h\colon \mathbb{R}^2 \to \mathbb{R}$ definida por

$$h(x_1, x_2) = -4\,x_1x_2 \tag{4.16}$$

é uma *forma bilinear* ($p = 2$) sobre $\mathbb{R}$ ($n = 1$).

Exemplo 4.14 Em geral, as funções $h\colon \mathbb{R}^2 \to \mathbb{R}$ definidas por

$$h(x_1, x_2) = a\,x_1x_2 \tag{4.17}$$

e onde a é uma constante real são *formas bilineares* ($p = 2$) sobre $\mathbb{R}$ ($n = 1$).

Exemplo 4.15 Mais geralmente, sendo $\mathbb{K}$ um corpo, a função $h\colon \mathbb{K}^p \to \mathbb{K}$ definida por

$$h(x_1, x_2, \dots, x_p) = a\,x_1x_2\cdots x_p = a\prod_{j=1}^{p} x_j \tag{4.18}$$

e onde a é uma constante de $\mathbb{K}$ é uma *forma p-linear* sobre $\mathbb{K}$ (verifique!). Observe que $a = h(1, 1, \dots, 1)$.

Exemplo 4.16 A função $h\colon (\mathbb{R}^2)^2 \to \mathbb{R}$ definida, para $\vec{x} = (x_1, x_2),\ \vec{y} = (y_1, y_2) \in \mathbb{R}^2$ por

$$h(\vec{x}, \vec{y}) = x_1y_1 - 2x_1y_2 + x_2y_1 + 3x_2y_2 \tag{4.19}$$

é uma *forma bilinear* sobre $\mathbb{R}^2$ (verifique!).

Exemplo 4.17 Mais geralmente, as funções $h\colon (\mathbb{R}^2)^2 \to \mathbb{R}$ definidas, para $\vec{x} = (x_1, x_2)$, $\vec{y} = (y_1, y_2) \in \mathbb{R}^2$ por

$$h(\vec{x}, \vec{y}) = \sum_{i,j=1}^{2} a_{ij}x_iy_j = X^{\mathrm{T}}AY = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{4.20}$$

(onde os a_{ij} são constantes reais) são *formas bilineares* sobre $\mathbb{R}^2$ (prove!). Observe que os a_{ij} são as imagens das 4 combinações possíveis dos 2 vectores da base canónica de $\mathbb{R}^2$ por h:

$$\begin{array}{ll} a_{11} = h((1,0),(1,0)) & a_{12} = h((1,0),(0,1)) \\ a_{21} = h((0,1),(1,0)) & a_{22} = h((0,1),(0,1)) \end{array}$$

Exemplo 4.18 Mais geralmente ainda, as funções $h\colon (\mathbb{K}^n)^2 \to \mathbb{K}$ definidas, para $\vec{x} = (x_1, x_2, \ldots, x_n),\ \vec{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{K}^n$ por

$$h(\vec{x}, \vec{y}) = \sum_{i,j=1}^{n} a_{ij}x_iy_j = X^{\mathrm{T}}AY = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{4.21}$$

(onde os a_{ij} são constantes de $\mathbb{K}$) são *formas bilineares* sobre $\mathbb{K}^n$. Observe que os a_{ij} são as imagens dos n^2 pares ordenados possíveis de vectores da base canónica de $\mathbb{K}^n$ por h. X e Y são matrizes-coluna e a matriz $A = [a_{ij}\ ] \in \mathbb{K}^{n,n}$ é chamada a *matriz da forma bilinear* h. $X^{\mathrm{T}}AY$ é uma matriz de tipo 1×1 que identificamos com o seu elemento escalar e é óbvio que esta matriz é simétrica, donde $X^{\mathrm{T}}AY = (X^{\mathrm{T}}AY)^{\mathrm{T}} = Y^{\mathrm{T}}A^{\mathrm{T}}X$.

Exemplo 4.19 As funções $h\colon (\mathbb{R}^2)^3 \to \mathbb{R}$ definidas, para quaisquer $\vec{x} = (x_1, x_2)$, $\vec{y} = (y_1, y_2),\ \vec{z} = (z_1, z_2) \in \mathbb{R}^2$ por

$$h(\vec{x}, \vec{y}, \vec{z}) = \sum_{i,j,k=1}^{2} a_{ijk}x_iy_jz_k \tag{4.22}$$

(onde os a_{ijk} são 2^3 constantes reais) são *formas trilineares* sobre $\mathbb{R}^2$ (verifique!). Observe que os a_{ijk} são as imagens dos $2^3 = 8$ ternos ordenados possíveis dos 2 vectores da base canónica de $\mathbb{R}^2$ por h.

Exemplo 4.20 Em geral, sendo $\mathbb{K}$ um corpo, as funções $h\colon (\mathbb{K}^2)^p \to \mathbb{K}$ definidas, para quaisquer $\vec{x}_1 = (x_{11}, x_{21}),\ \vec{x}_2 = (x_{12}, x_{22}), \ldots, \vec{x}_p = (x_{1p}, x_{2p}) \in \mathbb{K}^2$ por

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = \sum_{k_1,k_2,\ldots,k_p=1}^{2} a_{k_1k_2\cdots k_p}\, x_{k_11}x_{k_22}\cdots x_{k_pp} \tag{4.23}$$

(onde os $a_{k_1k_2\cdots k_p}$ são 2^p constantes de $\mathbb{K}$) são *formas p-lineares* sobre $\mathbb{K}^2$ (verifique!). Observe que os $a_{k_1k_2\cdots k_p}$ são as imagens das 2^p listas ordenadas possíveis de p vectores da base canónica de $\mathbb{K}^2$ por h.

Exemplo 4.21 Mais geralmente ainda, sendo $\mathbb{K}$ um corpo, as funções $h\colon (\mathbb{K}^n)^p \to \mathbb{K}$ definidas, para quaisquer

$$\vec{x}_1 = (x_{11}, x_{21}, \ldots, x_{n1}),\ \vec{x}_2 = (x_{12}, x_{22}, \ldots, x_{n2}), \ldots, \vec{x}_p = (x_{1p}, x_{2p}, \ldots, x_{np}) \in \mathbb{K}^n$$

por

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = \sum_{k_1, k_2, \ldots, k_p = 1}^{n} a_{k_1 k_2 \cdots k_p} \left(x_{k_1 1} x_{k_2 2} \cdots x_{k_p p} \right) \tag{4.24}$$

(onde os $a_{k_1 k_2 \cdots k_p}$ são n^p constantes de $\mathbb{K}$) são *formas p-lineares* sobre $\mathbb{K}^n$ (verifique!). Observe que os $a_{k_1 k_2 \cdots k_p}$ são as imagens das n^p listas ordenadas possíveis de p dos n vectores da base canónica de $\mathbb{K}^n$ por meio de h.

Exemplo 4.22 Seja I um intervalo compacto de $\mathbb{R}$ e $\mathscr{C}(I, \mathbb{R})$ o espaço das funções reais contínuas em I e considere, para quaisquer $f, g \in \mathscr{C}(I, \mathbb{R})$, a função $h\colon (\mathscr{C}(I, \mathbb{R}))^2 \to \mathbb{R}$ definida por

$$h(f, g) = \int_I f(x) g(x) \, \mathrm{d}x \tag{4.25}$$

Facilmente se verifica que h é uma *forma bilinear* sobre $\mathscr{C}(I, \mathbb{R})$.

Exemplo 4.23 Seja I um intervalo compacto de $\mathbb{R}$ e $\mathscr{C}(I, \mathbb{R})$ o espaço das funções reais contínuas em I e considere, para quaisquer funções $f_1, f_2, \ldots, f_p \in \mathscr{C}(I, \mathbb{R})$, a aplicação $h\colon (\mathscr{C}(I, \mathbb{R}))^p \to \mathbb{R}$ definida por

$$h(f_1, f_2, \ldots, f_p) = \int_I f_1(x) f_2(x) \cdots f_p(x) \, \mathrm{d}x \tag{4.26}$$

Facilmente se verifica que h é uma *forma p-linear* sobre $\mathscr{C}(I, \mathbb{R})$.

Exemplo 4.24 Se forem E, F e G espaços vectoriais sobre um corpo $\mathbb{K}$, $h\colon E^p \to F$ uma função p-linear sobre E com valores em F e $u\colon F \to G$ uma função linear, facilmente se prova que $g = u \circ h$ é uma função p-linear sobre E com valores em G

$$g(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = u(h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)) \tag{4.27}$$

$$\begin{tikzcd}
E^p \arrow[r, "h"] \arrow[dr, "g = u \circ h"'] & F \arrow[d, "u"] \\
& G
\end{tikzcd}$$

Fig. 4.3 – Composição de uma função linear u com uma função p-linear h.

Em particular, se dim$F = m$ e se $f = \left(\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\right)$ for uma base de F, as projecções

$$\mathrm{pr}_i\colon F \to \mathbb{K}; \vec{y} = \sum_{r=1}^{m} y_r \vec{f}_r \mapsto y_i, \text{ para } i = 1, 2, \ldots, m$$

são formas lineares de F em $\mathbb{K}$ e, portanto, as coordenadas $h_i = \mathrm{pr}_i \circ h\colon E^p \to \mathbb{K}$ da função $h\colon E^p \to F$ p-linear com valores em F são formas p-lineares h_i sobre E

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = \sum_{i=1}^{m} h_i(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)\vec{f}_i \tag{4.28}$$

$$\begin{tikzcd}
E^p \arrow[r, "h"] \arrow[dr, "h_i = \mathrm{pr}_i \circ h"'] & F \arrow[d, "\mathrm{pr}_i"] \\
 & \mathbb{K}
\end{tikzcd}$$

Fig. 4.4 – As Coordenadas h_i de uma função p-linear são formas p-lineares.

Reciprocamente, dadas m formas p-lineares $h_i\colon E^p \to \mathbb{K}$, a função $h\colon E^p \to F$ definida por

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = \sum_{i=1}^{m} h_i(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)\vec{f}_i$$

é uma função p-linear sobre E e com valores em F.

Exemplo 4.25 Considere-se a função $h\colon (\mathbb{R}^3)^2 \to \mathbb{R}^3$ definida para quaisquer vectores $\vec{x} = (x_1, x_2, x_3), \vec{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ por

$$h(\vec{x}, \vec{y}) = (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1) \tag{4.29}$$

Facilmente se comprova que h é uma *função bilinear* sobre $\mathbb{R}^3$, tomando valores também em $\mathbb{R}^3$ (trata-se, afinal, de uma lei de composição interna em $\mathbb{R}^3$).

Exemplo 4.26 Se forem E, F e G espaços vectoriais sobre um corpo $\mathbb{K}$, $h\colon E^p \to F$ uma função p-linear sobre E com valores em F e $u\colon G \to E$ uma função linear, então a função $f\colon G^p \to F$ definida por

$$f(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = h(u(\vec{x}_1), u(\vec{x}_2), \ldots, u(\vec{x}_p)) \tag{4.30}$$

é uma *aplicação p-linear* sobre G com valores em F: trata-se da composta de h com a função $u_p\colon G^p \to E^p; (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) \mapsto (u(\vec{x}_1), u(\vec{x}_2), \ldots, u(\vec{x}_p))$.

$$\begin{tikzcd} G^p \arrow[r, "u_p"] \arrow[dr, "f = h \circ u_p"'] & E^p \arrow[d, "h"] \\ & F \end{tikzcd}$$

Fig. 4.5 – Composição de uma função p-linear h com a função u_p.

4.4 Funções multilineares alternadas, simétricas e anti-simétricas

Nesta secção, vamos definir alguns tipos particulares de funções multilineares: as funções *alternadas, simétricas* e *anti-simétricas*.

Definição 4.5 – Funções p-lineares alternadas. – *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$, $p \geq 1$ um inteiro e $h\colon E^p \to F$ uma função p-linear sobre E com valores em F. Diz-se que h é **alternada** sse, para quaisquer índices $1 \leq i < j \leq p^{(5)}$, se tem*

$$\vec{x}_i = \vec{x}_j \Rightarrow h(\vec{x}_1, \ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots, \vec{x}_p) = \vec{o}_F \tag{4.31}$$

As funções p-lineares alternadas sobre E com valores em F formam um subespaço vectorial de $\mathscr{M}_p(E, F)$, que designaremos por $\mathscr{A}_p(E, F)$.

Portanto, uma função multilinear diz-se *alternada* sse é nula sempre que forem iguais dois quaisquer dos seus argumentos.

Definição 4.6 – Funções p-lineares simétricas e anti-simétricas. – *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$ e $p \geq 1$ um inteiro. Diz-se que uma função p-linear $h\colon E^p \to F$ sobre E com valores em F é:*

i) ***simétrica*** *sse para qualquer permutação $\sigma \in \mathfrak{S}_p$ e para toda a sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ de vectores de E, se tem*

$$h\big(\vec{x}_{\sigma_1}, \vec{x}_{\sigma_2}, \ldots, \vec{x}_{\sigma_p}\big) = h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) \tag{4.32}$$

ii) ***anti-simétrica*** *sse, nas mesmas condições, se tem*

$$h\big(\vec{x}_{\sigma_1}, \vec{x}_{\sigma_2}, \ldots, \vec{x}_{\sigma_p}\big) = \varepsilon(\sigma) h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) \tag{4.33}$$

As aplicações p-lineares simétricas sobre E com valores em F constituem um subespaço vectorial de $\mathscr{M}_p(E, F)$, que designaremos por $\mathscr{S}_p(E, F)$.

---

5 Observe que, se $p \geq 2$, o número de pares (i, j) com $1 \leq i < j \leq p$ é de $\binom{p}{2}$, ou seja, é igual ao número de transposições de ordem p.

Observação:

- Quando $p = 1$, observemos que toda a função p-linear é simultaneamente alternada, simétrica e anti-simétrica.

Exemplo 4.27 Atente na figura 4.3 e sejam E, F e G espaços vectoriais sobre um corpo $\mathbb{K}$ e $h\colon E^p \to F$ uma função p-linear sobre E com valores em F alternada (respectivamente simétrica, anti-simétrica). Para qualquer função linear $u\colon F \to G$, a função p-linear $g = u \circ h$ sobre E com valores em G definida por

$$g(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = u(h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p))$$

é também alternada (respectivamente simétrica, anti-simétrica).

Exemplo 4.28 Com referência à figura 4.5, sejam E, F e G espaços vectoriais sobre um corpo $\mathbb{K}$ e $h\colon E^p \to F$ uma função p-linear sobre E com valores em F alternada (respectivamente simétrica, anti-simétrica). Para qualquer função linear $u\colon G \to E$, a função $f\colon G^p \to F$ definida por

$$f(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = h(u(\vec{x}_1), u(\vec{x}_2), \dots, u(\vec{x}_p))$$

é uma aplicação p-linear alternada (respectivamente simétrica, anti-simétrica) sobre G com valores em F.

Podemos caracterizar as funções p-lineares simétricas de outras formas equivalentes à definição 4.6.*i*, como se mostra na

Proposição 4.5 – Caracterização das funções simétricas – *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$ e $p \geq 2$ um inteiro e $h\colon E^p \to F$ uma função p-linear sobre E com valores em F. Então são equivalentes as seguintes proposições:*

i) *h é simétrica.*

ii) *Para toda a sequência $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p)$ de p vectores de E e para todo o inteiro $1 \leq i < p$, tem-se*

$$h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{i+1}, \vec{x}_i, \dots, \vec{x}_p) = h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \vec{x}_{i+1}, \dots, \vec{x}_p) \qquad (4.34)$$

iii) *Para toda a sequência $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p)$ de p vectores de E e para quaisquer inteiros $1 \leq i < j \leq p$, tem-se*

$$h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_j, \dots, \vec{x}_i, \dots, \vec{x}_p) = h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_p) \qquad (4.35)$$

Demonstração:

As condições são necessárias, visto que as condições (4.34) e (4.35) resultam de (4.32) fazendo σ igual a uma transposição elementar ou a uma transposição. Mas elas são também suficientes, porque (proposições 4.1 e 4.3) qualquer permutação $\sigma \in \mathfrak{S}_p$ é uma composição de transposições (gerais ou elementares). $\square$

Podemos enunciar uma proposição semelhante, que se demonstra da mesma forma que a anterior, para as funções p-lineares *anti-simétricas*:

Proposição 4.6 – Caracterização das funções anti-simétricas – *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$, $p \geq 2$ um inteiro e $h\colon E^p \to F$ uma função p-linear sobre E com valores em F. Então, as seguintes proposições são equivalentes:*

i) *h é anti-simétrica.*

ii) *Para toda a sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ de p vectores de E e para todo o inteiro $1 \leq i < p$, tem-se*

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{i+1}, \vec{x}_i, \ldots, \vec{x}_p) = -h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \vec{x}_{i+1}, \ldots, \vec{x}_p) \tag{4.36}$$

iii) *Para toda a sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ de p vectores de E e para quaisquer inteiros $1 \leq i < j \leq p$, tem-se*

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_j, \ldots, \vec{x}_i, \ldots, \vec{x}_p) = -h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots, \vec{x}_p) \tag{4.37}$$

Observações:

- Devido a estas equivalências, alguns autores usam qualquer das condições anteriores como a própria definição de função *simétrica* (respectivamente, *anti-simétrica*).

- Como consequência das proposições anteriores, podemos afirmar:

– Numa aplicação p-linear *simétrica* é irrelevante a ordem dos seus p argumentos (o que significa que não se altera a função, quando executamos uma operação elementar do tipo 1 sobre a lista $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ dos seus argumentos – proposição 4.5.*iii*).

– Pelo contrário, uma função p-linear *anti-simétrica* muda de sinal quando se faz uma transposição de dois quaisquer dos seus argumentos (o que significa que a função muda de sinal, quando executamos uma qualquer operação elementar do tipo 1 sobre a lista $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ dos seus argumentos – proposição 4.6.*iii*).

Exemplo 4.29 A forma bilinear $h\colon (\mathbb{R}^2)^2 \to \mathbb{R}$ definida, para todo o $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$ e todo o $\vec{y} = (y_1, y_2) \in \mathbb{R}^2$, por

$$h(\vec{x}, \vec{y}) = 2x_1y_1 - 3x_1y_2 - 3x_2y_1 - x_2y_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -3 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = X^{\mathrm{T}}AY$$

é *simétrica*, visto que

$$h(\vec{y}, \vec{x}) = 2y_1x_1 - 3y_1x_2 - 3y_2x_1 - y_2x_2 = 2x_1y_1 - 3x_1y_2 - 3x_2y_1 - x_2y_2 = h(\vec{x}, \vec{y})$$

Observe que a matriz A é simétrica.

Exemplo 4.30 A função trilinear $h\colon (\mathbb{R}^2)^3 \to \mathbb{R}^2$ que transforma os terno de vectores $\vec{x} = (x_1, x_2), \vec{y} = (y_1, y_2), \vec{z} = (z_1, z_2) \in \mathbb{R}^2$ no vector

$$h(\vec{x}, \vec{y}, \vec{z}) = (3x_1y_1z_1 - 2x_1y_2z_2 - 2x_2y_1z_2 - 2x_2y_2z_1, x_1y_1z_1 + x_2y_1z_1 + x_1y_2z_1 + x_1y_1z_2)$$

é *simétrica*, visto que, trocando dois vectores consecutivos, não se altera o seu valor:

Assim, trocando $\vec{x}$ com $\vec{y}$ obtém-se

$$\begin{aligned} h(\vec{y}, \vec{x}, \vec{z}) &= (3y_1x_1z_1 - 2y_1x_2z_2 - 2y_2x_1z_2 - 2y_2x_2z_1, y_1x_1z_1 + y_2x_1z_1 + y_1x_2z_1 + y_1x_1z_2) \\ &= (3x_1y_1z_1 - 2x_1y_2z_2 - 2x_2y_1z_2 - 2x_2y_2z_1, x_1y_1z_1 + x_2y_1z_1 + x_1y_2z_1 + x_1y_1z_2) \\ &= h(\vec{x}, \vec{y}, \vec{z}) \end{aligned}$$

e, por fim, trocando $\vec{y}$ com $\vec{z}$,

$$\begin{aligned} h(\vec{x}, \vec{z}, \vec{y}) &= (3x_1z_1y_1 - 2x_1z_2y_2 - 2x_2z_1y_2 - 2x_2z_2y_1, x_1z_1y_1 + x_2z_1y_1 + x_1z_2y_1 + x_1z_1y_2) \\ &= (3x_1y_1z_1 - 2x_1y_2z_2 - 2x_2y_1z_2 - 2x_2y_2z_1, x_1y_1z_1 + x_2y_1z_1 + x_1y_2z_1 + x_1y_1z_2) \\ &= h(\vec{x}, \vec{y}, \vec{z}) \end{aligned}$$

Exemplo 4.31 Sendo $h\colon E^2 \to F$ uma função bilinear qualquer, a função $f\colon E^2 \to F$ definida por

$$f(\vec{x}, \vec{y}) = \frac{1}{2}(h(\vec{x}, \vec{y}) + h(\vec{y}, \vec{x})) \tag{4.38}$$

é uma função *simétrica*, visto que

$$f(\vec{y}, \vec{x}) = \frac{1}{2}(h(\vec{y}, \vec{x}) + h(\vec{x}, \vec{y})) = \frac{1}{2}(h(\vec{x}, \vec{y}) + h(\vec{y}, \vec{x})) = f(\vec{x}, \vec{y})$$

Exemplo 4.32 As funções dos exemplos 4.22 e 4.23 são *simétricas* (prove!).

Exemplo 4.33 Mais geralmente, sendo $h\colon E^p \to F$ uma função p-linear qualquer sobre E com valores em F, a função $f\colon E^p \to F$ definida por

$$f(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = k \sum_{\sigma \in \mathfrak{S}_p} h\big(\vec{x}_{\sigma_1}, \vec{x}_{\sigma_2}, \ldots, \vec{x}_{\sigma_p}\big) \tag{4.39}$$

onde k é um escalar constante e o somatório é estendido às $p!$ permutações de ordem p, é uma função *simétrica*, visto que para qualquer permutação $\omega \in \mathfrak{S}_p$ se tem, fazendo $\sigma' = \sigma \circ \omega$,

$$\begin{aligned} f\big(\vec{x}_{\omega_1}, \vec{x}_{\omega_2}, \ldots, \vec{x}_{\omega_p}\big) &= k \sum_{\sigma \in \mathfrak{S}_p} h\big(\vec{x}_{\sigma'_1}, \vec{x}_{\sigma'_2}, \ldots, \vec{x}_{\sigma'_p}\big) \\ &= k \sum_{\sigma' \in \mathfrak{S}_p} h\big(\vec{x}_{\sigma'_1}, \vec{x}_{\sigma'_2}, \ldots, \vec{x}_{\sigma'_p}\big) \\ &= f(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) \end{aligned}$$

Exemplo 4.34 A função bilinear $h\colon (\mathbb{R}^3)^2 \to \mathbb{R}^3$ do exemplo 4.25 é *anti-simétrica*. Para quaisquer vectores $\vec{x} = (x_1, x_2, x_3), \vec{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ tem-se

$$h(\vec{x}, \vec{y}) = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) \tag{4.40}$$

e trocando a ordem dos argumentos, vem

$$\begin{aligned} h(\vec{y}, \vec{x}) &= (y_2x_3 - y_3x_2, y_3x_1 - y_1x_3, y_1x_2 - y_2x_1) \\ &= (-(x_2y_3 - x_3y_2), -(x_3y_1 - x_1y_3), -(x_1y_2 - x_2y_1)) \\ &= -(x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) = -h(\vec{x}, \vec{y}) \end{aligned}$$

Exemplo 4.35 O leitor pode facilmente verificar que a forma trilinear $h\colon (\mathbb{R}^3)^3 \to \mathbb{R}$ definida por

$$h(\vec{x}, \vec{y}, \vec{z}) = x_1y_2z_3 + x_3y_1z_2 + x_2y_3x_1 - x_3y_2z_1 - x_2y_1z_3 - x_1y_3z_2$$

é *anti-simétrica*, visto que

$$h(\vec{x}, \vec{y}, \vec{z}) = -h(\vec{y}, \vec{x}, \vec{z}) = -h(\vec{x}, \vec{z}, \vec{y})$$

Exemplo 4.36 Sendo $h\colon E^2 \to F$ uma função bilinear qualquer, a função $f\colon E^2 \to F$ definida por

$$f(\vec{x}, \vec{y}) = \frac{1}{2}(h(\vec{x}, \vec{y}) - h(\vec{y}, \vec{x})) \tag{4.41}$$

é uma função *anti-simétrica*, visto que

$$f(\vec{y}, \vec{x}) = \frac{1}{2}(h(\vec{y}, \vec{x}) - h(\vec{x}, \vec{y})) = -\frac{1}{2}(h(\vec{x}, \vec{y}) - h(\vec{y}, \vec{x})) = -f(\vec{x}, \vec{y})$$

Exemplo 4.37 Mais geralmente, sendo $h\colon E^p \to F$ uma função p-linear qualquer, a função p-linear $f\colon E^p \to F$ definida por

$$f(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = k \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma)\, h\big(\vec{x}_{\sigma_1}, \vec{x}_{\sigma_2}, \dots, \vec{x}_{\sigma_p}\big) \tag{4.42}$$

onde k é um escalar constante e o somatório é estendido às $p!$ permutações de ordem p, é uma função anti-simétrica, visto que para qualquer permutação $\omega \in \mathfrak{S}_p$ se tem, fazendo $\sigma' = \sigma \circ \omega$ e atendendo a que $\varepsilon(\omega^{-1}) = \varepsilon(\omega)$,

$$\begin{aligned} f\big(\vec{x}_{\omega_1}, \vec{x}_{\omega_2}, \dots, \vec{x}_{\omega_p}\big) &= k \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma)\, h\big(\vec{x}_{\sigma'_1}, \vec{x}_{\sigma'_2}, \dots, \vec{x}_{\sigma'_p}\big) \\ &= k \sum_{\sigma' \in \mathfrak{S}_p} \varepsilon\big(\sigma' \circ \omega^{-1}\big)\, h\big(\vec{x}_{\sigma'_1}, \vec{x}_{\sigma'_2}, \dots, \vec{x}_{\sigma'_p}\big) \\ &= k \sum_{\sigma' \in \mathfrak{S}_p} \varepsilon(\sigma')\varepsilon\big(\omega^{-1}\big)\, h\big(\vec{x}_{\sigma'_1}, \vec{x}_{\sigma'_2}, \dots, \vec{x}_{\sigma'_p}\big) \\ &= \varepsilon(\omega)\, k \sum_{\sigma' \in \mathfrak{S}_p} \varepsilon(\sigma')\, h\big(\vec{x}_{\sigma'_1}, \vec{x}_{\sigma'_2}, \dots, \vec{x}_{\sigma'_p}\big) \\ &= \varepsilon(\omega) f(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) \end{aligned}$$

As funções multilineares *alternadas* gozam de importantes propriedades, que condensamos na seguinte

Proposição 4.7 – Propriedades das funções alternadas – *Sejam E e F espaços vectoriais sobre um corpo $\mathbb{K}$, $p \geq 2$ um inteiro e $h\colon E^p \to F$ uma função p-linear. Então:*

i) *Para que h seja alternada é necessário e suficiente que, para toda a sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ de p vectores de E contendo dois vectores consecutivos $\vec{x}_i$ e $\vec{x}_{i+1}$ iguais (onde $1 \leq i < p$), se tenha*

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \vec{x}_{i+1}, \ldots, \vec{x}_p) = \vec{o}_F \tag{4.43}$$

ii) *Se h é alternada, então h é anti-simétrica.*

iii) *Se o corpo $\mathbb{K}$ tiver característica diferente de 2, a recíproca da implicação anterior também é verdadeira, isto é, toda a função p-linear h anti-simétrica será alternada.*

iv) *Se h é alternada, não se altera o valor de h para a sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ de vectores de E somando a um qualquer deles $\vec{x}_i$ uma combinação linear $\sum\limits_{k \neq i} \beta_k \vec{x}_k$ dos outros vectores, ou seja,*

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_p) = h\Big(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i + \sum_{k \neq i} \beta_k \vec{x}_k, \ldots, \vec{x}_p\Big) \tag{4.44}$$

Em particular, uma operação elementar do tipo 3 executada sobre a sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ não altera o valor de h.

v) *Se h é alternada, então, para qualquer sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_p)$ de vectores de E, qualquer $1 \leq i \leq p$ e quaisquer $\alpha, \beta_k \in \mathbb{K}$, tem-se*

$$h\Big(\vec{x}_1, \vec{x}_2, \ldots, \alpha\vec{x}_i + \sum_{k \neq i} \beta_k \vec{x}_k, \ldots, \vec{x}_p\Big) = \alpha h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_p) \tag{4.45}$$

Em particular, a execução de uma operação de tipo 4 (substituição de um vector $\vec{x}_i$ por $\alpha\vec{x}_i + \beta\vec{x}_j$, onde $1 \leq i \neq j \leq p$ e $\alpha \neq 0$) dá

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots, \vec{x}_p) = \frac{1}{\alpha} h(\vec{x}_1, \vec{x}_2, \ldots, \alpha\vec{x}_i + \beta\vec{x}_j, \ldots, \vec{x}_j, \ldots, \vec{x}_p)$$

vi) *Se $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ é linearmente dependente e h é alternada, então:*

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = \vec{o}_F \tag{4.46}$$

vii) *Se h é alternada e se $n = \dim E$, então:*

- *Se $p > n$, h é a função nula, isto é, $\mathscr{A}_p(E, F) = \{\mathbb{O}_{FE^p}\}$.*

- *Se $p \leq n$ e $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ é uma base de E, então h é completamente determinada pelos seus valores nas $\binom{n}{p}$ sequências $(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p})$, com*

$1 \le i_1 < i_2 < \cdots < i_p \le n$. Exprimindo cada vector $\vec{x}_i$ através das suas coordenadas na base e por

$$\vec{x}_i = \sum_{k_i=1}^{n} x_{k_i i} \vec{e}_{k_i}, \text{ para } i = 1, 2, \dots, p$$

a função alternada h vem dada pela seguinte expressão:

$$h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} \left(\sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) x_{i_{\sigma_1} 1} x_{i_{\sigma_2} 2} \cdots x_{i_{\sigma_p} p} \right) h\left(\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{e}_{i_p}\right) \tag{4.47}$$

Na expressão do 2º membro, o somatório exterior tem $\binom{n}{p}$ parcelas e o somatório interior estende-se ao grupo simétrico $\mathfrak{S}_p$ das $p!$ permutações $(\sigma_1, \sigma_2, \dots, \sigma_p)$ de ordem p e em que $\varepsilon(\sigma)$ é o sinal de σ.

viii) Em particular, se $n = \dim E$ e h for n-linear alternada ($p = n$), então h será completamente determinada pelo seu valor $h(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ numa base $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ de E, vindo:

$$h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma_1 1} x_{\sigma_2 2} \cdots x_{\sigma_n n} \right) h(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) \tag{4.48}$$

onde x_{ki} é a k-ésima coordenada do vector $\vec{x}_i$ em relação à base $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$.

ix) Uma função n-linear alternada $h \ne \mathbb{O}_{FE^n}$ (não identicamente nula) sobre um espaço com dimensão n anula-se se e só se os seus n argumentos formarem uma sequência linearmente dependente.

$$h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = \vec{o}_F \Leftrightarrow (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \text{ é linearmente dependente} \tag{4.49}$$

Isto equivale ainda a dizer que uma sequência $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ é linearmente independente (é uma base de E) sse $h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \ne \vec{o}_F$.

Demonstração:

i) A condição apresentada é obviamente necessária, visto que existem dois vectores $\vec{x}_i$ e $\vec{x}_{i+1}$ iguais. E também é suficiente, como se prova a seguir:

- Seja $1 \le i < p$ e consideremos uma sequência $x = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p)$ de vectores de E e que h se anula sempre que dois vectores consecutivos de x sejam iguais; então será

$$h(\vec{x}_1, \dots, \vec{x}_i + \vec{x}_{i+1}, \vec{x}_i + \vec{x}_{i+1}, \dots, \vec{x}_p) = \vec{o}_F$$

Devido à multilinearidade de h, vem

$$\underbrace{h(\vec{x}_1, \dots, \vec{x}_i, \vec{x}_i, \dots, \vec{x}_p)}_{=\vec{o}_F} + h(\vec{x}_1, \dots, \vec{x}_i, \vec{x}_{i+1}, \dots, \vec{x}_p)$$
$$+ h(\vec{x}_1, \dots, \vec{x}_{i+1}, \vec{x}_i, \dots, \vec{x}_p) + \underbrace{h(\vec{x}_1, \dots, \vec{x}_{i+1}, \vec{x}_{i+1}, \dots, \vec{x}_p)}_{=\vec{o}_F} = \vec{o}_F$$

Mas ao 1ª e a 4ª parcelas do 1º membro da igualdade anterior são nulas por hipótese, donde resulta

$$h(\vec{x}_1, \dots, \vec{x}_i, \vec{x}_{i+1}, \dots, \vec{x}_p) = -h(\vec{x}_1, \dots, \vec{x}_{i+1}, \vec{x}_i, \dots, \vec{x}_p)$$

Daqui se segue que h é anti-simétrica, pela proposição 4.6.*ii*.

