Álgebra Linear

Licenciatura em Economia

Maria Joana Soares

setembro 2012

1

Matrizes

1.1 Conceitos básicos

Na disciplina de Álgebra Linear, as chamadas *matrizes* são objetos matemáticos que desempenham um papel fundamental, sendo usadas intensivamente. Neste primeiro capítulo, introduzimos o conceito de matriz, estudamos algumas matrizes especiais, definimos as operações básicas com matrizes e estudamos as propriedades dessas operações.

Definição 1.1. Uma *matriz de ordem (ou tipo)* $m \times n$ é simplesmente um quadro retangular de $m \times n$ números dispostos em m linhas e n colunas. Esses números, ditos *elementos* ou *entradas* da matriz, são representados ou entre parênteses curvos (sendo esta a notação que adotaremos neste curso) ou entre parênteses retos.

Nota: A não ser que algo seja dito em contrário, assumimos que os números que constituem a matriz são números reais, isto é, trabalharemos, essencialmente, com as chamadas *matrizes reais*; por vezes, no entanto, consideraremos *matrizes complexas*, ou seja, formadas por números complexos.[1]

O conjunto das matrizes de ordem $m \times n$ de elementos reais será denotado por $\mathbb{R}^{m \times n}$ e o conjunto das matrizes, da mesma ordem, de elementos complexos será designado por $\mathbb{C}^{m \times n}$.

É prática comum usar letras latinas maiúsculas, tais como $A, B, M, N, \ldots$, para denotar matrizes e usar a letra minúscula correspondente, com dois índices, para denotar os respetivos elementos: o primeiro índice indica a que linha pertence o elemento e o segundo refere-se à

[1] É possível definir matrizes com números pertencentes a outro tipo de conjuntos, mas tal está fora do âmbito deste curso.

coluna em que se situa o elemento. Por exemplo,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Também se escreve, em notação abreviada, $A = (a_{ij})_{m\times n}$ ou apenas $A = (a_{ij})$ se o tipo da matriz se deduzir pelo contexto. Se quisermos apenas indicar que A é uma matriz do tipo $m \times n$, escrevemos $A_{m\times n}$. Por vezes, é útil usar a notação $(A)_{ij}$ para designar o elemento situado na linha i e na coluna j da matriz A. Tal elemento é referido como o elemento de A na posição (i, j), ou apenas por elemento (i, j) de A.

Uma *submatriz* de uma dada matriz A é uma matriz obtida de A eliminando alguma(s) das suas linhas e/ou colunas. Por exemplo, se

$$A = \begin{pmatrix} 1 & 3 & 5 & 1 \\ 4 & 2 & 3 & 0 \\ -1 & 3 & 7 & 1 \end{pmatrix},$$

então

$$B = \begin{pmatrix} 4 & 3 & 0 \\ -1 & 7 & 1 \end{pmatrix}$$

é a submatriz de A obtida eliminado a sua primeira linha e a sua segunda coluna.

Definição 1.2. Seja A uma matriz de ordem $m \times n$.

- Se $m = n$, A diz-se uma matriz *quadrada*; se $m \neq n$, a matriz diz-se *retangular*. Quando A é uma matriz quadrada de ordem $n \times n$, é usual dizermos apenas que A é quadrada de ordem n.

- Se $m = 1$, A diz-se uma *matriz linha* ou *vetor linha*, e se $n = 1$, A diz-se uma *matriz coluna* ou *vetor coluna*. Estas matrizes são, geralmente, denotadas por letras minúsculas e em negrito, por exemplo, $\mathbf{b}, \mathbf{u}, \mathbf{v}, \ldots$ Também é usual, no caso de matrizes linha ou coluna, identificar os seus elementos apenas com um índice, por exemplo:

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \qquad \mathbf{u} = (u_1 \ u_2 \ \ldots u_n).$$

Dada uma matriz quadrada $A = (a_{ij})$ de ordem n, dizemos que os elementos $a_{11}, a_{22}, \ldots, a_{nn}$ constituem a *diagonal principal* de A, por vezes referida apenas como diagonal de A:

$$\begin{pmatrix} \underline{a_{11}} & a_{12} & \cdots & a_{1n} \\ a_{21} & \underline{a_{22}} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \underline{a_{nn}} \end{pmatrix}$$

Os elementos $a_{1n}, a_{2,n-1}, \ldots, a_{n1}$ formam a chamada *diagonal secundária* de A.

Definição 1.3 (Matrizes triangulares; matriz diagonal)**.** Uma matriz quadrada $A = (a_{ij})$ diz-se:

- *triangular inferior*, se os elementos situados acima da diagonal principal são todos nulos, i.e. se $a_{ij} = 0$ quando $i < j$;
- *triangular superior*, se os elementos situados abaixo da diagonal principal são todos nulos, i.e. se $a_{ij} = 0$ quando $i > j$;
- *diagonal*, se os elementos situados fora da diagonal principal são todos nulos, i.e. se $a_{ij} = 0$ quando $i \neq j$.

Um caso especialmente importante de uma matriz diagonal é o da matriz identidade de ordem n, I_n, que passamos a definir.

Definição 1.4 (Matriz identidade)**.** Dado $n \in \mathbb{N}$, chama-se *matriz identidade de ordem* n à matriz quadrada de ordem n, diagonal, e cujos elementos diagonais são todos iguais a 1. Esta matriz é denotada por I_n, ou apenas por I, se a ordem se deduzir pelo contexto. Tem-se, então,

$$I_n = (\delta_{ij})_{n \times n}$$

onde δ_{ij} designa o símbolo de Kronecker, definido, para $i, j = 1, \ldots, n$, por

$$\delta_{ij} = \begin{cases} 1, & \text{se} i = j \\ 0, & \text{se} i \neq j \end{cases}$$

Definição 1.5 (Matriz nula)**.** A matriz de ordem $m \times n$ cujos elementos são todos iguais a zero é chamada *matriz nula* e designada pelo símbolo $\mathbf{0}_{m \times n}$ ou, por vezes, simplesmente por $\mathbf{0}$, se a ordem for deduzida pelo contexto.

Exemplo 1.1. Considerem-se as matrizes

$$A = \begin{pmatrix} 1 & 2 & 0 & 5 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 7 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 2 & 1 & 3 & 4 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Então:

- A é triangular superior.
- B é triangular inferior.
- C é, simultaneamente, triangular superior, triangular inferior e diagonal, o mesmo acontecendo a D e a E.
- D é a matriz identidade de ordem 3.
- E é a matriz nula de ordem 3.

Definição 1.6 (Igualdade de matrizes)**.** Duas matrizes $A = (a_{ij})$ e $B = (b_{ij})$ são *iguais* se e só se forem da mesma ordem e os elementos nas posições correspondentes forem iguais, isto é, $a_{ij} = b_{ij}$ para cada escolha de i e j. Se A e B são iguais, escrevemos, como habitualmente $A = B$, escrevendo $A \neq B$ se elas não forem iguais.

Note-se que, de acordo com a definição anterior, se tem, por exemplo,

$$\begin{pmatrix} 1 & 2 & 3 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix},$$

embora os elementos que formam esses dois vetores sejam os mesmos.

1.2 Operações com matrizes

Agora que já conhecemos os principais conceitos básicos de matrizes, vamos aprender a "operar" com elas, isto é, vamos introduzir operações entre matrizes e estudar as suas propriedades.

1.2.1 Adição de matrizes

Definição 1.7 (Adição de matrizes)**.** Se $A = (a_{ij})$ e $B = (b_{ij})$ são duas matrizes da mesma ordem, a *soma* de A e B é uma matriz da mesma ordem, que denotaremos por $A+B$, obtida adicionando as entradas correspondentes. Isto é

$$(A+B)_{ij} = a_{ij} + b_{ij}; \quad \text{para cada } i \text{ e cada } j.$$

Exemplo 1.2. Por exemplo, tem-se:

$$\begin{pmatrix} -2 & 3 & 1 \\ 1 & 5 & -4 \end{pmatrix} + \begin{pmatrix} 2 & -1 & 1 \\ 3 & -2 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 2 \\ 4 & 3 & 2 \end{pmatrix}.$$

Definição 1.8 (Simétrica de uma matriz)**.** Seja A uma matriz. A matriz *simétrica de* A, denotada por $-A$, é a matriz obtida de A substituindo cada um dos seus elementos pelo seu simétrico, i.e.

$$(-A)_{ij} = -a_{ij}, \quad \text{para cada } i \text{ e cada } j. \tag{1.1}$$

Usaremos a notação, $A - B$ para designar a matriz $A + (-B)$.

Exemplo 1.3. Se

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & -1 & -1 \\ 2 & 0 & 4 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ -1 & 0 & 1 \end{pmatrix},$$

então

$$-A = \begin{pmatrix} -1 & -2 & -3 \\ 1 & 1 & 1 \\ -2 & 0 & -4 \end{pmatrix} \quad \text{e} \quad A - B = \begin{pmatrix} 0 & 0 & 0 \\ -4 & -4 & -4 \\ 3 & 0 & 3 \end{pmatrix}.$$

Propriedades da Adição de Matrizes

Para quaisquer matrizes A, B, C da mesma ordem, tem-se:

Comutatividade: $A + B = B + A$

Associatividade: $(A+B) + C = A + (B + C)$

Elemento Neutro: $A + \mathbf{0} = \mathbf{0} + A = A$

Elemento Simétrico : $A + (-A) = (-A) + A = \mathbf{0}$.

As propriedades anteriores são uma consequência imediata das definições da adição de matrizes, simétrica de uma matriz, matriz nula e das propriedades usuais da adição de números reais (ou complexos); a título de exemplo, demonstramos a comutatividade, ficando as restantes demonstrações ao cuidado dos alunos.

Comecemos por notar que, se $A = (a_{ij})$ e $B = (b_{ij})$ são duas matrizes $m \times n$, então, de acordo com a definição de soma de matrizes, ambas as matrizes $A + B$ e $B + A$ são também dessa mesma ordem. Além disso, tem-se:

$$(A+B)_{ij} \underset{①}{=} a_{ij} + b_{ij} \underset{②}{=} b_{ij} + a_{ij} \underset{③}{=} (B+A)_{ij}.$$

A justificação de cada uma das passagens ① − ③ é a seguinte:

① Definição de $A + B$.

② Comutatividade da adição de números reais (ou complexos).

③ Definição de $B + A$.

Podemos, portanto concluir, tendo em conta a definição de igualdade de matrizes, que $A + B = B + A$, como pretendíamos demonstrar. □

Nota: A associatividade da adição permite-nos escrever $A + B + C$, sem qualquer ambiguidade.

1.2.2 Multiplicação escalar

Definição 1.9 (Multiplicação Escalar)**.** Sejam $A = (a_{ij})$ uma matriz e α um número (usualmente designado por *escalar*). A multiplicação do escalar α pela matriz A é uma matriz da mesma ordem que A, designada por αA, e obtida multiplicando todos os elementos de A por α, isto é,

$$(\alpha A)_{ij} = \alpha a_{ij}, \quad \text{para cada } i \text{ e cada } j.$$

A operação de multiplicação de uma matriz por um escalar é designada simplesmente por *multiplicação escalar*.

Exemplo 1.4.

$$3\begin{pmatrix} 1 & 2 & -1 \\ 3 & 4 & 1 \\ -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 6 & -3 \\ 9 & 12 & 3 \\ -3 & 0 & 0 \end{pmatrix}.$$

Propriedades da Multiplicação Escalar

Para quaisquer matrizes $A = (a_{ij})$ e $B = (b_{ij})$ da mesma ordem e quaisquer escalares α, β, tem-se:

Distributividade (em relação à adição de matrizes): $\alpha(A + B) = \alpha A + \alpha B$

Distributividade (em relação à adição de escalares): $(\alpha + \beta)A = \alpha A + \beta A$

Associatividade Mista: $(\alpha\beta)\, A = \alpha\,(\beta A)$

Elemento Identidade : $1A = A$

Elmemento Absorvente : $0A = \mathbf{0}$

Demonstraremos a primeira das propriedades, ficando as restantes como exercício.

Tendo em conta a definição de adição de matrizes e de multiplicação de uma matriz por um escalar, é imediato concluir que $\alpha(A + B)$ e $\alpha A + \alpha B$ são matrizes da mesma ordem. Além disso, temos

$$(\alpha(A+B))_{ij} \underset{①}{=} \alpha(A+B)_{ij} \underset{②}{=} \alpha\,(a_{ij} + b_{ij}) \underset{③}{=} \alpha a_{ij} + \alpha b_{ij} \underset{④}{=} (\alpha A)_{ij} + (\alpha B)_{ij} \underset{⑤}{=} (\alpha A + \alpha B)_{ij}$$

① Definição de multiplicação de uma matriz por um escalar.

② Definição de adição de matrizes.

③ Distributividade da multiplicação em relação à adição (em $\mathbb{R}$ ou $\mathbb{C}$).

④ Definição de multiplicação de uma matriz por um escalar.

⑤ Definição de adição de matrizes.

Concluímos, portanto, que $\alpha(A + B) = \alpha A + \alpha B$.

1.2.3 Transposição e transconjugação

Definição 1.10 (Transposta de uma Matriz)**.** Seja A uma matriz de ordem $m \times n$. Define-se *transposta de* A, e designa-se por A^{T}, como sendo a matriz de ordem $n \times m$ obtida de A trocando as suas linhas com as colunas. Por outras palavras, se $A = (a_{ij})$, então

$$(A^{\mathsf{T}})_{ij} = a_{ji}.$$

Exemplo 1.5. Sendo

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 2 & -1 & 2 \\ 0 & 5 & -3 \\ 0 & 0 & 1 \end{pmatrix}$$

tem-se

$$A^\mathsf{T} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 3 & -1 & 5 & 0 \\ 4 & 2 & -3 & 1 \end{pmatrix}.$$

Segue-se uma lista de propriedades da transposição, cuja demonstração fica a cargo dos alunos.

Propriedades da Transposição

Para quaisquer matrizes $A = (a_{ij})$ e $B = (b_{ij})$ da mesma ordem e qualquer escalar α, tem-se:

- $(A^\mathsf{T})^\mathsf{T} = A$
- $(A + B)^\mathsf{T} = A^\mathsf{T} + B^\mathsf{T}$
- $(\alpha A)^\mathsf{T} = \alpha A^\mathsf{T}$

Definição 1.11 (Matriz simétrica; matriz antissimétrica)**.** Seja A uma matriz quadrada de ordem n. Então:

- A diz-se *simétrica*, se $A = A^\mathsf{T}$, isto é, se $a_{ij} = a_{ji}$.
- A diz-se *antissimétrica* se $A^\mathsf{T} = -A$, isto é, se $a_{ij} = -a_{ji}$.

Numa matriz simétrica, os elementos situados simetricamente em relação à diagonal principal são iguais (e numa matriz antissimétrica, são simétricos, sendo os da diagonal nulos).

Exemplo 1.6. A matriz

$$A = \begin{pmatrix} 1 & 2 & -1 & 4 \\ 2 & 0 & 3 & 1 \\ -1 & 3 & 5 & 7 \\ 4 & 1 & 7 & 9 \end{pmatrix}$$

é uma matriz simétrica e a matriz

$$B = \begin{pmatrix} 0 & 2 & -1 & 6 \\ -2 & 0 & 3 & 1 \\ 1 & -3 & 0 & -7 \\ -6 & -1 & 7 & 0 \end{pmatrix}$$

é antissimétrica.

As noções que daremos a seguir aplicam-se a matrizes complexas.[2]

Definição 1.12 (Conjugada e transconjugada de uma matriz)**.** Seja $A = (a_{ij}) \in \mathbb{C}^{m\times n}$.

- A matriz *conjugada* de A, denotada por $\overline{A}$, é a matriz obtida de A substituindo cada um dos seus elementos pelo seu conjugado.[3] Tem-se, então $(\overline{A})_{ij} = \overline{a_{ij}}$.
- A *transconjugada* de A, denotada por A^*, é a matriz transposta da conjugada de A, isto é, $A^* = (\overline{A})^\mathsf{T}$. Tem-se, portanto $(A^*)_{ij} = \overline{a_{ji}}$.

Definição 1.13 (Matriz hermiteana e matriz anti-hermiteana)**.** Seja A uma matriz quadrada de ordem n com elementos em $\mathbb{C}$. Então:

- A diz-se *hermiteana* ou *hermítica*, se $A = A^*$, isto é, se $a_{ij} = \overline{a_{ji}}$.
- A diz-se *anti-hermiteana* ou *anti-hermítica*, se $A = -A^*$, isto é, se $a_{ij} = -\overline{a_{ji}}$.

Exemplo 1.7. A matriz

$$\begin{pmatrix} 3 & 2+3i & 1-i \\ 2-3i & -5 & 4 \\ 1+i & 4 & 1 \end{pmatrix}$$

é hermítica. A matriz

$$\begin{pmatrix} 0 & 2+3i & -i \\ -2+3i & 0 & 4 \\ -i & -4 & 0 \end{pmatrix}$$

é anti-hermítica.

A transconjugação goza das seguintes propriedades, análogas à da transposição.

Propriedades da Transconjugação

Para quaisquer matrizes $A = (a_{ij}), B = (b_{ij}) \in \mathbb{C}^{m\times n}$ e qualquer escalar $\alpha \in \mathbb{C}$, tem-se:

- $(A^*)^* = A$
- $(A + B)^* = A^* + B^*$
- $(\alpha A)^* = \overline{\alpha} A^*$

[2] Como $\mathbb{R} \subset \mathbb{C}$, uma matriz $A \in \mathbb{C}^{m\times n}$ pode, eventualmente, ser formada apenas por números reais.

[3] Relembre que, se $z = x + ix \in \mathbb{C}$, o seu conjugado, $\overline{z}$, é dado por $\overline{z} = x - iy$.

1.2.4 Produto de matrizes

As operações de adição de matrizes e de multiplicação de uma matriz por um escalar são, de certa forma, muito naturais, isto é, "são aquilo que seria de esperar". O mesmo não se passa, contudo, com o produto de matrizes, sobre o qual nos debruçaremos nesta secção.[4].

Começamos por introduzir a seguinte definição.

Definição 1.14 (Matrizes encadeadas)**.** Duas matrizes A e B (dadas por esta ordem) dizem-se *encadeadas*, se o número de colunas de A for igual ao número de linhas de B. Assim, $A_{m\times n}$ e $B_{p\times q}$ são encadeadas se e só se $n = p$.

Definição 1.15 (Produto de matrizes)**.** Sejam $A_{m\times p}$ e $B_{p\times n}$ duas matrizes encadeadas. Então o produto da matriz A pela matriz B (por esta ordem) é uma matriz de ordem $m\times n$, denotada por AB, e cujo elemento na posição (i,j) $(i = 1,\ldots,m; j = 1,\ldots,n)$ é definido pela fórmula

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj}, \tag{1.2}$$

ou, usando uma notação mais compacta envolvendo o símbolo $\sum$ ("somatório"):

$$(AB)_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}. \tag{1.3}$$

Note que para formar o elemento da linha i, coluna j, da matriz produto AB, se utilizam os elementos da linha i da matriz A e os elementos da coluna j da matriz B; os elementos "correspondentes" são multiplicados e somam-se os produtos resultantes.

Exemplo 1.8. Sejam

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 2 & 1 & 3 & 0 \\ -1 & 1 & 4 & 2 \end{pmatrix}.$$

Como A é uma matriz 2×3 e B é uma matriz 3×4, é possível formar a matriz produto $C = AB$, a qual vai ser uma matriz de ordem 2×4. Para encontrar, por exemplo, o elemento situado na linha 2 e na coluna 3 de $C = AB$, selecionamos a linha 2 de A,

$$\begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$$

[4] A razão de definirmos o produto de uma forma um pouco "estranha" será compreendida mais à frente!

e a coluna 3 de B:

$$\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}.$$

Se multiplicarmos os elementos correspondentes (primeiro com primeiro, segundo com segundo e terceiro com terceiro) e somarmos os produtos obtidos, vem

$$c_{23} = 1 \times 2 + 1 \times 3 + (-1) \times 4 = 2 + 3 - 4 = 1.$$

Determine os restantes elementos da matriz AB e confira o seu resultado:

$$AB = \begin{pmatrix} 2 & 6 & 20 & 7 \\ 4 & 1 & 1 & -1 \end{pmatrix}.$$

Exemplo 1.9.

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 3 & 0 & -3 \end{pmatrix}.$$

O produto de um vetor coluna $m \times 1$ por um vetor linha $1 \times n$ resulta numa matriz $m \times n$. Em particular, o produto de um vetor $n \times 1$ por um vetor $1 \times n$ resulta numa matriz quadrada de ordem n.

Exemplo 1.10.

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \begin{pmatrix} 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 9 \\ 12 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}.$$

O produto de um vetor coluna $\mathbf{u}$ por uma matriz 1×1 com elemento k resulta num vetor coluna igual ao produto do escalar k por $\mathbf{u}$. De modo análogo se vê que o produto de uma matriz 1×1 com elemento k por um vetor linha $\mathbf{v}$ resulta num vetor linha igual ao produto do escalar k pelo vetor $\mathbf{v}$.

Exemplo 1.11.

$$\begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \end{pmatrix}.$$

O produto de um vetor linha $1 \times n$ por um vetor coluna $n \times 1$ origina uma matriz 1×1, a qual é, usualmente, identificada com o número que a constitui, isto é, é usual escrevermos

$$\begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = -2.$$

Exemplo 1.12.

$$\begin{pmatrix}5 & 6\\7 & 8\end{pmatrix}\begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}=\begin{pmatrix}23 & 34\\31 & 46\end{pmatrix}$$

e

$$\begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}\begin{pmatrix}5 & 6\\7 & 8\end{pmatrix}=\begin{pmatrix}19 & 22\\43 & 50\end{pmatrix}.$$

O produto de matrizes **não** é comutativo. Mesmo que AB e BA estejam definidas e sejam matrizes da mesma ordem, não se tem, necessariamente, $AB = BA$.

Exemplo 1.13.

$$\begin{pmatrix}1 & -1\\1 & -1\end{pmatrix}\begin{pmatrix}1 & -1\\1 & -1\end{pmatrix}=\begin{pmatrix}0 & 0\\0 & 0\end{pmatrix}.$$

Se $AB = \mathbf{0}$, **não** podemos concluir que $A = \mathbf{0}$ ou $B = \mathbf{0}$.

Exemplo 1.14. Sejam

$$A=\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix},\quad B=\begin{pmatrix}2 & 2\\2 & 2\end{pmatrix}\quad \text{e} \quad C=\begin{pmatrix}3 & 1\\1 & 3\end{pmatrix}.$$

Então, tem-se

$$AB=\begin{pmatrix}4 & 4\\4 & 4\end{pmatrix}=AC,$$

mas $B \neq C$.

Se A, B e C são matrizes dadas, tais que $AB = AC$, **não** podemos concluir que $B = C$ (mesmo que tenhamos $A \neq \mathbf{0}$).

Acabámos de constatar, através de exemplos, que algumas das propriedades usuais do produto de números, por exemplo, a comutatividade, a lei do anulamento do produto e a chamada lei do corte, não se "estendem" para o produto de matrizes. No entanto, são válidas algumas propriedades, que listamos de seguida.

Propriedades do Produto de Matrizes

Dadas matrizes A, B e C e um escalar α, têm-se as sguintes igualdades, desde que as operações envolvidas estejam definidas:

Associatividade: $(AB)C = A(BC)$

Distributividade à Esquerda: $(A+B)C = AC + BC$

Distributividade à Direita: $A(B+C) = AB + AC$

Elemento Identidade : $AI = A$ e $IA = A$.

Elemento Absorvente: $A\mathbf{0} = \mathbf{0}$ e $\mathbf{0}A = \mathbf{0}$.

Ordem Inversa da Transposta do Produto: $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$

Ordem Inversa da Transconjugada do Produto: $(AB)^* = B^*A^*$.

Associatividade Mista: $\alpha(AB) = (\alpha A)B = A(\alpha B)$

As demonstrações destas propriedades seguem o esquema usual já usado para demonstrar outras propriedades de operações com matrizes. Uma vez que se pretende estabelecer uma igualdade entre matrizes, teremos, primeiramente de mostrar que as matrizes em causa são da mesma ordem, provando, depois, que os elementos correspondentes são iguais. A título de exemplo, provemos que, sendo A uma matriz de tipo $m \times p$ e B uma matriz de tipo $p \times n$ (para que faça sentido efetuar o produto AB), se tem

$$(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}.$$

Sendo A de tipo $m \times p$ e B de tipo $p \times n$, a matriz AB será uma matriz do tipo $m \times n$ e, portanto, $(AB)^{\mathsf{T}}$ será do tipo $n \times m$. Mas, sendo A de tipo $m \times p$, A^{T} será do tipo $p \times m$; do mesmo modo, uma vez que B é do tipo $p \times n$, B^{T} será do tipo $n \times p$. Então, a matriz $B^{\mathsf{T}}A^{\mathsf{T}}$ será, tal como $(AB)^{\mathsf{T}}$, uma matriz do tipo $n \times m$. Vejamos agora que os elementos destas matrizes situados em posições correspondentes, são iguais. Tem-se

$$\begin{aligned}((AB)^{\mathsf{T}})_{ij} = (AB)_{ji} &= \sum_{k=1}^{p} a_{jk}b_{ki} = \sum_{k=1}^{p} b_{ki}a_{jk} \\ &= \sum_{k=1}^{p} (B^{\mathsf{T}})_{ik}(A^{\mathsf{T}})_{kj} = (B^{\mathsf{T}}A^{\mathsf{T}})_{ij},\end{aligned}$$

o que conclui a demonstração. (Justifique as diversas passagens!) $\square$

Nota: Tal como no caso da adição, a associatividade do produto de matrizes permite-nos escrever ABC sem qualquer ambiguidade.

Potências de matrizes quadradas

Definição 1.16 (Potência de uma matriz quadrada)**.** Dada uma matriz A, quadrada, e dado $k \in \mathbb{N}$, a potência k de A, que denotamos por A^k, é a matriz dada por

$$A^k = \underbrace{A.A \ldots .A}_{k \text{ vezes}}.$$

Naturalmente, quando $k = 1$, tem-se $A^1 = A$. Por uma questão de conveniência, convenciona-se que $A^0 = I$.

Produto de matrizes fracionadas em blocos

Para efetuar o produto de duas matrizes, pode ser muito útil, em certos casos, utilizar um método que consiste em considerar uma ou ambas as matrizes "fracionadas"em sub-matrizes – que, neste contexto, são vulgarmente designadas por *blocos* – tal como se descreve a seguir.

Produto de Matrizes Fracionadas

Suponhamos que duas matrizes A e B estão fracionadas em blocos como se indica abaixo:[a]

$$A = \left(\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1r} \\ \hline A_{21} & A_{22} & \ldots & A_{2r} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{s1} & A_{s2} & \cdots & A_{sr} \end{array}\right), \quad B = \left(\begin{array}{c|c|c|c} B_{11} & B_{12} & \cdots & B_{1t} \\ \hline B_{21} & B_{22} & \ldots & B_{2t} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline B_{r1} & B_{r2} & \cdots & B_{rt} \end{array}\right).$$

Note-se que o número de "colunas de blocos"de A é igual ao "número de linhas de blocos"de B, que os blocos que formam cada linha de blocos têm todos o mesmo número de linhas e que os blocos que formam cada coluna de blocos têm todos o mesmo número de colunas; suponhamos, além disso, que o número de linhas de cada bolco A_{ik} é igual ao número de colunas de cada bloco B_{kj}. Então, o produto AB pode formar-se combinando os blocos exatamente da mesma forma como combinamos os escalares no produto usual. Isto é, o bloco na posição (i, j) de AB pode ser obtido usando a seguinte fórmula

$$A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{ir}B_{rj}.$$

[a] As linhas horizontais e verticais da matriz são linhas "imaginárias"que nos ajudam a ver o fraccionamneto da matriz.

O resultado anterior decorre facilmente da forma como se efetua o produto de matrizes.

Exemplo 1.15. Considerem-se as seguintes matrizes fracionadas

$$A = \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right) = \left(\begin{array}{c|c} C & I \\ \hline I & \mathbf{0} \end{array}\right), \quad B = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 1 & 3 & 1 & 3 \\ 2 & 1 & 2 & 1 \end{array}\right) = \left(\begin{array}{c|c} I & \mathbf{0} \\ \hline C & C \end{array}\right),$$

onde $C = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$, I é a matriz identidade de ordem 2 e $\mathbf{0}$ a matriz nula de ordem 2. Então, usando a multiplicação por blocos, o produto pode calcular-se muito facilmente:

$$AB = \left(\begin{array}{c|c} C & I \\ \hline I & \mathbf{0} \end{array}\right) \left(\begin{array}{c|c} I & \mathbf{0} \\ \hline C & C \end{array}\right) = \left(\begin{array}{c|c} 2C & C \\ \hline I & \mathbf{0} \end{array}\right) = \left(\begin{array}{cc|cc} 2 & 6 & 1 & 3 \\ 4 & 2 & 2 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right),$$

uma vez que

$$CI+IC = C+C = 2C,\ \ C\mathbf{0}+IC = \mathbf{0}+C = C,\ \ I\,I+\mathbf{0}C = I+\mathbf{0} = I,\ \ I\mathbf{0}+\mathbf{0}C = \mathbf{0}+\mathbf{0} = \mathbf{0}.$$

O produto com matrizes fracionadas em blocos é também usado, com frequência, para estabelecer certos resultados.

Exemplo 1.16. Consideremos o produto de duas matrizes $A_{m\times p}$ e $B_{p\times n}$. Fracionemos A nos m blocos que são as suas linhas (blocos do tipo $1 \times p$), que designaremos por $\mathbf{l}_1(A), \mathbf{l}_2(A), \ldots, \mathbf{l}_m(A)$ e deixemos B por fracionar, ou, dito de outro modo, consideremos B como um único bloco, de tipo $p \times n$. Tem-se, então[5]

$$AB = \begin{pmatrix} \mathbf{l}_1(A) \\ \mathbf{l}_2(A) \\ \vdots \\ \mathbf{l}_m(A) \end{pmatrix} B = \begin{pmatrix} \mathbf{l}_1(A)\,B \\ \mathbf{l}_2(A)\,B \\ \vdots \\ \mathbf{l}_m(A)\,B \end{pmatrix}.$$

Se designarmos por $\mathbf{l}_1(AB), \ldots, \mathbf{l}_m(AB)$ as m linhas da matriz produto AB, podemos, de imediato concluir que

$$\mathbf{l}_i(AB) = \mathbf{l}_i(A)\,B,$$

[5]Neste caso não julgámos necessário separar as linhas da matriz por linhas horizontais, sendo claro qual é o fracionamento.

ou seja, que:

> para obtermos uma determinada linha i do produto AB, teremos apenas de multiplicar a linha i da matriz A pela matriz B, não sendo, portanto, necessário calcular toda a matriz AB.

De modo análogo, se designarmos por $\mathbf{c}_1(B), \ldots, \mathbf{c}_n(B)$ as colunas da matriz B, tem-se

$$AB = A\Big(\mathbf{c}_1(B)\ \ \mathbf{c}_2(B)\ \ \cdots\ \ \mathbf{c}_n(B)\Big) = \Big(A\,\mathbf{c}_1(B)\ \ A\,\mathbf{c}_2(B)\ \ \cdots\ \ A\,\mathbf{c}_n(B)\Big)$$

o que mostra que, sendo $\mathbf{c}_j(AB)$ a coluna j da matriz AB, se tem

$$\mathbf{c}_j(AB) = A\,\mathbf{c}_j(B).$$

Assim, tem-se:

> para obtermos uma dada coluna j da matriz produto AB, bastará multiplicar A pela coluna j de B.

Notação: De futuro, sendo A uma determinada matriz $m \times n$, usaremos a notação $\mathbf{a}_j$ (a letra minúscula correspondente à letra usada para a matriz, em negrito, indexada por j) para designar a sua j-ésima coluna. Assim,

$$A = \Big(\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_n\Big)$$

será o fracionamento de A nas suas colunas. Uma exceção à regra referida é o caso da matriz identidade de ordem n, I, cujas colunas são, geralmente, designadas por $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Seja A uma matriz $m \times n$ e seja $\mathbf{c}$ uma matriz coluna $n \times 1$. Então o produto $A\mathbf{c}$ pode ser obtido do seguinte modo, usando o fracionamento de A nas suas n colunas e o fracionamento de $\mathbf{c}$ nas suas linhas (cada uma apenas com um elemento)

$$A\mathbf{c} = \Big(\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_n\Big)\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n. \tag{1.4}$$

Definição 1.17. [Combinação linear de vetores] Dados m vetores $\mathbf{x}_1, \ldots, \mathbf{x}_m$ (todos do mesmo tipo), uma *combinação linear* desses vetores é um vetor da forma $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots + \alpha_m\mathbf{x}_m$, com $\alpha_1, \alpha_2, \ldots, \alpha_m$ escalares. Os escalares $\alpha_1, \ldots, \alpha_m$ são chamados os *coeficientes* da combinação linear.

Tem-se, então, o resultado enunciado no quadro seguinte.

O produto de uma matriz A por um vetor coluna $\mathbf{c}$ pode obter-se formando a combinação linear das colunas de A com os elementos de $\mathbf{c}$ como coeficientes.

1.3 Matrizes invertíveis

Definição 1.18 (Matriz invertível)**.** Uma matriz A, quadrada de ordem n, diz-se *invertível* ou *não singular*, se existir uma matriz C, quadrada de ordem n, tal que

$$AC = CA = I_n.$$

Nota: A exigência de que C seja quadrada de ordem n é, na realidade, redundante, pois tal decorre necessariamente do facto de ambos os produtos AC e CA estarem definidos.

Uma matriz que não seja invertível, diz-se *não invertível* ou *singular*.

O seguinte resultado mostra que a matriz C (se existir) é única.

Teorema 1.1. *Seja A uma matriz quadrada de ordem n. Então existe, no máximo, uma matriz C, quadrada de ordem n, tal que $AC = CA = I_n$.*

Dem: Sejam C_1 e C_2 matrizes quadradas de ordem n tais que $AC_1 = C_1A = I_n$ e $AC_2 = C_2A = I_n$. Então, temos

$$C_1 = C_1I_n = C_1(AC_2) = (C_1A)C_2 = I_nC_2 = C_2,$$

o que mostra que existe apenas uma matriz satisfazendo as condições exigidas. □

Definição 1.19 (Inversa de uma matriz)**.** Sendo A uma matriz invertível, à (única) matriz C que satisfaz $AC = CA = I$ chamamos *matriz inversa* de A. Esta matriz é designada pelo símbolo A^{-1}.

Observação importante: Pode provar-se resultado seguinte (neste momento não dispomos ainda de conhecimentos suficientes para o fazer, mas a demonstração será feita mais à frente neste curso):

Se A é uma matriz quadrada de ordem n e C é uma matriz quadrada da mesma ordem e tal que $AC = I$, então também $CA = I$, ou seja, A é invertível e $A^{-1} = C$.

Este resultado é muito útil porque nos diz que, para mostrar que uma matriz C é a inversa de uma matriz (quadrada) A, não teremos que mostrar que $AC = I$ e $CA = I$, **bastando apenas verificar uma das igualdades**.

Exemplo 1.17. A matriz $A = \begin{pmatrix} 6 & 5 \\ 5 & 4 \end{pmatrix}$ é invertível, sendo $A^{-1} = \begin{pmatrix} -4 & 5 \\ 5 & -6 \end{pmatrix}$. Com efeito, temos

$$\begin{pmatrix} 6 & 5 \\ 5 & 4 \end{pmatrix}\begin{pmatrix} -4 & 5 \\ 5 & -6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(Tendo em atenção a observação anterior, bastou-nos efetuar um dos produtos.)

