notas para a unidade curricular

Álgebra Linear CC

Carla Mendes

Departamento de Matemática
Universidade do Minho
2025/2026

Conteúdo

Introdução

O presente texto pretende ser um texto de apoio às aulas da unidade curricular Álgebra Linear CC.

1. Matrizes

1.1 Conceitos básicos

Neste capítulo introduz-se o conceito de matriz e estudam-se operações e propriedades relacionadas com matrizes. As matrizes são um objeto central no estudo de álgebra linear e são bastantes os contextos em áreas da matemática e suas aplicações em que o conceito de matriz se revelou ser fundamental. Por exemplo, para a representação e tratamento de informação que esteja dependente de parâmetros é frequente o recurso a quadros (tabelas) aos quais se dá a designação de matrizes.

Ao longo deste capítulo designamos por $\mathbb{K}$ o conjunto dos números reais ou o conjunto dos números complexos; quando necessário indicaremos explicitamente se nos referimos ao conjunto $\mathbb{R}$ dos números reais ou ao conjunto $\mathbb{C}$ dos números complexos. Aos elementos de $\mathbb{K}$ damos a designação de **escalares**.

Definição 1.1.1. *Sejam $m, n \in \mathbb{N}$. Chama-se* ***matriz do tipo*** $m \times n$ *(ou* ***de ordem*** $m \times n$*)* ***sobre*** $\mathbb{K}$ *a uma função $A : \{1, \ldots, m\} \times \{1, \ldots, n\} \to \mathbb{K}$ definida por $A(i, j) = a_{ij}$ e que se representa por um quadro em que os mn elementos a_{ij} são dispostos em m filas horizontais e n filas verticais do seguinte modo*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & a_{1\,n} \\ a_{21} & a_{22} & \cdots & a_{2\,n-1} & a_{2\,n} \\ a_{31} & a_{32} & \cdots & a_{3\,n-1} & a_{3\,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m-1\,1} & a_{m-1\,2} & \cdots & a_{m-1\,n-1} & a_{m-1\,n} \\ a_{m\,1} & a_{m\,2} & \cdots & a_{m\,n-1} & a_{m\,n} \end{bmatrix}.$$

Para cada $i \in \{1, \ldots, m\}$, chama-se **linha** i **da matriz** A *ao elemento* $(a_{i1}, \ldots, a_{in})$ *de* $\mathbb{K}^n$.
Para cada $j \in \{1, \ldots, n\}$, chama-se **coluna** j **da matriz** A *ao elemento* $(a_{1j}, \ldots, a_{mj})$ *de* $\mathbb{K}^m$.
Ao elemento a_{ij} de $\mathbb{K}$, $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$, chama-se **entrada** (i, j) *ou* **elemento da posição** (i, j) *da matriz A. Por vezes, representa-se a entrada (i, j) da matriz A por A_{ij}.*

Notação e terminologia:

- O conjunto das matrizes do tipo $m \times n$ sobre $\mathbb{K}$ representa-se por $\mathcal{M}_{m\times n}(\mathbb{K})$.
- O conjunto de todas as matrizes sobre $\mathbb{K}$ é representado por $\mathcal{M}(\mathbb{K})$.
- Em geral, representaremos as matrizes por letras maiúsculas e as suas entradas pela mesma letra, minúscula ou maiúscula, com índices que indicam a respetiva posição na matriz. Havendo ambiguidade na identificação da posição da matriz, coloca-se uma virgula a separar o índice da linha e o índice da coluna. Por exemplo, escreveremos $a_{2,34}$ ou $A_{2,34}$ para indicar o elemento na linha 2 e coluna 34 da matriz A.
- Se
$$A = \begin{bmatrix} a_{11} & \dots & a_{1\,n} \\ \vdots & \dots & \vdots \\ a_{m\,1} & \dots & a_{m\,n} \end{bmatrix} \in \mathcal{M}_{m\times n}(\mathbb{K}),$$
representaremos a matriz A abreviadamente de uma das seguintes formas: $A = [a_{ij}]_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$; $A = [a_{ij}]_{m\times n}$; $A = [a_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$. Quando o tipo da matriz for claro pelo contexto ou se não for importante para o estudo em questão, podemos escrever simplesmente $A = [a_{ij}]$.
- Uma matriz diz-se **real** ou **complexa** consoante os seus elementos sejam reais ou complexos.

Exemplo 1.1.2. *No lançamento de um objeto foram registados, relativamente à altura e distância atingidos pelo objeto, os valores descritos na tabela seguinte:*

$$\begin{array}{l} \textit{altura} \\ \textit{distancia} \end{array} \quad \begin{bmatrix} 100 & 200 & 300 & 450 \\ 253 & 337 & 395 & 451 \end{bmatrix}$$

A tabela anterior é um exemplo de uma matriz do tipo 2×4.

Exemplo 1.1.3. A matriz
$$A = \begin{bmatrix} 1 & 0 \\ 3 & 4 \\ 1 & -1 \end{bmatrix}$$
é uma matriz real do tipo 3×2, i.e. $A \in \mathcal{M}_{3\times 2}(\mathbb{K})$ (é claro que também temos $A \in \mathcal{M}_{3\times 2}(\mathbb{C})$). A linha 2 da matriz A é o elemento $(3, 4)$ de $\mathbb{K}^2$. A coluna 2 da matriz A é o elemento $(0, 4, -1)$ de $\mathbb{K}^3$. O elemento a_{32} (situado na linha 3 e coluna 2 da matriz) é o real -1.

Exemplo 1.1.4. Por $C = [c_{ij}]_{2\times 3}$, onde $c_{ij} = i^j$, para $i \in \{1, 2\}$ e $j \in \{1, 2, 3\}$, representa-se a matriz
$$\begin{bmatrix} 1 & 1^2 & 1^3 \\ 2 & 2^2 & 2^3 \end{bmatrix}.$$

Definição 1.1.5. *Sejam $m, n \in \mathbb{N}$. Uma matriz $A = [a_{ij}]_{m\times n}$ diz-se* ***matriz nula de ordem*** $m \times n$*, e representa-se por* $0_{m\times n}$ *(ou apenas por* 0*, caso não haja ambiguidade), se, para todo $i \in \{1, ..., m\}$ e para todo $j \in \{1, ..., n\}$, tem-se $a_{ij} = 0$.*

Exemplo 1.1.6. $0_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Definição 1.1.7. *Sejam $m, n, p, q \in \mathbb{N}$. Diz-se que as matrizes $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{p\times q}$ são* ***iguais****, e escreve-se $A = B$, se $m = p$, $n = q$ e $a_{ij} = b_{ij}$, quaisquer que sejam $i \in \{1, ..., m\}$ e $j \in \{1, ..., n\}$.*

Exemplo 1.1.8. *As matrizes* $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ *e* $B = [b_{ij}]_{3\times 3}$*, com* $b_{ij} = m.d.c.(i, j)$ *são matrizes iguais.*

Definição 1.1.9. *Sejam $m, n \in \mathbb{N}$ e $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Diz-se que:*

- *A é uma* ***matriz linha*** *se $m = 1$.*
- *A é uma* ***matriz coluna*** *se $n = 1$.*
- *A é uma* ***matriz quadrada*** *se $m = n$.*

Exemplo 1.1.10. *A matriz* $A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ *é uma matriz coluna (do tipo 3×1) e a matriz* $B = \begin{bmatrix} 5 & 6 & 7 & 8 \end{bmatrix}$ *é uma matriz linha (do tipo 1×4).*

Notação e terminologia:

- É usual representar matrizes coluna e matrizes linha por letras minúsculas; além disso, é usual omitir o índice 1 que é comum a todos os elementos. Por exemplo,
$$x = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} \qquad \text{e} \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$
representam uma matriz linha de ordem 4 e uma matriz coluna de ordem 3, respetivamente.

- O conjunto $\mathcal{M}_{n\times n}(\mathbb{K})$ das matrizes quadradas do tipo $n \times n$ também se representa por $\mathcal{M}_n(\mathbb{K})$. Uma matriz A pertencente a $\mathcal{M}_n(\mathbb{K})$ diz-se uma matriz quadrada de ordem n ou, simplesmente, uma matriz de ordem n e pode representar-se por $A = [a_{ij}]_n$.

Apresentam-se seguidamente alguns conceitos relativos a matrizes quadradas.

Definição 1.1.11. *Sejam $n \in \mathbb{N}$, $A \in \mathcal{M}_n(\mathbb{K})$ e $k \in \{1, \ldots, n\}$. Dá-se a designação de **submatriz principal de A de ordem k** a uma matriz quadrada de ordem k obtida de A retirando $n-k$ linhas da matriz A, digamos $i_1, i_2, \ldots, i_{n-k}$, e as $n-k$ colunas com os mesmos índices $i_1, i_2, \ldots, i_{n-k}$.*
*À submatriz principal de A obtida por remoção das últimas $n-k$ linhas e das últimas $n-k$ colunas dá-se a designação de **submatriz principal primária de A de ordem k** e representa-se por A_k.*

Exemplo 1.1.12. *Seja*

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \in \mathcal{M}_3(\mathbb{K}).$$

Então a matriz A tem:

- *uma submatriz principal de ordem 3: a própria matriz A;*
- *três submatrizes principais de ordem 2:*

$$\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix};$$

- *três submatrizes principais de ordem 1: $[a_{11}]$, $[a_{22}]$ e $[a_{33}]$.*

Tem-se $A_1 = [a_{11}]$, $A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ e $A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

Definição 1.1.13. *Sejam $n \in \mathbb{N}$ e $A = [a_{ij}]_n$ uma matriz quadrada sobre $\mathbb{K}$. Os elementos a_{ii}, $i \in \{1, ..., n\}$, designam-se por **elementos principais de A**. Diz-se que:*

- *os elementos $a_{11}, a_{22}, ..., a_{nn}$ se dispõem na **diagonal principal de A**;*
- *os elementos $a_{1n}, a_{2\,n-1}, ..., a_{n1}$ se dispõem na **diagonal secundária** de A;*
- *a **diagonal principal de A é não negativa** se, para todo $i \in \{1, \ldots, n\}$, $a_{ii} \geq 0$;*
- *a **diagonal principal de A é positiva** se, para todo $i \in \{1, \ldots, n\}$, $a_{ii} > 0$.*

Definição 1.1.14. Seja $n \in \mathbb{N}$. Uma matriz quadrada $A = [a_{ij}]_n$ diz-se:

- ***triangular superior*** se, para todos $i, j \in \{1, ..., n\}$,

$$i > j \Longrightarrow a_{ij} = 0;$$

- ***triangular inferior*** se, para todos $i, j \in \{1, ..., n\}$,

$$i < j \Longrightarrow a_{ij} = 0;$$

- ***diagonal*** se é simultaneamente triangular superior e triangular inferior, i.e., se, para todos $i, j \in \{1, ..., n\}$,

$$i \neq j \Longrightarrow a_{ij} = 0.$$

Exemplo 1.1.15. *Sejam*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 4 & 5 \end{bmatrix} \quad e \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

A matriz A é uma matriz triangular superior, B é uma matriz triangular inferior e C é uma matriz diagonal.

Notação: Uma matriz diagonal $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{K})$ pode representar-se abreviadamente por $A = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn})$.

Exemplo 1.1.16. *No exemplo anterior, tem-se $C = \operatorname{diag}(1, 0, 2, 3)$.*

Definição 1.1.17. *Uma matriz diagonal em que todos os elementos diagonais são iguais diz-se uma **matriz escalar**.*

Definição 1.1.18. *Dá-se a designação de **matriz identidade de ordem n**, e representa-se por I_n, à matriz escalar de ordem n em que todos os elementos diagonais são iguais a 1, i.e., para todos $i, j \in \{1, \ldots, n\}$, tem-se*

$$(I_n)_{ij} = \begin{cases} 1 & se\ i = j; \\ 0 & se\ i \neq j. \end{cases}$$

Usualmente, o elemento (i, j) da matriz I_n é representado por δ_{ij}, designado por ***s*ímbolo de Kronecker**.

Exemplo 1.1.19. $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$

1.2 Operações com matrizes

Nesta secção definem-se algumas operações envolvendo matrizes: adição de matrizes, multiplicação de um escalar por uma matriz e multiplicação de matrizes.

Definição 1.2.1. *Sejam $m, n \in \mathbb{N}$ e $A = [a_{ij}], B = [b_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$. Chama-se **matriz soma de** A **e** B, e representa-se por $A + B$, à matriz cuja entrada (i, j) é o elemento $a_{ij} + b_{ij}$, i.e.,*

$$A + B = [a_{ij} + b_{ij}]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} .$$

Exemplo 1.2.2. *Sejam $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & 0 & 4 \\ 2 & -1 & 3 \end{bmatrix}$. Então*

$$A + B = \begin{bmatrix} 1+3 & 2+0 & 3+4 \\ 2+2 & 1+(-1) & 0+3 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 7 \\ 4 & 0 & 3 \end{bmatrix}.$$

Como se poderá verificar no resultado seguinte, a adição de matrizes de $\mathcal{M}_{m\times n}(\mathbb{K})$ satisfaz propriedades semelhantes às da adição de elementos de $\mathbb{K}$.

Teorema 1.2.3. *Sejam $m, n \in \mathbb{N}$ e $A, B, C \in \mathcal{M}_{m\times n}(\mathbb{K})$. Então:*

1. $A + B = B + A$. *(comutatividade da adição em $\mathcal{M}_{m\times n}(\mathbb{K})$)*
2. $A + (B + C) = (A + B) + C$. *(associatividade da adição em $\mathcal{M}_{m\times n}(\mathbb{K})$)*
3. $0_{m\times n} + A = A = A + 0_{m\times n}$. *($0_{m\times n}$ elemento neutro da adição em $\mathcal{M}_{m\times n}(\mathbb{K})$)*
4. *existe uma matriz A' tal que $A + A' = 0_{m\times n} = A' + A$.*
   (existência de elemento oposto, para a adição, de qualquer $A \in \mathcal{M}_{m\times n}(\mathbb{K})$)

Demonstração. Demonstramos as propriedades *1.* e *4.*, deixando a prova de *2.* e *3.* como exercício.

1. Sejam $A = [a_{ij}], B = [b_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$. As matrizes $A + B$ e $B + A$ são ambas do tipo $m \times n$. Além disso,

$$A + B = [a_{ij} + b_{ij}]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} ,$$
$$B + A = [b_{ij} + a_{ij}]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

e, como a adição em $\mathbb{K}$ é comutativa, temos $a_{ij} + b_{ij} = b_{ij} + a_{ij}$, para quaisquer $i \in \{1, \ldots, m\}$ e $j \in \{1, \ldots, n\}$. Portanto, as matrizes $A + B$ e $B + A$ são iguais.

4. Sejam $A = [a_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $A' = [a'_{ij}]$ as matrizes de $\mathcal{M}_{m\times n}(\mathbb{K})$ tais que $a'_{ij} = -a_{ij}$. Então $A + A' \in \mathcal{M}_{m\times n}(\mathbb{K})$ e temos

$$A + A' = [a_{ij} + a'_{ij}]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} ,$$

onde $a_{ij} + a'_{ij} = a_{ij} + (-a_{ij}) = 0$, para quaisquer $i \in \{1, \ldots, m\}$ e $j \in \{1, \ldots, n\}$. Logo, $A + A' = 0_{m\times n}$. Por i), temos $A' + A = 0_{m\times n}$. □

Notação:

- Sendo A, B e C matrizes do mesmo tipo, podemos escrever, sem ambiguidade, $A+B+C$ para representar $(A+B)+C$ e $A+(B+C)$, atendendo à associatividade da adição.
- A matriz A' do teorema anterior representa-se por $-A$.
- Dadas duas matrizes A e B com a mesma ordem, representa-se por $A-B$ a soma de matrizes $A+(-B)$.

Definição 1.2.4. *Sejam $m,n \in \mathbb{N}$, $\alpha \in \mathbb{K}$ e $A=[a_{ij}]_{m\times n} \in \mathcal{M}_{m,n}(\mathbb{K})$. Chama--se **produto do escalar** α **pela matriz** A, e representa-se por αA, a matriz de $\mathcal{M}_{m\times n}(\mathbb{K})$ cujo elemento (i,j) é αa_{ij}, i.e.,*

$$\alpha A = [\alpha a_{ij}]_{\substack{i=1,\dots,m\\ j=1,\dots,n}} .$$

Exemplo 1.2.5. *Se* $A=\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & -1 & 3 \end{bmatrix}$ *então* $2A=\begin{bmatrix} 2 & 4 & 6 & 8 \\ 4 & 0 & -2 & 6 \end{bmatrix}$.

Observação: A matriz A' do Teorema 1.2.3 é a matriz $(-1)A$ e, tal como já foi referido, escrevemos $-A$ para representar esta matriz.

Teorema 1.2.6. *Sejam $m,n \in \mathbb{N}$, $A,B \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $\alpha,\beta \in \mathbb{K}$. Então:*

1. $(\alpha\beta)\,A = \alpha\,(\beta A)$.
2. $(\alpha+\beta)\,A = \alpha A + \beta A$.
3. $\alpha\,(A+B) = \alpha A + \alpha B$.
4. $0A = 0_{m\times n}$.
5. $1A = A$.
6. $(-\alpha)A = \alpha(-A) = -(\alpha A)$.

Demonstração. Mostramos a propriedade *1.*, ficando a prova das restantes propriedades como exercício.

Sejam $\alpha,\beta \in \mathbb{K}$ e $A=[a_{ij}] \in \mathcal{M}_{m,n}(\mathbb{K})$. Então $(\alpha\beta)A \in \mathcal{M}_{m,n}(\mathbb{K})$ e, uma vez que $(\beta A) \in \mathcal{M}_{m,n}(\mathbb{K})$, também temos $\alpha(\beta A) \in \mathcal{M}_{m,n}(\mathbb{K})$. Verifica-se ainda que

$$(\alpha\beta)A = [(\alpha\beta)a_{ij}]_{\substack{i=1,\dots,m\\ j=1,\dots,n}} ,$$
$$\alpha(\beta A) = [\alpha(\beta a_{ij})]_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$$

e, considerando que o produto de elementos de $\mathbb{K}$ é associativo, tem-se $(\alpha\beta)a_{ij} = \alpha(\beta a_{ij})$, para quaisquer $i \in \{1,...,m\}$, $j \in \{1,...,n\}$. Portanto, $(\alpha\beta)A = \alpha(\beta A)$. $\square$

Definição 1.2.7. *Sejam* $m, n, p \in \mathbb{N}$, $A = [a_i j] \in \mathcal{M}_{m\times p}(\mathbb{K})$, $B = [b_{ij}] \in \mathcal{M}_{p\times n}(\mathbb{K})$. *Designa-se por* ***produto de*** A ***por*** B, *e representa-se por* AB, *a matriz de* $\mathcal{M}_{m\times n}(\mathbb{K})$ *cuja entrada* (i, j) *é* $\sum_{k=1}^{p} a_{ik}b_{kj}$, *isto é,*

$$\begin{aligned} AB &= \left[\sum_{k=1}^{p} a_{ik}b_{kj}\right]_{\substack{i=1,\dots,m\\ j=1,\dots,n}} \\ &= [a_{i1}b_{1j} + a_{i2}b_{2j} + ... + a_{i\,p-1}b_{p-1\,j} + a_{ip}b_{pj}]_{\substack{i=1,\dots,m\\ j=1,\dots,n}} . \end{aligned}$$

Exemplo 1.2.8. *Sejam*

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix} \text{ e } B = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 2 \\ 1 & 1 & 2 & 1 \end{bmatrix}.$$

Então AB *é a matriz do tipo* 2×4 *sobre* $\mathbb{R}$ *tal que*

$$\begin{array}{ll} (AB)_{11} = 0\times 2+1\times 1+0\times 1 = 1, & (AB)_{12} = 0\times 0+1\times 2+0\times 1 = 2, \\ (AB)_{13} = 0\times 1+1\times 0+0\times 2 = 0, & (AB)_{14} = 0\times 0+1\times 2+0\times 1 = 2, \\ (AB)_{21} = 2\times 2+3\times 1+1\times 1 = 8, & (AB)_{22} = 2\times 0+3\times 2+1\times 1 = 7, \\ (AB)_{23} = 2\times 1+3\times 0+1\times 2 = 4, & (AB)_{24} = 2\times 0+3\times 2+1\times 1 = 7, \end{array}$$

i.e.,

$$AB = \begin{bmatrix} 1 & 2 & 0 & 2 \\ 8 & 7 & 4 & 7 \end{bmatrix}.$$

Observação: Contrariamente ao que sucede com a adição de matrizes, a multiplicação de matrizes não é, em geral, comutativa, tal como se pode verificar nos exemplos que a seguir se apresentam.

Exemplo 1.2.9. *Considerando as matrizes* A *e* B *do exemplo anterior, concluímos que* BA *não está definido, pois o número de colunas de* B *não coincide com o número de linhas de* A.

Exemplo 1.2.10. *Sejam*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ e } B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}.$$

Então

$$AB = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \neq \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} = BA.$$

Definição 1.2.11. *Sejam* $n \in \mathbb{N}$ *e* A *e* B *duas matrizes quadradas de ordem* n. *Diz-se que as matrizes* A *e* B *são* ***comutáveis*** *ou* ***permutáveis*** *se* $AB = BA$.

Embora o produto de matrizes não seja comutativo, existem outras propriedades que se prova serem válidas relativamente a esta operação.

Teorema 1.2.12. *Sejam $m, n, p \in \mathbb{N}$, $\alpha \in \mathbb{K}$ e $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ e B, C matrizes tais que as operações a seguir indicadas estejam definidas. Então:*

1. $(AB)\,C = A\,(BC)$. (*associatividade da multiplicação*)
2. $A\,(B + C) = AB + AC$.
   (*distributividade, à esquerda, da multiplicação em relação à adição*)
3. $(A + B)\,C = AC + BC$.
   (*distributividade, à direita, da multiplicação em relação à adição*)
4. $\alpha\,(AB) = (\alpha A)\,B = A\,(\alpha B)$.
5. $0_{p\times m}A = 0_{p\times n}$, $A0_{n\times p} = 0_{m\times p}$.
6. $AI_n = A$, $I_mA = A$.
7. se $m = n$, $\quad I_nA = AI_n = A$.

Demonstração. Provam-se as propriedades *1.* e *2.*, ficando a prova das restantes propriedades como exercício.

1. Sejam $A = [a_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$, $B = [b_{ij}] \in \mathcal{M}_{n\times p}(\mathbb{K})$ e $C = [c_{ij}] \in \mathcal{M}_{p\times q}(\mathbb{K})$. As matrizes $A(BC)$ e $(AB)C$ são ambas do tipo $m \times q$. Além disso, para quaisquer $i \in \{1, \ldots, m\}$ e $j \in \{1, \ldots, q\}$, tem-se

$$\begin{aligned}
(A(BC))_{ij} = \textstyle\sum_{k=1}^{n} a_{ik}(BC)_{kj} &= \textstyle\sum_{k=1}^{n} a_{ik} \sum_{t=1}^{p} b_{kt}c_{tj} \\
&= \textstyle\sum_{k=1}^{n} \sum_{t=1}^{p} a_{ik}b_{kt}c_{tj}, \\
((AB)C))_{ij} = \textstyle\sum_{t=1}^{p} (AB)_{it}C_{tj} &= \textstyle\sum_{t=1}^{p} (\sum_{k=1}^{n} a_{ik}b_{kt})c_{tj} \\
&= \textstyle\sum_{k=1}^{n} \sum_{t=1}^{p} a_{ik}b_{kt}c_{tj},
\end{aligned}$$

pelo que $(A(BC))_{ij} = ((AB)C))_{ij}$. Logo, $A(BC) = (AB)C$.

2. Sejam $A = [a_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$, $B = [b_{ij}] \in \mathcal{M}_{n\times p}(\mathbb{K})$ e $C = [c_{ij}] \in \mathcal{M}_{n\times p}(\mathbb{K})$. As matrizes $A(B+C)$ e $AB+AC$ são ambas do tipo $m\times p$. Além disso, para quaisquer $i \in \{1, \ldots, m\}$ e $j \in \{1, \ldots, p\}$, tem-se

$$\begin{aligned}
(A(B + C))_{ij} = \textstyle\sum_{k=1}^{n} a_{ik}(B + C)_{kj} &= \textstyle\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) \\
&= \textstyle\sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj} \\
&= (AB)_{ij} + (AC)_{ij} \\
&= (AB + AC)_{ij}.
\end{aligned}$$

Logo, $A(B + C) = AB + AC$. □

Observação: Sejam A, B e C matrizes tais que os produtos $(AB)C$ e $A(BC)$ estão definidos. Então, atendendo à associatividade da multiplicação, podemos escrever ABC para representar qualquer um dos produtos indicados.

Sejam $m, n \in \mathbb{N}$ e $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Atendendo à definição de multiplicação de matrizes, é simples concluir que a multiplicação de A por A está definida se e só se $m = n$. Neste caso, faz sentido a definição seguinte.