■ Por fim, sejam $1 \leq i < j \leq p$ dois índices e suponha-se que $\vec{x}_i = \vec{x}_j$. Se $j = i+1$, o resultado é imediato; se $i+1 < j \leq p$, trocando os vectores $\vec{x}_{i+1}$ e $\vec{x}_j$, os vectores $\vec{x}_i$ e $\vec{x}_j$ ficam consecutivos e por ser h anti-simétrica, vem visto que os vectores

$$h(\vec{x}_1, \dots, \vec{x}_i, \vec{x}_{i+1}, \dots, \vec{x}_j, \dots, \vec{x}_p) = -h\big(\vec{x}_1, \dots, \underbrace{\vec{x}_i, \vec{x}_j}_{\text{iguais}}, \dots, \vec{x}_{i+1}, \dots, \vec{x}_p\big) = -\vec{o}_F = \vec{o}_F$$

ii) Como h é alternada, teremos

$$\begin{aligned}\vec{o}_F &= h(\vec{x}_1, \dots, \vec{x}_i + \vec{x}_j, \dots, \vec{x}_i + \vec{x}_j, \dots, \vec{x}_p)\\ &= \underbrace{h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_i, \dots, \vec{x}_p)}_{=\vec{o}_F} + h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_p)\\ &\quad + h(\vec{x}_1, \dots, \vec{x}_j, \dots, \vec{x}_i, \dots, \vec{x}_p) + \underbrace{h(\vec{x}_1, \dots, \vec{x}_j, \dots, \vec{x}_j, \dots, \vec{x}_p)}_{=\vec{o}_F}\\ &= h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_p) + h(\vec{x}_1, \dots, \vec{x}_j, \dots, \vec{x}_i, \dots, \vec{x}_p)\end{aligned}$$

Daqui se segue imediatamente que h é anti-simétrica (proposição 4.6.*iii*):

$$h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_p) = -h(\vec{x}_1, \dots, \vec{x}_j, \dots, \vec{x}_i, \dots, \vec{x}_p)$$

iii) Suponha-se que h é anti-simétrica e calculemos $h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_p)$ com $\vec{x}_i = \vec{x}_j$.

$$\begin{aligned}&h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_i, \dots, \vec{x}_p) = -h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_i, \dots, \vec{x}_p)\\ &\Rightarrow h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_i, \dots, \vec{x}_p) + h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_i, \dots, \vec{x}_p) = \vec{o}_F\\ &\Rightarrow \underbrace{(1+1)}_{\neq 0} h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_i, \dots, \vec{x}_p) = \vec{o}_F\end{aligned}$$

Como o corpo $\mathbb{K}$ tem característica $\neq 2$, será $1+1 \neq 0$[6] e, portanto, h será alternada, porque:

$$h(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_i, \dots, \vec{x}_p) = \vec{o}_F$$

6 Como exemplo de corpo em que $1+1=0$ e no qual este resultado não é válido, considere o corpo $\mathbb{Z}_2 = \{0, 1\}$, com as operações de adição e produto seguintes

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

iv) Por h ser p-linear e alternada, tem-se:

$$\begin{aligned} h\Big(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i + \sum_{k\neq i}\beta_k\vec{x}_k, \ldots, \vec{x}_p\Big) &= h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_p) \\ &\quad + \sum_{k\neq i}\beta_k \underbrace{h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k, \ldots, \vec{x}_p)}_{\substack{=\vec{o}_F, \text{ porque o vector } \vec{x}_k \\ \text{é sempre igual a um dos restantes}}} \\ &= h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_p) \end{aligned}$$

O resultado enunciado para uma operação elementar do tipo 3 é apenas um caso particular desta proposição (em que, no somatório $\sum_{k\neq i}\beta_k\vec{x}_k$, todos os β_k são nulos à excepção de um).

v) Usando a alínea anterior e a p-linearidade de h, obtém-se imediatamente o resultado

$$\begin{aligned} h\Big(\vec{x}_1, \vec{x}_2, \ldots, \alpha\vec{x}_i + \sum_{k\neq i}\beta_k\vec{x}_k, \ldots, \vec{x}_p\Big) &= h(\vec{x}_1, \vec{x}_2, \ldots, \alpha\vec{x}_i, \ldots, \vec{x}_p) \\ &= \alpha h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_p) \end{aligned}$$

Quanto à operação de tipo 4, ela resulta de um caso particular da propriedade anterior:

$$h(\vec{x}_1, \vec{x}_2, \ldots, \alpha\vec{x}_i + \beta\vec{x}_j, \ldots, \vec{x}_j, \ldots, \vec{x}_p) = \alpha h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots, \vec{x}_p)$$

para o que basta, agora, multiplicar ambos os membros por $1/\alpha$.

vi) Se $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ é linearmente dependente, um dos seus vectores $\vec{x}_k$ será combinação linear dos restantes (ver proposição 1.9)

$$\vec{x}_k = \sum_{i\neq k}\alpha_i\vec{x}_i$$

O resultado é agora consequência da alínea anterior, para o que basta somar $\sum_{i\neq k}(-\alpha_i)\vec{x}_i$ ao vector $\vec{x}_k$, obtendo-se

$$h\Big(\vec{x}_1, \ldots, \vec{x}_k, \ldots, \vec{x}_p\Big) = h\Big(\vec{x}_1, \ldots, \vec{x}_k + \sum_{i\neq k}(-\alpha_i)\vec{x}_i, \ldots, \vec{x}_p\Big) = h\Big(\vec{x}_1, \ldots, \vec{o}_E, \ldots, \vec{x}_p\Big) = \vec{o}_F$$

vii) Se $p > n$, qualquer sequência de p vectores $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ será linearmente dependente (ver proposição 1.16.*i*) e o resultado anunciado é consequência imediata da alínea anterior.

Suponha-se $p \leq n$ e seja $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ uma base arbitrária de E. Exprimamos os vectores $(\vec{x}_i)_{1\leq i\leq p}$ através das respectivas coordenadas $(x_{ki})_{\substack{1\leq k\leq n \\ 1\leq i\leq p}}$ na base e:

$$\vec{x}_i = \sum_{k_i=1}^{n} x_{k_i i}\vec{e}_{k_i};\ i = 1, 2, \ldots, p$$

Como h é alternada, na expressão (4.15), anulam-se os valores de $h\big(\vec{e}_{k_1}, \vec{e}_{k_2}, \ldots, \vec{e}_{k_p}\big)$ sempre que haja índices repetidos em $(k_1, k_2, \ldots, k_p)$:

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = \sum_{1 \leq k_1 \neq k_2 \neq \cdots \neq k_p \leq n}^{n} x_{k_1 1} x_{k_2 2} \cdots x_{k_p p}\, h\big(\vec{e}_{k_1}, \vec{e}_{k_2}, \ldots, \vec{e}_{k_p}\big) \tag{4.50}$$

o que reduz o número de parcelas em (4.15) de n^p para apenas A_p^n (arranjos de n, p a p)

$$\begin{aligned} n \times (n-1) \times \cdots \times (n-(p-1)) &= \frac{n \times (n-1) \times \cdots \times (n-p+1) \times (n-p)!}{(n-p)!} \\ &= \frac{n!}{(n-p)!} = A_p^n = \binom{n}{p} p! \end{aligned}$$

Estas parcelas podem, por sua vez, associar-se em $\binom{n}{p}$ grupos de $p!$ parcelas nas quais os índices $(k_1, k_2, \ldots, k_p)$ são os mesmos, mas por ordem diferente e será possível, em cada um destes $\binom{n}{p}$ grupos de $p!$ parcelas, trocar a ordem dos vectores em $h\big(\vec{e}_{k_1}, \vec{e}_{k_2}, \ldots, \vec{e}_{k_p}\big)$ para $h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big)$, onde $1 \leq i_1 < i_2 < \cdots < i_p \leq n$ (ordem crescente dos p índices).

Portanto, para cada sequência $(k_1, k_2, \ldots, k_p)$, existe $\omega \in \mathfrak{S}_p$ tal que $i = k \circ \omega$ e $1 \leq i_1 < i_2 < \cdots < i_p \leq n$. Multiplicando ambos os membros à direita por $\sigma = \omega^{-1} \in \mathfrak{S}_p$, obtém-se $k = i \circ \sigma$ e, nesta igualdade, quando σ varre o grupo simétrico $\mathfrak{S}_p$, obtemos todas as sequências $(k_1, k_2, \ldots, k_p)$ que têm os mesmos p índices, mas por diferentes ordens. Como h é anti-simétrica (por ser alternada), vem

$$h\big(\vec{e}_{k_1}, \vec{e}_{k_2}, \ldots, \vec{e}_{k_p}\big) = h\big(\vec{e}_{i_{\sigma_1}}, \vec{e}_{i_{\sigma_2}}, \ldots, \vec{e}_{i_{\sigma_p}}\big) = \varepsilon(\sigma) h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big);\ \sigma \in \mathfrak{S}_p$$

Associando em (4.50) as $p!$ parcelas contendo o mesmo valor $h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big)$, ficamos com $\binom{n}{p}$ parcelas do tipo

$$\underbrace{\Bigg(\sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) x_{i_{\sigma_1} 1} x_{i_{\sigma_2} 2} \cdots x_{i_{\sigma_p} p}\Bigg)}_{p!\ \text{parcelas}} h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big);\ 1 \leq i_1 < i_2 < \cdots < i_p \leq n$$

Constatamos, portanto, que o somatório (4.50), que é composto de um total de $A_p^n = \binom{n}{p} p!$ parcelas, fica decomposto em $\binom{n}{p}$ parcelas do tipo anterior:

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = \overbrace{\underbrace{\sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n}}_{\binom{n}{p}\ \text{parcelas}} \underbrace{\Bigg(\sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) x_{i_{\sigma_1} 1} x_{i_{\sigma_2} 2} \cdots x_{i_{\sigma_p} p}\Bigg)}_{p!\ \text{parcelas}} h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big)}^{A_p^n = \binom{n}{p} p!\ \text{parcelas}}$$

viii) Se for $p = n$, o somatório exterior na expressão anterior terá apenas uma parcela (observe que $\binom{n}{n} = 1$) e, além disso, será $i = (i_1, i_2, \ldots, i_n) = (1, 2, \ldots, n) = \mathscr{I}_n$, o que implica $i \circ \sigma = \mathscr{I}_n \circ \sigma = \sigma$. Assim, para $p = n$, obtém-se

$$h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = \underbrace{\left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma_1 1} x_{\sigma_2 2} \cdots x_{\sigma_n n} \right)}_{n! \text{ parcelas}} h(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$$

Portanto, conclui-se que uma função n-linear alternada h sobre um espaço vectorial E com dimensão n com valores em F fica totalmente determinada pelo seu valor $h(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ numa base $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ de E.

ix) A expressão (4.48) implica que, se a função n-linear alternada h sobre um espaço E de dimensão n for nula numa base $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ qualquer de E, então h será nula em todos os pontos de E^n (será $h = \mathscr{O}_{FE^n}$). Portanto, se $x = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ é linearmente independente e $h \neq \mathscr{O}_{FE^n}$, deverá ser $h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \neq \vec{o}_F$ (caso contrário, h anulava-se na base x e seria $h = \mathscr{O}_{FE^n}$). Por outro lado, a alínea *v)* mostra que, se $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ for linearmente dependente, então $h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = \vec{o}_F$. Portanto, $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ é linearmente dependente sse $h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = \vec{o}_F$. □

Observações:

- A expressão (4.47) revela um dado muito importante: *as funções p-lineares alternadas sobre um espaço de dimensão $n \geq p$ ficam completamente determinadas pelos seus valores nos $\binom{n}{p}$ pontos de E^p da forma $\left(\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{e}_{i_p}\right)$, com $1 \leq i_1 < i_2 < \cdots < i_p \leq n$, e em que $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ é uma qualquer base de E.*
- *Em particular, pela expressão* (4.48)*, uma função n-linear alternada definida num espaço n-dimensional fica totalmente determinada pelo seu valor numa base $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ de E.*

Exemplo 4.38 Para melhor se compreender a estrutura da expressão (4.48), vejamos o exemplo de uma função bilinear alternada definida num espaço bidimensional E $(p = n = 2)$ e com valores num espaço F. O somatório (4.48) terá $2! = 2$ parcelas. Se for $e = (\vec{e}_1, \vec{e}_2)$ uma base de E, os argumentos $\vec{x}_1, \vec{x}_2$ expressar-se-ão na base e por

$$\begin{aligned} \vec{x}_1 &= x_{11}\vec{e}_1 + x_{21}\vec{e}_2 \\ \vec{x}_2 &= x_{12}\vec{e}_1 + x_{22}\vec{e}_2 \end{aligned}$$

e $h(\vec{x}_1, \vec{x}_2)$ terá então a seguinte expressão:

$$h(\vec{x}_1, \vec{x}_2) = (x_{11}x_{22} - x_{21}x_{12}) h(\vec{e}_1, \vec{e}_2)$$

Deste modo, h fica completamente determinada pelo valor $h(\vec{e}_1, \vec{e}_2) \in F$.

Exemplo 4.39 Para novo exemplo sobre a estrutura da expressão (4.48), vejamos o caso de uma função trilinear alternada definida num espaço tridimensional E $(p = n = 3)$ e com valores num espaço F. O somatório (4.48) terá $3! = 6$ parcelas. Se for $e = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ uma base de E, os argumentos $\vec{x}_1, \vec{x}_2, \vec{x}_3$ expressar-se-ão na base e por

$$\begin{aligned}\vec{x}_1 &= x_{11}\vec{e}_1 + x_{21}\vec{e}_2 + x_{31}\vec{e}_3\\ \vec{x}_2 &= x_{12}\vec{e}_1 + x_{22}\vec{e}_2 + x_{32}\vec{e}_3\\ \vec{x}_3 &= x_{13}\vec{e}_1 + x_{23}\vec{e}_2 + x_{33}\vec{e}_3\end{aligned}$$

e $h(\vec{x}_1, \vec{x}_2, \vec{x}_3)$ terá então a seguinte expressão:

$$(x_{11}x_{22}x_{33} + x_{21}x_{32}x_{13} + x_{31}x_{12}x_{23} - x_{31}x_{22}x_{13} - x_{11}x_{32}x_{23} - x_{21}x_{12}x_{33})h(\vec{e}_1, \vec{e}_2, \vec{e}_3)$$

É evidente que h fica completamente determinada pelo valor $h(\vec{e}_1, \vec{e}_2, \vec{e}_3) \in F$.

Exemplo 4.40 Ilustremos agora o uso da expressão (4.47) mais complexa, através do exemplo de uma função bilinear alternada ($p = 2$) definida num espaço tridimensional E ($n = 3$) e com valores num espaço F. O somatório exterior terá $\binom{3}{2} = 3$ parcelas e o interior ficará com $2! = 2$ parcelas, num total de $3 \times 2 = 6 = A_2^3$ parcelas. Se for $e = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ uma base de E, os argumentos $\vec{x}_1, \vec{x}_2$ expressar-se-ão na base e por

$$\begin{aligned}\vec{x}_1 &= x_{11}\vec{e}_1 + x_{21}\vec{e}_2 + x_{31}\vec{e}_3\\ \vec{x}_2 &= x_{12}\vec{e}_1 + x_{22}\vec{e}_2 + x_{32}\vec{e}_3\end{aligned}$$

e, para $h(\vec{x}_1, \vec{x}_2)$, virá a seguinte expressão:

$$(x_{11}x_{22} - x_{21}x_{12})h(\vec{e}_1, \vec{e}_2) + (x_{11}x_{32} - x_{31}x_{12})h(\vec{e}_1, \vec{e}_3) + (x_{21}x_{32} - x_{31}x_{22})h(\vec{e}_2, \vec{e}_3)$$

Neste caso, h fica determinada pelos 3 valores $h(\vec{e}_1, \vec{e}_2), h(\vec{e}_1, \vec{e}_3), h(\vec{e}_2, \vec{e}_3) \in F$, ou seja $h(\vec{e}_i, \vec{e}_j)$, com $1 \leq i < j \leq 3$.

4.5 Determinante numa base

Vamos, nesta secção, tratar das *formas n-lineares alternadas* definidas num espaço E com dimensão $n \geq 1$.

Sejam, então, $\dim E = n$ e $h, h' \colon E^n \to \mathbb{K}$ duas formas n-lineares alternadas. Para qualquer sequência $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ de vectores de E e qualquer base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$, temos, usando (4.48),

$$\begin{aligned}h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) &= \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma_1 1} x_{\sigma_2 2} \cdots x_{\sigma_n n}\right) h(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)\\ h'(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) &= \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma_1 1} x_{\sigma_2 2} \cdots x_{\sigma_n n}\right) h'(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)\end{aligned} \tag{4.51}$$

onde os $(x_{ij})_{1 \leq i,j \leq n}$ são as coordenadas dos vectores $(\vec{x}_j)_{1 \leq j \leq n}$ em relação à base e e os valores $h(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ e $h'(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ são *escalares* de $\mathbb{K}$. Se for $h \neq \mathbb{O}_{\mathbb{K}^{E^n}}$, será necessariamente,

$$h(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) \neq 0$$

e da 1ª equação (4.51) tiramos

$$\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)x_{\sigma_1 1}x_{\sigma_2 2}\cdots x_{\sigma_n n}=\frac{h(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_n)}{h(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)}$$

Substituindo esta expressão na 2ª igualdade (4.51), vem

$$h'(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_n)=\frac{h'(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)}{h(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)}\,h(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_n) \tag{4.52}$$

Concluimos, assim, que a função h' é múltipla de h, já que $\frac{h'(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)}{h(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)}$ é constante (independente de $\vec{x}_1,\vec{x}_2,\dots,\vec{x}_n$). Portanto:

Todas as formas n-lineares alternadas definidas num espaço n-dimensional são múltiplas das que não forem identicamente nulas. Isto significa que o espaço $\mathscr{A}(E,\mathbb{K})$ das formas n-lineares alternadas sobre um espaço vectorial n-dimensional E tem dimensão 1 e que qualquer forma n-linear $h\neq\mathscr{O}_{\mathbb{K}E^n}$ constitui uma base de $\mathscr{A}(E,\mathbb{K})$.

$$\dim(\mathscr{A}(E,\mathbb{K}))=1 \tag{4.53}$$

Se for $e'=\left(\vec{e}'_1,\vec{e}'_2,\dots,\vec{e}'_n\right)$ uma outra base de E, a igualdade (4.52) dá, fazendo $(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_n)=\left(\vec{e}'_1,\vec{e}'_2,\dots,\vec{e}'_n\right)$,

$$\frac{h'\left(\vec{e}'_1,\vec{e}'_2,\dots,\vec{e}'_n\right)}{h\left(\vec{e}'_1,\vec{e}'_2,\dots,\vec{e}'_n\right)}=\frac{h'(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)}{h(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)}$$

Portanto, o quociente $\frac{h'}{h}$ não depende da base de E em que é calculado, mas somente das funções h e h'.

A igualdade (4.48) mostra que para definir uma forma n-linear alternada sobre um espaço vectorial n-dimensional E, basta dar o seu valor numa base e deste espaço. Assim, poremos a seguinte

Definição 4.7 – Determinante numa base. – *Seja E um espaço vectorial de dimensão $n\geq 1$ sobre um corpo $\mathbb{K}$ e $e=(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)$ uma base de E. Chama-se **determinante na base** e à forma n-linear alternada (única) sobre E, $\det_e\colon E^n\to\mathbb{K}$, cujo valor na base e é igual a* 1. *Portanto, por definição,*

$$\det_e(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)=1 \tag{4.54}$$

e, da expressão (4.48) *resulta*

$$\begin{aligned}\det_e(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_n)&=\left(\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)x_{\sigma_1 1}x_{\sigma_2 2}\cdots x_{\sigma_n n}\right)\underbrace{\det_e(\vec{e}_1,\vec{e}_2,\dots,\vec{e}_n)}_{=1}\\&=\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)x_{\sigma_1 1}x_{\sigma_2 2}\cdots x_{\sigma_n n}\end{aligned} \tag{4.55}$$

em que os x_{ki} são as coordenadas do vector $\vec{x}_i$ na base e.

A definição anterior dá origem a uma forma n-linear alternada sobre E para cada base e deste espaço e todas estas formas são distintas de $\mathbb{O}_{\mathbb{K}E^n}$ (visto que não se anulam na base e).

Para as compararmos, seja $e' = \left(\vec{e}_1', \vec{e}_2', \ldots, \vec{e}_n'\right)$ uma segunda base de E. A igualdade (4.52) dá, fazendo $h' = \det_{e'}$ e $h = \det_e$ e atendendo ainda a (4.54),

$$\begin{aligned}\det_{e'}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) &= \frac{\det_{e'}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)}{\underbrace{\det_e(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)}_{=1}} \det_e(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) \\ &= \det_{e'}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)\det_e(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)\end{aligned}$$

Portanto,

$$\det_{e'}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = \det_e(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)\det_{e'}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) \tag{4.56}$$

Da última igualdade resulta ainda, fazendo $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = \left(\vec{e}_1', \vec{e}_2', \ldots, \vec{e}_n'\right)$,

$$1 = \det_{e'}\left(\vec{e}_1', \vec{e}_2', \ldots, \vec{e}_n'\right) = \det_e\left(\vec{e}_1', \vec{e}_2', \ldots, \vec{e}_n'\right)\det_{e'}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$$

Consequentemente, para qualquer par de bases(e, e') de E, tem-se

$$\det_{e'}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) = \frac{1}{\det_e\left(\vec{e}_1', \vec{e}_2', \ldots, \vec{e}_n'\right)} \tag{4.57}$$

4.6 Determinante de um endomorfismo. Determinante de uma matriz

Considere-se um espaço vectorial E de dimensão $n \geq 1$ e um endomorfismo $u\colon E \to E$. Para qualquer forma n-linear alternada $h\colon E^n \to \mathbb{K}$ e qualquer base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de E, vimos no exemplo 4.28 que a função $g\colon E^n \to \mathbb{K}$ definida por

$$g(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = h(u(\vec{x}_1), u(\vec{x}_2), \ldots, u(\vec{x}_n))$$

é uma forma n-linear alternada sobre E. A equação (4.52) mostra então que, se $h \neq \mathbb{O}_{\mathbb{K}E^n}$,

$$g(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = \frac{g(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)}{h(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)} h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$$

ou seja,

$$h(u(\vec{x}_1), u(\vec{x}_2), \ldots, u(\vec{x}_n)) = \frac{h(u(\vec{e}_1), u(\vec{e}_2), \ldots, u(\vec{e}_n))}{h(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)} h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$$

O quociente $\frac{h(u(\vec{e}_1), u(\vec{e}_2), \ldots, u(\vec{e}_n))}{h(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)}$ não depende da base e nem da forma n-linear alternada $h \neq \mathbb{O}_{\mathbb{K}E^n}$, sendo apenas função do endomorfismo u. Chamar-lhe-emos o *determinante* de u:

Definição 4.8 – Determinante de um endomorfismo. – *Seja E um espaço vectorial de dimensão $n \geq 1$ e $u\colon E \to E$ um endomorfismo de E. O* ***determinante*** *de u é o escalar (único), designado por* $\det(u)$, *tal que, para qualquer forma n-linear alternada não nula $h\colon E^n \to \mathbb{K}$*

sobre E e qualquer sequência $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ de vectores de E, se tem

$$h(u(\vec{x}_1), u(\vec{x}_2), \dots, u(\vec{x}_n)) = \det(u) \cdot h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \tag{4.58}$$

Podemos, agora, provar algumas das propriedades do determinante de um endomorfismo na seguinte

Proposição 4.8 – Propriedades do determinante de um endomorfismo – *Seja E um espaço vectorial de dimensão n sobre um corpo $\mathbb{K}$. então:*

i) *Se for $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ uma base de E, então para qualquer endomorfismo $u\colon E \to E$ e qualquer sequência $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ de vectores de E, tem-se*

$$\det_e(u(\vec{x}_1), u(\vec{x}_2), \dots, u(\vec{x}_n)) = \det(u) \cdot \det_e(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \tag{4.59}$$

Em particular, se $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$, será

$$\det(u) = \det_e(u(\vec{e}_1), u(\vec{e}_2), \dots, u(\vec{e}_n)) \tag{4.60}$$

ii) *O determinante da função identidade $\mathscr{I}_E$ é igual a 1*

$$\det(\mathscr{I}_E) = 1$$

Mais geralmente, para qualquer escalar $k \in \mathbb{K}$, o determinante da homotetia de razão k, $k\mathscr{I}_E\colon \vec{x} \mapsto k\vec{x}$ é dado por

$$\det(k\mathscr{I}_E) = k^n \tag{4.61}$$

iii) *O determinante da composta de dois endomorfismos de E é igual ao produto dos seus determinantes; por outras palavras, para todo o par (u, v) de endomorfismos de E, tem-se*

$$\det(v \circ u) = \det(v)\det(u) \tag{4.62}$$

iv) *Um endomorfismo $u\colon E \to E$ é regular (invertível) se e só se for $\det(u) \neq 0$. Neste caso, tem-se*

$$\det\left(u^{-1}\right) = (\det(u))^{-1} \tag{4.63}$$

Demonstração:

i) Basta considerar $h = \det_e$ em (4.58). Fazendo, por fim,

$$(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$$

em (4.59) e atendendo a que $\det_e(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = 1$, obtém-se

$$\det(u) = \det_e(u(\vec{e}_1), u(\vec{e}_2), \dots, u(\vec{e}_n))$$

ii) Basta fazer $u = \mathscr{I}_E$ na igualdade (4.60):

$$\det(\mathscr{I}_E) = \det_e(\mathscr{I}_E(\vec{e}_1), \mathscr{I}_E(\vec{e}_2), \ldots, \mathscr{I}_E(\vec{e}_n)) = \det_e(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) = 1$$

Para qualquer escalar $k \in \mathbb{K}$, tem-se por (4.60) e devido à multilinearidade de $\det_e$,

$$\det(k\mathscr{I}_E) = \det_e(k\vec{e}_1, k\vec{e}_2, \ldots, k\vec{e}_n) = k^n \det_e(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) = k^n$$

iii) Basta aplicar (4.60) ao endomorfismo composto $v \circ u$ e atender a (4.59)

$$\begin{aligned} \det(v \circ u) &= \det_e(v(u(\vec{e}_1)), v(u(\vec{e}_2)), \ldots, v(u(\vec{e}_n))) \\ &= \det(v)\det_e(u(\vec{e}_1), u(\vec{e}_2), \ldots, u(\vec{e}_n)) = \det(v)\det(u) \end{aligned}$$

iv) Se u é regular, então o endomorfismo inverso u^{-1} satisfaz a condição

$$u \circ u^{-1} = \mathscr{I}_E$$

e as duas alíneas anteriores implicam

$$\det(u)\det\big(u^{-1}\big) = 1$$

o que prova que $\det(u) \neq 0$ e que

$$\det\big(u^{-1}\big) = (\det(u))^{-1}$$

Reciprocamente, se $\det(u) \neq 0$ e se $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ é uma base de E, o escalar

$$\det_e(u(\vec{e}_1), u(\vec{e}_2), \ldots, u(\vec{e}_n)) = \det(u) \neq 0$$

isto mostra que $(u(\vec{e}_1), u(\vec{e}_2), \ldots, u(\vec{e}_n))$ é *linearmente independente* e que a característica c(u) é igual a n. Portanto u é invertível. □

Vamos, agora, definir a noção de determinante de uma matriz quadrada sobre um corpo $\mathbb{K}$:

Definição 4.9 – Determinante de uma matriz. – *Seja* $A = [\,a_{ij}\,] \in \mathbb{K}^{n,n}$ *uma matriz quadrada de ordem* $n \geq 1$ *sobre um corpo* $\mathbb{K}$. *Chama-se* ***determinante da matriz*** A *e designa-se por* $\det(A)$(7) *o determinante dos vectores-coluna de* A *na base canónica de* $\mathbb{K}^n$. *Como as coordenadas dos vectores-coluna de* A *na base canónica de* $\mathbb{K}^n$ *coincidem com as próprias componentes que compõem esses vectores, da igualdade* (4.55) *resulta*

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)\, a_{\sigma_1 1} a_{\sigma_2 2} \cdots a_{\sigma_n n} \tag{4.64}$$

7 Também se utilizam as notações $|A|$ e $|a_{ij}|$ para designar o determinante da matriz A, embora estas se prestem a alguma confusão com a notação usada para o *módulo*.

Observações:

- Se $n = 1$, a igualdade anterior dá para valor do determinante de $A = [\, a_{11} \,]$ o próprio escalar a_{11}:

$$\det([\, a_{11} \,]) = a_{11}$$

- A soma (4.64) estende-se às $n!$ permutações do grupo simétrico de ordem n e a definição mostra que a forma n-linear det é um elemento de $\mathscr{A}_n(\mathbb{K}^n, \mathbb{K})$.

- Se $n \geq 2$ e pelo corolário 4.4.3, das $n!$ parcelas, metade têm $\varepsilon(\sigma) = +1$ tendo a outra metade $\varepsilon(\sigma) = -1$.

- A ordem da matriz A diz-se também a ordem do $\det(A)$ e cada produto $a_{\sigma_1 1} a_{\sigma_2 2} \cdots a_{\sigma_n n}$ diz-se um *termo* do determinante ou da matriz (associado à permutação σ) e a paridade de σ diz-se igualmente paridade do termo.

- O termo $a_{11} a_{22} \cdots a_{nn}$ associado à permutação identidade diz-se *termo principal* do determinante.

- Usando esta terminologia, podemos afirmar que *o determinante da matriz A é a soma dos termos pares do determinante com os termos ímpares (que só existem para $n \geq 2$) multiplicados por -1.*

O determinante de uma matriz $A = [\, a_{ij} \,]$ pode também escrever-se indicando todos os elementos de A:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Convém observar que existe uma grande diferença entre a expressão

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

(com parêntesis rectos) que designa uma matriz e é um elemento de $\mathbb{K}^{n,n}$ e a expressão

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

(com traços verticais) que designa um escalar do corpo $\mathbb{K}$ e que é o determinante da matriz anterior. Convém ainda observar que o determinante pode também ser encarado como uma função $\det \colon \mathbb{K}^{n,n} \to \mathbb{K}; A \mapsto \det A$ (ou seja, um elemento de $\mathscr{F}(\mathbb{K}^{n,n}, \mathbb{K})$).

A definição de determinante de uma matriz quadrada permite-nos uma escrita alternativa para algumas das expressões das secções anteriores:

- A igualdade (4.47) pode reescrever-se na forma seguinte

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p) = \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} \begin{vmatrix} x_{i_1 1} & x_{i_1 2} & \cdots & x_{i_1 p} \\ x_{i_2 1} & x_{i_2 2} & \cdots & x_{i_2 p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i_p 1} & x_{i_p 2} & \cdots & x_{i_p p} \end{vmatrix} h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big) \tag{4.47.1}$$

onde figuram os determinantes das matrizes de ordem p cujas colunas contêm as coordenadas $i_1, i_2, \ldots, i_p$ dos vectores $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$ em relação à base $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de E. Se considerarmos a matriz $X = [\, x_{ki} \,] \in \mathbb{K}^{n,p}$ cujas colunas contêm as n coordenadas na base e dos vectores $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p)$, trata-se de formar os determinantes de todas as submatrizes de ordem p extraídas de X (e que são em número de $\binom{n}{p}$).

- A igualdade (4.48) pode ser escrita da seguinte forma

$$h(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = \begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{vmatrix} h(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) \tag{4.48.1}$$

onde figura o determinante da matriz quadrada de ordem n cujas colunas contêm as n coordenadas dos vectores $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ em relação à base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$.

- A igualdade (4.55) pode reescrever-se na forma seguinte

$$\text{det}_e(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = \begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{vmatrix} \tag{4.55.1}$$

onde figura o determinante da matriz quadrada de ordem n cujas colunas contêm as n coordenadas dos vectores $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ em relação à base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$.

- Considerando duas bases $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ e $e' = \big(\vec{e}\,'_1, \vec{e}\,'_2, \ldots, \vec{e}\,'_n\big)$ de E e atendendo à igualdade anterior, concluimos que $\text{det}_{e'}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ será o determinante da matriz cujas colunas são as coordenadas dos vectores $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ em relação à base e'; mas esta matriz é precisamente a matriz $T_{e'e}$ de mudança da base e' para a base e (ver equação (2.77)):

$$\text{det}_{e'}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) = \text{det}(T_{e'e})$$

Portanto, a igualdade (4.56) pode também ser escrita na forma alternativa:

$$\text{det}_{e'}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n) = \text{det}_e(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)\text{det}(T_{e'e}) \tag{4.56.1}$$

- Se for E um espaço vectorial de dimensão n sobre um corpo $\mathbb{K}$ e $u\colon E \to E$ um endomorfismo de E, a equação (4.60) escreve-se, para qualquer base $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de E,

$$\det(u) = \det_e(u(\vec{e}_1), u(\vec{e}_2), \ldots, u(\vec{e}_n))$$

Mas $\det_e(u(\vec{e}_1), u(\vec{e}_2), \ldots, u(\vec{e}_n))$ é o determinante da matriz cujas colunas são as coordenadas dos vectores da sequência $(u(\vec{e}_1), u(\vec{e}_2), \ldots, u(\vec{e}_n))$ em relação à base e. Porém, vimos no capítulo 3 – equação (3.18) – que esta matriz é precisamente a representação matricial $M_e(u)$ de u em relação à base e. Portanto,

$$\det(u) = \det(M_e(u)) \tag{4.65}$$

ou seja, *o determinante de um endomorfismo é igual ao determinante de qualquer das suas representações matriciais*.

4.7 Propriedades algébricas dos determinantes

Nesta secção, vamos estudar as principais propriedades do determinante de uma matriz quadrada, as quais, na sua maioria, resultam imediatamente das propriedades das formas n-lineares alternadas sobre um espaço n-dimensional já estudadas antes.

Vejamos o que resulta da expressão (4.64) para $n = 2$ e $n = 3$:

- Para $n = 2$, a soma (4.64) tem $2! = 2$ parcelas que são:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

e daqui resulta a seguinte mnemónica simples para o cálculo de determinantes de 2ª ordem

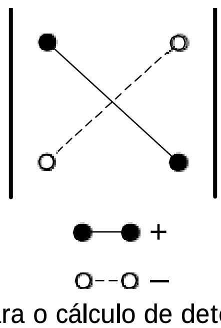


Fig. 4.6 – Mnemónica para o cálculo de determinantes de 2ª ordem.

- Para $n = 3$, a soma (4.64) tem $3! = 6$ parcelas que são:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}$$

e daqui resulta a seguinte mnemónica simples, chamada *regra de Sarrus*[8], para o cálculo de determinantes de 3ª ordem

---

8 *Sarrus, Pierre Frédérique*: matemático francês (Sainte Afrique 1798 – Sainte Afrique 1861).

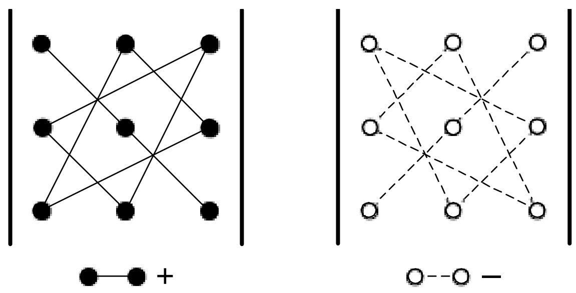


Fig. 4.7 – Regra de Sarrus, para o cálculo de determinantes de 3ª ordem.

De passagem, mencionemos duas variantes interessantes da regra de Sarrus:

- Na primeira, copiamos as duas primeiras colunas do determinante para a sua direita. Então, os termos pares obtêm-se a partir da diagonal principal e das duas diagonais paralelas à direita desta. Os termos ímpares obtêm-se a partir da diagonal secundária e das duas diagonais paralelas à direita desta, de acordo com a figura seguinte

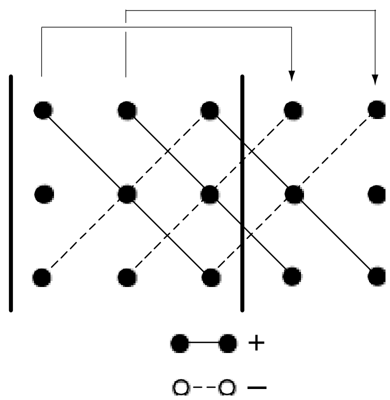


Fig. 4.8 – Uma variante da regra de Sarrus, para o cálculo de determinantes de 3ª ordem.