A matriz $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ não tem inversa. De facto, sendo $C = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ uma matriz de ordem 2, tem-se

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x+2z & y+2w \\ x+2z & y+2w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Leftrightarrow \begin{cases} x+2z=1 \\ x+2z=0 \\ y+2w=0 \\ y+2w=1 \end{cases}.$$

Como não é possível encontrar $x, z \in \mathbb{R}$ tais que se tenha simultaneamente $x + 2z = 1$ e $x + 2z = 0$ (nem $y, w \in \mathbb{R}$ tais que $y + 2w = 0$ e $4y + 2w = 1$), é imediato concluir que a matriz dada não tem, de facto, inversa.

Mais à frente neste curso (quando dispusermos de outras ferramentas ...) estudaremos critérios que nos permitem saber, com relativa facilidade, se uma dada matriz é ou não invertível e aprenderemos também a calcular inversas de matrizes invertíveis. Para já, vamos referir algumas propriedades das matrizes invertíveis; a demosntração destas propriedades é deixada como exercício.

Propriedades da Inversão de Matrizes

Se A e B são matrizes quadradas de ordem n, invertíveis, então são válidas as seguintes propriedades.

1. A^{-1} é invertível e $(A^{-1})^{-1} = A$.
2. Para qualquer $\alpha \neq 0$, a matriz αA é invertível e tem-se $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$.
3. A matriz produto AB é invertível e tem-se $(AB)^{-1} = B^{-1}A^{-1}$.
4. A matriz A^{T} é invertível e tem-se $(A^{\mathsf{T}})^{-1} = (A^{-1})^{\mathsf{T}}$.
5. A matriz A^* é invertível e tem-se $(A^*)^{-1} = (A^{-1})^*$

1.4 Exercícios

Exercício 1.1. Dê um exemplo de uma matriz (com mais de duas linhas e duas colunas) que seja:

(a) triangular superior (b) diagonal (c) simétrica (d) antissimétrica.

Exercício 1.2. Sendo

$$A = \begin{pmatrix} 2 & 1 & 3 \\ -2 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 2 \\ -3 & 1 & 1 \\ 2 & 1 & -1 \end{pmatrix} \quad \text{e} \quad C = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \end{pmatrix},$$

diga quais das operações abaixo estão definidas e, nesse caso, efectue essas operações:

(a) $A + B$ (b) $A + C$ (c) $2A - 3B$

(d) AC (e) CA (f) AC^{T}

(g) A^2 (h) C^2 (i) $2A^{\mathsf{T}} - 3B^{\mathsf{T}}$.

Exercício 1.3. Sejam A, B e C matrizes tais que

$$AB = \begin{pmatrix} 5 & 4 \\ 5 & 2 \end{pmatrix} \quad \text{e} \quad AC = \begin{pmatrix} 1 & 8 \\ -2 & 9 \end{pmatrix}.$$

Calcule $A(B + C)$, $B^{\mathsf{T}} A^{\mathsf{T}}$ e $(ABA)C$.

Exercício 1.4. Calcule

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

onde $x_1, x_2 \in \mathbb{R}$.

Exercício 1.5. Sejam

$$A = \begin{pmatrix} 1 & 2 & -4 & 6 & 1 \\ 2 & 1 & 4 & -2 & 2 \\ 1 & 0 & 1 & 0 & 2 \\ -3 & 2 & 1 & 1 & 1 \\ 2 & 4 & 1 & 5 & 4 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 1 & 1 & 3 & 5 \\ 2 & 1 & 0 & 4 & 3 \\ 1 & 0 & 1 & 3 & 7 \\ -3 & 2 & 1 & 9 & 1 \\ 2 & 0 & 1 & 6 & 6 \end{pmatrix}.$$

Calcule:

(a) a terceira linha da matriz AB;

(b) a segunda coluna da matriz AB;

(c) a matriz $A\mathbf{e}_2$, onde $\mathbf{e}_2$ designa a segunda coluna da matriz identidade (neste caso, de ordem 5);

(d) a matriz $\mathbf{e}_2^\mathsf{T} A$ onde $\mathbf{e}_2$ é a matriz referida na alínea anterior.

Exercício 1.6. Considere as matrizes seguintes

$$A = \begin{pmatrix} 2 & 3 & 1 & 4 \\ 1 & 0 & 0 & 0 \\ 2 & 1 & 3 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 2 & 1 & 0 & 4 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}$$

Determine:

(a) a segunda linha de AB;

(b) a terceira coluna de AB;

(c) a quarta linha de AB;

(d) a segunda coluna de AB.

Exercício 1.7. Seja A uma matriz de ordem $m \times n$ e sejam $\mathbf{e}_1, \ldots, \mathbf{e}_n$ as diversas colunas da matriz identidade de ordem n. A que é igual o produto $A\mathbf{e}_j$ ($j = 1, 2, \ldots, n$)?

Exercício 1.8. Sejam $A, B \in \mathbb{R}^{m \times n}$. Prove que:

(a) se $A\mathbf{x} = B\mathbf{x}$ para **todo** o vector $\mathbf{x} \in \mathbb{R}^{n \times 1}$, então $A = B$;

(b) se $A\mathbf{x} = \mathbf{0}$ para **todo** o vector $\mathbf{x} \in \mathbb{R}^{n \times 1}$, então A é a matriz nula.

Sugestão: Use o resultado do exercício anterior.

Exercício 1.9. Sejam

$$A = \begin{pmatrix} 1 & -2 & 1 & 3 \\ 2 & 2 & 1 & 1 \\ 3 & 0 & 1 & 2 \\ -1 & 2 & 0 & 0 \end{pmatrix} \quad \text{e} \quad \mathbf{c} = \begin{pmatrix} 2 \\ 3 \\ -1 \\ 1 \end{pmatrix}$$

Calcule $A\mathbf{c}$:

(a) do modo usual;

(b) formando uma combinção linear (adequada) das colunas de A.

Exercício 1.10. Sejam

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad D' = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{e} \quad A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Calcule: (a) $D\,D'$; (b) $D\,A$; (c) $A\,D$.

Nota: Este exercício pretende ilustrar os seguinte resultados (cuja demonstração, embora simples, omitimos):

1. Se $D = (d_{ij})$ e $D' = (d'_{ij})$ são duas matrizes diagonais da mesma ordem, então DD' também é diagonal e, além disso, tem-se $(DD')_{ii} = d_{ii}d'_{ii}$.
2. Sendo $A = (a_{ij})$ de ordem $m \times n$ e $D = (d_{ij})$ uma matriz $m \times m$, diagonal, a matriz DA obtém-se multiplicando cada uma das **linhas** de A pelo elemento diagonal da linha correspondente de D, i.e. tem-se $(DA)_{ij} = d_{ii}a_{ij}$.
3. Sendo $A = (a_{ij})$ de ordem $m \times m$ e $D = (d_{ij})$ uma matriz $n \times n$, diagonal, a matriz AD obtém-se multiplicando cada uma das **colunas** de A pelo elemento diagonal da coluna correspondente de D, i.e., tem-se $(AD)_{ij} = d_{jj}a_{ij}$.

Exercício 1.11. Sejam A e B duas matrizes quadradas de ordem n. Prove que:

(a) se A e B são simétricas (antissimétricas), então a matriz $A + B$ é simétrica (antissimétrica);

(b) se A é simétrica (antissimétrica), então A matriz αA é simétrica (antissimétrica), para qualquer escalar α;

(c) A matriz $A + A^{\mathsf{T}}$ é simétrica e a matriz $A - A^{\mathsf{T}}$ é antissimétrica;

(d) a matriz A pode decompor-se na soma de uma matriz simétrica com uma matriz antissimétrica;

(e) sendo A e B simétricas, a matriz AB é simétrica se e só se $AB = BA$;

(f) se A é simétrica e invertível, a sua inversa também é simétrica.

Exercício 1.12. Seja

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 1 & 3 & 3 & 2 \\ 2 & 4 & 3 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Mostre que a matriz

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -2 & 2 & -3 \\ 0 & 1 & -1 & 1 \\ -2 & 3 & -2 & 3 \end{pmatrix}$$

é a inversa de A.

Exercício 1.13. Demonstre as propriedades da inversão de matrizes enunciadas na pg. 18.

Exercício 1.14. Uma matriz quadrada A diz-se *idempotente*, se $A^2 = A$.

(a) Mostre que, se A é idempotente, então $A^k = A$, para qualquer $k \in \mathbb{N}$.

(b) Seja $A = I - X(X^\top X)^{-1}X^\top$, onde X é uma matriz $m \times n$ tal que $X^\top X$ é invertível e I é a matriz identidade (de ordem adequada).

(i) Qual será a ordem de I e de A?

(ii) Mostre que A é idempotente.

Exercício 1.15. (a) Mostre que a matriz

$$B = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

é idempotente.

(b) Seja

$$M = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1 & 1/3 & 1/3 & 1/3 \\ \hline 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \end{array}\right).$$

Calcule M^2 e M^3, usando o fracionamento indicado para M. A que será igual a matriz M^{300}?

Exercício 1.16. Considere a seguinte matriz diagonal

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Mostre que a matriz

$$D' = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix}$$

é a inversa de D.

Exercício 1.17. Mostre que se D é uma matriz diagonal de ordem n com elementos diagonais $d_{ii} \neq 0; i = 1, \ldots, n$, então D é invertível e D^{-1} é a matriz diagonal de elementos diagonais iguais a $\frac{1}{d_{ii}}; i = 1, 2, \ldots, n$.

2

Sistemas Lineares

2.1 Introdução

Muitos dos modelos matemáticos usados em Economia conduzem a um sistema de várias equações. Neste capítulo, debrucar-nos-emos sobre o estudo de uma classe particular de sistemas: aqules em que as equações envolvidas são lineares.

Comecemos por relembrar que uma equação linear nas variáveis (ou incógnitas) $x_1, x_2, \ldots, x_n$ é uma equação que possa ser escrita na forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{2.1}$$

onde $a_1, a_2, \ldots, a_n$ e b são números dados.[1] As constantes $a_1, a_2, \ldots a_n$ são os *coeficientes* das incógnitas e b é o *termo independente*.

Exemplo 2.1. A equação

$$2x_1 + x_2 - x_3 = 5$$

é uma equação linear nas variáveis x_1, x_2 e x_3. As equações

$$2x_1x_2 + 5x_3 = 4$$

e

$$\cos x_1 + x_2 - x_3 = 3$$

não são lineares.

[1] Neste curso, se nada for dito em contrário, quando nos referirmos a números, entendemos números reais.

Uma solução da equação (2.1) é uma lista ordenada de n números (isto é, aquilo a que chamamos um n-uplo ordenado) $(s_1, s_2, \ldots, s_n)$ tais que as substituições $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ nessa equação transformam a equação numa proposição verdadeira, isto é, tais que

$$a_1 s_1 + a_2 s_2 + \ldots + a_n s_n = b.$$

Exemplo 2.2. Consideremos a equação

$$2x_1 + x_2 = 5.$$

Então, $(1, 1)$ não é solução dessa equação, uma vez que $2 \times 1 + 1 = 3 \neq 5$, mas $(1, 3)$ é solução, já que $2 \times 1 + 3 = 5$. De facto, qualquer que seja $k \in \mathbb{R}$, $(k, 5 - 2k)$ é solução desta equação. Vemos, assim, que a equação admite uma infinidade de soluções.

Um sistema de equações lineares é uma colecção finita de equações lineares (todas nas mesmas incógnitas) consideradas em conjunto:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}, \tag{2.2}$$

onde a_{ij} e b_i são números dados. Os x_i's são as *variáveis* (ou *incógnitas*) que pretendemos determinar, a_{ij} é o coeficiente da incógnita x_j na i-ésima equação, sendo b_i o *termo independente* dessa mesma equação.

Uma solução do sistema é um n-uplo ordenado de números que é solução de todas as equações que constituem o sistema.

Comecemos por anlisar três casos muito simples de sistemas de duas equações em duas incógnitas x e y.[2]

Relembrando que uma equação linear

$$ax + by = c \quad (a \text{ e } b \text{ não simultaneamente nulos})$$

é a equação de uma reta no plano, vemos que determinar as soluções de um sistema de duas equações lineares nas duas incógnitas x e y corresponde a determinar os pontos de interseção de duas retas num plano. Existem, assim, três possibilidades distintas:

[2]Como temos apenas duas incógnitas, é mais simples e mais comum usar uma notação do tipo x, y para as incógnitas do que usar x_1, x_2; este tipo de procedimemto é usado, em geral, sempre que o número de incógnitas é pequeno.

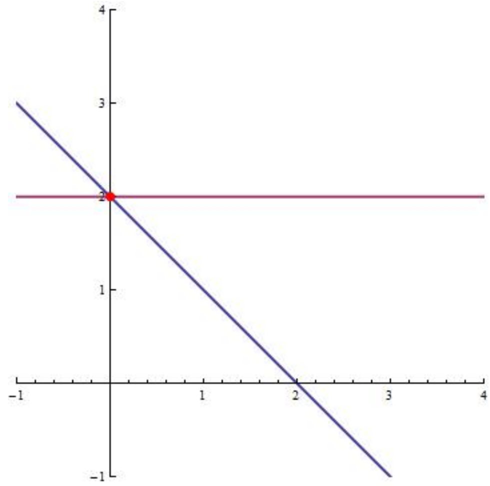

Figura 2.1: Uma única solução

1. as retas dadas são concorrentes num ponto, caso em que o sistema correspondente terá uma única solução;

2. as retas são estritamente paralelas, ou seja, não se intersectam, caso em que o sistema dado não terá solução;

3. as retas são coincidentes, caso em haverá uma infinidade de pontos comuns às duas retas, ou seja, o sistema em causa terá uma infinidade de soluções.

Exemplo 2.3. Consideremos o sistema

$$\begin{cases} x + y = 2 \\ \quad\; 2y = 4 \end{cases}.$$

Da segunda equação, $2y = 4$, resulta $y = 2$; sendo $y = 2$, ter-se-á, da primeira equação, que $x + 2 = 2$ ou seja, virá $x = 0$. Vemos, assim, que $(0, 2)$ é a (única) solução deste sistema. A interpretação geométrica deste sistema é dada na Fig. 2.1.

Exemplo 2.4. O sistema

$$\begin{cases} x + y = 2 \\ x + y = 4 \end{cases}$$

é um sistema sem solução, uma vez que não existem dois números cuja soma seja, simultaneamente, igual a 2 e a 4, ou seja, a exigência sobre as variáveis imposta por uma das equações é incompatível com a da outra equação; veja Fig. 2.2.

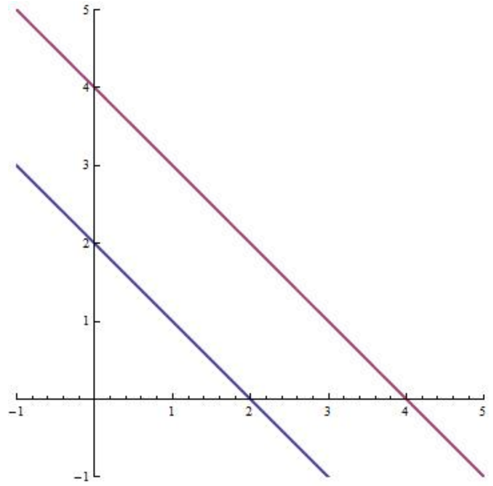

Figura 2.2: Sem solução

Exemplo 2.5. Considere-se o sistema

$$\begin{cases} x + y = 2 \\ 2x + 2y = 4 \end{cases}.$$

Verificamos facilmente que a "informação" fornecida pela primeira equação é idêntica à da segunda, uma vez que, quaisquer que sejam os números s_1 e s_2, se tem

$$2s_1 + 2s_2 = 4 \iff 2(s_1 + s_2) = 4 \iff s_1 + s_2 = 2,$$

pelo que qualquer par de números é solução da primeira equação se e só se também for solução da segunda.

Na prática, tudo se passa como se dispuséssemos apenas de uma equação nas duas incógnitas x e y, a equação $x + y = 2$, sendo fácil de verificar que será solução dessa equação qualquer par da forma $(2 - k, k)$, com $k \in \mathbb{R}$. Este sistema tem, portanto, uma infinidade de soluções; veja Fig. 2.3.

Aquilo que vimos aqui para o caso de sistemas de duas equações com duas incógnitas, é válido para para qualquer sistema de equações lineares, ou seja, pode provar-se que, dado um sistema de equações lineares, existem apenas três possibilidades, no que diz respeito às suas soluções:

1. o sistema não admite solução; diz-se, neste caso, que é *impossível* ou *inconsistente*;

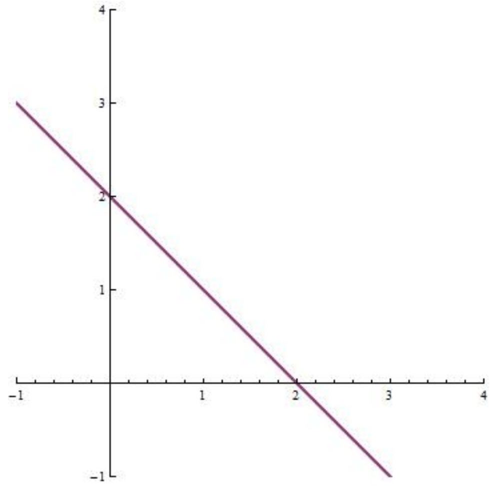

Figura 2.3: Uma infinidade de soluções

2. o sistema admite uma e uma só solução, caso em que se diz *possível e determinado*;

3. o sistema admite uma infinidade de soluções, dizendo-se, então, *possível e indeterminado*.

Quando lidamos com sistemas de equações lineares, estamos geralmente interessados, numa primeira fase, em saber a qual das categorias acima referidas pertence o sistema: é o que chamamos *discutir o sistema*. Sendo o sistema possível, pretendemos geralmente, determinar a sua solução, caso seja determinado, ou descrever o conjunto de todas as suas soluções, caso ele seja indeterminado: trata-se de *resolver o sistema*.

2.2 Eliminação Gaussiana

Definição 2.1. Dois sistemas de equações lineares dizem-se *e*quivalentes se e só se tiverem o mesmo conjunto de soluções.

O *método de eliminação de Gauss* (ou d*de eliminação Gaussiana*) é um processo sistemático de transformar um dado sistema num sistema *equivalente*, mas com uma forma que facilite a sua discussão e resolução.

O método baseia-se no uso das chamadas *operações elementares* de um sistema. Existem três tipos de operações elementares que podem ser efetuadas num sistema:

1. troca da ordem de equações;

2. multiplicação de uma equação por um número diferente de zero;

3. adição a uma equação de outra equação multiplicada por um número.

A importância das operações elementares de um sistema é estabelecida no seguinte teorema (cuja demonstração omitimos).

Teorema 2.1 (Princípio de Equivalência de Sistemas)**.** *Se, dado um sistema de equações lineares, efetuarmos sobre ele uma sequência finita de operações elementares, obtemos um sistema equivalente ao primeiro.*

Vejamos um exemplo muito simples de utilização de operações elementares para a redução de um sistema a uma forma adequada à sua resolução.

Exemplo 2.6. Consideremos o sistema

$$\left\{\begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & x_3 & = & 1 \\ 6x_1 & + & 2x_2 & + & x_3 & = & -1 \\ -2x_1 & + & 2x_2 & + & x_3 & = & 7 \end{array}\right. .$$

Comecemos por tentar "eliminar" a incógnita x_1 das segunda e terceira equações, adicionando-lhes múltiplos adequados da primeira.[3] Se adicionarmos à segunda equação a primeira multiplicada por -3 e adicionarmos a primeira equação à terceira, obtemos o seguinte sistema

$$\left\{\begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & x_3 & = & 1 \\ & - & x_2 & - & 2x_3 & = & -4 \\ & & 3x_2 & + & 2x_3 & = & 8 \end{array}\right. .$$

O nosso próximo objetivo é eliminar a incógnita x_2 da terceira equação, adicionando-lhe um múltiplo da segunda equação. Para tal, bastar-nos-á adicionar à terceira equação a segunda multiplicada por 3, resultando no seguinte sistema:

$$\left\{\begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & x_3 & = & 1 \\ & - & x_2 & - & 2x_3 & = & -4 \\ & & & - & 4x_3 & = & -4 \end{array}\right. .$$

Neste ponto, dizemos que o sistema foi *triangularizado* ou que tem a forma *triangular* (superior). Um sistema triangular superior resolve-se muito facilmente pelo chamado *método de substituição inversa*: da última equação (que envolve apenas a última incógnita) retira-se o

[3] Quando falamos em *múltiplo de uma equação*, queremos referir-nos ao produto dessa equação por um número qualquer não nulo, e não apenas ao produto dessa equação por um número inteiro!

valor da última incógnita; este valor é substituído na penúltima equação, sendo esta resolvida para obter o valor da penúltima incógnita; os valores destas incógnitas substituem-se então na equação anterior para retirar o valor da próxima incógnita, continuando-se, de forma análoga, até que todos os valores das incógnitas sejam encontrados. No nosso exemplo, da última equação

$$-4x_3 = -4$$

resulta $x_3 = 1$. Substituindo na segunda equação, vem

$$-x_2 - 2(1) = -4 \iff -x_2 = -4 + 2 \iff -x_2 = -2 \iff x_2 = 2.$$

Finalmente, substituindo os valores $x_2 = 2$ e $x_3 = 1$ na primeira equação, vem

$$2x_1 + 2 + 1 = 1 \iff 2x_1 = -2 \iff x_1 = -1.$$

Assim, vemos que o sistema dado é possível e determinado, sendo $(-1, 2, 1)$ a sua única solução.

2.2.1 Matriz simples e matriz ampliada de um sistema

Ao aplicarmos as operações elementares sobre um sistema verificamos que estas apenas afetam os coeficientes das incógnitas e os termos independentes, não havendo, portanto necessidade de escrever os símbolos "x_i"e "$=$"em cada passo. Por exemplo, se consideramos o seguinte sistema de equações

$$\left\{\begin{array}{rcrcrcr} 3x_1 & + & x_2 & - & x_3 & = & 4 \\ 5x_1 & & & + & 2x_3 & = & 5 \end{array}\right.,$$

bastará apenas considerar a seguinte matriz 2×4

$$\begin{pmatrix} 3 & 1 & -1 & 4 \\ 5 & 0 & 2 & 5 \end{pmatrix}.$$

Dado um sistema

$$\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right.,$$

à matriz formada pelos coeficientes das incógnitas

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

chamamos *matriz simples do sistema* e à matriz

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

chamamos *matriz ampliada do sistema*. Se A designar a matriz simples do sistema e $\mathbf{b}$ for a matriz coluna $m \times 1$ dos termos independentes, designamos a matriz ampliada do sistema por $(A\ \mathbf{b})$. Ao formar-se a matriz ampliada, é frequente separar-se, por uma linha vertical, a parte correspondente à matriz simples, da coluna dos termos independentes, para destacar o papel especial desta última coluna, isto é, a matriz ampliada do sistema é escrita na forma

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right),$$

usando-se, então, a notação $(A|\mathbf{b})$ para a designar.

Às operações elementares de sistemas correspondem, naturalmente, operações nas linhas da matriz ampliada. Por exemplo, à multiplicação de uma equação por um número diferente de zero corresponde a multiplicação (de todos os elementos) da linha respetiva por esse número.

Operações Elementares sobre Linhas

Dada uma matriz, chamam-se operações elementares sobre as linhas dessa matriz, as seguintes operações:

Tipo O1 Troca de duas linhas.

Tipo O2 Multiplicação de uma linha por um número diferente de zero.

Tipo O3 Substituição de uma linha pela sua soma com uma outra linha multiplicada por um número.

Tendo em conta o princípio de equivalência de sistemas, podemos afirmar que:

Se $(A|\mathbf{b})$ é a matriz ampliada de um sistema e $(E|\mathbf{c})$ é uma matriz obtida de $(A|\mathbf{b})$ por uma sequência de operações elementares sobre as suas linhas, então estas matrizes correspondem a sistemas equivalentes.

Definição 2.2. Uma matriz A diz-se *equivalente por linhas* a uma matriz B, se for possível converter A em B, usando um número finito de operações elementares sobre linhas. Note-se que, se A é uma matriz equivalente por linhas a B, também B é equivalente por linhas a A,[4] por isso poderemos simplesmente dizer que A e B são equivalentes por linhas. Escreve-se, então, $A \overset{\text{linhas}}{\sim} B$.

2.2.2 Matrizes em escada

A ideia do método de eliminação de Gauss é transformar a matriz ampliada $(A|\mathbf{b})$ de um sistema dado, por equivalência de linhas, numa outra matriz $(E|\mathbf{c})$ que tenha uma forma especialmente adequada à discussão e resolução do sistema. Mais precisamente, a forma a que pretendemos chegar é a da chamada *matriz em escada*. Para descrever essa forma, introduzimos primeiramente as seguintes definições.

Definição 2.3. Seja A uma matriz. Dada uma linha não nula de A (isto é, uma linha que não seja toda formada por zeros) chama-se *pivô* dessa linha ao primeiro, a contar da esquerda, elemento não nulo dessa linha. Uma linha nula não tem pivô.

Matriz em Escada

Diz-se que uma matriz é uma *matriz em escada* ou que *tem a forma em escada*, se forem satisfeitas as duas seguintes condições:

1. se uma linha da matriz for nula, então todas as linhas abaixo dela (caso existam) também são nulas, i.e. as linhas nulas da matriz (caso existam) ocupam a parte inferior da matriz;
2. se o pivô da linha i estiver na coluna j, então a linha seguinte (se existir) começa com, pelo menos, j zeros.

Exemplo 2.7. A matriz

$$A = \begin{pmatrix} \boxed{4} & 2 & 0 & 1 & -1 & 3 & 5 & 1 \\ 0 & 0 & \boxed{2} & 1 & 3 & -2 & 0 & 2 \\ 0 & 0 & 0 & \boxed{4} & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

é uma matriz em escada. Os pivôs são os elementos rodeados por um quadrado.

[4] Justifique esta afirmação!

A matriz

$$B = \begin{pmatrix} \boxed{4} & 2 & 0 & 1 & -1 & 3 & 5 & 1 \\ 0 & 0 & \boxed{2} & 1 & 3 & -2 & 0 & 2 \\ 0 & \boxed{3} & 0 & 4 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

não é uma matriz em escada, uma vez que o pivô da linha 2 está na coluna 3 e, na coluna 2 o elemento na linha 3 é não nulo. A matriz

$$C = \begin{pmatrix} \boxed{4} & 2 & 0 & 1 & -1 & 3 & 5 & 1 \\ 0 & 0 & \boxed{2} & 1 & 3 & -2 & 0 & 2 \\ 0 & 0 & 0 & \boxed{4} & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

também não é uma matriz em escada, já que a linha 4 é nula, mas a linha 5 não.

Antes de descrevermos, de uma forma genérica, o método de eliminação de Gauss para a conversão de uma matriz (em geral, matriz ampliada de um sistema), usando operações elementares sobre linhas, numa matriz em escada, vejamos alguns exemplos simples.

Exemplo 2.8. Considere-se a seguinte matriz

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 4 & 3 & 3 \\ 3 & 7 & 1 & 5 \end{pmatrix}. \tag{2.3}$$

A redução à forma em escada pode efetuar-se com se indica abaixo

$$\begin{pmatrix} \boxed{1} & 2 & 1 & 1 \\ 2 & 4 & 3 & 3 \\ 3 & 7 & 1 & 5 \end{pmatrix} \xrightarrow[\substack{L_2-2L_1 \\ L_3-3L_1}]{} \begin{pmatrix} \boxed{1} & 2 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & -2 & 2 \end{pmatrix} \xrightarrow[L_2 \leftrightarrow L_3]{} \begin{pmatrix} \boxed{1} & 2 & 1 & 1 \\ 0 & \boxed{1} & -2 & 2 \\ 0 & 0 & \boxed{1} & 1 \end{pmatrix}.$$

Nota: A notação usada para indicar as operações executadas é a seguinte: $L_i + kL_j$ significa que a i-ésima linha é substituída pela sua soma com a linha j multiplicada por k e $L_i \leftrightarrow L_j$ indica a troca das linhas i e j. Se quisermos indicar a multiplicaçao de uma dada linha i pelo escalar k, escereveríamos kL_i.

Exemplo 2.9. Consideremos agora a seguinte matriz

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 2 & 4 & 0 & 4 & 4 \\ 1 & 2 & 3 & 5 & 5 \\ 4 & 8 & 0 & 8 & 8 \end{pmatrix} \tag{2.4}$$

e vejamos como podemos reduzi-la à forma em escada.

$$\begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 2 & 4 & 0 & 4 & 4 \\ 1 & 2 & 3 & 5 & 5 \\ 4 & 8 & 0 & 8 & 8 \end{pmatrix} \xrightarrow[\substack{L_2-2L_1 \\ L_3-L_1 \\ L_4-4L_1}]{} \begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 0 & 0 & \boxed{\text{-2}} & -2 & -2 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & -4 & -4 & -4 \end{pmatrix} \xrightarrow[\substack{L_3+L2 \\ L_4-2L_2}]{} \begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 0 & 0 & \boxed{\text{-2}} & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Exemplo 2.10. Considere-se agora a seguinte matriz, bastante idêntica à do exemplo anterior:

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 2 & 4 & 0 & 4 & 4 \\ 1 & 2 & 3 & 5 & 5 \\ 4 & 8 & 0 & 8 & 6 \end{pmatrix}. \tag{2.5}$$

Temos, neste caso:

$$\begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 2 & 4 & 0 & 4 & 4 \\ 1 & 2 & 3 & 5 & 5 \\ 4 & 8 & 0 & 8 & 6 \end{pmatrix} \xrightarrow[\substack{L_2-2L_1 \\ L_3-L_1 \\ L_4-4L_1}]{} \begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 0 & 0 & \boxed{\text{-2}} & -2 & -2 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & -4 & -4 & -6 \end{pmatrix} \xrightarrow[\substack{L_3+L_2 \\ L_4-2L_2}]{} \begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 0 & 0 & \boxed{\text{-2}} & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}$$

$$\xrightarrow[L_3 \leftrightarrow L_4]{} \begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 0 & 0 & \boxed{\text{-2}} & -2 & -2 \\ 0 & 0 & 0 & 0 & \boxed{\text{-2}} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Exemplo 2.11. Como último exemplo, consideremos a matriz

$$\begin{pmatrix} 0 & 5 & 1 & -8 \\ 0 & 3 & -3 & 6 \\ 4 & -8 & 12 & 0 \end{pmatrix}. \tag{2.6}$$

Temos, neste caso

$$\begin{pmatrix} 0 & 5 & 1 & -8 \\ 0 & 3 & -3 & 6 \\ 4 & -8 & 12 & 0 \end{pmatrix} \xrightarrow[L_1 \leftrightarrow L_3]{} \begin{pmatrix} \boxed{4} & -8 & 12 & 0 \\ 0 & \boxed{3} & -3 & 6 \\ 0 & 5 & 1 & -8 \end{pmatrix} \xrightarrow[L_3-\frac{5}{3}L_2]{} \begin{pmatrix} \boxed{4} & -8 & 12 & 0 \\ 0 & \boxed{3} & -3 & 6 \\ 0 & 0 & \boxed{6} & -18 \end{pmatrix}.$$

Se estudarmos com cuidado cada um dos exemplos anteriores, não é difícil de concluir que o método que usámos – Método de Eliminação de Gauss – seguiu os passos descritos no quadro seguinte.

Método de Eliminação de Gauss

(Conversão de uma matriz numa matriz em escada)

Para reduzir uma matriz $A_{m\times n}$ à forma em escada, podemos seguir o seguinte processo:

1. Verificamos se a matriz é a matriz nula ou é uma matriz linha; se tal acontecer, ela está já na forma em escada e o processo termina; caso contrário, efetuamos os passos seguintes:

2. Começando da esquerda para a direita, procuramos a primeira coluna não nula da matriz;

3. Sendo j a coluna referida em **2.**, com eventual troca de linhas, colocamos na posição $(1, j)$ um elemento não nulo.

4. Adicionando múltiplos convenientes da linha 1 às linhas $2, \ldots, m$, anulamos todos os elementos da coluna j situados abaixo da posição 1.

5. "Desprezamos" a primeira linha da matriz obtida em **4.** e repetimos todo o processo (a partir de **1.**), aplicado à submatriz resultante.

2.3 Caraterística de uma matriz

Uma vez que há alguma flexibilidade na escolha das operações elementares por linhas usadas para converter uma dada matriz A numa outra matriz E com a forma em escada (nomeadamente, na escolha do elemento a colocar na posição de pivô, quando efetuamos troca de linhas), as entradas de E não estão definidas de forma única, a partir de A. No entanto, pode provar-se que a forma de E é única, no sentido em que *as posições dos pivôs em E estão univocamente determinadas pelas entradas da matriz A.* Isto significa, em particular, que o número de pivôs (que é igual ao número de linhas não nulas de E) também é univocamente determinado pela matriz A. Este número é chamado *caraterística* da matriz A e, como veremos posteriormente, é um número muito importante associado a uma matriz.

Caraterística de uma Matriz

Suponhamos que uma dada matriz A de ordem $m \times n$ é convertida, por operações elementares sobre linhas, numa matriz em escada E. A caraterística de A, que designaremos por car(A), é definida como

$$\begin{aligned} \text{car}(A) &= \text{número de pivôs de } E \\ &= \text{número de linhas não nulas de } E \\ &= \text{número de colunas básicas de } A, \end{aligned}$$

onde as *colunas básicas* (também referidas como *colunas principais*) de A são as colunas de A correspondentes às colunas de E que contêm os pivôs.

Retomando os Exemplos 2.8 – 2.11 vemos que: a matriz (2.3) do Exemplo 2.8 tem caraterística 3, sendo as suas 3 primeiras colunas, colunas principais; a matriz (2.16) do Exemplo 2.9 tem caraterística 2, sendo as suas colunas 1 e 3 as colunas principais; a matriz (2.5) do Exemplo 2.10 tem caraterística 3 e as suas colunas principais são as colunas 1, 3 e 5; finalmente, a matriz (2.6) considerada no Exemplo 2.11 tem caraterística 3 e as suas colunas principais são as colunas 1,2 e 3.

2.4 Resolução de sistemas com matriz em escada

Agora que já sabemos como transformar uma matriz dada numa matriz em escada, por meio de operações elementares sobre as suas linhas, vejamos como usar esse conhecimento para discutir e resolver um sistema.

Exemplo 2.12. Consideremos o sistema

$$\begin{cases} x_1 + 2x_2 + x_3 = 1 \\ 2x_2 + 4x_2 + 3x_3 = 3 \\ 3x_1 + 7x_2 + x_3 = 5 \end{cases}.$$

A matriz ampliada deste sistema é a matriz

$$(A|\mathbf{b}) = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 4 & 3 & 3 \\ 3 & 7 & 1 & 5 \end{array}\right)$$

que é precisamente a matriz (2.3) considerada no Exemplo 2.8 na pg. 34. Já vimos que essa matriz é equivalente por linhas à seguinte matriz em escada:

$$(E|\mathbf{c}) = \left(\begin{array}{ccc|c} \boxed{1} & 2 & 1 & 1 \\ 0 & \boxed{1} & -2 & 2 \\ 0 & 0 & \boxed{1} & 1 \end{array}\right),$$

a qual corresponde ao seguinte sistema (equivalente ao sistema dado):

$$\begin{cases} x_1 + 2x_2 + x_3 = 1 \\ \qquad x_2 - 2x_3 = 2 \\ \qquad\qquad x_3 = 1 \end{cases}.$$

É imediato concluir que, neste caso, vai ser possível encontrar uma e uma única solução para este sistema, por substituição inversa. A última equação diz-nos que $x_3 = 1$; substituindo este valor na segunda equação e resolvendo a equação resultante, vem $x_2 - 2(1) = 2$, de onde se obtém $x_2 = 4$; substituindo os valores de $x_2 = 4$ e $x_3 = 1$ na primeira equação, resulta $x_1 + 2(4) + 1 = 1$, de onde se retira o valor da variável x_1, $x_1 = -8$. Assim, o sistema dado é possível e determinado e a sua solução é $(-8, 4, 1)$.