Definição 1.2.13. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Chamamos* **potência de expoente k de** *A, com $k \in \mathbb{N}_0$, à matriz de $\mathcal{M}_n(\mathbb{K})$, que representamos por A^k, definida por*

$$A^k = \begin{cases} I_n & \text{se } k = 0 \\ A^{k-1}A & \text{se } k \in \mathbb{N} \end{cases}.$$

Teorema 1.2.14. *Sejam $n \in \mathbb{N}$, $A \in \mathcal{M}_n(\mathbb{K})$ e $k, l \in \mathbb{N}_0$. Então:*

1. $A^kA^l = A^{k+l}$.
2. $(A^k)^l = A^{kl}$.

Demonstração. Exercício. □

1.3 Matrizes invertíveis

Nesta secção, abordamos um tipo especial de matrizes: as *matrizes invertíveis*. Também conhecidas como *matrizes não singulares*, elas se caracterizam pela existência de uma matriz inversa que, ao ser multiplicada pela matriz original resulta na matriz identidade. Essa propriedade desempenha um papel relevante em diversas aplicações da álgebra linear.

Definição 1.3.1. *Seja $n \in \mathbb{N}$. Uma matriz $A \in \mathcal{M}_n(\mathbb{K})$ diz-se* ***invertível*** *se existe uma matriz $X \in \mathcal{M}_n(\mathbb{K})$ tal que $AX = XA = I_n$.*

Exemplo 1.3.2. *A matriz $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ é invertível, pois existe $X = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$ tal que*

$$AX = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \text{ e}$$
$$XA = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Teorema 1.3.3. *Seja $n \in \mathbb{N}$. Se $A \in \mathcal{M}_n(\mathbb{K})$ é uma matriz invertível, então existe uma e uma só matriz $A' \in \mathcal{M}_n(\mathbb{K})$ tal que $AA' = I_n = A'A$.*

Demonstração. (Existência) Se A é uma matriz invertível, então existe uma matriz A' tal que $AA' = I_n$ e $A'A = I_n$.
(Unicidade) Sejam X e Y matrizes de $\mathcal{M}_n(\mathbb{K})$ tais que $AX = XA = I_n$ e $AY = YA = I_n$. Então

$$X = XI_n = X\,(AY) = (XA)\,Y = I_nY = Y.$$ □

Definição 1.3.4. *Seja $A \in \mathcal{M}_n(\mathbb{K})$ uma matriz invertível. A única matriz $A' \in \mathcal{M}_n(\mathbb{K})$ tal que $A'A = I_n = AA'$ designa-se por* **matriz inversa** *de A e representa-se por A^{-1}.*

Exemplo 1.3.5. *Seja $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$. Do exemplo 7.1.8 sabe-se que A é invertível e tem-se*

$$A^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}.$$

Tal como se pode verificar no exemplo seguinte, nem toda a matriz quadrada é invertível.

Exemplo 1.3.6. *A matriz $A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ não é invertível. Com efeito, se admitirmos que existe $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tal que $AX = XA = I_2$, tem-se*

$$AX = \begin{bmatrix} a+c & b+d \\ -a-c & -b-d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a-b & a-b \\ c-d & c-d \end{bmatrix} = XA,$$

pelo que $0 = c - d = 1$. (contradição).

Definição 1.3.7. *Uma matriz quadrada que não admite inversa diz-se uma **matriz singular** ou **não invertível**.*

Dadas matrizes $A, A' \in \mathcal{M}_n(\mathbb{K})$, diz-se que A' é a inversa de A se ambas as igualdades $AA' = I_n$ e $A'A = I_n$ são satisfeitas. Contudo, sabendo que A é invertível, pode-se concluir que A' é a inversa de A verificando apenas uma das igualdades indicadas: $AA' = I_n$ ou $A'A = I_n$.

Teorema 1.3.8. *Sejam $n \in \mathbb{N}$, $A \in \mathcal{M}_n(\mathbb{K})$ uma matriz invertível e $A' \in \mathcal{M}_n(\mathbb{K})$ tal que $A'A = I_n$ (respectivamente, $AA' = I_n$). Então $A' = A^{-1}$ e, portanto, $AA' = I_n$ (respectivamente, $A'A = I_n$).*

Demonstração. Seja $n \in \mathbb{N}$ e admitamos que A é uma matriz invertível de ordem n e que A' é uma matriz quadrada de ordem n tal que $A'A = I_n$. Então, como,

$$A'A = I_n \Rightarrow A'AA^{-1} = I_nA^{-1} \Rightarrow A'I_n = A^{-1} \Rightarrow A' = A^{-1},$$

temos $A' = A^{-1}$ e, portanto, $AA' = I_n$. □

Teorema 1.3.9. *Sejam $n \in \mathbb{N}$ e $A, B, C \in \mathcal{M}_n(\mathbb{K})$. Se $AB = I_n$ e $CA = I_n$, então $B = C$, A é invertível e $A^{-1} = B = C$.*

Demonstração. Admitamos que $AB = I_n$ e $CA = I_n$. Então

$$C = CI_n = C(AB) = (CA)B = I_nB = B.$$

Portanto, A é invertível e do Teorema 1.3.3 segue que $A^{-1} = B$. □

Os dois resultados anteriores podem ser generalizados. De facto, se A e B são matrizes quadradas de ordem n tais que $AB = I_n$, então também se tem $BA = I_n$, pelo que A e B são matrizes invertíveis e $A = B^{-1}$ e $B = A^{-1}$. A prova desta generalização é apresentada no próximo capítulo.

A respeito de matrizes invertíveis prova-se também o resultado seguinte.

Teorema 1.3.10. *Sejam $n \in \mathbb{N}$ e $A, B \in \mathcal{M}_n(\mathbb{K})$ matrizes invertíveis. Então:*

1. *A^{-1} é invertível e $\left(A^{-1}\right)^{-1} = A$.*
2. *AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$.*

Demonstração. *1.* Imediata pela própria definição de matriz invertível.

2. Como

$$(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

e

$$\left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}A\right)B = B^{-1}I_nB = B^{-1}B = I_n,$$

conclui-se que a inversa de AB existe e é a matriz $B^{-1}A^{-1}$. □

Definição 1.3.11. *Sejam $p, n \in \mathbb{N}$ e $A, B \in \mathcal{M}_{p\times n}(\mathbb{K})$. Diz-se que a matriz A é **equivalente** à matriz B se existem matrizes invertíveis $P \in \mathcal{M}_p(\mathbb{K})$ e $Q \in \mathcal{M}_n(\mathbb{K})$ tais que $B = PAQ$.*

Como caso particular da definição anterior, temos a noção seguinte.

Definição 1.3.12. *Sejam $n \in \mathbb{N}$ e $A, B \in \mathcal{M}_n(\mathbb{K})$. Diz-se que a matriz A é **semelhante** à matriz B se existe uma matriz invertível $P \in \mathcal{M}_n(\mathbb{K})$ tal que $B = PAP^{-1}$.*

Note-se que se A é equivalente (semelhante) a B, então B é equivalente (semelhante a) A e, por isso, dizemos que A e B são equivalentes (semelhantes).

1.4 Transposição e conjugação

Definição 1.4.1. *Sejam $m, n \in \mathbb{N}$ e $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Chama-se **transposta de** A, e representa-se por A^T, à matriz de $\mathcal{M}_{n\times m}(\mathbb{K})$ cuja entrada (i, j) é a_{ji}, i.e., $A^T = [b_{ij}]_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$, onde $b_{ij} = a_{ji}$.*

Exemplo 1.4.2. Se $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$, então $A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$.

Teorema 1.4.3. *Sejam $\alpha \in \mathbb{K}$ e A, B matrizes sobre $\mathbb{K}$ tais que as operações seguintes estejam definidas. Então:*

1. $\left(A^T\right)^T = A$.
2. $(A+B)^T = A^T + B^T$.
3. $(\alpha A)^T = \alpha A^T$.
4. $(AB)^T = B^T A^T$.
5. $(A^k)^T = (A^T)^k$, *para todo* $k \in \mathbb{N}_0$.

Demonstração. Demonstramos a propriedade *4.*, ficando a prova das restantes propriedades como exercício.
Sejam $m, n, p \in \mathbb{N}$, $A = [a_{ij}]_{m\times p}$ e $B = [b_{ij}]_{p\times n}$. Então $(AB)^T$ e $B^T A^T$ são ambas matrizes do tipo $n \times m$. Além disso, para quaisquer $i \in \{1, \ldots, n\}$ e $j \in \{1, \ldots, m\}$, tem-se

$$(B^T A^T)_{ij} = \textstyle\sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki} = (AB)_{ji} = ((AB)^T)_{ij}.$$

Logo, $(AB)^T = B^T A^T$. □

Teorema 1.4.4. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se A é uma matriz invertível, então A^T é invertível e $(A^T)^{-1} = (A^{-1})^T$.*

Demonstração. Pela alínea *iv*) da proposição anterior, tem-se

$$\begin{aligned} \left(A^{-1}\right)^T A^T &= \left(AA^{-1}\right)^T = (I_n)^T = I_n \quad \text{ e} \\ A^T \left(A^{-1}\right)^T &= \left(A^{-1}A\right)^T = (I_n)^T = I_n. \end{aligned}$$

Logo, a matriz A^T é invertível e $(A^T)^{-1} = (A^{-1})^T$. □

Definição 1.4.5. *Sejam $m, n \in \mathbb{N}$ e $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Chama-se* ***conjugada de*** *A, e representa-se por $\overline{A}$, à matriz de $\mathcal{M}_{m\times n}(\mathbb{K})$ tal que, para cada $i \in \{1, ..., m\}$ e $j \in \{1, ..., n\}$, $(\overline{A})_{ij} = \overline{A_{ij}}$.*
Define-se a ***transconjugada*** *de A, e representa-se por A^*, como sendo a transposta da conjugada de A.*

Observação: Se $\mathbb{K} = \mathbb{R}$, tem-se $\overline{A} = A$ e, portanto, $A^* = A^T$.

Exemplo 1.4.6. *Seja*

$$A = \begin{bmatrix} 1-i & 0 & 3 \\ 2+3i & 4 & i \\ 0 & 0 & 6-4i \end{bmatrix}.$$

Então

$$\overline{A} = \begin{bmatrix} 1+i & 0 & 3 \\ 2-3i & 4 & -i \\ 0 & 0 & 6+4i \end{bmatrix} \quad e \quad A^* = \begin{bmatrix} 1+i & 2-3i & 0 \\ 0 & 4 & 0 \\ 3 & -i & 6+4i \end{bmatrix}.$$

Teorema 1.4.7. *Sejam A e B matrizes sobre $\mathbb{K}$ e $\alpha \in \mathbb{K}$. Então, sempre que as operações seguintes estejam definidas, tem-se:*

1. $(A^*)^* = A$.
2. $(A+B)^* = A^* + B^*$.
3. $(\alpha A)^* = \overline{\alpha} A^*$.
4. $(AB)^* = B^* A^*$.
5. $\left(A^k\right)^* = (A^*)^k$, *para todo* $k \in \mathbb{N}_0$.

Demonstração. 1. Sejam $A = [a_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $B = [b_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$. As matrizes $(A+B)^*$ e $A^* + B^*$ são ambas do tipo $n \times m$. Além disso, para quaisquer $i \in \{1,\ldots,n\}$ e $j \in \{1,\ldots,m\}$, tem-se

$$\begin{aligned} [(A+B)^*]_{ij} &= [(\overline{A+B})^T]_{ij} = (\overline{A+B})_{ji} \\ &= \overline{(A+B)_{ji}} = \overline{A_{ji} + B_{ji}} \\ &= \overline{A_{ji}} + \overline{B_{ji}} = (\overline{A})_{ji} + (\overline{B})_{ji} \\ &= ((\overline{A})^T)_{ij} + ((\overline{B})^T)_{ij} = A^*_{ij} + B^*_{ij}. \end{aligned}$$

Logo, $(A+B)^* = A^* + B^*$.

A prova das restantes propriedades é deixada como exercício. □

1.5 Matrizes simétricas e matrizes hermíticas

Definição 1.5.1. *Seja $n \in \mathbb{N}$. Uma matriz quadrada $A \in \mathcal{M}_n(\mathbb{K})$ diz-se:*

- ***simétrica*** *se $A^T = A$;*
- ***antissimétrica*** *se $A^T = -A$.*

Exemplo 1.5.2. *A matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ é uma matriz simétrica, mas a matriz $B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 7 & 5 & 6 \end{bmatrix}$ não é, uma vez que os elementos b_{13} e b_{31} não são iguais.*

Teorema 1.5.3. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Então*

1. $A + A^T$ *é uma matriz simétrica.*
2. $A - A^T$ *é uma matriz antissimétrica.*

Demonstração. Pelas alíneas *1.* e *2.* do Teorema 1.4.3, tem-se

$$\begin{aligned} (A + A^T)^T &= A^T + (A^T)^T = A^T + A = A + A^T, \\ (A - A^T)^T &= A^T - (A^T)^T = A^T - A = -(A - A^T). \end{aligned}$$

Logo, $A + A^T$ é uma matriz simétrica e $A - A^T$ é uma matriz antissimétrica. □

Teorema 1.5.4. *Seja $n \in \mathbb{N}$. Toda a matriz $A \in \mathcal{M}_n(\mathbb{K})$ pode ser expressa como a soma de uma matriz simétrica e de uma matriz antissimétrica.*

Demonstração. Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Tem-se

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T).$$

Pelo teorema anterior, $A + A^T$ é uma matriz simétrica e $A - A^T$ é uma matriz anstissimétrica. Logo, $\frac{1}{2}(A + A^T)$ é uma matriz simétrica e $\frac{1}{2}(A - A)^T$ é uma matriz antissimétrica. Portanto, toda a matriz A é a soma de matriz simétrica e de uma matriz antissimétrica. □

Teorema 1.5.5. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se A é uma matriz simétrica e invertível, então A^{-1} é uma matriz simétrica.*

Demonstração. Seja $A \in \mathcal{M}_n(\mathbb{K})$ uma matriz simétrica e invertível. Como A é uma matriz simétrica, temos $A^T = A$. Então,

$$\left(A^{-1}\right)^T = \left(A^T\right)^{-1} = A^{-1}$$

e, portanto, A^{-1} é uma matriz simétrica. □

Definição 1.5.6. *Seja $n \in \mathbb{N}$. Uma matriz $A \in \mathcal{M}_n(\mathbb{K})$ diz-se:*

- ***hermítica*** *se $A^* = A$.*
- ***anti-hermítica*** *se $A^* = -A$.*

Exemplo 1.5.7. *A matriz* $A = \begin{bmatrix} 1 & 0 & 2+3i \\ 0 & 2 & -i \\ 2-3i & i & 6 \end{bmatrix}$ *é uma matriz hermítica.*

No estudo de matrizes simétricas e de matrizes hermíticas destacam-se algumas matrizes especiais, as matrizes simétricas (hermíticas) definidas positivas, atendendo às suas aplicações práticas (tais como, por exemplo, estudo de cónicas e quádricas, estudo de máximos e mínimos).

Para simplificar a notação, no texto que se segue identificamos uma matriz $[a_{11}] \in \mathcal{M}_{11}(\mathbb{K})$ com o elemento a_{11}, sempre que nos referimos a operações e afirmações envolvendo o único elemento da matriz.

Definição 1.5.8. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz simétrica. Diz-se que a matriz A é:*

- ***simétrica semi-definida positiva*** *se $x^T Ax \geq 0$, para todo $x \in \mathcal{M}_{n\times 1}(\mathbb{R})$.*
- ***simétrica definida positiva*** *se $x^T Ax > 0$, para todo $x \in \mathcal{M}_{n\times 1}(\mathbb{R})\setminus\{0_{n\times 1}\}$.*

Exemplo 1.5.9. *Para qualquer $n \in \mathbb{N}$, a matriz I_n é definida positiva, pois, para qualquer $x \in \mathcal{M}_{n\times 1}(\mathbb{K}) \setminus \{0_{n\times 1}\}$, $x^T I_n x = x^T x > 0$.*

Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{C})$ e $x \in \mathcal{M}_{n\times 1}(\mathbb{C})$. Tem-se $x^T Ax \in \mathcal{M}_{1\times 1}(\mathbb{C})$ e identifica-se $x^T Ax = [c_{11}]$ com o número c_{11}. Se A é hermítica, tem-se $x^T Ax \in \mathbb{R}$; de facto, $\overline{x^T A\overline{x}} = \overline{x^T}\overline{A\overline{x}} = \overline{x}^T A^T x = \overline{x}^T A^T (x^T)^T = (x^T A\overline{x})^T = x^T A\overline{x}$.

Definição 1.5.10. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{C})$ uma matriz hermítica. Diz-se que a matriz A é:*

- ***hermítica semi-definida positiva*** *se $x^* Ax \geq 0$, para todo $x \in \mathcal{M}_{n\times 1}(\mathbb{C})$.*
- ***hermítica definida positiva*** *se $x^* Ax > 0$, para todo $x \in \mathcal{M}_{n\times 1}(\mathbb{C})\setminus\{0_{n\times 1}\}$.*

Teorema 1.5.11. *Sejam $n \in \mathbb{N}$, $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$) e $A \in \mathcal{M}_n(\mathbb{K})$ uma matriz simétrica (resp. hermítica).*

1. *Se A é uma matriz simétrica (resp. hermítica) semi-definida positiva, então a diagonal principal de A é não negativa.*
2. *Se A é uma matriz simétrica (resp. hermítica) definida positiva, então a diagonal principal de A é positiva.*

Demonstração. 1., 2.: Sejam $\mathbb{K} = \mathbb{R}$ e $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz simétrica semi-definida (resp. definida) positiva. Para cada $i \in \{1,\ldots,n\}$, seja $e_i \in \mathcal{M}_{n\times 1}(\mathbb{K})$ tal que $(e_i)_{j1} = 0$, para todo $j \in \{1,\ldots,n\} \setminus \{i\}$ e $(e_i)_{i1} = 1$. Então $e_i^T Ae_i = [a_{ii}]$ e o resultado segue da definição de matriz semi-definida (resp. definida) positiva.
A prova é similar quando $\mathbb{K} = \mathbb{C}$ e $A \in \mathcal{M}_n(\mathbb{C})$ é uma matriz hermítica semi-definida (resp. definida) positiva. □

A prova das propriedades seguintes é um exercício simples, pelo que é deixada ao cuidado do leitor.

Teorema 1.5.12. *Sejam $n \in \mathbb{N}$, $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$), $A, B \in \mathcal{M}_n(\mathbb{K})$ matrizes simétricas (resp. hermíticas) e $\alpha \in \mathbb{R}$.*

1. *Se A e B são matrizes simétricas (resp. hermíticas) definidas positivas, então $A + B$ é simétrica (resp. hermítica) definida positiva.*

2. *Se A é simétrica (resp. hermítica) definida positiva e $\alpha > 0$, então αA é simétrica (resp. hermítica) definida positiva.*

1.6 Matrizes ortogonais e matrizes unitárias

Definição 1.6.1. *Seja $n \in \mathbb{N}$. Uma matriz $A \in \mathcal{M}_n(\mathbb{K})$ diz-se:*

- ***ortogonal*** *se $AA^T = I_n = A^T A$.*
- ***unitária*** *se $AA^* = I_n = A^* A$.*

Observação: Se A é uma matriz ortogonal (resp. unitária), então A é uma matriz invertível e $A^{-1} = A^T$ (resp. $A^{-1} = A^*$).

Exemplo 1.6.2. *A matriz $A = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$ é uma matriz ortogonal, pois*

$$AA^T = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

e

$$A^T A = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Teorema 1.6.3. *Sejam $n \in \mathbb{N}$ e $A, B \in \mathcal{M}_n(\mathbb{K})$.*

1. *Se A é uma matriz ortogonal (resp. unitária), então A^{-1} é também uma matriz ortogonal (resp. unitária).*

2. *Se A e B são matrizes ortogonais (resp. unitárias), então AB é ortogonal (resp. unitária).*

Demonstração. 1. Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se A é uma matriz ortogonal, temos

$$AA^T = I_n = A^T A.$$

Então

$$A^{-1}(A^{-1})^T = A^{-1}(A^T)^{-1} = (A^T A)^{-1} = I_n^{-1} = I_n$$

e

$$(A^{-1})^T A^{-1} = (A^T)^{-1} A^{-1} = (AA^T)^{-1} = I_n^{-1} = I_n,$$

pelo que A^{-1} é também uma matriz ortogonal.

2. Exercício. □

1.7 Matrizes elementares

Nesta secção o estudo é dedicado a uma classe especial de matrizes, as *matrizes elementares*.

Para definirmos esta classe de matrizes, começamos por definir algumas operações que se podem efetuar sobre as linhas (colunas) de uma matriz, designadas por *operações elementares sobre linhas* (*colunas*).

Definição 1.7.1. *Sejam $m, n \in \mathbb{N}$ e $A \in \mathcal{M}_{m,n}(\mathbb{K})$. Definem-se como* ***operações elementares sobre linhas*** *da matriz A as seguintes operações:*

i) troca da linha i com a linha j;

ii) multiplicação da linha i por $\alpha \in \mathbb{K} \setminus \{0\}$;

iii) substituição da linha i pela sua soma com a linha j multiplicada por $\beta \in \mathbb{K}$, com $i \neq j$,

Analogamente, define-se **operação elementar sobre as colunas** de uma matriz, bastando substituir "linha" por "coluna" na definição anterior.

Ao longo do texto adotaremos as seguintes notações para as operações elementares sobre linhas:

- $A \underset{l_i \leftrightarrow l_j}{\longrightarrow} B$: para indicar que a matriz B é obtida da matriz A efetuando a troca das suas linhas i e j;
- $A \underset{l_i \to \alpha l_i}{\longrightarrow} B$: para indicar que a matriz B é obtida de A multiplicando a linha i da matriz A por $\alpha \in \mathbb{K} \setminus \{0\}$;
- $A \underset{l_i \to l_i + \beta l_j}{\longrightarrow} B$: para indicar que a matriz B é obtida de A substituindo a linha i da matriz A pela sua soma com a linha j multiplicada por $\beta \in \mathbb{K}$.

Para representar as operações elementares por colunas adota-se notação semelhante à anterior, mas escreve-se c_i para indicar a coluna i.

Teorema 1.7.2. *Sejam $m, n \in \mathbb{N}$ e $A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$. Se a matriz $B \in \mathcal{M}_{m\times n}(\mathbb{K})$ pode ser obtida da matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ efetuando uma operação elementar sobre linhas (resp., colunas), então a matriz A também pode ser obtida de B efetuando uma operação elementar sobre linhas (resp., colunas).*

Demonstração. No caso das operações elementares por linhas, basta ter em conta que:

- se $A \underset{l_i \leftrightarrow l_j}{\longrightarrow} B$, então $B \underset{l_j \leftrightarrow l_i}{\longrightarrow} A$;

- se $A \underset{l_i \to \alpha l_i}{\longrightarrow} B$, com $\alpha \neq 0$, então $B \underset{l_i \to \frac{1}{\alpha} l_i}{\longrightarrow} A$;

- se $A \underset{l_i \to l_i + \beta l_j}{\longrightarrow} B$, então $B \underset{l_i \to l_i - \beta l_j}{\longrightarrow} A$.

De forma análoga, justifica-se o resultado para o caso das operações elementares por colunas. □

Definição 1.7.3. *Sejam $m, n \in \mathbb{N}$ e $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$. Diz-se que A é* ***equivalente por linhas (resp., por colunas)*** *a B se B pode ser obtida a partir de A efetuando um número finito de operações elementares sobre as linhas (resp., colunas) de A.*

Com base no Teorema 1.7.2 é fácil concluir que se $A \in \mathcal{M}_{m,n}(\mathbb{K})$ é equivalente por linhas (resp., colunas) a $B \in \mathcal{M}_{m,n}(\mathbb{K})$, então B é equivalente por linhas (resp., colunas) a A e, por isso, podemos dizer apenas que A e B são equivalentes por linhas.

O resultado seguinte mostra-nos que podemos efetuar uma operação elementar sobre as linhas de uma matriz premultiplicando A (ou seja, multiplicando A à esquerda) por uma matriz adequada.

Teorema 1.7.4. *Sejam $n, m \in \mathbb{N}$, $\alpha, \beta \in \mathbb{K}$ e $A \in \mathcal{M}_{m,n}(\mathbb{K})$. Então:*

1. *Se E é a matriz quadrada de ordem m obtida da matriz identidade I_m trocando as suas linhas i e j, então a matriz EA é obtida da matriz A trocando as suas linhas i e j.*

2. *Se E é a matriz de ordem m obtida da matriz identidade I_m multiplicando a linha i por $\alpha \in \mathbb{K} \setminus \{0\}$, então EA é a matriz obtida de A multiplicando a linha i por α.*

3. *Se E é a matriz quadrada de ordem m obtida da matriz identidade I_m substituindo a linha na posição i pela sua soma com β vezes a linha na posição j, então EA é a matriz obtida da matriz A substituindo a linha na posição i pela sua soma com β vezes a linha na posição j, $i \neq j$.*

Demonstração. Exercício. □

Analogamente ao que acontece com as operações elementares sobre linhas, uma operação elementar sobre as colunas de uma matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ pode ser obtida posmultiplicando A (ou seja, multiplicando A à direita) por uma matriz obtida da matriz I_n efetuando nas suas colunas a mesma operação elementar que se pretende efetuar na matriz A.