- Na segunda variante, copiamos as duas primeiras linhas do determinante para baixo deste. Então, os termos pares obtêm-se a partir da diagonal principal e das duas diagonais paralelas abaixo desta. Os termos ímpares obtêm-se a partir da diagonal secundária e das duas diagonais paralelas abaixo desta (ver a figura seguinte)

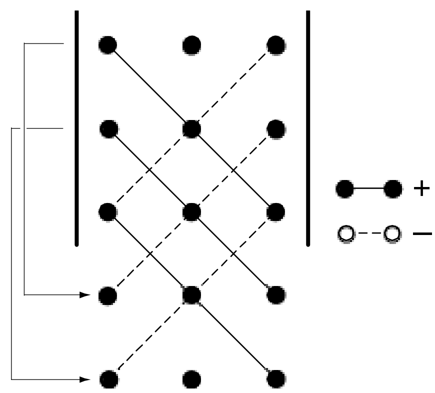


Fig. 4.9 – Outra variante da regra de Sarrus, para o cálculo de determinantes de 3ª ordem.

O uso de (4.64) para valores de n superiores a 3 torna-se impraticável, devido ao elevado número $n!$ de parcelas (para $n = 4$, seriam já 24 parcelas e, para $n = 10$, esse número elevar-se-ia a $10! = 3628800$): não existem, portanto, mnemónicas simples para $n > 3$ e o cálculo prático deste tipo de determinantes far-se-á por algoritmos que desenvolveremos mais à frente.

Nas páginas seguintes, vamos provar as mais importantes propriedades da função determinante de matriz, as quais são, na sua maioria, consequência imediata da matéria anterior respeitante às funções n-lineares alternadas em geral:

Proposição 4.9 *Para qualquer $n \geq 1$ e qualquer matriz $A \in \mathbb{K}^{n,n}$, tem-se*

$$\det\left(A^{\mathrm{T}}\right) = \det(A) \tag{4.66}$$

Demonstração:

Seja $B = A^{\mathrm{T}}$; então, será $b_{ij} = a_{ji}$. Se atendermos a (4.64), ter-se -á

$$\det(B) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) b_{\sigma_1 1} b_{\sigma_2 2} \cdots b_{\sigma_n n}$$

Seja ainda $\omega = \sigma^{-1}$ a permutação inversa de $\sigma \in \mathfrak{S}_n$ e alteremos a ordem dos factores $b_{\sigma_i i}$ em cada termo do determinante de B de modo a levar os referidos factores a ficarem ordenados por ordem crescente dos índices das linhas; é óbvio que as transposições que levam σ à permutação identidade transformarão esta na permutação ω, para a qual se tem em virtude de (4.14),

$$\varepsilon(\omega) = \varepsilon(\sigma)$$

Atendendo ainda a que a aplicação $\sigma \mapsto \omega$ de $\mathfrak{S}_n$ em si próprio é bijectiva e a que, portanto, quando σ percorre $\mathfrak{S}_n$, ω percorre também todo o $\mathfrak{S}_n$, vem

$$\begin{aligned} \det(B) &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) b_{\sigma_1 1} b_{\sigma_2 2} \cdots b_{\sigma_n n} = \sum_{\omega \in \mathfrak{S}_n} \varepsilon(\omega) b_{1\omega_1} b_{2\omega_2} \cdots b_{n\omega_n} \\ &= \sum_{\omega \in \mathfrak{S}_n} \varepsilon(\omega) a_{\omega_1 1} a_{\omega_2 2} \cdots a_{\omega_n n} = \det(A) \end{aligned}$$

o que termina a demonstração. □

Esta propriedade mostra que *o determinante de uma matriz é igual ao determinante dos seus* ***vectores-linha*** *na base canónica de* $\mathbb{K}^n$, que nos dá a seguinte expressão para o determinante de A, equivalente a (4.64),

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1\sigma_1} a_{2\sigma_2} \cdots a_{n\sigma_n} \tag{4.67}$$

Portanto, de ora em diante, escrevemos $\det(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ para designar o determinante da matriz cujas *colunas* são $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ ou cujas *linhas* são esses mesmos vectores (a que poderemos simplesmente chamar as *filas* da matriz).

Corolário 4.9.1 *Seja* $A \in \mathbb{C}^{n,n}$ *uma matriz complexa de ordem* n*. Então:*

i) $\det\big(\overline{A}\big) = \overline{\det(A)}$

ii) $\det\big(A^*\big) = \overline{\det(A)}$

iii) *Se* A *é hermitiana, então* $\det(A) \in \mathbb{R}$.

iv) *Se* A *é hemi-hermitiana, então* $\det(A) \in \mathbb{R}$ *ou* $\det(A) \in i\mathbb{R}$*, conforme* n *for par ou ímpar.*

Demonstração:

i) Resulta imediatamente da definição e das propriedades do operador conjugação em $\mathbb{C}$:

$$\det\big(\overline{A}\big) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)\overline{a}_{\sigma_1 1}\overline{a}_{\sigma_2 2}\cdots\overline{a}_{\sigma_n n} = \overline{\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)a_{\sigma_1 1}a_{\sigma_2 2}\cdots a_{\sigma_n n}} = \overline{\det(A)}$$

ii) Neste caso, tem-se

$$\det\big(A^*\big) = \det\Big(\big(\overline{A}\big)^{\mathrm{T}}\Big) = \det\big(\overline{A}\big) = \overline{\det(A)}$$

iii) Se A é *hermitiana*, tem-se

$$A^* = A \Rightarrow \det(A^*) = \det(A) \Rightarrow \overline{\det(A)} = \det(A) \Rightarrow \det(A) \in \mathbb{R}$$

iv) Se A é *hemi-hermitiana*, então se forem $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ as colunas de A,

$$\begin{aligned} A^* = -A &\Rightarrow \det(A^*) = \det(-A) = \det(-\vec{a}_1, -\vec{a}_2, \ldots, -\vec{a}_n) \\ &\Rightarrow \overline{\det(A)} = (-1)^n \det(A) \\ &\Rightarrow \begin{cases} n \text{ é par} & \Rightarrow \overline{\det(A)} = \det(A) \Rightarrow \det(A) \in \mathbb{R} \\ n \text{ é ímpar} & \Rightarrow \overline{\det(A)} = -\det(A) \Rightarrow \det(A) \in i\mathbb{R} \end{cases} \end{aligned}$$

o que termina a demonstração. □

A proposição seguinte resume algumas propriedades da função determinante:

Proposição 4.10 – Propriedades do determinante de uma matriz – *A função determinante (de uma matriz) goza das seguintes propriedades:*

i) $\det(I_n) = 1$ (4.68)

ii) *Operação elementar de tipo 1: para quaisquer* $1 \leq i < j \leq n$

$$\det(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots, \vec{x}_n) = -\det(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_j, \ldots, \vec{x}_i, \ldots, \vec{x}_n) \tag{4.69}$$

iii) *Multilinearidade: para qualquer* $1 \leq i \leq n$ *e* $\alpha \in \mathbb{K}$,

$$\begin{aligned} \text{det}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i + \vec{y}_i, \dots, \vec{x}_n) &= \text{det}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \dots, \vec{x}_n) + \text{det}(\vec{x}_1, \vec{x}_2, \dots, \vec{y}_i, \dots, \vec{x}_n) \\ \text{det}(\vec{x}_1, \vec{x}_2, \dots, \alpha\vec{x}_i, \dots, \vec{x}_n) &= \alpha\text{det}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \dots, \vec{x}_n) \end{aligned} \tag{4.70}$$

iv) *Para qualquer escalar* $k \in \mathbb{K}$ *e qualquer matriz* $A \in \mathbb{K}^{n,n}$,

$$\text{det}(kA) = k^n \text{det}(A) \tag{4.71}$$

Em particular,

$$\text{det}(-A) = (-1)^n \text{det}(A) \tag{4.72}$$

v) *Operação elementar de tipo 2: para qualquer* $1 \leq i \leq n$ *e* $\alpha \neq 0$,

$$\text{det}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \dots, \vec{x}_n) = \frac{1}{\alpha}\text{det}(\vec{x}_1, \vec{x}_2, \dots, \alpha\vec{x}_i, \dots, \vec{x}_n);\ \alpha \neq 0 \tag{4.73}$$

vi) *Operação elementar de tipo 3: para quaisquer* $1 \leq i \neq j \leq n$ *e sendo* $\beta \in \mathbb{K}$ *um escalar,*

$$\text{det}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_n) = \text{det}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i + \beta\vec{x}_j, \dots, \vec{x}_j, \dots, \vec{x}_n) \tag{4.74}$$

Mais geralmente, para qualquer $1 \leq i \leq n$ *e quaisquer escalares* β_k, *tem-se:*

$$\text{det}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \dots, \vec{x}_n) = \text{det}\Big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i + \sum_{k \neq i} \beta_k \vec{x}_k, \dots, \vec{x}_n\Big) \tag{4.75}$$

vii) *Anulamento do determinante de uma matriz:*

$$\text{det}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = 0 \Leftrightarrow (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) \textit{ é linearmente dependente} \tag{4.76}$$

Demonstração:

i) Da definição resulta, sendo $e = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ a base canónica de $\mathbb{K}^n$,

$$\text{det}(I_n) = \text{det}_e(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = 1$$

ii) É consequência de ser det uma forma n-linear alternada e, portanto, anti-simétrica.

iii) Resulta de ser det uma forma n-linear e, portanto, linear em cada um dos n vectores.

iv) É mais uma consequência da n-linearidade da função det:

$$\text{det}(kA) = \text{det}(k\vec{a}_1, k\vec{a}_2, \dots, k\vec{a}_n) = k^n \text{det}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n) = k^n \text{det}(A)$$

onde $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ são as colunas de A.Fazendo $k = -1$, obtém-se:

$$\text{det}(-A) = (-1)^n \text{det}(A)$$

v) Basta dividir por $\alpha \neq 0$ ambos os membros da segunda igualdade da alínea *ii* da presente proposição:

$$\det(\vec{x}_1, \vec{x}_2, \ldots, \alpha\vec{x}_i, \ldots, \vec{x}_n) = \alpha\det(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_n)$$

vi) É resultado de particularizar à função det a proposição 4.7.*iv*.

vii) Ver proposição 4.7.*ix*. □

Proposição 4.11 – Determinante do produto – *Sejam* $p, n \in \mathbb{N}$ *e* $A \in \mathbb{K}^{p,n}$ *e* $B \in \mathbb{K}^{n,p}$ *matrizes dos tipos* $p \times n$ *e* $n \times p$ *respectivamente. Então:*

i) *Se* $p > n$*, tem-se*

$$\det(AB) = 0 \tag{4.77}$$

ii) *Se* $p \leq n$*, então, usando as notações da secção 6 do capítulo 2 para designar submatrizes:*

$$\det(AB) = \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} \Big(\det(A[1, 2, \ldots, p; i_1, i_2, \ldots, i_p]) \det(B[i_1, i_2, \ldots, i_p; 1, 2, \ldots, p])\Big) \tag{4.78}$$

iii) *Se, em particular,* $p = n$ *então:*

$$\det(AB) = \det(A)\det(B) \tag{4.79}$$

Demonstração:

Sejam:

– $V \in \mathbb{K}^{p,p}$ a matriz produto de $A = [\, a_{ik} \,]$ por $B = [\, b_{kj} \,]$

$$V = AB$$

– $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ e $s = (\vec{s}_1, \vec{s}_2, \ldots, \vec{s}_p)$ as bases canónicas dos espaços vectoriais $\mathbb{K}^n$ e $\mathbb{K}^p$ respectivamente;

– $u\colon \mathbb{K}^n \to \mathbb{K}^p$ a função linear cuja matriz em relação ao par de bases canónicas (s, e) é A

$$M_{se}(u) = A \in \mathbb{K}^{p,n}$$

Fragmentemos B nas suas p colunas $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_p \in \mathbb{K}^n$ e calculemos o produto $V = AB$, por blocos. Obtemos, assim, as p colunas $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p \in \mathbb{K}^p$ da matriz V, pelas expressões:

$$\vec{v}_i = A\vec{b}_i;\ i = 1, 2, \ldots, p$$

Segundo o que vimos no capítulo 3 – equação (3.20) – isto equivale a dizer que as colunas de V são as imagens das colunas de B por meio de u:

$$\vec{v}_i = u\big(\vec{b}_i\big);\ i = 1, 2, \ldots, p$$

Tem-se, pois, por definição de determinante da matriz V,

$$\begin{aligned}\det(V) &= \det_s(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)\\ &= \det_s\Big(u\big(\vec{b}_1\big), u\big(\vec{b}_2\big), \dots, u\big(\vec{b}_p\big)\Big), \text{ onde } \vec{b}_k \in \mathbb{K}^n, \text{ para } k = 1, 2, \dots, p\end{aligned}$$

Se considerarmos, então, a forma p-linear alternada h sobre $\mathbb{K}^n$ definida por

$$h(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = \det_s\big(u(\vec{x}_1), u(\vec{x}_2), \dots, u(\vec{x}_p)\big),$$

poderemos escrever

$$\det(V) = h\big(\vec{b}_1, \vec{b}_2, \dots, \vec{b}_p\big)$$

Como se viu na proposição 4.7.*vii* tem-se:

i) Se $p > n$, será $h = \mathbb{O}_{\mathbb{K}E^p}$, donde

$$\det(V) = 0$$

ii) Para $p \leq n$, virá por (4.47.1),

$$\det(V) = h\Big(\vec{b}_1, \vec{b}_2, \dots, \vec{b}_p\Big) = \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} \underbrace{\begin{vmatrix} b_{i_1 1} & b_{i_1 2} & \cdots & b_{i_1 p} \\ b_{i_2 1} & b_{i_2 2} & \cdots & b_{i_2 p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{i_p 1} & b_{i_p 2} & \cdots & b_{i_p p} \end{vmatrix}}_{\det(B[i_1, i_2, \dots, i_p; 1, 2, \dots, p])} h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{e}_{i_p}\big) \tag{1}$$

Mas, da definição de h, obtém-se

$$h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{e}_{i_p}\big) = \det_s\big(u\big(\vec{e}_{i_1}\big), u\big(\vec{e}_{i_2}\big), \dots, u\big(\vec{e}_{i_p}\big)\big) \tag{2}$$

Porém, $\det_s\big(u\big(\vec{e}_{i_1}\big), u\big(\vec{e}_{i_2}\big), \dots, u\big(\vec{e}_{i_p}\big)\big)$ é igual ao determinante da matriz cujas colunas são as coordenadas em relação à base canónica s de $\mathbb{K}^p$ dos vectores $u\big(\vec{e}_{i_1}\big), u\big(\vec{e}_{i_2}\big), \dots, u\big(\vec{e}_{i_p}\big)$, ou seja, as colunas de índices $i_1, i_2, \dots, i_p$ da representação matricial $A = M_{se}(u)$ da função linear u:

$$\det_s\big(u\big(\vec{e}_{i_1}\big), u\big(\vec{e}_{i_2}\big), \dots, u\big(\vec{e}_{i_p}\big)\big) = \underbrace{\begin{vmatrix} a_{1i_1} & a_{1i_2} & \cdots & a_{1i_p} \\ a_{2i_1} & a_{2i_2} & \cdots & a_{2i_p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{pi_1} & a_{pi_2} & \cdots & a_{pi_p} \end{vmatrix}}_{\det(A[1,2,\dots,p; i_1, i_2, \dots, i_p])} \tag{3}$$

O resultado anunciado obtém-se, agora, substituindo (3) em (2) e, de seguida, em (1).

iii) A expressão (4.78) tem $\binom{n}{p}$ parcelas e, se $p = n$ terá apenas uma parcela, sendo

$$\begin{aligned}\det(A[1, 2, \dots, p; i_1, i_2, \dots, i_p]) &= \det(A)\\ \det(B[i_1, i_2, \dots, i_p; 1, 2, \dots, p]) &= \det(B)\end{aligned}$$

O resultado é imediato e confirma o já obtido em (4.62). □

Exemplo 4.41 Podemos verificar a expressão (4.78), usando as matrizes

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 5 & 2 & -4 \end{bmatrix} \text{ e } B = \begin{bmatrix} 3 & 0 \\ -2 & 5 \\ -1 & -4 \end{bmatrix}$$

Temos, calculando directamente o determinante de AB,

$$V = AB = \begin{bmatrix} 2 & -3 & 1 \\ 5 & 2 & -4 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ -2 & -5 \\ -1 & -4 \end{bmatrix} = \begin{bmatrix} 11 & 11 \\ 15 & 6 \end{bmatrix}$$

$$\det(V) = \begin{vmatrix} 11 & 11 \\ 15 & 6 \end{vmatrix} = 11 \begin{vmatrix} 1 & 1 \\ 15 & 6 \end{vmatrix} = 11(6-15) = -99$$

Por outro lado, usando (4.78), obtêm-se $\binom{3}{2} = 3$ parcelas

$$\begin{aligned} \det(V) &= \begin{vmatrix} 2 & -3 \\ 5 & 2 \end{vmatrix} \begin{vmatrix} 3 & 0 \\ -2 & -5 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ 5 & -4 \end{vmatrix} \begin{vmatrix} 3 & 0 \\ -1 & -4 \end{vmatrix} + \begin{vmatrix} -3 & 1 \\ 2 & -4 \end{vmatrix} \begin{vmatrix} -2 & -5 \\ -1 & -4 \end{vmatrix} \\ &= (4+15)(-15-0) + (-8-5)(-12-0) + (12-2)(8-5) \\ &= -19 \cdot 15 + 13 \cdot 12 + 10 \cdot 3 = -285 + 156 + 30 = -99 \end{aligned}$$

Corolário 4.11.1 *Seja $n \in \mathbb{N}$. Uma matriz $A \in \mathbb{K}^{n,n}$ é regular se e só se* $\det(A) \neq 0$ *e, neste caso,*

$$\det\left(A^{-1}\right) = (\det(A))^{-1} \tag{4.80}$$

Demonstração:

Já vimos que A tem inversa se e só se as suas linhas (e também as colunas) forem *linearmente independentes* e isto equivale à condição $\det(A) \neq 0$, por (4.76). Supondo satisfeita esta condição ter-se-á

$$AA^{-1} = I_n$$

Usando as equações (4.68) e (4.79), vem

$$\det(A)\det\left(A^{-1}\right) = 1$$

O resultado é, agora, imediato. □

A proposição seguinte dá conta de um importante resultado sobre determinantes e que será utilizado para provar o teorema de Laplace.

Proposição 4.12 – Determinante de uma matriz triangular – *Seja $r > 1$ um inteiro e A uma matriz triangular superior de submatrizes sobre um corpo $\mathbb{K}$*

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ O & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A_{rr} \end{bmatrix}$$

onde as matrizes $(A_{kk})_{1\leq k\leq r}$ *são quadradas de ordem* p_k *(sendo* A *de ordem* $n = \sum_{k=1}^{r} p_k$*). Nestas condições,*

$$\det(A) = \prod_{k=1}^{r} \det(A_{kk}) \tag{4.81}$$

Demonstração:

Bastará provar para $r = 2$, sendo o caso geral imediato, por indução em r.

Seja então

$$A = \begin{bmatrix} M & B \\ O_{n-p,p} & N \end{bmatrix} \in \mathbb{K}^{n,n}$$

uma matriz quadrada de ordem $n > 1$ e $M \in \mathbb{K}^{p,p}$ uma matriz quadrada de ordem $0 < p < n$. Claro está que N será quadrada de ordem $n - p > 0$ e $B \in \mathbb{K}^{p,n-p}$. Pretende-se mostrar que

$$\det(A) = \det(M)\det(N)$$

Seja $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ a base canónica de $\mathbb{K}^n$, $E = L(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_p)$ o subespaço de $\mathbb{K}^n$ gerado por $b = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_p)$ e $F = L(\vec{e}_{p+1}, \vec{e}_{p+2}, \ldots, \vec{e}_n)$ o subespaço de $\mathbb{K}^n$ gerado por $c = (\vec{e}_{p+1}, \vec{e}_{p+2}, \ldots, \vec{e}_n)$. É claro que E e F têm dimensões p e $n - p$, respectivamente, e que $\mathbb{K}^n$ é soma directa de E com F: $\mathbb{K}^n = E \oplus F$.

Sejam ainda:

- $u\colon \mathbb{K}^n \to \mathbb{K}^n;\ \vec{x} \mapsto A\vec{x}$ o endomorfismo de $\mathbb{K}^n$ representado por A em relação à base canónica $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ de $\mathbb{K}^n$.

- $v\colon E \to E;\ \vec{x} \mapsto M\vec{x}$ o endomorfismo de E representado por M em relação à base $b = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_p)$ de E.

- E, finalmente, $w\colon F \to F;\ \vec{x} \mapsto N\vec{x}$ o endomorfismo de F representado por N em relação à base $c = (\vec{e}_{p+1}, \vec{e}_{p+2}, \ldots, \vec{e}_n)$ de F.

Tem-se, portanto,

$$A = M_e(u),\ M = M_b(v) \text{ e } N = M_c(w)$$

Daqui se conclui que

$$\begin{aligned} \det(u) &= \det(M_e(u)) = \det(A) \\ \det(v) &= \det(M_b(v)) = \det(M) \\ \det(w) &= \det(M_c(w)) = \det(N) \end{aligned} \tag{1}$$

Dada a forma triangular da matriz $A = [\,a_{ij}\,]$ e as definições anteriores, podemos escrever, para $j = 1, 2, \dots, p$:

$$\begin{aligned} u(\vec{e}_j) &= \sum_{i=1}^{n} a_{ij}\vec{e}_i = \sum_{i=1}^{p} a_{ij}\vec{e}_i + \sum_{i=p+1}^{n} 0\vec{e}_i \\ &= v(\vec{e}_j);\ j = 1, 2, \dots, p \end{aligned} \tag{2}$$

Mas, para $j = p+1, p+2, \dots, n$,

$$\begin{aligned} u(\vec{e}_j) &= \sum_{i=1}^{n} a_{ij}\vec{e}_i = \sum_{i=1}^{p} a_{ij}\vec{e}_i + \sum_{i=p+1}^{n} a_{ij}\vec{e}_i \\ &= \underbrace{\sum_{i=1}^{p} a_{ij}\vec{e}_i}_{\in E} + w(\vec{e}_j) = ;\ j = p+1, p+2, \dots, n \end{aligned} \tag{3}$$

Da 1ª igualdade (1), obtém-se, atendendo a (2),

$$\begin{aligned} \det(A) &= \det(u) = \det_e(u(\vec{e}_1), u(\vec{e}_2), \dots, u(\vec{e}_n)) \\ &= \det_e(v(\vec{e}_1), v(\vec{e}_2), \dots, v(\vec{e}_p), u(\vec{e}_{p+1}), \dots, u(\vec{e}_n)) \end{aligned} \tag{4}$$

Considere-se, agora, a função g: $E^p \to \mathbb{K}$ definida por

$$g(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = \det_e(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p, u(\vec{e}_{p+1}), \dots, u(\vec{e}_n)) \tag{5}$$

A função g é uma forma p-linear alternada sobre E e, portanto, atendendo a (1)

$$g(v(\vec{e}_1), v(\vec{e}_2), \dots, v(\vec{e}_p)) = \det(v) g(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p) = \det(M) g(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p) \tag{6}$$

Substituindo $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p)$ por $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p)$ em (5), obtém-se

$$g(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p) = \det_e(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p, u(\vec{e}_{p+1}), \dots, u(\vec{e}_n))$$

mas, por ser $\det_e$ uma forma n-linear alternada, podemos atender a (3) e usar (4.75) para obter

$$\begin{aligned} g(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p) &= \det_e\left(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p, \sum_{i=1}^{p} a_{i,p+1}\vec{e}_i + w(\vec{e}_{p+1}), \dots, \sum_{i=1}^{p} a_{in}\vec{e}_i + w(\vec{e}_n)\right) \\ &= \det_e(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p, w(\vec{e}_{p+1}), \dots, w(\vec{e}_n)) \end{aligned} \tag{7}$$

Atendendo a (4), (5), (6) e (7), obtém-se:

$$\det(A) = g(v(\vec{e}_1), v(\vec{e}_2), \dots, v(\vec{e}_p)) = \det(M)\det_e(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p, w(\vec{e}_{p+1}), \dots, w(\vec{e}_n))$$

O resultado fica, portanto, demonstrado se pudermos provar que

$$\det_e(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p, w(\vec{e}_{p+1}), \dots, w(\vec{e}_n)) = \det(N) \tag{8}$$

Para tal, consideremos a função $h\colon F^{n-p} \to \mathbb{K}$, definida por

$$h(\vec{x}_{p+1}, \vec{x}_{p+2}, \ldots, \vec{x}_n) = \det{}_e(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_p, \vec{x}_{p+1}, \vec{x}_{p+2}, \ldots, \vec{x}_n)$$

A função h é uma forma $(n-p)$-linear alternada sobre F e, portanto, atendendo a (1),

$$\begin{aligned} h(w(\vec{e}_{p+1}), w(\vec{e}_{p+2}), \ldots, w(\vec{e}_n)) &= \det(w) h(\vec{e}_{p+1}, \vec{e}_{p+2}, \ldots, \vec{e}_n) \\ &= \det(N) h(\vec{e}_{p+1}, \vec{e}_{p+2}, \ldots, \vec{e}_n) \end{aligned}$$

Mas, da definição de h resulta

$$h(\vec{e}_{p+1}, \vec{e}_{p+2}, \ldots, \vec{e}_n) = \det{}_e(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_p, \vec{e}_{p+1}, \vec{e}_{p+2}, \ldots, \vec{e}_n) = 1$$

e portanto,

$$h(w(\vec{e}_{p+1}), w(\vec{e}_{p+2}), \ldots, w(\vec{e}_n)) = \det(N)$$

O primeiro membro de (8) vale então

$$\det{}_e(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_p, w(\vec{e}_{p+1}), \ldots, w(\vec{e}_n)) = h(w(\vec{e}_{p+1}), w(\vec{e}_{p+2}), \ldots, w(\vec{e}_n)) = \det(N)$$

como queríamos demonstrar. □

Corolário 4.12.1 *Seja $A = [\, a_{ij}\,] \in \mathbb{K}^{n,n}$ uma matriz triangular superior, isto é $a_{ij} = 0$ para $i > j$. Então*

$$\det(A) = \prod_{k=1}^{n} a_{kk} \tag{4.82}$$

Demonstração:

É consequência imediata da proposição anterior. Pode também verificar-se este resultado directamente: recordemos a expressão (4.64)

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma_1 1} a_{\sigma_2 2} \cdots a_{\sigma_n n}$$

Vamos ver que, das $n!$ parcelas deste somatório, *só uma* não é necessariamente nula:

Se $\sigma_1 > 1$, será $a_{\sigma_1 1} = 0$; Portanto, para que o termo associado a σ não seja nulo, tem que ser

$$\sigma_1 = 1$$

Quanto a σ_2, será $\sigma_2 \neq \sigma_1 = 1$; então, deverá ser $\sigma_2 \geq 2$ e, se $\sigma_2 > 2$, será $a_{\sigma_2 2} = 0$. Portanto, para que o termo associado a σ não seja nulo, tem que ser

$$\sigma_2 = 2$$

Continuando com este raciocínio, conclui-se que deverá ser $\sigma_k = k$, para todo o $1 \leq k \leq n$, ou seja, $\sigma = \mathscr{I}_n$. Assim, o único termo não necessariamente nulo é igual a $\prod\limits_{k=1}^{n} a_{kk}$ e que está associado à permutação identidade $\sigma = \mathscr{I}_n$ (termo principal). □

Observações:

- Usando a proposição 4.9, é fácil provar que a proposição 4.12 (e o seu corolário) são válidos igualmente para matrizes triangulares *inferiores*, visto que

$$\begin{vmatrix} A_{11} & O & \cdots & O \\ A_{21} & A_{22} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{vmatrix} = \det\left(\begin{bmatrix} A_{11} & O & \cdots & O \\ A_{21} & A_{22} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{bmatrix}^{\mathrm{T}}\right) = \begin{vmatrix} A_{11}^{\mathrm{T}} & A_{21}^{\mathrm{T}} & \cdots & A_{r1}^{\mathrm{T}} \\ O & A_{22}^{\mathrm{T}} & \cdots & A_{r2}^{\mathrm{T}} \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A_{rr}^{\mathrm{T}} \end{vmatrix}$$

$$= \prod_{k=1}^{r} \det\left(A_{kk}^{\mathrm{T}}\right) = \prod_{k=1}^{r} \det(A_{kk})$$

- Sejam $A \in \mathbb{K}^{p,p}$ e $B \in \mathbb{K}^{q,q}$ matrizes de ordens p e q sobre um corpo $\mathbb{K}$ e $U = \begin{bmatrix} A & C \\ D & B \end{bmatrix}$ (será obviamente $C \in \mathbb{K}^{p,q}$ e $D \in \mathbb{K}^{q,p}$). Então, supondo A invertível, podemos escrever

$$U = \begin{bmatrix} I_p & O_{p,q} \\ DA^{-1} & B - DA^{-1}C \end{bmatrix} \begin{bmatrix} A & C \\ O_{q,p} & I_q \end{bmatrix}$$

Então, atendendo à proposição 4.12, teremos

$$\det U = \det A \det\left(B - DA^{-1}C\right), \text{ com } \det A \neq 0$$

Do mesmo modo, se for B invertível, podemos escrever

$$U = \begin{bmatrix} A - CB^{-1}D & CB^{-1} \\ O_{q,p} & I_q \end{bmatrix} \begin{bmatrix} I_p & O_{p,q} \\ D & B \end{bmatrix}$$

Da proposição 4.12 resulta agora

$$\det U = \det B \det\left(A - CB^{-1}D\right), \text{ com } \det B \neq 0$$

4.8 Algoritmo de condensação para o cálculo de determinantes

Com base no que sabemos sobre o efeito das operações elementares sobre o valor do determinante de uma matriz e de ser fácil o cálculo do determinante de uma matriz triangular, podemos agora usar o algoritmo de condensação para levar o determinante à forma triangular superior, após o que o resultado é imediato, usando a igualdade (4.82). É este o primeiro dos métodos gerais para o cálculo de determinantes (em particular, de ordem superior a 3). Note-se que, neste caso, podem fazer-se operações elementares com linhas e também com colunas ao longo do procedimento, se isso nos for conveniente.

O quadro seguinte ilustra, em pseudo-código, o algoritmo de condensação vertical, adaptado ao cálculo de determinantes: o input é a matriz quadrada `a` de ordem `n` e o output é o seu determinante `d`.

```
% a é matriz quadrada de ordem n
n=Ordem(a); r=0; d=1;
Enquanto r < n
  % Busca do Pivot:
  i=r;
  Fazer
    i=i+1
  Enquanto a_{i,r+1} = 0 ∧ i < n;
  Se a_{i,r+1} ≠ 0 então                % a_{i,r+1} é pivot?
    r = r + 1;
    % Redução da coluna r de a:
    Se i ≠ r então
      Linha_i ↔ Linha_r;                % Operação de tipo 1
      d=-d                              % Anti-simetria
    FimSe;
    pivot=a_rr;
    d=pivot*d;                          % Multilinearidade
    Linha_r=Linha_r/pivot;              % Operação de tipo 2
    Para i=r+1 até n
      Linha_i=Linha_i-a_ir*Linha_r      % Operação de tipo 3
    FimPara
  Senão
    d=0;
    Devolver d
  FimSe
FimEnquanto;
% Saída do resultado:
  Saída: d
```

Na secção 4.16, apresenta-se a implementação do algoritmo anterior utilizando o software MATHEMATICA©, constituindo a função **`CondensacaoDet`** definida no *package* intitulado **``ALGA`Determinantes` ``**.