Observe-se que, neste caso, se tem:

- car(A) = número de pivôs de $E = 3$
- car$(A|\mathbf{b})$ = número de pivôs de $(E|\mathbf{c}) = 3$
- número de incógnitas = número de colunas de $A = 3$.

Exemplo 2.13. Consideremos agora o seguinte sistema, cuja matriz ampliada é a matriz do Exemplo 2.10 na pg. 35:

$$\begin{cases} x_1 + 2x_2 + x_3 + 3x_4 = 3 \\ 2x_1 + 4x_2 + 4x_4 = 4 \\ x_1 + 2x_2 + 3x_3 + 5x_4 = 5 \\ 4x_1 + 8x_2 + 8x_4 = 6 \end{cases}.$$

Retomando o Exemplo 2.10, vemos que a matriz ampliada deste sistema pode ser convertida na seguinte matriz em escada:

$$\left(\begin{array}{cccc|c} \boxed{1} & 2 & 1 & 3 & 3 \\ 0 & 0 & \boxed{-2} & -2 & -2 \\ 0 & 0 & 0 & 0 & \boxed{-2} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

A terceira linha desta matriz ampliada corresponde à equação

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -2,$$

a qual é, naturalmente, uma equação sem solução; assim, o sistema dado é impossível.

Note-se que, neste caso se tem:

- car$(A) = 2$
- car$(A|\mathbf{b}) = 3$.

Exemplo 2.14. Como último exemplo, considere-se o sistema nas variáveis x_1, x_2, x_3 e x_4, cuja matriz ampliada é a matriz do Exemplo 2.9 na pg. 35. Neste caso, após a redução à forma em escada, tem-se a seguinte matriz ampliada:

$$\left(\begin{array}{cccc|c} \boxed{1} & 2 & 1 & 3 & 3 \\ 0 & 0 & \boxed{-2} & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Não há agora nenhuma equação da forma $0x_1+0x_2+0x_3+0x_4 = k$ com $k \neq 0$, ou seja, não encontramos nenhuma equação sem solução; além disso, embora dispuséssemos, inicialmente, de 4 equações para determinar o valor das 4 variáveis, vemos que, na realidade, há apenas 2 equações com "informação relevante", pois duas delas foram transformadas numa identidade $0 = 0$, ou seja, o sistema dado é equivalente ao seguinte sistema *reduzido*:

$$\begin{cases} x_1 + 2x_2 + x_3 + 3x_3 = 3 \\ \qquad\qquad -2x_3 - 2x_4 = -2 \end{cases}.$$

Vemos, então, que não vai ser possível determinar univocamente o valor das 4 variáveis: duas delas vão ser *livres* ou *independentes* (isto é, vão poder tomar valores arbitrários), e as outras duas, ditas *principais* ou *dependentes*, irão tomar valores que dependem dos valores atribuídos às variáveis livres; por outras palavras, as variáveis dependentes vão ser função das variáveis livres. Note-se, que, neste caso se tem:

- car$(A) = 2$
- car$(A|\mathbf{b}) = 2$
- número de variáveis $= 4$
- número de variáveis livres$= 4 - 2 =$número de variáveis$-$ car(A).

Vamos considerar como variáveis principais aquelas que correspondem às colunas principais, ou, dito de outro modo, aquelas que estão associadas aos pivôs;[5] neste caso, serão principais as variáveis x_1 e x_3, sendo as variáveis x_2 e x_4 livres; atribuindo valores arbitrários a x_2 e x_4 isto é, considerando $x_2 = \alpha$ e $x_4 = \beta$, $\alpha, \beta \in \mathbb{R}$, e substituindo esses valores nas duas primeiras equações, tem-se

$$\begin{cases} x_1 + 2\alpha + x_3 + 3\beta = 3 \\ \qquad\quad -2x_3 - 2\beta = -2 \\ \qquad\qquad\qquad x_2 = \alpha \\ \qquad\qquad\qquad x_4 = \beta \end{cases}.$$

Passando para o lado direito das equações os termos com α e β, vem

$$\begin{cases} x_1 + x_3 = 3 - 2\alpha - 3\beta \\ \quad -2x_3 = -2 + 2\beta \\ \qquad x_2 = \alpha \\ \qquad x_4 = \beta \end{cases}$$

e as duas primeiras equações deste sistema estão prontas a ser resolvidas por substituição inversa:

$$-2x_3 = -2 + 2\beta \iff x_3 = -\frac{1}{2}(-2 + 2\beta) \iff x_3 = 1 - \beta;$$

substituindo a expressão de x_3 na primeira equação e resolvendo em ordem a x_1 virá:

$$x_1 + (1 - \beta) = 3 - 2\alpha - 3\beta \iff x_1 = 2 - 2\alpha - 2\beta$$

Vemos, então, que o conjunto de soluções do sistema dado é o conjunto

$$\{(2 - 2\alpha - 2\beta, \alpha, 1 - \beta, \beta) : \alpha, \beta \in \mathbb{R}\}.$$

Soluções particulares do sistema poderão ser encontradas atribuindo valores específicos a α e a β. Por exemplo, uma solução possível do sistema será $(0, 1, 1, 0)$, a qual corresponde à escolha $\alpha = 1$, $\beta = 0$.

Os exemplos que acabámos de considerar incluem cada um dos casos que nos podem surgir quando pretendemos discutir e resolver um sistema linear.

Descrevemos, de seguida, o procedimento sistemático que devemos adotar quando pretendermos discutir e, sendo possível, resolver, um sistema de equações lineares, com base no uso do processo de eliminação de Gauss.

[5] Isto não é estritamente necessário, mas, neste curso, por uma questão de simplicidade, seguiremos sempre esta metodologia.

Discussão e Resolução de um Sistema

(Redução à forma em escada por eliminação de Gauss + substituição inversa)

Seja dado um sistema de m equações lineares em n incógnitas e seja $(A|\mathbf{b})$ a respetiva matriz ampliada.

1. Para **discutir** esse sistema, podemos proceder do seguinte modo:

1.1 Convertemos a matriz ampliada do sistema numa matriz em escada $(E|\mathbf{c})$ de modo a determinarmos car(A) e car$(A|\mathbf{b})$.

1.2 Se car$(A) <$ car$(A|\mathbf{b})$, concluímos que o sistema é impossível.

1.3 Se car$(A) =$ car$(A|\mathbf{b})$, concluímos que o sistema é possível; nesse caso, ele será *determinado*, se tivermos car$(A) = n$ e *indeterminado*, se car$(A) < n$; neste último caso, o seu *grau de indeterminação* (ou seja, o número de variáveis livres) é dado por $n -$ car(A).

2. Se car$(A) =$ car$(A|\mathbf{b}) = r$ e pretendermos **resolver** o sistema, podemos proceder do seguinte modo, a partir do sistema cuja matriz ampliada é $(E|\mathbf{c})$.

2.1 Se $r = n$ (caso em que o sistema é determinado), usamos o método de substituição inversa, determinando o valor das incógnitas da última para a primeira.

2.2 Se $r < n$:

(i) Identificamos as r colunas de E com pivôs e consideramos as respetivas variáveis como variáveis principais, sendo as restantes $n-r$ variáveis, variáveis livres.

(ii) Supondo que as variáveis principais são as variáveis $x_{p_1}, \ldots, x_{p_r}$ e que $x_{\ell_1}, \ldots, x_{\ell_{n-r}}$ são as variáveis livres, fazemos $x_{\ell_1} = \alpha_1, x_{\ell_2} = \alpha_2, \ldots, x_{\ell_{n-r}} = \alpha_{n-r}$, com $\alpha_i \in \mathbb{R}$.

(iii) Substituímos as expressões acima nas primeiras r equações do sistema em escada, passamos esses termos para o lado direito das equações e resolvemos o sistema correspondente, em ordem às r variáveis principais, por substituição inversa; os valores das variáveis principais $x_{p_1}, \ldots, x_{p_r}$ virão dados em função de $\alpha_1, \ldots, \alpha_{n-r}$.

2.5 Método de Gauss-Jordan

O método de resolver sistemas descrito na secção anterior foi baseado na redução da matriz ampliada do sistema a uma matriz em escada. É possível, usando apenas operações elementares sobre linhas, converter uma matriz numa forma ainda mais simplificada, a chamada *forma em escada reduzida*.

Matriz em Escada Reduzida

Diz-se que uma matriz é uma *matriz em escada reduzida* ou que *tem a forma em escada reduzida*, se forem satisfeitas as seguintes condições:

1. a matriz tem a forma em escada;
2. os pivôs são todos iguais a 1;
3. as colunas que contêm um pivô têm todos os elementos, à excepção do pivô, iguais a zero.

Exemplo 2.15. A matriz

$$A = \begin{pmatrix} \boxed{1} & 2 & 0 & 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

tem a forma em escada reduzida, mas a matriz B abaixo, sendo embora muito idêntica a A, não é uma matriz em escada reduzida (porquê?).

$$B = \begin{pmatrix} \boxed{1} & 2 & 0 & 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & \boxed{1} & 0 & 2 & 1 & 3 & 1 \\ 0 & 0 & 0 & \boxed{1} & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

A matriz

$$C = \begin{pmatrix} \boxed{1} & 2 & 0 & 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & \boxed{2} & 0 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

também não é uma matriz em escada reduzida.

Vejamos um exemplo da redução à forma em escada reduzida de uma matriz A, usando operações elementares sobre linhas.

Exemplo 2.16. Consideremos a matriz já usada no Exemplo 2.10 na pg. 35:

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 2 & 4 & 0 & 4 & 4 \\ 1 & 2 & 3 & 5 & 5 \\ 2 & 4 & 0 & 4 & 4 \end{pmatrix}.$$

Tem-se

$$\begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 2 & 4 & 0 & 4 & 4 \\ 1 & 2 & 3 & 5 & 5 \\ 2 & 4 & 0 & 4 & 7 \end{pmatrix} \xrightarrow[\substack{L_2-2L_1 \\ L_3-L_1 \\ L_4-2L_1}]{} \begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 0 & 0 & -2 & -2 & -2 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & -2 & -2 & 1 \end{pmatrix} \xrightarrow[-\frac{1}{2}L_2]{} \begin{pmatrix} \boxed{1} & 2 & 1 & 3 & 3 \\ 0 & 0 & \boxed{1} & 1 & 1 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & -2 & -2 & 1 \end{pmatrix}$$

$$\xrightarrow[\substack{L_3-2L_2 \\ L_4+2L_2 \\ L_1-L_2}]{} \begin{pmatrix} \boxed{1} & 2 & 0 & 2 & 2 \\ 0 & 0 & \boxed{1} & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \xrightarrow[L_4 \leftrightarrow L_3]{} \begin{pmatrix} \boxed{1} & 2 & 0 & 2 & 2 \\ 0 & 0 & \boxed{1} & 1 & 1 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\frac{1}{3}L_3]{} \begin{pmatrix} \boxed{1} & 2 & 0 & 2 & 2 \\ 0 & 0 & \boxed{1} & 1 & 1 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[\substack{L_2-L_3 \\ L_1-2L_3}]{} \begin{pmatrix} \boxed{1} & 2 & 0 & 2 & 0 \\ 0 & 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

A redução à forma em escada reduzida de uma matriz por aplicação de operações elementares sobre as suas linhas constitui o chamado *método de Gauss-Jordan*. É fácil de perceber (através dos exemplos anteriores) de que forma podemos proceder para obter essa forma em escada reduzida. O processo é idêntico ao que descrevemos para obter a forma em escada, mas haverá que, adicionalmente:

1. garantir que os pivôs são iguais a 1, o que poderá ser sempre feito, multiplicando as linhas não nulas (i.e., as que têm pivô) por escalares convenientes;

2. garantir que, nas colunas que contêm um pivô, os elementos acima do pivô também são nulos, o que poderá ser feito adicionando múltiplos convenientes da linha de um pivô, não apenas às linhas situadas abaixo dessa linha, mas também às linhas situadas acima dessa linha.

Observação importante: Pode provar-se que, quando convertemos uma dada matriz A na forma em escada reduzida, usando operações elementares sobre linhas, mesmo que não sigamos sempre o mesmo processo, a matriz a que chegaremos será sempre a mesma, isto é, para além de ter sempre a mesma forma (como se passa na simples redução à forma em escada) tem os elementos nas diversas posições univocamente determinados. Por outras palavras, existe **uma única** matriz em escada reduzida, produzida a partir de A por operações elementares sobre linhas, ou seja, equivalente por linhas à matriz A. Neste curso, usamos a notação er(A) para designar tal matriz.

É fácil de ver (justifique!) que, se A for uma matriz quadrada de ordem n e tiver característica igual a n, a respetiva matriz em escada reduzida será a matriz identidade de ordem n, isto é, ter-se-á

$$\text{er}(A) = I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Naturalmente, o recíproco é também verdadeiro, isto é, se A for uma matriz cuja respetiva matriz em escada seja a matriz identidade de ordem n, então car(A) $= n$.

Resumindo, tem-se o resultado seguinte.

Sendo A quadrada de ordem n, tem-se

$$\text{car}(A) = n \iff \text{er}(A) = I_n.$$

Vejamos agora como resolver um sistema de equações, fazendo uso do método de Gauss-Jordan para reduzir a sua matriz ampliada à forma em escada reduzida.

Exemplo 2.17. Como primeiro exemplo, consideremos o sistema

$$\begin{cases} 2x_1 + x_2 + 2x_3 = 0 \\ x_1 - x_2 - x_3 = 2 \\ 3x_1 - x_2 - x_3 = 4 \end{cases}.$$

A sua matriz ampliada é a seguinte:

$$(A|\mathbf{b}) = \left(\begin{array}{ccc|c} 2 & 1 & 2 & 0 \\ 1 & -1 & -1 & 2 \\ 3 & -1 & -1 & 4 \end{array}\right).$$

Convertamos, então, esta matriz na forma em escada reduzida:

$$\left(\begin{array}{ccc|c} 2 & 1 & 2 & 0 \\ 1 & -1 & -1 & 2 \\ 3 & -1 & -1 & 4 \end{array}\right) \xrightarrow[\frac{1}{2}L_1]{} \left(\begin{array}{ccc|c} \boxed{1} & \frac{1}{2} & 1 & 0 \\ 1 & -1 & -1 & 2 \\ 3 & -1 & -1 & 4 \end{array}\right) \xrightarrow[\substack{L_2-L_1 \\ L_3-3L_1}]{} \left(\begin{array}{ccc|c} \boxed{1} & \frac{1}{2} & 1 & 0 \\ 0 & -\frac{3}{2} & -2 & 2 \\ 0 & -\frac{5}{2} & -4 & 4 \end{array}\right)$$

$$\xrightarrow[-\frac{2}{3}L_2]{} \left(\begin{array}{ccc|c} \boxed{1} & \frac{1}{2} & 1 & 0 \\ 0 & \boxed{1} & \frac{4}{3} & -\frac{4}{3} \\ 0 & -\frac{5}{2} & -4 & 4 \end{array}\right) \xrightarrow[\substack{L_3+\frac{5}{2}L_2 \\ L_1-\frac{1}{2}L_2}]{} \left(\begin{array}{ccc|c} \boxed{1} & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & \boxed{1} & \frac{4}{3} & -\frac{4}{3} \\ 0 & 0 & -\frac{2}{3} & \frac{2}{3} \end{array}\right) \xrightarrow[-\frac{3}{2}L_3]{} \left(\begin{array}{ccc|c} \boxed{1} & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & \boxed{1} & \frac{4}{3} & -\frac{4}{3} \\ 0 & 0 & \boxed{1} & -1 \end{array}\right)$$

$$\xrightarrow[\substack{L_2-\frac{4}{3}L_3 \\ L_1-\frac{1}{3}L_3}]{} \left(\begin{array}{ccc|c} \boxed{1} & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & -1 \end{array}\right) = \mathsf{er}\,(A|\mathbf{b}).$$

Vemos, então, que o sistema dado é equivalente ao sistema

$$\begin{cases} x_1 = 1 \\ x_2 = 0 \\ x_3 = -1 \end{cases},$$

o qual é, obviamente, possível e determinado, com solução $(1, 0, -1)$.

Exemplo 2.18. Seja agora dado o sistema

$$\begin{cases} x_1 - x_2 - x_3 = 1 \\ 2x_1 + x_2 - 5x_3 = 2 \\ -x_1 + x_2 = -1 \end{cases},$$

ao qual corresponde a seguinte matriz ampliada

$$(A|\mathbf{b}) = \left(\begin{array}{ccc|c} 1 & -1 & -1 & 1 \\ 2 & 1 & -5 & 2 \\ -1 & 2 & 0 & -1 \end{array}\right).$$

Tem-se, então

$$\left(\begin{array}{ccc|c} \boxed{1} & -1 & -1 & 1 \\ 2 & 1 & -5 & 2 \\ -1 & 2 & 0 & -1 \end{array}\right) \xrightarrow[\substack{L_2-2L_1 \\ L_3+L_1}]{} \left(\begin{array}{ccc|c} \boxed{1} & -1 & -1 & 1 \\ 0 & 3 & -3 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right) \xrightarrow[\frac{1}{3}L_2]{} \left(\begin{array}{ccc|c} \boxed{1} & -1 & -1 & 1 \\ 0 & \boxed{1} & -1 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right)$$

$$\xrightarrow[\substack{L_3-L_2 \\ L_1+L_2}]{} \left(\begin{array}{ccc|c} \boxed{1} & 0 & -2 & 1 \\ 0 & \boxed{1} & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) = \mathsf{er}\,(A|\mathbf{b}).$$

Neste caso, obtém-se o seguinte sistema, equivalente ao dado:

$$\begin{cases} x_1 - 2x_3 = 1 \\ \quad x_2 - x_3 = 0 \end{cases}.$$

A análise da matriz er $(A|\mathbf{b})$ mostra-nos que o sistema é possível, e *simplesmente indeterminado* (isto é, com grau de indeterminação igual a 1); as variáveis principais são x_1 e x_2, sendo x_3 variável livre. Fazendo $x_3 = \alpha$ ($\alpha \in \mathbb{R}$), substituindo nas equações do sistema reduzido e passando os termos com α para o lado direito do sistema, vem

$$\begin{cases} x_1 = 1 + 2\alpha \\ x_2 = \alpha \\ x_3 = \alpha \end{cases}.$$

Concluímos então que

$$\{(1 + 2\alpha, \alpha, \alpha) : \alpha \in \mathbb{R}\}$$

é o conjunto de todas as soluções do sistema.

Se efetuarmos uma contagem no número de operações envolvidas na resolução de um sistema por redução à forma em escada, usando eliminação de Gauss, seguida de substituição inversa, e o compararmos com o número de operações envolvidas na aplicação do método de Gauss-Jordan, baseado na redução à forma em escada reduzia, poderemos concluir que o primeiro método é (em geral) mais eficiente. No entanto, se pretendermos resolver vários sistemas que tenham em comum a mesma matriz simples (isto é, que difiram apenas nos termos independentes), o método de Gauss-Jordan poderá ser competitivo, se resolvermos os vários sistemas **em simultâneo**. Vejamos um exemplo da aplicação do método de Gauss-Jordan à resolução simultânea de vários sistemas com a mesma matriz simples.

Exemplo 2.19. Considerem-se os seguintes três sistemas, com a mesma matriz de coeficientes:

$$S_1 \equiv \begin{cases} 4x - 8y + 5z = 1 \\ 4x - 7y + 4z = 0 \\ 3x - 4y + 2z = 0 \end{cases}, \quad S_2 \equiv \begin{cases} 4x - 8y + 5z = 0 \\ 4x - 7y + 4z = 1 \\ 3x - 4y + 2z = 0 \end{cases}, \quad S_3 \equiv \begin{cases} 4x - 8y + 5z = 0 \\ 4x - 7y + 4z = 0 \\ 3x - 4y + 2z = 1 \end{cases}.$$

Em vez de consideramos um sistema de cada vez, formemos uma matriz ampliada da forma

$$(A|\mathbf{b_1}|\mathbf{b_2}|\mathbf{b_3})$$

onde $\mathbf{b_1}$, $\mathbf{b_2}$ e $\mathbf{b_3}$ são as colunas com os lados direitos dos sistemas S_1, S_2 e S_3 respetivamente e usemos o processo de Gauss-Jordan aplicado a essa matriz:

$$\left(\begin{array}{ccc|c|c|c} 4 & -8 & 5 & 1 & 0 & 0\\ 4 & -7 & 4 & 0 & 1 & 0\\ 3 & -4 & 2 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\frac{1}{4}L_1]{} \left(\begin{array}{ccc|c|c|c} \boxed{1} & -2 & \frac{5}{4} & \frac{1}{4} & 0 & 0\\ 4 & -7 & 4 & 0 & 1 & 0\\ 3 & -4 & 2 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\substack{L_2-4L_1\\ L_3-3L_1}]{} \left(\begin{array}{ccc|c|c|c} \boxed{1} & -2 & \frac{5}{4} & \frac{1}{4} & 0 & 0\\ 0 & \boxed{1} & -1 & -1 & 1 & 0\\ 0 & 2 & -\frac{7}{4} & -\frac{3}{4} & 0 & 1 \end{array}\right)$$

$$\xrightarrow[\substack{L_3-2L_2\\ L_1+2L_2}]{} \left(\begin{array}{ccc|c|c|c} \boxed{1} & 0 & -\frac{3}{4} & -\frac{7}{4} & 2 & 0\\ 0 & \boxed{1} & -1 & -1 & 1 & 0\\ 0 & 0 & \frac{1}{4} & \frac{5}{4} & -2 & 1 \end{array}\right) \xrightarrow[4L_3]{} \left(\begin{array}{ccc|c|c|c} \boxed{1} & 0 & -\frac{3}{4} & -\frac{7}{4} & 2 & 0\\ 0 & \boxed{1} & -1 & -1 & 1 & 0\\ 0 & 0 & \boxed{1} & 5 & -8 & 4 \end{array}\right)$$

$$\xrightarrow[\substack{L_2+L_3\\ L_1+\frac{3}{4}L_3}]{} \left(\begin{array}{ccc|c|c|c} \boxed{1} & 0 & 0 & 2 & -4 & 3\\ 0 & \boxed{1} & 0 & 4 & -7 & 4\\ 0 & 0 & \boxed{1} & 5 & -8 & 4 \end{array}\right).$$

As soluções dos três sistemas lêem-se, de imediato, da matriz obtida: elas são $(2, 4, 5)$ para o sistema S_1, $(-4, -7, -8)$ para o sistema S_2 e $(3, 4, 4)$ para o sistema S_3.

2.6 Sistemas homogéneos

Um sistema linear diz-se *homogéneo* se for da forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0\\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0\\ \qquad\qquad \vdots\\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}, \tag{2.7}$$

isto é, se os termos independentes de todas as equações forem iguais a zero. É imediato concluir que $x_1 = 0, x_2 = 0, \ldots, x_n = 0$, satisfazem todas as equações do sistema homogéneo (2.7) ou seja, que ele é **sempre possível**, admitindo (pelo menos) a solução $(0, 0, \ldots, 0)$, dita solução *nula* ou *trivial*. Uma outra forma de vermos que o sistema é sempre possível é a seguinte. Sendo $(A|\mathbf{0})$ a matriz ampliada do sistema e analisando em que consistem as operações elementares sobre linhas de uma matriz, é imediato concluir que, ao convertermos a matriz $(A|\mathbf{0})$ numa matriz em escada, a última coluna nula manter-seá sempre nula ao longo de todo o processo, pelo que, no final, teremos uma matriz da forma $(E|\mathbf{0})$. Isto significa que o número de pivôs de E é igual ao número de pivôs de $(E|\mathbf{0})$, ou dito de outro modo, que a característica da matriz simples coincide com a característica da matriz ampliada, condição que, como sabemos, nos garante que o sistema é possível.

Para sabermos se o sistema é determinado ou indeterminado, teremos, como para qualquer outro sistema, de comparar a característica da matriz A com o número de incógnitas. Se car$(A) = n$, o sistema será determinado, sendo a sua única solução o n-uplo $(0, 0, \ldots, 0)$. Se car$(A) < n$, o sistema será indeterminado, podendo resolver-se como habitualmente (identificando as variáveis principais e as variáveis livres e resolvendo o sistema em ordem às variáveis principais, em função dos valores das variáveis livres).

Em resumo, tem-se o resultado contido no quadro seguinte.

Sistemas Homogéneos

Considere-se um sistema homogéneo cuja matriz simples é $A_{m\times n}$. Então:

- Esse sistema é sempre possível, sendo o n-uplo $(0, 0, \ldots, 0)$ uma sua solução.
- Se car$(A) = n$, esse sistema é determinado (tendo apenaa a solução nula) e, se car$(A) < n$, esse sistema é indeterminado.

Nota: Uma vez que a coluna dos termos independentes de um sistema homogéneo se mantém sempre nula ao longo de todo o processo de eliminação Gaussiana, quando resolvemos um sistema deste tipo é usual fazer a eliminação apenas da matriz simples do sistema, omitindo a coluna dos termos independentes.

Exemplo 2.20. Considere-se o seguinte sistema homogéneo

$$\begin{cases} x_1 + 2x_2 - 2x_3 = 0 \\ 4x_1 + 8x_2 - 5x_3 = 0 \\ 3x_1 + x_2 - 6x_3 = 0 \end{cases}.$$

A matriz simples deste sistema é

$$A = \begin{pmatrix} 1 & 2 & -2 \\ 4 & 8 & -5 \\ 3 & 1 & -6 \end{pmatrix}.$$

Usando eliminação de Gauss, convertamos A na forma em escada:

$$\begin{pmatrix} \boxed{1} & 2 & -2 \\ 4 & 8 & -5 \\ 3 & 1 & -6 \end{pmatrix} \xrightarrow[\substack{L_2-4L_1 \\ L_3-3L_1}]{} \begin{pmatrix} \boxed{1} & 2 & -2 \\ 0 & 0 & 3 \\ 0 & -5 & 0 \end{pmatrix} \xrightarrow[L_2\leftrightarrow L_3]{} \begin{pmatrix} \boxed{1} & 2 & -2 \\ 0 & \boxed{-5} & 0 \\ 0 & 0 & \boxed{3} \end{pmatrix}.$$

Vemos, então, que car$(A) = 3 =$número de incógnitas, pelo que o sistema dado é determinado com única solução $(0, 0, 0)$.

Exemplo 2.21. Consideremos o sistema

$$\begin{cases} x_1 - x_2 + 2x_3 = 0 \\ 3x_1 + x_2 - 2x_3 = 0 \end{cases}.$$

Temos

$$A = \begin{pmatrix} \boxed{1} & -1 & 2 \\ 3 & 1 & -2 \end{pmatrix} \xrightarrow[L_2 - 3L_1]{} \begin{pmatrix} \boxed{1} & -1 & 2 \\ 0 & \boxed{4} & -8 \end{pmatrix}.$$

Como car$(A) = 2$ e temos três incógnitas, o sistema é simplesmente indeterminado. As variáveis principais são x_1 e x_2 e a variável livre é x_3. Fazendo $x_3 = \alpha$ ($\alpha \in \mathbb{R}$) e substituindo nas equações do sistema correspondente à última matriz obtida acima, vem

$$\begin{cases} x_1 - x_2 + 2\alpha = 0 \\ 4x_2 - 8\alpha = 0 \\ x_3 = \alpha \end{cases}.$$

Tem-se, então

$$4x_2 - 8\alpha = 0 \iff 4x_2 = 8\alpha \iff x_2 = 2\alpha.$$

Substituindo na primeira equação e resolvendo-a em ordem a x_1, virá

$$x_1 - 2\alpha + 2\alpha = 0 \iff x_1 = 0.$$

Assim, o sistema dado tem o seguinte conjunto de soluções

$$\{(0, 2\alpha, \alpha) : \alpha \in \mathbb{R}\}.$$

2.7 Sistemas com matriz quadrada

Pela sua especial importância, salientamos, no quadro seguinte, alguns resultados para o caso particular de sistemas em que o número de equações é igual ao número de incógnitas (ou seja, em que a matriz do sistema é uma matriz quadrada). Deixamos ao cuidado do aluno justificar cada uma das afirmações, que são simples aplicação de resultados já estudados.

Sistemas com Matrizes Quadradas

Considere-se um sistema de n equações lineares em n incógnitas cuja matriz ampliada é $(A|\mathbf{b})$. Então, as seguintes condições são equivalentes:

- O sistema é possível e determinado.
- car$(A) = n$.
- er$(A) = I_n$.
- O sistema homogéneo correspondente (isto é, o sistema homogéneo com A com matriz simples) tem apenas a solução nula.

2.8 Representação matricial de sistemas

Consideremos novamente um sistema de m equações lineares em n incógintas

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}.$$

Tendo em conta a definição de produto de matrizes (em particular, relembarndo como se calcula o produto de uma matriz por um vetor) e a definição de igualdade de matrizes, vemos que o sistema anterior pode expressar-se como a seguinte *equação matricial*

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\mathbf{b}},$$

a qual pode ser escrita, numa forma compacta, como

$$A\mathbf{x} = \mathbf{b},$$

onde A é a matriz dos coeficientes do sistema (matriz simples do sistema), $\mathbf{x}$ é o vetor coluna formado pelas incógnitas e $\mathbf{b}$ é o vetor coluna dos termos independentes.

Uma solução do sistema é, como dissemos, um n-uplo $(s_1, s_2, \ldots, s_n)$ que satisfaz simultaneamente todas as equações do sistema, isto é, que verifica

$$\begin{cases} a_{11}s_1 + a_{12}x_2 + \cdots + a_{1n}s_n = b_1 \\ a_{21}s_1 + a_{22}x_2 + \cdots + a_{2n}s_n = b_2 \\ \vdots \\ a_{m1}s_1 + a_{m2}s_2 + \cdots + a_{mn}s_n = b_m \end{cases},$$

ou seja, tal que $A\mathbf{s} = \mathbf{b}$, onde $\mathbf{s}$ designa o vetor coluna $\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}$.

Neste contexto, faz, portanto, mais sentido considerar a solução do sistema, não como o n-uplo $(s_1, \ldots, s_n)$, mas sim como o vector coluna $\mathbf{s}$ tal que $A\mathbf{s} = \mathbf{b}$.

Neste curso, referir-nos-emos às soluções de sistemas, quer na forma de n-uplos ordenados, quer na forma de vectores coluna com n elementos, conforme seja mais conveniente.

Então, temos

$$\begin{aligned} A\mathbf{x} = \mathbf{b} \text{ tem solução } &\iff \exists \mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} : \; A\mathbf{s} = \mathbf{b} \\ &\iff \exists s_1, \ldots, s_n \in \mathbb{R} : \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} = \mathbf{b} \\ &\iff \exists s_1, \ldots, s_n \in \mathbb{R} : s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + \cdots + s_n\mathbf{a}_n = \mathbf{b}. \end{aligned}$$

Venos assim que:

Um sistema $A\mathbf{x} = \mathbf{b}$ tem solução se e só se $\mathbf{b}$ for uma combinação linear das colunas de A.

2.9 Matrizes invertíveis (de novo)

2.9.1 Caracterização em termos da caraterística

No primeiro capítulo referimos que uma matriz A, quadrada de ordem n, se diz invertível ou não-singular, se existir uma matriz C, quadrada de ordem n, tal que $AC = CA = I$ e que uma matriz que não seja invertível também é designada por singular. Também mostrámos que a matriz C atrás referida, se existir, é única; tal matriz é chamada a inversa de A e designada por A^{-1}. Demos também exemplos de matrizes invertíveis e de matrizes não-invertíveis. Vamos ver agora que, dada uma certa matriz, é possível saber se essa matriz é ou não invertível, desde que conheçamos a sua caraterística. Mais precisamente, tem-se o seguinte resultado importante:

Teorema 2.2. *Seja A uma matriz quadrada de ordem n, Então, tem-se:*

$$A \text{ é invertível} \iff \text{car}(A) = n.$$

Dem: $\Longrightarrow$)

Seja A invertível. Então, existe uma matriz C tal que $AC = I$. Vejamos que esta matriz C é única.[6] Sejam C_1 e C_2 matrizes tais que $AC_1 = I$ e $AC_2 = I$. Então, segue-se que $AC_1 = AC_2$. Como, por hipótese, existe A^{-1}, multiplicando ambos os membros da igualdade anterior por A^{-1}, vem $A^{-1}AC_1 = A^{-1}AC_2$ ou seja, vem $IC_1 = IC_2$ ou ainda, $C_1 = C_2$.

Designando as colunas de C por $\mathbf{c}_1, \ldots, \mathbf{c}_n$ e por $\mathbf{e}_1, \ldots, \mathbf{e}_n$ as colunas da matriz identidade (seguindo a convenção usual), a igualdade $AC = I$ pode escrever-se como

$$A\Big(\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n\Big) = \Big(\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n\Big),$$

ou ainda, como

$$\Big(A\mathbf{c}_1 \ A\mathbf{c}_2 \ \cdots \ A\mathbf{c}_n\Big) = \Big(\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n\Big).$$

A igualdade anterior e o facto de existir apenas uma matriz C tal que $AC = I$ mostram que cada um dos sistemas $A\mathbf{x}_j = \mathbf{e}_j; j = 1, \ldots, n$ admite uma e uma só solução. Como vimos, isto implica que terá de ser car$(A) = n$.

$\Longleftarrow$) Suponhamos agora que car$(A) = n$ e provemos que A tem inversa. Novamente, usando os resultados que conhecemos para sistemas de equações lineares, o facto de termos car$(A) = n$ garante-nos que haverá solução única para cada um dos sistemas $A\mathbf{x}_j = \mathbf{e}_j; j = 1, \ldots, n$, ou seja, será possível encontrar uma (e uma só) matriz $C = \Big(\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n\Big)$ tal que $AC = I$. Essa matriz terá como colunas cada uma das soluções dos sistemas $A\mathbf{x}_j = \mathbf{e}_j$.

Vamos mostrar agora que a matriz C assim obtida também satisfaz $CA = I$, ou seja, é a inversa de A. [7]

---

[6] Sabemos que existe uma única matriz que satisfaz simultaneamente $AC = I$ e $CA = I$, mas é necessário mostrar que também só existe uma matriz C que satisfaz apenas a condição $AC = I$.

[7] Não queremos aqui invocar o resultado, já que ainda não foi demonstrado, usado no Capítulo 1, que diz que para mostrar que uma matriz C é a inversa de uma matriz (quadrada) A, não teremos que mostrar que $AC = I$ e $CA = I$, bastando apenas verificar uma das igualdades.

Suponhamos que não, isto é, suponhamos que $CA \neq I$ e designemos por X a matriz $CA - I$, isto é, seja $X = CA - I$. Então, $X \neq \mathbf{0}$ e além disso, tem-se

$$AX = A(CA - I) = (AC)A - A = IA - A = A - A = \mathbf{0}.$$

Fracionando X em colunas e fazendo o mesmo para a matriz nula, vemos que a equação anterior nos diz que

$$A\Big(\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{x}_n\Big) = \Big(A\mathbf{x}_1\ A\mathbf{x}_2\ \cdots\ A\mathbf{x}_n\Big) = \Big(\mathbf{0}\ \mathbf{0}\ \cdots\ \mathbf{0}\Big).$$

Mas, se X é uma matriz não nula, ela terá (pelo menos) uma coluna não nula. Isto significa que teremos, para um certo j, $A\mathbf{x}_j = \mathbf{0}$, com $\mathbf{x}_j \neq \mathbf{0}$, ou seja, que o sistema homogéneo $A\mathbf{x} = \mathbf{0}$ terá uma solução para além da solução trivial. Como sabemos, isto não é possível quando car$(A) = n$.