Definição 1.7.5. *Seja $n \in \mathbb{N}$. Chamamos* **matriz elementar sobre linhas (respetivamente, colunas)** *de $\mathcal{M}_n(\mathbb{K})$ a toda a matriz que pode ser obtida da matriz identidade I_n por aplicação de uma operação elementar sobre as suas linhas (respetivamente, colunas).*

Teorema 1.7.6. *Toda a matriz elementar sobre linhas (respetivamente, colunas) é também uma matriz elementar sobre colunas (respetivamente, linhas).*

Demonstração. Facilmente se prova o resultado enunciado. De facto, tem-se:

- $I_n \underset{l_i \leftrightarrow l_j}{\longrightarrow} E$ se e só se $I_n \underset{c_i \leftrightarrow c_j}{\longrightarrow} E$;
- $I_n \underset{l_i \to \alpha l_i}{\longrightarrow} E$ se e só se $I_n \underset{c_i \to \alpha c_i}{\longrightarrow} E$;
- $I_n \underset{l_i \to l_i + \beta l_j}{\longrightarrow} E$ se e só se $I_n \underset{c_i \to c_i + \beta c_j}{\longrightarrow} E$; □

A respeito de matrizes elementares é também conveniente observar que toda a matriz elementar sobre linhas (respetivamente, colunas) é uma matriz invertível.

Teorema 1.7.7. *Seja $n \in \mathbb{N}$. Toda a matriz elementar $E \in \mathcal{M}_n(\mathbb{K})$ é invertível e, para quaisquer $i, j \in \{1, \ldots, n\}$, tem-se:*

1) Se $i \neq j$ e $I_n \underset{l_i \leftrightarrow l_j}{\longrightarrow} E$, então $I_n \underset{l_i \leftrightarrow l_j}{\longrightarrow} E^{-1}$;

2) Se $\alpha \in \mathbb{K} \setminus \{0\}$ e $I_n \underset{l_i \to \alpha l_i}{\longrightarrow} E$, então $I_n \underset{l_i \to \frac{1}{\alpha} l_i}{\longrightarrow} E^{-1}$;

3) Se $i \neq j$, $\beta \in \mathbb{K}$ e $I_n \underset{l_i \to l_i + \beta l_j}{\longrightarrow} E$, então $I_n \underset{l_i \to l_i + (-\beta) l_j}{\longrightarrow} E^{-1}$

Demonstração. Exercício. □

1.8 Formas de escada e característica de uma matriz

Definição 1.8.1. *Sejam $m, n \in \mathbb{N}$. Uma matriz $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ diz-se uma* ***matriz em forma de escada*** *se satisfaz as seguintes condições*

- *se o primeiro elemento não nulo numa linha está na coluna j, então a linha seguinte começa com pelo menos j elementos nulos;*
- *se houver linhas totalmente constituídas por zeros, elas aparecem depois das outras.*

Ao primeiro elemento não nulo de cada linha de uma matriz em forma de escada dá-se a designação de ***pivot****.*

Exemplo 1.8.2.

1. *As matrizes $\mathbf{0}_{m \times n}$ e I_n são matrizes em forma de escada.*
2. *A matriz*
$$\begin{bmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
está em forma de escada.

3. *A matriz*
$$\begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$
embora muito parecida com a matriz anterior, não está em forma de escada (o primeiro elemento não nulo na linha 3 *está na coluna* 4 *e a linha* 4 *não começa com* 4 *elementos nulos).*

4. *A matriz*
$$\begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
não está em forma de escada (o primeiro elemento não nulo na linha 1 *está na coluna* 3 *e a linha* 2 *não começa com* 3 *elementos nulos).*

Teorema 1.8.3. *Sejam $m, n \in \mathbb{N}$. Toda a matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ é equivalente por linhas a uma matriz em forma de escada.*

Demonstração. Sejam $m, n \in \mathbb{N}$ e $A \in \mathcal{M}_{m\times n}(\mathbb{K})$.

Se $A = 0_{m\times n}$, então A é uma matriz em forma de escada.

Consideremos, agora, o caso em que $A \neq 0_{m\times n}$ e façamos a prova do resultado por indução em m.

Base de indução: Se $m = 1$, então A é uma matriz em forma de escada.

Passo de indução: Seja $k \in \mathbb{N}$. Admitamos que toda a matriz $B \in \mathcal{M}_{k,n}(\mathbb{K})$ é equivalente por linhas a uma matriz em forma de escada.
Seja $A \in \mathcal{M}_{k+1,n}(\mathbb{K})$. Mostremos que A é equivalente por linhas a uma matriz em forma de escada. Uma vez que A é uma matriz não nula, existe pelo menos uma coluna com um elemento não nulo. Seja $k \in \{1, \ldots, n\}$ o menor indíce tal que a coluna k tem um elemento não nulo e coloque-se na entrada $(1, k)$ um elemento não nulo, trocando, se necessário, a linha 1 com uma linha que tenha um elemento não nulo na coluna k. Obtem-se, assim, uma matriz da forma

$$B = \begin{bmatrix} 0 & \ldots & 0 & b_{1k} & b_{1,k+1} & \ldots & b_{1n} \\ 0 & \ldots & 0 & b_{2k} & b_{2,k+1} & \ldots & b_{2n} \\ & \ddots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & b_{mn} & b_{m,k+1} & \ldots & b_{mn} \end{bmatrix}.$$

Agora, para $i = 2, 3, \ldots m$, soma-se à linha i a linha 1 multiplicada por $-\frac{b_{ik}}{b_{1k}}$. Esta combinação de operações elementares transforma a matriz B na matriz

$$C = \begin{bmatrix} 0 & \ldots & 0 & b_{1k} & b_{1,k+1} & \ldots & b_{1n} \\ 0 & \ldots & 0 & 0 & c_{2,k+1} & \ldots & c_{2n} \\ & \ddots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & c_{m,k+1} & \ldots & c_{mn} \end{bmatrix}.$$

Por hipótese de indução, a matriz

$$D = \begin{bmatrix} 0 & \dots & 0 & 0 & c_{2,k+1} & \dots & c_{2n} \\ & \ddots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & c_{m,k+1} & \dots & c_{mn} \end{bmatrix}.$$

é equivalente por linhas a uma matriz E em forma de escada. Assim, aplicando a C as operações elementares que permitem a partir de D obter a matriz E (com a respetiva adequação no indice das linhas), obtem-se uma matriz em forma de escada equivalente por linhas à matriz a matriz C e, portanto, a matriz A também é equivalente por linhas a uma matriz em forma de escada. $\square$

Apresenta-se seguidamente um processo, baseado na demonstração anterior, que permite obter uma matriz em forma de escada equivalente por linhas a uma dada matriz. A este processo dá-se a designação de **método de eliminação de Gauss**.

Método de eliminação de Gauss

Sejam $m, n \in \mathbb{N}$ e $A = [a_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$. Então a matriz A pode ser reduzida a uma matriz em escada, efectuando operações elementares sobre as suas linhas, de acordo com o seguinte processo.

Para k de 1 até m:

i) Procurar a primeira coluna com elementos não nulos na linha k ou abaixo desta.

ii) Se não existir a coluna indicada em i), dá-se o processo por terminado.

iii) Caso exista a coluna referida em i) e se j_k é essa coluna, $j_k \in \{1, ..., n\}$, assegura-se que o elemento na linha k desta coluna é não nulo, trocando, se necessário, a linha k com alguma linha que esteja abaixo; representemos por $a_{kj_k}^{(k)}$ esse elemento.

iv) Para cada $i \in \{k+1, ..., m\}$, adiciona-se à linha i a linha k multiplicada por $-\frac{a_{ij_k}^{(k)}}{a_{kj_k}^{(k)}}$, onde $a_{ij_k}^{(k)}$ representa o elemento na linha i e coluna j_k da matriz que foi obtida após a aplicação dos passos anteriores.

Terminado o processo a matriz que se obtém é uma matriz em escada.

Aos elementos $a_{1j_1}^{(1)}, a_{2j_2}^{(2)}, ...$ referidos no processo anterior dá-se a designação de ***pivots da eliminação***.

Exemplo 1.8.4. *Consideremos a matriz*

$$A = \begin{bmatrix} 0 & 1 & 2 & 2 \\ 1 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \end{bmatrix}.$$

Aplicando o método de eliminação de Gauss a esta matriz, temos

$$\begin{bmatrix} 0 & 1 & 2 & 2 \\ 1 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \end{bmatrix} \xrightarrow[l_1 \leftrightarrow l_4]{} \begin{bmatrix} 1 & 3 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 \end{bmatrix} \to$$

$$\xrightarrow[\substack{l_2 \to l_2 - l_1 \\ l_3 \to l_3 - 3l_1}]{} \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & -2 & -1 & -1 \\ 0 & -8 & -2 & -2 \\ 0 & 1 & 2 & 2 \end{bmatrix} \xrightarrow[\substack{l_3 \to l_3 - 4l_2 \\ l_4 \to l_4 + \frac{1}{2}l_2}]{} \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & -2 & -1 & -1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & \frac{3}{2} & \frac{3}{2} \end{bmatrix} \to$$

$$\xrightarrow[l_4 \to l_4 - \frac{3}{4}l_3]{} \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & -2 & -1 & -1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A última matriz obtida é uma matriz em forma de escada e equivalente por linhas à matriz A.

Observação: Pela Proposição 1.7.6, toda a transformação que pode ser feita numa matriz por meio de operações elementares sobre linhas também pode ser realizada por meio de operações elementares sobre colunas. Por conseguinte, toda a matriz pode ser transformada numa matriz em escada por meio de operações elementares sobre colunas.

O método de eliminação de Gauss, quando aplicado a uma matriz A, permite obter uma matriz em escada equivalente por linhas à matriz inicial. Porém, a matriz em escada que é obtida no final do processo pode não ser sempre a mesma, uma vez que que há alguma flexibilidade na escolha das transformações elementares a efetuar sobre a matriz A (nomeadamente, na escolha das linhas que se trocam para colocar um elemento na posição de pivot). No entanto, embora não seja possível garantir a unicidade da matriz em escada que é obtida por aplicação do método de eliminação de Gauss, prova-se que o número de pivots usados no método de eliminação de Gauss, que é igual ao número de linhas não nulas da matriz em escada obtida de A, é univocamente determinado pelas entradas da matriz A. Com efeito, quaisquer matrizes em escada equivalentes por linhas a uma matriz A têm o mesmo número de linhas não nulas, pelo que faz sentido considerar a definição seguinte.

Definição 1.8.5. *Sejam $m, n \in \mathbb{N}$ e $A \in \mathcal{M}_{m,n}(\mathbb{K})$. Designa-se por **característica** da matriz A, e representa-se por car(A), o número de linhas não nulas de qualquer matriz em forma de escada que seja equivalente por linhas a A.*

Exemplo 1.8.6. *Como vimos no exemplo anterior, a matriz*

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

é equivalente por linhas à seguinte matriz em forma de escada

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

Então, como U tem 3 linhas não nulas, temos $car(A) = 3$.

O processo descrito no método de eliminação de Gauss pode ser complementado com outras operações elementares de forma a que uma matriz seja transformada até se obter uma *matriz em forma de escada reduzida*.

Definição 1.8.7. *Sejam $m, n \in \mathbb{K}$ e $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Diz-se que A é uma **matriz em forma de escada reduzida** (ou que está em forma de escada reduzida) se satisfaz as seguintes condições:*

- *A é uma matriz em escada;*
- *se uma linha tem elementos não nulos, então o primeiro elemento não nulo da linha é igual a 1;*
- *se o primeiro elemento não nulo de uma linha i está na coluna j, então todos os elementos da coluna j, com excepção do elemento que está na linha i, são iguais a zero.*

Exemplo 1.8.8.

1. *As matrizes $0_{m\times n}$ e I_n são matrizes em forma de escada reduzida.*
2. *A matriz*
$$\begin{bmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
está em forma de escada reduzida.
3. *A matriz*
$$\begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$
embora muito parecida com a matriz anterior, não está em forma de escada reduzida.
4. *A matriz* $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ *não está em forma de escada reduzida.*

Seguindo uma prova similar à do teorema anterior, prova-se que toda a matriz pode ser transformada numa matriz em forma de escada reduzida.

Teorema 1.8.9. *Sejam $m, n \in \mathbb{N}$. Toda a matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ é equivalente por linhas a uma matriz em forma de escada reduzida.*

Demonstração. Sejam $m, n \in \mathbb{N}$ e $A \in \mathcal{M}_{m\times n}(\mathbb{K})$.

Se $A = 0_{m\times n}$, então A é uma matriz em forma de escada.

Consideremos, agora, o caso em que $A \neq 0_{m\times n}$ e façamos a prova do resultado por indução em m.

Base de indução: Se $m = 1$, a matriz A também é equivalente por linhas a uma matriz em forma de escada reduzida. De facto, para obter uma matriz em forma de escada reduzida equivalente por linhas à matriz A, basta multiplicar a matriz A por $\frac{1}{a_{1k}}$, onde k é o indíce mais pequeno entre os indíces das colunas que têm elementos não nulos.

Passo de indução: Seja $k \in \mathbb{N}$. Admitamos que toda a matriz $B \in \mathcal{M}_{k,n}(\mathbb{K})$ é equivalente por linhas a uma matriz em forma de escada reduzida.
Seja $A \in \mathcal{M}_{k+1,n}(\mathbb{K})$. Mostremos que A é equivalente por linhas a uma matriz em forma de escada reduzida. Uma vez que A é uma matriz não nula, existe pelo menos uma coluna com um elemento não nulo. Seja $k \in \{1, \ldots, n\}$ o menor indíce tal que a coluna k tem um elemento não nulo e coloque-se na entrada $(1, k)$ um elemento não nulo, trocando, se necessário, a linha 1 com uma linha que tenha um elemento não nulo na coluna k. Seguidamente, multiplique-se a linha 1 por $\frac{1}{a_{1k}}$. Obtem-se, assim, uma matriz da forma

$$B = \begin{bmatrix} 0 & \ldots & 0 & 1 & b_{1,k+1} & \ldots & b_{1n} \\ 0 & \ldots & 0 & b_{2k} & b_{2,k+1} & \ldots & b_{2n} \\ & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & b_{mn} & b_{m,k+1} & \ldots & b_{mn} \end{bmatrix}.$$

Agora, para $i = 2, 3, \ldots m$, soma-se à linha i a linha 1 multiplicada por $-b_{ik}$. Esta combinação de operações elementares transforma a matriz B na matriz

$$C = \begin{bmatrix} 0 & \ldots & 0 & 1 & b_{1,k+1} & \ldots & b_{1n} \\ 0 & \ldots & 0 & 0 & c_{2,k+1} & \ldots & c_{2n} \\ & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & c_{m,k+1} & \ldots & c_{mn} \end{bmatrix}.$$

Por hipótese de indução, a matriz

$$D = \begin{bmatrix} 0 & \ldots & 0 & 0 & c_{2,k+1} & \ldots & c_{2n} \\ & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & c_{m,k+1} & \ldots & c_{mn} \end{bmatrix}.$$

é equivalente por linhas a uma matriz E em forma de escada reduzida. Assim, aplicando a C as operações elementares que permitem a partir de D obter a matriz

E (com a respetiva adequação no indice das linhas), obtem-se uma matriz em forma de escada reduzida equivalente por linhas à matriz a matriz C e, portanto, a matriz A também é equivalente por linhas a uma matriz em forma de escada reduzida. $\square$

Ao processo de transformar uma dada matriz numa matriz em forma de escada reduzida dá-se a designação de ***condensação*** da matriz. O método a seguir descrito para condensação de uma matriz é designado por método de ***eliminação de Gauss-Jordan***, por ser uma extensão do método de eliminação de Gauss.

Método de eliminação de Gauss-Jordan

Sejam $m, n \in \mathbb{N}$ e $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Se A é a matriz nula, então a matriz já está na forma de escada reduzida. Caso A não seja a matriz nula, então A pode ser transformada numa matriz em escada reduzida, efectuando operações elementares sobre as suas linhas, de acordo com o seguinte processo: aplicando o método de eliminação de Gauss à matriz A obtemos uma matriz em escada A' equivalente a A. Uma vez obtida a matriz em escada A', anulam-se todos os elementos não nulos que estejam acima dos pivots ${a_{kj_k}}^{(k)}$ da matriz A'; para tal, para cada $i \in \{1, \ldots, k-1\}$, adiciona-se à linha i da matriz A' a linha k multiplicada por $-\frac{a_{ij_k}^{(k)}}{{a_{kj_k}}^{(k)}}$. Finalmente, multiplica-se cada linha não nula da matriz pelo inverso do pivot dessa linha. Terminado o processo, obtém-se uma matriz em escada reduzida e equivalente por linhas à matriz A.

É conveniente observar que uma matriz A é equivalente por linhas a uma única matriz em forma de escada reduzida - a esta matriz em escada reduzida dá-se a designação de **forma de Hermite** de A.

Exemplo 1.8.10. *Consideremos a matriz a seguir indicada*

$$A = \begin{bmatrix} 0 & 1 & 2 & 2 & 0 & 2 \\ 0 & 1 & 2 & -2 & 0 & 1 \\ 0 & 3 & 6 & 2 & 1 & -1 \\ 0 & 0 & 0 & 2 & 0 & 1 \end{bmatrix}.$$

Aplicando o método descrito anteriormente de forma a transformar a matriz A numa matriz em forma de escada reduzida, obtemos

$$\begin{bmatrix} 0 & 1 & 2 & 2 & 0 & 2 \\ 0 & 1 & 2 & -2 & 0 & 1 \\ 0 & 3 & 6 & 2 & 1 & -1 \\ 0 & 0 & 0 & 2 & 0 & 1 \end{bmatrix} \xrightarrow[\substack{l_2 \to l_2 - l_1 \\ l_3 \to l_3 - 3l_1}]{} \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 2 \\ 0 & 0 & 0 & -4 & 0 & -1 \\ 0 & 0 & 0 & -4 & 1 & -7 \\ 0 & 0 & 0 & 2 & 0 & 1 \end{bmatrix} \to$$

$$\xrightarrow[\substack{l_3 \to l_3 - l_2 \\ l_4 \to l_4 + \frac{1}{2}l_2}]{} \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 2 \\ 0 & 0 & 0 & -4 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \xrightarrow[\substack{l_3 \to l_3 + 12l_4 \\ l_2 \to l_2 + 2l_4 \\ l_1 \to l_1 - 4l_4}]{} \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \to$$

$$\xrightarrow[l_1 \to l_1 + \frac{1}{4}l_2]{} \begin{bmatrix} 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \xrightarrow[\substack{l_2 \to -\frac{1}{4}l_2 \\ l_4 \to 2l_4}]{} \begin{bmatrix} 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

A última matriz obtida é a forma de Hermite de A.

Observação: Pelo Teorema 1.7.6, a transformação de uma matriz A numa matriz em forma de escada reduzida também pode ser obtida por meio de operações elementares sobre colunas.

2. Sistemas de Equações Lineares

2.1 Conceitos básicos

São frequentes os problemas cuja resolução envolve sistemas de equações lineares. Embora alguns destes sistemas sejam simples de resolver, outros há que, devido às suas dimensões e complexidade, requerem métodos sistemáticos para a sua resolução. Neste capítulo iremos estudar um desses métodos.

Tal como no capítulo anterior, representamos por $\mathbb{K}$ o conjunto $\mathbb{R}$ ou o conjunto $\mathbb{C}$.

Definição 2.1.1. *Sejam $n \in \mathbb{N}$. Uma **equação linear** nas incógnitas $x_1, x_2, \ldots, x_n$, sobre $\mathbb{K}$, é uma equação do tipo*

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b,$$

com $a_1, a_2, \ldots, a_n, b \in \mathbb{K}$.

*Os elementos a_i $(i \in \{1, 2, ..., n\})$ designam-se por **coeficientes** da equação e o elemento b designa-se por **termo independente** da equação.*

Definição 2.1.2. *Sejam $n, m \in \mathbb{N}$. Dá-se o nome de **sistema de m equações lineares em n incógnitas** $x_1, x_2, ..., x_n$, sobre $\mathbb{K}$, a uma coleção de equações lineares*

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

onde $a_{ij}, b_i \in \mathbb{K}$, para todo $i \in \{1, 2, ..., m\}$ e $j \in \{1, 2, ..., n\}$.

*O sistema (S) diz-se **homogéneo** se $b_1 = b_2 = \ldots = b_m = 0$, i.e., se os termos independentes de (S) são todos nulos.*

*Chama-se **solução de** (S) a qualquer n-uplo $(\alpha_1, \alpha_2, ..., \alpha_n)$ de $\mathbb{K}^n$ tal que, para todo $i \in \{1, ..., m\}$, $a_{i1}\alpha_1 + a_{i2}\alpha_2 + ... + a_{in}\alpha_n = b_i$.*

*Representa-se por $\mathrm{Sol}_{(S)}$ o **conjunto de soluções de** (S), i.e.,*

$$\mathrm{Sol}_{(S)} = \{(\alpha_1, \alpha_2, ..., \alpha_n) \in \mathbb{K}^n : a_{i1}\alpha_1 + a_{i2}\alpha_2 + ... + a_{in}\alpha_n = b_i, \forall i \in \{1, ..., m\}\}.$$

Definição 2.1.3. *Se* (S) *e* (S') *são sistemas de equações lineares sobre* $\mathbb{K}$ *com o mesmo conjunto de soluções, diz-se que* (S) *e* (S') *são* ***equivalentes****.*

Definição 2.1.4. *Um sistema (S) de* m *equações lineares em* n *incógnitas e de coeficientes em* $\mathbb{K}$ *diz-se:*

- ***impossível*** *se não tem solução i.e., se* $\mathrm{Sol}_{(S)} = \emptyset$*;*
- ***possível*** *se tem, pelo menos, uma solução, i.e., se* $\mathrm{Sol}_{(S)} \neq \emptyset$*;*
- ***possível determinado*** *se o sistema tem uma única solução;*
- ***possível indeterminado*** *caso o sistema tenha mais do que uma solução.*

Observação: Um sistema (S) de m equações lineares em n incógnitas que seja homogéneo é sempre possível, uma vez que $(0, 0, ..., 0) \in \mathbb{K}^n$ é solução de (S); a esta solução dá-se a designação de ***solução trivial***.

Dado um sistema (S) de equações lineares entende-se por:

- ***discutir o sistema***, verificar se (S) é possível e, neste caso, se é determinado ou indeterminado;
- ***resolver o sistema***, determinar o conjunto de soluções do sistema.

Exemplo 2.1.5. *A equação*

$$x_1 \;+\; x_2 \;-\; 8x_3 \;=\; 0$$

é equivalente a

$$x_1 \;=\; -x_2 \;+\; 8x_3.$$

Uma vez que x_2 *e* x_3 *são arbitrários, este sistema é possível e indeterminado. Para obtermos uma solução diferente da trivial podemos considerar, por exemplo,* $x_2 = 1$ *e* $x_3 = 1$*, donde resulta* $x_1 = 7$*.*

Exemplo 2.1.6. *Consideremos o seguinte sistema de 3 equações lineares em 3 incógnitas:*

$$\left\{\begin{array}{rcrcrcrl} -x & - & 5y & - & 2z & = & 2 & (i) \\ 2x & - & 2y & + & z & = & 0 & (ii) \\ 3x & + & 3y & + & 3z & = & -1 & (iii). \end{array}\right.$$

Se adicionarmos (i) *e* (iii)*, obtemos obtemos a equação*

$$2x - 2y + z = 1,$$

o que não é consistente com a equação (ii)*, pelo que o sistema não admite nenhuma solução, i.e., é um sistema impossível.*

Exemplo 2.1.7. *Dado o sistema de* 3 *equações lineares em* 3 *incógnitas*

$$\begin{cases} 0x_1 & + & 1x_2 & - & 8x_3 & = & -17 & \textit{(i)} \\ 1x_1 & + & 0x_2 & + & 1x_3 & = & 10 & \textit{(ii)} \\ 1x_1 & - & 1x_2 & + & 0x_3 & = & 0 & \textit{(iii)} \end{cases}$$

vamos determinar o seu conjunto de soluções.
Se trocarmos a equação (i) com a equação (ii), obtemos

$$\begin{cases} 1x_1 & + & 0x_2 & + & 1x_3 & = & 10 & \textit{(i)} \\ 0x_1 & + & 1x_2 & - & 8x_3 & = & -17 & \textit{(ii)} \\ 1x_1 & - & 1x_2 & + & 0x_3 & = & 0 & \textit{(iii)} \end{cases}$$

Agora, se substituirmos a equação (iii) por (iii)-(i) ficamos com

$$\begin{cases} 1x_1 & + & 0x_2 & + & 1x_3 & = & 10 & \textit{(i)} \\ 0x_1 & + & 1x_2 & - & 8x_3 & = & -17 & \textit{(ii)} \\ 0x_1 & - & 1x_2 & - & 1x_3 & = & -10 & \textit{(iii)} \end{cases}$$

Substituindo a equação (iii) por (iii)+(ii) tem-se

$$\begin{cases} 1x_1 & + & 0x_2 & + & 1x_3 & = & 10 & \textit{(i)} \\ 0x_1 & + & 1x_2 & - & 8x_3 & = & -17 & \textit{(ii)} \\ 0x_1 & + & 0x_2 & - & 9x_3 & = & -27 & \textit{(iii)} \end{cases}$$

Finalmente, se multiplicarmos a equação (iii) por $-1/9$*, obtém-se*

$$\begin{cases} 1x_1 & + & 0x_2 & + & 1x_3 & = & 10 & \textit{(i)} \\ 0x_1 & + & 1x_2 & - & 8x_3 & = & -17 & \textit{(ii)} \\ 0x_1 & + & 0x_2 & + & 1x_3 & = & 3 & \textit{(iii)} \end{cases}$$

O último sistema é de resolução mais simples do que o sistema inicial: substituindo x_3 *por* 3 *em (ii) obtém-se* $x_2 = 7$ *e da equação (i), por substituição de* x_3 *e* x_2*, resulta que* $x_1 = 7$*. Assim, uma vez que o sistema admite solução e que esta é única, concluímos que o sistema é possível e determinado.*

Em cada um dos exemplos anteriores, o sistema inicial foi sucessivamente transformado noutros sistemas efectuando apenas as seguintes operações sobre equações:

i) troca da equação i com a equação j;

ii) multiplicação da equação i por $\alpha \in \mathbb{K} \setminus \{0\}$;

iii) substituição da equação i pela sua soma com a equação j multiplicada por $\beta \in \mathbb{K}$, com $i \neq j$.