Exemplo 4.42 A título de exemplo, calculemos o determinante da seguinte matriz A de 5ª ordem, pelo método de condensação vertical:

$$\det A = \begin{vmatrix} -1 & 2 & 0 & -2 & 3 \\ 4 & 3 & 1 & -2 & 2 \\ -3 & 0 & 2 & 4 & 1 \\ 2 & -5 & 1 & 0 & 3 \\ 4 & 1 & -1 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 0 & 2 & -3 \\ 4 & 3 & 1 & -2 & 2 \\ 3 & 0 & -2 & -4 & -1 \\ 2 & -5 & 1 & 0 & 3 \\ 4 & 1 & -1 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 0 & 2 & -3 \\ 0 & 11 & 1 & -10 & 14 \\ 0 & 6 & -2 & -10 & 8 \\ 0 & -1 & 1 & -4 & 9 \\ 0 & 9 & -1 & -6 & 12 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & -2 & 0 & 2 & -3 \\ 0 & 1 & -1 & 4 & -9 \\ 0 & 6 & -2 & -10 & 8 \\ 0 & 11 & 1 & -10 & 14 \\ 0 & 9 & -1 & -6 & 12 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 0 & 2 & -3 \\ 0 & 1 & -1 & 4 & -9 \\ 0 & 0 & 4 & -34 & 62 \\ 0 & 0 & 12 & -54 & 113 \\ 0 & 0 & 8 & -42 & 93 \end{vmatrix} = 2\begin{vmatrix} 1 & -2 & 0 & 2 & -3 \\ 0 & 1 & -1 & 4 & -9 \\ 0 & 0 & 2 & -17 & 31 \\ 0 & 0 & 12 & -54 & 113 \\ 0 & 0 & 8 & -42 & 93 \end{vmatrix}$$

$$= 2\begin{vmatrix} 1 & -2 & 0 & 2 & -3 \\ 0 & 1 & -1 & 4 & -9 \\ 0 & 0 & 2 & -17 & 31 \\ 0 & 0 & 0 & 48 & -73 \\ 0 & 0 & 0 & 26 & -31 \end{vmatrix} = \frac{2}{24}\begin{vmatrix} 1 & -2 & 0 & 2 & -3 \\ 0 & 1 & -1 & 4 & -9 \\ 0 & 0 & 2 & -17 & 31 \\ 0 & 0 & 0 & 48 & -73 \\ 0 & 0 & 0 & 0 & 205 \end{vmatrix}$$

$$= \frac{1}{12}(1 \cdot 1 \cdot 2 \cdot 48 \cdot 205) = 1640$$

4.9 Teorema de Laplace

Na presente secção, vamos enunciar um teorema que nos vai conduzir a um novo método geral (recursivo) para o cálculo de determinantes: trata-se do teorema de Laplace[9]:

Proposição 4.13 – Teorema de Laplace (geral) – *Sejam $n > 1$ e $0 < p < n$ números inteiros e $A \in \mathbb{K}^{n,n}$ uma matriz quadrada de ordem n sobre um corpo $\mathbb{K}$. Para quaisquer p colunas de A, com índices $1 \leq j_1 < j_2 < \cdots < j_p \leq n$ e usando as notações da secção 6 do capítulo 2 para as submatrizes de A, o determinante de A pode ser calculado por*

$$\sum_{1\leq i_1<\cdots<i_p\leq n} (-1)^{\sum\limits_{k=1}^{p}(i_k+j_k)} \det(A[i_1,\ldots,i_p;j_1,\ldots,j_p])\det(A(i_1,\ldots,i_p;j_1,\ldots,j_p)) \qquad (4.83.1)$$

O segundo membro da igualdade anterior é chamado o ***desenvolvimento de Laplace de*** $\det(A)$ ***segundo as colunas*** $j_1, j_2, \ldots, j_p$. *Existem $\binom{n}{p}$ desenvolvimentos de Laplace para o determinante de A segundo p colunas (correspondentes a outras tantas sequências estritamente crescentes $j_1 < j_2 < \cdots < j_p$ de índices das colunas) tendo o somatório na expressão* (4.83.1) *um número de parcelas também igual a $\binom{n}{p}$.*

Demonstração:

Seja $e = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$ a base canónica de $\mathbb{K}^n$ e $(\vec{a}_k)_{1\leq k\leq n}$ as colunas de $A = [\,a_{ij}\,] \in \mathbb{K}^{n,n}$. Fixemos p colunas em A, de índices $1 \leq j_1 < j_2 < \cdots < j_p \leq n$ e defina-se uma função $h\colon (\mathbb{K}^n)^p \to \mathbb{K}$ mediante a igualdade

$$h\big(\vec{x}_{j_1}, \vec{x}_{j_2}, \ldots, \vec{x}_{j_p}\big) = \det_e\big(\vec{a}_1, \ldots, \vec{a}_{j_1-1}, \vec{x}_{j_1}, \vec{a}_{j_1+1}, \ldots, \vec{a}_{j_2-1}, \vec{x}_{j_2}, \vec{a}_{j_2+1}, \ldots, \vec{a}_{j_p-1}, \vec{x}_{j_p}, \vec{a}_{j_p+1}, \ldots, \vec{a}_n\big)$$

A função h é uma forma p-linear alternada sobre $\mathbb{K}^n$ e a definição de h mostra que

$$\det(A) = \det_e(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) = h\big(\vec{a}_{j_1}, \vec{a}_{j_2}, \ldots, \vec{a}_{j_p}\big)$$

Atendendo a $(4.47.1)$, obtém-se, fazendo uso das notações introduzidas na secção 6 do capítulo 2 para a designação das submatrizes de A,

$$\begin{aligned}\det(A) &= \sum_{1\leq i_1<i_2<\cdots<i_p\leq n} \begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_p} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_pj_1} & a_{i_pj_2} & \cdots & a_{i_pj_p} \end{vmatrix} h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big) \\ &= \sum_{1\leq i_1<i_2<\cdots<i_p\leq n} \det(A[i_1,\ldots,i_p;j_1,\ldots,j_p])\, h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big) \qquad (1)\end{aligned}$$

Recorrendo à definição da forma p-linear alternada h, tem-se:

$$h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big) = \det_e\big(\vec{a}_1, \ldots, \vec{a}_{j_1-1}, \vec{e}_{i_1}, \vec{a}_{j_1+1}, \ldots, \vec{a}_{j_2-1}, \vec{e}_{i_2}, \vec{a}_{j_2+1}, \ldots, \vec{a}_{j_p-1}, \vec{e}_{i_p}, \vec{a}_{j_p+1}, \ldots, \vec{a}_n\big)$$

---

9 *Laplace, Pierre-Simon, marquis de*: astrónomo, matemático e físico francês (Beaumont-en-Auge 1749 – Paris 1827).

Transpondo sucessivamente $\vec{e}_{i_1}$ (que está na coluna j_1) com todos os vectores anteriores, leva-se aquele vector para a coluna 1 (após $j_1 - 1$ transposições elementares); repetindo o processo para $\vec{e}_{i_2}$, podemos deslocar este para a coluna 2, após $j_2 - 2$ transposições e assim sucessivamente até levar o vector $\vec{e}_{i_p}$ para a coluna p (executando mais $j_p - p$ transposições elementares). Terminado este processo e atendendo à anti-simetria da forma n-linear alternada $\det_e$, levamos os vectores $\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}$ para as p primeiras colunas, *sem alterar a ordem dos restantes* $n-p$ vectores $\vec{a}_k$; fazendo $s_j = \sum_{k=1}^{p} j_k - \sum_{k=1}^{p} k$, obtém-se:

$$h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big) =$$
$$(-1)^{s_j}\det_e\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}, \vec{a}_1, \ldots, \vec{a}_{j_1-1}, \vec{a}_{j_1+1}, \ldots, \vec{a}_{j_2-1}, \vec{a}_{j_2+1}, \ldots, \vec{a}_{j_p-1}, \vec{a}_{j_p+1}, \ldots, \vec{a}_n\big)$$

Porém, $\det_e$ é também função n-linear alternada (logo, anti-simétrica) das *linhas* da matriz e, como $i_1 < i_2 < \cdots < i_p$, podemos praticar com as linhas $i_1, i_2, \cdots, i_p$ um processo de transposições elementares sucessivas semelhante ao usado para as colunas $j_1, j_2, \cdots, j_p$ e que leve aquelas linhas (que contêm as componentes iguais a 1 nos vectores $\vec{e}_{i_k}$) a ocupar as p primeiras posições; O número total de transposições a realizar será, como antes, de $s_i = \sum_{k=1}^{p} i_k - \sum_{k=1}^{p} k$. Usando as notações da secção 6 do capítulo 2 para designar as submatrizes de A e após este procedimento, somos conduzidos, usando (4.81), às igualdades:

$$\begin{aligned}
h\big(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_p}\big) &= (-1)^{s_i}(-1)^{s_j}\begin{vmatrix} I_p & A[i_1, \ldots, i_p; j_1, \ldots, j_p) \\ O_{n-p,p} & A(i_1, \ldots, i_p; j_1, \ldots, j_p) \end{vmatrix} \\
&= (-1)^{s_i+s_j}\begin{vmatrix} I_p & A[i_1, \ldots, i_p; j_1, \ldots, j_p) \\ O_{n-p,p} & A(i_1, \ldots, i_p; j_1, \ldots, j_p) \end{vmatrix} \\
&= (-1)^{s_i+s_j}\overbrace{\det(I_p)}^{=1}\det(A(i_1, \ldots, i_p; j_1, \ldots, j_p)) \\
&= (-1)^{s_i+s_j}\det(A(i_1, \ldots, i_p; j_1, \ldots, j_p)) \qquad (2)
\end{aligned}$$

Mas

$$\begin{aligned}
s_i + s_j &= \left(\sum_{k=1}^{p} i_k - \sum_{k=1}^{p} k\right) + \left(\sum_{k=1}^{p} j_k - \sum_{k=1}^{p} k\right) \\
&= \sum_{k=1}^{p}(i_k + j_k) - 2\sum_{k=1}^{p} k = \sum_{k=1}^{p}(i_k + j_k) - p(p+1)
\end{aligned}$$

donde:

$$(-1)^{s_i+s_j} = (-1)^{\sum_{k=1}^{p}(i_k+j_k) - p(p+1)} = \frac{(-1)^{\sum_{k=1}^{p}(i_k+j_k)}}{\underbrace{(-1)^{\overbrace{p(p+1)}^{\text{par}}}}_{=1}} = (-1)^{\sum_{k=1}^{p}(i_k+j_k)}$$

Substituindo o valor obtido para $(-1)^{s_i+s_j}$ na expressão (2) de $h\big(\vec{e}_{i_1},\vec{e}_{i_2},\dots,\vec{e}_{i_p}\big)$, obtemos

$$h\big(\vec{e}_{i_1},\vec{e}_{i_2},\dots,\vec{e}_{i_p}\big) = (-1)^{\sum_{k=1}^{p}(i_k+j_k)}\det(A(i_1,\dots,i_p;j_1,\dots,j_p))$$

Substituindo esta igualdade em (1), obtém-se finalmente (4.83.1). $\square$

Observações:

- Considerando que p pode tomar valores entre 1 e $n-1$ e usando o binómio de Newton, o número de formas de aplicar o teorema de Laplace a quaisquer colunas de A será

$$\sum_{p=1}^{n-1}\binom{n}{p} = \sum_{p=0}^{n}\binom{n}{p} - 2 = (1+1)^n - 2 = 2^n - 2 = 2\big(2^{n-1}-1\big)$$

Considerando ainda que o desenvolvimento de Laplace por determinadas p colunas coincide com o desenvolvimento pelas $n-p$ colunas complementares (visto que $\binom{n}{p}=\binom{n}{n-p}$ e que menores complementares têm a mesma paridade), o número de desenvolvimentos de Laplace *distintos* por colunas de A é, de facto, metade do total calculado acima:

$$\text{Número de desenvolvimentos por colunas} = 2^{n-1}-1 \tag{4.84}$$

- Por ser $\det(A^{\mathrm{T}}) = \det(A)$, é também válida uma fórmula semelhante a (4.83.1), mas dizendo respeito a p linhas de A: é o *desenvolvimento de Laplace de* $\det(A)$ *segundo as linhas* $i_1,i_2,\dots,i_p$; portanto, fixando quaisquer p índices de linhas $1\le i_1<i_2<\cdots<i_p\le n$, o determinante de $A\in\mathbb{K}^{n,n}$ é dado pela expressão:

$$\sum_{1\le j_1<\cdots<j_p\le n}(-1)^{\sum_{k=1}^{p}(i_k+j_k)}\det(A[i_1,\dots,i_p;j_1,\dots,j_p])\det(A(i_1,\dots,i_p;j_1,\dots,j_p)) \tag{4.83.2}$$

Existem $\binom{n}{p}$ desenvolvimentos de Laplace para o determinante de A segundo p linhas (correspondentes às sequências estritamente crescentes $i_1<i_2<\cdots<i_p$ de índices das linhas) tendo o somatório na expressão (4.83.2) um número de parcelas também igual a $\binom{n}{p}$; Quanto ao número de desenvolvimentos de Laplace *distintos* por linhas, ele é igualmente de $2^{n-1}-1$.

- A expressão $\det(A[i_1,\dots,i_p;j_1,\dots,j_p])$ que é o determinante da submatriz (de ordem p) de A formada pelas linhas de índices $i_1,\dots,i_p$ e pela colunas de índices $j_1,\dots,j_p$ é chamada o *menor* de ordem p formado por aquelas filas de A e o valor $u=\sum_{k=1}^{p}(i_k+j_k)$ define a *paridade* do menor (o menor diz-se *par* ou *ímpar* consoante aquela soma u for par ou ímpar).

- Se $(i_1,\dots,i_p)=(j_1,\dots,j_p)$, o menor $\det(A[i_1,\dots,i_p;j_1,\dots,j_p])$ diz-se *principal*. É formado por elementos de A pertencentes a linhas e colunas com iguais índices (o que significa que a diagonal principal de $A[i_1,\dots,i_p;j_1,\dots,j_p]$ é uma parte da diagonal principal de A).

Existem $\binom{n}{p}$ menores principais de ordem p num determinante de ordem n e é óbvio que todo o menor principal é *par*, visto que $u = \sum_{k=1}^{p}(i_k + i_k) = 2\sum_{k=1}^{p} i_k$ é par.

- O valor $\det(A(i_1, \ldots, i_p; j_1, \ldots, j_p))$ que é o determinante da submatriz de A (de ordem $n - p$) obtida eliminando as linhas de índices $i_1, \ldots, i_p$ e as colunas de índices $j_1, \ldots, j_p$ é chamada o *menor complementar* do menor $\det(A[i_1, \ldots, i_p; j_1, \ldots, j_p])$. Um menor e o seu menor complementar *têm necessariamente a mesma paridade*, visto que a soma das paridades dos dois é $2(1 + 2 + \cdots + n) = n(n + 1)$ que é sempre um número par (se as paridades fossem distintas, a soma referida seria ímpar!). O complementar de um menor principal de ordem p é outro menor principal (de ordem $n - p$).

- Chama-se *cofactor* ou *complemento algébrico* de um menor $\det(A[i_1, \ldots, i_p; j_1, \ldots, j_p])$ ao produto de $(-1)^u$ pelo respectivo menor complementar $\det(A(i_1, \ldots, i_p; j_1, \ldots, j_p))$; usaremos a notação $\mathrm{cof}(\det(A[i_1, \ldots, i_p; j_1, \ldots, j_p]))$

$$\mathrm{cof}(\det(A[i_1, \ldots, i_p; j_1, \ldots, j_p])) = (-1)^{\sum_{k=1}^{p}(i_k + j_k)} \det(A(i_1, \ldots, i_p; j_1, \ldots, j_p)) \tag{4.85.1}$$

Um menor de ordem $p = 1$ será $\det(A[i; j])$ e é obviamente igual ao elemento a_{ij} da linha i e da coluna j de A; a sua paridade é dada por $i + j$ e o menor complementar $\det(A(i; j))$ é de ordem $n - 1$. O cofactor de a_{ij} será

$$\mathrm{cof}(a_{ij}) = (-1)^{i+j} \det(A(i; j)) \tag{4.85.2}$$

em que $A(i; j)$ designa a submatriz de ordem $n - 1$ obtida de A por eliminação da linha i e da coluna j.

O valor da paridade $(-1)^{i+j}$ é exclusivamente função da posição (i, j) do elemento a_{ij} de A e o cálculo deste valor pode ser feito facilmente usando a mnemónica ilustrada na figura seguinte

$$\begin{vmatrix} + & - & + & - & \cdots & * \\ - & + & - & + & \cdots & * \\ + & - & + & - & \cdots & * \\ - & + & - & + & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & * & \cdots & + \end{vmatrix}$$

Fig. 4.10 – Mnemónica para determinar a paridade dos menores de 1ª ordem.

- Usando a terminologia que acabámos de introduzir, podemos enunciar o teorema de Laplace (nas suas duas variantes) do seguinte modo:

O determinante de uma matriz de ordem n é igual à soma dos produtos dos $\binom{n}{p}$ menores de ordem p contidos em p filas paralelas (linhas ou colunas) pelos respectivos cofactores.

Este enunciado traduz as duas fórmulas seguintes (desenvolvimento de Laplace do determinante de A por p colunas ou por p linhas, respectivamente):

$$\sum_{1\leq i_1<\cdots<i_p\leq n} \det(A[i_1,\ldots,i_p;j_1,\ldots,j_p])\,\mathrm{cof}(\det(A[i_1,\ldots,i_p;j_1,\ldots,j_p])); 1\leq j_1<j_2<\cdots<j_p\leq n$$

$$\sum_{1\leq j_1<\cdots<j_p\leq n} \det(A[i_1,\ldots,i_p;j_1,\ldots,j_p])\;\mathrm{cof}(\det(A[i_1,\ldots,i_p;j_1,\ldots,j_p])); 1\leq i_1<i_2<\cdots<i_p\leq n$$

▪ A função **`LaplaceDet`** do *package* **``ALGA`Determinantes` ``** implementada em linguagem MATHEMATICA© na secção 4.16 usa as fórmulas de Laplace para calcular um determinante de ordem arbitrária.

Exemplo 4.43 Determinemos, de novo, o valor do determinante da matriz do exemplo 4.42, através do teorema de Laplace, escolhendo $p=2$ e desenvolvendo pelos $\binom{5}{2}=\frac{5!}{2!3!}=10$ menores de 2ª ordem contidos nas colunas $j_1=1$ e $j_2=3$, por exemplo:

$$\det(A)=\begin{vmatrix}-1&2&0&-2&3\\4&3&1&-2&2\\-3&0&2&4&1\\2&-5&1&0&3\\4&1&-1&2&0\end{vmatrix}=$$

$$\begin{aligned}
&=-\begin{vmatrix}-1&0\\4&1\end{vmatrix}\begin{vmatrix}0&4&1\\-5&0&3\\1&2&0\end{vmatrix}+\begin{vmatrix}-1&0\\-3&2\end{vmatrix}\begin{vmatrix}3&-2&2\\-5&0&3\\1&2&0\end{vmatrix}\\
&\quad-\begin{vmatrix}-1&0\\2&1\end{vmatrix}\begin{vmatrix}3&-2&2\\0&4&1\\1&2&0\end{vmatrix}+\begin{vmatrix}-1&0\\4&-1\end{vmatrix}\begin{vmatrix}3&-2&2\\0&4&1\\-5&0&3\end{vmatrix}\\
&\quad-\begin{vmatrix}4&1\\-3&2\end{vmatrix}\begin{vmatrix}2&-2&3\\-5&0&3\\1&2&0\end{vmatrix}+\begin{vmatrix}4&1\\2&1\end{vmatrix}\begin{vmatrix}2&-2&3\\0&4&1\\1&2&0\end{vmatrix}-\begin{vmatrix}4&1\\4&-1\end{vmatrix}\begin{vmatrix}2&-2&3\\0&4&1\\-5&0&3\end{vmatrix}\\
&\quad-\begin{vmatrix}-3&2\\2&1\end{vmatrix}\begin{vmatrix}2&-2&3\\3&-2&2\\1&2&0\end{vmatrix}+\begin{vmatrix}-3&2\\4&-1\end{vmatrix}\begin{vmatrix}2&-2&3\\3&-2&2\\-5&0&3\end{vmatrix}-\begin{vmatrix}2&1\\4&-1\end{vmatrix}\begin{vmatrix}2&-2&3\\3&-2&2\\0&4&1\end{vmatrix}\\
&=-(-1)\cdot 2+(-2)(-44)-(-1)(-16)+1\cdot(86)-11\cdot(-48)+2\cdot(-18)-(-8)\cdot 94-(-7)\cdot 12\\
&\quad+(-5)(-4)-(-6)\cdot 22=2+88-16+86+528-36+752+84+20+132=1640
\end{aligned}$$

Vamos mencionar, em seguida, um caso particular do teorema de Laplace, por vezes chamado *teorema de Laplace restrito* e que se obtém imediatamente do teorema geral, fazendo aí $p=1$:

Corolário 4.13.1 – Teorema de Laplace restrito – *Sejam $n>1$ e $A\in\mathbb{K}^{n,n}$ uma matriz quadrada de ordem n sobre um corpo $\mathbb{K}$. Então:*

i) *Para qualquer $1 \leq j \leq n$, o determinante de A pode ser calculado por:*

$$\det(A) = \sum_{i=1}^{n} a_{ij} \operatorname{cof}(a_{ij}); 1 \leq j \leq n \tag{4.86.1}$$

O 2º membro da expressão anterior é conhecido por ***desenvolvimento de Laplace do determinante de*** A ***pela coluna*** j.

ii) *Para qualquer $1 \leq i \leq n$, o determinante de A é também dado por:*

$$\det(A) = \sum_{j=1}^{n} a_{ij} \operatorname{cof}(a_{ij}) 1 \leq i \leq n \tag{4.86.2}$$

O 2º membro da expressão anterior é conhecido por ***desenvolvimento de Laplace do determinante de*** A ***pela linha*** i.

Demonstração:

Basta considerar $p = 1$ em $(4.83.1)$ e $(4.83.2)$ e atender a que $\det(A[i;j]) = a_{ij}$. □

Observações:

- As expressões $(4.86.1)$ e $(4.86.2)$ dão uma definição *recursiva* do $\det(A)$, visto exprimirem um determinante de ordem n em termos de determinantes de ordem $n-1$. Quando se aplica o teorema de Laplace restrito, a ordem dos determinantes baixa de n para $n-1$ (embora o número de determinantes a calcular passe de 1 para n). Frisemos que convém aplicar o teorema à fila (linha ou coluna) que contenha mais zeros, visto que, sempre que $a_{ij} = 0$ e sendo este o elemento absorvente do produto em $\mathbb{K}$, o produto $a_{ij} \operatorname{cof}(a_{ij})$ anula-se e, assim, poupamos o trabalho de calcular o cofactor de a_{ij} (que envolveria o cálculo de um determinante de ordem $n-1$): isto significa que o número de parcelas não nulas em $(4.86.1)$ ou $(4.86.2)$ pode, na prática, ser inferior a n.

- É óbvio que o teorema de Laplace restrito conduz a um *algoritmo recursivo* para calcular determinantes de qualquer ordem n, segundo a seguinte fórmula de recorrência:

$$A = [\, a_{ij} \,]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \Rightarrow \det(A) = \begin{cases} a_{11} & \text{, se } n = 1 \\ \sum\limits_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A(i;j)), & \text{se } n > 1 \end{cases}$$

Aplica-se sucessivamente o teorema de Laplace restrito à fila mais conveniente de acordo com o critério descrito anteriormente, até que a ordem dos determinantes envolvidos seja calculável por qualquer das mnemónicas referidas nas figuras 4.6 e 4.7, ou seja, até que os determinantes a calcular sejam de 2ª ou 3ª ordem (podia mesmo levar-se a recorrência até se obter $n = 1$). Na secção 4.16 deste capítulo, implementamos na linguagem MATHEMATICA© a função **`RecursivoDet`** como parte do *package* **``ALGA`Determinantes` ``** e usando o algoritmo que acabámos de descrever.

Exemplo 4.44 Calculemos o valor do determinante da matriz A do exemplo 4.42, por recorrência, através do teorema de Laplace restrito, começando por aplicá-lo à linha 1 de A (que tem um zero), o que conduz a determinantes de 4ª ordem:

$$
\begin{aligned}
\det(A) &= \begin{vmatrix} -1 & 2 & 0 & -2 & 3 \\ 4 & 3 & 1 & -2 & 2 \\ -3 & 0 & 2 & 4 & 1 \\ 2 & -5 & 1 & 0 & 3 \\ 4 & 1 & -1 & 2 & 0 \end{vmatrix} \\
&= -1\begin{vmatrix} 3 & 1 & -2 & 2 \\ 0 & 2 & 4 & 1 \\ -5 & 1 & 0 & 3 \\ 1 & -1 & 2 & 0 \end{vmatrix} - 2\begin{vmatrix} 4 & 1 & -2 & 2 \\ -3 & 2 & 4 & 1 \\ 2 & 1 & 0 & 3 \\ 4 & -1 & 2 & 0 \end{vmatrix} \\
&+ 2\begin{vmatrix} 4 & 3 & 1 & 2 \\ -3 & 0 & 2 & 1 \\ 2 & -5 & 1 & 3 \\ 4 & 1 & -1 & 0 \end{vmatrix} + 3\begin{vmatrix} 4 & 3 & 1 & -2 \\ -3 & 0 & 2 & 4 \\ 2 & -5 & 1 & 0 \\ 4 & 1 & -1 & 2 \end{vmatrix}
\end{aligned}
$$

Aplicando, de novo, o teorema de Laplace restrito a cada um dos determinantes de 4ª ordem anteriores e desenvolvendo pela linha 2, para o primeiro; pela coluna 4, para o segundo; pela linha 2, para o terceiro e pela coluna 2 para o quarto, obtém-se (usando a regra de Sarrus da figura 4.7 ou uma das suas variantes para os determinantes de 3ª ordem obtidos):

$$
\begin{vmatrix} 3 & 1 & -2 & 2 \\ 0 & 2 & 4 & 1 \\ -5 & 1 & 0 & 3 \\ 1 & -1 & 2 & 0 \end{vmatrix} = 2\begin{vmatrix} 3 & -2 & 2 \\ -5 & 0 & 3 \\ 1 & 2 & 0 \end{vmatrix} - 4\begin{vmatrix} 3 & 1 & 2 \\ -5 & 1 & 3 \\ 1 & -1 & 0 \end{vmatrix} + 1\begin{vmatrix} 3 & 1 & -2 \\ -5 & 1 & 0 \\ 1 & -1 & 2 \end{vmatrix} = -160
$$

$$
\begin{vmatrix} 4 & 1 & -2 & 2 \\ -3 & 2 & 4 & 1 \\ 2 & 1 & 0 & 3 \\ 4 & -1 & 2 & 0 \end{vmatrix} = -2\begin{vmatrix} -3 & 2 & 4 \\ 2 & 1 & 0 \\ 4 & -1 & 2 \end{vmatrix} + 1\begin{vmatrix} 4 & 1 & -2 \\ 2 & 1 & 0 \\ 4 & -1 & 2 \end{vmatrix} - 3\begin{vmatrix} 4 & 1 & -2 \\ -3 & 2 & 4 \\ 4 & -1 & 2 \end{vmatrix} = -100
$$

$$
\begin{vmatrix} 4 & 3 & 1 & 2 \\ -3 & 0 & 2 & 1 \\ 2 & -5 & 1 & 3 \\ 4 & 1 & -1 & 0 \end{vmatrix} = 3\begin{vmatrix} 3 & 1 & 2 \\ -5 & 1 & 3 \\ 1 & -1 & 0 \end{vmatrix} - 2\begin{vmatrix} 4 & 3 & 2 \\ 2 & -5 & 3 \\ 4 & 1 & 0 \end{vmatrix} + 1\begin{vmatrix} 4 & 3 & 1 \\ 2 & -5 & 1 \\ 4 & 1 & -1 \end{vmatrix} = -20
$$

$$
\begin{vmatrix} 4 & 3 & 1 & -2 \\ -3 & 0 & 2 & 4 \\ 2 & -5 & 1 & 0 \\ 4 & 1 & -1 & 2 \end{vmatrix} = -3\begin{vmatrix} -3 & 2 & 4 \\ 2 & 1 & 0 \\ 4 & -1 & 2 \end{vmatrix} + 5\begin{vmatrix} 4 & 1 & -2 \\ -3 & 2 & 4 \\ 4 & -1 & 2 \end{vmatrix} + 1\begin{vmatrix} 4 & 1 & -2 \\ -3 & 2 & 4 \\ 2 & 1 & 0 \end{vmatrix} = 440
$$

Substituindo estes resultados no desenvolvimento inicial, vem

$$\det(A) = -1(-160) - 2(-100) + 2(-20) + 3 \times 440 = 160 + 200 - 40 + 1320 = 1640$$

Exemplo 4.45 Mostremos que, para todo o inteiro $n \geq 2$ e $(x_1, x_2, x_3, \ldots, x_{n-1}, x_n) \in \mathbb{K}^n$, o seguinte determinante V_n – dito de *Vandermonde*[10] – é dado por:

$$V_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ x_1 & x_2 & x_3 & \cdots & x_{n-1} & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_{n-1}^2 & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^{n-2} & x_2^{n-2} & x_3^{n-2} & \cdots & x_{n-1}^{n-2} & x_n^{n-2} \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_{n-1}^{n-1} & x_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

O produto no 2º membro tem $\binom{n}{2} = \frac{1}{2}(n-1)n$ factores.

Consideremos a sucessão de funções escalares $v_n\colon \mathbb{K} \to \mathbb{K}$ definida, para todo o $n \geq 2$ e $x \in \mathbb{K}$, por:

$$v_n(x) = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ x_1 & x_2 & x_3 & \cdots & x_{n-1} & x \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_{n-1}^2 & x^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^{n-2} & x_2^{n-2} & x_3^{n-2} & \cdots & x_{n-1}^{n-2} & x^{n-2} \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_{n-1}^{n-1} & x^{n-1} \end{vmatrix}; \; n \geq 2 \wedge x \in \mathbb{K}$$

É claro que o determinante de Vandermonde é dado por $V_n = v_n(x_n)$ e o teorema de Laplace restrito aplicado à última coluna do determinante anterior mostra que v_n constitui um polinómio em x de grau $n-1$ (em particular, v_2 é o polinómio de grau 1 definido por $v_2(x) = x - x_1$) e que o coeficiente de maior grau deste polinómio é o cofactor de x^{n-1}, ou seja, $v_{n-1}(x_{n-1})$. Além disso, tem-se $v_n(x_k) = 0$, para $1 \leq k \leq n-1$, visto que o determinante fica com a última coluna igual à coluna k, pelo que as raízes do polinómio v_n são $\{x_1, x_2, x_3, \ldots, x_{n-1}\}$. Decompondo o polinómio v_n em factores lineares, podemos escrever, para $n \geq 3$ e $x = x_n$,

$$v_n(x_n) = v_{n-1}(x_{n-1})\big((x_n - x_1)(x_n - x_2)\cdots(x_n - x_{n-1})\big) = v_{n-1}(x_{n-1}) \prod_{i=1}^{n-1}(x_n - x_i)$$

---

10 *Vandermonde, Alexandre-Théophile*: matemático francês (Paris, 1735 – Paris 1796).

Aplicando recursivamente a igualdade anterior, vem, para os $v_k(x_k)$, com $2 \leq k \leq n$,

$$\begin{cases} v_n(x_n) = v_{n-1}(x_{n-1}) \displaystyle\prod_{i=1}^{n-1}(x_n - x_i) \\ v_{n-1}(x_{n-1}) = v_{n-2}(x_{n-2}) \displaystyle\prod_{i=1}^{n-2}(x_{n-1} - x_i) \\ v_{n-2}(x_{n-2}) = v_{n-3}(x_{n-3}) \displaystyle\prod_{i=1}^{n-3}(x_{n-2} - x_i) \\ \quad \ldots\ldots\ldots\ldots \\ v_3(x_3) = v_2(x_2) \displaystyle\prod_{i=1}^{2}(x_3 - x_i) \\ v_2(x_2) = x_2 - x_1 = \displaystyle\prod_{i=1}^{1}(x_2 - x_i) \end{cases}$$

Substituindo sucessivamente $v_k(x_k)$ em $v_{k+1}(x_{k+1})$, para $k = 2, 3, \ldots, n-1$, obtemos a expressão seguinte para $v_n(x_n)$:

$$v_n(x_n) = \prod_{i=1}^{1}(x_2 - x_i) \prod_{i=1}^{2}(x_3 - x_i) \cdots \prod_{i=1}^{n-3}(x_{n-2} - x_i) \prod_{i=1}^{n-2}(x_{n-1} - x_i) \prod_{i=1}^{n-1}(x_n - x_i)$$

Observe que o número total de factores na expressão anterior é de

$$1 + 2 + \cdots + (n-3) + (n-2) + (n-1) = \frac{1}{2}(n-1)n = \binom{n}{2}$$

Podemos, por fim, escrever:

$$V_n = v_n(x_n) = \prod_{j=2}^{n} \left(\prod_{i=1}^{j-1} (x_j - x_i) \right) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

4.10 Método abreviado para o cálculo de determinantes

Vimos que, ao usarmos o teorema de Laplace restrito, há conveniência em fazer o desenvolvimento do determinante pela fila com o maior número possível de zeros, porque assim se minimiza o número de cofactores a calcular: o ideal seria que apenas um elemento da fila não fosse nulo! Esta observação conduz à ideia óbvia de usar primeiro o método de condensação para anular todos os elementos de uma fila à excepção de um e, de seguida, aplicar o teorema de Laplace restrito a essa fila. Tal algoritmo constitui o chamado *método misto* ou *abreviado* para o cálculo de determinantes (na prática, de ordem superior a 3). Neste algoritmo, para calcularmos um determinante de ordem n, agimos do modo seguinte:

- Por condensação, anulamos todos os elementos de uma fila qualquer (convém a que já tiver mais zeros à partida), com excepção do pivot (é claro que, se isto não for possível, o determinante é nulo).

- Aplicamos o teorema de Laplace restrito à fila referida no ponto anterior, o que nos conduzirá a um determinante apenas e de ordem $n-1$.

- Repetimos os dois passos anteriores para o determinante de ordem $n-1$ obtido acima e assim sucessivamente, até que o determinante a calcular seja de 2ª ou 3ª ordem e possa já ser calculado pelas mnemónicas das figuras 4.6 ou 4.7.

Exemplo 4.46 Usemos o algoritmo anterior para repetir o cálculo do determinante de 5ª ordem mencionado no exemplo 4.42

$$\det(A)=\begin{vmatrix}-1 & 2 & 0 & -2 & 3\\ 4 & 3 & 1 & -2 & 2\\ -3 & 0 & 2 & 4 & 1\\ 2 & -5 & 1 & 0 & 3\\ 4 & 1 & -1 & 2 & 0\end{vmatrix}$$

Por condensação horizontal (operações com colunas), anulemos os elementos da linha 3, usando $a_{35}=1$ como pivot, através das operações de tipo 3 (não alteram o valor do determinante) seguintes: $C_1\to C_1+3C_5, C_3\to C_3-2C_5, C_4\to C_4-4C_5$ e aplicando, de seguida, o teorema de Laplace restrito à linha 3:

$$\begin{vmatrix}-1 & 2 & 0 & -2 & 3\\ 4 & 3 & 1 & -2 & 2\\ -3 & 0 & 2 & 4 & 1\\ 2 & -5 & 1 & 0 & 3\\ 4 & 1 & -1 & 2 & 0\end{vmatrix}=\begin{vmatrix}8 & 2 & -6 & -14 & 3\\ 10 & 3 & -3 & -10 & 2\\ 0 & 0 & 0 & 0 & 1\\ 11 & -5 & -5 & -12 & 3\\ 4 & 1 & -1 & 2 & 0\end{vmatrix}=\begin{vmatrix}8 & 2 & -6 & -14\\ 10 & 3 & -3 & -10\\ 11 & -5 & -5 & -12\\ 4 & 1 & -1 & 2\end{vmatrix}$$

Após o procedimento anterior, usemos operações sobre linhas (condensação vertical) para anular os elementos da 2ª coluna, tomando $a_{42}=1$ como pivot (note que há conveniência em que o pivot seja igual a 1, para que se sigam operações de tipo 3); as operações de tipo 3 serão: $L_1\to L_1-2L_4, L_2\to L_2-3L_4, L_3\to L_3+5L_4$. O uso do teorema de Laplace restrito na 2ª coluna seguido da regra de Sarrus completam o cálculo (em vez de usarmos a regra de Sarrus, podíamos, claro está, baixar a ordem do determinante para 2, através da operação $C_3\to C_3-8C_1$ e de mais uma aplicação do teorema de Laplace, desta feita à linha 2):

$$\begin{vmatrix}8 & 2 & -6 & -14\\ 10 & 3 & -3 & -10\\ 11 & -5 & -5 & -12\\ 4 & 1 & -1 & 2\end{vmatrix}=\begin{vmatrix}0 & 0 & -4 & -18\\ -2 & 0 & 0 & -16\\ 31 & 0 & -10 & -2\\ 4 & 1 & -1 & 2\end{vmatrix}=\begin{vmatrix}0 & -4 & -18\\ -2 & 0 & -16\\ 31 & -10 & -2\end{vmatrix}=4\begin{vmatrix}0 & -2 & -9\\ -1 & 0 & -8\\ 31 & -10 & -2\end{vmatrix}=1640$$

4.11 Aplicação ao cálculo da característica de uma matriz

É possível calcular a característica de uma matriz, atendendo à teoria dos determinantes. Comecemos por provar a

Proposição 4.14 *Seja $A\in\mathbb{K}^{m,n}$ uma matriz não nula, de característica r:*

i) Qualquer submatriz M extraída de A tem característica inferior ou igual a r.

ii) Toda a matriz quadrada ***regular*** *extraída de A é de ordem inferior ou igual a r.*

iii) Existe uma matriz quadrada ***regular*** *extraída de A e de ordem r.*

Demonstração:

i) Sejam $1 \leq p \leq m$ e $1 \leq q \leq n$ e $M = A[i_1, i_2, \ldots, i_p; j_1, j_2, \ldots, j_q] \in \mathbb{K}^{p,q}$ uma submatriz de tipo $p \times q$ extraída de A. Considere-se a matriz

$$M' = A[1, 2, \ldots, m; j_1, j_2, \ldots, j_q] \in \mathbb{K}^{m,q}$$

que é composta pelas q colunas de A que contribuem para M (note que M é submatriz de M'); o espaço gerado pelas colunas de M' é subespaço do espaço gerado pelas colunas de A (ambos subespaços de $\mathbb{K}^m$), o que significa que a característica de M' é inferior ou igual à de A:

$$\mathrm{c}(M') \leq r$$

Mas o espaço gerado pelas linhas de M é subespaço do espaço gerado pelas linhas de M' (ambos subespaços de $\mathbb{K}^q$), donde se deduz ser a característica de M é inferior ou igual à de M':

$$\mathrm{c}(M) \leq \mathrm{c}(M')$$

Das duas relações anteriores se conclui o resultado pretendido $\mathrm{c}(M) \leq r$.

ii) Se M é matriz *regular* de ordem p extraída de A, teremos $\mathrm{c}(M) = p$ e, pela alínea anterior, será $p = \mathrm{c}(M) \leq r$, como diz o enunciado.

iii) Como A tem característica $r \leq \min\{m,n\}$, será essa a dimensão do espaço $U \subset \mathbb{K}^m$ gerado pela colunas $(\vec{a}_j)_{1 \leq j \leq n}$ de A. Como se viu na proposição 1.16.*vi* do capítulo 1, a sequência $(\vec{a}_j)_{1 \leq j \leq n}$ das colunas de A contém uma *base* $(\vec{a}_{j_k})_{1 \leq k \leq r}$ de U formada por r vectores, em que $1 \leq j_1 < j_2 < \cdots < j_r \leq n$.