Vemos, portanto, que não poderá ser $CA - I \neq \mathbf{0}$, ou seja que teremos de ter $CA = I$, tal como pretendíamos mostrar. □

Estamos agora em condições, finalmente, de demonstrar o resultado referido (e utilizado) no Capítulo 1, que diz que para mostrar que uma matriz C é a inversa de uma matriz quadrada A, não teremos que mostrar que $AC = I$ e $CA = I$, bastando apenas verificar uma das igualdades.

Proposição 2.1. *Se A é uma matriz quadrada de ordem n e C é uma matriz quadrada da mesma ordem e tal que $AC = I$, então também $CA = I$, ou seja, A é invertível e $A^{-1} = C$.*

Dem: Comecemos por observar que $AC = I$ implica que C é invertível. De facto, se C não fosse invertível, seria car$(C) < n$, o que implicaria que o sistema homogéneo $A\mathbf{x} = \mathbf{0}$ teria solução não nula, i.e., existiria um vetor $\mathbf{u} \neq \mathbf{0}$ e tal que $C\mathbf{u} = \mathbf{0}$. Mas, se assim fosse, viria

$$\mathbf{u} = I\mathbf{u} = AC\mathbf{u} = A(C\mathbf{u}) = A\mathbf{0} = \mathbf{0},$$

o que contradiz o facto de ser $\mathbf{u} \neq 0$. Sabendo que C é invertível e sendo C^{-1} a sua inversa, podemos escrever

$$AC = I \Longrightarrow ACC^{-1} = C^{-1} \Longrightarrow AI = C^{-1} \Longrightarrow A = C^{-1} \Longrightarrow CA = I.$$

□

2.9.2 Cálculo de inversas – Método de Gauss-Jordan

Acabámos de ver que, se uma dada matriz A é invertível, cada uma das colunas da matriz inversa de A é a solução do sistema de equações cuja matriz simples é A e cujo lado direito é a correspondente coluna da matriz identidade. Por outras palavras, a coluna j de A^{-1} é a (única) solução do sistema $A\mathbf{x}_j = \mathbf{e}_j$. Assim sendo, se pretendermos calcular a inversa de A, poderemos usar o método de Gauss-Jordan para resolver os n sistemas $A\mathbf{x}_j = \mathbf{e}_j$; $j = 1, \ldots, n$, em simultâneo. Note-se que, sendo A invertível, se tem car$(A) = n$, pelo que será er$(A) = I_n$. Temos, então, o procedimento descrito no quadro seguinte, para calcular a inversa de uma dada matriz invertível.

Cálculo da Inversa

(Método de Gauss-Jordan)

Para determinar a inversa de uma dada matriz A invertível, podemos aplicar o método de Gauss-Jordan, convertendo a matriz ampliada

$$\Big(A \,|\, \mathbf{e}_1\, \mathbf{e}_2 \, \cdots \mathbf{e}_n \Big)$$

na forma em escada reduzida

$$\Big(I_n \,|\, \mathbf{c}_1\, \mathbf{c}_2 \, \cdots \mathbf{c}_n \Big).$$

As diversas colunas da inversa serão, então, as colunas $\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_n$.
De uma forma compacta, fácil de lembrar:

$$(\, A \,|\, I \,) \xrightarrow{\text{Gauss-Jordan}} (\, I \,|\, A^{-1} \,).$$

Naturalmente, se ao tentarmos aplicar o método anterior, virmos que a matriz do lado esquerdo, A, não pode ser convertida na matriz identidade, teremos de concluir que car$(A) < n$ e que A não tem inversa.

2.10 Exercícios

Exercício 2.1. Identifique quais das seguintes matrizes são matrizes em escada. Para as que o forem, indique quais os pivôs.

(a) $A = \begin{pmatrix} 1 & 0 & 0 & 0 & 5 \\ 0 & 0 & 3 & 0 & 4 \\ 0 & 0 & 0 & 2 & 0 \end{pmatrix}$ (b) $B = \begin{pmatrix} 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -4 \\ 0 & 0 & 2 & -2 & 3 \end{pmatrix}$ (c) $C = \begin{pmatrix} 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

(d) $D = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ (e) $E = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ (f) $F = \begin{pmatrix} 1 & 2 & 3 & -1 & 1 \\ 0 & 1 & 2 & 5 & 2 \\ 0 & 0 & 1 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$

Exercício 2.2. Converta cada uma das matrizes dadas na forma em escada, usando o método de eliminação de Gauss. Indique, então, qual a caraterística de cada uma das matrizes e quais são as suas colunas principais.

(a) $A = \begin{pmatrix} 0 & 5 & 0 & 0 \\ 0 & 3 & 0 & 2 \\ 1 & 6 & 4 & 2 \end{pmatrix}$ (b) $B = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 5 & 7 & 2 \\ 3 & 6 & 6 & 0 \end{pmatrix}$ (c) $C = \begin{pmatrix} 0 & 5 & 10 & 25 \\ 2 & 6 & 4 & 0 \\ 0 & 3 & 4 & 0 \end{pmatrix}$

(d) $D = \begin{pmatrix} 0 & 2 & 1 & 3 & 1 \\ 0 & 4 & -1 & 3 & 8 \\ 0 & 2 & 3 & 5 & 7 \\ 0 & 4 & 2 & 6 & 2 \\ 0 & 6 & 1 & 7 & -3 \end{pmatrix}$ (e) $E = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 6 & 8 \\ 2 & 6 & 0 \\ 1 & 2 & 5 \\ 3 & 8 & 6 \end{pmatrix}$

Exercício 2.3. Considere uma matriz A de ordem $m \times n$. Justifique as seguintes afirmações:

(a) car$(A) \leq m$ e car$(A) \leq n$.

(b) Se uma das linhas de A é nula, então car$(A) < m$.

(c) Se uma das linhas de A é um múltiplo de outra linha, então car$(A) < m$.

(d) Se uma das colunas de A é formada toda por zeros, então car$(A) < n$.

Exercício 2.4. Discuta, em função dos valores do parâmetro α, a característica da seguinte matriz

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & \alpha^2 - 5 \end{pmatrix}.$$

Exercício 2.5. Para cada um dos sistemas seguintes, forme a respetiva matriz ampliada e reduza-a à forma em escada; indique qual a caraterística da matriz simples e qual a caraterística da matriz ampliada do sistema e, tendo em conta essa informação, classifique o sistema (possível/impossível; sendo possível, determinado/indeterminado); se o sistema for possível, resolva-o.

(a) $\begin{cases} x_1 - x_2 - 2x_3 = -6 \\ 3x_1 + x_2 - 2x_3 = -6 \\ -2x_1 - 2x_2 + x_3 = 2 \end{cases}$ (b) $\begin{cases} x_1 - x_2 - 2x_3 = -1 \\ -2x_1 + x_2 + x_3 = 2 \\ 3x_1 + 2x_2 + 9x_3 = 4 \end{cases}$ (c) $\begin{cases} x_1 + 2x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 1 \\ 2x_1 - x_2 + 5x_3 = 3 \end{cases}$

(d) $\begin{cases} x - 2y = -1 \\ 4x + y = 14 \\ 3x - 4y = 1 \end{cases}$ (e) $\begin{cases} x - y + 3z = 4 \\ 2x - 2y + z = 3 \\ -x + y + z = 0 \end{cases}$ (f) $\begin{cases} x_1 + 2x_2 + x_3 + 2x_4 = 3 \\ 2x_1 + 4x_2 + x_3 + 3x_4 = 4 \\ 3x_1 + 6x_2 + x_3 + 4x_4 = 5 \end{cases}$

(g) $\begin{cases} x-y+z=1 \\ x-y-z=2 \\ x+y-z=3 \\ x+y+z=4 \end{cases}$ (h) $\begin{cases} x-y+z=1 \\ x-y-z=2 \\ x+y-z=3 \\ x+y+z=2 \end{cases}$

Exercício 2.6. Discuta, em função dos valores da constante $k \in \mathbb{R}$, cada um dos sistemas de equações seguintes; resolva cada um dos sitemas para **um** dos valores de k que o tornam possível:

(a) $\begin{cases} x+y=k \\ 3x-ky=2 \end{cases}$ (b) $\begin{cases} x+y+z=k \\ kx+y+2z=2 \\ x+ky+z=4 \end{cases}$ (c) $\begin{cases} x_1-2x_2+3x_3=1 \\ 2x_1+kx_2+6x_3=6 \\ -x_1+3x_2+(k-3)x_3=0 \end{cases}$

Exercício 2.7. Diga quais das seguintes matrizes são matrizes em escada reduzida e, para as que não forem, reduza-as a essa forma, usando o método de Gauss-Jordan.

(a) $\begin{pmatrix} 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 & 3 \\ 1 & 2 & 4 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Exercício 2.8. Em cada uma das alíneas seguintes, suponha a matriz ampliada de um sistema foi transformada, usando operações elementares sobre linhas, na matriz com a forma em escada reduzida apresentada. Em cada caso, discuta o sistema e resolva-o (nos casos em que for possível).

(a) $\begin{pmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 5 & 0 & 2 & 8 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 7 & 0 & 0 & -8 & -3 \\ 0 & 0 & 1 & 0 & 6 & 5 \\ 0 & 0 & 0 & 1 & 3 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ (d) $\begin{pmatrix} 0 & 1 & 2 & 0 & -2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Exercício 2.9. Use o método de Gauss-Jordan para resolver, em simultâneo, três sistemas de equações S_1,S_2 e S_3 cuja matriz simples seja

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ -1 & -2 & 0 \end{pmatrix}$$

e cujos lados direitos, sejam, respectivamente

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{b}_2 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} \ \text{e} \ \mathbf{b}_3 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

Exercício 2.10. Resolva os seguintes sistemas homogéneos:

(a) $\begin{cases} -3x + y + z + w = 0 \\ x + y - 3z + w = 0 \\ x + y + z - 3w = 0 \end{cases}$ (b) $\begin{cases} 2x + y - 3z = 0 \\ 3x - y + 5z = 0 \\ 4x + y + 3z = 0 \end{cases}$

Exercício 2.11. Considere, para $k \in \mathbb{R}$, a matriz

$$A_k = \begin{pmatrix} 1 & 1 & -1 \\ 1 & k-1 & k-3 \\ -1 & 1-k & 2 \end{pmatrix}.$$

(a) Discuta, em função do valor do parâmetro k, o sistema $A_k\mathbf{x} = \mathbf{b}_k$, onde $\mathbf{b}_k = \begin{pmatrix} 1 \\ 4 \\ 2k-6 \end{pmatrix}$.

(b) Resolva o sistema homogéneo $A_k\mathbf{x} = \mathbf{0}$, para o caso $k = 2$.

Exercício 2.12. Mostre que se $\mathbf{u}$ e $\mathbf{v}$ são soluções de um sistema homogéneo $A\mathbf{x} = \mathbf{0}$, então:

(a) A soma $\mathbf{u} + \mathbf{v}$ também é solução do sistema.

(b) Para qualquer $\alpha \in \mathbb{R}$, $\alpha\mathbf{u}$ também é solução do sistema.

(c) Use o resultado da alínea anterior para concluir que: *Se um sistema homogéneo tem uma solução não nula, então tem uma infinidade de soluções*. (Tal como afirmámos, mas não demonstrámos!)

Exercício 2.13. (a) Considere um sistema $A\mathbf{x} = \mathbf{b}$ que seja possível e seja $\mathbf{s}_0$ uma sua solução. Mostre que toda a solução $\mathbf{s}$ do sistema pode ser escrita como $\mathbf{s} = \mathbf{s}_0 + \mathbf{u}$, em que $\mathbf{u}$ é uma solução do sistema homogéneo associado, $A\mathbf{x} = \mathbf{0}$.

(b) Use o resultado da alínea anterior e o resultado da alínea c) do exercício anterior para estabelecer o resultado seguinte (que temos vindo a usar, sem demonstração): *Um sistema possível ou tem apenas uma solução ou tem um número infinito de soluções.*

Exercício 2.14. Determine ângulos α, β e γ, com $0 \leq \alpha < 2\pi$, $0 \leq \beta < 2\pi$ e $0 \leq \gamma < \pi$ tais que

$$\begin{cases} 2\,\text{sen}\,\alpha - \cos\beta + 3\tan\gamma = 3 \\ 4\,\text{sen}\,\alpha + 2\cos\beta - 2\tan\gamma = 2 \\ 6\,\text{sen}\,\alpha - 3\cos\beta + \tan\gamma = 9 \end{cases}$$

Sugetão: Efetue uma mudança de variáveis, $x = \text{sen}\,\alpha$, $y = \cos\beta$ e $z = \tan\gamma$, para converter o sistema dado num sistema linear.

Exercício 2.15. Seja

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 1 & 3 & 3 & 2 \\ 2 & 4 & 3 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

(a) Mostre que a matriz

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -2 & 2 & -3 \\ 0 & 1 & -1 & 1 \\ -2 & 3 & -2 & 3 \end{pmatrix}.$$

é a inversa de A.

(b) Baseando-se apenas na informação fornecida pela alínea anterior, diga:

(i) qual é a caraterística de A;

(ii) se $\begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix}$ é uma possível solução do sistema homogéneo $A\mathbf{x} = \mathbf{0}$;

(iii) qual é a inversa da matriz A^{T};

(iv) qual é a inversa de $2A$.

Exercício 2.16. Nas alíneas seguintes, determine se as matrizes consideradas têm inversa e, em caso afirmativo, calcule-a, usando o método de Gauss-Jordan.

(a) $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 3 & 5 \end{pmatrix}$ (b) $B = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 0 & 1 \\ 3 & 1 & 2 \end{pmatrix}$ (c) $C = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & 0 \\ 2 & 2 & 3 \end{pmatrix}$

Exercício 2.17. Considere um sistema $A\mathbf{x} = \mathbf{b}$, onde A é uma matriz quadrada de ordem n. Mostre que esse sistema tem uma e uma só solução se e só se a matriz A for invertível, sendo a solução dada por $A^{-1}\mathbf{b}$.

3

Determinantes

3.1 Conceitos básicos

Neste capítulo, vamos ver que é possível associar um número a uma matriz quadrada – o chamado *determinante* da matriz – o qual (para além de outras aplicações importantes) nos vai fornecer um critério para saber se essa matriz é ou não invertível.

Antes de introduzirmos o conceito de determinante, precisamos de algumas definições.

Definição 3.1 (Permutação)**.** Sendo $n \in \mathbb{N}$, uma permutação do conjunto $\{1, 2, \ldots, n\}$ é uma aplicação bijectiva deste conjunto nele próprio. O conjunto de todas as permutações de $\{1, 2, \ldots, n\}$ é designado por S_n.

Uma permutação $\sigma \in S_n$ é frequentemente representada na forma

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma_1 & \sigma_2 & \cdots & \sigma_n \end{pmatrix},$$

ou, mais simplesmente, como

$$\sigma = (\sigma_1, \sigma_2, \cdots, \sigma_n)$$

onde $\sigma_1 = \sigma(1), \sigma_2 = \sigma(2), \ldots, \sigma_n = \sigma(n)$.

Por exemplo,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}$$

ou

$$(2, 4, 5, 1, 3)$$

representam a mesma permutação.

Como a permutação se identifica com o n-uplo ordenado $(\sigma_1, \ldots, \sigma_n)$, costumamos referir-nos a $\sigma_1, \ldots, \sigma_n$ como *elementos* da permutação.

Definição 3.2. Dizemos que dois elementos σ_i e σ_j de uma permutação $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ estão em *inversão*, se estão fora da ordem natural, isto é, se $i < j$ e $\sigma_i > \sigma_j$. O *número de inversões* de uma permutação, designado por $\nu(\sigma)$, é o número total de pares de números que estão em inversão nessa permutação.

Por exemplo, relativamente à permutação anterior, 2 e 1, 4 e 1, 4 e 3 , 5 e 1 e 5 e 3 estão em inversão e portanto, o número de inversões dessa permutação é 5.

Definição 3.3. Diz-se que uma permutação é *par*, se o seu número de inversões for par e diremos que ela é *ímpar*, se esse número for ímpar.

A permutação considerada acima é, portanto, uma permutação ímpar. O número

$$\text{sgn}(\sigma) := (-1)^{\nu(\sigma)} = \begin{cases} 1, & \text{se } \sigma \text{ é par} \\ -1, & \text{se } \sigma \text{ é ímpar,} \end{cases}$$

é chamado *sinal* da permutação σ.

É fácil de mostrar (tal pode ser feito por indução sobre n) que o número total de permutações do conjunto $\{1, 2, \ldots, n\}$, isto é, o número de elementos de S_n, é igual a $n!$.

Estamos agora em condições de introduzir a definição de determinante de uma matriz quadrada.

Determinante de uma matriz

Seja $A = (a_{ij})$ uma matriz quadrada de ordem n. O *determinante* de A, designado por $\det A$ ou $|A|$, é o número dado por

$$\det A = \sum_{\sigma \in S_n} \text{sgn}(\sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{3.1}$$

Note-se que cada um dos produtos $a_{1\sigma(1)}\, a_{2\sigma(2)} \cdots a_{n\sigma(n)}$ na fórmula acima é obtido com n elementos da matriz A, escolhidos de modo a que não haja dois elementos pertencentes, quer à mesma linha, quer à mesma coluna. Um produto desse tipo é chamado um *produto elementar* da matriz A. Vemos assim, que o determinante de A é uma soma algébrica de produtos elementares de A.

Vamos analisar em pormenor a fórmula (3.1), para os casos em que $n = 1$, $n = 2$ e $n = 3$.

Caso $n = 1$

Se $A = (a_{11})$ é uma matriz 1×1, então, usando a fórmula (3.1), vemos que

$$\det(a_{11}) = a_{11}, \tag{3.2}$$

uma vez que existe uma única permutação $\sigma \in S_1$, que é a permutação $\sigma = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, a qual é, naturalmente, par.

Nota: Neste caso, não é conveniente usarmos a notação $|a_{11}|$ para o determinante, para não se confundir com o módulo.

Caso $n = 2$

Existem duas permutações possíveis do conjunto $\{1, 2\}$,

$$\sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad \text{e} \quad \sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

A primeira é uma permutação par e a segunda é ímpar. Então, temos

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= \operatorname{sgn}(\sigma_1) a_{1\sigma_1(1)} a_{2\sigma_1(2)} + \operatorname{sgn}(\sigma_2) a_{1\sigma_2(1)} a_{2\sigma_2(2)} \\ &= a_{11}a_{22} - a_{12}a_{21}. \end{aligned} \tag{3.3}$$

Exemplo 3.1.

$$\det \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} = 2 \times 5 - 3 \times 4 = 10 - 12 = -2.$$

Caso $n = 3$

Para $n = 3$, temos $3! = 6$ permutações do conjunto $\{1, 2, 3\}$, a saber:

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

$$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \qquad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \qquad \sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

As permutações σ_1, σ_2 e σ_3 são pares e as permutações σ_4, σ_5 e σ_6 são ímpares. Então, tem-se

$$\begin{aligned}
\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\
&\quad - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \\
&= \underbrace{\left(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}\right)}_{\mathcal{A}} - \\
&\quad - \underbrace{\left(a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32}\right)}_{\mathcal{S}}
\end{aligned} \tag{3.4}$$

Cada uma das três parcelas de $\mathcal{A}$ é obtida com os produtos dos elementos assinalados abaixo:

$$\begin{pmatrix} a_{11} & * & * \\ * & a_{22} & * \\ * & * & a_{33} \end{pmatrix}, \qquad \begin{pmatrix} * & a_{12} & * \\ * & * & a_{23} \\ a_{31} & * & * \end{pmatrix}, \qquad \begin{pmatrix} * & * & a_{13} \\ a_{21} & * & * \\ * & a_{32} & * \end{pmatrix}$$

De modo análogo, tem-se para $\mathcal{S}$:

$$\begin{pmatrix} * & * & a_{13} \\ * & a_{22} & * \\ a_{31} & * & * \end{pmatrix}, \qquad \begin{pmatrix} * & a_{12} & * \\ a_{21} & * & * \\ * & * & a_{33} \end{pmatrix}, \qquad \begin{pmatrix} a_{11} & * & * \\ * & * & a_{23} \\ * & a_{32} & * \end{pmatrix}$$

Com base nas figuras acima, podemos enunciar o modo de calcular um determinante de uma matriz de ordem 3, na forma de uma regra, espécie de mnemónica, conhecida por *regra de Sarrus*.

Regra de Sarrus

(Cálculo do determinante de uma matriz de ordem 3)

O determinante de uma matriz de ordem 3 é uma diferença de duas parcelas $\mathcal{A}$ e $\mathcal{S}$, em que cada uma delas - aditivo $\mathcal{A}$ e subtrativo $\mathcal{S}$ - é soma de três parcelas obtidas como produtos de três elementos de A; os elementos usados no aditivo são os da diagonal de A e os que estão nos vértices dos triângulos cuja base é paralela a essa diagonal; os elementos usados no subtrativo são os da diagonal secundária e os dos vértices dos triângulos de base paralela a essa diagonal.

Exemplo 3.2. Vejamos um exemplo da aplicação da regra de Sarrus. Consideremos a matriz quadrada de ordem 3:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 3 \\ 5 & 2 & 1 \end{pmatrix}.$$

Temos

$$\mathcal{A} = 1 \times (-1) \times 1 + 2 \times 3 \times 5 + 0 \times 2 \times 1 = -1 + 30 + 0 = 29,$$

$$\mathcal{S} = 1 \times (-1) \times 5 + 2 \times 3 \times 1 + 0 \times 2 \times 1 = -5 + 6 + 0 = 1.$$

Então,

$$\det A = \mathcal{A} - \mathcal{S} = 29 - 1 = 28.$$

Quando n cresce, $n!$ cresce muito rapidamente. Sendo este o número de elementos de S_n, a fórmula (3.1) que define o determinante de uma matriz A de ordem n torna-se de difícil utilização, mesmo quando n não é muito grande. No entanto, o determinante goza de algumas propriedades que nos vão ajudar a calculá-lo de forma mais simples.

Nota: Com um certo abuso de linguagem, falaremos muitas vezes, em linhas, colunas e ordem de um determinante, querendo referir-nos, naturalmente, às linhas, colunas e ordem da respetiva matriz.

3.2 Propiedades dos determinantes

Começamos por enunciar o seguinte resultado, cuja demonstração omitimos; ver, e.g.[**?**, p.463].

Teorema 3.1. *Se A é uma matriz quadrada, então*

$$\det A = \det A^{\mathsf{T}}.$$

O teorema anterior é muito importante, porque nos garante que todas as propriedades que se demonstrem para linhas são também válidas para colunas, não necessitando de uma demonstração separada.

Por vezes, referimo-nos a *filas* de um determinante para designar as linhas ou colunas da respetiva matriz.

3.2.1 Determinantes de matrizes especiais

As duas proposições seguintes referem casos em que a "forma" especial da matriz nos permite, de imediato, calcular o seu determinante.

Proposição 3.1. *Se A tem uma linha (ou coluna) nula, então* $\det A = 0$.

Dem: Seja $A = (a_{ij})$ e suponhamos que uma sua determinada linha k é nula, isto é, é toda formada por zeros. Então, tem-se:

$$\begin{aligned}\det A &= \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{k\sigma(k)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots \ 0 \ \cdots a_{n\sigma(n)} \\ &= 0.\end{aligned}$$

□

Proposição 3.2. *O determinante de uma matriz triangular é igual ao produto dos seus elementos diagonais.*

Dem: Consideremos o caso em que a matriz é triangular inferior, i.e. seja $A = (a_{ij})$ tal que $a_{ij} = 0$ se $j > i$ (o que equivale a afirmar que os elementos de A situados acima da diagonal principal são todos nulos).

Analisemos os diversos termos $a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)}$ envolvidos na expansão do determinante, dada por (3.1), quando σ percorre o conjunto S_n.

Comecemos por notar que qualquer termo que corresponda a uma permutação σ para a qual seja $\sigma(1) \neq 1$ será nulo, uma vez que envolve um elemento acima da diagonal; sendo $\sigma(1) = 1$ (o que significa que $\sigma(2)$ só poderá, então, tomar valores de 2 a n, porque σ é uma permutação do conjunto $\{1, 2, \ldots n\}$), se $\sigma(2) \neq 2$, também o termo será nulo; e assim sucessivamente, de modo que podemos concluir que o único termo, na expansão do determinante, que não é necessariamente nulo é o termo $a_{11} \ldots a_{nn}$; como este termo está associado a uma permutação par, concluímos que $\det A = a_{11} \ldots a_{nn}$, como pretendíamos mostrar.

Para estabelecer o resultado para a o caso em que a matriz é triangular superior basta invocar o Teorema 3.1. □

Exemplo 3.3. Calculemos o determinante da matriz

$$A = \begin{pmatrix} 3 & 1 & 5 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

usando a regra de Sarrus e confirmemos que $\det(A) = 3 \times 2 \times 2 = 12$.

Tem-se, com a notação usual,

$$\mathcal{A} = 3 \times 2 \times 2 + 1 \times 2 \times \mathbf{0} + \mathbf{0} \times \mathbf{0} \times 5 = 12$$

e

$$\mathcal{S} = \mathbf{0} \times 2 \times 5 + \mathbf{0} \times 2 \times 3 + \mathbf{0} \times 1 \times 2 = 0,$$

pelo que

$$\det A = \mathcal{A} - \mathcal{S} = 12 - 0 = 12.$$

Vemos que, à exceção de $a_{11}a_{22}a_{33}$, cada termo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ contém sempre (pelo menos) um elemento situado abaixo da diagonal (esses elementos estão salientados a negrito), sendo portanto nulo.

Como consequência imediata da proposição anterior, tem-se o resultado seguinte.

Corolário 3.1. *O determinante de uma matriz diagonal é igual ao produto dos seus elementos diagonais.*

Dem: Basta lembrar que uma matriz diagonal é uma matriz triangular. □

3.2.2 Determinantes e operações elementares

Quando trabalhamos com matrizes não necessariamente ligadas a sistemas de equações, faz sentido definir operações elementares sobre as suas colunas, de um modo totalmente análogo ao que fizemos para linhas; ou seja, podemos falar em troca de colunas, multiplicação de uma coluna por um escalar não nulo e substituição de uma coluna pela sua soma com outra coluna multiplicada por um escalar.

Vamos ver agora como as operações elementares sobre as linhas ou colunas de uma matriz afetam o respetivo determinante. Uma vez mais, lembramos que nos bastará estabelecer as propriedades relativas a linhas.

Nota: Nas proposições seguintes, A designa sempre uma matriz quadrada de ordem n e $L_1, L_2, \ldots, L_n$ as suas linhas.

Proposição 3.3. *Se B se obtém de A por troca de duas das suas linhas (ou colunas), então*

$$\det B = -\det A.$$

Dizemos, informalmente, que a troca de linhas (ou colunas) troca o sinal do determinante.

Dem: ver, e.g. [**?**, p.463]. □

Corolário 3.2. *Se A tem duas linhas (ou duas colunas) iguais, então* $\det A = 0$.

Dem: Suponhamos que $L_k = L_j$ para k e j fixos $(1 \leq k < j \leq n)$. Então, tem-se

$$\det A = \det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_j \\ \vdots \\ L_n \end{pmatrix} \underset{(L_k = L_j)}{=} \det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_k \\ \vdots \\ L_n \end{pmatrix} \underset{(\text{Troca linhas } k \text{ e } j)}{=} -\det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_k \\ \vdots \\ L_n \end{pmatrix}$$

$$\underset{(L_k = L_j)}{=} -\det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_j \\ \vdots \\ L_n \end{pmatrix} = -\det A,$$

o que implica que $\det A = 0$, como queríamos provar. □

Proposição 3.4. *Se multiplicarmos uma linha (ou coluna) de A por um número α, o determinante vem multiplicado por α.*

Dem: Seja $A = (a_{ij})$ e suponhamos que $B = (b_{ij})$ se obtém de A multiplicando uma determinada linha k, arbitrária, mas fixada, por α. Isto significa que, para $i, j = 1, 2, \ldots, n$, se tem $b_{ij} = a_{ij}$, se $i \neq k$, e $b_{kj} = \alpha a_{kj}$. Então, usando a definição de determinante, vem

$$\begin{aligned} \det B &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{1\sigma(1)} \ldots b_{k\sigma(k)} \ldots b_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \ldots \alpha a_{k\sigma(k)} \ldots a_{n\sigma(n)} \\ &= \alpha \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \ldots a_{k\sigma(k)} \ldots a_{n\sigma(n)} \\ &= \alpha \det A, \end{aligned}$$

como pretendíamos mostrar. □

Nota: O resultado referido na proposição anterior é válido mesmo que o escalar α seja igual a zero; no entanto, apenas quando $\alpha \neq 0$ estaremos a fazer uma operação elementar (do tipo O2).

Exemplo 3.4. Sejam $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ e $B = \begin{pmatrix} 5 & 10 \\ 4 & 3 \end{pmatrix}$. Note-se que B se obtém de A multiplicando a sua primeira linha pelo número 5. Calculemos $\det A$ e $\det B$ através da definição e vejamos que, de facto $\det B = 5 \det A$. Tem-se:

$$\det A = 1 \times 3 - 4 \times 2 = 3 - 8 = -5$$

e

$$\det B = 5 \times 3 - 4 \times 10 = 5(1 \times 3 - 4 \times 2) = 5 \times (-5) = 5 \det A.$$

A proposição anterior tem o seguinte corolário imediato.

Corolário 3.3. *Se A é uma matriz quadrada de ordem n e α é um escalar, então*

$$\det(\alpha A) = \alpha^n \det A.$$

Antes de vermos o que sucede ao determinante de uma matriz quando efetuamos na matriz uma operação elementar de tipo O3 – adição a uma linha (coluna) de outra linha (coluna) multiplicada por um escalar – vamos dar dois resultados auxiliares.

Proposição 3.5. *Suponhamos que um dada linha k de A ($1 \leq k \leq n$) se decompõe na forma*

$$L_k = \begin{pmatrix} a_{k1} & \dots & a_{kn} \end{pmatrix} = \begin{pmatrix} a'_{k1} & \dots & a'_{kn} \end{pmatrix} + \begin{pmatrix} a''_{k1} & \dots & a''_{kn} \end{pmatrix} = L'_k + L''_k.$$

Então,

$$\det A = \det \begin{pmatrix} L_1 \\ \vdots \\ L'_k + L''_k \\ \vdots \\ L_n \end{pmatrix} = \det \begin{pmatrix} L_1 \\ \vdots \\ L'_k \\ \vdots \\ L_n \end{pmatrix} + \det \begin{pmatrix} L_1 \\ \vdots \\ L''_k \\ \vdots \\ L_n \end{pmatrix}.$$

Dem: A demonstração é análoga à demonstração da Proposição 3.4 e é deixada ao cuidado dos alunos. □

Nota: Naturalmente, existe um resultado idêntico ao anterior relativo a colunas.

Exemplo 3.5.

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 1 & -2 & 4 \\ 2 & 3 & 4 \end{pmatrix} + \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 2 & 3 & 4 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 & 3 \\ 5 & -1 & 6 \\ 2 & 3 & 4 \end{pmatrix}.$$

Verifique, calculando cada um dos determinantes envolvidos.

Proposição 3.6. *Se A tem duas linhas (ou duas colunas) proporcionais, então* $\det A = 0$.

Dem: Este resultado segue-se facilmente da Proposição 3.4 e do Corolário 3.2. Com efeito, suponhamos que $L_j = \alpha L_k$ para k e j fixos e onde assumimos, sem perda de generalidade, que $k < j$. Então

$$\det A = \det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_j \\ \vdots \\ L_n \end{pmatrix} = \det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ \alpha L_k \\ \vdots \\ L_n \end{pmatrix} = \alpha \det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_k \\ \vdots \\ L_n \end{pmatrix} = \alpha 0 = 0.$$

□

Proposição 3.7. *O determinante de uma matriz não se altera quando se adiciona a uma linha (coluna) dessa matriz outra linha (coluna) multiplicada por um escalar.*

Dem: Suponhamos que a linha j de A é substituída pela sua soma com a linha k (com $k < j$, sem perda de generalidade), multiplicada por α . Então, tem-se

$$\det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_j + \alpha L_k \\ \vdots \\ L_n \end{pmatrix} = \det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_j \\ \vdots \\ L_n \end{pmatrix} + \det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ \alpha L_k \\ \vdots \\ L_n \end{pmatrix} = \det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_j \\ \vdots \\ L_n \end{pmatrix} + 0 = \det \begin{pmatrix} L_1 \\ \vdots \\ L_k \\ \vdots \\ L_j \\ \vdots \\ L_n \end{pmatrix}.$$

Justifique todas as passagens. □

3.2.3 Cálculo de determinantes usando operações elementares

Como sabemos, é sempre possível converter uma dada matriz A na forma em escada, usando operações elementares. Além disso, se A for quadrada, ao reduzirmos A à forma em escada, obtemos uma matriz triangular superior. Por outro lado, o determinante de uma matriz triangular é muito simples de calcular: basta multiplicar os elementos da sua diagonal.

Os factos anteriores indicam-nos que poderemos adotar o seguinte procedimento para calcular o determinante de uma dada matriz A:

1. Convertemos A na forma em escada, usando operações elementares.
2. Calculamos o determinante da matriz triangular resultante (multiplicando os elementos da sua diagonal).
3. Relacionamos o valor do determinante de A com o do determinante da matriz triangular obtida, tendo em conta as operações do tipo O1 ou tipo O2 que usámos na conversão da matriz à forma em escada.

Vejamos um exemplo de aplicação deste tipo de procedimento.

Exemplo 3.6.

$$\begin{vmatrix} 1 & -2 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & -4 & 2 & 3 \\ 1 & 1 & 3 & 3 \end{vmatrix} \underset{(\mathrm{O3})}{=} \begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 0 & 3 & 4 \\ 0 & -2 & 0 & 2 \\ 0 & 3 & 1 & 2 \end{vmatrix} \underset{(\mathrm{O1})}{=} - \begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & -2 & 0 & 2 \\ 0 & 0 & 3 & 4 \\ 0 & 3 & 1 & 2 \end{vmatrix}$$

$$\underset{(\mathrm{O2})}{=} -(-2) \begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 3 & 4 \\ 0 & 3 & 1 & 2 \end{vmatrix} \underset{(\mathrm{O3})}{=} 2 \begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 1 & 5 \end{vmatrix}$$

$$\underset{(\mathrm{O1})}{=} -2 \begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 3 & 4 \end{vmatrix} \underset{(\mathrm{O3})}{=} -2 \begin{vmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & -11 \end{vmatrix}$$

$$\underset{(\text{Det. matriz } \Delta)}{=} -2\,(1 \times 1 \times 1 \times (-11)) = 22.$$

3.3 Determinante e invertibilidade

As proposições 3.3 e 3.4 mostram-nos que duas das operações elementares – troca de linhas ou colunas e multiplicação de uma linha ou de uma coluna por um escalar não nulo (operações de tipo O1 e tipo O2, respetivamente) – afetam o determinante da respetiva matriz, tendo como efeito multiplicá-lo por um número **diferente de zero**: o número pelo qual multiplicamos a linha ou coluna, no caso de uma operação do tipo O2, ou o número -1, para uma operação de tipo O1. Por outro lado, a Proposição 3.7 garante-nos que o terceiro tipo de operação elementar – adição a uma linha (coluna) de outra linha (coluna) multiplicada por um número – não altera o valor do determinante. Temos, pois, o resultado contido no quadro seguinte.

Se B se obtém de A por uma sequência finita de operações elementares sobre as suas linhas (ou colunas), então

$$\det B = K \det A,$$

para um certo $K \neq 0$. Em particular, tem-se

$$\det B = 0 \iff \det A = 0.$$

Como consequência do resultado enunciado acima, têm-se os seguintes resultados importantes.