A estas operações damos a designação de ***operações elementares sobre equações***.
É simples de verificar que se (S') for um sistema obtido a partir de um sistema (S) efectuando uma operação elementar sobre as equações de (S), então os dois sistemas são equivalentes.

Um sistema (S) de equações lineares e de coeficientes em $\mathbb{K}$

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

pode ser representado pela equação matricial

$$Ax = b$$

onde

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

As matrizes A, x e b designam-se, respetivamente, por ***matriz simples de*** (S), ***matriz incógnita de*** (S) e ***matriz dos termos independentes de*** (S). Tendo em conta que as incógnitas têm um papel secundário na resolução de um sistema e que as operações elementares sobre as equações envolvem apenas os coeficientes e os termos independentes, o sistema (S) pode ainda ser representado, de uma forma mais abreviada, pela matriz

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

à qual se dá a designação de ***matriz ampliada de*** (S) e que se representa por $[A \,|\, b]$. O sistema (S) fica completamente representado por esta matriz, uma vez que cada linha da matriz representa uma equação de (S).

Exemplo 2.1.8. *Se considerarmos o sistema* (S) *apresentado no exemplo* 2.1.7*, a matriz simples, a matriz incógnita e a matriz dos termos independentes deste sistema são, respetivamente,*

$$A = \begin{bmatrix} 0 & 1 & -8 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} -17 \\ 10 \\ 0 \end{bmatrix}.$$

A matriz ampliada associada ao sistema é a matriz

$$A = \left[\begin{array}{ccc|c} 0 & 1 & -8 & -17 \\ 1 & 0 & 1 & 10 \\ 1 & -1 & 0 & 0 \end{array}\right].$$

Dado um sistema (S) de m equações lineares em n incógnitas e de coeficientes em $\mathbb{K}$, o seu conjunto de soluções pode ser determinado através da resolução da equação matricial $Ax = b$ que lhe está associada. De facto, $(\alpha_1, \alpha_2, ..., \alpha_n) \in \mathbb{K}^n$ é solução de (S) se e só se

$$A \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = b.$$

Nos exemplos 2.1.5, 2.1.6 e 2.1.7, os sistemas foram resolvidos recorrendo apenas a operações elementares sobre equações. Efectuar uma destas operações sobre as equações de um sistema (S) corresponde, em termos matriciais, a efectuar operações análogas sobre as linhas da matriz ampliada $[A \,|\, b]$ associada ao sistema. Mais precisamente:

- trocar a equação i com a equação j no sistema (S) corresponde a trocar a linha i com a linha j da matriz $[A \,|\, b]$;
- multiplicar a equação i do sistema (S) por $\alpha \in \mathbb{K} \setminus \{0\}$ corresponde a multiplicar a linha i da matriz $[A \,|\, b]$ por α;
- substituir a equação i do sistema (S) pela sua soma com a equação j multiplicada por $\beta \in \mathbb{K}$ corresponde a substituir a linha i da matriz $[A \,|\, b]$ pela sua soma com a linha j multiplicada por $\beta \in \mathbb{K}$, com $i \neq j$.

É simples verificar que, da mesma forma que as operações elementares sobre as equações de um sistema não alteram o seu conjunto de soluções, as operações elementares sobre as linhas da sua matriz ampliada também não alteram a solução da equação matricial associada ao sistema.

Teorema 2.1.9. *Sejam $m, n \in \mathbb{N}$ e $Ax = b$ a equação matricial de um sistema de m equações lineares em n incógnitas, sobre $\mathbb{K}$. Se a matriz $[A' \,|\, b']$ é obtida de $[A \,|\, b]$ efectuando uma operação elementar sobre as linhas, então $A'x = b'$ e $Ax = b$ têm o mesmo conjunto de soluções.*

Demonstração. Cada operação elementar sobre as linhas da matriz ampliada $[A \,|\, b]$ corresponde a multiplicar (à esquerda) ambos os membros da equação $Ax = b$ por uma matriz elementar E. Assim, tendo em conta que $A' = EA$, $b' = Eb$ e que as matrizes elementares são invertíveis, o resultado é imediato. De facto, dado $c \in \mathcal{M}_{n\times 1}(\mathbb{K})$,

$$Ac = b \Rightarrow EAc = Eb \Rightarrow A'c = b'$$

e

$$A'c = b' \Rightarrow EAc = Eb \Rightarrow E^{-1}EAc = E^{-1}Eb \Rightarrow Ac = b. \qquad \square$$

Observação: No primeiro capítulo observámos que toda a matriz que é equivalente por linhas a uma matriz A também é equivalente por colunas à matriz A. Porém, o resultado anterior não é válido para todas as operações elementares por colunas. Com efeito, quando se aplicam operações elementares sobre matrizes tendo por objetivo a resolução de sistemas, a única operação sobre colunas que pode ser aplicada é a troca de colunas, sendo que neste caso tem de se proceder a uma troca de incógnitas no sistema final que seja coerente com a troca de colunas efetuada. Por este motivo, na resolução de sistemas optaremos por recorrer apenas a operações elementares sobre linhas.

2.2 Discussão e resolução de sistemas

Nesta secção, vemos como aplicar o método de eliminação de Gauss na discussão e resolução de sistemas. Este método transforma o sistema original em um sistema equivalente, porém de mais fácil resolução.

No Exemplo 2.1.7, o sistema

$$\left\{\begin{array}{rcrcrcr} 0x_1 & + & 1x_2 & - & 8x_3 & = & -17 \\ 1x_1 & + & 0x_2 & + & 1x_3 & = & 10 \\ 1x_1 & - & 1x_2 & + & 0x_3 & = & 0 \end{array}\right.$$

foi sucessivamente transformado, efectuando operações elementares sobre equações, até obtermos o sistema

$$\left\{\begin{array}{rcrcrcr} 1x_1 & + & 0x_2 & + & 1x_3 & = & 10 \\ 0x_1 & + & 1x_2 & - & 8x_3 & = & -17 \\ 0x_1 & + & 0x_2 & + & 1x_3 & = & 3 \end{array}\right.$$

o qual é de resolução bastante mais simples. Considerando a representação do sistema em termos de matrizes, esta transformação corresponde a efectuar sucessivas operações elementares sobre as linhas da matriz ampliada do sistema

$$\left[\begin{array}{rrr|r} 0 & 1 & -8 & -17 \\ 1 & 0 & 1 & 10 \\ 1 & -1 & 0 & 0 \end{array}\right]$$

até obtermos a matriz

$$\left[\begin{array}{rrr|r} 1 & 0 & 1 & 10 \\ 0 & 1 & -8 & -17 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

Esta última matriz tem a particularidade de ter a forma de escada.

O processo que foi adoptado na resolução do sistema do Exemplo 2.1.7, e que permitiu reduzir a matriz ampliada deste sistema a uma matriz em escada, é o método de eliminação de Gauss descrito no capítulo anterior.

O método de eliminação de Gauss permite simplificar e sistematizar a resolução de sistemas de equações lineares. Um passo elementar deste método, quando aplicado à matriz ampliada $[A|b]$ de um sistema $Ax = b$, consiste em adicionar a uma certa equação um múltiplo de outra, de forma a que na equação obtida seja nulo o coeficiente de certa incógnita. Diz-se que se eliminou essa incógnita da equação. Os passos elementares são conduzidos de maneira a eliminar a incógnita x_1 de todas as equações a partir da 2ª equação, depois eliminar a incógnita x_2 de todas as equações a partir da 3ª equação, etc. Quando termina o método de eliminação de Gauss, obtemos uma matriz em escada $[U|c]$. O sistema correspondente a esta matriz, $Ux = c$, é de resolução mais simples e há a garantia de ser equivalente ao sistema $Ax = b$, uma vez que $[U|c]$ é obtida de $[A|b]$ efectuando apenas operações elementares sobre linhas. Quando se obtém o sistema correspondente à matriz $[U|c]$, é fácil verificar se o sistema é possível ou impossível. Se o sistema for possível, resolve-se de baixo para cima, escrevendo, se necessário, as ***incógnitas básicas*** (as que estão a multiplicar pelos pivots) em função das ***livres*** (as restantes variáveis).

Exemplo 2.2.1. *Consideremos o seguinte sistema de 4 equações lineares em 4 incógnitas*

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & + & x_3 & + & x_4 & = & 1 \\ x_1 & + & x_2 & & & & & = & 2 \\ 2x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 \\ & & x_2 & + & 2x_3 & + & x_4 & = & 2 \end{array}\right. .$$

Aplicando o método de eliminação de Gauss à matriz ampliada do sistema, temos

$$\left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 2 \\ 2 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 & 2 \end{array}\right] \xrightarrow[\substack{l_2 \to l_2 - l_1 \\ l_3 \to l_3 - 2l_1}]{} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & -3 & -1 & -1 & -1 \\ 0 & 1 & 2 & 1 & 2 \end{array}\right] \to$$

$$\xrightarrow[\substack{l_3 \to l_3 - 3l_2 \\ l_4 \to l_4 + l_2}]{} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 2 & 2 & -4 \\ 0 & 0 & 1 & 0 & 3 \end{array}\right] \xrightarrow[l_4 \to l_4 - \frac{1}{2}l_3]{} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 2 & 2 & -4 \\ 0 & 0 & 0 & -1 & 5 \end{array}\right].$$

Desta forma obtém-se o sistema

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & + & x_3 & + & x_4 & = & 1 \\ & & -x_2 & - & x_3 & - & x_4 & = & 1 \\ & & & & 2x_3 & + & 2x_4 & = & -4 \\ & & & & & - & x_4 & = & 5 \end{array}\right.$$

equivalente ao inicial, mas de mais fácil resolução. Da última equação obtém-se $x_4 = -5$*, substituindo* x_4 *na 3ª equação temos* $x_3 = 3$*, donde* $x_2 = 1$ *e* $x_1 = 1$.

Exemplo 2.2.2. *Consideremos o seguinte de 4 equações lineares em 4 incógnitas*

$$\left\{\begin{array}{rcrcrcrcr} 2x_1 & + & 4x_2 & + & 6x_3 & + & 8x_4 & = & 1 \\ -2x_1 & - & 2x_2 & - & 4x_3 & - & 6x_4 & = & 1 \\ 4x_1 & + & 8x_2 & + & 14x_3 & + & 8x_4 & = & -6 \\ 4x_1 & + & 8x_2 & + & 16x_3 & + & 8x_4 & = & -6 \end{array}\right. .$$

Aplicando o método de eliminação de Gauss à matriz ampliada do sistema, temos

$$\left[\begin{array}{rrrr|r} 2 & 4 & 6 & 8 & 1 \\ -2 & -2 & -4 & -6 & 1 \\ 4 & 8 & 14 & 8 & -6 \\ 4 & 12 & 16 & 20 & 6 \end{array}\right] \xrightarrow[\substack{l_2 \to l_2 + l_1 \\ l_3 \to l_3 - 2l_1 \\ l_4 \to l_4 - 2l_1}]{} \left[\begin{array}{rrrr|r} 2 & 4 & 6 & 8 & 1 \\ 0 & 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & -8 & -8 \\ 0 & 4 & 4 & 4 & 4 \end{array}\right] \rightarrow$$

$$\xrightarrow[l_4 \to l_4 - 2l_2]{} \left[\begin{array}{rrrr|r} 2 & 4 & 6 & 8 & 1 \\ 0 & 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & -8 & -8 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

O sistema correspondente à última matriz, e equivalente ao sistema inicial, é o sistema

$$\left\{\begin{array}{rcrcrcrcr} 2x_1 & + & 4x_2 & + & 6x_3 & + & 8x_4 & = & 1 \\ & & 2x_2 & + & 2x_3 & + & 2x_4 & = & 2 \\ & & & & 2x_3 & - & 8x_4 & = & -8 \end{array}\right.$$

Resolvendo o sistema por substituição inversa, obtem-se

$$Sol_{(S)} = \{(a_1, a_2, a_3, a_4) \mid a_1 = 6a_4 + \frac{5}{2}, a_2 = -5a_4 + 5, a_3 = 4a_4 - 4\}.$$

Como vimos nos exemplos anteriores, quando se aplica o método de eliminação de Gauss na resolução de um sistema de equações lineares, a matriz ampliada do sistema é transformada, por meio de operações elementares sobre linhas, numa matriz em forma de escada. Mas, como foi referido no capítulo anterior, o processo descrito no método de eliminação de Gauss pode ser complementado com outras operações elementares, de forma a obter uma matriz equivalente por linhas à matriz ampliada e que esteja em forma de *escada reduzida*. Este método, conhecido por método de eliminação de Gauss-Jordan, quando aplicado na resolução de sistemas de equações lineares, permite obter sistemas equivalentes aos sistema dados e que, de uma forma geral, terão uma resolução ainda mais simplificada.

Exemplo 2.2.3. *Consideremos de novo o sistema indicado no Exemplo 2.2.1.*

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & + & x_3 & + & x_4 & = & 1 \\ x_1 & + & x_2 & & & & & = & 2 \\ 2x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 \\ & & x_2 & + & 2x_3 & + & x_4 & = & 2 \end{array}\right. .$$

Nesse mesmo exemplo vimos que, aplicando o método de eliminação de Gauss à matriz ampliada deste sistema, é possível obter a matriz em escada

$$\left[\begin{array}{rrrr|r} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 2 & 2 & -4 \\ 0 & 0 & 0 & -1 & 5 \end{array}\right].$$

Partindo desta última matriz e seguindo o processo descrito no Teorema 1.8.9 temos

$$\left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 2 & 2 & -4 \\ 0 & 0 & 0 & -1 & 5 \end{array}\right] \xrightarrow[\substack{l_3 \to l_3 + 2l_4 \\ l_2 \to l_2 - l_4 \\ l_1 \to l_1 + l_4}]{} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 6 \\ 0 & -1 & -1 & 0 & -4 \\ 0 & 0 & 2 & 0 & 6 \\ 0 & 0 & 0 & -1 & 5 \end{array}\right] \to$$

$$\xrightarrow[\substack{l_2 \to l_2 + \frac{1}{2}l_3 \\ l_1 \to l_1 - \frac{1}{2}l_3}]{} \left[\begin{array}{cccc|c} 1 & 2 & 0 & 0 & 3 \\ 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 2 & 0 & 6 \\ 0 & 0 & 0 & -1 & 5 \end{array}\right] \xrightarrow[l_1 \to l_1 + 2l_2]{} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 2 & 0 & 6 \\ 0 & 0 & 0 & -1 & 5 \end{array}\right] \to$$

$$\xrightarrow[\substack{l_2 \to -l_2 \\ l_3 \to \frac{1}{2}l_3 \\ l_4 \to -l_4}]{} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -5 \end{array}\right].$$

A última matriz obtida, em forma de escada reduzida, é equivalente por linhas à matriz $[A \,|\, b]$*. Por conseguinte, o sistema inicial* (S) *é equivalente ao sistema*

$$\begin{cases} x_1 = 1 \\ x_2 = 1 \\ x_3 = 3 \\ x_4 = -5 \end{cases},$$

donde obtemos $Sol_{(S)} = \{(1, 1, 3, -5)\}$.

Seguidamente debruçamo-nos sobre um outro tipo de problema que também já foi referido no início do capítulo - a discussão de sistemas de equações lineares. Como iremos ver, a discussão de um sistema de equações lineares pode ser feita recorrendo à característica da sua matriz simples e da sua matriz ampliada, sem que seja necessário determinar o seu conjunto de soluções. Aplicando o método de eliminação de Gauss à matriz ampliada do sistema, representada por $[A \,|\, b]$, podemos determinar a característica dessa matriz e a característica da matriz simples do sistema, ou seja, da matriz A. Sendo $[U|c]$ a matriz em escada obtida a partir de $[A|b]$ por aplicação do método de eliminação de Gauss, então car$([A|b])$ é o número de linhas não nulas de $[U|c]$ e car(A) é o número de linhas não nulas de U. Note-se que se tem sempre car$(A) \leq$ car$([A|b])$.

Sendo (S) um sistema a m equações lineares e n incógnitas, $[A|b]$ a matriz ampliada do sistema e designando por r a característica de A, tem-se um dos seguintes casos:

1) $r = m = n$; 2) $r = m < n$; 3) $r < m$.

No primeiro caso, a matriz $[U \,|\, c]$ tem a forma

$$[U \,|\, c] = \left[\begin{array}{cccc|c} u_{11} & u_{12} & \dots & u_{1n} & c_1 \\ 0 & u_{22} & \dots & u_{2n} & c_2 \\ 0 & 0 & \dots & u_{3n} & c_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & u_{nn} & c_n \end{array}\right]$$

que corresponde ao sistema

$$\left\{\begin{array}{ccccccccc} u_{11}x_1 & + & u_{12}x_2 & + & \dots & + & u_{1n}x_n & = & c_1 \\ & & u_{22}x_2 & + & \dots & + & u_{2n}x_n & = & c_2 \\ & & & & & & & \vdots & \\ & & & & & & u_{nn}x_n & = & c_n \end{array}\right.$$

em que $u_{ii} \neq 0$, para todo $i \in \{1, ..., n\}$.
A partir da última equação deste sistema obtemos $x_n = \frac{c_n}{u_{nn}}$ e, por substituição inversa nas equações anteriores, obtemos sucessivamente $x_{n-1}, x_{n-2}, ..., x_1$. Neste caso, o sistema é possível determinado.

No segundo caso, em que temos $r = m < n$, a matriz $[U \mid c]$ será da forma

$$[U \mid c] = \left[\begin{array}{ccccccccccc|c} 0 & \dots & u_{1j_1} & \dots & u_{1j_2} & \dots & u_{1j_3} & \dots & u_{1j_m} & \dots & u_{1n} & c_1 \\ 0 & \dots & 0 & \dots & u_{2j_2} & \dots & u_{2j_3} & \dots & u_{2j_m} & \dots & u_{2n} & c_2 \\ 0 & \dots & 0 & \dots & 0 & \dots & u_{3j_3} & \dots & u_{3j_m} & \dots & u_{3n} & c_3 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & u_{mj_m} & \dots & u_{mn} & c_n \end{array}\right]$$

onde, para todo $k \in \{1, ..., m\}$, $u_{kj_k} \neq 0$ e, para todo $j < j_k$, $u_{kj} = 0$.

No sistema correspondente a $[U \mid c]$ são livres as incógnitas x_j, com $j \in \{1, ..., n\} \setminus \{j_1, j_2, ..., j_m\}$, e obtemos as restantes incógnitas em função destas. Desta forma, o sistema é possível indeterminado.

No caso em que $r < m$, temos

$$[U \mid c] = \left[\begin{array}{ccccccccccc|c} 0 & \dots & u_{1j_1} & \dots & u_{1j_2} & \dots & u_{1j_3} & \dots & u_{1j_m} & \dots & u_{1n} & c_1 \\ 0 & \dots & 0 & \dots & u_{2j_2} & \dots & u_{2j_3} & \dots & u_{2j_m} & \dots & u_{2n} & c_2 \\ 0 & \dots & 0 & \dots & 0 & \dots & u_{3j_3} & \dots & u_{3j_m} & \dots & u_{3n} & c_3 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & u_{rj_m} & \dots & u_{rn} & c_r \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & c_{r+1} \\ . & . & . & . & . & . & . & . & . & . & . & . \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & c_m \end{array}\right]$$

Se $c_i \neq 0$, para algum $i \in \{r+1, ..., m\}$, o sistema é obviamente impossível.

Se $c_{r+1} = \cdots = c_m = 0$, então as últimas $m-r$ equações do sistema correspondente à matriz $[U|c]$ são identidades, e o sistema é equivalente ao das r primeiras equações. Este sistema está num dos casos estudados antes: de facto, se $r = n$, o sistema enquadra-se no primeiro caso, e se $r < n$, temos um sistema do tipo que foi estudado no segundo caso.

Do que foi observado conclui-se o seguinte:

Teorema 2.2.4. *Sejam $m, n \in \mathbb{N}$ e $Ax = b$ um sistema de equações lineares, com $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Então:*

- *o sistema é possível se e só se $car(A) = car([A|b])$;*
- *o sistema é possível determinado se e só se $car(A) = car([A|b]) = n$;*
- *o sistema é possível indeterminado se e só se $car(A) = car([A|b]) < n$.*

Exemplo 2.2.5. *Consideremos o sistema de 4 equações lineares em 4 incógnitas*

$$\begin{cases} + 2x_2 - x_3 = 1 \\ x_1 + x_2 = 2 \\ -x_1 + x_2 + x_3 + x_4 = 1 \\ 2x_2 + x_3 + x_4 = 2 \end{cases}$$

Aplicando o método de eliminação de Gauss à matriz ampliada do sistema, temos

$$\left[\begin{array}{rrrr|r} 0 & 2 & -1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 2 \\ -1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 & 2 \end{array}\right] \xrightarrow[l_1 \leftrightarrow l_2]{} \left[\begin{array}{rrrr|r} 1 & 1 & 0 & 0 & 2 \\ 0 & 2 & -1 & 0 & 1 \\ -1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 & 2 \end{array}\right] \rightarrow$$

$$\xrightarrow[l_3 \to l_3 + l_1]{} \left[\begin{array}{rrrr|r} 1 & 1 & 0 & 0 & 2 \\ 0 & 2 & -1 & 0 & 1 \\ 0 & 2 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 & 2 \end{array}\right] \xrightarrow[\substack{l_3 \to l_3 - l_2 \\ l_4 \to l_4 - l_2}]{} \left[\begin{array}{rrrr|r} 1 & 1 & 0 & 0 & 2 \\ 0 & 2 & -1 & 0 & 1 \\ 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 2 & 1 & 1 \end{array}\right] \rightarrow$$

$$\xrightarrow[l_4 \to l_4 - l_3]{} \left[\begin{array}{rrrr|r} 1 & 1 & 0 & 0 & 2 \\ 0 & 2 & -1 & 0 & 1 \\ 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right]$$

O sistema correspondente à última matriz, e equivalente ao inicial, é o sistema

$$\begin{cases} x_1 + x_2 = 2 \\ + 2x_2 - x_3 = 1 \\ 2x_3 + x_4 = 2 \\ 0 = -1 \end{cases}$$

que obviamente é impossível.

Exemplo 2.2.6. *Consideremos o sistema do Exemplo 2.2.1. Aplicando o método de eliminação de Gauss à matriz ampliada deste sistema obtemos a matriz*

$$\left[\begin{array}{rrrr|r} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 2 & 2 & -4 \\ 0 & 0 & 0 & -1 & 5 \end{array}\right]$$

à qual corresponde o sistema

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & + & x_3 & + & x_4 & = & 1 \\ & & -x_2 & - & x_3 & - & x_4 & = & 1 \\ & & & & 2x_3 & + & 2x_4 & = & -4 \\ & & & & & - & x_4 & = & 5 \end{array}\right.$$

Uma vez que $car(A) = car(A|b) = 4$, o sistema é possível e determinado. De facto, da última equação resulta que $x_4 = -5$ e substituindo sucessivamente nas equações anteriores, obtemos $x_3 = 3$, $x_2 = 1$, $x_1 = 1$.