Seja, agora, $M' = A[1, 2, \ldots, m; j_1, j_2, \ldots, j_r] \in \mathbb{K}^{m,r}$ a matriz extraída de A cujas colunas são os vectores da base referida. As colunas de M' são linearmente independentes e a característica de M' é igual a r. Isto implica que o subespaço $V \subset \mathbb{K}^r$ gerado pela sequência das linhas $x = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m)$ de M' tem dimensão r e, de novo pela proposição 1.16.*vi*, x conterá uma *base* $(\vec{x}_{ik})_{1 \leq k \leq r}$ de V formada por r vectores. Desta forma, a matriz

$$M = A[i_1, i_2, \ldots, i_r; j_1, j_2, \ldots, j_r]$$

é a matriz quadrada *regular* de ordem r e extraída de A, cuja existência o enunciado garante. $\square$

Observações:

- Numa matriz A não nula de característica r, todas as submatrizes quadradas regulares têm ordem, quando muito, igual a r: r representa, assim, um *majorante* das ordens das submatrizes quadradas regulares de A.

- Mas existe uma submatriz quadrada regular de ordem r, o que implica que r é o *máximo* do conjunto das ordens das submatrizes quadradas regulares de A:

$$r = \max\{p\colon A[i_1, i_2, \ldots, i_p; j_1, j_2, \ldots, j_p] \text{ é regular}\}$$

- Isto significa que submatrizes quadradas extraídas de A e de ordem $p > r$:

  – Ou não existem, quando $r = \min\{m,n\}$.

– Ou são todas singulares, se $r < \min\{m,n\}$ e em número de $\sum_{p=r+1}^{\min\{m,n\}} \left(\binom{m}{p} \times \binom{n}{p}\right)$.

- Em particular, as submatrizes quadradas extraídas de A e de ordem $p = r + 1$:

– Ou não existem, quando $r = \min\{m,n\}$.

– Ou são todas singulares, se $r < \min\{m,n\}$ e em número de $\binom{m}{r+1} \times \binom{n}{r+1}$.

Vamos, agora, definir a noção de *matriz principal* extraída de outra e a de *matriz orlada*, que intervirão na proposição 4.15 que caracterizará as matrizes principais.

Definição 4.10 – Matriz principal – *Seja $A \in \mathbb{K}^{m,n}$ uma matriz não nula de tipo $m \times n$ e de característica r. Chama-se **matriz principal extraída de** A a qualquer submatriz*

$$P = A[i_1, i_2, \ldots, i_r; j_1, j_2, \ldots, j_r] \in \mathbb{K}^{r,r}$$

*quadrada **regular** de ordem $p = r$ extraída de A (observe que, sendo $A \neq O_{m,n}$, será $r > 0$).*

*As linhas de índices $i_1, i_2, \ldots, i_r$ são chamadas **linhas principais** de A e as colunas de índices $j_1, j_2, \ldots, j_r$ são **colunas principais** de A.*

*Por outras palavras, uma matriz principal extraída de A é uma submatriz P quadrada **regular** de A e de ordem p **máxima** (igual a r). Em geral, é possível extrair várias matrizes principais de A.*

Definição 4.11 – Matriz orlada – *Seja $A \in \mathbb{K}^{m,n}$ uma matriz de tipo $m \times n$ e p e q dois inteiros tais que $0 < p < m$ e $0 < q < n$. Seja ainda $P = A[i_1, i_2, \ldots, i_p; j_1, j_2, \ldots, j_q] \in \mathbb{K}^{p,q}$ uma submatriz de tipo $p \times q$ extraída de A. Chama-se **matriz orlada associada a** P a toda a matriz M de tipo $(p+1) \times (q+1)$ da forma*

$$M = A[\{i_0\} \cup \{i_1, i_2, \ldots, i_p\}; \{j_0\} \cup \{j_1, j_2, \ldots, j_q\}] \in \mathbb{K}^{p+1,q+1}$$

em que $i_0 \in [1, m] \setminus \{i_1, i_2, \ldots, i_p\}$ e $j_0 \in [1, n] \setminus \{j_1, j_2, \ldots, j_q\}$; isto significa que uma matriz orlada associada a P contém uma linha e uma coluna adicionais extraídas das linhas e colunas de A que não contribuíram para a formação de P. O número de matrizes orladas associadas a P é de $(m-p) \times (n-q)$ (uma por cada par (i_0, j_0), com $i_0 \in [1, m] \setminus \{i_1, i_2, \ldots, i_p\}$ e $j_0 \in [1, n] \setminus \{j_1, j_2, \ldots, j_q\}$).

Proposição 4.15 – Caracterização das matrizes principais – *Seja $A \in \mathbb{K}^{m,n}$ uma matriz não nula e P uma matriz quadrada **regular** extraída de A. A matriz P é principal se e só se não existem matrizes orladas associadas a P que sejam regulares.*

Demonstração:

Seja r a característica de A e p a ordem da matriz *regular*

$$P = A[i_1, i_2, \dots, i_p; j_1, j_2, \dots, j_p]$$

extraída de A. Observe-se que a proposição 4.14.*ii implica que* $p \leq r$.

▪ Se a matriz P é principal, então $p = r$ e, devido à proposição 4.14.*ii*, não existem matrizes regulares de ordem estritamente superior a p e extraídas de A. Em particular, não existem matrizes orladas (que terão ordem $p+1 > p$) regulares associadas a P.

▪ Reciprocamente, suponha-se que $P = A[i_1, i_2, \dots, i_p; j_1, j_2, \dots, j_p] \in \mathbb{K}^{p,p}$ não é principal, isto é, que P é regular e $p < r$.

Sejam $(\vec{a}_{j_k})_{1 \leq k \leq p}$ as p colunas de A que contribuem para P e $M \in \mathbb{K}^{m,p}$ a matriz

$$M = A[1, 2, \dots, m; j_1, j_2, \dots, j_p] \in \mathbb{K}^{m,p}$$

que tem por colunas estes vectores. A matriz P é uma matriz regular extraída de M; a característica de M é, portanto, superior ou igual a p e, de facto, é mesmo igual a p, visto que M tem só p colunas, o que prova que os $(\vec{a}_{j_k})_{1 \leq k \leq p}$ são linearmente independentes. Como $p < r$, existe um vector-coluna $\vec{a}_{j_0}$, com $j_0 \in [1, n] \setminus \{j_1, j_2, \dots, j_p\}$, tal que

$$\vec{a}_{j_0} \in \mathbb{K}^m \setminus L(\vec{a}_{j_k})_{1 \leq k \leq p}$$

Pela proposição 1.12, a sequência $(\vec{a}_k)_{k \in \{j_0\} \cup \{j_1, j_2, \dots, j_p\}}$ é ainda linearmente independente e a característica da matriz

$$M' = A[1, 2, \dots, m; \{j_0\} \cup \{j_1, j_2, \dots, j_p\}] \in \mathbb{K}^{m,p+1}$$

que tem por colunas os vectores desta sequência é igual a $p+1$. Como a matriz P é regular, as p linhas $\left(\vec{x}_{i_1}, \vec{x}_{i_2}, \dots, \vec{x}_{i_p}\right)$ de M' representadas em P são linearmente independentes em $\mathbb{K}^{p+1}$ (ver observações na pág. 28 do capítulo 1) e, como o espaço gerado pelas m linhas de M' tem dimensão $p+1$, existe forçosamente um vector-linha de M', $\vec{x}_{i_0}$, com $i_0 \in [1, m] \setminus \{i_1, i_2, \dots, i_p\}$, tal que

$$\vec{x}_{i_0} \in \mathbb{K}^{p+1} \setminus L(\vec{x}_{i_k})_{1 \leq k \leq p}$$

A proposição 1.12 assegura então que a sequência $(\vec{x}_k)_{k \in \{i_0\} \cup \{i_1, i_2, \dots, i_p\}}$ de vectores de $\mathbb{K}^{p+1}$ é ainda linearmente independente. Isto mostra que a matriz

$$M'' = A[\{i_0\} \cup \{i_1, i_2, \dots, i_p\}; \{j_0\} \cup \{j_1, j_2, \dots, j_p\}] \in \mathbb{K}^{p+1,p+1}$$

(quadrada de ordem $p+1$), formada por estes vectores é *regular* e é uma matriz orlada associada a P, o que completa a demonstração. □

Observações:

Do que ficou exposto, se conclui que:

▪ Se a matriz $A \in \mathbb{K}^{m,n}$ é nula, a sua característica é igual a 0 (não existem matrizes principais extraídas de A).

- Se a matriz $A \in \mathbb{K}^{m,n} \setminus \{O_{m,n}\}$, o que ficou exposto mostra que a característica de A é a ordem de qualquer matriz principal P extraída de A, ou seja, é a ordem p de uma submatriz regular P, para a qual não existam matrizes orladas *regulares* associadas.

- Quando encontramos uma submatriz M regular (necessariamente de ordem $p \leq r$) extraída de A, *a vantagem do critério das matrizes orladas é diminuir o número de submatrizes de ordens estritamente superiores a p cuja singularidade deve ser verificada* (calculando os determinantes respectivos). Este número baixa de $\sum_{k=p+1}^{\min\{m,n\}} \left(\binom{m}{k} \times \binom{n}{k}\right)$ (total de submatrizes quadradas de ordens $k > p$) para apenas $(m-p) \times (n-p)$ (total de matrizes orladas associadas a M).

- Na secção 4.16, apresenta-se a função **`Caracteristica`** como parte do *package* **`ALGA`Determinantes``** e que determina a característica de uma matriz, pelo método da matriz principal. No package existem ainda funções para a determinação de um par de filas principais da matriz – função **`FilasPrincipais`** – e para determinar uma matriz principal – função **`MatrizPrincipal`**.

Exemplo 4.47 Considere-se a matriz real, de tipo 4×5,

$$A = \begin{bmatrix} 2 & 1 & -1 & 3 & 2 \\ 1 & 1 & -4 & 2 & -2 \\ 1 & 0 & 3 & 1 & 4 \\ 0 & 1 & -7 & 1 & -6 \end{bmatrix}$$

e tomemos a submatriz $P = A[1,2;1,2]$ de ordem $p = 2$, calculando o seu determinante

$$\det(P) = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1 \neq 0$$

A matriz P é regular e tem $(m-p)(n-p) = (4-2)(5-2) = 6$ matrizes orladas associadas, cujos determinantes são

$$\begin{vmatrix} 2 & 1 & -1 \\ 1 & 1 & -4 \\ 1 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 2 \\ 1 & 1 & -2 \\ 1 & 0 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 1 & -1 \\ 1 & 1 & -4 \\ 0 & 1 & -7 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 2 \\ 1 & 1 & -2 \\ 0 & 1 & -6 \end{vmatrix} = 0$$

Como não existem matrizes orladas *regulares* associadas a P, podemos concluir que $r = \mathrm{c}(A) = p = 2$ e que $P = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ é uma matriz principal extraída de A. As duas primeiras linhas e colunas de A são as principais.

Exemplo 4.48 Considere-se a matriz real, de tipo 4×3,

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 3 & 2 & 1 \\ 1 & 4 & -1 \\ -2 & 1 & 0 \end{bmatrix}$$

e tomemos a submatriz $P = A[1,2,3;1,2,3]$ de ordem $p = 3$, calculando o seu determinante

$$\det(P) = \begin{vmatrix} 2 & -1 & 0 \\ 3 & 2 & 1 \\ 1 & 4 & -1 \end{vmatrix} = -16 \neq 0$$

A matriz P é regular e tem $(m-p)(n-p) = (4-3)(3-3) = 0$ matrizes orladas associadas. Como não existem matrizes orladas *regulares* (e, neste caso, nem singulares...) associadas a P, podemos concluir que $r = \mathrm{c}(A) = p = 3$ e que $P = \begin{bmatrix} 2 & -1 & 0 \\ 3 & 2 & 1 \\ 1 & 4 & -1 \end{bmatrix}$ é uma matriz principal extraída de A. As linhas e as colunas principais de A são as 3 primeiras.

4.12 Aplicação aos sistemas de equações lineares. Regra de Cramer

A teoria dos determinantes pode, como vamos ver, ser aplicada ao estudo e resolução dos sistemas de equações lineares. Para tal, comecemos por estabelecer a seguinte

Definição 4.12 – Matrizes características de um sistema linear – *Seja $AX = B$ um sistema linear de m equações a n incógnitas sobre um corpo $\mathbb{K}$, r a característica de $A \in \mathbb{K}^{m,n} \setminus \{O_{m,n}\}$ e $A' = \big[A \,\big|\, B\big] \in \mathbb{K}^{m,n+1}$ a sua matriz completa.*

*Sejam ainda $P = A[i_1, \dots, i_r; j_1, \dots, j_r] \in \mathbb{K}^{r,r}$ uma matriz principal extraída de A e $S = \{1, 2, \dots, m\} \setminus \{i_1, \dots, i_r\}$ o complementar de $\{i_1, \dots, i_r\}$ em $\{1, 2, \dots, m\}$. Chamaremos **matriz característica** associada a P e relativa à equação k a cada uma das $m - r$ matrizes $(C_k)_{k \in S}$ (de ordem $r + 1$) da forma*

$$C_k = A'[\{i_1, \dots, i_r\} \cup \{k\}; j_1, \dots, j_r, n+1];\ k \in \{1, 2, \dots, m\} \setminus \{i_1, \dots, i_r\}$$

*Se $r = m$, é $S = \emptyset$ e não existem matrizes características associadas a P; o número de matrizes características é de $m - r$. O determinante de C_k é chamado **determinante característico relativo à equação** k; as equações correspondentes aos índices $i_1, \dots, i_r$ das linhas presentes em P são chamadas **equações principais** do sistema e as incógnitas correspondentes às colunas $j_1, \dots, j_r$ presentes em P são as **incógnitas principais**, dizendo-se das restantes $n - r$ incógnitas que são **secundárias**.*

Podemos, agora, enunciar uma condição necessária e suficiente, baseada em determinantes, para que um sistema de equações lineares $AX = B$ seja possível e que constitui o *teorema de Rouché*(11):

---

11 *Rouché, Eugène*: matemático francês (Sommières 1832 – Lunel 1910).

Proposição 4.16 – Teorema de Rouché – *Seja $AX = B$ um sistema linear de m equações a n incógnitas sobre um corpo $\mathbb{K}$ e $A \neq O_{m,n}$. Nestas condições e sendo $r = \mathrm{c}(A)$, tem-se $0 < r \leq m$ e são equivalentes as proposições:*

i) *O sistema $AX = B$ é possível.*

ii) *Existe uma matriz principal $P \in \mathbb{K}^{r,r}$ extraída de A, tal que são nulos os $m - r$ determinantes característicos associados a P (ou não existem, se $r = m$).*

iii) *Para toda a matriz principal $P \in \mathbb{K}^{r,r}$ extraída de A, são nulos os $m - r$ determinantes característicos associados a P (ou não existem, se $r = m$) .*

Demonstração:

Sejam $A' = \left[A \,|\, B\right]$ a matriz ampliada do sistema, $r = \mathrm{c}(A)$ e $s = \mathrm{c}(A')$. Vimos no capítulo 2 que o sistema é possível sse $\mathrm{c}(A') = \mathrm{c}(A)$. Portanto, podemos substituir a condição *i)* pela condição:

i′) $\mathrm{c}(A') = \mathrm{c}(A)$.

- Mostremos que *i′) $\Rightarrow$ iii)*: suponhamos que a característica de A é igual à de A'; então toda a matriz principal P extraída de A é também matriz principal extraída de A', o que significa que todas as matrizes orladas de P em A' têm determinantes nulos e, em particular, são nulos os determinantes característicos associados a P.

- *iii) $\Rightarrow$ ii)*: como $A \neq O_{m,n}$, pela proposição 4.14.*iii* existe uma matriz principal $P \in \mathbb{K}^{r,r}$ de A e, por hipótese, serão nulos todos os determinantes característicos de P.

- *ii) $\Rightarrow$ i′)*: seja $P \in \mathbb{K}^{r,r}$ uma matriz principal extraída de A e tal que são nulos todos determinantes característicos associados a P e mostremos que P é também matriz principal extraída de A'. Se for $U \in \mathbb{K}^{r+1,r+1}$ uma matriz orlada associada a P em A', U está numa de duas situações possíveis:

- A coluna $r + 1$ de U é extraída de B. Neste caso a matriz orlada U é uma matriz característica associada a P e é nulo o seu determinante, por hipótese.

- A coluna $r + 1$ de U é extraída de A. Neste caso, $U \in \mathbb{K}^{r+1,r+1}$ é uma matriz extraída de A e, como P é principal extraída de A, será nulo o determinante de U.

São, portanto, singulares todas as matrizes U orladas de P em A' o que implica que P é matriz principal extraída de A', ou seja $\mathrm{c}(A') = \mathrm{c}(A)$, terminando a demonstração. □

Os seguintes exemplos ilustram o uso da proposição anterior para estudar a existência de solução dos sistemas lineares.

Exemplo 4.49 Considere-se o sistema

$$\begin{cases} 2x + y - z + 3w = 0 \\ x + y - 4z + 2w = 1 \\ x + 3z + w \qquad = -1 \\ y - 7z + w \qquad = 2 \end{cases}$$

Fica

$$A = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 1 & 1 & -4 & 2 \\ 1 & 0 & 3 & 1 \\ 0 & 1 & -7 & 1 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \end{bmatrix}$$

Tomemos $P = A[1, 2; 1, 2]$ e calculemos o seu determinante

$$\det(P) = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1 \neq 0$$

A matriz P tem ordem $p = 2$ e tem $(m - p)(n - p) = 4$ matrizes orladas associadas, cujos determinantes são

$$\begin{vmatrix} 2 & 1 & -1 \\ 1 & 1 & -4 \\ 1 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 1 & -1 \\ 1 & 1 & -4 \\ 0 & 1 & -7 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{vmatrix} = 0$$

Daqui resulta que $r = \mathrm{c}(A) = p = 2$ e que a matriz $P = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ é uma matriz principal extraída de A.

No caso presente, existem $m - r = 4 - 2 = 2$ determinantes característicos (relativos às 2ª e 3ª equações) e valem

$$\begin{vmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{vmatrix} = 0$$

Da proposição 4.16.*ii*, podemos então concluir que o sistema dado é *possível* (e *indeterminado* de grau $n - r = 4 - 2 = 2$). As duas primeiras equações são principais (por dizerem respeito às linhas presentes em P) e as incógnitas x, y são também principais (por corresponderem às colunas representadas em P); z, w são secundárias.

Exemplo 4.50 Considere-se o sistema

$$\begin{cases} 2x + 3y - 2z = 5 \\ x - 2y + 3z \;\; = 2 \\ 4x - y + 4z \;\; = 1 \end{cases}$$

Neste caso, tem-se

$$A = \begin{bmatrix} 2 & 3 & -2 \\ 1 & -2 & 3 \\ 4 & -1 & 4 \end{bmatrix} \text{ e } B = \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}$$

Façamos $P = A[1,2;1,2]$ e calculemos $\det(P)$ e ainda o determinante da única matriz orlada associada a P (neste caso, trata-se da própria matriz A)

$$\det(P) = \begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7 \neq 0 \wedge \det(A) = 0$$

Podemos concluir que $r = \mathrm{c}(A) = 2$ e que $P = \begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix}$ é uma matriz principal extraída de A. Neste caso, existe apenas um $(m - r = 1)$ determinante característico (relativo à 3ª equação) e vale

$$\begin{vmatrix} 2 & 3 & 5 \\ 1 & -2 & 2 \\ 4 & -1 & 1 \end{vmatrix} = 56 \neq 0$$

Por fim, a proposição 4.16.*iii* permite concluir que o sistema é *impossível*.

Vamos, agora, usar a teoria dos determinantes para a resolução de sistema de equações lineares. Consideraremos dois casos:

No primeiro, vamos considerar apenas sistemas $AX = B$, com A matriz quadrada regular (igual número de equações e de incógnitas e com uma só solução), chamados *sistemas de Cramer*[(12)].

No segundo caso, vamos estudar os sistemas $AX = B$, com $r = \mathrm{c}(A) = m < n$ (número de equações estritamente inferior ao número de incógnitas e indeterminados de grau $n - r$).

Proposição 4.17 – Regra de Cramer – *Seja $n > 0$ um inteiro e considere-se um sistema de n equações lineares a n incógnitas $x_1, x_2, \ldots, x_n$ sobre um corpo $\mathbb{K}$ da forma $AX = B$, onde a matriz quadrada A é* ***regular*** *(sistema de Cramer), com colunas $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n \in \mathbb{K}^n$, e $\vec{b} \in \mathbb{K}^n$ é o vector-coluna B. Nestas condições, o sistema tem uma e uma só solução que é dada por:*

$$x_k = \frac{\det\left(\vec{a}_1, \vec{a}_2, \ldots \vec{a}_{k-1}, \vec{b}, \vec{a}_{k+1}, \ldots, \vec{a}_n\right)}{\det(A)}; \; k = 1, 2, \ldots, n \tag{4.87}$$

Demonstração:

Como A (a matriz simples do sistema) é quadrada e regular, tem-se $\mathrm{c}\left(\left[A \,|\, B\right]\right) = \mathrm{c}(A) = n$ e o sistema é *determinado*, existindo uma e uma só solução. Para a determinar, recorramos à forma vectorial do sistema: se $(x_1, x_2, \ldots, x_n)$ for a solução, existirá uma e uma só forma de exprimir o vector $\vec{b}$ nas colunas $\vec{a}_i$ de A:

$$\sum_{i=1}^{n} x_i \vec{a}_i = \vec{b}$$

---

12 *Cramer, Gabriel*: matemático suíço (Genève 1704 – Bagnols 1752).

Mas a função det é n-*linear alternada* e, portanto, para todo o $1 \leq k \leq n$, será

$$\det\left(\vec{a}_1, \vec{a}_2, \ldots \vec{a}_{k-1}, \vec{b}, \vec{a}_{k+1}, \ldots, \vec{a}_n\right) = \det\left(\vec{a}_1, \vec{a}_2, \ldots \vec{a}_{k-1}, \sum_{i=1}^{n} x_i\vec{a}_i, \vec{a}_{k+1}, \ldots, \vec{a}_n\right)$$

$$= \sum_{i=1}^{n} x_i \underbrace{\det(\vec{a}_1, \vec{a}_2, \ldots \vec{a}_{k-1}, \vec{a}_i, \vec{a}_{k+1}, \ldots, \vec{a}_n)}_{=\delta_{ik}\det(A)} = x_k\det(A)$$

Dividindo ambos membros por $\det(A) \neq 0$, obtém-se imediatamente o resultado pretendido.

□

Observações:

- A primeira nota que faremos consiste em chamar a atenção para o facto de a regra de Cramer permitir o cálculo de cada incógnita independentemente das restantes, ao contrário do método de condensação que é um método global que só permite calcular o vector solução na sua totalidade (não é possível calcular uma incógnita apenas).

- O cálculo do vector solução $(x_1, x_2, \ldots, x_n)$ pela regra de Cramer acarreta o cálculo de $n+1$ determinantes de ordem n (um por cada incógnita e ainda o determinante de A), o que pode ser muito trabalhoso, para valores elevados de n. Para estes casos, é preferível usar o método de condensação, reservando o uso da regra de Cramer, para os casos em que se pretendem apenas alguma(s) incógnita(s).

Exemplo 4.51 Considere-se o sistema de equações lineares

$$\begin{cases} x + y + 2z - w &= 2 \\ x + 2y + 2z &= -1 \\ -x - 4z + 3w &= 3 \\ -x - 3y - z - w &= 1 \end{cases}$$

Tem-se, usando o método de condensação,

$$\det(A) = \begin{vmatrix} 1 & 1 & 2 & -1 \\ 1 & 2 & 2 & 0 \\ -1 & 0 & -4 & 3 \\ -1 & -3 & -1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 2 \\ 0 & -2 & 1 & -2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = -1 \neq 0$$

Daqui se conclui que A é quadrada e regular: estamos, pois, perante um sistema de Cramer. Para calcular as incógnitas x, y, z, w, teremos que calcular mais quatro determinantes de 4ª ordem, para as matrizes obtidas de A mediante a substituição das colunas 1 a 4 por

$B = [\,2\ \ -1\ \ 3\ \ 1\,]^{\mathrm{T}}$.

$$\det(\Delta_x) = \begin{vmatrix} 2 & 1 & 2 & -1 \\ -1 & 2 & 2 & 0 \\ 3 & 0 & -4 & 3 \\ 1 & -3 & -1 & -1 \end{vmatrix} = -15; \qquad \det(\Delta_y) = \begin{vmatrix} 1 & 2 & 2 & -1 \\ 1 & -1 & 2 & 0 \\ -1 & 3 & -4 & 3 \\ -1 & 1 & -1 & -1 \end{vmatrix} = 5$$

$$\det(\Delta_z) = \begin{vmatrix} 1 & 1 & 2 & -1 \\ 1 & 2 & -1 & 0 \\ -1 & 0 & 3 & 3 \\ -1 & -3 & 1 & -1 \end{vmatrix} = 3; \qquad \det(\Delta_w) = \begin{vmatrix} 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & -1 \\ -1 & 0 & -4 & 3 \\ -1 & -3 & -1 & 1 \end{vmatrix} = -2$$

A solução do sistema, segundo (4.87), será o vector $(15, -5, -3, 2)$. Observemos que a determinação da solução envolveu o cálculo de 5 determinantes de 4ª ordem, para os quais tivémos que utilizar um dos métodos gerais de cálculo (por exemplo, 5 condensações). Ora, o método de Gauss-Jordan, permite-nos com *uma só condensação* determinar a solução!

Vamos, agora, abordar o segundo caso mencionado em cima:

Proposição 4.18 – Regra de Cramer para sistemas indeterminados – *Seja $AX = B$ um sistema com m equações e n incógnitas e com $r = \mathrm{c}(A) = m$. No capítulo 1, vimos que um tal sistema é sempre possível e a condição $\mathrm{c}(A) = m$ implica $m \leq n$. O caso $r = n$ é o que foi já tratado na proposição anterior (regra de Cramer).*

Suponha-se então que $m = r < n$ e sejam $(\vec{a}_k)_{k\in[1,n]}$ as colunas de A (que são vectores de $\mathbb{K}^r$). Neste caso, o sistema é indeterminado de grau $n - r$ e a solução pode ser calculada do seguinte modo: seja

$$P = A[1, 2, \ldots, r; j_1, \ldots, j_r] \in \mathbb{K}^{r,r}$$

(onde $1 \leq j_1 < j_2 < \ldots < j_r \leq n$) uma submatriz principal de A e $\vec{a}_{j_1}, \vec{a}_{j_2}, \ldots, \vec{a}_{j_r} \in \mathbb{K}^r$ as r colunas linearmente independentes de P. Introduzindo os $n - r$ parâmetros escalares arbitrários $(t_k)_{k\in[1,n]\setminus\{j_1,\ldots,j_r\}}$, seja

$$\vec{b}' = \vec{b} - \sum_{k\in[1,n]\setminus\{j_1,\ldots,j_r\}} t_k \vec{a}_k$$

A solução geral do sistema será dada por:

$$\begin{cases} x_{j_k} = \dfrac{\det\left(\vec{a}_{j_1}, \vec{a}_{j_2}, \ldots, \vec{a}_{j_{k-1}}, \vec{b}', \vec{a}_{j_{k+1}}, \ldots, \vec{a}_{j_r}\right)}{\det(P)}; \ k = 1, 2, \ldots, r \\ x_k = t_k; \ k \in [1, n] \setminus \{j_1, \ldots, j_r\} \end{cases} \tag{4.88}$$

Demonstração:

Sendo $P = A[1, 2, \ldots, r; j_1, \ldots, j_r] \in \mathbb{K}^{r,r}$ uma matriz principal de A, isso significa que $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ são as incógnitas principais e as restantes $n - r$ são as secundárias. Introduzamos os $n - r$ parâmetros escalares $t_k = x_k$, para $k \in [1, n] \setminus \{j_1, \ldots, j_r\}$ e determinemos as soluções do sistema em função desses parâmetros, recorrendo à forma vectorial do sistema: para qualquer solução $(x_1, x_2, \ldots, x_n)$ do sistema será:

$$\sum_{s=1}^{r} x_{j_s}\vec{a}_{j_s} + \sum_{k\in[1,n]\setminus\{j_1,\ldots,j_r\}} t_k\vec{a}_k = \vec{b} \Rightarrow \sum_{s=1}^{r} x_{j_s}\vec{a}_{j_s} = \vec{b} - \sum_{k\in[1,n]\setminus\{j_1,\ldots,j_r\}} t_k\vec{a}_k = \vec{b}'$$

Mas a função det é n-*linear alternada* e, portanto, para todo o $1 \leq k \leq r$, será

$$\begin{aligned}\det\Big(\vec{a}_{j_1}, \vec{a}_{j_2}, \ldots, \vec{a}_{j_{k-1}}, \vec{b}', \vec{a}_{j_{k+1}}, \ldots, \vec{a}_{j_r}\Big) &= \det\Big(\vec{a}_{j_1}, \vec{a}_{j_2}, \ldots, \vec{a}_{j_{k-1}}, \sum_{s=1}^{r} x_{j_s}\vec{a}_{j_s}, \vec{a}_{j_{k+1}}, \ldots, \vec{a}_{j_r}\Big)\\ &= \sum_{s=1}^{r} x_{j_s} \underbrace{\det(\vec{a}_{j_1}, \vec{a}_{j_2}, \ldots, \vec{a}_{j_{k-1}}, \vec{a}_{j_s}, \vec{a}_{j_{k+1}}, \ldots, \vec{a}_{j_r})}_{=\delta_{sk}\det(P)}\\ &= x_{j_k}\det(P)\end{aligned}$$

Basta agora dividir ambos membros por $\det(P) \neq 0$, para obter o resultado pretendido.

$\square$

Observações:

- A equação (4.88) é em tudo semelhante à regra de Cramer, mas agora, para calcular uma incógnita principal, calcula-se o determinante da matriz que resulta de substituir em P a coluna relativa a essa incógnita principal por $\vec{b}' = \vec{b} - \sum_{k\in[1,n]\setminus\{j_1,\ldots,j_r\}} t_k\vec{a}_k$.

- Observemos também que, uma vez que podem existir várias matrizes principais extraídas de A constituídas por diferentes escolhas de $r < n$ colunas de A, o conjunto das r incógnitas principais é, até certo ponto, arbitrário (num sistema duplamente indeterminado com 5 incógnitas, poderão ser principais as incógnitas (x_1, x_2, x_3), (x_1, x_3, x_4) ou ainda (x_2, x_3, x_5), por exemplo).

- Podemos resolver um sistema $AX = B$ de m equações a n incógnitas, segundo o seguinte procedimento:

Determinamos uma matriz principal P extraída de A, o que conduz à característica r de A, que será a ordem de P. Para a procura de P socorremo-nos da proposição 4.15.

– Se $r = m = n$, o sistema é determinado (sistema de Cramer) e resolve-se pela regra de Cramer da proposição 4.17.

– Se $r = m < n$, o sistema é indeterminado de grau $n - r$ e resolve-se, pela regra de Cramer da proposição 4.18.

– Se $r < m$, então calculamos os $m - r$ determinantes característicos associados a P. Se algum for $\neq 0$, o sistema é impossível (teorema de Rouché); se todos são nulos, então o sistema é possível (de novo, o teorema de Rouché) e é equivalente ao sistema de r equações a n incógnitas formado pelas equações principais (as equações correspondentes às linhas principais de A: as linhas das quais se extraiu P). Este sistema está numa situações anteriores e resolve-se pela regra de Cramer da proposição 4.17, se $r = n$; ou pela regra de Cramer da proposição 4.18, se $r < n$.

O seguinte esquema resume a discussão anterior, em que r é a ordem de uma matriz principal P extraída de A (igual à característica de A):

$$\begin{cases} r=m \Rightarrow \begin{array}{l} \text{sist}^{\text{a}} \text{ possível} \\ \text{(0 determinantes} \\ \text{característicos)} \end{array} \begin{cases} r=n \Rightarrow \text{det}^{\text{o}} \text{ (prop. 4.17)} \\ r<n \Rightarrow \text{indet}^{\text{o}} \text{ de grau } n-r \text{ (prop. 4.18)} \end{cases} \\ r<m \Rightarrow \begin{array}{c} m-r \text{ determinantes} \\ \text{característicos} \end{array} \begin{cases} \text{Algum é } \neq 0 \Rightarrow \text{sist}^{\text{a}} \text{ impossível} \\ \text{Todos nulos } \Rightarrow \begin{array}{c} r \text{ equações} \\ \text{principais} \end{array} \begin{cases} r=n \Rightarrow \text{det}^{\text{o}} \text{ (prop. 4.17)} \\ r<n \Rightarrow \text{indet}^{\text{o}} \text{ de grau } n-r \text{ (prop. 4.18)} \end{cases} \end{cases} \end{cases}$$

■ O *package* **`ALGA`Determinantes``** contém a função **`Cramer`** que implementa o método anterior para resolver qualquer sistema linear (ou constatar a inexistência de solução).