Determinante e Caraterística

Teorema 3.2. *Seja A uma matriz quadrada de ordem n. Então, tem-se*

$$\det A = 0 \iff \mathrm{car}(A) < n, \tag{3.5}$$

ou, de forma equivalente,

$$\det A \neq 0 \iff \mathrm{car}(A) = n. \tag{3.6}$$

Dem: Para demonstrar (3.5), recordemos que A pode ser sempre convertida na forma em escada reduzida, $\mathrm{er}(A)$, por uma sequência de operações elementares. Então, tem-se

$$\det A = 0 \iff \det \mathrm{er}(A) = 0.$$

Mas, como A é quadrada, a matriz em escada reduzida $\mathrm{er}(A)$ é uma matriz triangular, pelo que o seu determinante será igual a zero se e só se houver (pelo menos) um elemento nulo na diagonal, ou seja, se e só se o número de pivôs de $\mathrm{er}(A)$ for inferior a n. Como esse número de pivôs é a caraterística de A, o resultado está estabelecido. $\square$

Temos também o seguinte resultado importante.

Determinante e Invertibilidade

Teorema 3.3. *Seja A uma matriz quadrada de ordem n. Então, tem-se*

$$A \text{ é inverível} \iff \det A \neq 0.$$

Dem: Basta lembrar que A é invertível se e só se $\mathrm{car}(A) = n$ (resultado estabelecido no capítulo anterior) e invocar o resultado (3.6). $\square$

3.4 Teorema de Laplace

Na secção anterior vimos um processo de calcular determinantes, usando operações elementares, que é útil quando a matriz em causa tem uma ordem "grande". O chamado *Teorema de Laplace* fornece-nos um processo alternativo de cálculo de tais determinantes, mostrando-nos como um determinante de ordem n se pode calcular à custa de (no máximo) n determinantes de ordem $n-1$. O teorema pode ser aplicado sucessivamente até que os determinantes obtidos tenham uma ordem "razoável"(2 ou 3, por exemplo).

Antes de enunciar esse importante teorema, introduzamos algumas definições.

Definição 3.4. Seja A uma matriz quadrada de ordem n. O *menor do elemento* a_{ij} ($1 \leq i, j \leq n$) é o determinante da submatriz obtida de A suprimindo-lhe a linha i e a coluna j, isto é, a linha e a coluna a que pertence a_{ij}. O menor de a_{ij} será denotado por M_{ij}.

Exemplo 3.7. Seja $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 5 \\ 4 & 1 & 4 \end{pmatrix}$. Então, tem-se, por exemplo

$$M_{11} = \det \begin{pmatrix} 0 & 5 \\ 1 & 4 \end{pmatrix} = -5, \quad M_{23} = \det \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix} = -7.$$

Nota:

- Naturalmente, os menores só estão definidos para matrizes de ordem $n \geq 2$. Até ao final deste capítulo, e para evitar repetir essa informação, assumimos que as matrizes consideradas têm ordem $n \geq 2$.
- O menor do elemento a_{ij} não depende de a_{ij}, pois a_{ij} não faz parte da submatriz que usamos para calcular M_{ij}, mas apenas da sua posição na matriz, pelo que seria mais correto falarmos em "menor da posição (i, j)"ou "menor (i, j)". No entanto, o uso da designação "menor de um elemento"é habitual e facilita o enunciado de alguns resultados, pelo que será aqui adotado.

Definição 3.5. Chama-se *complemento algébrico* ou *cofator* do elemento a_{ij} ao produto do seu menor por $(-1)^{i+j}$. O complemento algébrico de a_{ij} será denotado por A_{ij}. Tem-se, assim

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

Exemplo 3.8. Considerando novamente a matriz do exemplo anterior, tem-se

$$A_{11} = (-1)^{1+1} M_{11} = M_{11} = -5, \quad A_{23} = (-1)^{2+3} M_{23} = -M_{23} = 7.$$

Teorema de Laplace

Teorema 3.4. *O determinante de uma matriz é igual à soma dos produtos que se obtêm multiplicando os elementos de uma sua linha (ou coluna) pelos respetivos complementos algébricos, isto é, tem-se, para* $i, j = 1, \ldots, n$*:*

$$\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \tag{3.7}$$

e

$$\det A = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}. \tag{3.8}$$

Dem: Ver, e.g., [**?**, pp.132]. $\square$

Exemplo 3.9. Seja

$$A = \begin{pmatrix} 0 & 0 & 0 & 2 \\ 5 & 1 & 2 & 4 \\ 3 & 7 & 1 & 0 \\ 0 & 2 & -1 & 4 \end{pmatrix}.$$

Vamos usar o Teorema de Laplace, fazendo a expansão do determinante ao longo da primeira linha. Temos

$$\det A = 2(-1)^{1+4} \begin{vmatrix} 5 & 1 & 2 \\ 3 & 7 & 1 \\ 0 & 2 & -1 \end{vmatrix} = (-2) \begin{vmatrix} 5 & 1 & 2 \\ 3 & 7 & 1 \\ 0 & 2 & -1 \end{vmatrix}.$$

Aplicando novamente o Teorema de Laplace, ao longo da terceira linha do determinante de ordem 3 obtido, e calculando os determinantes de ordem 2 que daí resultam, usando a fórmula (3.3), vem

$$\begin{vmatrix} 5 & 1 & 2 \\ 3 & 7 & 1 \\ 0 & 2 & -1 \end{vmatrix} = 2(-1)^{3+2} \begin{vmatrix} 5 & 2 \\ 3 & 1 \end{vmatrix} + (-1)(-1)^{3+3} \begin{vmatrix} 5 & 1 \\ 3 & 7 \end{vmatrix} = (-2)(5-6) - (35-3) = 2 - 32 = -30$$

Então, tem-se, $\det A = (-2)(-30) = 60$.

Nota:

1. Quando usarmos o Teorema de Laplace, devemos fazê-lo expandindo ao longo de uma linha ou de uma coluna da matriz com o maior número de zeros possível.

2. Para calcular um determinante, podemos combinar o uso de operações elementares com o uso do Teorema de Laplace (começando, por exemplo, por converter todos os elementos de uma dada coluna, abaixo da posição 1, a zero e fazendo depois a expansão, pelo Teorema de Laplace, ao longo dessa coluna).

O seguinte teorema é uma consequência simples do Teorema de Laplace.

Teorema 3.5. *Seja A uma matriz quadrada de ordem n. Então, a soma dos produtos dos elementos de uma dada linha (coluna) pelos complementos algébricos dos elementos homólogos de outra linha (coluna) é igual a zero. Isto é, tem-se, para $i, j = 1, \ldots, n$:*

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = 0; \text{ se } i \neq j, \tag{3.9}$$

e

$$a_{1k}A_{1p} + a_{2k}A_{2p} + \cdots + a_{nk}A_{np} = 0; \text{ se } k \neq p. \tag{3.10}$$

Dem: Para estabelecer (3.9) basta notar que a expressão

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$$

é igual à expansão, usando o Teorema de Laplace ao longo da linha j, de uma matriz obtida de A substituindo a linha j por uma linha igual à linha i, ou seja, de uma matriz com duas linhas iguais. Como sabemos, esse determinante é igual a zero, o que estabelece o resultado. A demonstração de (3.10) é totalmente análoga. □

Combinando os resultados (3.7) e (3.9) podemos escrever

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = \begin{cases} \det A, & \text{se } i = j, \\ 0, & \text{se } i \neq j \end{cases} \tag{3.11}$$

ou, de uma forma mais compacta, como

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = \delta_{ij} \det A,$$

onde δ_{ij}, chamado *símbolo de Kronecker*, é definido como

$$\delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j. \end{cases}$$

De modo análogo, os resultados (3.8) e (3.10) escrevem-se como

$$a_{1k}A_{1p} + a_{2k}A_{2p} + \cdots + a_{nk}A_{np} = \delta_{kp} \det A.$$

Definição 3.6 (Matriz dos cofatores e matriz adjunta de uma matriz)**.** Dada uma certa matriz A, quadrada de ordem n, chama-se matriz dos cofatores a matriz (da mesma ordem) cujo elemento na posição (i,j) é o cofator A_{ij}. A transposta da matriz dos cofatores é chamada *matriz adjunta* de A e designada por adj A.Tem-se, assim

$$\operatorname{adj} A = (A_{ij})^{\mathsf{T}} = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}. \tag{3.12}$$

Exemplo 3.10. Considere-se novamente a matriz $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 5 \\ 4 & 1 & 4 \end{pmatrix}$. Se calcularmos todos os complementos algébricos dos elementos de A, vem:

$$\begin{array}{lll} A_{11} = -5, & A_{12} = 12, & A_{13} = 2 \\ A_{21} = -9, & A_{22} = 8, & A_{23} = 7 \\ A_{31} = 10, & A_{32} = -7, & A_{33} = -4 \end{array}$$

pelo que

$$\operatorname{adj} A = \begin{pmatrix} -5 & 12 & 2 \\ -9 & 8 & 7 \\ 10 & -7 & -4 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} -5 & -9 & 10 \\ 12 & 8 & -7 \\ 2 & 7 & -4 \end{pmatrix}.$$

O teorema seguinte fornece-nos um novo processo de calcular a inversa de uma dada matriz invertível A.

Teorema 3.6. *Dada A quadrada de ordem n, tem-se:*

1. $$A \operatorname{adj} A = (\det A) I_n. \tag{3.13}$$

2. *Se A é invertível, então*

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A. \tag{3.14}$$

Dem:

1. Usando a definição de adjunta e tendo em conta o resultado (3.11), vemos que

$$A\,\text{adj}\,A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \det A \end{pmatrix} = (\det A)I_n,$$

o que estabelece o resultado (3.13).

2. Se A é invertível, sabemos que $\det A \neq 0$. Então, multiplicando ambos os membros de (3.13) por $\frac{1}{\det A}$, vem

$$A\frac{1}{\det A}\,\text{adj}\,A = I_n,$$

o que mostra que a matriz $\frac{1}{\det A}\,\text{adj}\,A$ é a inversa de A.

$\square$

Nota: Embora não tenhamos definido a "divisão" de uma matriz por um escalar, é muito frequente escrevermos $\frac{\text{adj}\,A}{\det A}$ com o significado $\frac{1}{\det A}\,\text{adj}(A)$ e dizer, então, de um modo informal que: *"A inversa de uma matriz (invertível) se pode obter dividindo a adjunta pelo determinante"*.

Exemplo 3.11. Considerando de novo a matriz A do exemplo anterior, se calcularmos o seu determinante vem $\det A = 17$. Então, a inversa de A é a matriz

$$A^{-1} = \begin{pmatrix} -\frac{5}{17} & -\frac{9}{17} & \frac{10}{17} \\ \frac{12}{17} & \frac{8}{17} & -\frac{7}{17} \\ \frac{2}{17} & \frac{7}{17} & -\frac{4}{17}. \end{pmatrix}.$$

3.5 Regra de Cramer

Vamos agora ver como podemos usar determinates para calcular a solução de um sistema cuja matriz seja quadrada e invertível.

Definição 3.7. Um sistema $A\mathbf{x} = \mathbf{b}$ diz-se um *sistema de Cramer* se a matriz A do sistema for quadrada e $\det A \neq 0$.

Um sistema de Cramer (com uma matriz A de ordem n) tem solução única, já que, sendo $\det A \neq 0$, será $\text{car}(A) = n$, o que sabemos ser condição suficiente para garantir que o sistema é possível e determinado. Além disso, como $\det A \neq 0$, a matriz A é invertível e, portanto, tem-se:

$$A\mathbf{x} = \mathbf{b} \iff A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \iff \mathbf{x} = A^{-1}\mathbf{b}.$$

Regra de Cramer

Teorema 3.7. *Seja $A\mathbf{x} = \mathbf{b}$ um sistema de Cramer, com A de ordem n. Então, o valor de cada uma das incógnitas x_i é dado por*

$$x_i = \frac{\det B_i}{\det A}; i = 1, \ldots, n,$$

onde B_i é a matriz obtida de A substituindo a sua coluna i pela coluna $\mathbf{b}$ dos termos independentes.

Dem: Tem-se, como vimos, $\mathbf{x} = A^{-1}\mathbf{b}$, ou seja, usando a expressão de A^{-1} em termos da matriz adjunta de A:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1i} & A_{2i} & \cdots & A_{ni} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix}.$$

Da igualdade anterior, obtemos

$$x_i = \frac{1}{\det A}\left(A_{1i}b_1 + A_{2i}b_2 + \cdots + A_{ni}b_n\right); i = 1, 2, \ldots, n. \tag{3.15}$$

Seja B_i a matriz obtida de A substituindo a sua coluna i pela coluna $\mathbf{b}$ dos termos independentes, i.e.

$$B_i = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{pmatrix}.$$

Se calcularmos det B_i usando o Teorema de Laplace ao longo da coluna i, vemos que

$$\begin{aligned} \det B_i &= b_1\,A_{1i} + b_2\,A_{2i} + \cdots + b_n\,A_{ni} \\ &= A_{1i}b_1 + A_{2i}b_2 + \cdots + A_{ni}b_n. \end{aligned} \tag{3.16}$$

O resultado pretendido segue-se de imediato, combinando (3.15) e (3.16). □

3.6 Determinante do produto de matrizes

Embora não demonstremos o próximo teorema, ele é bastante importante e útil.

Teorema 3.8 (Determinante do produto de matrizes)**.** *Se A e B são matrizes quadradas da mesma ordem, então*

$$\det(AB) = \det A\ \det B.$$

Dem: Veja, e.g. [**?**, p.467].

Nota: O Exercício 3.2 do final do capítulo mostra que um resultado do mesmo tipo não é válido para a soma de matrizes.

Corolário 3.4. *Se A é invertível, então*

$$\det(A^{-1}) = \frac{1}{\det A}.$$

Dem: Como $AA^{-1} = I$, podemos concluir que $\det(AA^{-1}) = \det I = 1$ (recorde-se que I é diagonal, com todos os elementos diagonais iguais a 1 e tenha em atenção o resultado do Corolário 3.1 na pg. 67); usando o resultado do teorema anterior, vem $\det A \det(A^{-1}) = 1$ e o resultado pretendido segue-se, de imediato. □

3.7 Exercícios

Exercício 3.1. Calcule o determinante de cada uma das seguintes matrizes:

(a) $\begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$ $(\theta \in \mathbb{R})$ (c) $\begin{pmatrix} 1 & 2 & 2 \\ 3 & -1 & 0 \\ -1 & 1 & 4 \end{pmatrix}$.

Exercício 3.2. Sejam

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 5 \\ 2 & 1 \end{pmatrix}.$$

Calcule $\det A + \det B$ e $\det(A + B)$ e compare os resultados.

Exercício 3.3. Indique qual o determinante de cada uma das seguintes matrizes:

(a) $A = \begin{pmatrix} 3 & -2 & 1 & 4 \\ 0 & 2 & 5 & 9 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ (b) $B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 4 & 4 & 3 \end{pmatrix}$

(c) $C = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 3 & -4 & 5 & 7 \\ 2 & 1 & 2 & 1 \\ 3 & 1 & 3 & 3 \end{pmatrix}$ (d) $D = \begin{pmatrix} 1 & 10 & 5 & 1 \\ 2 & 20 & 3 & 1 \\ 3 & 30 & -1 & 4 \\ 4 & 40 & 2 & -8 \end{pmatrix}$

Exercício 3.4. Calcule o determinante das seguintes matrizes, fazendo uso de operações elementares:

(a) $\begin{pmatrix} 1 & 2 & 0 & 1 \\ -1 & 1 & 3 & 5 \\ 2 & 1 & 1 & 1 \\ 0 & 3 & 1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 0 & 1 & 2 \\ 2 & 1 & -1 & 1 \\ 2 & -1 & 0 & 2 \\ -2 & 1 & 3 & 4 \end{pmatrix}$. (c) $\begin{pmatrix} 1 & 1 & -1 & 0 \\ 4 & 1 & 2 & 1 \\ 2 & -1 & 1 & 2 \\ -2 & 3 & 1 & 5 \end{pmatrix}$.

Exercício 3.5. Sabendo que $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = -5$, indique qual o valor de cada um dos seguintes determinantes:

(a) $\begin{vmatrix} d & e & f \\ g & h & i \\ a & b & c \end{vmatrix}$ (b) $\begin{vmatrix} 3a & 3b & 3c \\ -d & -e & -f \\ 4g & 4h & 4i \end{vmatrix}$ (c) $\begin{vmatrix} a+g & b+h & c+i \\ d & e & f \\ g & h & i \end{vmatrix}$

(d) $\begin{vmatrix} -3a & -3b & -3c \\ d & e & f \\ g-4d & h-4e & i-4f \end{vmatrix}$ (e) $\begin{vmatrix} b & e & h \\ a & d & g \\ c & f & i \end{vmatrix}$.

Exercício 3.6. Prove que se todos os elementos de uma matriz quadrada A são números inteiros, então $\det A$ é um número inteiro.

Exercício 3.7. ([**?**, p.160]) Os números 20604, 53227, 25755, 20927 e 78421 são divisíveis por 17. Justifique que o mesmo sucede ao determinante

$$\begin{vmatrix} 2 & 0 & 6 & 0 & 4 \\ 5 & 3 & 2 & 2 & 7 \\ 2 & 5 & 7 & 5 & 5 \\ 2 & 0 & 9 & 2 & 7 \\ 7 & 8 & 4 & 2 & 1 \end{vmatrix},$$

sem calcular o seu valor.

Exercício 3.8. (a) Mostre que

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (b-a)(c-a)(c-b).$$

(b) Tendo em conta o resultado da alínea anterior, diga, justificando, se a matriz

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 7 \\ 4 & 9 & 49 \end{pmatrix}$$

é uma matriz invertível.

Exercício 3.9. Seja $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ e suponha que $\det A = 5$. Determine:

(a) $\det(3A)$ (b) $\det(A^{-1})$ c) $\det((2A)^{-1})$

Exercício 3.10. Calcule os seguintes determinantes, fazendo uso do Teorema de Laplace:

(a) $\begin{vmatrix} 1 & 3 & 1 & 1 \\ 2 & -1 & 0 & 2 \\ 3 & 2 & 1 & 2 \\ 0 & 1 & 0 & 4 \end{vmatrix}$ (b) $\begin{vmatrix} 2 & 0 & 1 & 0 \\ 4 & -1 & 0 & 1 \\ 5 & 0 & 1 & -2 \\ 7 & 1 & 0 & 1 \end{vmatrix}$ (c) $\begin{vmatrix} 1 & 0 & 4 & 0 \\ 2 & 3 & 3 & 5 \\ 0 & 3 & -5 & 5 \\ 4 & 0 & -2 & 7 \end{vmatrix}$.

Exercício 3.11. Seja $A = \begin{pmatrix} 1 & 3 & 4 \\ 6 & 3 & 1 \\ -2 & 1 & 2 \end{pmatrix}$

(a) Calcule os menores e os complementos algébricos de todos os elementos de A.

(b) Calcule $\det A$.

(c) Calcule $\operatorname{adj} A$.

(d) Calcule A^{-1}.

Exercício 3.12. Suponha que uma dada matriz A se pode decompor num produto da forma $A = LU$, onde

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{pmatrix} \quad \text{e} \quad U = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}.$$

Determine a expressão do determinante de A.

Exercício 3.13. Seja A uma matriz quadrada de ordem 3 e suponha que $\det A = 2$. Diga se as afirmações seguintes são verdadeiras ou falsas, justificando convenientemente:

(a) $\text{car}(A) = 2$.

(b) É possível encontrar uma matriz X, quadrada de ordem 3, tal que

$$AX = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix}.$$

(c) $\text{er}(A) = I_3$.

Exercício 3.14. Seja $A_{n\times n}$ invertível. Mostre que:

(a) A matriz $\text{adj}\, A$ também invertível, sendo $(\text{adj}\, A)^{-1} = \frac{1}{\det A}A = \det(A^{-1})A$.

(b) $\text{adj}\, A = \text{adj}(A^{-1})$.

(c) $\det(\text{adj}\, A) = (\det A)^{n-1}$.

Exercício 3.15. Verifique que cada um dos sistemas seguintes é um sistema de Cramer e determine a respetiva solução, usando a regra de Cramer.

(a) $\begin{cases} x_1 + 2x_2 + x_3 = 5 \\ 2x_1 + 2x_2 + x_3 = 6 \\ x_1 + 2x_2 + 3x_3 = 9 \end{cases}$ (b) $\begin{cases} x_1 + x_2 = 0 \\ x_2 + x_3 - 2x_4 = 1 \\ x_1 + 2x_3 + x_4 = 0 \\ x_1 + x_2 + x_4 = 0 \end{cases}.$

Exercício 3.16. Considere o sistema

$$\begin{cases} 2x_1 + x_2 - x_3 = 6 \\ x_1 - x_2 - x_3 = 2 \\ x_1 + x_2 + 3x_3 = 0 \end{cases}$$

e seja A a sua matriz simples.

(a) Calcule $\det A$.

(b) Justifique que o sistema é possível e determinado e determine o valor da incógnita x_1 (sem resolver totalmente o sistema).

(c) Justifique que A é invertível e determine o elemento na posição $(2, 3)$ da matriz A^{-1}, sem calcular A^{-1}.

4

Formas Quadráticas

4.1 Conceitos básicos

Definição 4.1. Uma função $f : \mathbb{R}^n \to \mathbb{R}$ diz-se um *monómio* se puder escrever-se na forma

$$f(x_1, \ldots, x_n) = c\, x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \tag{4.1}$$

onde $c \in \mathbb{R}$ e $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{N} \cup \{0\}$.. O grau do monómio (4.1) é igual a $\alpha_1 + \alpha_2 + \ldots + \alpha_n$.

Exemplo 4.1.

1. $f(x_1, x_2) = 3\, x_1\, x_2$ é um monómio de grau 2.

2. $f(x_1, x_2, x_3) = 5\, x_1^2\, x_2\, x_3^2$ é um monómio de grau 5.

Definição 4.2. Uma *forma quadrática* em $\mathbb{R}^n$ é uma soma de monómios de grau 2 (definidos em $\mathbb{R}^n$), i.e. é uma função $Q : \mathbb{R}^n \to \mathbb{R}$ da forma

$$Q(x_1, x_2, \ldots, x_n) = \sum_{i \leq j} q_{ij} x_i x_j. \tag{4.2}$$

Exemplo 4.2.

1. A função $Q_1(x_1, x_2) = 3x_1^2 + 5x_1x_2 + 4x_2^2$ é uma forma quadrática em $\mathbb{R}^2$.

2. A função $Q_2(x_1, x_2, x_3) = 3x_1^2 + 2x_1x_2 + 5x_1x_3 - x_2^2 + 3x_2x_3 + x_3^2$ é uma forma quadrática em $\mathbb{R}^3$.

4.2 Matriz de uma forma quadrática

Considere-se uma qualquer forma quadrática em $\mathbb{R}^2$

$$Q(x_1, x_2) = q_{11}x_1^2 + q_{12}x_1x_2 + q_{22}x_2^2.$$

É muito simples de verificar (verifique!) que esta forma quadrática pode escrever-se como

$$Q(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} q_{11} & q_{12} \\ 0 & q_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

isto é, na forma

$$\mathbf{x}^{\mathsf{T}} A \mathbf{x} \tag{4.3}$$

onde $A = \begin{pmatrix} q_{11} & q_{12} \\ 0 & q_{22} \end{pmatrix}$ e $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Na verdade, é fácil de verificar que poderemos escrever $Q(x_1, x_2)$ na forma (4.3), usando muitas outras matrizes $A = (a_{ij})$: basta que os seus coeficientes satisfaçam $a_{11} = q_{11}$, $a_{22} = q_{22}$ e $a_{12} + a_{21} = q_{12}$. Em particular, se desejarmos que a matriz utilizada seja *simétrica*, bastará considerar

$$a_{11} = q_{11}, \; a_{22} = q_{22} \quad \text{e} \quad a_{12} = a_{21} = \frac{1}{2}q_{12}$$

(sendo essa a única escolha possível que satisfaz a exigência de se usar uma matriz simétrica). De modo análogo, dada uma qualquer forma quadrática em $\mathbb{R}^3$

$$Q(x_1, x_2, x_3) = q_{11}x_1^2 + q_{12}x_1x_2 + q_{13}x_1x_3 + q_{22}x_2^2 + +q_{23}x_2x_3 + q_{33}x_3^2$$

existe uma (única) maneira de escrevê-la na forma $Q(x_1, x_2, x_3) = \mathbf{x}^{\mathsf{T}} A \mathbf{x}$ com A uma matriz quadrada de ordem 3, simétrica, a qual é a seguinte:

$$Q(x_1, x_2, x_3) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} q_{11} & \frac{1}{2}q_{12} & \frac{1}{2}q_{13} \\ \frac{1}{2}q_{12} & q_{22} & \frac{1}{2}q_{23} \\ \frac{1}{2}q_{13} & \frac{1}{2}q_{23} & q_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Exemplo 4.3. Consideremos de novo as formas quadráticas do Exemplo 4.2 e escrevamo-las na forma $\mathbf{x}^{\mathsf{T}} A \mathbf{x}$, usando matrizes simétricas.

1. $Q_1(x_1, x_2, x_3) = 3x_1^2 + 5x_1x_2 + 4x_2^2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3 & \frac{5}{2} \\ \frac{5}{2} & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$

2.
$$\begin{aligned} Q_2(x_1,x_2,x_3) &= 3x_1^2+2x_1x_2+5x_1x_3-x_2^2+3x_2x_3+x_3^2 \\ &= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 3 & 1 & \frac{5}{2} \\ 1 & -1 & \frac{3}{2} \\ \frac{5}{2} & \frac{3}{2} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \end{aligned}$$

De um modo mais geral, tem-se o resultado contido no quadro seguinte (cuja demonstração omitimos):

Matriz de uma forma quadrática

Teorema 4.1. *Toda a forma quadrática*

$$Q(x_1,x_2,\ldots,x_n)=\sum_{i\le j} q_{ij}x^i x^j$$

pode escrever-se de (modo único) como

$$Q(x_1,x_2,\ldots,x_n)=\mathbf{x}^{\mathsf{T}}A\mathbf{x} \tag{4.4}$$

com A uma matriz quadrada de ordem n simétrica e $\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, *sendo os coeficientes da matriz $A=(a_{ij})$ dados por*

$$a_{ii}=q_{ii}, \quad a_{ij}=a_{ji}=\frac{1}{2}q_{ij} \quad (\text{se } i<j). \tag{4.5}$$

Reciprocamente, dada uma matriz simétrica A, a expressão (4.4) define uma forma quadrática.

À única matriz simétrica associada, pela fórmula (4.4), à forma quadrática Q, chamamos matriz dessa forma quadrática.

4.3 Matrizes e formas quadráticas definidas positivas/negativas

Se considerarmos uma qualquer forma quadrática

$$Q(x_1,\ldots,x_n)=\sum_{i\le j} q_{ij}x_i x_j$$

é imediato concluir que se tem sempre $Q(x_1, x_2, \ldots, x_n) = 0$ quando $(x_1, x_2, \ldots, x_n) = (0, 0, \ldots, 0)$. Os valores que a forma quadrática pode assumir quando $(x_1, x_2, \ldots, x_n) \neq (0, 0, \ldots, 0)$ classificam a forma quadrática, de acordo com a definição seguinte.

Definição 4.3. Uma forma quadrática $Q(x_1, \ldots, x_n)$ diz-se:

- *definida positiva* (d.p.) se verificar

$$Q(x_1, \ldots, x_n) > 0, \quad \text{para qualquer } (x_1, \ldots, x_n) \neq (0, 0, \ldots, 0).$$

- *definida negativa* (d.n.) se verificar

$$Q(x_1, \ldots, x_n) < 0, \quad \text{para qualquer } (x_1, \ldots, x_n) \neq (0, 0, \ldots, 0).$$

- *semidefinida positiva* (s.d.p.) se verificar

$$Q(x_1, \ldots, x_n) \geq 0, \quad \text{para qualquer } (x_1, \ldots, x_n) \neq (0, 0, \ldots, 0).$$

- *semidefinida negativa* (s.d.n.) se verificar

$$Q(x_1, \ldots, x_n) \leq 0, \quad \text{para qualquer } (x_1, \ldots, x_n) \neq (0, 0, \ldots, 0).$$

Uma forma quadrática que não satisfaça nenhuma das condições anteriores, i.e. que assuma um valor positivo para (pelo menos) um n-uplo $(x_1, \ldots, x_n)$ e um valor negativo para (pelo menos) outro n-uplo é dita *indefinida*.

Nota: No que se segue, $\mathbf{x}$ designa sempre o vetor coluna $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ e $\mathbf{0}$ o vetor nulo $\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ com n componentes.

Uma matriz quadrada de ordem n, simétrica, "herda" a classificação em definida positiva, semidefinida positiva, definida negativa, semidefinida negativa ou indefinida, da correspondente forma quadrática $Q(x_1, \ldots, x_n) = \mathbf{x}^{\mathsf{T}} A \mathbf{x}$. Mais precisamente, temos a seguinte definição.

Definição 4.4. Seja A uma matriz quadrada de ordem n, simétrica. Essa matriz diz-se:

- *definida positiva* (d.p.), se verificar

$$\mathbf{x}^{\mathsf{T}} A \mathbf{x} > 0, \ \forall \mathbf{x} \neq \mathbf{0}$$

- *definida negativa* (d.n.), se verificar

$$\mathbf{x}^\mathsf{T} A\mathbf{x} < 0, \ \forall \mathbf{x} \neq \mathbf{0}$$

- *semidefinida positiva* (s.d.p.), se verificar

$$\mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0, \ \forall \mathbf{x} \neq \mathbf{0}$$

- *semidefinida negativa* (s.d.n.), se verificar

$$\mathbf{x}^\mathsf{T} A\mathbf{x} \leq 0, \ \forall \mathbf{x} \neq \mathbf{0}$$

Uma matriz simétrica que não satisfaça nenhuma das condições anteriores é dita *indefinida*.

Nota: Naturalmente, uma matriz (forma quadrática) que seja definida (positva ou negativa) é também uma matriz (forma quadrática) semidefinida.

Exemplo 4.4.

1. A forma quadrática em $\mathbb{R}^2$ definida por

   $$Q(x_1, x_2) = x_1^2 + x_2^2$$

   é definida positiva, uma vez que, se $(x_1, x_2) \neq (0,0)$, se tem $x_1^2 + x_2^2 > 0$. Assim, a matriz $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ é uma matriz definida positiva.

2. A forma quadrática em $\mathbb{R}^2$ definida por

   $$Q(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2$$

   é semidefinida positiva (mas não definida positiva). Com efeito, temos $x_1^2 + 2x_1x_2 + x_2^2 = (x_1 + x_2)^2 \geq 0, \forall (x_1, x_2) \neq (0,0)$, tendo-se $x_1^2 + x_2^2 = 0 \iff x_1 = -x_2$. Assim, a matriz

   $$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

   é semidefinida positiva.

3. A forma quadrática
$$Q(x_1, x_2) = x_1^2 - x_2^2$$
é indefinida. Com efeito, tem-se
$$x_1^2 - x_2^2 = (x_1 - x_2)(x_1 + x_2)$$
pelo que, por exemplo, $Q(1, 0) > 0$ e $Q(0, 1) < 0$. A matriz
$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
é indefinida.

Vamos agora enunciar alguns teoremas que, no seu conjunto, nos permitem classificar uma dada matriz simétrica (e, consequentemente, uma forma quadrática) num dos tipos acima referidos: definida positiva, definida negativa, semidefinida positiva, semidefinida negativa ou indefinida.

Antes disso, vamos introduzir algumas notações e definições.

Definição 4.5 (Menores principais e menores principais liderantes)**.** Dada uma matriz A quadrada de ordem n, qualquer sua submatriz quadrada cuja diagonal faça parte da diagonal de A é chamada uma submatriz *principal*. Os determinantes das submatrizes principais são chamados *menores principais* de A.

Uma submatriz principal contida nas primeiras k linhas e k colunas de A $(1 \leq k \leq n)$ é chamada uma submatriz principal *liderante* de A e designada por A_k; os determinantes das matrizes principais liderantes chamam-se *menores principais liderantes*.

Exemplo 4.5. Uma matriz A quadrada de ordem 3 tem 3 menores principais liderantes:
$$\det(a_{11}), \quad \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{e} \quad \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$
e 7 menores principais. Os restantes menores principais são:
$$\det(a_{22}), \quad \det(a_{33}), \quad \det\begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} \quad \text{e} \quad \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}.$$

O teorema seguinte estabelece critérios para saber se uma dada matriz é ou não definida (positiva ou negativa) em termos da análise dos seus menores principais liderantes.

Teorema 4.2. *Seja A uma matriz quadrada de ordem n, simétrica. Então, A é:*

- *definida positiva se e só se os seus menores principais liderantes forem todos positivos.*
- *definida negativa se e só se os seus menores principais liderantes de ordem par forem todos positivos e os seus menores principais liderantes de ordem ímpar forem todos negativos.*

Dada uma matriz simétrica A, se ela não verificar nenhuma das condições anteriores, podemos apenas concluir que ela não é definida, mas não podemos, de imediato, saber se ela é indefinida ou se é semidefinida. O teorema seguinte dá-nos condições **suficientes**, mas não necessárias, para que A seja indefinida.

Teorema 4.3. *Seja A uma matriz quadrada de ordem n, simétrica. Se A verificar uma das duas condições seguintes, então A é indefinida:*

1. *Existe (pelo menos) um menor principal liderante de ordem par negativo;*
2. *Existem (pelo menos) dois menores liderantes de ordem ímpar tais que um deles é positivo e outro é negativo.*

No caso em que nenhum dos teoremas anteriores permita concluir sobre o tipo da matriz A, será necessário recorrer a um outro critério, mais trabalhoso, envolvendo a análise de todos os menores principais da matriz, e não apenas dos liderantes.

Teorema 4.4. *Seja A uma matriz quadrada de ordem n, simétrica. Então A é:*

- *semidefinida positiva se e só se todos os seus menores principais forem ≥ 0;*
- *semidefinida negativa se e só se todos os seus menores principais de ordem par forem ≥ 0 e os de ordem ímpar forem ≤ 0.*

Exemplo 4.6. Seja A uma matriz quadrada de ordem 4, simétrica.

1. Se $\det A_1 > 0,\ \det A_2 > 0,\ \det A_3 > 0$ e $\det A_4 > 0$, então o Teorema 4.2 permite-nos concluir que A é d.p.
2. Se $\det A_1 < 0,\ \det A_2 > 0,\ \det A_3 < 0$ e $\det A_4 > 0$, então podemos concluir que A é d.n., usando novamente o Teorema 4.2.
3. Se $\det A_1 = 0,\ \det A_2 < 0,\ \det A_3 < 0$ e $\det A_4 > 0$, então, usando o Teorema 4.3, podemos concluir que A é indefinida.

4. Se $\det A_1 < 0$, $\det A_2 = 0$, $\det A_3 > 0$ e $\det A_4 > 0$, então, usando o Teorema 4.3, podemos concluir que A é indefinida.

5. Se $\det A_1 > 0$, $\det A_2 = 0$, $\det A_3 > 0$ e $\det A_4 > 0$, então o Teorema 4.2 garante-nos que A não é definida (por causa de $\det A_2$). Não estamos nas condições do Teorema 4.3 para podermos concluir que A é indefinida. Assim, a informação fornecida não é suficiente para saber de que tipo é a matriz, podendo esta ser semidefinida positiva ou indefinida. Seria, portanto, necessário analisar os restantes 11 menores principais de A: se eles fossem todos não negativos, A seria semidefinida negativa; se algum deles fosse negativo, A seria indefinida.

4.4 Exercícios

Exercício 4.1. Escreva cada uma das formas quadráticas seguintes na forma $\mathbf{x}^\mathsf{T} A\mathbf{x}$, com A uma matriz simétrica.

(a) $Q(x_1, x_2) = 5x_1^2 + 4x_1x_2 + x_3^2$.