Exemplo 2.2.7. *Consideremos o sistema (S) de 3 equações lineares em 3 incógnitas*

$$\left\{\begin{array}{rcrcrcr} x_1 & - & 3x_2 & + & 2x_3 & = & 1 \\ 2x_1 & & & + & 4x_3 & = & 8 \\ -x_1 & + & 4x_2 & - & 2x_3 & = & 0 \end{array}\right. .$$

Como

$$\left[\begin{array}{rrr|r} 1 & -3 & 2 & 1 \\ 2 & 0 & 4 & 8 \\ -1 & 4 & -2 & 0 \end{array}\right] \xrightarrow[\substack{l_2 \to l_2 - 2l_1 \\ l_3 \to l_3 + l_1}]{} \left[\begin{array}{rrr|r} 1 & -3 & 2 & 1 \\ 0 & 6 & 0 & 6 \\ 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow[l_3 \to l_3 - \frac{1}{6}l_2]{} \left[\begin{array}{rrr|r} 1 & -3 & 2 & 1 \\ 0 & 6 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

concluímos que $c(A) = c(A|b) = 2 < 3$, pelo que o sistema é possível indeterminado. O sistema correspondente à última matriz é dado por

$$\left\{\begin{array}{rcrcrcr} x_1 & - & 3x_2 & + & 2x_3 & = & 1 \\ & & 6x_2 & & & = & 6 \end{array}\right.$$

onde x_3 é arbitrário, $x_2 = 1$, $x_1 = 4 - 2x_3$. Assim,

$$Sol_{(S)} = \{(4 - 2a, 1, a) \in \mathbb{R}^3 : a \in \mathbb{R}\}.$$

No caso particular dos sistemas homogéneos, já havíamos observado que estes sistemas são sempre possíveis. Agora, com base no teorema anterior, sabe-se também que o sistema $Ax = 0$, com $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, é determinado se e só $car(A) = n$.

Definição 2.2.8. *Sejam $m, n \in \mathbb{N}$ e $Ax = b$, com $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, um sistema de equações lineares. Dá-se a designação de* ***sistema homogéneo associado a*** *$Ax = b$ ao sistema $Ax = 0$.*

O conjunto de soluções de um sistema e do sistema homogéneo associado estão relacionados de acordo com o estabelecido no teorema seguinte.

Teorema 2.2.9. *Sejam $m, n \in \mathbb{N}$, $Ax = b$ um sistema de equações lineares, com $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, e $y \in \mathcal{M}_{n\times 1}(\mathbb{K})$ uma solução do sistema. Então $w \in \mathcal{M}_{n\times 1}(\mathbb{K})$ é solução de $Ax = b$ se e só se $w = y + z$, onde $z \in \mathcal{M}_{n\times 1}(\mathbb{K})$ é solução de $Ax = 0$.*

Demonstração. Seja $y \in \mathcal{M}_{n\times 1}(\mathbb{K})$ uma solução de $Ax = b$.
Suponhamos que $z \in \mathcal{M}_{n\times 1}(\mathbb{K})$ é uma solução do sistema homogéneo $Ax = 0$. Então

$$A\,(y+z) = Ay + Az = b + 0 = b,$$

pelo que $w = y + z$ é solução de $Ax = b$.
Reciprocamente, suponhamos que w é uma solução de $Ax = b$. Então

$$A\,(w-y) = Aw - Ay = b - b = 0.$$

Por conseguinte, w pode escrever-se como a soma de y com uma solução do sistema $Ax = 0$, pois $w = y + (w - y)$ onde $w - y$ é solução de $Ax = 0$. □

Do teorema anterior segue de imediato o resultado seguinte.

Corolário 2.2.10. *Sejam $m, n \in \mathbb{N}$ e $Ax = b$ um sistema de equações lineares, com $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Então, se for possível, o sistema $Ax = b$ é determinado se e só se o sistema homogéneo associado é determinado (i.e., admite como única solução $0_{n\times 1}$).*

2.3 Inversão de matrizes

A determinação da inversa de uma matriz, caso exista, corresponde à resolução de um sistema de equações lineares. Assim sendo, o método de eliminação de Gauss-Jordan, para além de nos facultar um algoritmo para a resolução de sistemas, dá-nos também um processo para o cálculo da inversa de matrizes invertíveis.

Antes de vermos de que forma o método de eliminação de Gauss-Jordan pode ser aplicado no cálculo da inversa de uma matriz, apresentamos uma caracterização de matrizes invertíveis que nos permite decidir sobre a invertibilidade de uma matriz através do cálculo da sua característica.

Teorema 2.3.1. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Então A é invertível se e só se $car(A) = n$.*

Demonstração. Suponha-se que A é invertível. Então

$$Ax = 0 \Rightarrow A^{-1}(Ax) = 0 \Rightarrow (A^{-1}A)x = 0 \Rightarrow x = 0$$

Logo, como o sistema $Ax = 0$ é determinado, temos $car(A) = n$.

Reciprocamente, suponhamos que $car(A) = n$ e mostremos que A é invertível. Por definição, uma matriz $A \in \mathcal{M}_n(\mathbb{K})$ é invertível se existir uma matriz $X = [x_{ij}] \in \mathcal{M}_n(\mathbb{K})$ tal que $AX = I_n = XA$. A existência de $X \in \mathcal{M}_n(\mathbb{K})$ tal que $AX = I_n$ é equivalente à existência de

$$x_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix}, \quad x_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix}, \quad \ldots, \quad x_n = \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{bmatrix}$$

tais que

$$Ax_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad Ax_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad Ax_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad (*)$$

Como $car(A) = n$, cada um dos sistemas indicado em $(*)$ é possível e determinado, o que significa que existe $X \in \mathcal{M}_n(\mathbb{K})$ tal que $AX = I_n$. Para podermos concluir que X é a inversa de A falta verificar que $XA = I_n$. Para tal, comecemos por mostrar que se $B \in \mathcal{M}_{n\times n}(\mathbb{K})$ é uma matriz tal que $AB = 0$, então $B = 0$. De facto, se representarmos por b_i a coluna i de B, $i \in \{1, ..., n\}$, então de $AB = 0$ segue que, para todo $i \in \{1, ..., n\}$, $Ab_i = 0$. Como $car(A) = n$, o sistema $Ax = 0$ é determinado e, portanto, $b_i = 0$, para todo $i \in \{1, ..., n\}$. Logo, $B = 0$.
Assim, de

$$A(XA - I_n) = A(XA) - A = (AX)A - A = I_nA - A = 0,$$

concluímos que $XA - I_n = 0$, i.e., $XA = I_n$. Logo, A é invertível. □

Já sabemos, pelo Teorema 1.3.8, que se $A, X \in \mathcal{M}_n(\mathbb{K})$ são matrizes tais que A é invertível e $AX = I_n$, então $XA = I_n$. Logo, sendo A uma matriz invertível, para determinarmos a sua inversa basta resolver os sistemas referidos em $(*)$. Além disso, uma vez que a matriz simples destes sistemas é a mesma, podemos optar pela resolução de todos os sistemas em simultâneo. Tal é conseguido aplicando o método de eliminação de Gauss-Jordan à matriz $[A \,|\, I_n]$; a matriz obtida após a aplicação deste método é a matriz $[I_n \,|A^{-1}]$.

Exemplo 2.3.2. *Seja*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 5 & 6 & 14 \end{bmatrix}.$$

Então, de

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 5 & 6 & 14 & 0 & 0 & 1 \end{array}\right] \xrightarrow[l_3 \to l_3 - 5l_1]{} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & -4 & -1 & -5 & 0 & 1 \end{array}\right] \xrightarrow[l_3 \to l_3 + 2l_2]{}$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & 0 & 3 & -5 & 2 & 1 \end{array}\right] \xrightarrow[\substack{l_1 \to l_1 - l_3 \\ l_2 \to l_2 - \frac{2}{3}l_3}]{} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 6 & -2 & -1 \\ 0 & 2 & 0 & \frac{10}{3} & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 0 & 3 & -5 & 2 & 1 \end{array}\right] \to$$

$$\xrightarrow[l_1 \to l_1 - l_2]{} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{8}{3} & -\frac{5}{3} & -\frac{1}{3} \\ 0 & 2 & 0 & \frac{10}{3} & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 0 & 3 & -5 & 2 & 1 \end{array}\right] \xrightarrow[\substack{l_2 \to \frac{1}{2}l_2 \\ l_3 \to \frac{1}{3}l_3}]{} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{18}{3} & -\frac{5}{3} & -\frac{1}{3} \\ 0 & 1 & 0 & \frac{10}{6} & -\frac{1}{6} & -\frac{2}{6} \\ 0 & 0 & 1 & -\frac{5}{3} & \frac{2}{3} & \frac{1}{3} \end{array}\right],$$

concluímos que a matriz A admite inversa e que

$$A^{-1} = \begin{bmatrix} \frac{18}{3} & -\frac{5}{3} & -\frac{1}{3} \\ \frac{10}{6} & -\frac{1}{6} & -\frac{2}{6} \\ -\frac{5}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

Sejam $A, B \in \mathcal{M}_n(\mathbb{K})$. Já foi estabelecido anteriormente que se A é uma matriz invertível e $AB = I_n$, então $BA = I_n$. O teorema seguinte generaliza este resultado.

Teorema 2.3.3. *Sejam $n \in \mathbb{N}$ e $A, B \in \mathcal{M}_n(\mathbb{K})$. Se $AB = I_n$, então A e B são matrizes invertíveis, $B^{-1} = A$ e $BA = I_n$.*

Demonstração. Admitamos que AB é invertível. Então

$$Bx = 0 \Rightarrow ABx = 0 \Rightarrow x = 0.$$

Assim, o sistema $Bx = 0$ é possível e determinado, pelo que $car(B) = n$. Logo, B é invertível. Provemos, agora, que A é invertível. De facto, como $AB = I_n$ e B é invertível, tem-se

$$A = AI_n = A(BB^{-1}) = (AB)B^{-1} = I_nB^{-1} = B^{-1}.$$

Como a matriz B^{-1} é invertível, então A é uma matriz invertível.
Considerando que $B^{-1} = A$, também se tem $BA = I_n$. $\square$

Teorema 2.3.4. *Sejam $n \in \mathbb{N}$ e $A, B \in \mathcal{M}_n(\mathbb{K})$. Então AB é invertível se e só se A e B são invertíveis.*

Demonstração. Foi provado, no Teorema 1.3.10, que se A e B são invertíveis, então AB é invertível, pelo que resta provar a implicação contrária.
Para provar o resultado recíproco, comecemos por mostar que se AB é invertível, então A é invertível. No sentido de se fazer esta prova por redução ao absurdo, admitamos que AB é invertível e que A não é invertível. Então

$$car(AB) = n \text{ e } car(A) < n.$$

Sejam A' uma matriz em forma de escada equivalente por linhas à matriz A e $E_p, \ldots, E_1$ matrizes elementares tais que $A' = E_p \cdots E_1A$. Logo $A'B = E_p \cdots E_1AB$, donde segue que

$$car(AB) = car(E_1 \cdots E_p(AB)) = car((E_p \cdots E_1A)B) = car(A'B).$$

Como $car(AB) = n$, também se tem $car(A'B) = n$. No entanto, como $car(A) < n$, a matriz A' tem, pelo menos uma linha nula, pelo que $A'B$ também tem pelo menos uma linha nula e, portanto, $car(A'B) < n$ (contradição).

Mostremos, agora, que se AB é invertível, então B é invertível. De facto, como AB é invertível, existe $X \in \mathcal{M}_n(\mathbb{K})$ tal que $(AB)X = I_n$. Então $A(BX) = I_n$ e do resultado anterior segue que BX é invertível. Consequentemente, pelo que foi provado anteriormente, B é invertível. $\square$

3. Espaços Vetoriais $\mathbb{R}^n$

Muitas estruturas matemáticas têm propriedades em comum que podem ser generalizadas através de um estudo mais abstrato. Um dos principais objetos de estudo da Álgebra Linear são os *Espaços Vetoriais*. Esta noção é uma generalização da estrutura associada ao conjunto formado por todos os segmentos orientados com origem num determinado ponto, munido das operações de adição de segmentos orientados e multiplicação de um escalar por um segmento. Os elementos deste conjunto são conhecidos por vetores, termo que que dá origem à denominação *espaço vetorial*. Além disso, os pontos e segmentos orientados no plano são geralmente identificados com elementos de $\mathbb{R}^2$, onde a adição de vetores de $\mathbb{R}^2$ e a multiplicação de números reais por vetores de $\mathbb{R}^2$ possuem uma interpretação geométrica natural. De modo análogo, no espaço tridimensional, os pontos e segmentos orientados correspondem a elementos de $\mathbb{R}^3$, e as operações de adição de vetores e multiplicação de escalares por vetores também têm significado geométrico. Estas duas estruturas servem de motivação para o estudo dos chamados espaços vetorias $\mathbb{R}^n$.

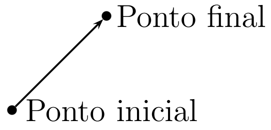
3.1 Dos segmentos orientados aos espaços vetoriais $\mathbb{R}^n$

As quantidades físicas mais simples ficam completamente definidas pela sua magnitude, ou seja, por um único valor numérico, eventualmente associado a uma unidade de medida adequada. A uma quantidade deste tipo dá-se a designação de **escalar**. Quantidades tais como temperatura, comprimento e volume são exemplos de escalares.

Há outro tipo de quantidades físicas, tais como força e deslocamento, que só ficam completamente definidas se, para além da sua magnitude, for indicada uma direção e um sentido; a este tipo de quantidade dá-se a designação de **vetor**.
A magnitude de um vetor é especificada por um número real não negativo que pode estar associado a uma unidade de medida (a magnitude dos vetores e dos escalares também pode ser especificada por um número complexo, mas neste curso consideraremos apenas números reais). A direção de um vetor é especificada pela sua relação com determinadas linhas e/ou planos de referência. O sentido do vetor é determinado pela ordem de dois pontos.

No plano ou no espaço, os vetores podem ser representados geometricamente por setas: o comprimento da seta é proporcional à magnitude do vetor e a direção e o sentido da seta indicam a direção e o sentido do vetor.

A origem da seta é designada **ponto inicial** do vetor e a extremidade da seta é o **ponto final** do vetor.



Se um vetor **v** tem ponto inicial A e ponto final B, denotamos o vetor por $v = \overrightarrow{AB}$ quando pretendemos explicitar os pontos inicial e final.

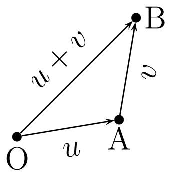
Nas aplicações ocorrem dois tipos de vetores: os vetores **fixos** e os **livres**. Um vetor **fixo** é um vetor cujo efeito físico depende da localização do ponto inicial, além da magnitude, direção e sentido. Um vetor **livre** é um vetor cujo efeito físico depende apenas da magnitude, direção e sentido. Neste curso trataremos apenas de vetores livres e passamos a denominá-los apenas de vetores. Atendendo a que os vetores livres não são afetados por translação, dizemos que dois vetores u e v são **iguais** (ou **equivalentes**) se têm a mesma magnitude, a mesma direção e o mesmo sentido.

O vetor cujos ponto inicial e ponto terminal coincidem tem comprimento zero, pelo que denominaremos este vetor de **vetor zero** ou **vetor nulo**. O vetor nulo não tem direção nem sentido e, portanto, convencionamos que tem a direção ou o sentido que forem convenientes.

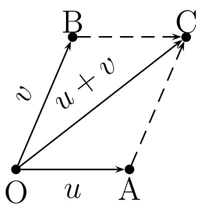
Existem várias operações que podem ser efetuadas com vetores.

Adição de vetores

Sejam O, A e B são pontos do plano ou do espaço. Supondo que $u = \overrightarrow{OA}$ e $v = \overrightarrow{AB}$, então a soma dos vetores u e v, representada por $u + v$, corresponde ao vetor $\overrightarrow{OB}$.



À regra de adição de vetores apresentada anteriormente dá-se a designação da *Regra do Triângulo para a Adição Vetorial.*
De forma equivalente, a adição de vetores pode ser definida com base na *Regra do Paralelograma para a Adição Vetorial*: se O, A, B, C são pontos do plano ou do espaço e u e v são vetores tais $u = \overrightarrow{OA} = \overrightarrow{BC}$ e $v = \overrightarrow{OB} = \overrightarrow{AC}$, então $u + v = \overrightarrow{OC}$.

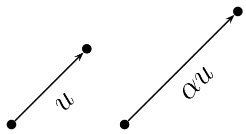


Mutiplicação de um escalar por um vetor

Sejam $\alpha \in \mathbb{R}$ e u um vetor do plano ou do espaço. O produto do escalar α pelo vetor u é o vetor αu definido da seguinte forma:

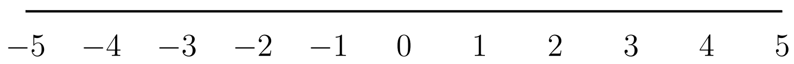
(a) se $\alpha = 0$ ou $u = 0$, então $\alpha u = 0$;

(b) se $\alpha \neq 0$ e $u \neq 0$, αu é o vetor com a mesma direção de u, mas cuja magnitude é $|\alpha|$ vezes a magnitude de u e cujo sentido é o mesmo de u se $\alpha > 0$ e o oposto de u se $\alpha < 0$.



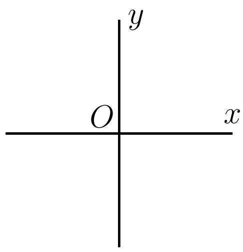
Embora a representação geométrica seja extremamente útil para a intuição e justificação de diversos conceitos vetoriais, em muitas situações torna-se necessário descrever os vetores de forma algebricamente precisa, de modo a possibilitar o estudo sistemático da álgebra vetorial por métodos analíticos.

A descrição analítica dos vetores resulta de forma natural da descrição geométrica, desde que se introduza um sistema de coordenadas. A ideia de utilizar um número real para localizar um ponto sobre uma reta já era conhecida pelos antigos gregos.

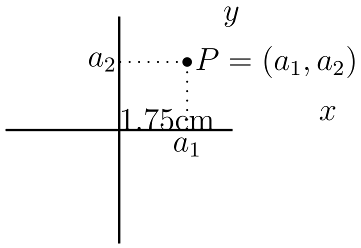


Cada número real é representado por exatamente um ponto P na reta e cada ponto P da reta representa um e um só número real.

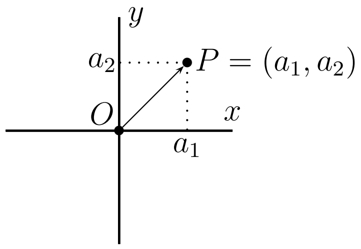
Em 1637, Descartes generalizou a ideia anterior, utilizando um par de números reais para localizar um ponto P no plano. Para tal, foram definidos sistemas de coordenadas cartesianas. Um **sistema de coordenadas cartesianas do plano** consiste de duas retas perpendiculares que usualmente são designadas por **eixo x** e **eixo y**.



O ponto de interseção dos dois eixos, representado por O, é denominado a **origem** do sistema de coordenadas.
Denotando por $\mathbb{R}^2$ o conjunto $\{(x, y) : x, y \in \mathbb{R}\}$ de todos os pares de números reais e considerando um sistema de coordenadas do plano, podemos estabelecer uma correspondência bijetiva entre os pontos do plano e $\mathbb{R}^2$; a cada ponto P está associado um e um só par ordenado (a_1, a_2) e, reciprocamente, a cada par de números reais está associado um e um só ponto.

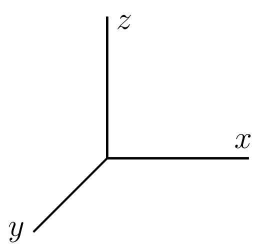


Aos números a_1, a_2 damos a designação de **coordenadas** de P e escrevemos $P = (a_1, a_2)$ para indicar as coordenadas de P no plano.
Um par de números reais pode ter outras interpretações, em particular poderá ser usado para representar um deslocamento. Se um vetor v representar um deslocamento com início na origem de um sistema de coordenadas cartesianas no plano e com término num ponto P, o vetor v fica completamente determinado pelas coordenadas do ponto P. Ao vetor v, usualmente representado por $\overrightarrow{OP}$, dá-se a designação de **vetor posição de** P.

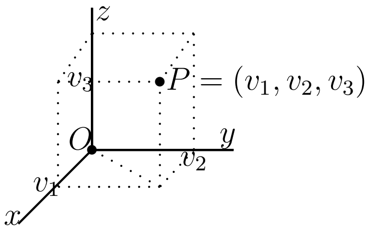


Se v é o vetor posição de $P = (a_1, a_2)$, isto é, se $v = \overrightarrow{OP}$, escrevemos $v = (a_1, a_2)$ para identificar o vetor. De um modo geral, qualquer deslocamento entre dois pontos de um plano pode ser representado por um par de números reais; este par deverá medir o deslocamento que é efetuado em cada uma das direções. Se A e B são pontos cujas coordenadas no sistema de coordenadas cartesianas são, respetivamente, $A = (a_1, a_2)$ e $B = (b_1, b_2)$, então o vetor $\overrightarrow{AB}$, correspondente ao deslocamento de A até B, é representado pelo par $(b_1 - a_1, b_2 - a_2)$, e escreve-se $\overrightarrow{AB} = (b_1 - a_1, b_2 - a_2)$. Dois deslocamentos que tenham a mesma direção, sentido e magnitude são considerados equivalentes, pelo que o deslocamento correspondente a $\overrightarrow{AB}$ é equivalente ao deslocamento com início na origem do sistema de coordenadas e término no ponto de coordenadas $(b_1 - a_1, b_2 - a_2)$.

Analogamente ao que sucede no espaço bidimensional, todo o triplo de números reais (a_1, a_2, a_3) pode ser identificado com um ponto P no espaço tridimensional. Neste caso consideramos um **sistema de coordenadas cartesianas do espaço** definido por três eixos perpendiculares, geralmente designados por **eixo** x, **eixo** y e **eixo** z.



Desta forma, é possível estabelecer uma bijeção entre o conjunto de todos os pontos do espaço e o conjunto $\mathbb{R}^3 = \{(a_1, a_2, a_3) : a_1, a_2, a_3 \in \mathbb{R}\}$ de todos os triplos de números reais; cada ponto P do espaço é identificado com um e um só triplo de números reais (a_1, a_2, a_3). Aos números a_1, a_2, a_3 damos a designação de **coordenadas** de P e escrevemos $P = (a_1, a_2, a_3)$ para indicar as coordenadas de P no espaço.



Dados dois pontos A e B no espaço, de coordenadas (a_1, a_2, a_3) e (b_1, b_2, b_3), o deslocamento de A até B também pode ser representado por um triplo de reais, mais precisamente, por $(b_1 - a_1, b_2 - a_2, b_3 - a_3)$.

O uso de pares e triplos de números reais para representar pontos e vetores no plano e no espaço já era comum nos séculos XVIII e XIX. No entanto, no final do século XIX e início do século XX, matemáticos e físicos passaram a reconhecer a importância de considerar espaços de dimensão superior. Um exemplo notável é o trabalho de Albert Einstein, que adicionou uma componente temporal t às três componentes espaciais (x, y, z), formando assim um quádruplo (x, y, z, t) para descrever um ponto no universo espaço-tempo de quatro dimensões. Matemáticos como A. Cayley e H. G. Grassmann também fizeram referência a espaços de dimensão superior e recorreram a n-uplos $(a_1, a_2, \ldots, a_n)$ de números reais para localizar pontos em um espaço n-dimensional. Essas sequências ordenadas de números reais podem ser interpretadas não apenas como posições de pontos, mas também como vetores de deslocamento ou posição, além de servirem para modelar diversos fenômenos em áreas como a economia.

3.2 Espaços vetoriais reais $\mathbb{R}^n$

Nesta secção, introduzimos os espaços vetoriais $\mathbb{R}^n$ que constituem o modelo algébrico fundamental para o estudo dos vetores em qualquer dimensão. Nesses espaços, as operações vetoriais são definidas de forma puramente numérica, o que permite desenvolver uma teoria unificada e independente da interpretação geométrica, mas que preserva toda a intuição construída a partir da representação de vetores como segmentos orientados no plano e no espaço.

3.2.1 Definição

Dado $n \in \mathbb{N}$, define-se

$$\mathbb{R}^n = \{(x_1, x_2, ..., x_n) : x_i \in \mathbb{R}, \forall i = 1, 2, ..., n\}$$

como o conjunto dos n-uplos ordenados de elementos de $\mathbb{R}$.

O vetor $(0, 0, \ldots, 0)$ de $\mathbb{R}^n$ é designado por **vetor nulo** de $\mathbb{R}^n$ e é representado por $0_{\mathbb{R}^n}$.

Teorema 3.2.1. *Sejam* $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ *e* $\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ *as operações definidas, respetivamente, por*

- $x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$,
  para todos $x = (x_1, x_2, ..., x_n)\,, y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$,
- $\alpha \cdot x = (\alpha x_1, \alpha x_2, ..., \alpha x_n)$,
  para todos $\alpha \in \mathbb{R}$ *e* $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$.