Exemplo 4.52 Vejamos o sistema

$$\begin{cases} 2x-2y+z-3v+w=-1 \\ x+2y+4z-3w \qquad = 4 \\ -3x-2z+5v-w \quad = -1 \end{cases}$$

Tem-se

$$A = \begin{bmatrix} 2 & -2 & 1 & -3 & 1 \\ 1 & 2 & 4 & 0 & -3 \\ -3 & 0 & -2 & 5 & -1 \end{bmatrix} \text{ e } B = \begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix}$$

A matriz

$$P = A[1,2,3;1,3,4] = \begin{array}{c} \begin{array}{ccc} x & z & v \end{array} \\ \begin{bmatrix} 2 & 1 & -3 \\ 1 & 4 & 0 \\ -3 & -2 & 5 \end{bmatrix} \end{array}$$

é tal que $\det(P) = 5 \neq 0$ e seja $p = 3$ a sua ordem (e característica). Não existem matrizes orladas associadas a P ($(m-p)(n-p) = 0$); deste modo, P é uma matriz *principal* de A e $r = p = m = 3 < n$, sendo o sistema *possível* e *duplamente indeterminado* ($n - r = 5 - 3$).

As incógnitas principais são x, z e v (cujas colunas contribuíram para a formação de P); y e w são secundárias. Pela proposição 4.18, para determinar a solução geral do sistema, calculamos os 3 determinantes seguintes introduzindo os parâmetros reais arbitrários a e b, com

$a = y$ e $b = w$,

$$\det(\Delta_x) = \begin{vmatrix} -1+2a-b & 1 & -3 \\ 4-2a+3b & 4 & 0 \\ -1+b & -2 & 5 \end{vmatrix} = -28+38a-5b$$
$$\det(\Delta_z) = \begin{vmatrix} 2 & -1+2a-b & -3 \\ 1 & 4-2a+3b & 0 \\ -3 & -1+b & 5 \end{vmatrix} = 12-12a+5b$$
$$\det(\Delta_v) = \begin{vmatrix} 2 & 1 & -1+2a-b \\ 1 & 4 & 4-2a+3b \\ -3 & -2 & -1+b \end{vmatrix} = -13+18a$$

A solução geral do sistema será então, para qualquer $(a', b') = \frac{1}{5}(a, b) \in \mathbb{R}^2$,

$$\begin{aligned}(x, y, z, v, w) &= \left(\frac{1}{5}(-28+38a-5b), a, \frac{1}{5}(12-12a+5b), \frac{1}{5}(-13+18a), b\right) \\ &= \frac{1}{5}(-28, 0, 12, -13, 0) + \frac{1}{5}a(38, 5, -12, 18, 0) + \frac{1}{5}b(-5, 0, 5, 0, 5) \\ &= \frac{1}{5}(-28, 0, 12, -13, 0) + a'(38, 5, -12, 18, 0) + b'(-5, 0, 5, 0, 5)\end{aligned}$$

Se escolhessemos outra matriz principal (por exemplo $P' = A[1, 2, 3; 1, 2, 3]$, em que $\det(P') = 18 \neq 0$), chegaríamos a uma solução, com x, y e z como incógnitas principais e em função de $(a, b) \in \mathbb{R}^2$ onde $a = v$ e $b = w$.

Exemplo 4.53 Considere-se o sistema

$$\begin{cases} x + 3y - 3z &= 1 \\ x + 10y - 10z &= 1 \\ 2x - y + z &= 2 \end{cases}$$

Tem-se

$$A = \begin{bmatrix} 1 & 3 & -3 \\ 1 & 10 & -10 \\ 2 & -1 & 1 \end{bmatrix} \text{ e } B = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

A matriz

$$P = A[1, 2; 1, 2] = \begin{array}{c} \begin{array}{cc} x & y \end{array} \\ \begin{bmatrix} 1 & 3 \\ 1 & 10 \end{bmatrix} \end{array}$$

é tal que $\det(P) = 7 \neq 0$ e seja p a sua ordem. Existe uma só matriz ($(m-p)(n-p) = 1$) orlada singular associada a P e extraída de A, que é a própria matriz A:

$$\det(A) = \begin{vmatrix} 1 & 3 & -3 \\ 1 & 10 & -10 \\ 2 & -1 & 1 \end{vmatrix} = 0$$

Deste modo, P é uma matriz *principal* de A e $r = p = 2 < m = 3$. Para verificar se o sistema é possível, vamos calcular o único ($m - r = 1$) determinante característico associado a P (relativo à 3ª equação)

$$C_3 = \begin{vmatrix} 1 & 3 & 1 \\ 1 & 10 & 1 \\ 2 & -1 & 2 \end{vmatrix} = 0$$

Como o único determinante característico é nulo, o sistema é *possível* e *simplesmente indeterminado* ($n - r = 3 - 2$) sendo equivalente ao seguinte sistema formado pelas 2 equações principais (a 1ª e a 2ª, cujos coeficientes contribuíram para P):

$$\begin{cases} x + 3y - 3z &= 1 \\ x + 10y - 10z &= 1 \end{cases}$$

As incógnitas principais são x e y (cujas colunas contribuíram para a formação de P); z é secundária. Pela proposição 4.18, para determinar a solução geral do sistema, calculamos os 2 determinantes seguintes introduzindo o parâmetro real arbitrário a, com $a = z$,

$$\det(\Delta_x) = \begin{vmatrix} 1 + 3a & 3 \\ 1 + 10a & 10 \end{vmatrix} = 7$$

$$\det(\Delta_y) = \begin{vmatrix} 1 & 1 + 3a \\ 1 & 1 + 10a \end{vmatrix} = 7a$$

O vector solução geral do sistema será então, para qualquer $a \in \mathbb{R}$,

$$(x, y, z) = \left(\frac{7}{7}, \frac{7a}{7}, a \right) = (1, a, a) = (1, 0, 0) + a(0, 1, 1); a \in \mathbb{R}$$

Vamos terminar esta secção, enunciando uma importante proposição relativa a sistemas homogéneos e que resulta do que vimos no capítulo 2

Proposição 4.19 *Seja n um inteiro > 0 e $A \in \mathbb{K}^{n,n}$ uma matriz de ordem n sobre $\mathbb{K}$. O sistema homogéneo $AX = O$, de n equações a n incógnitas, tem solução X não trivial (isto é, $X \neq O$) se e só se* $\det(A) = 0$.

Demonstração:

Um sistema homogéneo tem sempre a solução trivial $X = O$ e dizer que tem solução não trivial equivale a dizer que é indeterminado. Como vimos na capítulo 2, isto acontece sse a característica r da matriz A é estritamente inferior a n; por sua vez, esta condição equivale a $\det(A) = 0$. □

O seguinte corolário mostra uma consequência desta proposição que nos dá uma condição suficiente (mas não necessária) para que uma lista de funções seja linearmente independente:

Corolário 4.19.1 *Seja $I \subset \mathbb{R}$ um intervalo aberto, m um inteiro positivo e $\mathbb{K}$ um dos corpos $\mathbb{R}$ ou $\mathbb{C}$. Consideremos, no espaço vectorial $\mathscr{C}^{m-1}(I, \mathbb{K})$ sobre o corpo $\mathbb{K}$ constituído pelas funções (reais ou complexas) $m - 1$ vezes continuamente diferenciáveis em I, a lista de funções*

$$f = (f_1, f_2, \ldots, f_m).$$

Se o determinante

$$W_f(x) = \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) & \cdots & f_m(x) \\ f_1'(x) & f_2'(x) & f_3'(x) & \cdots & f_m'(x) \\ f_1''(x) & f_2''(x) & f_3''(x) & \cdots & f_m''(x) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_1^{(m-1)}(x) & f_2^{(m-1)}(x) & f_3^{(m-1)}(x) & \cdots & f_m^{(m-1)}(x) \end{vmatrix}$$

chamado **wronskiano**[13] *de* $(f_1, f_2, \ldots, f_m)$ *não for identicamente nulo em* I*, então a sequência* f *é linearmente independente.*

Demonstração:

Mostraremos que, se f é linearmente dependente, então $W_f(x) = 0$, para todo o $x \in I$: a dependência linear da sequência f equivale à existência de escalares $k_1, k_2, \ldots, k_m \in \mathbb{K}$ não todos nulos tais que, para todo o $x \in I$,

$$\sum_{i=1}^{m} k_i f_i(x) = 0.$$

Como as funções f_i são $m-1$ vezes diferenciáveis em I, derivando sucessivamente até à ordem $m-1$, obtém-se, para qualquer $x \in I$:

$$\begin{cases} \sum_{i=1}^{m} k_i f_i(x) = 0 \\ \sum_{i=1}^{m} k_i f_i'(x) = 0 \\ \sum_{i=1}^{m} k_i f_i''(x) = 0 \\ \cdots \\ \sum_{i=1}^{m} k_i f_i^{(m-1)}(x) = 0 \end{cases}$$

13 Em homenagem a *Wrónski, Józef Maria Hoëné*: matemático e filósofo polaco (Wolsztyn 1778 – Neuilly 1853).

Para cada $x \in I$, obtivémos assim um sistema de m equações lineares homogéneas nas m incógnitas k_i, com solução $(k_1, k_2, \ldots, k_m)$ não nula. O resultado é agora consequência imediata da proposição 4.19. □

Exemplo 4.54 A sequência de funções $f = (x, \text{sin}x, \text{cos}x)$ é linearmente independente, pois o respectivo wronskiano não é identicamente nulo em $I = \mathbb{R}$:

$$W_f(x) = \begin{vmatrix} x & \text{sin}x & \text{cos}x \\ 1 & \text{cos}x & -\text{sin}x \\ 0 & -\text{sin}x & -\text{cos}x \end{vmatrix} = -x$$

Exemplo 4.55 Seja $\omega > 0$. A sequência f de funções complexas definidas para $t \in \mathbb{R}$ por

$$f = (e^{-i2\omega t}, e^{-i\omega t}, 1, e^{i\omega t}, e^{i2\omega t})$$

é também linearmente independente, porque, como veremos, o seu wronskiano é negativo (e, neste caso, independente do argumento t):

$$\begin{aligned} W_f(t) &= \begin{vmatrix} e^{-i2\omega t} & e^{-i\omega t} & 1 & e^{i\omega t} & e^{i2\omega t} \\ -i2\omega e^{-i2\omega t} & -i\omega e^{-i\omega t} & 0 & i\omega e^{i\omega t} & i2\omega e^{i2\omega t} \\ -4\omega^2 e^{-i2\omega t} & -\omega^2 e^{-i\omega t} & 0 & -\omega^2 e^{i\omega t} & -4\omega^2 e^{i2\omega t} \\ i8\omega^3 e^{-i2\omega t} & i\omega^3 e^{-i\omega t} & 0 & -i\omega^3 e^{i\omega t} & -i8\omega^3 e^{i2\omega t} \\ 16\omega^4 e^{-i2\omega t} & \omega^4 e^{-i\omega t} & 0 & \omega^4 e^{i\omega t} & 16\omega^4 e^{i2\omega t} \end{vmatrix} \\ &= -\omega^{10} \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ -4 & -1 & 0 & -1 & -4 \\ 8 & 1 & 0 & -1 & -8 \\ 16 & 1 & 0 & 1 & 16 \end{vmatrix} \\ &= -4\omega^{10} \begin{vmatrix} -1 & -1 & 1 & 1 \\ -2 & -1 & -1 & -2 \\ 4 & 1 & -1 & -4 \\ 8 & 1 & 1 & 8 \end{vmatrix} = -4\omega^{10} \begin{vmatrix} 0 & -1 & 1 & 1 \\ -4 & -1 & -1 & -2 \\ 0 & 1 & -1 & -4 \\ 16 & 1 & 1 & 8 \end{vmatrix} \end{aligned}$$

$$\begin{aligned} W_f(t) &= -4\omega^{10} \left(4 \begin{vmatrix} -1 & 1 & 1 \\ 1 & -1 & -4 \\ 1 & 1 & 8 \end{vmatrix} - 16 \begin{vmatrix} -1 & 1 & 1 \\ -1 & -1 & -2 \\ 1 & -1 & -4 \end{vmatrix} \right) \\ &= -4\omega^{10}(-4 \times 6 + 16 \times 6) = -288\omega^{10} < 0 \end{aligned}$$

Exempo 4.56 Note-se que a recíproca do corolário 4.19.1 não é verdadeira: o facto de o wronskiano de $f = (f_1, f_2, \ldots, f_m)$ ser identicamente nulo em I não implica que f seja linearmente dependente em I, como mostramos no exemplo seguinte: sejam $f_1, f_2 \colon \mathbb{R} \to \mathbb{R}$ as funções definidas por

$$f_1(x) = x^2 \text{ e } f_2(x) = |x|x$$

e cujos gráficos se mostram na figura 4.11, juntamente com as derivadas de primeira ordem.

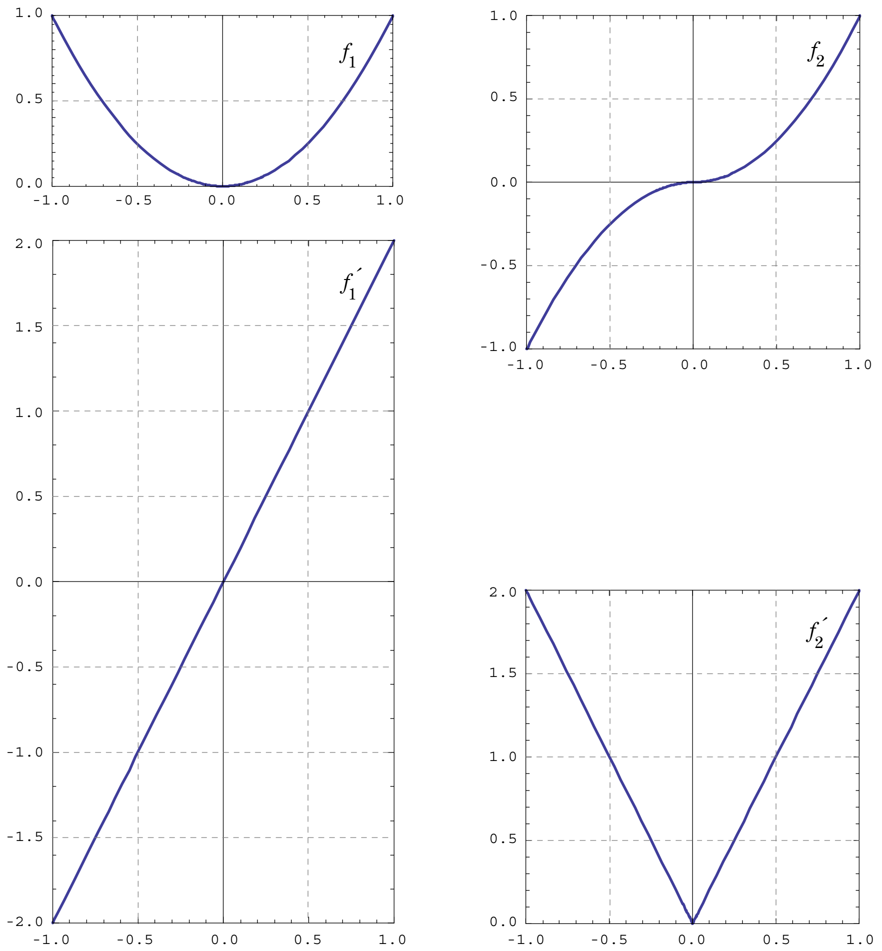


Fig. 4.11 – Gráficos de f_1, f_2, f_1' e f_2'.

As funções $f_1, f_2 \in \mathscr{C}^1(\mathbb{R}, \mathbb{R})$ e tem-se, para todo o $x \in \mathbb{R}$,

$$f_1'(x) = 2x \text{ e } f_2'(x) = 2|x|$$

Observemos que a expressão $2|x|$ para a derivada de f_2 é válida no ponto $x = 0$, visto que a derivada de f_2 na origem é

$$f_2'(0) = \lim_{x \to 0} \frac{f_2(x) - f_2(0)}{x} = \lim_{x \to 0} \frac{|x|x}{x} = \lim_{x \to 0} |x| = 0 = (2|x|)_{x=0}$$

O wronskiano de $f = (f_1, f_2)$ é, para qualquer $x \in \mathbb{R}$,

$$W_f(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{vmatrix} = \begin{vmatrix} x^2 & |x|x \\ 2x & 2|x| \end{vmatrix} = 2|x|x^2 - 2|x|x^2 = 0$$

Assim, $W_f(x)$ é identicamente nulo em $\mathbb{R}$, mas f é linearmente independente em $\mathbb{R}$, como se prova facilmente a seguir: se c_1, c_2 são reais tais que

$$\forall_{x \in \mathbb{R}} \; c_1 f_1(x) + c_2 f_2(x) = 0,$$

teremos, fazendo $x = 1$ e $x = -1$ respectivamente,

$$\begin{cases} c_1 + c_2 = 0 \\ c_1 - c_2 = 0 \end{cases}$$

A única solução do sistema homogéneo anterior é $c_1 = c_2 = 0$, o que prova que f é linearmente independente, mostrando, assim, que a recíproca do corolário 4.19.1 não é verdadeira: quando $W_f(x) = 0$ no intervalo aberto I, *f pode ser ou não linearmente dependente*. Todavia, demonstra-se que:

Seja $\mathbb{K}$ o corpo $\mathbb{R}$ ou $\mathbb{C}$, *$I \subset \mathbb{R}$ um intervalo aberto* e $f = (f_1, f_2, \ldots, f_m)$ *uma lista de funções de I em $\mathbb{K}$ pertencentes a $\mathscr{C}^{m-1}(I, \mathbb{K})$. Se o wronskiano de f é identicamente nulo em I, então existe pelo menos um subintervalo aberto $J \subset I$ tal que a lista das **restrições** das funções f_i ao intervalo J é linearmente dependente.*

Por exemplo, no caso das funções (f_1, f_2) apresentado neste exemplo, as suas restrições ao intervalo $\mathbb{R}^-$ (e também ao intervalo $\mathbb{R}^+$) são linearmente dependentes, o que é óbvio.

Exemplo 4.57 A sequência de funções $f = (\cos^2 x, \sin^2 x, 1)$ é linearmente dependente, visto que $\cos^2 x + \sin^2 x - 1 = 0$. Portanto, pelo corolário 4.19.1, o wronskiano de f deverá ser nulo, para todo o $x \in \mathbb{R}$ e, de facto, assim é:

$$W_f(x) = \begin{vmatrix} \cos^2 x & \sin^2 x & 1 \\ -\sin 2x & \sin 2x & 0 \\ -2\cos 2x & 2\cos 2x & 0 \end{vmatrix} = -2\sin 2x \cos 2x + 2\sin 2x \cos 2x = 0$$

4.13 Matriz adjunta e inversão de matrizes quadradas

O teorema de Laplace restrito tem aplicação no cálculo da inversa de uma matriz quadrada regular. Para analisarmos esta questão, devemos definir a noção de *matriz adjunta* de uma matriz quadrada (regular ou não).

Definição 4.13 – Matriz adjunta – *Seja $A = [\,a_{ij}\,] \in \mathbb{K}^{n,n}$ uma matriz de ordem $n > 1$. A matriz $\widehat{A} \in \mathbb{K}^{n,n}$ cujos elementos são os cofactores dos a_{ij} é chamada **matriz complementar** da matriz A. A transposta de $\widehat{A}$ chama-se **matriz adjunta** de A e designa-se por* $\operatorname{adj}(A) \in \mathbb{K}^{n,n}$*:*

$$\widehat{A} = [\,\operatorname{cof}(a_{ij})\,]; \quad \operatorname{adj}(A) = \left(\widehat{A}\right)^{\mathrm{T}} \tag{4.89}$$

Daqui resulta que a relação entre os elementos de $B = [\,b_{ij}\,] = \operatorname{adj}(A)$ e de $A = [\,a_{ij}\,]$ é:

$$b_{ij} = \operatorname{cof}(a_{ji});\ 1 \leq i \leq n \wedge 1 \leq j \leq n \tag{4.90}$$

Proposição 4.20 *Seja $A = [\,a_{ij}\,] \in \mathbb{K}^{n,n}$ uma matriz de ordem $n > 1$. Então:*

i) $$\widehat{\left(A^{\mathrm{T}}\right)} = \left(\widehat{A}\right)^{\mathrm{T}} \tag{4.91}$$

ii) $\text{adj}(A^{\mathrm{T}}) = (\text{adj}(A))^{\mathrm{T}}$ (4.92)

Demonstração:

i) Para provar a primeira igualdade, determinemos o elemento da linha i e coluna j de ambos os membros e verifiquemos que são iguais:

$$\begin{aligned} \big(\widehat{(A^{\mathrm{T}})}\big)_{ij} &= \text{cof}\big((A^{\mathrm{T}})_{ij}\big) = \text{cof}(a_{ji}) \\ \big((\widehat{A})^{\mathrm{T}}\big)_{ij} &= \big(\widehat{A}\big)_{ji} = \text{cof}(a_{ji}) \end{aligned}$$

ii) Basta fazer uso da igualdade (4.91) e da definição de adjunta:

$$\text{adj}\big(A^{\mathrm{T}}\big) = \Big(\widehat{(A^{\mathrm{T}})}\Big)^{\mathrm{T}} = \Big((\widehat{A})^{\mathrm{T}}\Big)^{\mathrm{T}} = (\text{adj}(A))^{\mathrm{T}}$$ □

O teorema de Laplace restrito vai implicar uma importante propriedade da matriz adjunta, com base na qual é depois imediato o cálculo da inversa de uma matriz regular A.

Proposição 4.21 – Propriedade fundamental da adjunta – *Seja* $A = [\,a_{ij}\,] \in \mathbb{K}^{n,n}$ *uma matriz de ordem* $n > 1$ *e* adj(A) *a sua matriz adjunta. Então:*

i) A *e a sua matriz adjunta são permutáveis e o seu produto é uma matriz escalar*

$$\text{adj}(A)A = A\text{adj}(A) = \det(A)I_n \tag{4.93}$$

ii) *Se* A *é singular, fica*

$$\text{adj}(A)A = A\text{adj}(A) = O_n \tag{4.94}$$

iii) *Se* A *é regular, a* ***inversa*** *de* A *vem dada por*

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A) \tag{4.95}$$

Esta igualdade mostra que o elemento c_{ij} *da linha* i *e da coluna* j *da matriz* $C = A^{-1}$ *inversa de* A *se calcula por:*

$$c_{ij} = \frac{\text{cof}(a_{ji})}{\det(A)};\ 1 \leq i, j \leq n \tag{4.96}$$

Demonstração:

i) A igualdade matricial adj$(A)A = \det(A)I_n$ equivale às n^2 igualdades escalares

$$\sum_{k=1}^{n} a_{kj}\text{cof}(a_{ki}) = \delta_{ij}\det(A);\ 1 \leq i, j \leq n \tag{1}$$

e a igualdade matricial $A\text{adj}(A) = \det(A)I_n$ equivale às n^2 igualdades escalares

$$\sum_{k=1}^{n} a_{ik}\text{cof}(a_{jk}) = \delta_{ij}\det(A);\ 1 \leq i, j \leq n \tag{2}$$

Provemos as igualdades (1):

- Se $j = i$, então (1) transforma-se em

$$\sum_{k=1}^{n} a_{ki}\text{cof}(a_{ki}) = \underbrace{\delta_{ii}}_{=1}\det(A) = \det(A);\ 1 \leq i \leq n$$

o que constitui exactamente o desenvolvimento pelo teorema de Laplace restrito de determinante de A pela coluna i.

- Se $j \neq i$, sejam $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ os vectores-coluna de A e consideremos a matriz $A' = [\,a'_{ij}\,]$ obtida de A substituindo nesta a coluna i pela coluna j

$$A' = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_j, \ldots, \vec{a}_j, \ldots, \vec{a}_n]$$

Observemos que os cofactores dos elementos da coluna i de A' são iguais aos da mesma coluna de A (visto que os citados cofactores não dependem dos elementos da coluna i e que os restantes elementos de A' são iguais aos de A). Isto permite concluir que os elementos a'_{ki} da coluna i de A' satisfazem as igualdades:

$$\begin{array}{l} a'_{ki} = a_{kj} \\ \text{cof}(a'_{ki}) = \text{cof}(a_{ki}) \end{array}\ ;\ 1 \leq k \leq n,$$

Como det é uma forma alternada será $\det(A') = 0$ (colunas i e j de A' iguais, por construção). Usando o teorema de Laplace restrito para calcular $\det(A')$ (que é nulo!) segundo a coluna i, obtém-se

$$0 = \det(A') = \sum_{k=1}^{n} a'_{ki}\text{cof}(a'_{ki}) = \sum_{k=1}^{n} a_{kj}\text{cof}(a_{ki});\ j \neq i$$

o que termina a demonstração para $j \neq i$. Provámos, assim, a igualdade $\text{adj}(A)A = \det(A)I_n$ e, para provar $A\text{adj}(A) = \det(A)I_n$, basta atender a (4.92) e aplicar a igualdade que ora demonstrámos à matriz A^{T}:

$$\begin{aligned} A\text{adj}(A) &= \big((A\text{adj}(A))^{\mathrm{T}}\big)^{\mathrm{T}} = \big((\text{adj}(A))^{\mathrm{T}}A^{\mathrm{T}}\big)^{\mathrm{T}} = \big(\text{adj}\big(A^{\mathrm{T}}\big)A^{\mathrm{T}}\big)^{\mathrm{T}} \\ &= \big(\det\big(A^{\mathrm{T}}\big)I_n\big)^{\mathrm{T}} = \det\big(A^{\mathrm{T}}\big)I_n^{\mathrm{T}} = \det(A)I_n \end{aligned}$$

Também seria possível provar a igualdade $A\text{adj}(A) = \det(A)I_n$ provando directamente as igualdades (2), por meio de raciocínio semelhante ao que usámos, mas aplicado a uma linha de A.

ii) Se A é *singular*, será $\det(A) = 0$ e (4.94) resultará imediatamente de (4.93).

iii) Se A é *regular*, será $\det(A) \neq 0$ e multiplicando (4.93) por $\frac{1}{\det(A)}$, obtém-se

$$\Big(\frac{1}{\det(A)}\text{adj}(A)\Big)A = A\Big(\frac{1}{\det(A)}\text{adj}(A)\Big) = I_n$$

e estas igualdades mostram que a matriz $\frac{1}{\det(A)}\text{adj}(A)$ é a *inversa* de A

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A);\ \det(A) \neq 0$$

que é a igualdade (4.95) desejada. Quanto a (4.96) é obviamente equivalente a (4.95). □

Observações:

- Podemos resumir a propriedade fundamental da matriz adjunta dizendo que:

– É nula a soma dos produtos dos elementos de uma fila de um determinante pelos cofactores dos elementos correspondentes de qualquer outra fila paralela.

– É igual ao determinante da matriz a soma dos produtos dos elementos de uma sua fila pelos seus próprios cofactores (teorema de Laplace restrito).

- Comparando (4.95) com o método de condensação do capítulo 2, mais uma vez se verifica que o método de condensação é global, isto é, não é possível calcular por condensação apenas um elemento da matriz inversa enquanto que (4.96) nos dá qualquer elemento da inversa sem calcularmos os restantes.

- A fórmula (4.95) implica o cálculo de n^2 determinantes de ordem $n-1$ (para o cálculo da adjunta) e ainda um determinante de ordem n o que mostra que, para valores de $n > 3$, o método de condensação é muito menos trabalhoso.

- O *package* **`ALGA`Determinantes``**, exposto na secção 4.16, implementa as funções **`Adj`** e **`MatrizInversa`** para calcular a adjunta e a inversa de uma matriz quadrada, usando a matriz adjunta.

Exemplo 4.58 Consideremos a matriz

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & -1 & 4 \\ 3 & -2 & 1 \end{bmatrix}$$

O determinante de A é

$$\det(A) = 39 \neq 0$$

A matriz é regular e, para calcular a inversa, determinemos a adjunta

$$\text{adj}(A) = \left(\widehat{A}\right)^{\mathrm{T}} = \begin{bmatrix} 7 & 10 & -1 \\ 1 & 7 & 11 \\ 10 & -8 & -7 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 7 & 1 & 10 \\ 10 & 7 & -8 \\ -1 & 11 & -7 \end{bmatrix}$$

Por fim, (4.95) dá a inversa

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A) = \frac{1}{39}\begin{bmatrix} 7 & 1 & 10 \\ 10 & 7 & -8 \\ -1 & 11 & -7 \end{bmatrix}$$

Podemos verificar a propriedade fundamental da matriz adjunta

$$\text{adj}(A)A = \begin{bmatrix} 7 & 1 & 10 \\ 10 & 7 & -8 \\ -1 & 11 & -7 \end{bmatrix} \begin{bmatrix} 1 & 3 & -2 \\ 2 & -1 & 4 \\ 3 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 39 & 0 & 0 \\ 0 & 39 & 0 \\ 0 & 0 & 39 \end{bmatrix} = 39 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A\text{adj}(A) = \begin{bmatrix} 1 & 3 & -2 \\ 2 & -1 & 4 \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} 7 & 1 & 10 \\ 10 & 7 & -8 \\ -1 & 11 & -7 \end{bmatrix} = \begin{bmatrix} 39 & 0 & 0 \\ 0 & 39 & 0 \\ 0 & 0 & 39 \end{bmatrix} = 39 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Exemplo 4.59 Consideremos a matriz

$$A = \begin{bmatrix} 5 & -1 & 0 & 2 & 3 \\ 1 & 0 & 1 & 2 & -1 \\ -1 & 2 & 0 & 3 & 1 \\ 0 & 2 & -3 & 2 & 0 \\ 1 & -1 & 0 & 1 & 0 \end{bmatrix}$$

O determinante de A pode ser calculado pelo método misto: primeiro, condensamos a 3ª coluna (a fila da matriz que contém mais zeros) com 1 como pivot; de seguida, desenvolvemos pelo teorema de Laplace restrito segundo essa coluna:

$$\det(A) = \begin{vmatrix} 5 & -1 & 0 & 2 & 3 \\ 1 & 0 & 1 & 2 & -1 \\ -1 & 2 & 0 & 3 & 1 \\ 0 & 2 & -3 & 2 & 0 \\ 1 & -1 & 0 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 5 & -1 & 0 & 2 & 3 \\ 1 & 0 & 1 & 2 & -1 \\ -1 & 2 & 0 & 3 & 1 \\ 3 & 2 & 0 & 8 & -3 \\ 1 & -1 & 0 & 1 & 0 \end{vmatrix} = -\begin{vmatrix} 5 & -1 & 2 & 3 \\ -1 & 2 & 3 & 1 \\ 3 & 2 & 8 & -3 \\ 1 & -1 & 1 & 0 \end{vmatrix}$$

Neste último determinante, condensemos, agora, a 4ª linha com o seu primeiro elemento 1 como pivot e apliquemos o teorema de Laplace restrito a essa linha, seguido da regra de Sarrus:

$$\det(A) = -\begin{vmatrix} 5 & 4 & -3 & 3 \\ -1 & 1 & 4 & 1 \\ 3 & 5 & 5 & -3 \\ 1 & 0 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 4 & -3 & 3 \\ 1 & 4 & 1 \\ 5 & 5 & -3 \end{vmatrix} = -137 \neq 0$$

O resultado mostra que A é matriz regular; determinemos apenas o elemento c_{23} da linha 2 e coluna 3 da inversa A^{-1}:

$$c_{23} = \; = \frac{\text{cof}(a_{32})}{\det(A)} = \frac{6}{-137} = -\frac{6}{137}$$

Em que usámos o método misto para calcular $\text{cof}(a_{32})$:

$$\text{cof}(a_{32}) = (-1)^{3+2}\det(A(3;2)) = -\begin{vmatrix} 5 & 0 & 2 & 3 \\ 1 & 1 & 2 & -1 \\ 0 & -3 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{vmatrix} = -\begin{vmatrix} 8 & 3 & 8 & 0 \\ 1 & 1 & 2 & -1 \\ 0 & -3 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{vmatrix}$$

$$= \begin{vmatrix} 8 & 3 & 8 \\ 0 & -3 & 2 \\ 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 8 & 3 & 0 \\ 0 & -3 & 2 \\ 1 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 3 & 0 \\ -3 & 2 \end{vmatrix} = 6$$

4.14 Fórmula de Cauchy

O teorema de Laplace restrito vai permitir a dedução de uma fórmula devida a Cauchy[(14)] e que permite exprimir um determinante de ordem n num determinante de ordem $n-1$ e em $(n-1)^2$ determinantes de ordem $n-2$. É o que faremos na demonstração da

Proposição 4.22 – Fórmula de Cauchy – *Seja $n > 2$ um inteiro e $A \in \mathbb{K}^{n,n}$ uma matriz quadrada de ordem n e fragmentemos A antes da última linha e da última coluna:*

$$A = \left[\begin{array}{c|c} A' & C \\ \hline L & a_{nn} \end{array}\right]$$

Então, o determinante de A é dado por:

$$\det(A) = a_{nn}\det(A(n;n)) - \sum_{i,j=1}^{n-1} (-1)^{i+j}\det(A'(i;j))a_{in}a_{nj} \tag{4.97}$$

ou ainda:

$$\det(A) = a_{nn}\det(A') - L\,\text{adj}(A')C \tag{4.98}$$

Demonstração:

Comecemos por desenvolver o determinante de A pela coluna n, usando o teorema de Laplace restrito e separemos a parcela contendo a_{nn}:

$$\det(A) = a_{nn}\det(A(n;n)) + \sum_{i=1}^{n-1} a_{in}(-1)^{i+n}\det(A(i;n)) \tag{1}$$

As $n-1$ matrizes $(A(i;n))_{1\leq i\leq n-1}$, de ordem $n-1$, contêm os elementos $(a_{nj})_{1\leq j\leq n-1}$ da linha n de A e que constituem sempre a linha $n-1$ dos $A(i;n)$.

---

14 *Cauchy, Augustin Louis*: matemático francês (Paris 1789 – Sceaux 1857).