(b) $Q(x_1, x_2, x_3) = 2x_1^2 + 7x_1x_2 - x_1x_3 + 8x_2^2 + 4x_2x_3 - 2x_3^2$.

(c) $Q(x_1, x_2, x_3, x_4) = x_1^2 + 7x_2^2 + 3x_3^2 - x_4^2$.

Exercício 4.2. Seja A uma matriz quadrada de ordem n.

(a) Quantos menores principais liderantes tem a matriz A?

(b) Quantos menores principais de ordem k $(1 \leq k \leq n)$ tem A?

(c) Quantos menores principais tem A, se $n = 5$? E se $n = 6$?

Exercício 4.3. Classifique cada uma das matrizes seguintes (dizendo se é d.p., d.n., s.d.p., s.d.n. ou indefinida).

a) $\begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}$ b) $\begin{pmatrix} -3 & -1 \\ -1 & -2 \end{pmatrix}$ c) $\begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$ d) $\begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix}$

e) $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 5 \\ 0 & 5 & 6 \end{pmatrix}$ f) $\begin{pmatrix} 2 & 1 & -1 \\ 1 & 4 & -2 \\ -1 & -2 & 4 \end{pmatrix}$ g) $\begin{pmatrix} -2 & 4 & -1 \\ 4 & -2 & -1 \\ -1 & -1 & -2 \end{pmatrix}$ h) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$.

Exercício 4.4. Seja A uma matriz quadrada de ordem n, simétrica. Mostre que, se A é definida positiva, então os elementos da diagonal de A são todos positivos.

Sugestão: Calcule $\mathbf{x}^\mathsf{T} A$x para $\mathbf{x} = \mathbf{e}_i; i = 1, \ldots, n$.

Mostre, através de um contra-exemplo, que o recíproco do resultado anterior não é verdadeiro.

Exercício 4.5. Mostre que toda a matriz simétrica definida positiva (ou definida negativa) é invertível.

5

Valores e Vetores Próprios

5.1 Definição de valores e vetores próprios

De um modo geral, dada uma matriz A, quadrada de ordem n e um vetor não nulo $\mathbf{x} \in \mathbb{R}^{n\times 1}$, ao efeturamos o produto $A\mathbf{x}$ obtemos um novo vetor que não é um múltiplo de $\mathbf{x}$. Por exemplo, tem-se

$$\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix},$$

e é evidente que o vetor $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ não é um múltiplo de $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Mas, por vezes, pode acontecer que $A\mathbf{x}$ seja um múltiplo de $\mathbf{x}$. Por exemplo, dada a matriz $A = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$ e o vetor $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ tem-se

$$A\mathbf{x} = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Neste caso, diremos que o vetor $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ é um vetor próprio de A associado ao valor próprio 3. Mais geralmente, temos a seguinte definição.

Definição 5.1. Seja $A \in \mathbb{R}^{n\times n}$. Um escalar $\lambda \in \mathbb{R}$ diz-se um *valor próprio* de A se e só se existir um vetor não nulo $\mathbf{x} \in \mathbb{R}^{n\times 1}$ tal que

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Neste caso, dizemos que $\mathbf{x}$ é um *vetor próprio* associado ao valor próprio λ.

Valores e vetores próprios para matrizes complexas são definidos de modo análogo; se A é uma matriz quadrada de ordem n com elementos em $\mathbb{C}$, um escalar $\lambda \in \mathbb{C}$ será um valor próprio de A se e só se existir um vetor não nulo $\mathbf{x}$ (com n componentes em $\mathbb{C}$) tal que $A\mathbf{x} = \lambda\mathbf{x}$. Nestas condições, $\mathbf{x}$ diz-se um vetor próprio de A associado ao valor próprio λ.

Como os números reais são elementos de $\mathbb{C}$, uma matriz $n \times n$ de elementos reais pode ser sempre vista como uma matriz de números complexos, pelo que, dada uma matriz com números reais, teremos de saber se procuramos apenas valores próprios reais e vetores próprios com elementos reais, ou se aceitamos também valores próprios complexos e vetores próprios em com entradas complexas.

Embora os valores próprios complexos possam ser importantes, neste curso, a não ser que algo seja dito em contrário, dada uma matriz real, quadrada de ordem n, estaremos apenas interessados nos seus valores próprios reais e em vetores próprios de $\mathbb{R}^{n\times 1}$.

O teorema seguinte é muito importante porque nos fornece um processo de cálculo dos valores próprios de uma matriz.

Teorema 5.1. *Seja $A \in \mathbb{R}^{n\times n}$. Então, λ é um valor próprio de A se e só se*

$$\det(A - \lambda I) = 0.$$

Dem:

$$\begin{aligned}
\lambda \text{ é valor próprio de } A &\iff \exists \mathbf{x} \in \mathbb{R}^{n\times 1}, \mathbf{x} \neq \mathbf{0} : A\mathbf{x} = \lambda\mathbf{x} \\
&\iff \exists \mathbf{x} \in \mathbb{R}^{n\times 1}, \mathbf{x} \neq \mathbf{0} : A\mathbf{x} = \lambda I\mathbf{x} \\
&\iff \exists \mathbf{x} \in \mathbb{R}^{n\times 1}, \mathbf{x} \neq \mathbf{0} : (A - \lambda I)\mathbf{x} = \mathbf{0} \\
&\iff \text{ O sistema } (A - \lambda I)\mathbf{x} = \mathbf{0} \text{ tem uma solução não nula} \\
&\iff \text{car}(A - \lambda I) < n \iff \det(A - \lambda I) = 0;
\end{aligned}$$

veja a Equação 3.6 na pg. 72. □

Corolário 5.1. *Seja A uma matriz quadrada de ordem n. Então, tem-se*

$$\det A = 0 \iff \lambda = 0 \textit{ é um valor próprio de } A.$$

Dem:

$$\lambda = 0 \text{ é um valor próprio de } A \iff \det(A - 0I) = 0 \iff \det A = 0.$$

□

É então imediato obter o resultado seguinte.

Corolário 5.2. *Seja A uma matriz quadrada de ordem n. Então, A é uma matriz invertível se e só se $\lambda = 0$ não for valor próprio de A.*

Vejamos um exemplo de cálculo de valores e vetores próprios de uma matriz.

Exemplo 5.1. Seja

$$A = \begin{pmatrix} 7 & -4 \\ 5 & -2 \end{pmatrix}$$

e comecemos por calcular os seus valores próprios, usando o Teorema 5.1. Temos

$$\begin{aligned} \det(A - \lambda I) = 0 &\iff \begin{vmatrix} 7-\lambda & -4 \\ 5 & -2-\lambda \end{vmatrix} = 0 \\ &\iff \lambda^2 - 5\lambda + 6 = 0 \iff \lambda = 2 \text{ ou } \lambda = 3. \end{aligned}$$

Concluímos, assim, que A tem dois valores próprios, $\lambda_1 = 2$ e $\lambda_2 = 3$.

Para determinarmos os vetores próprios associados ao valor próprio $\lambda_1 = 2$, teremos de encontrar as soluções não nulas do sistema homogéneo

$$(A - 2I)\mathbf{x} = \mathbf{0}.$$

Mas,

$$A - 2I = \begin{pmatrix} 5 & -4 \\ 5 & -4 \end{pmatrix}.$$

Podemos converter a matriz anterior na seguinte matriz com a forma em escada

$$\begin{pmatrix} 5 & -4 \\ 0 & 0 \end{pmatrix}.$$

Resolvendo o sistema homogéneo da forma habitual, vemos que ele tem como soluções todos os elementos do conjunto $\left\{ \begin{pmatrix} 4/5\,\alpha \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\}$. Assim, os vetores próprios associados ao valor próprio $\lambda_1 = 2$ serão todos os vetores não nulos do conjunto anterior, ou seja, serão os vetores do conjunto

$$\left\{ \begin{pmatrix} 4/5\,\alpha \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\} \setminus \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

De modo análogo se verifica que os vetores próprios associados ao valor $\lambda_2 = 3$ serão os vetores do conjunto

$$\left\{ \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\} \setminus \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Definição 5.2. Sendo A uma matriz quadrada de ordem n, pode provar-se que

$$p_A(\lambda) = \det(A - \lambda I) \tag{5.1}$$

é um polinómio na variável λ de grau exatamente igual a n. Este polinómio $p_A(\lambda)$ é chamado *polinómio caraterístico* de A. A equação

$$p_A(\lambda) = 0 \tag{5.2}$$

é chamada *equação caraterística* de A.

Vemos, assim, que os valores próprios de A são as raízes (reais) da sua equação caraterística ou, dito de outro modo, são os zeros (reais) do seu polinómio caraterístico.

As raízes de $p_A(\lambda) = 0$ podem, naturalmente, ser raízes múltiplas.

Definição 5.3. Seja λ um valor próprio de uma matriz quadrada A de ordem n. À multiplicidade de λ, enquanto raiz da equação caraterística de A, chamamos *multiplicidade algébrica* do valor próprio λ. Denotaremos a multiplicidade algébrica de λ por $\text{ma}(\lambda)$.

Exemplo 5.2. Seja $A = \begin{pmatrix} -4 & 2 \\ -1 & -1 \end{pmatrix}$. Temos

$$p_A(\lambda) = \begin{vmatrix} -4-\lambda & 2 \\ -1 & -1-\lambda \end{vmatrix} = \lambda^2 + 5\lambda + 6.$$

Então

$$p_A(\lambda) = 0 \iff \lambda^2 + 5\lambda + 6 = 0 \iff \lambda = -3 \text{ ou } \lambda = 2,$$

ou seja, A tem dois valores próprios distintos, $\lambda_1 = -3$ e $\lambda_2 = 2$, cada um deles com multiplicidade algébrica igual a 1.

Exemplo 5.3. Consideremos, agora, a matriz $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}$. Neste caso, temos,

$$p_A(\lambda) = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & 3-\lambda & 0 \\ 0 & 1 & 3-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 3-\lambda & 0 \\ 1 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda)^2.$$

Então,

$$p_A(\lambda) = 0 \iff 1-\lambda = 0 \text{ ou } (3-\lambda)^2 = 0 \iff \lambda = 1 \text{ ou } \lambda = 3.$$

Assim, A tem um valor próprio $\lambda_1 = 1$, com multiplicidade algébrica igual a 1, e um valor próprio $\lambda_2 = 3$, de multiplicidade algébrica igual a 2.

Neste caso, por vezes dizemos que A tem 3 valores próprios: $\lambda_1 = 1$ e $\lambda_2 = \lambda_3 = 3$, isto é, contamos o valor próprio de multiplicidade algébrica igual a 2 como 2 valores próprios (iguais).

Exemplo 5.4. Seja $A = \begin{pmatrix} 1 & 3 & 4 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}$. Neste caso, tem-se

$$p_A(\lambda) = \begin{vmatrix} 1-\lambda & 3 & 4 \\ 0 & 2-\lambda & -1 \\ 0 & 1 & 2-\lambda \end{vmatrix} = (1-\lambda) \begin{vmatrix} 2-\lambda & -1 \\ 1 & 2-\lambda \end{vmatrix} = (1-\lambda)(\lambda^2 - 4\lambda + 5).$$

Então

$$p_A(\lambda) = 0 \iff 1 - \lambda = 0 \;\text{ ou }\; \lambda^2 - 4\lambda + 5 = 0.$$

Como $\Delta = (-4)^2 - 4 \times 1 \times 5 = 16 - 20 = -4 < 0$, a equação $\lambda^2 - 4\lambda + 5 = 0$ não tem raízes reais. Podemos, portanto dizer que A tem apenas uma valor próprio, $\lambda = 1$, de multiplicidade algébrica igual a 1.

Note-se que, se considerássemos a matriz A como uma matriz de números complexos e aceitássemos valores próprios complexos para A, então A teria três valores próprios $\lambda_1 = 1$, $\lambda_2 = 2+i$ e $\lambda_3 = 2-i$.

5.2 Diagonalização de matrizes

Definição 5.4. Sejam A e B matrizes quadradas de ordem n. Dizemos que A é *semelhante* a B, se existir uma matriz invertível S tal que

$$B = S^{-1}AS.$$

Note-se que, se A é semelhante a B, também B é semelhante a A (porquê?) e, por isso, podemos dizer que A e B *são semelhantes* (em vez de dizer que A é semelhante a B ou que B é semelhante a A).

Definição 5.5. Uma matriz A quadrada de ordem n diz-se *diagonalizável*, se for semelhante a uma matriz diagonal, i.e. se existir uma matriz invertível S e uma matriz diagonal D, tais que

$$S^{-1}AS = D.$$

Neste caso, dizemos que a matriz S *diagonaliza* A ou é uma matriz *diagonalizante* para A.

O que nos vai interessar agora é responder à questão de saber quando uma dada matriz A é ou não diagonalizável.

Começamos por referir (sem demonstrar) um caso em que se sabe, de imediato, que A não é diagonalizável.

Teorema 5.2. *Seja A quadrada de ordem n e suponhamos que A tem k valores próprios distintos, $\lambda_1, \lambda_2, \ldots, \lambda_k$. Se* ma$(\lambda_1)+ma(\lambda_2)+\cdots+ma(\lambda_k) < n$, *então A não é diagonalizável*

Nota: Quando referirmos que A não é diagonalizável, queremos dizer que não existe uma matriz **real** S, quadrada de ordem n e invertível tal que $S^1AS = D$, com D diagonal. No entanto, A pode ser diagonalizável se aceitarmos trabalhar com matrizes complexas.

Suponhamos, agora, que ma$(\lambda_1)+\cdots+$ma$(\lambda_k) = n$, isto é, que A tem n valores próprios reais, quando contamos os valores próprios de multiplicidade m_k como m_k valores próprios. O teorema seguinte diz-nos em que condições A será uma matriz diagonalizável.

Nota: No que se segue, por uma questão de simplicidade, usamos a notação diag$(d_1, d_2, \ldots, d_n)$ para designar uma matriz quadrada de ordem n diagonal com os elementos $d_1, d_2, \ldots, d_n$ (por esta ordem) na diagonal.

Teorema 5.3. *Seja A uma matriz quadrada de ordem n.*

1. *Suponhamos que A tem n valores próprios reais, não necessariamente distintos, $\lambda_1, \lambda_2, \ldots, \lambda_n$, com vetores próprios associados $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, respetivamente. Seja S a matriz que tem esses vetores próprios como colunas, i.e.*

$$S = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{pmatrix}. \tag{5.3}$$

   Se a matriz S for invertível, então tem-se

$$S^{-1}AS = D, \tag{5.4}$$

   onde D é a matriz diagonal cujos elementos diagonais são os valores próprios de A, i.e. $D =$ diag$(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Por outras palavras, nestas condições A é diagonalizável, sendo semelhante à matriz diagonal cujos elementos diagonais são os seus valores próprios e, além disso, a matriz que tem os vetores próprios como colunas é uma matriz diagonalizante para A.

2. *Reciprocamente, suponhamos que A é diagonalizável, isto é, que temos $S^{-1}AS = D$ para uma certa matriz $S \in \mathbb{R}^{n\times n}$, invertível e com $D =$ diag$(d_1, d_2, \ldots, d_n)$. Nesse caso, os elementos da diagonal de D são valores próprios de A, tendo como vetores próprios associados as colunas de S.*

Dem:

1. Como $A\mathbf{v}_i = \lambda_i \mathbf{v}_i; i = 1, \ldots, n$, tem-se

$$\begin{aligned}
AS &= A\Big(\mathbf{v}_1\ \mathbf{v}_2\ \cdots \mathbf{v}_n\Big) \\
&= \Big(A\mathbf{v}_1\ A\mathbf{v}_2\ \cdots A\mathbf{v}_n\Big) \\
&= \Big(\lambda_1\mathbf{v}_1\ \ \lambda_2\mathbf{v}_2\ \ \cdots\ \ \lambda_n\mathbf{v}_n\Big) \\
&= \Big(\mathbf{v}_1\ \mathbf{v}_2\ \cdots \mathbf{v}_n\Big)\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \\
&= SD.
\end{aligned}$$

   Como, por hipótese, S é invertível, podemos multiplicar ambos os membros da igualdade

$$AS = SD,$$

   à esquerda, por S^{-1}, obtendo-se

$$S^{-1}AS = S^{-1}SD = D,$$

   como queríamos mostrar.

2. Suponhamos agora que

$$S^{-1}AS = D,$$

   com $D = \text{diag}(d_1, d_2, \ldots, d_n)$. Multiplicando ambos os membros da igualdade $S^{-1}AS = D$, à esquerda, pela matriz S, obtemos

$$AS = SD.$$

   Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ as colunas de S. A igualdade $AS = SD$ escreve-se, então como

$$A\Big(\mathbf{v}_1\ \ \mathbf{v}_2\ \ \cdots\ \ \mathbf{v}_n\Big) = \Big(\mathbf{v}_1\ \ \mathbf{v}_2\ \ \cdots\ \ \mathbf{v}_n\Big)\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix},$$

   ou seja, temos

$$\Big(A\mathbf{v}_1\ \ A\mathbf{v}_2\ \ \cdots\ \ A\mathbf{v}_n\Big) = \Big(d_1\mathbf{v}_1\ \ d_2\mathbf{v}_2\ \ \cdots\ \ d_n\mathbf{v}_n\Big).$$

Segue-se, então, que

$$A\mathbf{v}_1 = d_1\mathbf{v}_1,\ A\mathbf{v}_2 = d_2\mathbf{v}_2,\ \dots,\ A\mathbf{v}_n = d_n\mathbf{v}_n.$$

Como todos os vetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ são não nulos (note-se que são as colunas de uma matriz invertível), as igualdades acima mostram que $d_1, d_2, \dots, d_n$ são valores próprios de A com vetores próprios associados $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ (as colunas de S), como se pretendia mostrar.

□

O teorema seguinte, que daremos sem demonstração, estabelece uma **condição suficiente** (mas não necessária) para que uma matriz seja diagonalizável.

Teorema 5.4. *Seja A uma matriz quadrada de ordem n. Suponhamos que A tem n valores próprios $\lambda_1, \dots, \lambda_n$* **distintos** *e sejam $\mathbf{v}_1, \dots, \mathbf{v}_n$ vetores próprios associados a $\lambda_1, \dots, \lambda_n$, respetivamente. Então, a matriz com esses vetores próprios como colunas é uma matriz invertível e, portanto, A é diagonalizável.*

Quando A tem n valores próprios reais (contando um valor próprio de multiplicidade m_k como m_k valores próprios), mas não está nas condições do teorema anterior, isto é, quando há valores próprios múltiplos, A pode ou não ser diagonalizável. Tal dependerá, de acordo com o Teorema 5.3, de ser possível ou não encontrar n vetores próprios de A que formem uma matriz quadrada de ordem n, invertível (isto é, uma matriz de caraterística n). Para responder à questão de saber se tal é possível ou não, vamos precisar de um novo conceito, que passamos a apresentar.

Definição 5.6. Seja λ um valor próprio de A. Então, o sistema homogéneo $(A - \lambda I)\mathbf{x} = \mathbf{0}$ é, como sabemos, um sistema indeterminado. Chamamos *multiplicidade geométrica* de λ, e denotamos por mg(λ), ao grau de indeterminação deste sistema, isto é,

$$\text{mg}(\lambda) = n - \text{car}(A - \lambda I).$$

Pode mostrar-se que a multiplicidade geométrica de um valor próprio nunca pode exceder a sua multiplicidade algébrica.

Além disso, mostra-se que uma matriz A (que tenha n valores próprios reais) será diagonalizável se e só se a multiplicidade geométrica de cada valor próprio for igual à sua multiplicidade algébrica.

Vamos ver exemplos que ilustram como podemos proceder, nos diversos casos, para decidir se A é ou não diagonalizável e para encontrar a matriz diagonalizante, em caso afirmativo.

Exemplo 5.5. A matriz considerada no Exemplo 5.4 não é uma matriz diagonalizável, já que apenas temos um valor próprio λ_1 e ma$(\lambda_1) = 1 < 3$.

Exemplo 5.6. Consideremos novamente a matriz $A = \begin{pmatrix} 7 & -4 \\ 5 & -2 \end{pmatrix}$ estudada no Exemplo 5.1, a qual vimos ter dois valores próprios distintos, $\lambda_1 = 2$ e $\lambda_2 = 3$. Vimos também que os vetores próprios associados ao valor próprio $\lambda_1 = 2$ são os vetores do conjunto

$$\left\{ \begin{pmatrix} 4/5\,\alpha \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\} \setminus \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

e que os vetores próprios associados ao valor $\lambda_2 = 3$ serão os vetores do conjunto

$$\left\{ \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\} \setminus \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

De acordo com o Teorema 5.4, bastará escolher um vetor próprio $\mathbf{v}_1$ associado ao valor próprio λ_1 e um vetor próprio $\mathbf{v}_2$ associado ao valor próprio λ_2, sendo a matriz formada por esses vetores uma matriz diagonalizante para A. Por exemplo, escolhendo $\mathbf{v}_1 = \begin{pmatrix} 4/5 \\ 1 \end{pmatrix}$ e $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, obter-se-á a matriz

$$S = \begin{pmatrix} 4/5 & 1 \\ 1 & 1 \end{pmatrix},$$

como matriz diagonalizante para A.

Apenas por curiosidade, vamos verificar que, de facto, assim é. É fácil de ver que car$(S) = 2$, pelo que S é invertível; se determinarmos a inversa de S, vemos que ela é a matriz $\begin{pmatrix} -5 & 5 \\ 5 & -4 \end{pmatrix}$. Então, tem-se

$$S^{-1}AS = \begin{pmatrix} -5 & 5 \\ 5 & -4 \end{pmatrix} \begin{pmatrix} 7 & -4 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} 4/5 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Exemplo 5.7. Consideremos agora a matriz do Exemplo 5.3,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix},$$

que, como vimos, tem valores próprio $\lambda_1 = 1$ e $\lambda_2 = 3$, tais que ma$(\lambda_1) = 1$ e ma$(\lambda_2) = 2$. Como ma$(\lambda_1) +$ ma$(\lambda_2) = 3 = n$, esta matriz pode ou não ser diagonalizável. Para podermos saber se sim ou não, teremos de encontrar a multiplicidade geométrica dos valores próprios.

Se resolvermos o sistema homogéneo $(A - \lambda_1 I)\mathbf{x} = \mathbf{0}$, vemos que ele é um sistema simplesmente indeterminado e que as suas soluções formam o conjunto

$$\left\{ \begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \right\}. \tag{5.5}$$

Por outro lado, se considerarmos o sistema $(A - \lambda_2 I)\mathbf{x} = \mathbf{0}$, isto é, o sistema cuja matriz simples é

$$\begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

é imediato concluir que ele é simplesmente indeterminado (logo, $\text{mg}(\lambda_2) = 1$) e tem por soluções os vetores do conjunto

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\}. \tag{5.6}$$

Neste caso, A não vai ser diagonalizável, já que $\text{mg}(\lambda_2) < \text{ma}(\lambda_2)$.

De facto, é imediato reconhecer que não vai ser possível encontrar três vetores próprios de A que formem uma matriz invertível: uma vez que os vetores próprios são, ou da forma $\begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix}$ ou da forma $\begin{pmatrix} 0 \\ 0 \\ \alpha \end{pmatrix}$, ao escolhermos três vetores próprios, um deles será sempre um múltiplo de outro, pelo que a matriz por eles formada terá duas colunas proporcionais, ou seja, terá determinante igual a zero, não sendo portanto invertível.

Exemplo 5.8. Consideremos a matriz

$$A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$$

Neste caso, se calcularmos o seu polinómio caraterístico, vem (após algum trabalho):

$$p_A(\lambda) = (1 - \lambda)(2 - \lambda)^2,$$

pelo que concluímos que A tem dois valores próprios $\lambda_1 = 1$ e $\lambda = 2$, tais que $\text{ma}(\lambda_1) = 1$ e $\text{ma}(\lambda_2) = 2$. Como $\text{ma}(\lambda_1) + \text{ma}(\lambda_2) = 3 = n$, a matriz A pode ou não ser diagonalizável.

O sistema $A - \lambda_1 \mathbf{x} = \mathbf{0}$ é um sistema simplesmente indeterminado, cujas soluções formam o conjunto

$$\left\{ \begin{pmatrix} \alpha \\ -\frac{1}{3}\alpha \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\}. \tag{5.7}$$

(Verifique!). Se considerarmos o sistema $A - \lambda_2 \mathbf{x} = \mathbf{0}$, vemos que ele é duplamente indeterminado, tendo por soluções os vetores do conjunto

$$\left\{ \begin{pmatrix} 2\alpha + 2\beta \\ \alpha \\ \beta \end{pmatrix} : \alpha, \beta \in \mathbb{R} \right\}. \tag{5.8}$$

Neste caso, temos $\mathsf{mg}(\lambda_2) = \mathsf{ma}(\lambda_2)$, o mesmo se passando, naturalmente, com o valor próprio λ_1.[1] Logo, A é diagonalizável, isto é, vai ser possível encontrar 3 vetores próprios de A, $\mathbf{v}_1, \mathbf{v}_2$ e $\mathbf{v}_3$ que, dispostos em colunas de uma matriz, formem uma matriz invertível. Vejamos como podemos proceder.

Vamos escolher para $\mathbf{v}_1$ um vetor próprio associado ao valor próprio λ_1, isto é, um vetor não nulo do conjunto (5.7), por exemplo, o vetor que se obtém quando tomamos $\alpha = 1$: $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1/3 \\ 1 \end{pmatrix}$. Do conjunto (5.8), se escolhermos $\mathbf{v}_2$ fazendo $\alpha = 1$ e $\beta = 0$ e $\mathbf{v}_3$ tomando $\alpha = 0$ e $\beta = 1$, obtemos os dois vetores próprios (associados a $\lambda_2 = 3$), $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ e $\mathbf{v}_3 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$. Obtemos, então, a seguinte matriz S:

$$S = \begin{pmatrix} 1 & 2 & 2 \\ -1/3 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

É fácil de verificar que esta matriz tem caraterística 3, sendo por isso, invertível; sendo formada por vetores próprios de A, ela será uma matriz diagonalizante para A.

Vamos ver que, de facto, assim é. Calculando a inversa de S, obtemos

$$S^{-1} = \begin{pmatrix} -3 & 6 & 6 \\ -1 & 3 & 2 \\ 3 & -6 & 5 \end{pmatrix}.$$

[1] Todos os valores próprios simples, i.e., como multiplicidade algébrica igual a 1, têm multiplicidade geométrica também igual a 1, já que, como dissemos, para qualquer valor próprio λ tem-se sempre $\mathsf{mg}(\lambda) \leq \mathsf{ma}(\lambda)$ e, além disso, $\mathsf{mg}(\lambda) \geq 1$.

Tem-se, então

$$S^{-1}AS = \begin{pmatrix} -3 & 6 & 6 \\ -1 & 3 & 2 \\ 3 & -6 & 5 \end{pmatrix} \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \\ -1/3 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

.

5.3 Valores próprios de matrizes simétricas

Existem alguns resultados especiais relativos aos valores e vetores próprios de matrizes simétricas, que referimos agora.

Teorema 5.5. *Seja A uma matriz quadrada de ordem n, simétrica. Então:*

1. *O polinómio caraterístico de A, $p_A(\lambda)$, tem apenas zeros reais, isto é, A tem n valores próprios reais (contando cada valor próprio de A com multiplicidade m_k como m_k valores próprios).*

2. *Se $\mathbf{x}$ e $\mathbf{y}$ são vetores próprios associados a dois valores próprios distintos, λ e μ, respetivamente, então*
$$\mathbf{x}^{\mathsf{T}}\mathbf{y} = 0.$$

Dem: Vamos demonstrar apenas a afirmação 2.; a demonstração de 1. pode ser vista, por exemplo, em [**?**].

Sejam, então, $\mathbf{x}$ e $\mathbf{y}$ vetores próprios associados a dois valores próprios distintos, λ e μ, respetivamente. Isto significa que $\mathbf{x} \neq \mathbf{0}$, $\mathbf{y} \neq \mathbf{0}$ e

$$A\mathbf{x} = \lambda\mathbf{x} \quad \text{e} \quad A\mathbf{y} = \mu\mathbf{y},$$

Multiplicando estas duas igualdades, à esquerda, por $\mathbf{y}^{\mathsf{T}}$ e $\mathbf{x}^{\mathsf{T}}$, respetivamente, obtém-se

$$\mathbf{y}^{\mathsf{T}}A\mathbf{x} = \lambda\mathbf{y}^{\mathsf{T}}\mathbf{x} \quad \text{e} \quad \mathbf{x}^{\mathsf{T}}A\mathbf{y} = \mu\mathbf{x}^{\mathsf{T}}\mathbf{y}.$$

Transpondo ambos os membros da primeira destas igualdades, e atendendo a que A é simétrica (i.e. $A = A^{\mathsf{T}}$), obtém-se

$$\mathbf{x}^{\mathsf{T}}A\mathbf{y} = \lambda\mathbf{x}^{\mathsf{T}}\mathbf{y}.$$

Temos, então, que

$$\lambda\mathbf{x}^{\mathsf{T}}\mathbf{y} = \mu\mathbf{x}^{\mathsf{T}}\mathbf{y},$$

ou seja, que

$$(\lambda - \mu)\mathbf{x}^{\mathsf{T}}\mathbf{y} = 0.$$

Como $\lambda \neq \mu$, segue-se que terá de ser $\mathbf{x}^{\mathsf{T}}\mathbf{y} = 0$, como se queria demonstrar. $\square$

Definição 5.7. Uma matriz Q, quadrada de ordem n, diz-se *ortogonal* se e só se satisfizer

$$Q^{\mathsf{T}}Q = I_n.$$

Por outras palavras, uma matriz ortogonal tem como inversa a sua transposta.

Seja Q uma matriz quadrada e sejam $\mathbf{q}_1, \ldots, \mathbf{q}_n$ as suas colunas. Então Q será ortogonal se e só se tivermos

$$\begin{pmatrix} \mathbf{q}_1^{\mathsf{T}} \\ \mathbf{q}_2^{\mathsf{T}} \\ \vdots \\ \mathbf{q}_n^{\mathsf{T}} \end{pmatrix} \begin{pmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

ou seja, Q será ortogonal se e só se as suas colunas satisfizerem

$$\mathbf{q}_i^{\mathsf{T}}\mathbf{q}_j = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i = j \end{cases}.$$

Teorema 5.6 (Teorema espetral)**.** *Seja A uma matriz quadrada de ordem n, simétrica. Então existe uma matriz ortogonal Q tal que*

$$Q^{\mathsf{T}}AQ = D,$$

onde D é a matriz diagonal cujos elementos diagonais são os valores próprios de A.

Dem: Vamos fazer a demonstração no caso particular em que A tem n valores próprios distintos; a demonstração no caso geral pode ser vista, por exemplo, em [**?**].

Se A tem n valores próprios distintos, $\lambda_1, \ldots, \lambda_n$, como esses valores próprios são reais, estamos nas condições do Teorema 5.4; formando uma matriz Q com vetores próprios $\mathbf{v}_1, \ldots, \mathbf{v}_n$ associados a esses valores próprios como colunas, essa matriz será uma matriz invertível e diagonalizante para A. Mas, de acordo com o teorema anterior, ter-se-á

$$\mathbf{v}_i^{\mathsf{T}}\mathbf{v}_j = 0, \quad \text{para } i \neq j.$$

Por outro lado, é sempre possível escolher os vetores próprios de tal modo que $\mathbf{v}_i^{\mathsf{T}}\mathbf{v}_i = 1$: de facto, como qualquer vetor próprio $\mathbf{x}_i$ é não nulo, ter-se-á sempre

$$\mathbf{x}_i^{\mathsf{T}}\mathbf{x}_i = x_{1i}^2 + \cdots + x_{ni}^2 > 0,$$

pelo que poderemos começar por considerar vetores próprios quaisquer $\mathbf{x}_1, \ldots, \mathbf{x}_n$ e escolher, então, $\mathbf{v}_i = \dfrac{1}{\sqrt{\mathbf{x}_i^\mathsf{T}\mathbf{x}_i}}\mathbf{x}_i$. Com esta escolha de vetores próprios, a matriz $Q = \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}$ que diagonaliza A será, então, uma matriz ortogonal, pelo que teremos

$$Q^{-1}AQ = Q^\mathsf{T}AQ = D.$$

□

Exemplo 5.9. Consideremos a seguinte matriz simétrica $A = \begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}$. Temos

$$p_A(\lambda) = \begin{vmatrix}2-\lambda & 1\\ 1 & 2-\lambda\end{vmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3$$

Emtão,

$$p_A(\lambda) = 0 \iff \lambda^2 - 4\lambda + 3 = 0 \iff \lambda = 1 \quad \text{ou} \quad \lambda = 3.$$

Resolvendo o sistema homogéneo $(A - 1I)\mathbf{x} = \mathbf{0}$, obtém-se o seguinte conjunto de soluções

$$\left\{\begin{pmatrix}-\alpha\\ \alpha\end{pmatrix} : \alpha \in \mathbb{R}\right\}.$$

Por outro lado, o sistema $(A - 3I)\mathbf{x} = \mathbf{0}$ tem como soluções os vetores do conjunto

$$\left\{\begin{pmatrix}\alpha\\ \alpha\end{pmatrix} : \alpha \in \mathbb{R}\right\}.$$

Escolhamos, então, um vetor não nulo de cada um destes conjuntos, isto é, um vetor próprio associado ao valor próprio $\lambda_1 = 1$ e um vetor próprio associado ao valor próprio $\lambda_2 = 3$. Por exemplo, consideremos $\mathbf{x}_1 = \begin{pmatrix}-1\\ 1\end{pmatrix}$ e $\mathbf{x}_2 = \begin{pmatrix}1\\ 1\end{pmatrix}$. Temos

$$\mathbf{x}_1^\mathsf{T}\mathbf{x}_1 = \begin{pmatrix}-1 & 1\end{pmatrix}\begin{pmatrix}-1\\ 1\end{pmatrix} = 2,$$

pelo que devemos considerar

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}}\mathbf{x}_1 = \begin{pmatrix}-1/\sqrt{2}\\ 1/\sqrt{2}\end{pmatrix}.$$

Temos, também, $\mathbf{x}_2^\mathsf{T}\mathbf{x}_2 = 2$, pelo que devemos considerar

$$\mathbf{v}_2 = \begin{pmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{pmatrix}.$$

Obtemos, então, a seguinte matriz Q, ortogonal e diagonalizante para A:

$$Q = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Q é, de facto, ortogonal, já que

$$Q^{\mathsf{T}}Q = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Além disso, Q diagonaliza A:

$$Q^{\mathsf{T}}AQ = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

5.4 Formas quadráticas e valores próprios

Seja $Q(x_1, \ldots, x_n)$ uma forma quadrática e escrevamo-la como

$$Q(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}A\mathbf{x}$$

, com A uma matriz simétrica.

Então, é possível classificar a matriz A (e portanto, a forma Q) em d.p., d.n., s.d.p,.s.d.n. ou indefinida, em função dos valores próprios de A, de acordo com o seguinte teorema.

Teorema 5.7. *Seja A uma matriz simétrica de ordem n e sejam $\lambda_1, \ldots, \lambda_n$ os seus valores próprios (não necessariamente distintos). Então:*

1. *A é definida positiva se e só se todos os seus v.p.'s são positivos;*
2. *A é definida negativa se e só se todos os seus v.p.'s são negativos;*
3. *A é semidefinida positiva se e só todos seus v.p.'s são não negativos;*
4. *A é semidefinida negativa se e só se todos os seus v.p.'s são não positivos;*
5. *A é indefinida se e só se A tem (pelo menos) um valor próprio positivo e (pelo menos) um valor próprio negativo.*

Dem:

Vamos demonstrar 1., sendo as restantes demonstrações totalmente idênticas.