Então, são válidas as seguintes propriedades:

(1) $\forall_{x,y\in\mathbb{R}^n} \quad x + y = y + x$;

(2) $\forall_{x,y,z\in\mathbb{R}^n} \quad x + (y + z) = (x + y) + z$;

(3) $\forall_{x\in\mathbb{R}^n} \quad x + 0_{\mathbb{R}^n} = x = 0_{\mathbb{R}^n} + x$;

(4) $\forall_{x\in\mathbb{R}^n}\, \exists_{x'\in\mathbb{R}^n} \quad x + x' = 0_{\mathbb{R}^n} = x' + x$;

(5) $\forall_{x,y\in\mathbb{R}^n}\, \forall_{\alpha\in\mathbb{R}} \quad \alpha \cdot (x + y) = \alpha \cdot x + \alpha{\cdot}y$;

(6) $\forall_{x\in\mathbb{R}^n}\, \forall_{\alpha,\beta\in\mathbb{R}} \quad (\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$;

(7) $\forall_{x\in\mathbb{R}^n}\, \forall_{\alpha,\beta\in\mathbb{R}} \quad (\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)\,$;

(8) $\forall_{x\in\mathbb{R}^n} \quad 1 \cdot x = x$.

Definição 3.2.2. *Ao quádruplo* $(\mathbb{R}^n, +, \cdot, \mathbb{R})$*, onde* $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ *e* $\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ *são as operações definidas, respetivamente, por*

- $x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$,
  para todos $x = (x_1, x_2, ..., x_n)\,, y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$,
- $\alpha \cdot x = (\alpha x_1, \alpha x_2, ..., \alpha x_n)$,
  para todos $\alpha \in \mathbb{R}$ *e* $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$,

dá-se a designação de ***espaço vetorial real*** $\mathbb{R}^n$ *(ou* ***espaço vetorial*** $\mathbb{R}^n$*).*
Os elementos de $\mathbb{R}^n$ *são designados por* ***vetores*** *e os elementos de* $\mathbb{R}$ *designam-se por* ***escalares****.*

Terminologia e notação: Consideremos o quádruplo $(\mathbb{R}^n, +, \cdot, \mathbb{R})$ definido anteriormente.

- Não havendo ambiguidade, o espaço vetorial $(\mathbb{R}^n, +, \cdot, \mathbb{R})$ será referido apenas por $\mathbb{R}^n$.
- A operação $+$ definida em $\mathbb{R}^n$ designa-se por ***adição de vetores*** e a operação $\cdot$ por ***multiplicação de um escalar por um vetor***.
- Para cada $x \in \mathbb{R}^n$ e para cada $\alpha \in \mathbb{R}$, é usual escrever αx para representar $\alpha \cdot x$.
- Para cada $x \in \mathbb{R}^n$, ao elemento x' determinado na condição (4), chama-se ***simétrico de*** x e representa-se por $-x$.
- Dados $x, y \in \mathbb{R}^n$, representa-se por $x - y$ o vetor $x + (-y)$.

3.2.2 Subespaços vetoriais

Definição 3.2.3. *Um subconjunto V de $\mathbb{R}^n$ diz-se um* ***subespaço vetorial*** *do espaço vetorial $\mathbb{R}^n$, e escreve-se $V \leq \mathbb{R}^n$, se:*

i) $V \neq \emptyset$; *ii)* $\forall_{x,y \in V}$, $x + y \in V$; *iii)* $\forall_{x \in V}$, $\forall_{\alpha \in \mathbb{R}}$, $\alpha \cdot x \in V$.

Terminologia: Um suconjunto de $\mathbb{R}^n$ que satisfaça a condição *ii)* diz-se **fechado para a adição de vetores**. A condição *iii)* descreve-se dizendo que o conjunto V é **fechado para a multiplicação de números reais por vetores**.

Teorema 3.2.4. *Seja V um subespaço vetorial do espaço vetorial $\mathbb{R}^n$. Então $0_{\mathbb{R}^n} \in V$.*

Demonstração. Se V é um subespaço vetorial do espaço vetorial $\mathbb{R}^n$, então existe $x \in V$, pelo que, pelas condições *ii)* e *iii)*, $x + (-x) = 0_{\mathbb{R}^n} \in V$. □

Observação: Caso um subconjunto de $\mathbb{R}^n$ não contenha o vetor nulo de $\mathbb{R}^n$, ele não poderá ser um subespaço de $\mathbb{R}^n$. Assim, a definição anterior pode ser enunciada de forma equivalente substituindo a condição $V \neq \emptyset$ por $0_{\mathbb{R}^n} \in V$.

Exemplo 3.2.5. *O conjunto $V = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}$ é subespaço vetorial do espaço vetorial $\mathbb{R}^2$. De facto, V é um subconjunto de $\mathbb{R}^2$ e*
i) $V \neq \emptyset$, pois $(0,0) \in V$;
ii) para quaisquer $x = (x_1, x_2), y = (y_1, y_2) \in V$, temos $x + y \in V$, uma vez que $x_1, y_1 \in \mathbb{R}$, $x_2 = 0$ e $y_2 = 0$, pelo que $x + y = (x_1 + y_1, x_2 + y_2) \in \mathbb{R}^2$ e $x_2 + y_2 = 0 + 0 = 0$;
iii) para quaisquer $x = (x_1, x_2) \in V$ e $\alpha \in \mathbb{R}$, temos $\alpha \cdot x \in V$, pois $x \in \mathbb{R}$ e $x_2 = 0$, pelo que $\alpha \cdot x = (\alpha \cdot x_1, \alpha \cdot x_2) \in \mathbb{R}^2$ e $\alpha \cdot x_2 = \alpha \cdot 0 = 0$.

Exemplo 3.2.6. *O conjunto $V = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 2\}$ não é subespaço vetorial do espaço vetorial $\mathbb{R}^2$, pois existem $x = (1, 2)$ e $y = (0, 2)$ tais que $x, y \in V$ e $x + y = (1, 4) \notin V$.*

Seguidamente, vamos estudar formas de construir subespaços vetoriais do espaço vetorial $\mathbb{R}^n$ a partir de outros subsespaços vetoriais dados.

Exemplo 3.2.7. *No espaço vetorial $\mathbb{R}^3$, consideremos os subespaços vetoriais*

$$V = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\} \quad e \quad W = \{(x, y, z) \in \mathbb{R}^3 : y = 2z\}.$$

Então

$$\begin{aligned} V \cap W &= \{(x, y, z) \in \mathbb{R}^3 : (x, y, z) \in V \text{ e } (x, y, z) \in W\} \\ &= \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0 \text{ e } y = 2z\} \\ &= \{(x, y, z) \in \mathbb{R}^3 : x = -3z \text{ e } y = 2z\} \\ &= \{(-3z, 2z, z) \in \mathbb{R}^3 : z \in \mathbb{R}\} \end{aligned}$$

e é simples verificar que $V \cap W$ é um subespaço vetorial de $\mathbb{R}^3$.

De forma geral, a interseção de quaisquer dois subespaços vetoriais de um espaço vetorial $\mathbb{R}^n$ é também um subespaço vetorial de $\mathbb{R}^n$.

Teorema 3.2.8. *Sejam V e W subespaços vetoriais do espaço vetorial $\mathbb{R}^n$. Então $V \cap W$ é um subespaço vetorial de $\mathbb{R}^n$.*

Demonstração. Sejam V, W subespaços vetoriais do espaço vetorial $\mathbb{R}^n$. Como V e W são subespaços vetoriais de $\mathbb{R}^n$, tem-se $V, W \subseteq \mathbb{R}^n$, pelo que $V \cap W \subseteq \mathbb{R}^n$. Além disso, verifica-se que:

i) $V \cap W \neq \varnothing$, pois V e W são subespaços vetoriais de $\mathbb{R}^n$, pelo que $0_{\mathbb{R}^n} \in V$ e $0_{\mathbb{R}^n} \in W$ e, portanto, $0_{\mathbb{R}^n} \in V \cap W$.

ii) Para quaisquer $x, y \in V \cap W$, tem-se $x + y \in V \cap W$. De facto, se $x, y \in V \cap W$, então $x, y \in V$ e $x, y \in W$. Logo, $x + y \in V$ e $x + y \in W$, uma vez que V e W são subespaços vetoriais de $\mathbb{R}^n$ e, portanto, $x + y \in V \cap W$.

iii) Para quaisquer $x \in V \cap W$ e $\alpha \in \mathbb{R}$, $\alpha x \in V \cap W$. Se $x \in V \cap W$, então $x \in V$ e $x \in W$. Logo, atendendo a que V e W são subespaços vetoriais de $\mathbb{R}^n$, tem-se $\alpha x \in V$ e $\alpha x \in W$. Assim, $\alpha x \in V \cap W$.

De *i)*, *ii)* e *iii)* conclui-se que $V \cap W$ é subespaço vetorial de $\mathbb{R}^n$. □

O resultado anterior pode ser generalizado a qualquer família não vazia de subuniuniversos de $\mathbb{R}^n$.

Teorema 3.2.9. *Seja I um conjunto. Se $\{V_i : i \in I\}$ é uma família não vazia de subespaços vetoriais do espaço vetorial $\mathbb{R}^n$, então $\bigcap_{i \in I} V_i$ é um subespaço vetorial de $\mathbb{R}^n$.*

Demonstração. Exercício. □

Na sequência dos resultados anteriores, coloca-se a questão se a união de dois subespaços vetoriais de um espaço vetorial $\mathbb{R}^n$ também será um subespaço vetorial de $\mathbb{R}^n$.

Exemplo 3.2.10. *Considerando, novamente, os subespaços vetoriais*

$$V = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\} \quad W = \{(x, y, z) \in \mathbb{R}^3 : y = 2z\}$$

do espaço vetorial $\mathbb{R}^3$, tem-se

$$V \cup W \;=\; \{(x, y, z) \in \mathbb{R}^3 : (x, y, z) \in V \text{ ou } (x, y, z) \in W\}$$

Facilmente se verifica que $V \cup W$ não é um subespaço vetorial de $\mathbb{R}^3$, uma vez que $(2, 0, -2) \in V \subseteq V \cup W$, $(0, 4, 2) \in W \subseteq V \cup W$, mas $(2, 0, -2) + (0, 4, 2) = (2, 4, 0) \notin V \cup W$.

Como mostra o exemplo anterior, a união de dois subespaços vetoriais de um espaço vetorial real $\mathbb{R}^n$ nem sempre é um subespaço vetorial de $\mathbb{R}^n$. De facto, tal só se verifica nas condições seguintes.

Teorema 3.2.11. *Sejam V, W subespaços vetoriais do espaço vetorial $\mathbb{R}^n$. Então $V \cup W$ é um subespaço vetorial de $\mathbb{R}^n$ se e só se $V \subseteq W$ ou $W \subseteq V$.*

Demonstração. Suponhamos que $V \subseteq W$ ou que $W \subseteq V$. Então $V \cup W = W$ ou $V \cup W = V$, respetivamente. Logo, $V \cup W$ é um subespaço vetorial de $\mathbb{R}^n$.
Reciprocamente, admitamos que $V \cup W$ é um subespaço vetorial de $\mathbb{R}^n$ e mostremos que $V \subseteq W$ ou $W \subseteq V$. No sentido de fazermos esta prova, admitamos que $V \nsubseteq W$. Então existe $v \in V$ tal que $v \notin W$. Como $v \in V$, então $v \in V \cup W$. Para todo $w \in W$, também temos $w \in V \cup W$. Sendo $V \cup W$ um subespaço vetorial de $\mathbb{R}^n$, segue que $w + v \in V \cup W$. Logo, $w + v \in V$ ou $w + v \in W$. Como $-v \in V$ e $-w \in W$, resulta que $(w + v) + (-v) \in V$ ou $(-w) + (w + v) \in W$, ou seja, $w \in V$ ou $v \in W$. Atendendo a que $v \notin W$, temos, então, $w \in V$. Fica assim provado que todo o elemento de W é também elemento de V, ou seja, provou-se que $W \subseteq V$. $\square$

Definição 3.2.12. *Sejam V, W subespaços vetoriais do espaço vetorial $\mathbb{R}^n$. Designa-se por **soma dos subespaços** V e W, e representa-se por $V + W$, o conjunto $\{v + w \in \mathbb{R}^n : v \in V \text{ e } w \in W\}$.*

Observação:

- Se V e W são subespaços vetoriais do espaço vetorial $\mathbb{R}^n$, então $V \subseteq V + W$. De facto, se W é subespaço vetorial de $\mathbb{R}^n$, tem-se $0_{\mathbb{R}^n} \in W$. Então, considerando que, para todo $v \in V$, $v = v + 0_{\mathbb{R}^n}$, temos $v \in V + W$. Assim, fica provado que $V \subseteq V + W$. De modo análogo, prova-se que $W \subseteq V + W$.
- Da definição de soma de subespaços vetoriais e da comutatividade da adição de vetores também é imediato que, sendo V e W subespaços vetoriais do espaço vetorial $\mathbb{R}^n$, temos $V + W = W + V$.

Exemplo 3.2.13. *Consideremos, no espaço vetorial real $\mathbb{R}^4$, os subespaços*

$$\begin{array}{rcl} U & = & \{(s,t,u,v) \in \mathbb{R}^4 : s = 0, u = 0\}, \\ V & = & \{(a,b,c,d) \in \mathbb{R}^4 : b + d = 0 \text{ e } b - d = 0\}, \\ W & = & \{(x,0,y,2x) \in \mathbb{R}^4 : x, y \in \mathbb{R}\}. \end{array}$$

Então

$$U = \{(0,t,0,v) \in \mathbb{R}^4 : t, v \in \mathbb{R}\} \quad \text{e} \quad V = \{(a,0,c,0) \in \mathbb{R}^4 : a, c \in \mathbb{R}\}$$

e, portanto,

$$\begin{array}{rcl} V + U & = & \{(a,t,c,v) \in \mathbb{R}^4 : a, c, t, v \in \mathbb{R}\}, \\ V + W & = & \{(a + x, 0, c + y, 2x) \in \mathbb{R}^4 : a, c, x, y \in \mathbb{R}\}, \\ U + W & = & \{(x, t, y, v + 2x) \in \mathbb{R}^4 : x, t, v, y \in \mathbb{R}\}. \end{array}$$

Facilmente se verifica que qualquer um destes conjuntos é um subespaço vetorial do espaço vetorial real $\mathbb{R}^4$.

Generalizando o exemplo anterior, prova-se que a soma de quaisquer dois subespaços de um espaço vetorial $\mathbb{R}^n$ é ainda um subespaço vetorial de $\mathbb{R}^n$.

Teorema 3.2.14. *Sejam V e W subespaços vetoriais do espaço vetorial $\mathbb{R}^n$. Então $V+W$ é um subespaço vetorial de $\mathbb{R}^n$.*

Demonstração. Sejam V e W subespaços vetoriais do espaço vetorial $\mathbb{R}^n$. Então $V+W \subseteq \mathbb{R}^n$, pois, para qualquer $x \in V+W$, tem-se $x = v+w$, com $v \in V$ e $w \in W$. Por conseguinte, como $V, W \subseteq \mathbb{R}^n$, segue que $v, w \in \mathbb{R}^n$. Logo, $v+w \in \mathbb{R}^n$. O conjunto $V+W$ também satisfaz as condições *1)*, *2)* e *3)* da definição de subespaço vetorial:

i) $V+W \neq \emptyset$, uma vez que $V \neq \emptyset$ e $W \neq \emptyset$.

ii) Dados $x, y \in V+W$, tem-se

$$x = v_1 + w_1 \text{ e } y = v_2 + w_2, \text{ com } v_1, v_2 \in V \text{ e } w_1, w_2 \in W.$$

Logo, recorrendo às propriedades associativa e comutativa da adição de vetores,

$$x + y = (v_1 + w_1) + (v_2 + w_2) = (v_1 + v_2) + (w_1 + w_2).$$

Como V é subespaço vetorial de $\mathbb{R}^n$ e $v_1, v_2 \in V$, o vetor $v_1 + v_2$ é um elemento de V. De modo análogo, concluímos que $w_1 + w_2$ é um elemento de W. Logo, $x + y \in V + W$.

iii) Dado $x \in V+W$ tem-se $x = v+w$, com $v \in V$ e $w \in W$. Logo, para qualquer $\alpha \in \mathbb{R}$,

$$\alpha x = \alpha(v + w) = \alpha v + \alpha w.$$

Uma vez que V e W são subespaços vetoriais de $\mathbb{R}^n$, temos que $\alpha v \in V$ e $\alpha w \in W$. Assim, $\alpha x \in V + W$.

De *1)*, *2)* e *3)* conclui-se que $V+W$ é subespaço vetorial de $\mathbb{R}^n$. □

Definição 3.2.15. *Sejam V, W subespaços vetoriais do espaço vetorial $\mathbb{R}^n$.*
*Diz-se que $V+W$ é uma **soma direta** se $V \cap W = \{0_{\mathbb{R}^n}\}$.*
*Diz-se que $\mathbb{R}^n$ **é soma direta de** V **e** W, e escreve-se $\mathbb{R}^n = V \oplus W$, se $\mathbb{R}^n = V+W$ e a soma $V+W$ é direta. Caso $\mathbb{R}^n$ seja soma direta de V e W, diz-se que V é **suplementar** de W relativamente a $\mathbb{R}^n$ (e que W é suplementar de V relativamente a $\mathbb{R}^n$ ou que V e W são suplementares).*

Exemplo 3.2.16. *Consideremos o espaço vetorial real $\mathbb{R}^4$ e os subespaços*

$$\begin{array}{rcl} U & = & \{(s,t,u,v) \in \mathbb{R}^4 : s = 0, u = 0\}, \\ V & = & \{(a,b,c,d) \in \mathbb{R}^4 : b + d = 0 \text{ e } b - d = 0\}, \\ W & = & \{(x,0,y,2x) \in \mathbb{R}^4 : x, y \in \mathbb{R}\}. \end{array}$$

Então

$$\begin{array}{rcl} U \cap V & = & \{(x,y,z,w) \in \mathbb{R}^4 : x = 0, z = 0, y - w = 0, y + w = 0\} \\ & = & \{(0,0,0,0)\}, \\ V \cap W & = & \{(a,b,c,d) \in \mathbb{R}^4 : b + d = 0, b - d = 0, b = 0, d = 2a\} \\ & = & \{(0,0,c,0) \mid c \in \mathbb{R}\}, \\ U \cap W & = & \{(a,b,c,d) \in \mathbb{R}^4 : a = 0, c = 0, b = 0, d = 2a\} \\ & = & \{(0,0,0,0)\}. \end{array}$$

Logo, $V+W$ não é uma soma direta e $U+V$ e $U+W$ são somas diretas.
Uma vez que $U+W=\mathbb{R}^4$ e $U\cap W=\{0_{\mathbb{R}^4}\}$, então $\mathbb{R}^4$ é soma direta de U e W; o espaço W é um suplementar de U relativamente a $\mathbb{R}^4$.
Atendendo a que $\mathbb{R}^4 \nsubseteq V+W$, pois $(0,1,0,0)\notin V+W$, concluímos que $\mathbb{R}^4$ não é soma direta de V e W.

Como podemos verificar no exemplo que se segue, existem subespaços do espaço vetorial $\mathbb{R}^n$ que admitem mais do que um suplementar relativamente a $\mathbb{R}^n$.

Exemplo 3.2.17. *Consideremos no espaço vetorial real $\mathbb{R}^3$, os subespaços*

$$\begin{aligned} V_1 &= \{(a,b,c)\in\mathbb{R}^3 : b=c=0\},\\ V_2 &= \{(x,y,z)\in\mathbb{R}^3 : x=0\},\\ V_3 &= \{(s,t,u)\in\mathbb{R}^3 : s=t\}. \end{aligned}$$

Tem-se $\mathbb{R}^3=V_1\oplus V_2$ e $\mathbb{R}^3=V_1\oplus V_3$ e $V_2\neq V_3$.

Teorema 3.2.18. *Sejam V e W subespaços vetoriais do espaço vetorial real $\mathbb{R}^n$. Então:*

1. *$\mathbb{R}^n$ é soma direta de V e W se e só se cada vetor de $\mathbb{R}^n$ se escreve, de modo único, na forma $v+w$ com $v\in V$ e $w\in W$.*

2. *$\mathbb{R}^n$ é soma direta de V e W se e só se $\mathbb{R}^n=V+W$ e $0_{\mathbb{R}^n}$ se escreve, de modo único, na forma $v+w$ com $v\in V$ e $w\in W$.*

Demonstração. 1. Sejam V e W subespaços vetoriais do espaço vetorial $\mathbb{R}^n$. Admitamos que $\mathbb{R}^n$ é soma direta de V e W, ou seja, que $\mathbb{R}^n=V+W$ e $U\cap W=\{0_{\mathbb{R}^n}\}$. Vamos mostrar que todo o elemento de $\mathbb{R}^n$ se escreve, de modo único, na forma $v+w$ com $v\in V$ e $w\in W$.
Uma vez que $\mathbb{R}^n=V+W$, então, para todo $x\in\mathbb{R}^n$, existem $v\in V$ e $w\in W$ tais que $x=v+w$. Além disso, se admitirmos que x é um elemento de $\mathbb{R}^n$ tal $x=v+w$ com $v\in V$ e $w\in W$ e $x=v'+w'$ com $v'\in V$ e $w'\in W$, prova-se que $v=v'$ e $w=w'$. De facto, como $v+w=v'+w'$, vem que $v-v'=w-w'$. Como $v,v'\in V$ e $w,w'\in W$, então $v-v',w-w'\in V\cap W=\{0_{\mathbb{R}^n}\}$. Logo, $v-v'=0_{\mathbb{R}^n}$ e $w-w'=0_{\mathbb{R}^n}$, pelo que $v=v'$ e $w=w'$.
Reciprocamente, suponhamos que cada elemento de $\mathbb{R}^n$ se escreve, de modo único, na forma $v+w$ com $v\in V$ e $w\in W$. Logo, $\mathbb{R}^n=V+W$. Para concluir que $\mathbb{R}^n$ é soma direta de V e W resta provar que $V\cap W=\{0_{\mathbb{R}^n}\}$. Uma vez que $V\cap W$ é um subespaço vetorial de $\mathbb{R}^n$, é claro que $\{0_{\mathbb{R}^n}\}\subseteq V\cap W$. Também se prova que $V\cap W\subseteq\{0_{\mathbb{R}^n}\}$. De facto, dado $x\in V\cap W$, tem-se $x=x+0_{\mathbb{R}^n}$ com $x\in V$ e $0_{\mathbb{R}^n}\in W$ e $x=0_{\mathbb{R}^n}+x$ com $0_{\mathbb{R}^n}\in V$ e $x\in W$. Então, como $x\in\mathbb{R}^n$ e todo o vetor de $\mathbb{R}^n$ se escreve de, modo único, como soma de um elemento de V com um elemento de W, temos $x=0_{\mathbb{R}^n}$. Portanto, $V\cap W\subseteq\{0_{\mathbb{R}^n}\}$.

2. Exercício. □

As noções de soma e de soma direta de subespaços podem ser generalizadas a mais de dois subespaços da forma seguinte.

Definição 3.2.19. *Sejam $r \in \mathbb{N}\backslash\{1\}$ e V_1, V_2, ..., V_r subespaços vetoriais do espaço vetorial real $\mathbb{R}^n$.*
Designa-se por ***soma dos subespaços*** *V_1, V_2, ..., V_r, e representa-se por $V_1 + V_2 + \ldots + V_r$, o conjunto*

$$V_1 + V_2 + \ldots + V_n = \{x_1 + x_2 + \ldots + x_r : x_1 \in V_1, x_2 \in V_2, \ldots, x_r \in V_n\}.$$

Diz-se que $V_1 + V_2 + \ldots + V_r$ é uma ***soma direta*** *se*

$$V_i \cap (V_1 + \ldots + V_{i-1} + V_{i+1} + \ldots + V_r) = \{0_{\mathbb{R}^n}\}, \text{ para todo } i \in \{1, \ldots, r\}.$$

Diz-se que $\mathbb{R}^n$ ***é soma direta de*** *$V_1, V_2, \ldots, V_r$, e escreve-se $\mathbb{R}^n = V_1 \oplus V_2 \oplus \ldots \oplus V_r$, se*

i) $\mathbb{R}^n = V_1 + V_2 + \ldots + V_r$;

ii) $V_1 + V_2 + \ldots + V_r$ *é uma soma direta.*

Os dois últimos resultados podem também ser generalizados a somas de mais de dois subespaços.