Aplicando agora o teorema de Laplace restrito à linha $n-1$ (a última) de cada uma dessas matrizes $A(i;n)$ e observando que $A(i,n;j,n) = A'(i;j)$, temos, para todo o $1 \leq i \leq n-1$:

$$\det(A(i;n)) = \sum_{j=1}^{n-1} a_{nj}(-1)^{n-1+j}\det(A(i,n;j,n)) = \sum_{j=1}^{n-1} a_{nj}(-1)^{n-1+j}\det(A'(i;j)) \quad (2)$$

Substituindo (2) em (1), obtém-se a primeira das igualdades desejadas

$$\begin{aligned}\det(A) &= a_{nn}\det(A(n;n)) + \sum_{i=1}^{n-1} a_{in}(-1)^{i+n}\left(\sum_{j=1}^{n-1} a_{nj}(-1)^{n-1+j}\det(A'(i;j))\right)\\ &= a_{nn}\det(A(n;n)) + \sum_{i,j=1}^{n-1}(-1)^{i+j+2n-1}\det(A'(i;j))a_{in}a_{nj}\\ &= a_{nn}\det(A(n;n)) - \sum_{i,j=1}^{n-1}(-1)^{i+j}\det(A'(i;j))a_{in}a_{nj}\end{aligned}$$

Quanto a (4.98), observemos que:

- $A' = A(n;n)$.
- $(-1)^{i+j}\det(A'(i;j))$ é o cofactor de a_{ij} em relação à matriz A' e, consequentemente, a matriz $\left[(-1)^{i+j}\det(A'(i;j))\right]_{1\leq i,j\leq n-1}$, de ordem $n-2$, é a matriz $\widehat{A'}$ complementar de A'.
- $(a_{in})_{1\leq i\leq n-1}$ constitui a coluna C.
- $(a_{nj})_{1\leq j\leq n-1}$ constitui a linha L.

Deste modo, fica

$$\begin{aligned}\det(A) &= a_{nn}\det(A') - C^{\mathrm{T}}\widehat{A'}L^{\mathrm{T}} = \left(a_{nn}\det(A') - C^{\mathrm{T}}\widehat{A'}L^{\mathrm{T}}\right)^{\mathrm{T}}\\ &= a_{nn}\det(A') - L\left(\widehat{A'}\right)^{\mathrm{T}}C = a_{nn}\det(A') - L\,\mathrm{adj}(A')C\end{aligned}$$

o que termina a demonstração. $\square$

Observações:

- A fórmula de Cauchy permite reduzir o cálculo de um determinante de ordem n ao cálculo de um determinante de ordem $n-1$ (o determinante de A') e de mais $(n-1)^2$ determinantes de ordem $n-2$ (no cálculo de adj(A')).

■ Podemos calcular o número de termos na fórmula de Cauchy, para o que se indicam os números de parcelas de cada subexpressão naquela fórmula:

Expressão	$a_{nn}\text{det}(A(n;n))$	$\sum_{i,j=1}^{n-1}(-1)^{i+j}\text{det}(A'(i;j))a_{in}a_{nj}$	$\text{det}(A)$
nº de parcelas	$(n-1)!$	$(n-1)^2(n-2)!$	$n!$

O total de parcelas é:

$$(n-1)!+(n-1)^2(n-2)! = (n-1)!+(n-1)(n-1)! = (n-1)!(1+n-1) = n!$$

Reencontrámos, assim, o número de parcelas de (4.64), como era de esperar.

■ Quando houver um elemento nulo no determinante de A podemos, por duas operações elementares de tipo 1, levá-lo à posição de a_{nn} o que fará com que a primeira parcela da fórmula de Cauchy se anule, poupando o cálculo de det(A').

■ Quando a matriz A for simétrica, é $L = C^{\mathrm{T}}$ e a fórmula de Cauchy transforma-se em

$$\text{det}(A) = a_{nn}\text{det}(A') - C^{\mathrm{T}}\text{adj}(A')C$$

A segunda parcela da expressão no 2º membro da igualdade anterior é uma forma quadrática em C, com matriz igual a adj(A').

■ A fórmula de Cauchy é susceptível de uma generalização que consiste em usar uma linha i e uma coluna j quaisquer da matriz A para aplicar o teorema de Laplace restrito. Fixemos um par $(i,j)_{1\le i,j\le n}$: trocando sucessivamente a linha i com cada uma das seguintes ($n-i$ trocas) e a coluna j com cada uma das seguintes ($n-j$ trocas), levamos o elemento a_{ij} a ocupar a posição (n,n) mantendo a ordem das linhas de índices $r \neq i$ e das restantes colunas de índices $s \neq j$; assim, devido à anti-simetria, fica

$$\text{det}(A) = (-1)^{(n-i)+(n-j)}\left|\begin{array}{c|c} A(i;j) & A(i;j] \\ \hline A[i;j) & a_{ij}\end{array}\right|$$

Aplicando agora a fórmula de Cauchy (4.98) ao determinante anterior e observando que $(-1)^{(n-i)+(n-j)} = (-1)^{2n-(i+j)} = (-1)^{2n}\times(-1)^{2(i+j)}\times(-1)^{-(i+j)} = (-1)^{i+j}$, vem

$$\begin{aligned}\text{det}(A) &= (-1)^{i+j}(a_{ij}\text{det}(A(i;j)) - A[i;j)\,\text{adj}(A(i;j))\,A(i;j]\,) \\ &= a_{ij}\text{cof}(a_{ij}) - (-1)^{i+j}A[i;j)\,\text{adj}(A(i;j))\,A(i;j]\end{aligned}$$

Obtivemos, portanto, uma forma mais geral da fórmula de Cauchy (4.98), válida agora para quaisquer $1 \le i,j \le n$ e da qual (4.98) é um caso particular (com $i=j=n$):

$$\text{det}(A) = a_{ij}\text{cof}(a_{ij}) - (-1)^{i+j}A[i;j)\,\text{adj}(A(i;j))\,A(i;j] \tag{4.98.1}$$

■ O *package* **`ALGA`Determinantes``** fornece as implementações (4.98) e (4.98.1) da fórmula de Cauchy nas funções denominadas **`Cauchy`** (ver secção 4.16).

Exemplo 4.60 Calculemos de novo o determinante do exemplo 4.42

$$\det(A) = \left|\begin{array}{cccc|c} -1 & 2 & 0 & -2 & 3 \\ 4 & 3 & 1 & -2 & 2 \\ -3 & 0 & 2 & 4 & 1 \\ 2 & -5 & 1 & 0 & 3 \\ \hline 4 & 1 & -1 & 2 & 0 \end{array}\right|$$

usando, desta vez, a fórmula de Cauchy e começando por calcular adj(A')

$$\text{adj}(A') = \left(\widehat{A'}\right)^{\mathrm{T}} = \begin{bmatrix} -52 & -6 & 74 & -76 \\ 28 & 18 & 34 & 4 \\ -12 & 6 & 54 & 12 \\ -4 & -30 & 50 & -28 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} -52 & 28 & -12 & -4 \\ -6 & 18 & 6 & -30 \\ 74 & 34 & 54 & 50 \\ -76 & 4 & 12 & -28 \end{bmatrix}$$

A fórmula de Cauchy dá, finalmente,

$$\det(A) = 0 \begin{vmatrix} -1 & 2 & 0 & -2 \\ 4 & 3 & 1 & -2 \\ -3 & 0 & 2 & 4 \\ 2 & -5 & 1 & 0 \end{vmatrix} - [\,4 \;\; 1 \;\; -1 \;\; 2\,] \begin{bmatrix} -52 & 28 & -12 & -4 \\ -6 & 18 & 6 & -30 \\ 74 & 34 & 54 & 50 \\ -76 & 4 & 12 & -28 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \\ 3 \end{bmatrix}$$

$$= -[\,-440 \;\; 104 \;\; -72 \;\; -152\,] \begin{bmatrix} 3 \\ 2 \\ 1 \\ 3 \end{bmatrix} = 1640$$

4.15 Derivada de um determinante

Se os elementos de uma matriz real quadrada forem funções reais (respectivamente, complexas) de uma variável real, o determinante dessa matriz será uma função real (respectivamente, complexa) da mesma variável. Vamos, nesta secção, estudar a diferenciabilidade desta função, supondo a diferenciabilidade das funções elementos da matriz.

Proposição 4.23 – Derivada de um determinante – *Seja $I \subset \mathbb{R}$ um intervalo aberto em $\mathbb{R}$, $n \geq 1$ um número inteiro e $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Seja ainda $(\vec{a}_j)_{1 \leq j \leq n} \colon I \to \mathbb{K}^n; x \mapsto \vec{a}_j(x)$ uma família de n funções vectoriais definidas em I e com valores em $\mathbb{K}^n$ diferenciáveis num ponto $x_0 \in I$. Nestas condições, a função $\Delta \colon I \to \mathbb{K}$ definida por*

$$\Delta(x) = \det(\vec{a}_1(x), \vec{a}_2(x), \ldots, \vec{a}_j(x), \ldots, \vec{a}_n(x)) \tag{4.99}$$

é diferenciável em x_0 e tem-se

$$\Delta'(x_0) = \sum_{j=1}^{n} \det\bigl(\vec{a}_1(x_0), \vec{a}_2(x_0), \ldots, \vec{a}_j'(x_0), \ldots, \vec{a}_n(x_0)\bigr) \tag{4.100}$$

Demonstração:

Seja $\mathbb{K}$ um dos corpos $\mathbb{R}$ ou $\mathbb{C}$ e, antes de mais, observe-se que dizer que as n funções vectoriais $\vec{a}_j\colon I \to \mathbb{K}^n$ são diferenciáveis em x_0 equivale a dizer que são diferenciáveis nesse ponto as suas n funções componentes escalares $(a_{ij})_{1\le i,j\le n}$ de variável real, tendo-se

$$\vec{a}'_j(x_0) = \big(a'_{1j}(x_0), a'_{2j}(x_0), \dots, a'_{nj}(x_0)\big);\ 1 \le j \le n$$

Mas, segundo (4.64),

$$\Delta(x) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\, a_{\sigma_1 1}(x)\, a_{\sigma_2 2}(x)\cdots a_{\sigma_n n}(x)$$

Todas as funções que intervêm na expressão anterior estão definidas em I e com valores em $\mathbb{K}$ e são diferenciáveis em x_0, donde Δ é também diferenciável neste ponto e tem-se, atendendo às regras de derivação da soma e do produto:

$$\begin{aligned}
\Delta'(x_0) &= \sum_{\sigma\in\mathfrak{S}_n} \Big(\varepsilon(\sigma)\, a'_{\sigma_1 1}(x_0)\, a_{\sigma_2 2}(x_0)\cdots a_{\sigma_n n}(x_0) + \varepsilon(\sigma)\, a_{\sigma_1 1}(x_0)\, a'_{\sigma_2 2}(x_0)\cdots a_{\sigma_n n}(x_0) + \\
&\qquad \cdots + \varepsilon(\sigma)\, a_{\sigma_1 1}(x_0)\, a_{\sigma_2 2}(x_0)\cdots a'_{\sigma_n n}(x_0)\Big) \\
&= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\, a'_{\sigma_1 1}(x_0)\, a_{\sigma_2 2}(x_0)\cdots a_{\sigma_n n}(x_0) + \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\, a_{\sigma_1 1}(x_0)\, a'_{\sigma_2 2}(x_0)\cdots a_{\sigma_n n}(x_0) + \\
&\qquad \cdots + \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\, a_{\sigma_1 1}(x_0)\, a_{\sigma_2 2}(x_0)\cdots a'_{\sigma_n n}(x_0) \\
&= \det\big(\vec{a}'_1(x_0), \vec{a}_2(x_0), \dots, \dots, \vec{a}_n(x_0)\big) + \det\big(\vec{a}_1(x_0), \vec{a}'_2(x_0), \dots, \dots, \vec{a}_n(x_0)\big) + \\
&\qquad \cdots + \det\big(\vec{a}_1(x_0), \vec{a}_2(x_0), \dots, \dots, \vec{a}'_n(x_0)\big) \\
&= \sum_{j=1}^{n} \det\big(\vec{a}_1(x_0), \vec{a}_2(x_0), \dots, \vec{a}'_j(x_0), \dots, \vec{a}_n(x_0)\big)
\end{aligned}$$

Esta é a expressão que queríamos provar. □

Observações:

- A expressão (4.100) diz-nos que a derivada de um determinante de ordem n é a soma dos n determinantes obtidos derivando sucessivamente as colunas $1, 2, \dots, n$ e deixando inalteradas as restantes $n-1$ colunas:

$$\Delta'(x_0) = \begin{vmatrix} a'_{11}(x_0) & a_{12}(x_0) & \cdots & a_{1n}(x_0) \\ a'_{21}(x_0) & a_{22}(x_0) & \cdots & a_{2n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ a'_{n1}(x_0) & a_{n2}(x_0) & \cdots & a_{nn}(x_0) \end{vmatrix} + \begin{vmatrix} a_{11}(x_0) & a'_{12}(x_0) & \cdots & a_{1n}(x_0) \\ a_{21}(x_0) & a'_{22}(x_0) & \cdots & a_{2n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(x_0) & a'_{n2}(x_0) & \cdots & a_{nn}(x_0) \end{vmatrix} +$$
$$\cdots + \begin{vmatrix} a_{11}(x_0) & a_{12}(x_0) & \cdots & a'_{1n}(x_0) \\ a_{21}(x_0) & a_{22}(x_0) & \cdots & a'_{2n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(x_0) & a_{n2}(x_0) & \cdots & a'_{nn}(x_0) \end{vmatrix}$$

■ Como o determinante de uma matriz é também o determinante dos seus vectores-linha na base canónica de $\mathbb{K}^n$, podemos também derivar um determinante *derivando por linhas*:

$$\Delta'(x_0) = \begin{vmatrix} a'_{11}(x_0) & a'_{12}(x_0) & \cdots & a'_{1n}(x_0) \\ a_{21}(x_0) & a_{22}(x_0) & \cdots & a_{2n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(x_0) & a_{n2}(x_0) & \cdots & a_{nn}(x_0) \end{vmatrix} + \begin{vmatrix} a_{11}(x_0) & a_{12}(x_0) & \cdots & a_{1n}(x_0) \\ a'_{21}(x_0) & a'_{22}(x_0) & \cdots & a'_{2n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(x_0) & a_{n2}(x_0) & \cdots & a_{nn}(x_0) \end{vmatrix} +$$
$$\cdots + \begin{vmatrix} a_{11}(x_0) & a_{12}(x_0) & \cdots & a_{1n}(x_0) \\ a_{21}(x_0) & a_{22}(x_0) & \cdots & a_{2n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ a'_{n1}(x_0) & a'_{n2}(x_0) & \cdots & a'_{nn}(x_0) \end{vmatrix}$$

■ A partir da regra anterior, podem calcular-se as derivadas de ordem superior, aplicando-a à função Δ' e assim sucessivamente.

Exemplo 4.61 A derivada da função $\Delta\colon \mathbb{R} \to \mathbb{R}$ definida pelo determinante de 2ª ordem

$$\Delta(x) = \begin{vmatrix} \sin(x^2) & \frac{x}{1+x^2} \\ \ln(1+x^2) & \cos(x) \end{vmatrix}$$

poderá ser calculada por (4.100), derivando por colunas, vindo:

$$\begin{aligned} \Delta'(x) &= \begin{vmatrix} (\sin(x^2))' & \frac{x}{1+x^2} \\ (\ln(1+x^2))' & \cos(x) \end{vmatrix} + \begin{vmatrix} \sin(x^2) & \left(\frac{x}{1+x^2}\right)' \\ \ln(1+x^2) & (\cos(x))' \end{vmatrix} \\ &= \begin{vmatrix} 2x\cos(x^2) & \frac{x}{1+x^2} \\ \frac{2x}{1+x^2} & \cos(x) \end{vmatrix} + \begin{vmatrix} \sin(x^2) & \frac{1-x^2}{(1+x^2)^2} \\ \ln(1+x^2) & -\sin(x) \end{vmatrix} \\ &= 2x\cos(x)\cos\left(x^2\right) - \frac{2x^2}{(1+x^2)^2} - \sin(x)\sin\left(x^2\right) - \frac{1-x^2}{(1+x^2)^2}\ln\left(1+x^2\right) \end{aligned}$$

4.16 Anexos: determinantes e o MATHEMATICA©

MATHEMATICA©

Neste anexo apresentamos a implementação de um *package* para o MATHEMATICA© que é útil para resolver alguns problemas relacionados com a teoria das funções multilineares e determinantes. Algumas das funções necessárias existem já no MATHEMATICA©, nomeadamente as seguintes:

- **Permutations[s]**

Determina as $n!$ permutações do conjunto (lista) **s**, com n elementos.

- **Det[mat]**

Calcula o determinante da matriz quadrada **mat**.

- **KSubsets[s,p]**

Determina todos os subconjuntos formados por **p** elementos do conjunto **s** de n elementos (em número de $\binom{n}{p}$ se $0 \leq p \leq n$, ou 0 se $p > n$).

- **Cofactor[mat,{i,j}]**

Determina o cofactor do elemento da linha **i** e coluna **j** da matriz **mat**.

Esta função está implementada no *package* **DiscreteMath`Combinatorica`**; no entanto, o *package* **ALGA`Determinantes`** define uma função **Cof** mais geral do que esta. Além desta, este package implementa as seguintes funções adicionais:

- **PermutacaoQ[lista]**

Determina se a lista dada é uma permutação de $\{1, 2, \ldots, n\}$, para algum n.

- **TransposicaoQ[lista]**

Determina se a lista é uma transposição.

- **ElementarQ[lista]**

Determina se a lista é uma transposição elementar.

- **PermInv[p]**

Determina a permutação inversa de **p**.

- **Composicao[p1,p2]**

Determina a composta de **p1** com **p2** (**p1** $\circ$ **p2**).

- **Decomposicao[p]**

Devolve uma lista de transposições, cuja composição (pela ordem resultante) é igual a **p**.

- **Epsilon[p]**

Determina a paridade ou sinal da permutação **p**. Equivale à função **Signature** já existente.

- **`CondensacaoDet[mat]`**

Calcula o determinante da matriz quadrada **`mat`**, pelo método de condensação.

- **`MatrizComplementar[mat,linhas,colunas]`**

Determina a submatriz de **`mat`** obtida por eliminação da linhas indicadas na lista **`linhas`** e das colunas indicadas na lista **`colunas`**.

- **`MenorComplementar[mat,linhas,colunas]`**

Determina o menor complementar do menor da matriz **`mat`** formado pelas linhas e colunas indicadas nas listas **`linhas`** e **`colunas`**.

- **`Paridade[linhas,colunas]`**

Determina a paridade de um menor formado pelas filas indicadas nas listas **`linhas`** e **`colunas`**.

- **`Cof[mat,linhas,colunas]`**

Determina o cofactor do menor da matriz **`mat`** formado pelas linhas e colunas indicadas nas listas **`linhas`** e **`colunas`**.

- **`LaplaceDet[mat,filas,Colunas->`**`Opção`**`]`**

Calcula o determinante de **`mat`** aplicando o teorema de Laplace às filas indicadas em **`filas`**. A opção **`Colunas->True`** ou **`Colunas->False`** indica se essas filas são linhas ou colunas (por defeito é **`Colunas->True`**).

- **`RecursivoDet[mat]`**

Calcula o determinante de **`mat`**, aplicando recursivamente o teorema de Laplace restrito à primeira coluna do determinante.

- **`Adj[mat]`**

Determina a matriz adjunta da matriz quadrada **`mat`**.

- **`MatrizInversa[mat]`**

Determina a matriz inversa da matriz quadrada **`mat`**, através da matriz adjunta.

- **`FilasPrincipais[mat]`**

Devolve o par formado pelas listas de linhas e de colunas principais de **`mat`**.

- **`MatRank[mat]`**

Determina a característica da matriz **`mat`** através de uma matriz principal. É equivalente à função **`MatrixRank`** existente no MATHEMATICA©.

- **`MatrizPrincipal[mat]`**

Determina uma matriz principal de **`mat`**.

- **`Cramer[mat,vec,simb]`**

Determina o vector solução geral do sistema linear de matriz simples **`mat`** e vector de termos independentes **`vec`**, exprimindo-a, no caso de indeterminação, nos necessários parâmetros arbitrários **`simb[1]`**, **`simb[2]`**, etc.

- **`Cauchy[mat,{i,j}]`**

Calcula o determinante de **`mat`**, através da fórmula de Cauchy (4.98.1) e usando a linha **`i`** e a coluna **`j`**.

- **`Wronskian[lista]`**

Calcula o wronskiano da lista de funções puras **`lista`**.

Segue-se a listagem do *package* **`ALGA`Determinantes``**:

```
(* Package by Carlos Ribeiro, Março 2009 *)
(* Contexto : ALGA`Determinantes` *)
(* Versão : 3.7 *)
(* Versão do Mathematica : 7.0 *)

BeginPackage["ALGA`Determinantes`", {"Combinatorica`"}]

(* ------------------------------ HELP ON-LINE ------------------------------ *)

PermutacaoQ::usage = "PermutacaoQ[p] determina se a lista p é uma permutação."

TransposicaoQ::usage = "TransposicaoQ[p] determina se a lista p é uma transposição."

ElementarQ::usage = "ElementarQ[p] determina se a lista p é uma transposição elementar."

PermInv::usage = "PermInv[p] calcula a permutação inversa de p."

Composicao::usage = "Composicao[p1, p2] calcula a permutação composta (p1 o p2)."

Decomposicao::usage = "Decomposicao[p] devolve uma lista {t1,t2,...,tn} de \
transposições tais que p = t1 o t2 o...o tn."

Epsilon::usage = "Epsilon[p] determina a paridade da permutação p."

DefDet::usage = "DefDet[mat] devolve o determinante da matriz quadrada mat, calculado \
pela soma dos termos da matriz multiplicados pela paridade."

Opr1::usage = "Opr1[A,i,j] Matriz que resulta de trocar as linhas i e j na matriz A."

Opr2::usage = "Opr2[A,i,x] Matriz que resulta de multiplicar a linha i da matriz A \
pelo escalar x não-nulo."

Opr3::usage = "Opr3[A,i,j,y] matriz obtida de A, substituindo a linha i pela sua soma \
com o produto do escalar y pela linha j."

CondensacaoDet::usage = "CondensacaoDet[mat] calcula o determinante da matriz quadrada \
mat, usando o método de condensação."

MatrizComplementar::usage =
"MatrizComplementar[mat, linhas, colunas] devolve a submatriz obtida de mat eliminando \
as filas indicadas nas listas linhas e colunas."
```

```
MenorComplementar::usage =
"MenorComplementar[mat, linhas, colunas] devolve o menor complementar da submatriz \
de mat cujas filas se indicam nas listas linhas e colunas."

Paridade::usage =
"Paridade[linhas, colunas] calcula a paridade do menor constituído pelas linhas e colunas \
indicadas nos argumentos."

Cof::usage =
"Cof[mat, linhas, colunas] calcula o cofactor da submatriz de mat cujas linhas e colunas \
se indicam nos 2° e 3° argumentos."

LaplaceDet::usage =
"LaplaceDet[mat,filas] devolve o valor do determinante de mat, calculando-o pelo teorema \
de Laplace, usando os menores contidos nas filas da lista filas. A opção Colunas -> False \
indica que as filas indicadas são linhas; A opção Colunas -> True indica que as filas \
indicadas são colunas; por defeito, o desenvolvimento é por colunas."

Colunas::usage = "A opção Colunas -> False indica que as filas indicadas na função LaplaceDet \
são linhas; A opção Colunas -> True indica que as filas indicadas são colunas; por defeito, o \
desenvolvimento é por colunas."

RecursivoDet::usage = "RecursivoDet[mat] calcula o determinante de mat, com um algoritmo \
recursivo, baseado no teorema de Laplace restrito."

Adj::usage = "Adj[mat] calcula a matriz adjunta de mat."

MatrizInversa::usage = "MatrizInversa[mat] calcula a matriz inversa de mat."

FilasPrincipais::usage = "FilasPrincipais[mat] devolve o par formado pelas listas de linhas \
e de colunas principais da matriz mat."

MatRank::usage = "MatRank[mat] devolve a característica de mat."

MatrizPrincipal::usage = "MatrizPrincipal[mat] devolve uma matriz principal de mat."

Cramer::usage = "Cramer[mat, vec, simb] devolve o vector solução geral do sistema de matriz \
simples mat e termos independentes vec, usando a regra de Cramer e os símbolos simb[1], \
simb[2], etc, como parâmetros arbitrários (quando indeterminado)."

Cauchy::usage = "Cauchy[mat, {i, j}] devolve o determinante de mat, calculado pela regra de \
Cauchy, a partir do elemento na posição (i,j)."

Wronskian::usage = "Wronskian[lista,var] devolve o Wronskiano da lista de funções puras \
indicadas, usando a variável var como argumento."

Begin["`Private`"]
(* ------------------------------- DEFAULTS ------------------------------- *)

Options[LaplaceDet] = {Colunas -> True};

(* -------------------------- MENSAGENS DE ERRO --------------------------- *)

Opr2::escalarnulo = "Operação do tipo 2 inválida. O 3° argumento não pode ser nulo.";

MatrizComplementar::fila = "Comprimento da lista `1` é incompatível com a matriz dada.";
MatrizComplementar::errfilas = "Índice(s) inválidos em `1`.";

MenorComplementar::rect = Paridade::rect = Adj::rect = MatrizInversa::rect = "Matriz \
rectangular encontrada; matriz quadrada esperada.";

MatrizInversa::sing = "Matriz singular.";
```

```
Cramer::errdim = "Número de linhas da matriz simples é diferente da dimensão do vector dos \
termos independentes.";
Cramer::imp = "Sistema sem solução.";

Cauchy::rect1 = "Matriz rectangular ou de ordem inferior a 3 encontrada.";
Cauchy::domain := "fila(s) indicadas fora do domínio admissível.";

(* ------------------------------ IMPLEMENTAÇÃO ------------------------------ *)

(* Implementação da função PermutacaoQ *)
PermutacaoQ[p_List] := p != {} && Sort[p] == Range[Length[p]]

(* Implementação da função TransposicaoQ *)
TransposicaoQ[p_?PermutacaoQ] :=
  Module[{r = Range[Length[p]]-p},
    Length[p]-Count[r, 0] == 2
  ]

(* Implementação da função ElementarQ *)
ElementarQ[p_?PermutacaoQ] :=
  Module[{n = Range[Length[p]], nz},
    nz = Sort[Complement[n, Flatten[Position[n-p, 0]]]];
    Length[nz] == 2 && nz[[2]] == nz[[1]]+1
  ]

(* Implementação da função PermInv *)
PermInv[p_?PermutacaoQ] :=
  Module[{s = {}},
    Do[
      s = Join[s, Flatten[Position[p, i]]],
      {i, Length[p]}
    ];
    s
  ]

(* Implementação da função Composicao *)
Composicao[p1_?PermutacaoQ, p2_?PermutacaoQ] :=
  Module[{s = {}},
    Do[
      s = Join[s, {p1[[p2[[i]]]]}],
      {i, Length[p1]}
    ];
    s
  ]

(* Implementação da função Decomposicao *)
Decomposicao[p_?PermutacaoQ] :=
  Module[{j, id = Range[Length[p]], id1, tmp, dec = {}},
    Do[
      j = Position[id, p[[i]]][[1,1]];
      If[j != i,
        tmp = id[[i]];
        id[[i]] = id[[j]];
        id[[j]] = tmp;
        id1 = Range[Length[p]];
        tmp = id1[[i]];
        id1[[i]] = id1[[j]];
        id1[[j]] = tmp;
        dec = Join[dec, {id1}]
      ],
      {i, Length[p]-1}
    ];
    dec
```

```
  ]
(* Implementação da função Epsilon *)
Epsilon[p_?PermutacaoQ] := (-1)^Length[Decomposicao[p]]

(* Implementação da função DefDet *)
DefDet[a_?MatrixQ]:=
  Module[{n = Length[a], sn},
    sn = Permutations[Range[n]];
    Sum[Signature[sn[[i]]]*Product[a[[sn[[i,j]], j]], {j, n}], {i, Length[sn]}]
    ]

(* Implementação das 3 operações elementares *)
Opr1[a_?MatrixQ, i_Integer, j_Integer] :=
  Module[{temp = a[[i]], b = a},
    b[[i]] = b[[j]]; b[[j]] = temp; b
  ]

Opr2[a_?MatrixQ, i_Integer, x_] :=
  Module[{b = a},
    If[x == 0,
      Message[Opr2::escalarnulo],
      b[[i]] = x b[[i]]; b
    ]
  ]

Opr3[a_?MatrixQ, i_Integer, j_Integer, x_] :=
  Module[{b = a},
    b[[i]] = b[[i]] + x b[[j]]; b
  ]

(* Implementação da função CondensacaoDet *)
CondensacaoDet[a_?MatrixQ] :=
  Module[{n = Length[a], b = a, r = 0, i, d = 1},
    While[r < n,
      i = r;
      While[
        i++;
        b[[i, r+1]] == 0 && i < n,
        Null
      ];
      If[b[[i, r+1]] == 0,
        Return[0],
        r++;
        If[i != r, b = Opr1[b, r, i]; d *= -1];
        d *= b[[r, r]]; b = Opr2[b, r, 1/b[[r, r]]];
        For[i = r+1, i <= n, i++,
          b = Opr3[b, i, r, -b[[i, r]]]
        ]
      ]
    ];
    d
  ]

(* Implementação da função MatrizComplementar *)
MatrizComplementar[a_?MatrixQ, linhas_List, colunas_List] :=
  Module[{l = Union[linhas], c = Union[colunas], m = Range[Length[a]],
    n = Range[Dimensions[a][[2]]]},
    Which[
      Union[m, l] != m, Message[MatrizComplementar::errfilas, linhas],
      Union[n, c] != n, Message[MatrizComplementar::errfilas, colunas],
      Length[l] >= Length[a], Message[MatrizComplementar::fila, linhas],
      Length[c] >= Dimensions[a][[2]], Message[MatrizComplementar::fila, colunas],
```

```
      True, a[[Complement[m, l], Complement[n, c]]]
    ]
  ]

(* Implementação da função MenorComplementar *)
MenorComplementar[a_?MatrixQ, linhas_List, colunas_List] :=
  Module[{submat},
    If[Length[a]==Dimensions[a][[2]],
      submat = MatrizComplementar[a, linhas, colunas];
      If[Length[submat]==Dimensions[submat][[2]],
        Det[submat],
        Message[MenorComplementar::rect]
      ],
      Message[MenorComplementar::rect]
    ]
  ]

(* Implementação da função Paridade *)
Paridade[linhas_List, colunas_List] :=
  Module[{l = Union[linhas], c = Union[colunas]},
    (-1)^(Sum[l[[i]],{i, Length[l]}]+Sum[c[[i]],{i, Length[c]}])
  ]

(* Implementação da função Cof *)
Cof[a_?MatrixQ, linhas_List, colunas_List] :=
  Paridade[linhas,colunas]*MenorComplementar[a, linhas, colunas]

(* Implementação da função LaplaceDet *)
LaplaceDet[a_?MatrixQ, filas_List, opts___?OptionQ] :=
  Module[{n = Length[a], fs = Union[filas], seq, cols},
    seq = KSubsets[Range[n], Length[fs]];
    cols = Colunas /. Flatten[{opts, Options[LaplaceDet]}];
    If[cols,
      Sum[Cof[a, seq[[i]], fs]*Det[a[[seq[[i]], fs]]], {i, Length[seq]}],
      Sum[Cof[a, fs, seq[[i]]]*Det[a[[fs, seq[[i]]]]], {i, Length[seq]}]
    ]
  ]

(* Implementação da função RecursivoDet *)
RecursivoDet[a_?MatrixQ] :=
  Module[{n = Length[a]},
    If[n == 1,
      a[[1, 1]],
      Sum[a[[i, 1]]*(-1)^(i+1)*RecursivoDet[Drop[a, {i}, {1}]], {i, n}]
    ]
  ]

(* Implementação da função Adj *)
Adj[a_?MatrixQ] :=
  Module[{n = Length[a]},
    If[Length[a]==Dimensions[a][[2]],
      Transpose[Table[Cof[a,{i},{j}],{i,n},{j,n}]],
      Message[Adj::rect]
    ]
  ]

(* Implementação da função MatrizInversa *)
MatrizInversa[a_?MatrixQ] :=
  Module[{d},
    If[Length[a]==Dimensions[a][[2]],
      If[(d = Det[a]) == 0,
        Message[MatrizInversa::sing],
```

```
        Adj[a]/d
      ],
      Message[MatrizInversa::rect]
    ]
  ]

(* 1ª Implementação da função FilasPrincipais *)
FilasPrincipais[a_?MatrixQ] :=
  Module[{m = Length[a], n = Dimensions[a][[2]], p, u, v, i, j},
    p = Min[m, n]+1;
    While[p > 1,
      p--;
      u = KSubsets[Range[m], p];
      v = KSubsets[Range[n], p];
      i = 0;
      While[i < Length[u],
        i++;
        j = 0;
        While[j < Length[v],
          j++;
          If[Simplify[Det[a[[u[[i]],v[[j]]]]]] != 0, Return[{u[[i]],v[[j]]}]]
        ]
      ]
    ];
    {{},{}}
  ]

(* Implementação da função MatRank *)
MatRank[a_?MatrixQ]:= Length[FilasPrincipais[a][[1]]]

(* Implementação da função MatrizPrincipal *)
MatrizPrincipal[a_?MatrixQ] :=
  Module[{p = FilasPrincipais[a]},
    a[[p[[1]], p[[2]]]]
  ]

(* Implementação da função Cramer *)
Cramer[a_?MatrixQ, b_?VectorQ, x_Symbol] :=
  Module[{m = Length[a], n = Dimensions[a][[2]], eqp, vrp, eqs, vrs,
    r, poss = True, i = 0, teste, sol, matp, d, bc, matpx},
    If[m != Length[b],
      Message[Cramer::errdim],
      {eqp, vrp} = FilasPrincipais[a];
      eqs = Complement[Range[m],eqp];
      vrs = Complement[Range[n],vrp];
      r = Length[eqp];
      While[poss && i < m-r,
        i++;
        teste = Union[eqp, {eqs[[i]]}];
        poss = Simplify[Det[Join[Transpose[a[[teste, vrp]]], {b[[teste]]}]] == 0]
      ];
      If[poss,
        sol = Array[x, {n}];
        If[r > 0,
          matp = Transpose[a[[eqp, vrp]]];
          d = Det[matp];
          bc = b[[eqp]]-a[[eqp, vrs]].sol[[vrs]];
          Do[
            matpx = matp;
            matpx[[i]] = bc;
            sol[[vrp[[i]]]] = Det[matpx]/d,
            {i, r}
```

```
          ]
        ];
        Simplify[sol],
        Message[Cramer::imp]
      ]
    ]
  ]

(* Implementação da função Cauchy *)
Cauchy[a_?MatrixQ, filas_List]:=
  Module[{n = Length[a], i, j},
    If[Length[a]==Dimensions[a][[2]] && n > 2,
      f = Range[n];
      {i, j} = filas;
      If[MemberQ[f, i] && MemberQ[f, j],
        a[[i, j]]*Cof[a, {i}, {j}]-(-1)^(i+j)*
          Drop[a[[i]], {j}].Adj[Drop[a, {i}, {j}]].Drop[a, {i}][[All, j]],
        Message[Cauchy::domain]
      ],
      Message[Cauchy::rect1]
    ]
  ]

(* Implementação da função Wronskian *)
Wronskian[seq_List,x_]:=Module[{n=Length[seq]},
  Det[Table[D[seq[[j]][x],{x,i}],{i,0,n-1},{j,1,n}]]//FullSimplify
]

End[]

EndPackage[]
```

O *package* **`ALGA`Determinantes``** deverá ser previamente carregado em memória, o que se consegue por meio de um dos comandos seguintes:

```
<<ALGA`Determinantes`
Get["ALGA`Determinantes`"]
Needs["ALGA`Determinantes`"]
DeclarePackage["ALGA`Determinantes`"]
```

Nas páginas seguintes apresentamos um notebook ilustrativo do uso das funções implementadas no *package* anterior, bem como de algumas funções internas ao sistema MATHEMATICA©.