Como A é simétrica, sabemos que existe uma matriz ortogonal Q tal que $Q^\mathsf{T}AQ = D$, com $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Mas, $Q^\mathsf{T}AQ = D \Rightarrow A = QDQ^\mathsf{T}$. Tem-se, então, para qualquer vetor $\mathbf{x} \in \mathbb{R}^{n\times 1}$:

$$\mathbf{x}^\mathsf{T}A\mathbf{x} = \mathbf{x}^\mathsf{T}\left(QDQ^\mathsf{T}\right)\mathbf{x} = \mathbf{x}^\mathsf{T}(Q^\mathsf{T})^\mathsf{T}DQ^\mathsf{T}\mathbf{x}$$

$$= (Q^\mathsf{T}\mathbf{x})^\mathsf{T}D(Q^\mathsf{T}\mathbf{x}) = \mathbf{y}^\mathsf{T}D\mathbf{y},$$

com

$$\mathbf{y} = Q^\mathsf{T}\mathbf{x}.$$

Note-se que, sendo Q invertível, teremos

$$\mathbf{x} \neq \mathbf{0} \iff \mathbf{y} = Q^\mathsf{T}\mathbf{x} \neq \mathbf{0}.$$

Então, tem-se

$$\begin{aligned} A \text{ é d.p.} &\iff \mathbf{x}^\mathsf{T}A\mathbf{x} > 0, \forall \mathbf{x} \neq \mathbf{0} \\ &\iff \mathbf{y}^\mathsf{T}D\mathbf{y} > 0,\ \forall \mathbf{y} \neq 0 \\ &\iff \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 > 0,\ \forall \mathbf{y} \neq 0, \\ &\iff \lambda_1 > 0, \lambda_2 > 0, \ldots, \lambda_n > 0. \end{aligned}$$

como se pretendia mostrar. □

5.5 Exercícios

Exercício 5.1. Considere a matriz $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$. Sem resolver a sua equação caraterística, indique quais dos vetores considerados nas alíneas seguintes são vetores próprios de A e, para os que o forem, identifique o valor próprio a que estão associados.

(a) $\begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$; (b) $\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$; (c) $\begin{pmatrix} 0 \\ 5 \\ 0 \end{pmatrix}$; (d) $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Exercício 5.2. Determine os valores próprios e os vetores próprios de cada uma das matrizes seguintes:

(a) $A = \begin{pmatrix} 4 & -4 \\ -4 & -2 \end{pmatrix}$; (b) $A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 3 & 2 \\ -2 & -3 & 10 \end{pmatrix}$; (c) $A = \begin{pmatrix} 3 & 2 & 4 & 1 \\ 0 & 2 & 2 & -1 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Exercício 5.3. Seja A uma matriz quadrada e seja λ um valor próprio de A com vetor próprio associado $\mathbf{x}$. Prove que:

(a) Dado $\alpha \in \mathbb{R}$, $\alpha\lambda$ é um valor próprio da matriz αA, com vetor próprio associado $\mathbf{x}$.

(b) λ^2 é um valor próprio da matriz A^2, com vetor próprio associado $\mathbf{x}$.

(c) Dado $\beta \in \mathbb{R}$, $\lambda - \beta$ é um valor próprio da matriz $A - \beta I$, com vetor próprio associado $\mathbf{x}$-

(d) Se A é invertível, então $\frac{1}{\lambda}$ é um valor próprio da matriz A^{-1} com vetor próprio associado $\mathbf{x}$.

Exercício 5.4. Justifique a seguinte afirmação: "*Os valores próprios de uma matriz triangular são iguais aos elementos da sua diagonal principal.*"

Exercício 5.5. Considere a matriz $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$.

(a) Seja B a matriz obtida de A trocando a ordem das suas linhas. Calcule os valores próprios de A e de B e compare-os.

(b) Seja C a matriz obtida de A multiplicando a sua primeira linha por 3. Calcule os valores próprios de C e compare-os com os de A.

(c) Seja D a matriz obtida de A adicionando à sua primeira linha, a segunda linha multiplicada por -1. Determine os valores próprios de D e compare-os com os de A.

Nota: O objectivo deste exercício é mostrar que, em geral, as operações elementares sobre as linhas de uma matriz **alteram** os valores próprios da matriz (e é fácil de ver que o mesmo se passa com as operações elementares sobre as colunas), pelo que, dada uma certa matriz A, **não podemos** começar por transformar A numa matriz mais simples B e usar $\det(B - \lambda I) = 0$ para calcular os valores próprios de A. Podemos, no entanto, efetuar operações elementares sobre as linhas ou colunas da matriz $(A - \lambda I)$ quando estamos a calcular $p_A(\lambda) = \det(A - \lambda I)$.

Exercício 5.6. Seja $A = \begin{pmatrix} 2 & 3 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}$.

(a) Determine os valores próprios e vetores próprios de A.

(b) Justifique que A é diagonalizável e determine uma matriz S que a diagonaliza.

Exercício 5.7. Para cada uma das matrizes simétricas dadas nas alíneas seguintes, encontre uma matriz ortogonal que as diagonaliza.

(a) $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. (b) $B = \begin{pmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}$.

Exercício 5.8. (a) Seja D uma matriz diagonal, $D = \text{diag}(d_1, \ldots, d_n)$. Mostre que, para $k \in \mathbb{N}$, se tem $D^k = \text{diag}(d_1^k, \ldots, d_n^k)$.

(b) Mostre que, se $A = SBS^{-1}$, então, para qualquer $k \in \mathbb{N}$, se tem $A^k = SB^kS^{-1}$.

(c) Mostre que, se $S^{-1}AS = D$, com $D = \text{diag}(d_1, \ldots, d_n)$, então, para qualquer $k \in \mathbb{N}$, $A^k = S\,\text{diag}(d_1^k, \ldots, d_n^k)S^{-1}$.

(d) Use o resultado da alínea anterior para calcular A^5, onde A é a matriz considerada no Exemplo 5.8.

Exercício 5.9. Seja A uma matriz quadrada de ordem 3 cujos valores próprios são $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \frac{1}{3}$ e $\lambda_3 = \frac{1}{4}$. Considere as sucessivas potências de A: A, $A^2, A^3, A^4, \ldots$ Para que matriz tende A^k quando $k \to \infty$? Justifique.

6

Espaços Euclideanos

Neste capítulo começaremos por recordar algumas definições e operações básicas com vetores no plano (espaço bi-dimensional, $\mathbb{R}^2$) e no espaço usual (espaço tri-dimensional, $\mathbb{R}^3$), que os alunos já conhecem do ensino secundário. Da disciplina de Cálculo para a Economia os alunos já devem também estar familiarizados com o conjunto $\mathbb{R}^n$, para outros valores de $n \in \mathbb{N}$.[1] O objetivo deste capítulo é estudar um pouco mais pormenorizadamente o conjunto $\mathbb{R}^n$.

6.1 O espaço Euclideano $\mathbb{R}^2$

Suponhamos fixado um sistema de eixos no plano, o qual, por simplicidade, escolhemos como ortogonal e monométrico (também dito ortonormado (o.n.)), i.e. constituído por duas retas orientadas (*eixos*) perpendiculares entre si e com a mesma unidade de comprimento em ambas as retas. O ponto de interseção das duas retas é a origem do sistema de eixos, vulgarmente designado por O ou $\mathbf{0}$. Como sabemos, cada ponto P do plano fica totalmente identificado por um par ordenado de números reais (x_1, x_2) – as chamadas *coordenadas* desse ponto –, sendo x_1 a *abcissa* e x_2 a *ordenada* de P e, reciprocamente, a todo o par de números reais (x_1, x_2) corresponde um e um só ponto do plano; ver Figura 6.1.

O conjunto de todos os pares ordenados de números reais denota-se por $\mathbb{R}^2$, isto é, tem-se

$$\mathbb{R}^2 = \big\{(x_1, x_2) : x_1, x_2 \in \mathbb{R}\big\}.$$

Vemos assim, que é possível estabelecer uma correspondência biunívoca entre o conjunto $\mathbb{R}^2$ e o conjunto de pontos do plano e é por isso que falaremos muitas vezes, com algum abuso de linguagem, no plano $\mathbb{R}^2$.

[1] Que assumimos que os alunos conheciam, aliás, quando falámos nas formas quadráticas!

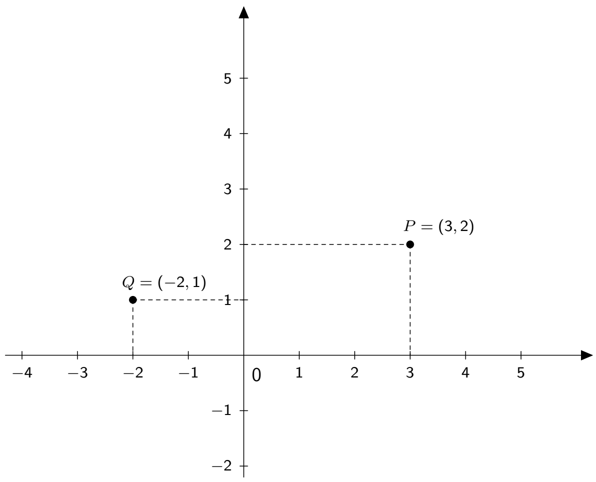

Figura 6.1: Coordenadas de pontos no plano

Um elemento qualquer (x_1, x_2) de $\mathbb{R}^2$ pode também ser identificado, geometricamente, com o segmento orientado $[O, P]$ que une a origem do sistema de eixos ao ponto P, de coordenadas (x_1, x_2); ver Fig. 6.2.

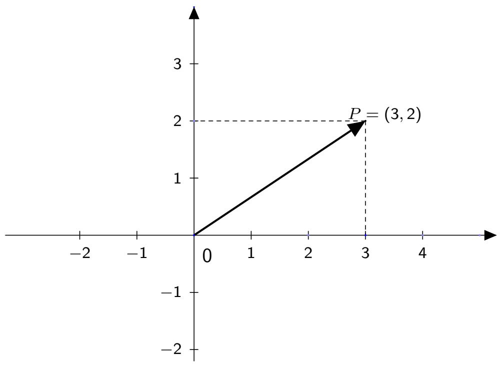

Figura 6.2: O segmento orientado $[O, P]$, onde $P = (3, 2)$.

Assim sendo, também podemos identificar o conjunto $\mathbb{R}^2$ com o conjunto de todos os segmentos orientados que têm como ponto inicial a origem do sistema de eixos.[2]

Vamos identificar todos os segmentos orientados que tenham a mesma direção, o mesmo

[2]Ao par $(0, 0)$ correspoderá o segmento orientado nulo, ou seja, aquele cuja origem (ponto inicial) e extremidade (ponto final) coincidem; o segmento orientado nulo tem direção e sentido indeterminados e comprimento igual a zero.

sentido e o mesmo comprimento, independentemente de qual seja o seu ponto inicial, com um único objeto, a que chamaremos *vetor livre* (ou simplesmente *vetor*). Assim, um vetor é caraterizado por uma direção, sentido e comprimento.[3] Qualquer um dos segmentos orientados que caraterizam o vetor pode ser usado para o representar, isto é, pode ser escolhido como um *representante* desse vetor. Os vetores serão vulgarmente designados por letras minúsculas em negrito, por exemplo, $\mathbf{x}, \mathbf{y}, \mathbf{x}, \mathbf{y}$, etc.

O conjunto $\mathbb{R}^2$ pode também ser identificado com o conjunto dos vetores do plano: a um elemento (x_1, x_2) de $\mathbb{R}^2$ faremos corresponder o (único) vetor livre que tem como representante o segmento orienatdo $[O, P]$, onde P é o ponto de coordenadas (x_1, x_2); reciprocamente, dado um vetor, se escolhermos para seu representante o segmento orientado cujo ponto inicial está na origem do sistema de eixos, as coordenadas (x_1, x_2) do ponto final desse segmento orientado definirão um e um só elemento de $\mathbb{R}^2$; ver Figura 6.3.

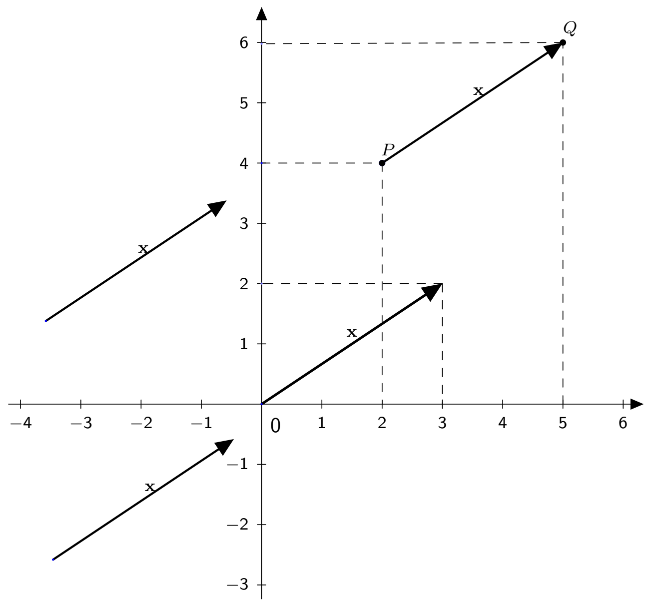

Figura 6.3: Vários representantes do vetor $\mathbf{x} = (3, 2)$, $\quad \mathbf{x} = Q - P, \quad Q = P + \mathbf{x}$

Dados dois pontos do plano , $P = (p_1, p_2)$ e $Q = (q_1, q_2)$, o vetor $\mathbf{x}$ que corresponde ao segmento orientado $[P, Q]$ tem como componentes $(q_1 - p_1, q_2 - p_2)$ (já que são essas, precisamente, as coordenadas do ponto extremidade do segmento orientado que se obtém quando transladamos o segmento orientado $[P, Q]$ para a origem); ver Figura 6.3. É nesse

[3] O único vetor que não tem direção e e sentido definidos é o *vetor nulo*, correspondente ao segmento orientado nulo; este vetor fica, no entanto, totalmente caraterizado pelo seu comprimento, igual a zero.

sentido que passamos a identificar a diferença de dois pontos Q e P do plano com um vetor do plano, e que escrevemos $\mathbf{x} = Q - P$.

Sendo a diferença de dois pontos do plano um vetor, faz sentido também falar na soma de um ponto com um vetor. Ao escrevermos $Q = P + \mathbf{x}$, queremos dizer que o ponto Q é o ponto extremidade do segmento orientado que representa $\mathbf{x}$ e que tem como origem o ponto P; ver novamente Figura 6.3.

• Adição de vetores

Dados dois vetores do plano, $\mathbf{x} = (x_1, x_2)$ e $\mathbf{y} = (y_1, y_2)$, a *soma* de $\mathbf{x}$ com $\mathbf{y}$, denotada por $\mathbf{x} + \mathbf{y}$, é definida como o vetor que se obtém somando as componentes respetivas de $\mathbf{x}$ e $\mathbf{y}$, i.e.

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2).$$

Geometricamente, facilmente se verifica que a soma de dois vetores pode obter-se pela conhecida *regra do paralelogramo*, ilustrada na Figura 6.4.

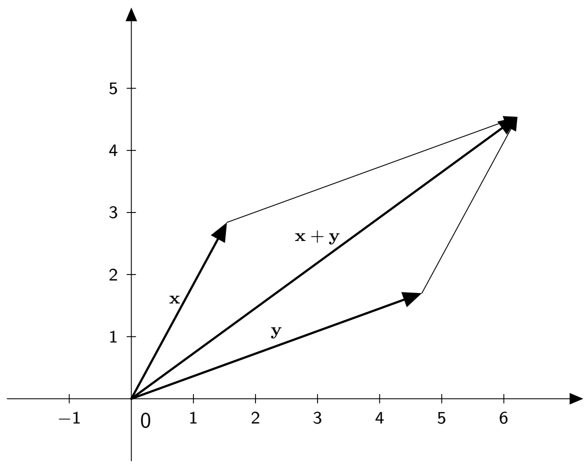

Figura 6.4: Adição de vetores no plano

• Simétrico de um vetor

Dado um vetor $\mathbf{x} = (x_1, x_2)$, o *simétrico* de $\mathbf{x}$, denotado por $-\mathbf{x}$, é o vetor obtido substituindo as componentes de $\mathbf{x}$ pelos seus simétricos, i.e.,

$$-\mathbf{x} = (-x_1, -x_2).$$

O simétrico de um vetor $\mathbf{x}$ tem a mesma direção e o mesmo comprimento de $\mathbf{x}$, mas sentido contrário, isto é, se $\mathbf{x}$ é representado por um segmento orientado $[P, Q]$,então $-\mathbf{x}$ será representado por $[Q, P]$.

A soma de um vetor $\mathbf{x}$ com o seu simétrico resulta no vetor nulo.

Por diferença de dois vetores $\mathbf{x}$ e $\mathbf{y}$, $\mathbf{x}-\mathbf{y}$, entende-se, simplesmente, a soma de $\mathbf{x}$ com o simétrico de $\mathbf{y}$, i.e.

$$\mathbf{x}-\mathbf{y}=\mathbf{x}+(-\mathbf{y}).$$

Uma vez que $\mathbf{y}+(\mathbf{x}-\mathbf{y})=\mathbf{x}$, isto é, que o vetor $\mathbf{x}-\mathbf{y}$ é tal que a sua soma com $\mathbf{y}$ é o vetor $\mathbf{x}$, então, geometricamente, a diferença de vetores pode ser obtida simplesmente *completando o triângulo*; veja a Figura 6.5.

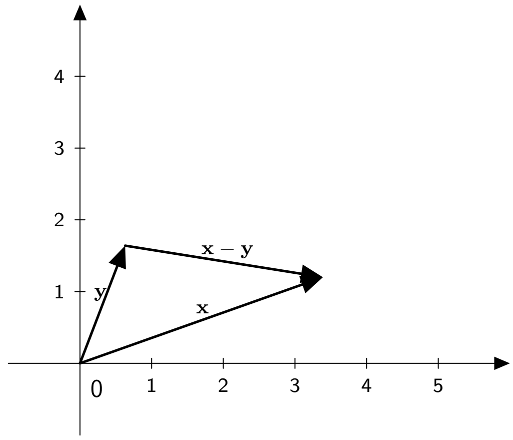

Figura 6.5: Diferença de vetores no plano

• Produto de um escalar por um vetor

Dado um vetor $\mathbf{x}=(x_1,x_2)$ e um número $\alpha\in\mathbb{R}$ (a que nos referiremos, como já vem sendo habitual neste curso, como um *escalar*), define-se o produto de α por $\mathbf{x}$ e denota-se por $\alpha\mathbf{x}$ como sendo o vetor que se obtém de $\mathbf{x}$ multiplicando as suas componentes por α, isto é

$$\alpha\mathbf{x}=(\alpha x_1,\alpha x_2).$$

Geometricamente, o produto do escalar $\alpha\neq 0$ pelo vetor $\mathbf{x}\neq\mathbf{0}$ é um vetor que:

- tem a mesma direção de $\mathbf{x}$;
- tem o mesmo sentido de $\mathbf{x}$, se $\alpha>0$, e sentido contrário, se $\alpha<0$;
- tem um comprimento igual ao produto de $|\alpha|$ pelo comprimento de $\mathbf{x}$.

O produto do escalar $\alpha=0$ por um qualquer vetor $\mathbf{x}$ é o vetor nulo e o produto de qualquer esclar α pelo vetor nulo é também o vetor nulo.

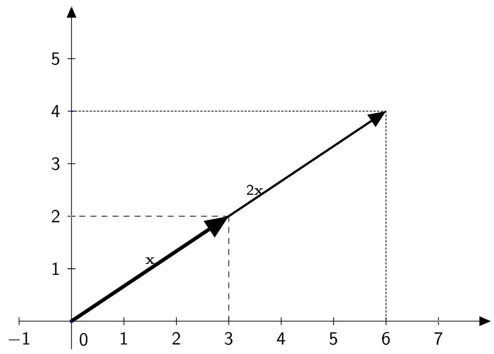

Figura 6.6: Produto de um escalar por um vetor

O conjunto $\mathbb{R}^2$, quando nele consideramos a adição e multiplicação escalar atrás definidas, chama-se espaço Euclideano de dimensão 2 (ou a duas dimensões).

• **Norma (Euclideana) de um vetor**

Dado um vetor $\mathbf{x} = (x_1, x_2)$, a *norma* (Euclideana) de $\mathbf{x}$, denotada por $\|\mathbf{x}\|$, é a definida por

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}.$$

Tendo em atenção o Teorema de Pitágoras, facilmente se vê que a norma do vetor $\mathbf{x} = (x_1, x_2)$ é simplesmente o comprimento desse vetor; veja a Figura 6.27.

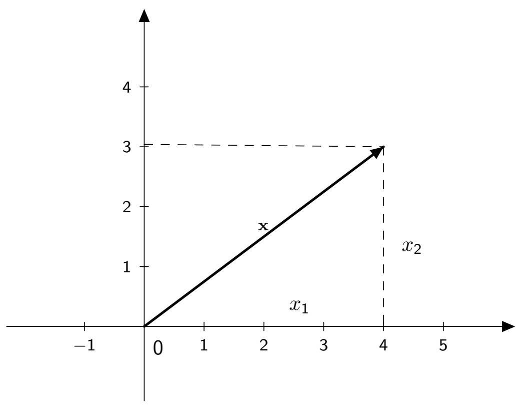

Figura 6.7: $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}$

• **Distância entre dois pontos**

Dados dois pontos $P = (p_1, p_2)$ e $Q = (q_1, q_2)$ do plano, a distância entre esses pontos é o comprimento do vetor $Q - P$, ou seja, do vetor $\mathbf{x} = (q_1 - p_1, q_2 - p_2)$. Assim, temos que a distância entre os pontos $P = (p_1, p_2)$ e $Q = (q_1, q_2)$ é dada por

$$distR2d(P, Q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2} \tag{6.1}$$

• Produto interno (ou escalar) de dois vetores

Dados dois vetores $\mathbf{x} = (x_1, x_2)$ e $\mathbf{y} = (y_1, y_2)$, o *produto interno* (usual ou Euclideano) desses vetores, denotado por $\mathbf{x} \cdot \mathbf{y}$,[4] é o número dado por

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2. \tag{6.2}$$

O produto interno de dois vetores também é chamado *produto escalar* desses vetores, para acentuar que o resultado é um escalar (i.e. um número real).

O produto interno está relacionado com a norma de um vetor: sendo $\mathbf{x} = (x_1, x_2)$ um vetor de $\mathbb{R}^2$, tem-se

$$\mathbf{x} \cdot \mathbf{x} = x_1^2 + x_2^2 = \|x\|^2$$

de onde se conclui de imediato que

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}. \tag{6.3}$$

• Propriedades do produto interno

Sendo $\mathbf{x}, \mathbf{y}, \mathbf{z}$ vetores quaisquer de $\mathbb{R}^2$ e sendo α um escalar qualquer, tem-se:

P1 $\mathbf{x} \cdot \mathbf{x} \geq 0$ e $\mathbf{x} \cdot \mathbf{x} = \mathbf{0} \iff \mathbf{x} = \mathbf{0}$

P2 $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$

P3 $(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{x} \cdot \mathbf{z}$ e $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$

P4 $\alpha(\mathbf{x} \cdot \mathbf{y}) = (\alpha\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\alpha\mathbf{y})$

• Ângulo de dois vetores

Dados dois vetores $\mathbf{x}$ e $\mathbf{y}$ em $\mathbb{R}^2$, não nulos, entende-se por *ângulo entre esses dois vetores* o ângulo convexo definido pelos dois segmentos orientados que representam esses vetores e que têm origem na origem do sistema de eixos.[5] Note-se que falamos de ângulo no sentido da sua medida e que não estamos a considerar ângulos orientados, ou seja, consideramos o ângulo entre $\mathbf{x}$ e $\mathbf{y}$ como o mesmo que o ângulo entre $\mathbf{y}$ e $\mathbf{x}$; ver Figura 6.11.

---

[4] Outra notação usual é $\langle \mathbf{x}, \mathbf{y} \rangle$.

[5] Ou em qualquer outro ponto do plano, desde que ambos tenham a mesma origem.

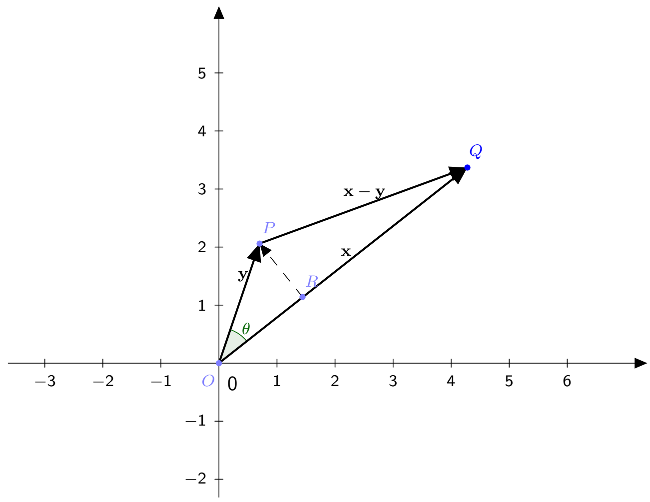

Figura 6.8: Ângulo de dois vetores

O chamado *Teorema dos cossenos* aplicado ao triângulo de vértices O, P e Q diz-nos que

$$\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta.$$

Mas,

$$\begin{aligned}
&\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta \\
&\qquad \iff (x_1-y_1)^2 + (x_2-y_2)^2 = x_1^2 + x_2^2 + y_1^2 + y_2^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta \\
&\qquad \iff x_1^2 + y_1^2 - 2x_1y_1 + x_2^2 + y_2^2 - 2x_2y_2 = x_1^2 + x_2^2 + y_1^2 + y_2^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta \\
&\qquad \iff -2x_1y_1 - 2x_2y_2 = -2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta \\
&\qquad \iff x_1y_1 + x_2y_2 = \|\mathbf{x}\|\|\mathbf{y}\|\cos\theta \\
&\qquad \iff \cos\theta = \frac{x_1y_1 + x_2y_2}{\|\mathbf{x}\|\|\mathbf{y}\|}
\end{aligned}$$

Temos, assim

$$\cos\theta = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}, \tag{6.4}$$

de onde se pode obter o valor de θ. Então, dois vetores não nulos $\mathbf{x}$ e $\mathbf{y}$ serão perpendiculares (i.e. formarão um ângulo θ igual a $\pi/2$) se e só se tivermos $\mathbf{x}\cdot\mathbf{y} = 0$. Quando $\mathbf{x} = \mathbf{0}$ ou $\mathbf{y} = \mathbf{0}$, ter-se-á também $\mathbf{x}\cdot\mathbf{y} = 0$. Em conclusão, temos:

$$\mathbf{x}\cdot\mathbf{y} = 0 \iff \mathbf{x} = \mathbf{0} \text{ ou } \mathbf{y} = \mathbf{0} \text{ ou } \mathbf{x} \text{ e } \mathbf{y} \text{ são perpendiculares.}$$

Em qualquer dos casos anteriores, dizemos que $\mathbf{x}$e $\mathbf{y}$ são *ortogonais*, e escrevemos $\mathbf{x}\perp\mathbf{y}$. Temos, pois

$$\mathbf{x}\perp\mathbf{y} \iff \mathbf{x}\cdot\mathbf{y} = 0. \tag{6.5}$$

6.2 O espaço $\mathbb{R}^3$

O conjunto de todos os ternos ordenados de números reais é denotado por $\mathbb{R}^3$, isto é, tem-se

$$\mathbb{R}^3 = \big\{(x_1, x_2, x_3) : x_1, x_2, x_3 \in \mathbb{R}\big\}.$$

De modo totalmente análogo ao que fizemos para $\mathbb{R}^2$, podemos identificar $\mathbb{R}^3$ com o conjunto dos pontos do espaço usual, no qual se fixou um sistema de três eixos cartesianos intersetando-se num mesmo ponto O (por simplicidade, escolhidos como perpendiculares 2 a 2, e com a mesma medida de comprimento em todos eles). Também podemos identificar $\mathbb{R}^3$ com o conjunto dos segmentos orientados no espaço com origem no ponto O ou ainda com o conjunto de todos os vetores livres do espaço (definidos por uma direção, sentido e comprimento).

A adição de vetores em $\mathbb{R}^3$ e a multiplicação de um escalar por um vetor de $\mathbb{R}^3$ são definidas de modo análogo ao que fizemos em $\mathbb{R}^2$. Mais precisamente, sendo $\mathbf{x} = (x_1, x_2, x_3)$ e $\mathbf{y} = (y_1, y_2, y_3)$ dois vetores arbitrários de $\mathbb{R}^3$, a soma de $\mathbf{x}$ com $\mathbf{y}$ é o vetor definido por

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$$

e o produto do escalar $\alpha \in \mathbb{R}$ pelo vetor $\mathbf{x}$ é o vetor denotado por $\alpha\mathbf{x}$ e dado por

$$\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \alpha x_3).$$

A adição de vetores em $\mathbb{R}^3$ e a multiplicação de números reais por vetores de $\mathbb{R}^3$ tem interpretações geométricas análogas às de $\mathbb{R}^2$.

O conjunto $\mathbb{R}^3$ com a adição e multiplicação por um escalar acima definidas é chamado *espaço Euclideano de dimensão 3* (ou a três dimensões).

Também de modo totalmente análogo ao que é feito para $\mathbb{R}^2$, se define o simétrico de um vetor $\mathbf{x} = (x_1, x_2, x_3)$ de $\mathbb{R}^3$, como sendo o vetor denotado $-\mathbf{x}$ e dado por

$$-\mathbf{x} = (-x_1, -x_2, -x_3),$$

sendo o significado geométrico de $-\mathbf{x}$ idêntico ao do caso $\mathbb{R}^2$.

A norma (Euclideana) de um vetor $\mathbf{x} = (x_1, x_2, x_3)$ de $\mathbb{R}^3$ é definida por

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2} \tag{6.6}$$

e a sua interpretação como comprimento do vetor $\mathbf{x}$ é também válida neste caso.

A distância entre dois pontos $P = (p_1, p_2, p_3)$ e $Q = (q_1, q_2, q_3)$ é dada por

$$d(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2} \tag{6.7}$$

O produto interno (Euclideano) ou produto escalar de dois vetores $\mathbf{x} = (x_1, x_2, x_3)$ e $\mathbf{y} = (y_1, y_2, y_3)$, denotado por $\mathbf{x} \cdot \mathbf{y}$ é definido por

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 \tag{6.8}$$

e o ângulo θ entre dois vetores não nulos $\mathbf{x} = (x_1, x_2, x_3)$ e $\mathbf{y} = (y_1, y_2, y_3)$ pode ser obtido usando a fórmula

$$\theta = \text{arccos}\, \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}, \theta \in [0, \pi]. \tag{6.9}$$

O produto interno em $\mathbb{R}^3$ goza das mesmas propiedade P1-P4 enunciadas para $\mathbb{R}^2$. Além disso, tem-se também

$$\mathbf{x} \perp \mathbf{y} \iff \mathbf{x} \cdot \mathbf{y} = 0. \tag{6.10}$$

6.2.1 Produto vetorial em $\mathbb{R}^3$

Em $\mathbb{R}^3$ (não havendo análogo em $\mathbb{R}^2$) pode definir-se outro tipo de produto entre vetores, o chamado *produto vetorial*, que tem bastantes aplicações.

Definição 6.1 (Produto vetorial)**.** Dados $\mathbf{x} = (x_1, x_2, x_3)$ e $\mathbf{y} = (y_1, y_2, y_3)$ em $\mathbb{R}^3$, o *produto vetorial* desses dois vetores, denotado por $\mathbf{x} \times \mathbf{y}$[6] é o **vetor** dado por

$$\mathbf{x} \times \mathbf{y} = \big(x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1\big). \tag{6.11}$$

A expressão das componentes do vetor $\mathbf{x} \times \mathbf{y}$ será fácil de fixar se procedermos do seguinte modo. Considerando os vetores

$$\mathbf{e}_1 = (1, 0, 0),\ \mathbf{e}_2 = (0, 1, 0) \ \text{ e } \ \mathbf{e}_3 = (0, 0, 1),$$

qualquer vetor $\mathbf{x} = (x_1, x_2, x_3)$ de $\mathbb{R}^3$ se pode escrever como combinação linear desses vetores:

$$\mathbf{x} = (x_1, x_2, x_3) = x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3.$$

---

[6] Também denotado, frequentemente, por $\mathbf{x} \wedge \mathbf{y}$.

Temos, assim, duas maneira distintas de representar o mesmo vetor $\mathbf{x}$: (x_1, x_2, x_3) ou $x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$. [7] De modo análogo, temos

$$\mathbf{y} = (y_1, y_2, y_3) = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3.$$

Usando este tipo de notação, é fácil de ver que o vetor $\mathbf{x} \times \mathbf{y}$ é o vetor que se obtém quando se desenvolve o seguinte determinante (simbólico)

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix},$$

usando o Teorema de Laplace, ao longo da sua primeira linha. Com efeito, tem-se

$$\begin{aligned} \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} &= \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \mathbf{e}_3 \\ &= (x_2y_3 - x_3y_2)\mathbf{e}_1 - (x_1y_3 - x_3y_1)\mathbf{e}_2 + (x_1y_2 - x_2y_1)\mathbf{e}_3 \\ &= \big(x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1\big) \\ &= \mathbf{x} \times \mathbf{y}. \end{aligned}$$

As seguintes propriedades do produto vetorial obtêm-se facilmente usando as propriedades conhecidas para os determinantes, sendo a sua demonstração deixada como exercício.

Proposição 6.1. *Dados quaisquer vetores $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ em $\mathbb{R}^3$ e dado um qualquer escalar $\alpha \in \mathbb{R}$, tem-se:*

1. $\mathbf{x} \times \mathbf{y} = -(\mathbf{y} \times \mathbf{x})$
2. $(\alpha\mathbf{x}) \times \mathbf{y} = \mathbf{x} \times (\alpha\mathbf{y}) = \alpha(\mathbf{x} \times \mathbf{y})$
3. $(\mathbf{x} + \mathbf{y}) \times \mathbf{z} = \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z}$
4. $\mathbf{x} \times (\mathbf{y} + \mathbf{z}) = \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z}$
5. $\mathbf{x} \times \mathbf{y} = \mathbf{0} \iff \mathbf{x} = \mathbf{0}$ *ou* $\mathbf{y} = \mathbf{0}$ *ou* $\mathbf{x}$ *e* $\mathbf{y}$ *são paralelos, i.e.* $\mathbf{y} = \lambda\mathbf{x}$ *para* $\lambda \in \mathbb{R}$*; em particular,* $\mathbf{x} \times \mathbf{x} = \mathbf{0}$.

[7] Os vetores $\mathbf{e}_1, \mathbf{e}_2$ e $\mathbf{e}_3$ são muitas vezes designados por $\mathbf{i}, \mathbf{j}$ e $\mathbf{k}$, respetivamente, sendo, então, $\mathbf{x}$ escrito como $x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$.