Teorema 3.2.20. *Sejam $r \in \mathbb{N} \setminus \{1\}$ e $V_1, V_2, \ldots, V_r$ subespaços vetoriais do espaço vetorial real $\mathbb{R}^n$. Então $V_1 + V_2 + \ldots + V_r$ é um subespaço vetorial de $\mathbb{R}^n$.*

Teorema 3.2.21. *Sejam $r \in \mathbb{N} \setminus \{1\}$ e $V_1, V_2, \ldots, V_r$ subespaços vetoriais do espaço vetorial real $\mathbb{R}^n$. Então são equivalentes as três afirmações seguintes:*

i) *$\mathbb{R}^n$ é soma direta de $V_1, V_2, \ldots, V_r$.*

ii) *$\mathbb{R}^n = V_1 + V_2 + \ldots + V_r$ e cada elemento de $\mathbb{R}^n$ escreve-se, de modo único, na forma $v_1 + v_2 + \ldots + v_r$ com $v_1 \in V_1$, $v_2 \in V_2$, ..., $v_r \in V_r$.*

iii) *$\mathbb{R}^n = V_1 + V_2 + \ldots + V_r$ e o vetor $0_{\mathbb{R}^n}$ escreve-se, de modo único, na forma $v_1 + v_2 + \ldots + v_r$ com $v_1 \in V_1$, $v_2 \in V_2$, ..., $v_r \in V_r$.*

3.2.3 Combinação linear de vetores

Definição 3.2.22. Seja S um subconjunto não vazio de $\mathbb{R}^n$. Diz-se que:

- $v \in \mathbb{R}^n$ é ***combinação linear dos elementos*** $v_1, v_2, ..., v_k \in \mathbb{R}^n$, $k \in \mathbb{N}$, se existem escalares $\alpha_1, \alpha_2, ..., \alpha_k \in \mathbb{R}$ tais que

$$v = \sum_{i=1}^{k} \alpha_i v_i = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_k v_k.$$

  Neste caso, aos escalares $\alpha_1, \alpha_2, \ldots, \alpha_k$ dá-se a designação de ***coeficientes*** da combinação linear.

- $v \in \mathbb{R}^n$ é ***combinação linear de elementos de S*** se existem v_1, v_2, ..., $v_k \in S$ tais que v é combinação linear destes elementos.

Exemplo 3.2.23. *No espaço vetorial real $\mathbb{R}^3$, consideremos os vetores $v_1 = (1,2,0)$, $v_2 = (-2,3,4)$ e $v_3 = (-1,12,8)$. O vetor v_3 é combinação linear de v_1 e v_2, pois*

$$(-1,12,8) = 3(1,2,0) + 2(-2,3,4).$$

Exemplo 3.2.24. *No espaço vetorial real $\mathbb{R}^3$, consideremos os vetores $v_1 = (1,2,0)$, $v_2 = (-2,3,0)$ e $v_3 = (-1,2,2)$. O vetor v_3 não é combinação linear de v_1 e v_2, pois*

$$(-1,2,2) \neq \alpha(1,2,0) + \beta(-2,3,0), \text{ para quaisquer } \alpha, \beta \in \mathbb{R}.$$

Exemplo 3.2.25. *No espaço vetorial real $\mathbb{R}^n$, temos*

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

para qualquer $(x_1, x_2, ..., x_n) \in \mathbb{R}^n$. Assim, qualquer vetor de $\mathbb{R}^n$ é combinação linear dos vetores

$$e_1 = (1,0,0,...,0)\,, e_2 = (0,1,0,...,0)\,, \ldots, e_n = (0,0,...,0,1)\,.$$

Exemplo 3.2.26. *No espaço vetorial real $\mathbb{R}^2$, todo o vetor (x,y) é combinação linear dos vetores $(1,0)\,,\ (0,1)$ e $(1,1)$, uma vez que*

$$(x,y) = (x-1)\,(1,0) + (y-1)\,(0,1) + 1\,(1,1)\,.$$

3.2.4 Subespaço gerado por um conjunto de vetores

Teorema 3.2.27. *Seja S um subconjunto não vazio de $\mathbb{R}^n$. Então,*

1. *o conjunto*

$$V = \{x : x \text{ é combinação linear de elementos de } S\}$$

   é um subespaço vetorial de $\mathbb{R}^n$;

2. *$S \subseteq V$;*

3. *V é o menor subespaço vetorial de $\mathbb{R}^n$ que contém S.*

Demonstração. Sejam S e V conjuntos nas condições indicadas no enunciado do teorema.

1. Mostremos que V é subespaço vetorial de $\mathbb{R}^n$. Claramente, V é um subconjunto de $\mathbb{R}^n$, pois $S \subseteq \mathbb{R}^n$ e, como $\mathbb{R}^n$ é espaço vetorial, toda a combinação linear de elementos de S é um elemento de $\mathbb{R}^n$. Assim, todo o elemento de V é também elemento de $\mathbb{R}^n$. Além disso, verifica-se que:

i) $V \neq \emptyset$, pois $0_{\mathbb{R}^n} \in V$. De facto, como $S \neq \emptyset$, existe $x \in S$. Logo, como $0_{\mathbb{R}^n} = 0x$, $0_{\mathbb{R}^n}$ é combinação linear de elementos de S e, portanto, $0_{\mathbb{R}^n} \in V$.

ii) Para quaisquer $x, y \in V$, tem-se $x + y \in V$. De facto, se $x, y \in V$, existem $x_1, x_2, ..., x_k,\ y_1, y_2, ..., y_m \in S$ e $\alpha_1, \alpha_2, ..., \alpha_k,\ \beta_1, \beta_2, ..., \beta_m \in \mathbb{R}$ tais que

$$x = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_k x_k \quad \text{e} \quad y = \beta_1 y_1 + \beta_2 y_2 + ... + \beta_m y_m.$$

Assim,

$$x + y = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_k x_k + \beta_1 y_1 + \beta_2 y_2 + ... + \beta_m y_m,$$

i.e., $x + y$ é combinação linear de $k + m$ elementos de S. Portanto, $x + y \in V$.

iii) Para quaisquer $\alpha \in \mathbb{R}$ e $x \in V$, tem-se $\alpha x \in V$. Com efeito, se $x \in V$, existem $x_1, x_2, ..., x_k \in S$ e $\alpha_1, \alpha_2, ..., \alpha_k \in \mathbb{R}$ tais que $x = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_k x_k$. Então

$$\begin{aligned} \alpha x &= \alpha (\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_k x_k) \\ &= \alpha (\alpha_1 x_1) + \alpha (\alpha_2 x_2) + ... + \alpha (\alpha_k x_k) \\ &= (\alpha\alpha_1) x_1 + (\alpha\alpha_2) x_2 + ... + (\alpha\alpha_k) x_k, \end{aligned}$$

com $\alpha\alpha_1, \alpha\alpha_2, \ldots, \alpha\alpha_k \in \mathbb{R}$. Portanto, αx é combinação linear de elementos de S, pelo que $\alpha x \in V$.

De *i)*, *ii)* e *iii)* conclui-se que V é subespaço vetorial de $\mathbb{R}^n$.

2. Para todo $x \in S$, $x = 1x$ e, portanto, $x \in V$. Logo, $S \subseteq V$.

3. Pretendemos mostrar que se U é um subespaço de $\mathbb{R}^n$ tal que $S \subseteq U$, então $V \subseteq U$. De facto, se $x \in V$, existem $x_1, x_2, ..., x_k \in S$ e $\alpha_1, \alpha_2, ..., \alpha_k \in \mathbb{R}$ tais que

$$x = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_k x_k.$$

Então, se admitirmos que U é um subespaço vetorial de $\mathbb{R}^n$ tal que $S \subseteq U$, também temos $x_1, x_2, ..., x_k \in U$ e, consequentemente, $x \in U$. Logo, $V \subseteq U$. □

Definição 3.2.28. *Seja S um subconjunto não vazio de $\mathbb{R}^n$. Ao subespaço vetorial do espaço vetorial $\mathbb{R}^n$ definido por*

$$V = \{x \in \mathbb{R}^n : x \text{ é combinação linear de elementos de } S\}$$

*chama-se **subespaço de $\mathbb{R}^n$ gerado por** S, e representa-se por $\langle S \rangle$. Ao conjunto S chamamos **conjunto gerador** de V.*

Convenciona-se que $\langle \emptyset \rangle = \{0_{\mathbb{R}^n}\}$.

Notação: Se S é um subconjunto de $\mathbb{R}^n$ tal que $S = \{x_1, x_2, ..., x_k\}$, onde $k \in \mathbb{N}$, pode-se representar o subespaço de $\mathbb{R}^n$ gerado por S por $\langle x_1, x_2, ..., x_k \rangle$ em vez de $\langle \{x_1, x_2, ..., x_k\} \rangle$.

Exemplo 3.2.29. *Na sequência do exemplo 3.2.26, tem-se*

$$\mathbb{R}^2 = \langle (1, 0), (0, 1), (1, 1) \rangle.$$

Exemplo 3.2.30. *Consideremos o espaço vetorial real* $\mathbb{R}^n$ *e sejam*

$$e_1 = (1, 0, 0, ..., 0), e_2 = (0, 1, 0, ..., 0), \ldots, e_n = (0, 0, ..., 0, 1).$$

Com base no exemplo 3.2.25, podemos afirmar que

$$\mathbb{R}^n = \langle e_1, e_2, ..., e_n \rangle .$$

Exemplo 3.2.31. *Considere-se, em* $\mathbb{R}^3$*, o conjunto* $V = \{(x, y, z) \in \mathbb{R}^3 : z = x - y\}$*. Verifica-se facilmente que* V *é um subespaço vetorial de* $\mathbb{R}^3$*. Uma vez que*

$$\begin{aligned} V &= \{(x, y, x - y) \in \mathbb{R}^3 : x, y \in \mathbb{R}\} \\ &= \{x \cdot (1, 0, 1) + y \cdot (0, 1, -1) \in \mathbb{R}^3 : x, y \in \mathbb{R}\} \\ &= \langle (1, 0, 1), (0, 1, -1) \rangle , \end{aligned}$$

conclui-se que $\{(1, 0, 1), (0, 1, -1)\}$ *é um conjunto gerador de* V*.*

Teorema 3.2.32. *Sejam* $k \in \mathbb{N}$*,* $v_1, ..., v_k, v \in \mathbb{R}^n$ *tais que* v *é combinação linear de* $v_1, ..., v_k$*. Então,*

$$\langle v_1, ..., v_k, v \rangle = \langle v_1, ..., v_k \rangle .$$

Demonstração. Sejam $V = \langle v_1, ..., v_k \rangle$ e $W = \langle v_1, ..., v_k, v \rangle$. Para provar a igualdade $V = W$, vamos mostrar que $V \subseteq W$ e $W \subseteq V$.

($V \subseteq W$): Do teorema anterior segue que $\{v_1, ..., v_k, v\} \subseteq W$. Logo, como V é o menor subespaço de $\mathbb{R}^n$ que contém $\{v_1, ..., v_k\}$ e W é um subespaço de $\mathbb{R}^n$ que também contém este conjunto, temos $V \subseteq W$.

($W \subseteq V$): Por hipótese, v é combinação linear de $v_1, ..., v_k$, pelo que $v \in V$. Logo, como $\{v_1, ..., v_k, v\} \subseteq V$ e W é o menor subespaço de $\mathbb{R}^n$ que contém este conjunto, temos $W \subseteq V$. □

Teorema 3.2.33. *Sejam* $k \in \mathbb{N}$*,* v_1*,...,*$v_k \in \mathbb{R}^n$*,* $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_k \in \mathbb{R}$ *e* $\alpha_i \in \mathbb{R} \setminus \{0\}$*. Então*

$$\langle v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_k \rangle = \langle v_1, \ldots, v_{i-1}, \sum_{j=1}^{k} \alpha_j v_j, v_{i+1}, \ldots, v_k \rangle.$$

Demonstração. Exercício. □

Teorema 3.2.34. *Todo subespaço vetorial do espaço vetorial* $\mathbb{R}^n$ *admite um conjunto gerador finito.*

Demonstração. Seja V um subespaço vetorial de $\mathbb{R}^n$.

Se $V = \{0_{\mathbb{R}^n}\}$, então $V = \langle \emptyset \rangle$ e o conjunto gerador é o conjunto vazio, que é finito. Se $V \neq \{0_{\mathbb{R}^n}\}$, existe $v_1 \in V$. Se $V = \langle v_1 \rangle$, então V é gerado pelo conjunto $\{v_1\}$, que é finito. Caso contrário, isto é, se $V \neq \langle v_1 \rangle$, consideremos $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tais que $v_1 = \alpha_1 e_1 + \ldots + \alpha_n e_n$, onde, para cada $i \in \{1, \ldots, n\}$, e_i é o n-uplo cujo elemento na coordenada i é 1 e todos os outros elementos são zero. Note-se que existem $\alpha_1, \ldots, \alpha_n$ nas condições indicadas, pois $< e_1, \ldots, e_n >= \mathbb{R}^n$ e $v_1 \in \mathbb{R}^n$. Como $v_1 \neq 0_{\mathbb{R}^n}$, tem-se $\alpha_{i_1} \neq 0$, para algum $i_1 \in \{1, \ldots, n\}$. Então, pelo teorema anterior,

$$< e_1, \ldots, e_{i_1-1}, e_{i_1}, e_{i_1+1}, \ldots, e_n >=< e_1, \ldots, e_{i_1-1}, v_1, e_{i_1+1}, \ldots, e_n > .$$

Consideremos, agora, $v_2 \in V \backslash < v_1 >$. Se $V = \langle v_1, v_2 \rangle$, então V é gerado pelo conjunto $\{v_1, v_2\}$, que é finito. Caso contrário, consideremos reais $\beta_1, \ldots, \beta_{i_1-1}$,$\beta_{i_1}$,$\beta_{i_1+1}$, $\ldots, \beta_n$ tais que

$$v_2 = \beta_1 e_1 + \ldots + \beta_{i_1-1} e_{i_1-1} + \beta_{i_1} v_1 + \beta_{i_1+1} e_{i_1+1} + \beta_n e_n.$$

Como $v_2 \in V \backslash < v_1 >$, temos $v_2 \neq \beta v_1$, para todo $\beta \in \mathbb{R}$ e, por conseguinte, tem-se $\beta_{i_2} \neq 0$, para algum $i_2 \in \{1, \ldots, n\} \setminus \{i_1\}$. Assim,

$$\begin{aligned} & < e_1, \ldots, e_{i_1-1}, e_{i_1}, e_{i_1+1}, \ldots, e_{i_2-1}, e_{i_2}, e_{i_2+1}, \ldots, e_n > \\ = \; & < e_1, \ldots, e_{i_1-1}, v_1, e_{i_1+1}, \ldots, \ldots, e_{i_2-1}, v_2, e_{i_2+1}, \ldots, e_n > . \end{aligned}$$

Seguidamente considera-se $v_3 \in V \backslash < v_1, v_2 >$ e repete-se o processo. Este processo termina, uma vez que $\{e_1, \ldots, e_n\}$ é finito e se tivermos $v_1, \ldots, v_n \in V$ tais que $< e_1, \ldots, e_n >=< v_1, \ldots, v_n >$, então $V = \mathbb{R}^n$ e, portanto, não existe $v_{n+1} \in V \backslash < v_1, \ldots, v_n >$.
Logo, V admite um conjunto gerador finito. □

3.2.5 Dependência e independência linear

Definição 3.2.35. *Seja $k \in \mathbb{N}$. Uma sequência $(v_1, ..., v_k)$ de vetores de $\mathbb{R}^n$ diz-se **linearmente independente** se, para quaisquer $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$,*

$$\alpha_1 v_1 + ... + \alpha_k v_k = 0_{\mathbb{R}^n} \Rightarrow \alpha_1 = \cdots = \alpha_k = 0.$$

*Caso contrário, isto é, se existirem escalares $\alpha_1, \ldots, \alpha_k$ não todos nulos tais que $\alpha_1 v_1 + ... + \alpha_k v_k = 0_{\mathbb{R}^n}$, a sequência $(v_1, ..., v_k)$ diz-se **linearmente dependente**.*

Observação:

- Note-se que se $v_1, \ldots, v_k$ são vetores de $\mathbb{R}^n$, então é sempre possível escrever $0_{\mathbb{R}^n}$ como combinação linear destes vetores, pois $0_{\mathbb{R}^n} = 0v_1 + \ldots + 0v_k$. Logo, a sequência $(v_1, \ldots, v_k)$ é linearmente independente se e só se $0v_1 + \ldots + 0v_k$ é a única forma de escrever $0_{\mathbb{R}^n}$ como combinação linear de $v_1, \ldots, v_k$.
- Se $v = 0_{\mathbb{R}^n}$, então (v) é linearmente dependente pois $0_{\mathbb{R}^n} = 1 \cdot 0_{\mathbb{R}^n}$ e $1 \neq 0$.
- Se $v \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, é simples verificar que (v) é linearmente independente. De facto, dado $\alpha \in \mathbb{R}$, se $\alpha v = 0_{\mathbb{R}^n}$, então $\alpha = 0$ ou $v = 0_{\mathbb{R}^n}$. Como, por hipótese, $v \neq 0_{\mathbb{R}^n}$, então $\alpha = 0$. Logo, (v) é linearmente independente.

Exemplo 3.2.36. *No espaço vetorial real $\mathbb{R}^4$, a sequência de vetores*

$$((1,0,0,0),(1,0,-1,1),(2,1,0,1))$$

é linearmente independente, pois, para quaisquer $\alpha,\beta,\gamma \in \mathbb{R}$,

$$\begin{aligned}
&\alpha(1,0,0,0)+\beta(1,0,-1,1)+\gamma(2,1,0,1)=(0,0,0,0)\\
&\Rightarrow (\alpha+\beta+2\gamma,\gamma,-\beta,\beta+\gamma)=(0,0,0,0)\\
&\Rightarrow \alpha+\beta+2\gamma=0,\gamma=0,-\beta=0,\beta+\gamma=0\\
&\Rightarrow \alpha=\beta=\gamma=0
\end{aligned}$$

Exemplo 3.2.37. *No espaço vetorial real $\mathbb{R}^2$, a sequência de vetores*

$$((1,0),(0,1),(1,1))$$

é linearmente dependente, pois

$$1\,(1,0)+1\,(0,1)+(-1)\,(1,1)=(0,0)\,.$$

Note-se que o vetor nulo pode ser escrito como combinação linear dos três vetores indicados utilizando escalares não nulos.

Exemplo 3.2.38. *No espaço vetorial real $\mathbb{R}^n$, a sequência de vetores $(e_1,\ldots,e_n)$, onde cada e_i é o n-uplo cujo elemento na coordenada i é 1 e todos os outros elementos são zero, é linearmente independente. De facto,*

$$\begin{aligned}
&\alpha_1e_1+\ldots+\alpha_ne_n=0_{\mathbb{R}^n}\\
\Rightarrow\ &(\alpha_1,\ldots,\alpha_n)=(0,\ldots,0)\ .\\
\Rightarrow\ &\alpha_1=\cdots=\alpha_n=0
\end{aligned}$$

Teorema 3.2.39. *Sejam $k\in\mathbb{N}$ e $v_1,\ldots,v_k\in\mathbb{R}^n$. A sequência de vetores $(v_1,\ldots,v_k)$ é linearmente independente se e só se qualquer combinação linear de $v_1,\ldots,v_k$ tem coeficientes únicos, i.e., se e só se, para quaisquer $\alpha_1,\ldots,\alpha_k,\beta_1,\ldots,\beta_k\in\mathbb{R}$,*

$$\alpha_1v_1+\ldots+\alpha_kv_k=\beta_1v_1+\ldots+\beta_kv_k\Rightarrow\alpha_1=\beta_1\ e\ \ldots\ e\ \alpha_k=\beta_k.$$

Demonstração. $\Rightarrow$) Admitamos que $(v_1,\ldots,v_k)$ é linearmente independente e que $\alpha_1,\ldots,\alpha_k,\beta_1,\ldots,\beta_k\in\mathbb{R}$ são tais que

$$\alpha_1v_1+\ldots+\alpha_kv_k=\beta_1v_1+\ldots+\beta_kv_k.$$

Então

$$(\alpha_1-\beta_1)v_1+\ldots+(\alpha_k-\beta_k)v_k=0_{\mathbb{R}^n}$$

e, uma vez que a sequência $(v_1,\ldots,v_k)$ é linearmente independente, da igualdade anterior resulta que

$$\alpha_1-\beta_1=0,\ \ldots,\ \alpha_k-\beta_k=0$$

i.e.

$$\alpha_1=\beta_1,\ \ldots,\ \alpha_k=\beta_k.$$

$\Leftarrow$) Suponhamos que qualquer combinação linear dos vetores $v_1, \ldots, v_k$ tem coeficientes únicos, i.e., para quaisquer $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k \in \mathbb{R}$,

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = \beta_1 v_1 + \ldots + \beta_k v_k \Rightarrow \alpha_1 = \beta_1 \text{ e } \ldots \text{ e } \alpha_k = \beta_k.$$

Então a sequência $(v_1, \ldots, v_k)$ é linearmente independente, pois, dados escalares $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tais que

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0_{\mathbb{R}^n},$$

temos

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0 v_1 + \cdots + 0 v_k,$$

donde segue que $\alpha_1 = \cdots = \alpha_k = 0$. ☐

Teorema 3.2.40. *Sejam $k \in \mathbb{N} \setminus \{1\}$ e $v_1, \ldots, v_k$ elementos de $\mathbb{R}^n$. A sequência $(v_1, \ldots, v_k)$ é linearmente dependente se e só se existe $i \in \{1, \ldots, k\}$ tal que v_i é combinação linear de $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$.*

Demonstração. $\Rightarrow$) Sejam $v_1, ..., v_k$, com $k \in \mathbb{N} \setminus \{1\}$, vetores de $\mathbb{R}^n$ tais que $(v_1, \ldots, v_k)$ é linearmente dependente. Então existem $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tais que $\alpha_i \neq 0$, para algum $i \in \{1, \ldots, k\}$, e

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0_{\mathbb{R}^n}.$$

Como $\alpha_i \neq 0$, existe $\alpha_i^{-1} \in \mathbb{R}$ e da igualdade anterior resulta que

$$\alpha_i^{-1} \left(\alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_i v_i + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k\right) = \alpha_i^{-1} 0_{\mathbb{R}^n},$$

i.e.,

$$\left(\alpha_i^{-1}\alpha_1\right) v_1 + \cdots + \left(\alpha_i^{-1}\alpha_{i-1}\right) v_{i-1} + v_i + \left(\alpha_i^{-1}\alpha_{i+1}\right) v_{i+1} + \cdots + \left(\alpha_i^{-1}\alpha_k\right) v_k = 0_{\mathbb{R}^n}.$$

Assim,

$$v_i = \left(-\alpha_i^{-1}\alpha_1\right) v_1 + \cdots + \left(-\alpha_i^{-1}\alpha_{i-1}\right) v_{i-1} + \left(-\alpha_i^{-1}\alpha_{i+1}\right) v_{i+1} + \cdots + \left(-\alpha_i^{-1}\alpha_k\right) v_k,$$

e portanto, v_i é combinação linear dos restantes vetores.

$\Leftarrow$) Suponhamos que existem $i \in \{1, \ldots, k\}$ e $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_k \in \mathbb{R}$ tais que

$$v_i = \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k.$$

Então,

$$\alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + (-1)\, v_i + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k = 0_{\mathbb{R}^n},$$

i.e.,

$$\alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_i v_i + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k = 0_{\mathbb{R}^n},$$

com $\alpha_i = -1 \neq 0$. Logo, $(v_1, \ldots, v_k)$ é linearmente dependente. ☐

Corolário 3.2.41. *Seja $k \in \mathbb{N}$. Se $v_1, \ldots, v_k$ são vetores de $\mathbb{R}^n$ tais que $v_i = 0_{\mathbb{R}^n}$, para algum $i \in \{1, \ldots, k\}$, então a sequência de vetores $(v_1, \ldots, v_k)$ é linearmente dependente.*

Teorema 3.2.42. *Sejam $k \in \mathbb{N}$ e $v_1, \ldots, v_k, v$ vetores de $\mathbb{R}^n$. São válidas as propriedades seguintes:*

1. *se a sequência de vetores $(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_k)$ é linearmente independente, então a sequência $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$ também é linearmente independente;*

2. *se a sequência de vetores $(v_1, \ldots, v_k)$ é linearmente dependente, então a sequência $(v_1, \ldots, v_k, v)$ também é linearmente dependente.*

3. *se a sequência de vetores $(v_1, \ldots, v_k)$ é linearmente independente e a sequência $(v_1, \ldots, v_k, v)$ é linearmente dependente, então v é combinação linear dos vetores $v_1, \ldots, v_k$.*

Demonstração. 1. Admitamos que a sequência $(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_k)$ é linearmente independente. Mostremos que a sequência $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$ também é linearmente independente. De facto, para quaisquer escalares $\alpha_1, \ldots, \alpha_{i-1}$, α_{i+1}, ..., $\alpha_k \in \mathbb{R}$ tais que

$$\alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k = 0_{\mathbb{R}^n},$$

temos

$$\alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + 0_{\mathbb{K}} v_i + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k = 0_{\mathbb{R}^n}.$$

Então, como a sequência $(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_k)$ é linearmente independente, tem-se $\alpha_1 = \cdots = \alpha_{i-1} = \alpha_{i+1} = \cdots = \alpha_k = 0$. Logo, a sequência

$$(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$$

é linearmente independente.
2. Admitamos que a sequência de vetores $(v_1, \ldots, v_n)$ é linearmente dependente. Então existe $i \in \{1, \ldots, k\}$ tal que $\alpha_i \neq 0$ e

$$\alpha_1 v_1 + \cdots + \alpha_i v_i + \cdots + \alpha_k v_k = 0_{\mathbb{R}^n}.$$

Logo,

$$\alpha_1 v_1 + \cdots + \alpha_i v_i + \cdots + \alpha_k v_k + 0v = 0_{\mathbb{R}^n},$$

com $\alpha_i \neq 0$. Portanto, $(v_1, \ldots, v_k, v)$ é linearmente dependente.