Determinantes

4.16.1. Permutações

- **Com o MATHEMATICA, podemos trabalhar no Grupo simétrico de qualquer ordem**

```
<< ALGA`Determinantes`
```

A seguinte lista não é uma permutação

```
?PermutacaoQ
```

PermutacaoQ[p] determina se a lista p é uma permutação.

```
PermutacaoQ[{2, 4, 1}]
```

```
False
```

Mas esta já será

```
PermutacaoQ[{3, 4, 2, 1}]
```

```
True
```

Podemos definir uma função para gerar aleatoriamente uma permutação com uma ordem também aleatória entre dois extremos dados n1 e n2

```
PermutacaoRandom[n1_Integer, n2_Integer] :=
 Module[{n, p}, n = RandomInteger[{n1, n2}];
  While[p = RandomInteger[{1, n}, n];
   ! PermutacaoQ[p], Null]; p]
```

E usá-la para gerar uma permutação de ordem aleatória entre 5 e 10

```
PermutacaoRandom[5, 10]
```

```
{5, 7, 6, 4, 1, 2, 3}
```

Algumas transposições junto com outras permutações que o não são

```
{TransposicaoQ[{5, 7, 3, 1, 4, 6, 2}],
 TransposicaoQ[{1, 2, 4, 3}],
 TransposicaoQ[{1, 5, 3, 4, 2, 6}],
 TransposicaoQ[{2, 3, 1}]}
```

```
{False, True, True, False}
```

Das transposições anteriores, só a segunda é elementar

```
{ElementarQ[{5, 7, 3, 1, 4, 6, 2}],
 ElementarQ[{1, 2, 4, 3}],
 ElementarQ[{1, 5, 3, 4, 2, 6}], ElementarQ[{2, 3, 1}]}
```

```
{False, True, False, False}
```

Podemos calcular a permutação inversa

```
p = {4, 2, 3, 1, 6, 5}; q = PermInv[p]
```

```
{4, 2, 3, 1, 6, 5}
```

E verificar o resultado

```
Composicao[p, q] == Range[Length[p]]
```

```
True
```

Construção do grupo simétrico de ordem 10

```
s10 = Permutations[Range[10]];
```

Podemos agora obter duas permutações aleatórias de ordem 10

```
{p = s10[[RandomInteger[{1, 10!}]]],
 q = s10[[RandomInteger[{1, 10!}]]]}
```

```
{{2, 3, 8, 5, 7, 4, 6, 9, 10, 1}, {2, 4, 7, 10, 6, 8, 1, 9, 5, 3}}
```

E compô-las

```
?Composicao
```

Composicao[p1, p2] calcula a permutação composta (p1 o p2).

```
r = Composicao[p, q]
```

```
{3, 5, 6, 1, 4, 9, 2, 10, 7, 8}
```

Determinemos as paridades e verifique-se o teorema da paridade do produto

```
?Epsilon
```

Epsilon[p] determina a paridade da permutação p.

```
{Epsilon[p], Epsilon[q], Epsilon[r]}
```

```
{1, 1, 1}
```

Determinemos as paridades pela função `Signature[]` do sistema

```
{Signature[p], Signature[q], Signature[r]}
```

```
{1, 1, 1}
```

Por fim, façamos a decomposição de uma permutação em produto de transposições

```
?Decomposicao
```

Decomposicao[p] devolve uma lista {t1,t2,...,tn} de transposições tais que p = t1 o t2 o...o tn.

```
d = Decomposicao[p]
```

```
{{2, 1, 3, 4, 5, 6, 7, 8, 9, 10}, {1, 3, 2, 4, 5, 6, 7, 8, 9, 10},
 {1, 2, 8, 4, 5, 6, 7, 3, 9, 10}, {1, 2, 3, 5, 4, 6, 7, 8, 9, 10},
 {1, 2, 3, 4, 7, 6, 5, 8, 9, 10}, {1, 2, 3, 4, 5, 7, 6, 8, 9, 10},
 {1, 2, 3, 4, 5, 6, 7, 9, 8, 10}, {1, 2, 3, 4, 5, 6, 7, 8, 10, 9}}
```

Verifiquemos que se trata, de facto, de transposições

```
Table[TransposicaoQ[d[[i]]], {i, Length[d]}]
```

```
{True, True, True, True, True, True, True, True}
```

Verifiquemos que a composição das transposições anteriores (pela ordem obtida) dá como resultado a permutação p

```
Fold[Composicao, d[[1]], Drop[d, 1]]
```

```
{2, 3, 8, 5, 7, 4, 6, 9, 10, 1}
```

4.16.2. Determinantes (com matriz real)

- **As funções** `Det[]`, `DefDet[]`, `CondensacaoDet[]`, `RecursivoDet[]` **e** `Cauchy[]` **permitem o cálculo de determinantes**

```
?DefDet
```

DefDet[mat] devolve o determinante da matriz quadrada
mat, calculado pela soma dos termos da matriz multiplicados pela paridade.

```
?CondensacaoDet
```

CondensacaoDet[mat] calcula o determinante da matriz quadrada mat, usando o método de condensação.

```
?RecursivoDet
```

RecursivoDet[mat] calcula o determinante de mat, com um algoritmo recursivo, baseado no teorema de Laplace restrito.

```
?Cauchy
```

Cauchy[mat, {i, j}] devolve o determinante de mat, calculado pela regra de Cauchy, a partir do elemento na posição (i,j).

Defina-se uma matriz real aleatória

```
(a = RandomInteger[{-9, 9}, {7, 7}]) // MatrixForm
```

$$\begin{pmatrix} 3 & -2 & 8 & -8 & -7 & -8 & -5 \\ 2 & -9 & -9 & 1 & 0 & 7 & -6 \\ 4 & 4 & -6 & -8 & -7 & -5 & -5 \\ -5 & -8 & -4 & -5 & 3 & 8 & -1 \\ -6 & -9 & 1 & 4 & 9 & -4 & 9 \\ -9 & -3 & -8 & -6 & -6 & 6 & -2 \\ 4 & -5 & -9 & 4 & -8 & -2 & 6 \end{pmatrix}$$

O determinante será, pelas 5 funções

```
{Det[a], DefDet[a], CondensacaoDet[a],
 RecursivoDet[a], Cauchy[a, {6, 3}]}
```

```
{29 585 486, 29 585 486, 29 585 486, 29 585 486, 29 585 486}
```

- **As funções** `MatrizComplementar[]`, `MenorComplementar[]`, `Paridade[]` **e** `Cof[]` **permitem calcular menores complementares, paridades e cofactores**

```
?MatrizComplementar
```

MatrizComplementar[mat, linhas, colunas] devolve a
submatriz obtida de mat eliminando as filas indicadas nas listas linhas e colunas.

```
?MenorComplementar
```

MenorComplementar[mat, linhas, colunas] devolve o menor
complementar da submatriz de mat cujas filas se indicam nas listas linhas e colunas.

```
? Paridade
```

Paridade[linhas, colunas] calcula a paridade do menor constituído pelas linhas e colunas indicadas nos argumentos.

```
? Cof
```

Cof[mat, linhas, colunas] calcula o cofactor da submatriz de mat cujas linhas e colunas se indicam nos 2º e 3º argumentos.

A submatriz complementar de a obtida eliminando as linhas 2, 4, 6 e as colunas 1, 3, 5 é

```
MatrizComplementar[a, {2, 4, 6}, {1, 3, 5}] //
 MatrixForm
```

$$\begin{pmatrix} -2 & -8 & -8 & -5 \\ 4 & -8 & -5 & -5 \\ -9 & 4 & -4 & 9 \\ -5 & 4 & -2 & 6 \end{pmatrix}$$

O menor complementar do menor contido nas linhas 2, 4, 6 e as colunas 1, 3, 5 é o determinante da matriz anterior

```
{m = MenorComplementar[a, {2, 4, 6}, {1, 3, 5}], Det[%]}
```

```
{432, 432}
```

A paridade será

```
s = Paridade[{2, 4, 6}, {1, 3, 5}]
```

```
-1
```

E o cofactor é calculado por

```
{Cof[a, {2, 4, 6}, {1, 3, 5}], s * m}
```

```
{-432, -432}
```

- **As funções `LaplaceDet[]`, `RecursivoDet[]` e `Cauchy[]` calculam determinantes pelo teorema de Laplace geral (ou restrito, se a lista fornecida for singular), por um algoritmo recursivo e pela fórmula de Cauchy; as funções `Adj[]` e `MatrizInversa[]` calculam a adjunta e a inversa de uma matriz.**

O teorema de Laplace aplicado às linhas 2, 4 e 7 e ainda às colunas 1 e 4 dá

```
{LaplaceDet[a, {2, 4, 7}, Colunas → False],
 LaplaceDet[a, {1, 4}]}
```

```
{29 585 486, 29 585 486}
```

O teorema de Laplace restrito aplicado recursivamente fornece também o determinante

```
RecursivoDet[a]
```

```
29 585 486
```

A fórmula de Cauchy calcula também o determinante

```
Cauchy[a, {3, 6}]
```

```
29 585 486
```

A adjunta de a é a matriz

```
(b = Adj[a]) // MatrixForm
```

$$\begin{pmatrix} -110\,278 & -537\,436 & 673\,320 & 2\,826\,692 & -1\,373\,054 & -3\,175\,200 & 1\,404\,062 \\ -1\,395\,366 & -1\,158\,856 & 797\,432 & -500\,684 & -734\,698 & 136\,094 & -593\,170 \\ 1\,407\,056 & -763\,550 & -2\,059\,438 & 595\,414 & -900\,064 & -332\,850 & 31\,180 \\ -108\,108 & 2\,092\,098 & -1\,838\,698 & -4\,014\,292 & 424\,082 & 1\,394\,400 & -370\,612 \\ -1\,166\,809 & 285\,799 & 1\,613\,917 & 921\,187 & 1\,125\,365 & -1\,478\,582 & -1\,368\,988 \\ -468\,076 & -1\,066\,376 & -1\,568\,186 & 2\,837\,736 & -2\,273\,218 & -767\,088 & 863\,826 \\ -618\,401 & -3\,121\,873 & -18\,551 & 3\,441\,779 & -586\,949 & -1\,425\,802 & 2\,257\,034 \end{pmatrix}$$

A matriz a e a sua adjunta permutam

```
a.b == b.a
```

```
True
```

E a adjunta verifica a igualdade fundamental: `a.Adj[a] == Det[a]*IdentityMatrix[Length[a]]`

```
a.b // MatrixForm
```

$$\begin{pmatrix} 29\,585\,486 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 29\,585\,486 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 29\,585\,486 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 29\,585\,486 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 29\,585\,486 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 29\,585\,486 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 29\,585\,486 \end{pmatrix}$$

Que é igual a

```
Det[a] IdentityMatrix[Length[a]] // MatrixForm
```

$$\begin{pmatrix} 29\,585\,486 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 29\,585\,486 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 29\,585\,486 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 29\,585\,486 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 29\,585\,486 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 29\,585\,486 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 29\,585\,486 \end{pmatrix}$$

A inversa através da adjunta é

```
(c = MatrizInversa[a]) // MatrixForm
```

$$\begin{pmatrix} -\frac{7877}{2\,113\,249} & -\frac{268\,718}{14\,792\,743} & \frac{336\,660}{14\,792\,743} & \frac{1\,413\,346}{14\,792\,743} & -\frac{686\,527}{14\,792\,743} & -\frac{226\,800}{2\,113\,249} & \frac{702\,031}{14\,792\,743} \\ -\frac{99\,669}{2\,113\,249} & -\frac{579\,428}{14\,792\,743} & \frac{398\,716}{14\,792\,743} & -\frac{250\,342}{14\,792\,743} & -\frac{367\,349}{14\,792\,743} & \frac{9721}{2\,113\,249} & -\frac{296\,585}{14\,792\,743} \\ \frac{100\,504}{2\,113\,249} & -\frac{381\,775}{14\,792\,743} & -\frac{1\,029\,719}{14\,792\,743} & \frac{297\,707}{14\,792\,743} & -\frac{450\,032}{14\,792\,743} & -\frac{23\,775}{2\,113\,249} & \frac{15\,590}{14\,792\,743} \\ -\frac{7722}{2\,113\,249} & \frac{1\,046\,049}{14\,792\,743} & -\frac{919\,349}{14\,792\,743} & -\frac{2\,007\,146}{14\,792\,743} & \frac{212\,041}{14\,792\,743} & \frac{99\,600}{2\,113\,249} & -\frac{185\,306}{14\,792\,743} \\ -\frac{166\,687}{4\,226\,498} & \frac{285\,799}{29\,585\,486} & \frac{1\,613\,917}{29\,585\,486} & \frac{921\,187}{29\,585\,486} & \frac{1\,125\,365}{29\,585\,486} & -\frac{105\,613}{2\,113\,249} & -\frac{684\,494}{14\,792\,743} \\ -\frac{33\,434}{2\,113\,249} & -\frac{533\,188}{14\,792\,743} & -\frac{784\,093}{14\,792\,743} & \frac{1\,418\,868}{14\,792\,743} & -\frac{1\,136\,609}{14\,792\,743} & -\frac{54\,792}{2\,113\,249} & \frac{431\,913}{14\,792\,743} \\ -\frac{88\,343}{4\,226\,498} & -\frac{3\,121\,873}{29\,585\,486} & -\frac{18\,551}{29\,585\,486} & \frac{3\,441\,779}{29\,585\,486} & -\frac{586\,949}{29\,585\,486} & -\frac{101\,843}{2\,113\,249} & \frac{1\,128\,517}{14\,792\,743} \end{pmatrix}$$

Verificação de que se trata da matriz inversa

```
a.c // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

4.16.3. Determinantes (com matriz complexa)

- **As funções** `Det[]`**,** `CondensacaoDet[]`**,** `RecursivoDet[]` **e** `Cauchy[]` **permitem o cálculo de determinantes**

Defina-se uma matriz complexa aleatória

```
(a = Table[
    RandomInteger[{-9, 9}] + RandomInteger[{-6, 6}] i,
    {4}, {4}]) // MatrixForm
```

$$\begin{pmatrix} -4+3i & 4-5i & 2-4i & 4+i \\ -3+i & 9+5i & 8-2i & -3+5i \\ -1+i & 6+i & -1+2i & 4+6i \\ 3-5i & -5+5i & 3+3i & 7-3i \end{pmatrix}$$

O determinante será, pelas 4 funções

```
{Det[a], CondensacaoDet[a],
 RecursivoDet[a], Cauchy[a, {3, 1}]}
```

$\{3952-1608\,i,\ 3952-1608\,i,\ 3952-1608\,i,\ 3952-1608\,i\}$

As funções `MatrizComplementar[]`, `MenorComplementar[]`, `Paridade[]` e `Cof[]` permitem calcular menores complementares, paridades e cofactores

A submatriz complementar de a obtida eliminando as linhas 2 e 4 e as colunas 1 e 3 é

```
MatrizComplementar[a, {2, 4}, {1, 3}] // MatrixForm
```

$$\begin{pmatrix} 4 - 5\,i & 4 + i \\ 6 + i & 4 + 6\,i \end{pmatrix}$$

O menor complementar do menor contido nas linhas 2, 4 e as colunas 1, 3 é o determinante da matriz anterior

```
{m = MenorComplementar[a, {2, 4}, {1, 3}], Det[%]}
```

```
{23 - 6 i, 23 - 6 i}
```

A paridade será

```
s = Paridade[{2, 4}, {1, 3}]
```

```
1
```

E o cofactor é calculado por

```
{Cof[a, {2, 4}, {1, 3}], s * m}
```

```
{23 - 6 i, 23 - 6 i}
```

- **As funções** `LaplaceDet[]`, `RecursivoDet[]` **e** `Cauchy[]` **calculam determinantes pelo teorema de Laplace geral (ou restrito, se a lista fornecida for singular), por um algoritmo recursivo e pela fórmula de Cauchy; as funções** `Adj[]` **e** `MatrizInversa[]` **calculam a adjunta e a inversa de uma matriz.**

```
? LaplaceDet
```

LaplaceDet[mat,filas] devolve o valor do determinante de mat, calculando–o pelo teorema de Laplace, usando os menores contidos nas filas da lista filas. A opção Colunas –> False indica que as filas indicadas são linhas; A opção Colunas –> True indica que as filas indicadas são colunas; por defeito, o desenvolvimento é por colunas.

```
? Adj
```

Adj[mat] calcula a matriz adjunta de mat.

```
? MatrizInversa
```

MatrizInversa[mat] calcula a matriz inversa de mat.

O teorema de Laplace aplicado às linhas 1, 3 e 4 e ainda às colunas 1 e 4 dá

```
{LaplaceDet[a, {1, 3, 4}, Colunas → False],
 LaplaceDet[a, {1, 4}]}
```

```
{3952 - 1608 ⅈ, 3952 - 1608 ⅈ}
```

O teorema de Laplace restrito aplicado recursivamente fornece também o determinante

```
RecursivoDet[a]
```

```
3952 - 1608 ⅈ
```

A fórmula de Cauchy calcula também o determinante

```
Cauchy[a, {2, 2}]
```

```
3952 - 1608 ⅈ
```

A adjunta de a é a matriz

```
(b = Adj[a]) // MatrixForm
```

$$\begin{pmatrix} -888+98\,i & 187-469\,i & -116-280\,i & 80+30\,i \\ -436+176\,i & 22-198\,i & 184-424\,i & -256-164\,i \\ 174+54\,i & 362-56\,i & -468-44\,i & 182+166\,i \\ 184-258\,i & -59+21\,i & 284-56\,i & 268-154\,i \end{pmatrix}$$

A matriz a e a sua adjunta permutam

```
a.b == b.a
```

```
True
```

E a adjunta verifica a propriedade fundamental: `a.Adj[a] == Det[a]*IdentityMatrix[Length[a]]`

```
a.b // MatrixForm
```

$$\begin{pmatrix} 3952-1608\,i & 0 & 0 & 0 \\ 0 & 3952-1608\,i & 0 & 0 \\ 0 & 0 & 3952-1608\,i & 0 \\ 0 & 0 & 0 & 3952-1608\,i \end{pmatrix}$$

Que é igual a

```
Det[a] IdentityMatrix[Length[a]] // MatrixForm
```

$$\begin{pmatrix} 3952-1608\,i & 0 & 0 & 0 \\ 0 & 3952-1608\,i & 0 & 0 \\ 0 & 0 & 3952-1608\,i & 0 \\ 0 & 0 & 0 & 3952-1608\,i \end{pmatrix}$$

A inversa através da adjunta é

```
(c = MatrizInversa[a]) // MatrixForm
```

$$\begin{pmatrix} -\frac{229185}{1137748}-\frac{32519\,i}{568874} & \frac{186647}{2275496}-\frac{194099\,i}{2275496} & -\frac{128}{284437}-\frac{40409\,i}{568874} & \frac{16745}{1137748}+\frac{7725\,i}{568874} \\ -\frac{31345}{284437}-\frac{173\,i}{568874} & \frac{25333}{1137748}-\frac{46695\,i}{1137748} & \frac{22015}{284437}-\frac{21559\,i}{284437} & -\frac{23375}{568874}-\frac{16559\,i}{284437} \\ \frac{37551}{1137748}+\frac{30825\,i}{1137748} & \frac{47521}{568874}+\frac{22549\,i}{1137748} & -\frac{55587}{568874}-\frac{28951\,i}{568874} & \frac{28271}{1137748}+\frac{59293\,i}{1137748} \\ \frac{71377}{1137748}-\frac{22617\,i}{568874} & -\frac{33367}{2275496}-\frac{1485\,i}{2275496} & \frac{18944}{284437}+\frac{7355\,i}{568874} & \frac{81673}{1137748}-\frac{2776\,i}{284437} \end{pmatrix}$$

Que verifica a definição de inversa

```
a.c // MatrixForm
```

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

4.16.4. Submatrizes principais e característica de uma matriz

- **As funções `FilasPrincipais[]`, `MatrizPrincipal[]` e `MatRank[]` permitem saber as filas principais de uma matriz, uma matriz principal e a característica de uma matriz qualquer, usando determinantes.**

```
?FilasPrincipais
```

FilasPrincipais[mat] devolve o par formado pelas listas de linhas e de colunas principais da matriz mat.

```
?MatrizPrincipal
```

MatrizPrincipal[mat] devolve uma matriz principal de mat.

```
? MatRank
```

MatRank[mat] devolve a característica de mat.

Defina-se uma matriz

```
(a = {{3, -2, -1, 0}, {-6, 4, 0, -3},
     {-3, 2, -1, -3}, {0, 0, -2, -3}}) // MatrixForm
```

$$\begin{pmatrix} 3 & -2 & -1 & 0 \\ -6 & 4 & 0 & -3 \\ -3 & 2 & -1 & -3 \\ 0 & 0 & -2 & -3 \end{pmatrix}$$

As filas principais são dadas num par de listas

```
FilasPrincipais[a]
```

```
{{1, 2}, {1, 3}}
```

Extraindo uma submatriz principal

```
(p = MatrizPrincipal[a]) // MatrixForm
```

$$\begin{pmatrix} 3 & -1 \\ -6 & 0 \end{pmatrix}$$

Verificação de que a matriz anterior é principal: ela é regular

```
Det[p] != 0
```

```
True
```

...mas todas as submatrizes de ordem 3 são singulares

```
gl = KSubsets[Range[Length[a]], 3];
gc = KSubsets[Range[Dimensions[a][[2]]], 3];
```

```
Table[Det[a[[gl[[i]], gc[[j]]]]],
 {i, Length[gl]}, {j, Length[gc]}]
```

```
{{0, 0, 0, 0}, {0, 0, 0, 0}, {0, 0, 0, 0}, {0, 0, 0, 0}}
```

...e o mesmo se passa com as de ordem 4

```
gl = KSubsets[Range[Length[a]], 4];
gc = KSubsets[Range[Dimensions[a][[2]]], 4];
```

```
Table[Det[a[[gl[[i]], gc[[j]]]]],
 {i, Length[gl]}, {j, Length[gc]}]
```

```
{{0}}
```

A característica calculada pela função implementada no *package* e pela função existente no sistema MATHEMATICA:

```
{MatRank[a], MatrixRank[a]}
```

```
{2, 2}
```

4.16.5. Regra de Cramer

Caso de um sistema de Cramer

```
?Cramer
```

Cramer[mat, vec, simb] devolve o vector solução geral do sistema de matriz simples mat e termos independentes vec, usando a regra de Cramer e os símbolos simb[1], simb[2], etc, como parâmetros arbitrários (quando indeterminado).

Neste exemplo, a matriz simples e o vector dos termos independentes são

```
(a = {{2, 1, -3}, {3, -2, 2}, {5, -3, -1}}) // MatrixForm
```

$$\begin{pmatrix} 2 & 1 & -3 \\ 3 & -2 & 2 \\ 5 & -3 & -1 \end{pmatrix}$$

```
b = {5, 5, 16};
```

O sistema é determinado (sistema de Cramer) e a solução é

```
sol = Cramer[a, b, x]
```

```
{1, -3, -2}
```

Verificação da solução

```
Simplify[a.sol == b]
```

```
True
```

Caso de um sistema indeterminado

Neste outro exemplo, a matriz simples e o vector dos termos independentes são

```
(a = {{1, 2, -1, 3}, {2, 4, 4, 3}, {3, 6, -1, 8}}) //
 MatrixForm
```

$$\begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 4 & 3 \\ 3 & 6 & -1 & 8 \end{pmatrix}$$

```
b = {3, 9, 10};
```

O sistema é duplamente indeterminado e a solução geral, sendo principais a 1ª e a 2ª incógnitas, será

```
sol = Cramer[a, b, x]
```

$$\left\{\frac{1}{2}\,(7 - 4\,x[2] - 5\,x[4]),\ x[2],\ \frac{1}{2}\,(1 + x[4]),\ x[4]\right\}$$

As colunas principais da matriz simples são a 1ª e 3ª, confirmando serem estas as incógnitas principais escolhidas pelo algoritmo anterior. As equações principais foram as duas primeiras.

```
FilasPrincipais[a]
```

```
{{1, 2}, {1, 3}}
```

Verificação da solução

```
Simplify[a.sol == b]
```

```
True
```

Caso de um sistema impossível

Neste último exemplo, a matriz simples e o vector dos termos independentes são

```
(a = {{1, 5, 4, -13}, {3, -1, 2, 5}, {2, 2, 3, -4}}) //
 MatrixForm
```

$$\begin{pmatrix} 1 & 5 & 4 & -13 \\ 3 & -1 & 2 & 5 \\ 2 & 2 & 3 & -4 \end{pmatrix}$$

```
b = {3, 2, 1};
```

Neste caso, o sistema é impossível

```
Cramer[a, b, x]
```

Cramer::imp : Sistema sem solução.

Caso de um sistema indeterminado sobre o corpo dos complexos

Neste último exemplo, a matriz simples e o vector dos termos independentes são complexos

```
(a = {{1 + I, 3 - 2 I, 1 - I}, {2 I, -2 + 3 I, 1 + I},
    {-6 - 2 I, -12, -4 + 4 I}}) // MatrixForm
```

$$\begin{pmatrix} 1+i & 3-2i & 1-i \\ 2i & -2+3i & 1+i \\ -6-2i & -12 & -4+4i \end{pmatrix}$$

```
b = {6 - I, -9 + 4 I, -20 - 16 I}
```

$\{6-i, -9+4i, -20-16i\}$

Neste caso, o sistema é simplesmente indeterminado e foram seleccionadas as duas primeiras incógnitas para principais

```
sol = Cramer[a, b, x]
```

$$\left\{\left(-\frac{4}{53}+\frac{14\,i}{53}\right)(5+2\,x[3]),\ \left(-\frac{2}{53}+\frac{7\,i}{53}\right)((1-16\,i)+(1+i)\,x[3]),\ x[3]\right\}$$

Verificação da solução

```
Simplify[a.sol == b]
```

```
True
```

4.16.6. Uso da função `Wronskiano[]`

```
Remove["Global`*"]
```

? Wronskiano

Wronskiano[lista,var] devolve o Wronskiano da lista de funções puras indicadas, usando a variável var como argumento.

```
f = {Exp[-I 3 ω #] &, Exp[-I 2 ω #] &, Exp[-I ω #] &,
   1 &, Exp[I ω #] &, Exp[I 2 ω #] &, Exp[I 3 ω #] &};
```

Wronskiano não idênticamente nulo: lista `f` linearmente independente

```
Wronskiano[f, x]
```

$24\,883\,200\, i\, \omega^{21}$

```
f = {Sin[#]^2 &, Cos[#]^2 &, 1 &};
```

Wronskiano idênticamente nulo, porque a lista `f` é linearmente dependente

```
Wronskiano[f, x]
```

```
0
```

```
f = {Sin[#] &, Cos[#] &, 1 &};
```

Wronskiano não idênticamente nulo: lista `f` linearmente independente

```
Wronskiano[f, x]
```

```
-1
```

Bibliografia

[1] AITKEN, A. C. *Determinants and Matrices*. 1944.

[2] ALBERT, A. A. *Fundamental concepts of higher algebra*, University of Chicago Press.

[3] ANTON, H. e RORRES, C., *Álgebra Linear com Aplicações*. Bookman, Porto Alegre, 2001.

[4] APOSTOL, T. M. *Calculus*, vol. 2. John Wiley & Sons, New York, 1969.

[5] BIRKHOFF, G. and MAC LANE, S. *Algebra*. Macmillan, New York, 1965.

[6] BIRKHOFF, G. and MAC LANE, S. *A survey of Modern Algebra*. Macmillan, New York, 1965.

[7] CARAÇA, B. J. *Lições de Álgebra e Análise*, vol. I, 1956.

[8] CHAMBADAL, L. et OVAERT, J. L. *Cours de Mathématiques – algèbre I*. Gauthier-Villars, Paris.

[9] CHAMBADAL, L. et OVAERT, J. L. *Cours de Mathématiques – algèbre II*. Gauthier-Villars, Paris.

[10] CONWAY, JOHN H. and SMITH, DEREK A. *On Quaternions and Octonions: their geometry, arithmetic and symmetry.* A. K. Peters Ltd, 2003.

[11] CULLEN, CHARLES G. *Linear Algebra with Applications*. Addison-Wesley, Reading, Mass., 1997.

[12] DIAS AGUDO, F. R. *Introdução à Álgebra Linear e Geometria Analítica*. Livraria Escolar Editora, Lisboa, 1964.

[13] DIEUDONNÉ, J. *Algèbre linéaire et géométrie élémentaire*. Hermann, Paris, 1964.

[14] EFIMOV, N. V. and ROZENDORN, E. R. *Linear Algebra and Multi-dimensional Geometry*. MIR Publishers, Moscow, 1975.

[15] GANTMACHER, F. R. *The Theory of Matrices* (vols. 1 e 2), Chelsea, New York, 1959.

[16] GEL'FAND, I. M. *Lectures on Linear Algebra*, Interscience, New York, 1961.

[17] GUERREIRO, J. S. *Curso de Matemáticas Gerais, Volume 1 – Conjuntos. Noções de Álgebra*, Livraria Escolar Editora, Lisboa, 1973.

[18] HALMOS, Paul R. *Finite Dimensional Vector Spaces*. Springer-Verlag, New York, 1987.

[19] HALMOS, Paul R. *Linear Algebra Problem Book*. The Mathematical Association of America, Washington, 1995.

[20] HAMILTON, WILLIAM ROWAN. *Elements of Quaternions*, 1969.

[21] HANSELMAN, DUANE and LITTLEFIELD, BRUCE. *The Student Edition of MATLAB*. Prentice Hall, New Jersey.

[22] HOFFMAN, K. and KUNZE, R. *Linear Algebra*. 1961.

[23] JACOBSON, N. *Lectures in Abstract Algebra*. Van Nostrand, Princeton, 1951.

[24] KOLMOGOROV, A. N. e FOMIN, S. V. *Elements of Function Theory and Functional Analysis*. MIR Publishers, Moscow, 1975.

[25] LAY, DAVID. *Linear Algebra and its Applications*. 3rd Edition, Addison Wesley, 2003.

[26] LEITE, F. S. e SARAIVA, J. *Generalization of the De Moivre Formulas for Quaternions and Octonions.* Estudos de Matemática in the honour of Luís de Albuquerque, Universidade de Coimbra, 1987.

[27] LIPSCHUTZ, SEYMOUR. *Álgebra Linear*. McGraw-Hill do Brasil, São Paulo, 1968.

[28] MAEDER, ROMAN. *The Mathematica Programmer*. Academic Press, Inc., New York, 1994.

[29] MAGALHÃES, LUÍS T. *Álgebra Linear como introdução à Matemática Aplicada*. Texto Editora, Lisboa, 1990.

[30] MATOS, ISABEL T. *Tópicos de Álgebra Linear*, no site do ISEL em http://www.deetc.isel.ipl.pt/paginaspessoais/isabelteixeira/Alg/documentos/

AL_ApontamentosTeoria.pdf.

[31] MAZZOLA, G., MILMEISTER, G., WEISSMANN, J. *Comprehensive Mathematics for Computer Scientists 1 & 2*, Springer-Verlag, Berlin Heidelberg New York, 2004.

[32] MC DUFFEE, C. C. *The Theory of Matrices*, 1946.

[33] MESSIAS, MANUEL J. A. *Sebenta de Álgebra Linear*. Reprografia do ISEL, Lisboa, 2000.

[34] MONTEIRO, ANTÓNIO. *Álgebra Linear e Geometria Analítica*. Editora McGraw-Hill de Portugal, L^{da}, Lisboa, 2001.

[35] POOLE, DAVID. *Linear Algebra: a modern introduction*. Thomson Brooks/Cole, USA, Canada, Mexico, Spain, 2006.

[36] RIBEIRO, CARLOS M. *Álgebra Linear e Geometria Analítica, Volume 1*, no site do ISEL em http://www.deetc.isel.ipl.pt/paginaspessoais/carlosribeiro/Apontamentos.htm.

[37] RIBEIRO, CARLOS M. *Álgebra Linear e Geometria Analítica, Volume 2*, no site do ISEL em http://www.deetc.isel.ipl.pt/paginaspessoais/carlosribeiro/Apontamentos.htm.

[38] RIBEIRO, CARLOS M. *Álgebra Linear e Geometria Analítica, Volume 3*, no site do ISEL em http://www.deetc.isel.ipl.pt/paginaspessoais/carlosribeiro/Apontamentos.htm.

[39] RIBEIRO, CARLOS M. *Álgebra Linear e Geometria Analítica, Apêndices*, no site do ISEL em http://www.deetc.isel.ipl.pt/paginaspessoais/carlosribeiro/Apontamentos.htm.

[40] RIBEIRO, CARLOS M. *Problemas de Álgebra Linear*, no site do ISEL em http://www.deetc.isel.ipl.pt/paginaspessoais/carlosribeiro/Apontamentos.htm.

[41] RIBEIRO, CARLOS M. *Testes, Exames e suas Correcções*, no site do ISEL em http://www.deetc.isel.ipl.pt/paginaspessoais/carlosribeiro/ProvasAnt.htm.

[42] SERGE LANG. *Algebra*, Addison-Wesley, Reading, Mass., 1965.

[43] SOMMERVILLE, D. M. Y. *Analytical Geometry of Three Dimensions*, 1951.

[44] STOLL, ROBERT R. *Sets, Logic and Axiomatic Theories.* W. H. Freeman and Company, San Francisco, 1974.

[45] SUPPES, P. *Axiomatic Set Theory.* Dover Publications, Inc., New York, 1972.

[46] VÁRIOS AUTORES. *Dictionnaire des Mathématiques: algèbre, analyse, géométrie*. Encyclopædia Universalis Albin Michel, Paris, 1997.

[47] WOLFRAM STEPHEN. *The* MATHEMATICA *Book*, Fourth Edition. Wolfram Media – Cambridge University Press.

ISEL - DEETC

Versão 3.7, Março de 2009

ISEL - DEETC