Nota: Convém chamar a atenção de que, em geral, $(\mathbf{x} \times \mathbf{y}) \times \mathbf{z} \neq \mathbf{x} \times (\mathbf{y} \times \mathbf{z})$ ou seja: *O produto vetorial não goza da propriedade associativa.*

• Interpretação geométrica do produto vetorial

Proposição 6.2. *Sejam $\mathbf{x}, \mathbf{y}$ vetores não nulos de $\mathbb{R}^3$ e seja $\mathbf{z} = \mathbf{x} \times \mathbf{y}$. Então, tem-se:*

1. $\mathbf{z}$ *é ortogonal a $\mathbf{x}$ e a $\mathbf{y}$.*
2. *Os vetores $\mathbf{x}$, $\mathbf{y}$ e $\mathbf{z}$ formam um sistema direto, isto é, o sentido de $\mathbf{z}$ é como se indica na Figura 6.9 ; ver nota abaixo.*
3. $\|\mathbf{x} \times \mathbf{y}\|^2 = \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2$.
4. $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\|\|\mathbf{y}\| \operatorname{sen}\theta$, *onde θ é o ângulo entre $\mathbf{x}$ e $\mathbf{y}$.*

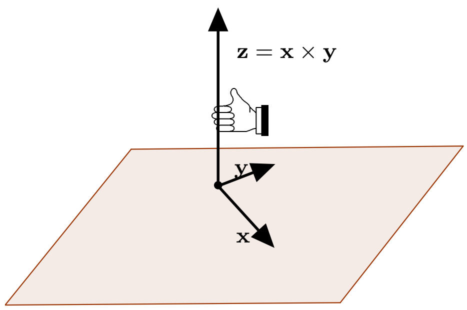

Figura 6.9: Direção e sentido $\mathbf{x} \times \mathbf{y}$

Nota: Uma regra usada para saber o sentido de $\mathbf{z} = \mathbf{x} \times \mathbf{y}$ é usualmente designada por **Regra da mão direita**: *colocando-se a mão direita com o polegar ao longo de $\mathbf{z}$ e apontando no sentido de $\mathbf{z}$, os restantes dedos, ao fecharem, rodarão de $\mathbf{x}$ para $\mathbf{y}$.*

Dem:

1. Temos

$$\begin{aligned}(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x} &= (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) \cdot (x_1, x_2, x_3) \\ &= (x_2y_3 - x_3y_2)x_1 + (x_3y_1 - x_1y_3)x_2 + (x_1y_2 - x_2y_1)x_3 \\ &= x_1x_2y_3 - x_1x_3y_2 + x_2x_3y_1 - x_1x_2y_3 + x_1x_3y_2 - x_2x_3y_1 = 0.\end{aligned}$$

De modo análogo se mostra que $(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{y} = 0$.

2. Não demonstraremos esta propriedade.

3. Temos
$$\|\mathbf{x}\times\mathbf{y}\|^2 = (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - y_1x_2)^2$$
e
$$\|\mathbf{x}\|^2\|\mathbf{y}\|^2 - (\mathbf{x}\cdot\mathbf{y})^2 = (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2.$$
Para finalizar a demonstração, teremos apenas de "desenvolver"os produtos nos lados direitos das duas fórmulas anteriores e mostrar que eles dão o mesmo resultado.

4. Temos
$$\begin{aligned}\|\mathbf{x}\times\mathbf{y}\|^2 &= \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - (\mathbf{x}\cdot\mathbf{y})^2\\ &= \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - (\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta)^2\\ &= \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - \|\mathbf{x}\|^2\|\mathbf{y}\|^2\cos^2\theta\\ &= \|\mathbf{x}\|^2\|\mathbf{y}\|^2\left(1-\cos^2\theta\right)\\ &= \|\mathbf{x}\|^2\|\mathbf{y}\|^2\operatorname{sen}^2\theta\end{aligned}$$
Como $0 \leq \theta \leq \pi$, tem-se $\operatorname{sen}\theta \geq 0$ e, portanto, a igualdade acima é equivalente a
$$\|\mathbf{x}\times\mathbf{y}\| = \|\mathbf{x}\|\|\mathbf{y}\|\operatorname{sen}\theta.$$

Esta última propriedade pode expressar-se do seguinte modo: *O comprimento do vetor* $\mathbf{x}\times\mathbf{y}$ *é igual à área do paralelogramo definido pelos dois vetores* $\mathbf{x}$ *e* $\mathbf{y}$. Observe a Figura 6.10 e note que que $\|\mathbf{y}\|\operatorname{sen}\theta$ é igual à altura h do paralelogramo.

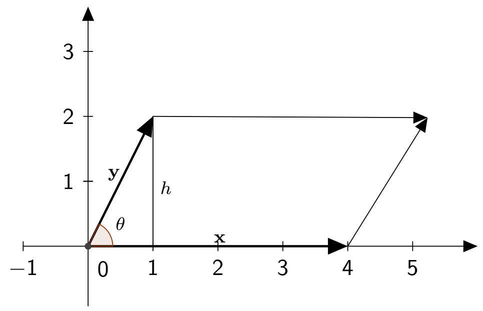

Figura 6.10: Interpretação de $\|\mathbf{x}\times\mathbf{y}\|$

6.2.2 Retas no espaço tridimensional

• Reta definida por um ponto e uma direção

Recordemos que uma reta no espaço $\mathbb{R}^3$ fica totalmente definida por um ponto $P_0 = (x_0, y_0, z_0)$ e por uma direção – a de um certo vetor $\mathbf{v} = (v_1, v_2, v_3)$, não nulo.

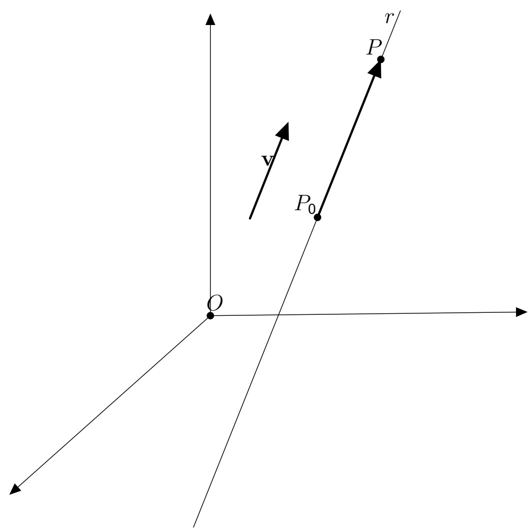

Figura 6.11: Reta que passa por P_0 e tem a direção de $\mathbf{v}$

Um ponto $P = (x, y, z)$[8] estará sobre a reta r que passa por P e tem a direção de $\mathbf{v}$ se e só se tivermos

$$P - P_0 = \lambda \mathbf{v}, \text{ para } \lambda \in \mathbb{R}. \tag{6.12}$$

Esta equação pode reescrever-se como

$$P = P_0 + \lambda \mathbf{v}, \text{ para } \lambda \in \mathbb{R}, \tag{6.13}$$

ou ainda como

$$(x, y, z) = (x_0, y_0, z_0) + \lambda(v_1, v_2, v_3) \text{ para } \lambda \in \mathbb{R}. \tag{6.14}$$

A qualquer uma das formas (6.12) – (6.14) damos o nome de *equação vetorial* da reta r.

O vetor $\mathbf{v}$ é chamado *vetor diretor* da reta r e as suas coordenadas, v_1, v_2, v_3, (ou as de qualquer outro vetor que tenha a mesma direção) dizem-se *parâmetros directores* de r. Os cossenos dos ângulos α, β e γ que $\mathbf{v}$ forma com $\mathbf{e}_1 = (1, 0, 0), \mathbf{e}_2 = (0, 1, 0)$ e $\mathbf{e}_3 = (0, 0, 1)$, respetivamente, recebem o nome de *cossenos directores* da recta r.

Da equação vetorial, obtêm-se, de imediato, as chamadas *equações paramétricas* da reta:

$$\begin{cases} x = x_0 + \lambda\, v_1 \\ y = y_0 + \lambda\, v_2 \\ z = z_0 + \lambda\, v_3 \end{cases}, \quad \lambda \in \mathbb{R}. \tag{6.15}$$

Eliminado o parâmetro λ das equações paramétricas, obtemos as chamadas *equações normais* da reta:

$$\frac{x - x_0}{v_1} = \frac{y - y_0}{v_2} = \frac{z - z_0}{v_3} \tag{6.16}$$

[8]Preferimos, aqui, usar a notação mais usual (x, y, z) para pontos em $\mathbb{R}^3$, em vez da notação que temos vindo a adotar de (x_1, x_2, x_3), por ser esta mais familiar para os alunos.

.

Nota: Na equação anterior, se algum dos denominadores for igual a zero, deve entender-se que o respetivo numerador também é zero; isto é, por exemplo

$$\frac{x-1}{2} = \frac{y-3}{0} = \frac{z-2}{4}$$

deve ser interpretado como

$$\frac{x-1}{2} = \frac{z-2}{4} \text{ e } y = 3.$$

Exemplo 6.1. Seja r a reta que passa pelo ponto $P_0 = (1, 2, -1)$ e tem a direção de $\mathbf{v} = (3, -1, 5)$.

- A equação vetorial de r é : $(x, y, z) = (1, 2, -1) + \lambda(3, -1, 5), \lambda \in \mathbb{R}$.

- As equações paramétricas de r são: $\begin{cases} x = 1 + 3\lambda \\ y = 2 - \lambda \\ z = -1 + 5\lambda \end{cases}, \ \lambda \in \mathbb{R}$

- As equações normais de r são: $\dfrac{x-1}{3} = 2 - y = \dfrac{z+1}{5}$.

• Reta definida por dois pontos

Dados os pontos $P_0 = (x_0, y_0, z_0)$ e $P_1 = (x_1, y_1, z_1)$, o problema da determinação da equação da reta definida por esses pontos reduz-se facilmente ao caso anterior, tomando para $\mathbf{v}$ o vector $\mathbf{v} = P_1 - P_0$.

6.2.3 Planos em $\mathbb{R}^3$

• Plano definido por um ponto e (perpendicular a) uma direção

Sejam dados um ponto $P_0 = (x_0, y_0, z_0)$ e um vector não nulo $\mathbf{v} = (a, b, c)$ e suponhamos que se pretende determinar a equação do plano α que passa por P_0 e é perpendicular ao vector $\mathbf{v}$; ver Fig. 6.12.

Um ponto $P(x, y, z)$ pertencerá ao plano α se e só se $(P - P_0)$ for ortogonal a $\mathbf{v}$, i.e. se e só se tivermos

$$(P - P_0) \cdot \mathbf{v} = 0. \tag{6.17}$$

A equação anterior é designada por *equação vetorial do plano que passa por P_0 e é ortogonal a $\mathbf{v}$*. Dessa equação, obtém-se, de imediato

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

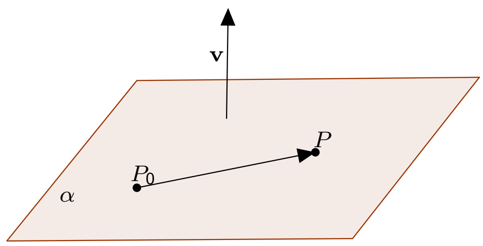

Figura 6.12: Plano que passa por P_0 e é perpendicular a $\mathbf{v}$

ou

$$ax + by + cz - (ax_0 + by_0 + cz_0) = 0.$$

A equação anterior pode ser reescrita como

$$ax + by + cz = d, \tag{6.18}$$

onde a, b, c são os parâmetros diretores de uma reta perpendicular ao plano – reta essa que se chama *eixo do plano* – e onde $d = ax_0 + by_0 + cz_0$. Tal equação é conhecida como *equação cartesiana* ou *equação geral* do plano α.

Reciprocamente, é fácil de verificar que toda a equação da forma $ax + by + cz = d$ é a equação de um plano perpendicular ao vector $\mathbf{v} = (a, b, c)$, isto é, é a equação de um plano cujo eixo tem (a, b, c) como parâmetros diretores.

- **Plano definido por um ponto e duas direções**

Dados um ponto P_0 e duas direções distintas – definidas à custa de dois vetores $\mathbf{x}$ e $\mathbf{y}$ não paralelos – para determinarmos a equação do plano que contém P_0 e é paralelo a $\mathbf{x}$ e a $\mathbf{y}$, bastará determinar o vetor $\mathbf{v} = \mathbf{x} \times \mathbf{y}$, o qual será ortogonal ao plano pretendido, e caímos, então, no caso anteriormente estudado: plano defindo por um ponto e perpendicular a uma direção.

Em alternativa, poderemos determinar outras formas da equação do plano.

Um ponto $P = (x, y, z)$ estará no plano que passa por $P_0 = (x_0, y_0, z_0)$ e é paralelo aos vetores (não paralelos) $\mathbf{u} = (u_1, u_2, u_3)$ e $\mathbf{v} = (v_1, v_2, v_3)$ se e só se tivermos (veja a Figura 6.2.3):

$$P - P_0 = \lambda \mathbf{u} + \mu \mathbf{v}, \quad \lambda, \mu \in \mathbb{R}, \tag{6.19}$$

ou seja, se e só se tivermos

$$\begin{cases} x = x_0 + \lambda u_1 + \mu v_1 \\ y = y_0 + \lambda u_2 + \mu v_2 \\ z = z_0 + \lambda u_3 + \mu v_3 \end{cases}, \quad \lambda, \mu \in \mathbb{R}. \tag{6.20}$$

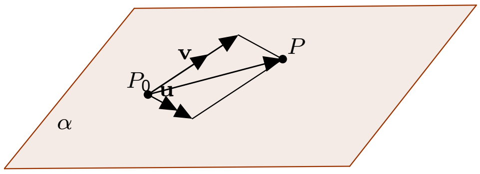

Figura 6.13: Plano definido por um ponto e duas direções

A equação (6.19) é a *equação vetorial do plano definido por um ponto* P_0 *e duas direções* $\mathbf{u}$ *e* $\mathbf{v}$ e (6.20) são as suas *equações paramétricas.*

Exemplo 6.2. Seja α o plano que passa por $P_0 = (1, -2, 0)$ e é paralelo aos vetores $\mathbf{u} = (1, 2, 3)$ e $\mathbf{v} = (-1, 0, 1)$. Podemos calcular o vetor $\mathbf{w} = \mathbf{u} \times \mathbf{v}$, vindo

$$\mathbf{w} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 2 & 3 \\ -1 & 0 & 1 \end{vmatrix} = 2\mathbf{e}_1 - 4\mathbf{e}_2 + 2\mathbf{e}_3 = (2, -4, 2).$$

Então, temos $d = 2 \times 1 - 4 \times (-2) + 2 \times 0 = 10$, pelo que a equação cartesiana do plano α é:

$$2x - 4y + 2z = 10.$$

As equações paramétricas de α são:

$$\begin{cases} x = 1 + \lambda - \mu \\ y = -2 + 2\lambda \\ z = 3\lambda + \mu \end{cases}, \quad \lambda, \mu \in \mathbb{R}.$$

Note que, das equações paramétricas, obtemos

$$2x - 4y + 2z = 2(1 + \lambda - \mu) - 4(-2 + 2\lambda) + 2(3\lambda + \mu) = 10.$$

- **Plano definido por três pontos não colineares**

Dados três pontos A, B, C não colineares, a determinação da equação do plano que os contém reduz-se ao caso anterior, tomando, por exemplo, $P_0 = A$ e $\mathbf{x} = B - A$ e $\mathbf{y} = C - A$.

6.3 Espaço Euclideano $\mathbb{R}^n$

Os espaços Eucideanos $\mathbb{R}^2$ e $\mathbb{R}^3$ podem ser vistos como dois primeiros exemplos do espaço Euclideano $\mathbb{R}^n$ com $n \in \mathbb{N}$.

Dado $n \in \mathbb{N}$, considere-se o conjunto $\mathbb{R}^n$ de todos os n-uplos ordenados de elementos reais, isto é, o conjunto de elementos da forma $= (x_1, \ldots, x_n)$, com $x_i \in \mathbb{R}$:

$$\mathbb{R}^n = \big\{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R}; i = 1, \ldots, n\big\}. \tag{6.21}$$

Note-se que dois n-uplos $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ e $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ são considerados iguais se e só se $x_1 = y_1$, $x_2 = y_2$, $\ldots$, $x_n = y_n$, isto é, por exemplo $(1, 2, 3) \neq (2, 1, 3)$.

Dado $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, chamamos a $x_1, x_2, \ldots, x_n$ as *componentes* de $\mathbf{x}$.

Definição 6.2 (Soma de vetores de $\mathbb{R}^n$)**.** Dados dois elementos $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$ em $\mathbb{R}^n$, chamamos *soma de* $\mathbf{x}$ *e* $\mathbf{y}$ ao elemento de $\mathbb{R}^n$, que denotaremos por $\mathbf{x} + \mathbf{y}$, obtido somando as componentes correspondentes de $\mathbf{x}$ e $\mathbf{y}$, isto é:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n). \tag{6.22}$$

Definição 6.3 (Produto de um escalar por um elemento de $\mathbb{R}^n$)**.** Dados $\alpha \in \mathbb{R}$ e $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, chamamos *produto do escalar* α *por* $\mathbf{x}$ ao elemento de $\mathbb{R}^n$, que designamos por $\alpha\mathbf{x}$, obtido multiplicando todas as componentes de $\mathbf{x}$ por α, isto é:

$$\alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n). \tag{6.23}$$

Temos, assim duas operações: uma, que chamaremos de *adição*, que a cada par de elementos $\mathbf{x}$ e $\mathbf{y}$ de $\mathbb{R}^n$ associa um novo elemento de $\mathbb{R}^n$, dado pela sua soma $\mathbf{x} + \mathbf{y}$, definida por (6.22) e outra, que chamaremos de *multiplicação escalar*, que a cada escalar α e a cada elemento $\mathbf{x} \in \mathbb{R}^n$ associa o elemento obtido como produto do escalar α por $\mathbf{x}$, $\alpha\mathbf{x}$, obtido usando a definição (6.25).

Definição 6.4 (Vetor nulo de $\mathbb{R}^n$)**.** Designamos por $\mathbf{0}_n$ ou, mais simplesmente por $\mathbf{0}$, o elemento de $\mathbb{R}^n$ com todas as componentes nulas, isto é, $\mathbf{0} = (0, 0, \ldots, 0, 0)$. Este elemento é designado por *vetor nulo* de $\mathbb{R}^n$.

Definição 6.5 (Simétrico de um elemento de $\mathbb{R}^n$)**.** Sendo $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ um determinado elemento de $\mathbb{R}^n$, denotaremos por $-\mathbf{x}$ o elemento obtido substituindo cada componente de $\mathbf{x}$ pelo seu simétrico, i.e.

$$-(x_1, \ldots, x_n) = (-x_1, \ldots, -x_n). \tag{6.24}$$

Este elemento é chamado o *simétrico* de $\mathbf{x}$.

Usando as propriedades conhecidas para a adição e para a multiplicação de números reais, vemos facilmente que a adição e multiplicação escalar acima definidas gozam das seguintes propriedades:

A1 a adição é comutativa: $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.

A2 a adição é associativa: $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n, \quad (\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.

A3 o elemento $\mathbf{0} = (0, \ldots, 0)$ é *neutro* para a adição, isto é, dado qualquer $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, tem-se $(x_1, \ldots, x_n) + (0, \ldots, 0) = (0, \ldots, 0) + (x_1, \ldots, x_n) = (x_1, \ldots, x_n)$, ou seja, tem-se $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$.

A4 dado um elemento qualquer $\mathbf{x} = (x_1, \ldots, x_n)$ de $\mathbb{R}^n$, o elemento $-\mathbf{x} = (-x_1, \ldots, -x_n)$ satisfaz $(x_1, \ldots, x_n) + (-x_1, \ldots, x_n) = (-x_1, \ldots, -x_n) + (x_1, \ldots, x_n) = (0, \ldots, 0)$, ou seja, tem-se $\mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$.

MEA1 a multiplicação escalar é distributiva em relação à adição em $\mathbb{R}$: $\forall \alpha, \beta \in \mathbb{R}, \ \forall \mathbf{x} \in \mathbb{R}^n, \quad (\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$.

MEA2 a multiplicação escalar é distributiva em relação à adição em $\mathbb{R}^n$: $\forall \alpha \in \mathbb{R}, \ \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$.

ME1 a multiplicação escalar satisfaz a associatividade mista: $\forall \alpha, \beta \in \mathbb{R}, \forall \mathbf{x} \in \mathbb{R}^n, \quad \alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$.

ME2 $\forall \mathbf{x} \in \mathbb{R}^n, \quad 1\mathbf{x} = \mathbf{x}$.

O conjunto $\mathbb{R}^n$ com a adição e multiplicação por um escalar acima definidas é um exemplo de uma estrutura algébrica conhecida por *espaço vetorial* (real) e é chamado *espaço Euclideano de dimensão* n (ou a n dimensões). Os elementos de $\mathbb{R}^n$ são usualmente designados por *vetores*, podendo também ser referidos como *pontos* de $\mathbb{R}^n$.

Nota: Se tivermos em atenção a maneira como definimos a adição de vetores coluna de $\mathbb{R}^{n\times 1}$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

e a forma como definimos a adição em $\mathbb{R}^n$

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

verificamos que essas operações são totalmente idênticas: o que fazemos é apenas somar os elementos nas posições correspondentes. O mesmo se passa relativamente à multiplicação por um escalar. Vemos, assim que, no fundo $\mathbb{R}^{n\times 1}$ e $\mathbb{R}^n$ são "o mesmo"espaço, apenas diferindo na forma como apresentamos os seus elementos. Por esse motivo, identificaremos $\mathbb{R}^{n\times 1}$ com $\mathbb{R}^n$ e usaremos indistintamente a notação de vetor coluna $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ ou de n-uplo $\mathbf{x} = (x_1, \ldots, x_n)$, para um certo vetor $\mathbf{x} \in \mathbb{R}^n$, conforme seja mais conveniente.

6.3.1 Produto escalar em $\mathbb{R}^n$

Para estender as noções de distância, comprimento (norma) ângulo para o espaço $\mathbb{R}^n$, começamos por introduzir a seguinte generalização do produto escalar (Euclideano), dado para os casos $n = 2$ e $n = 3$.

Definição 6.6 (Produto escalar de dois vetores)**.** Dados dois vetores arbitrários $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$ de $\mathbb{R}^n$, o *produto escalar* (Euclideano) desses vetores, representado por $\mathbf{x} \cdot \mathbf{y}$, é o escalar definido por

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{k=1}^{n} x_k y_k. \tag{6.25}$$

Nota: Se escrevermos os vetores $\mathbf{x}$ e $\mathbf{y}$ na notação de coluna, o produto escalar $\mathbf{x} \cdot \mathbf{y}$ pode ser escrito como

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^{\mathsf{T}} \mathbf{y}. \tag{6.26}$$

O produto escalar goza das seguintes propriedades:

Proposição 6.3. *Sendo $\mathbf{x}, \mathbf{y}$ e $\mathbf{z}$ vetores arbitrários de $\mathbb{R}^n$ e α um escalar qualquer em $\mathbb{R}$, tem-se:*

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$
2. $(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$
3. $(\alpha \mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\alpha \mathbf{y}) = \alpha (\mathbf{x} \cdot \mathbf{y})$
4. $\mathbf{x} \cdot \mathbf{x} \geq 0$ *e* $\mathbf{x} \cdot \mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$.

Dem: Como exercício.

6.3.2 Norma Euclideana em $\mathbb{R}^n$

Definição 6.7 (Norma de um vetor)**.** Dado um vetor $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ de $\mathbb{R}^n$, a *norma* (Euclideana) de $\mathbf{x}$ (ou *comprimento* de $\mathbf{x}$), denotada por $\|\mathbf{x}\|$, é o número dado por

$$\|\mathbf{x}\| = (\mathbf{x} \cdot \mathbf{x})^{1/2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}. \tag{6.27}$$

O seguinte teorema, que daremos sem demonstração, é uma das desigualdades mais importante de Álgebra Linear.

Teorema 6.1 (Desigualdade de Cauchy-Schwarz)**.** *Se* $\mathbf{x} = (x_1, \ldots, x_n)$ *e* $\mathbf{y} = (y_1, \ldots, y_n)$ *são dois vetores quaisquer de* $\mathbb{R}^n$*, tem-se*

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

Na seguinte proposição estão enunciadas as propriedades da norma.

Proposição 6.4. *Dados quaisquer dois vetores* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *e sendo* $\alpha \in \mathbb{R}$ *um escalar qualquer, tem-se:*

1. $\|\mathbf{x}\| \geq 0 \quad$ *e* $\quad \|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$
2. $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

Dem:

1. É uma consequência imediata da propriedade 4. do produto interno.
2. Temos
$$\|\alpha \mathbf{x}\|^2 = (\alpha \mathbf{x})^\mathsf{T} (\alpha \mathbf{x}) = (\alpha \mathbf{x}^\mathsf{T})(\alpha \mathbf{x}) = \alpha^2 (\mathbf{x}^\mathsf{T} \mathbf{x}) = |\alpha|^2 \|\mathbf{x}\|^2.$$
O resultado segue-se de imediato, tomando a raiz quadrada de ambos os membros da igualdade anterior.
3. Temos
$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y})^\mathsf{T} (\mathbf{x} + \mathbf{y}) = (\mathbf{x}^\mathsf{T} + \mathbf{y}^\mathsf{T})(\mathbf{x} + \mathbf{y}) \\ &= \mathbf{x}^\mathsf{T}\mathbf{x} + \mathbf{x}^\mathsf{T}\mathbf{y} + \mathbf{y}^\mathsf{T}\mathbf{x} + \mathbf{y}^\mathsf{T}\mathbf{y} \\ &= \|\mathbf{x}\|^2 + 2(\mathbf{x} \cdot \mathbf{y}) + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2|\mathbf{x} \cdot \mathbf{y}| + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2, \end{aligned}$$
onde, na penúltima passagem, fizemos uso da desigualdade de Cauchy-Schwarz. Tomando a raiz quadrada de ambos os membros (e, tendo em conta que as quantidades envolvidas são não negativas), obtém-se o resultado pretendido.

$\square$

6.3.3 Ângulo entre vetores de $\mathbb{R}^n$

Definição 6.8 (Ângulo entre vetores)**.** Dados dois vetores não nulos $\mathbf{x}$ e $\mathbf{y}$ em $\mathbb{R}^n$, o *ângulo* θ entre esses vetores é definido por

$$\theta = \text{arccos}\,\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}, \quad 0 \leq \theta \leq \pi \tag{6.28}$$

Nota: A definição anterior faz sentido, uma vez que, pela desigualdade de Cauchy-Schwarz, temos sempre

$$\frac{|\mathbf{x}\cdot\mathbf{y}|}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1$$

ou seja, temos

$$-1 \leq \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1.$$

Definição 6.9 (Ortogonalidade)**.** Dados dois vetores $\mathbf{x}$ e $\mathbf{y}$ em $\mathbb{R}^n$, dizemos que eles são *ortogonais*, e escrevemos $\mathbf{x}\perp\mathbf{y}$, se e só se

$$\mathbf{x}\cdot\mathbf{y} = 0. \tag{6.29}$$

É, então, imediato reconhecer que temos

$$\mathbf{x}\perp\mathbf{y} \iff \mathbf{x} = \mathbf{0} \text{ ou } \mathbf{y} = \mathbf{0} \text{ ou } \theta = \frac{\pi}{2}, \tag{6.30}$$

onde θ designa o ângulo entre $\mathbf{x}$ e $\mathbf{y}$, se esses vetores forem ambos não nulos.

6.3.4 Distância Euclideana

Definição 6.10. Dados dois pontos $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ e $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ de $\mathbb{R}^n$, a distância (Euclideana) entre esses pontos, denotada por $d(\mathbf{x}, \mathbf{y})$, é definida por

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2}. \tag{6.31}$$

Como consequência imediata da definição da distância Euclideana e das propriedades da norma Euclideana enunciadas na Proposição 6.4, obtêm-se os resultados contidos na proposição seguinte.

Proposição 6.5. *Dados quaisquer vetores* $\mathbf{x}$, $\mathbf{y}$ *e* $\mathbf{z}$ *de* $\mathbb{R}^n$, *tem-se:*

1. $d(\mathbf{x},\mathbf{y}) \geq 0$ *e* $d(\mathbf{x},\mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$
2. $d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$
3. $d(\mathbf{x},\mathbf{z}) \leq d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z})$.

6.3.5 Retas e Hiperplanos em $\mathbb{R}^n$

Definição 6.11 (Reta em $\mathbb{R}^n$)**.** Dados um ponto $P_0 = (p_1, p_2, \ldots, p_n)$ e um vetor não nulo $\mathbf{v} = (v_1, \ldots, v_n)$ de $\mathbb{R}^n$, chama-se *reta que passa por P_0 e tem a direção de $\mathbf{v}$* ao conjunto de pontos $P = (x_1, x_2, \ldots, x_n)$ de $\mathbb{R}^n$ que satisfazem

$$P - P_0 = \lambda \mathbf{v}, \qquad \text{para } \lambda \in \mathbb{R}. \tag{6.32}$$

A equação (6.32) – chamada *equação vetorial da reta* – pode, naturalmente, ser escrita noutras formas; por exemplo, separando as diversas componentes dos dois lados da igualdade anterior, obtêm-se as chamadas *equações paramétricas da reta*:

$$\begin{cases} x_1 - p_1 = \lambda v_1 \\ x_2 - p_2 = \lambda v_2 \\ \quad\vdots \\ x_n - p_n = \lambda v_n \end{cases}, \lambda \in \mathbb{R},$$

de onde se obtêm as seguintes equações, chamadas *equações normais da reta*:

$$\frac{x_1 - p_1}{v_1} = \frac{x_2 - p_2}{v_2} = \cdots = \frac{x_n - p_n}{v_n}.$$

Definição 6.12 (Hiperplano em $\mathbb{R}^n$)**.** Dados um ponto $P_0 = (p_1, p_2, \ldots, p_n)$ e um vetor não nulo $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ de $\mathbb{R}^n$, chama-se *hiperplano que passa por P_0 e é ortogonal a $\mathbf{a}$* ao conjunto de pontos $P = (x_1, \ldots, x_n)$ de $\mathbb{R}^n$ que satisfazem

$$(P - P_0) \cdot \mathbf{v} = 0. \tag{6.33}$$

A equação (6.33) pode escrever-se como

$$a_1(x_1 - p_1) + a_2(x_2 - p_2) + \cdots + a_n(x_n - p_n) = 0,$$

ou seja, como

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = d, \tag{6.34}$$

onde $d = a_1p_1 + a_2p_2 + \cdots + a_np_n$.

Nota:

1. A definição de reta em $\mathbb{R}^n$ coincide, naturalmente, com a definição usual de reta (que passa por um ponto e é paralela a uma dada direção), nos casos $n = 2$ ou $n = 3$.[9]

2. Os hiperplanos de $\mathbb{R}^3$ são planos (definidos por um ponto e ortogonais a uma dada direção).

3. Os hiperplanos de $\mathbb{R}^2$ são retas (definidas por um ponto e ortogonais a uma dada direção).

[9] Não fizemos revisão de retas em $\mathbb{R}^2$, por ser este um assunto amplamente estudado no Ensino Secundário.

6.4 Exercícios

Exercício 6.1. Determine o valor de k para o qual os vetores $\mathbf{x} = (1, -2, 2)$ e $\mathbf{y} = (-1, 0, k)$ são ortogonais.

Exercício 6.2. Considere os três pontos $A = (1, 7, 3)$, $B = (0, 7 - 1)$ e $C = (-1, 6, 2)$ do espaço $\mathbb{R}^3$.

(a) Use o produto escalar para mostrar que os três pontos são os vértices de um triângulo retângulo.

(b) Calcule a área desse triângulo.

Exercício 6.3. Determine um vetor de comprimento igual a 1, perpendicular aos vetores $\mathbf{x} = (2, -1, 3)$ e $\mathbf{y} = (-4, 3, -5)$.

Exercício 6.4. Considere três pontos $A = (1, -1, 4)$, $B = (2, 0, 1)$ e $C = (0, 2, 3)$ em $\mathbb{R}^3$.

(a) Mostre que eles não são colineares (i.e. não estão sobre uma mesma reta).

(b) Determine um vetor perpendicular ao plano definido por esses três pontos.

(c) Calcule a área do triângulo cujos vértices são A, B e C.

Exercício 6.5. Determine a equação vetorial, as equações paramétricas e as equações normais da reta que:

(a) passa pelo ponto $P_0 = (3, 4, -5)$ e é paralela ao vetor $\mathbf{v} = (1, 2, 3)$;

(b) passa pela origem e pelo ponto $P = (-6, 3, 5)$;

(c) passa pelo ponto $P = (1, 1, 1)$ e é perpendicular ao plano XOY;

(d) passa pelo ponto $P = (1, 2, 0)$ e é paralela à reta de equações normais

$$\frac{x+1}{7} = \frac{2-y}{3} = \frac{z}{2}.$$

Exercício 6.6. Determine os pontos onde a reta que passa pelos pontos $A = (-6, 6, -4)$ e $B = (12, -6, 2)$ interseta os planos coordenados.

Exercício 6.7. Determine uma equação do plano que:

(a) passa pelo ponto $P = (1, 3, 1)$ e é perpendicular ao vetor $\mathbf{a} = (2, 1, -1)$;

(b) passa pelo ponto $P = (1, -1, 5)$ e é paralelo ao plano de equação $2x - 3y + 4z = 0$;

(c) passa pela origem e pelos pontos $P = (1, 1, 1)$ e $Q = (1, -1, 3)$;

(d) passa pelo ponto $P = (2, 4, 6)$ e contém a reta r de equações paramétricas

$$\begin{cases} x = 7 - 3\lambda \\ y = 3 + 4\lambda \\ z = 5 + 2\lambda \end{cases}, \quad \lambda \in \mathbb{R}.$$

Exercício 6.8. Considere o plano α de equação $2x - y + z = 0$ e o ponto $P = (3, 1, 5)$.

(a) Verifique que o ponto P não pertence ao plano α.

(b) Determine a equação da reta r que passa por P e é perpendicular a α.

(c) Determine as coordenadas do ponto de interseção da reta r (determinada em b)) com o plano α.

(d) Calcule a distância do ponto P ao plano α.

Exercício 6.9. Determine a equação do plano que passa pelo ponto $P = (1, 3, -2)$ e contém a reta de interseção dos planos $x + y + z = 5$ e $3x - y = 4$.

Exercício 6.10. Mostre que a reta de interseção dos planos $x + y - z = 0$ e $x - y - 5z = -7$ é paralela à reta

$$\frac{x+3}{3} = \frac{y-1}{-2} = z - 5.$$

Exercício 6.11. Determine se os seguintes vetores de $\mathbb{R}^4$ são ortogonais:

(a) $\mathbf{x} = (-4, -6, -10, 1)$ e $\mathbf{y} = (2, 1, -2, 9)$

(b) $\mathbf{x} = (0, 3, -2, 1)$ e $\mathbf{y} = (5, 2, -1, 0)$

Exercício 6.12. Sejam $\mathbf{x}$ e $\mathbf{y}$ dois vetores em $\mathbb{R}^n$, ortogonais.

(a) Mostre que

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

Que resultado conhecido é este, no caso $n = 2$?

(b) Mostre que

$$\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x} - \mathbf{y}\|.$$

Interprete este resultado geometricamente, no caso $n = 2$.

Exercício 6.13. (a) Determine a equação do hiperplano de $\mathbb{R}^2$ que passa pelo ponto $P_0 = (1, 1)$ e é ortogonal ao vetor $\mathbf{a} = (2, 3)$.

(b) Reescreva a equação anterior na forma mais usual $x_2 = mx_1 + b$ de uma reta em $\mathbb{R}^2$ (correspondente à forma $y = mx + b$, se designarmos por (x, y), em vez de (x_1, x_2), as coordenadas de um ponto genérico dessa reta). Qual é o declive dessa reta?

(c) Determine a tangente do ângulo que o vetor $\mathbf{a} = (2, 3)$ faz com o semieixo positivo OX (ou seja, o declive de uma reta paralela ao vetor $\mathbf{a}$). Que relação tem este valor com o declive da reta obtida na alínea anterior?

Exercício 6.14. Determine a equação do hiperplano de $\mathbb{R}^4$ que passa pelo ponto $P_0 = (1, -2, 1, 3)$ e é ortogonal ao vetor $\mathbf{a} = (2, 3, 1, 4)$.

Exercício 6.15. Considere a reta r de $\mathbb{R}^4$ que passa pelo ponto $P_0 = (1, 2, 3, 4)$ e tem a direção do vetor $\mathbf{v} = (5, 6, 7, 8)$. Diga se o ponto $Q = (11, 14, 17, 18)$ pertence a r.