3. Admitamos que a sequência de vetores $(v_1, \ldots, v_k)$ é linearmente independente e que a sequência $(v_1, \ldots, v_k, v)$ é linearmente dependente. Como a sequência $(v_1, \ldots, v_k, v)$ é linearmente dependente, existem escalares $\alpha_1, \ldots, \alpha_k, \alpha$ não todos nulos tais que

$$\alpha_1 v_1 + \cdots + \alpha_k v_k + \alpha v = 0.$$

Como a sequência $(v_1, \ldots, v_k)$ é linearmente independente, então $\alpha \neq 0$. De facto, se $\alpha = 0$, ter-se-ia

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0,$$

com $\alpha_i \neq 0$, para algum $i \in \{1, \ldots, k\}$ (contradição). Como $\alpha \neq 0$, então

$$v = (-\alpha^{-1}\alpha_1)v_1 + \cdots + (-\alpha^{-1}\alpha_k)v_k$$

e, portanto, v é combinação linear de $v_1, \ldots, v_k$. □

Teorema 3.2.43 (Teorema de Steinitz). *Sejam V um subespaço vetorial de $\mathbb{R}^n$, $r, p \in \mathbb{N}$, e $S = \{v_1, \ldots, v_r\}$ e $W = \{w_1, \ldots, w_p\}$ subconjuntos de $\mathbb{R}^n$ com, respetivamente, r e p vetores.*
Se $V =< S >$ e $w_1, \ldots, w_p$ são vetores de V tais que $(w_1, ..., w_p)$ é linearmente independente, então $p \leq r$ e é possível substituir p dos vetores de S por $w_1, ..., w_p$ de forma a obter um subconjunto S' de $\mathbb{R}^n$ tal que $V =< S' >$.

Demonstração. Sejam $r \in \mathbb{N}$, V um subespaço vetorial de $\mathbb{R}^n$ e $S = \{v_1, ..., v_r\}$ um subconjunto de $\mathbb{R}^n$ com r vetores tal que $< S >= V$. Pretendemos mostrar que se $w_1, \ldots, w_p$, $p \in \mathbb{N}$, são vetores de V tais que $(w_1, \ldots, w_p)$ é linearmente independente, então $p \leq r$ e é possível substituir p dos vetores de S por $w_1, \ldots, w_p$ de forma a obter um subconjunto S' de $\mathbb{R}^n$ tal que $V =< S' >$.

A prova é feita recorrendo ao método de indução matemática aplicado a p (o número de vetores linearmente independentes).

Caso $p = 1$: Suponhamos que $w_1 \in V$ e que (w_1) é linearmente independente.

1) É óbvio que $1 \leq r$, qualquer que seja $r \in \mathbb{N}$.

2) Como (w_1) é linearmente independente, então $w_1 \neq 0_{\mathbb{R}^n}$. Por outro lado, como $w_1 \in V$, w_1 é combinação linear de v_1, v_2, …, v_r e, portanto, existem $\lambda_1, ..., \lambda_r \in \mathbb{R}$, com $\lambda_i \neq 0$, para algum $i \in \{1, \ldots, r\}$, tais que

$$w_1 = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r.$$

Suponhamos, sem perda de generalidade, que $\lambda_1 \neq 0$. Então, pelo Teorema 3.2.33,

$$V = \langle v_1, v_2, \ldots, v_r \rangle = \langle w_1, v_2, \ldots, v_r \rangle .$$

Logo, o resultado é válido para $p = 1$.

Passo de indução: Admita-se, por hipótese de indução, que as afirmações são verdadeiras para $k \in \mathbb{N}$, ou seja, admita-se que se $w_1, \ldots, w_k \in \mathbb{R}^n$ são k vetores de V tais que $(w_1, \ldots, w_k)$ é linearmente independente, então

$h_1)$ $k \leq r$;

$h_2)$ é possível substituir k dos vetores de S por $w_1, \ldots, w_k$ de forma a obter um subconjunto S' de $\mathbb{R}^n$ tal que $V =< S' >$.

Pretendemos mostrar que o resultado é válido para $k + 1$, ou seja, temos de provar que se $w_1, w_2, \ldots, w_k, w_{k+1}$ são $k + 1$ vetores de V tais que $(w_1, \ldots, w_k, w_{k+1})$ é linearmente independente, então

$t_1)$ $k + 1 \leq r$;

$t_2)$ é possível substituir $k + 1$ dos vetores de S por $w_1, \ldots, w_k, w_{k+1}$ de forma a obter um subconjunto S'' de $\mathbb{R}^n$ tal que $V =< S'' >$.

De facto, se admitirmos que a sequência $(w_1, \ldots, w_k, w_{k+1})$ é linearmente independente, então, pelo Teorema 3.2.42, a sequência $(w_1, \ldots, w_k)$ é linearmente independente. Logo,

t_1) Por h_1), temos $k \leq r$. Se admitirmos que $k = r$, então por h_2) segue que $\langle v_1, \ldots, v_r\rangle = \langle w_1, \ldots w_k\rangle$, pelo que w_{k+1} é combinação linear de $w_1, \ldots w_k$, o que contaria a hipótese de que $(w_1, \ldots, w_k, w_{k+1})$ é linearmente independente. Assim, $k < r$ e, portanto, $k + 1 \leq r$.

t_2) Por h_2) é possível substituir k dos vetores de S por $w_1, w_2, \ldots, w_k$ de forma a obter um subconjunto S' de $\mathbb{R}^n$ tal que $\mathbb{R}^n = \langle S'\rangle$. Suponhamos, sem perda de generalidade, que

$$S' = \{w_1, \ldots, w_k, v_{k+1}, v_{k+2}, \ldots, v_r\}.$$

Logo, w_{k+1} é combinação linear dos vetores $w_1, \ldots, w_k,\ v_{k+1}, v_{k+2}, \ldots, v_r$, isto é, existem $\lambda_1, \ldots, \lambda_k, \lambda_{k+1}, \ldots, \lambda_r$ tais que

$$w_{k+1} = \lambda_1 w_1 + \cdots + \lambda_k w_k + \lambda_{k+1} v_{k+1} + \lambda_{k+2} v_{k+2} + \cdots \lambda_r v_r.$$

Como $(w_1, \ldots, w_k, w_{k+1})$ é linearmente independente, temos $\lambda_i \neq 0$, para algum $i \in \{k+1, \ldots, r\}$, caso contrário viria

$$w_{k+1} = \lambda_1 w_1 + \cdots + \lambda_k w_k$$

o que contraria a hipótese de que a sequência $(w_1, \ldots, w_k, w_{k+1})$ é linearmente independente. Suponhamos, sem perda de generalidade, que $\lambda_{k+1} \neq 0$. Então, pelo Teorema 3.2.33,

$$\langle w_1, \ldots, w_k, v_{k+1}, v_{k+2}, \ldots, v_r\rangle = \langle w_1, \ldots, w_k, w_{k+1}, v_{k+2} \ldots, v_r\rangle.$$

Logo, $S'' = \{w_1, \ldots, w_k, w_{k+1}, v_{k+2}, \ldots, v_r\}$ é um conjunto obtido de S substituindo $k + 1$ dos vetores de S por $w_1, \ldots, w_k, w_{k+1}$ e tal que $V = \langle S''\rangle$. $\square$

Observação: Do teorema anterior, resulta que se um subespaço vetorial V do espaço vetorial $\mathbb{R}^n$ é gerado por r vetores, então qualquer sequência de vetores de V que seja linearmente independente nunca pode ter mais do que r vetores. Em particular, qualquer sequência de vetores de $\mathbb{R}^n$ que seja linearmente independente não pode ter mais do que n vetores.

Exemplo 3.2.44. *No espaço vetorial* $\mathbb{R}^4$, *consideremos os vetores* $w_1 = (-2, 0, 0, 2)$, $w_2 = (1, 0, 2, -1)$, $u_1 = (0, 0, 0, 1)$, $u_2 = (0, 0, 1, 0)$, $u_3 = (1, 0, 0, 0)$. *Seja* U *o subespaço vetorial de* $\mathbb{R}^4$ *gerado por* $S = \{u_1, u_2, u_3\}$, *i.e.,* $U =< u_1, u_2, u_3 >$.
Tem-se

$$\begin{aligned} w_1 &= 2u_1 + 0u_2 + (-2)u_3, \\ w_2 &= (-1)u_1 + 2u_2 + 1u_3 \end{aligned}$$

e, portanto, w_1 *e* w_2 *são vetores de* U. *Além disso, é simples verificar que a sequência* (w_1, w_2) *é linearmente independente. Logo, pelo teorema anterior, é possível substituir dois dos vetores de* $S = \{u_1, u_2, u_3\}$ *por* w_1 *e* w_2 *de forma a obter um conjunto* S' *tal que* $U = \langle S'\rangle$. *A substituição faz-se vetor a vetor seguindo o processo descrito* *na demonstração do referido teorema. Note-se, no entanto, que pode haver mais de uma maneira de efectuar a substituição e o conjunto S' obtido no final do processo pode não ser único. Porém, todo o conjunto S' obtido pelo processo indicado no teorema anterior gera o mesmo subespaço que o conjunto $\{u_1, u_2, u_3\}$. Vamos ver duas formas de efectuar essa substituição.*

1) Temos $w_1 = 2u_1 + 0u_2 + (-2)u_3$ e $2 \in \mathbb{R} \setminus \{0\}$, logo, pelo Teorema 3.2.33,

$$< u_1, u_2, u_3 >=< w_1, u_2, u_3 > .$$

Como w_2 é combinação linear de u_1, u_2, u_3, então também é combinação linear de w_1, u_2, u_3. De facto, tem-se

$$u_1 = \frac{1}{2}w_1 + 0u_2 + 1u_3,$$

pelo que

$$w_2 = (-1)u_1 + 2u_2 + 1u_3 = -\frac{1}{2}w_1 + 2u_2 + 0u_3$$

onde $2 \in \mathbb{R} \setminus \{0\}$. Então, novamente pelo Teorema 3.2.33, segue que

$$< w_1, u_2, u_3 >=< w_1, w_2, u_3 > .$$

Logo $U =< u_1, u_2, u_3 >=< w_1, w_2, u_3 >$.

2) Uma vez que $w_1 = 2u_1 + 0u_2 + (-2)u_3$ e $-2 \in \mathbb{R} \setminus \{0\}$, então, pelo Teorema 3.2.33, tem-se

$$< u_1, u_2, u_3 >=< u_1, u_2, w_1 > .$$

Como w_2 é combinação linear de u_1, u_2, u_3, então também é combinação linear de u_1, u_2, w_1. De facto, como

$$u_3 = 1u_1 + 0u_2 - \frac{1}{2}w_1,$$

segue que

$$w_2 = 0u_1 + 2u_2 - \frac{1}{2}w_1,$$

com $2 \in \mathbb{R} \setminus \{0\}$. Logo, pelo Teorema 3.2.33,

$$< u_1, u_2, w_1 >=< u_1, w_2, w_1 >$$

e, portanto, $U =< u_1, u_2, u_3 >=< u_1, w_2, w_1 >$.

Teorema 3.2.45. *Sejam $r, p \in \mathbb{N}$ tais que $p \leq r$ e $S = \{v_1, \ldots, v_r\}$ e $W = \{w_1, \ldots, w_p\}$ subconjuntos de $\mathbb{R}^n$ com, respetivamente, r e p vetores e tais que as sequências $(v_1, \ldots, v_r)$ e $(w_1, \ldots, w_p)$ são linearmente independentes. Se S' é um conjunto que se obtém de S substituindo p dos vetores de S por $w_1, \ldots, w_p$ e $< S' >=< S >$, então uma sequência formada pelos vetores de S' é linearmente independente.*

Demonstração. A prova pode ser feita por indução sobre p. $\square$

3.2.6 Bases e dimensão

Como foi referido anteriormente, qualquer subespaço vetorial de $\mathbb{R}^n$ é gerado por algum conjunto finito de vetores de $\mathbb{R}^n$. Além disso, qualquer sequência de vetores de $\mathbb{R}^n$ que seja linearmente independente pode ter, no máximo, n vetores. Assim, qualquer sequência de vetores que gere um subespaço de $\mathbb{R}^n$ e que seja linearmente independente tem de ser finita. Tal motiva a definição seguinte.

Definição 3.2.46. *Sejam $r \in \mathbb{N}$, $(v_1, \ldots, v_r)$ uma sequência de vetores de $\mathbb{R}^n$ e V um subespaço vetorial não nulo de $\mathbb{R}^n$. Diz-se que a sequência $(v_1, \ldots, v_r)$ é uma* ***base*** *de V se:*

i) a sequência $(v_1, \ldots, v_r)$ é linearmente independente;

ii) $\{v_1, \ldots, v_r\}$ é um conjunto gerador de V.

Convenciona-se que $(v_i)_{i\in\emptyset}$ é a única base do espaço $\{0_{\mathbb{R}^n}\}$.

Observação: Uma vez que uma base é definida como sendo uma sequência, duas bases com os mesmos elementos ordenados de forma diferente são distintas.

Definição 3.2.47. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, $r \in \mathbb{N}$, $\mathcal{B} = (v_1, \ldots, v_r)$ uma base de V e $v \in V$. Chamam-se* ***componentes*** *ou* ***coordenadas*** *de v na base $(v_1, \ldots, v_r)$ aos coeficientes $\alpha_1, \ldots, \alpha_r$ da combinação linear $v = \alpha_1 v_1 + \ldots + \alpha_r v_r$.*

Exemplo 3.2.48. *Considerando o espaço vetorial $\mathbb{R}^n$, a sequência $(e_1, \ldots, e_n)$, onde cada e_i é o n-uplo cujo elemento na coordenada i é 1 e todos os outros elementos são zero, é uma base de $\mathbb{R}^n$. De facto, de exemplos anteriores sabemos que a sequência $(e_1, \ldots, e_n)$ é linearmente independente e que $\{e_1, \ldots, e_n\}$ é um conjunto gerador de $\mathbb{R}^n$. À base $(e_1, \ldots, e_n)$ dá-se a designação de* **base canónica** *de $\mathbb{R}^n$.*

Exemplo 3.2.49. *A sequência $((1,0), (0,1), (1,1))$ não é uma base do espaço vetorial real $\mathbb{R}^2$, pois, embora $\{((1,0), (0,1), (1,1)\}$ seja um conjunto gerador de $\mathbb{R}^2$, a sequência $((1,0), (0,1), (1,1))$ é linearmente dependente.*

Teorema 3.2.50. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, $r \in \mathbb{N}$ e $(v_1, \ldots, v_r)$ uma sequência de vetores de V. Então $(v_1, \ldots, v_r)$ é uma base de V se e só se todo o elemento de V se escreve, de modo único, como combinação linear dos vetores $v_1, \ldots, v_r$.*

Demonstração. $\Rightarrow$) Sejam $r \in \mathbb{N}$ e $(v_1, \ldots, v_r)$ uma base de V. Seja $v \in V$. Como $\{v_1, \ldots, v_r\}$ é um conjunto gerador de V, existem α_1, …, $\alpha_r \in \mathbb{R}$ tais que $v = \alpha_1 v_1 + \ldots + \alpha_r v_r$. Além disso, como a sequência $(v_1, \ldots, v_r)$ é linearmente independente, sabemos, pelo Teorema 3.2.59, que v se escreve de modo único como combinação linear destes vetores. Logo, cada vetor de V escreve-se de modo único como combinação linear dos vetores $v_1, \ldots, v_r$.

$\Leftarrow$) Reciprocamente, admitamos que cada vetor de V se escreve de modo único como combinação linear dos vetores $v_1, \ldots, v_r$. Então, pelo Teorema 3.2.59, a sequência $(v_1, \ldots, v_r)$ é linearmente independente. Como todo o vetor de $\mathbb{R}^n$ se escreve como combinação linear destes vetores, também temos $V = \langle v_1, \ldots, v_r \rangle$. Logo $(v_1, \ldots, v_r)$ é uma base de V. □

Teorema 3.2.51. *Sejam $V \neq \{0_{\mathbb{R}^n}\}$ um subespaço do espaço vetorial $\mathbb{R}^n$ e $v_1, \ldots, v_p$ vetores de $\mathbb{R}^n$ tais que $V =< v_1, \ldots, v_p >$. Então existem $i_1, \ldots, i_r \in \{1, \ldots, p\}$, tais que $(v_{i_1}, \ldots, v_{i_r})$ é uma base de V.*

Demonstração. Sejam $V \neq \{0_{\mathbb{R}^n}\}$ um subespaço do espaço vetorial $\mathbb{R}^n$ e $v_1, \ldots, v_p$ vetores de $\mathbb{R}^n$ tais que $V =< v_1, \ldots, v_p >$. Se a sequência $(v_1, \ldots, v_p)$ é linearmente independente, então ela é, por definição, uma base de V. Caso contrário, pelo Teorema 3.2.40, sabe-se que existe $i \in \{1, \ldots, p\}$ tal que v_i é combinação linear de $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_p$. Consequentemente, pelo Teorema 3.2.32,

$$\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_p\}$$

é um conjunto gerador de V. Se $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_p)$ é linearmente independente, então esta sequência é uma base de V. Caso contrário, repete-se o procedimento, removendo vetores dependentes, até se obter uma sequência linearmente independente que ainda gera V. Como $V \neq \{0_{\mathbb{R}^n}\}$, existe pelo menos um vetor não nulo entre $v_1, \ldots, v_p$, garantindo que o processo não elimina todos os vetores. Além disso, como o número de vetores é finito, o processo termina após um número finito de passos. Portanto, existem $i_1, \ldots, i_r \in \{1, \ldots, p\}$, tais que $(v_{i_1}, \ldots, v_{i_r})$ é uma base de V. □

Exemplo 3.2.52. *No espaço vetorial real $\mathbb{R}^3$, consideremos os vetores*

$$u_1 = (-1, 0, 0), u_2 = (1, 0, 1), u_3 = (0, 0, 2), u_4 = (1, -1, 1), u_5 = (1, 1, 0).$$

É simples verificar que $\{u_1, u_2, u_3, u_4, u_5\}$ é um conjunto gerador de $\mathbb{R}^3$.
De facto, dados $(a, b, c) \in \mathbb{R}^3$ e $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5 \in \mathbb{R}$, tem-se

$$(a, b, c) = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_4 u_4 + \alpha_5 u_5 \text{ se e só se } \begin{cases} \alpha_1 &= -a + b + \alpha_2 + 2\alpha_4 \\ \alpha_3 &= (c - \alpha_2 - \alpha_4)/2 \\ \alpha_5 &= b + \alpha_4 \end{cases}.$$

Assim,

$$(a, b, c) = (-a + b + \alpha_2 + 2\alpha_4)u_1 + \alpha_2 u_2 + ((c - \alpha_2 - \alpha_4)/2)u_3 + \alpha_4 u_4 + (b + \alpha_4)u_5.$$

Em particular,

$$(0, 0, 0) = (\alpha_2 + 2\alpha_4)u_1 + \alpha_2 u_2 + ((-\alpha_2 - \alpha_4)/2)u_3 + \alpha_4 u_4 + \alpha_4 u_5,$$

para quaisquer $\alpha_2, \alpha_4 \in \mathbb{R}$ e, portanto, a sequência $(u_1, u_2, u_3, u_4, u_5)$ é linearmente dependente. Tomando, por exemplo, $\alpha_2 = -1$ e $\alpha_4 = 1$, tem-se

$$u_4 = -u_1 + u_2 + 0u_3 - u_5,$$

pelo que $< u_1, u_2, u_3, u_4, u_5 >=< u_1, u_2, u_3, u_5 >$.
Agora, para quaisquer $\alpha_1, \alpha_2, \alpha_3, \alpha_5 \in \mathbb{R}$,

$$\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_5 u_5 = (0,0,0) \Rightarrow \begin{cases} \alpha_1 &= \alpha_2 \\ \alpha_3 &= -(\alpha_2/2) \\ \alpha_5 &= 0 \end{cases}$$

e, portanto,

$$(0,0,0) = \alpha_2 u_1 + \alpha_2 u_2 + (-\alpha_2/2)u_3 + 0u_5,$$

para todo $\alpha_2 \in \mathbb{R}$. *Logo* (u_1, u_2, u_3, u_5) *é linearmente dependente. Tomando, por exemplo,* $\alpha_2 = 1$, *segue que*

$$u_1 = -u_2 + (1/2)u_3 + 0u_5,$$

pelo que $< u_1, u_2, u_3, u_5 >=< u_2, u_3, u_5 >$.
Para quaisquer $\alpha_2, \alpha_3, \alpha_5 \in \mathbb{R}$,

$$\alpha_2 u_2 + \alpha_3 u_3 + \alpha_5 u_5 = (0,0,0) \Rightarrow \alpha_2 = \alpha_3 = \alpha_5 = 0.$$

Assim, a sequência (u_2, u_3, u_5) *é linearmente independente. Portanto,* (u_2, u_3, u_5) *é uma base de* $\mathbb{R}^3$.

Teorema 3.2.53. *Todo subespaço vetorial do espaço vetorial* $\mathbb{R}^n$ *admite uma base.*

Demonstração. Seja V um subespaço vetorial de $\mathbb{R}^n$.
Se $V = \{0_{\mathbb{R}^n}\}$, então, por convenção, a sequência vazia ( ) é uma base de V.
Se $V \neq \{0_{\mathbb{R}^n}\}$, o resultado segue dos teoremas 3.2.34 e 3.2.51. $\square$

Teorema 3.2.54. *Sejam* $V \neq \{0_{\mathbb{R}^n}\}$ *um subespaço vetorial de* $\mathbb{R}^n$ *e* $(w_1, \ldots, w_p)$, $p \in \mathbb{N}$, *uma sequência de vetores de* V *linearmente independente. Então existe uma base de* V *da qual fazem parte os vetores* $w_1, \ldots, w_p$.

Demonstração. Sejam $V \neq \{0_{\mathbb{R}^n}\}$ um subespaço vetorial de $\mathbb{R}^n$ e $(w_1, \ldots, w_p)$ uma sequência de vetores de V linearmente independente.
Pelo teorema anterior, V admite uma base; seja $(v_1, \ldots, v_r)$, $r \in \mathbb{N}$, uma dessas bases. Então $S = \{v_1, \ldots, v_r\}$ é um conjunto gerador de V com r elementos distintos.
Logo, pelo Teorema 3.2.43, temos $p \leq r$ e é possível substituir p dos elementos de S pelos vetores $w_1, \ldots, w_p$ de forma a obter um conjunto S' gerador de V.
Se $p = r$, temos $S' = \{w_1, \ldots, w_p\}$ e, portanto, $(w_1, \ldots, w_p)$ é uma base de V.
Se $p < r$, suponhamos, sem perda de generalidade que $S' = \{w_1, \ldots, w_p, v_{p+1}, \ldots, v_r\}$.
Uma vez que as sequências $(v_1, \ldots, v_r)$ e $(w_1, \ldots, w_p)$ são linearmente independentes e $< S >=< S' >$, do Teorema 3.2.45 conclui-se que $(w_1, \ldots, w_p, v_{p+1}, \ldots, v_r)$ é linearmente independente. Logo $(w_1, \ldots, w_p, v_{p+1}, \ldots, v_r)$ é uma base de V. $\square$

Exemplo 3.2.55. *Consideremos, no espaço vetorial real* $\mathbb{R}^3$, *os vetores* $w_1 = (0, 1, -1)$ *e* $w_2 = (1, -1, -1)$. *A sequência* (w_1, w_2) *é linearmente independente, logo existe uma base de* $\mathbb{R}^3$ *da qual fazem parte estes vetores. Vamos determinar uma dessas bases seguindo o processo descrito na demonstração anterior.*

Sendo $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$, *a sequência* (e_1, e_2, e_3) *é uma base de* $\mathbb{R}^3$. *Uma vez que* $w_1 = e_2 - e_3$, *tem-se* $< e_1, e_2, e_3 >=< e_1, w_1, e_3 >$. *Agora, como* $w_2 = e_1 - e_2 - e_3$ *e* $e_2 = w_1 + e_3$, *vem* $w_2 = e_1 - w_1 - 2e_3$, *pelo que* $< e_1, w_1, e_3 >=< e_1, w_1, w_2 >$. *Logo,* $< e_1, e_2, e_3 >=< e_1, w_1, w_2 >$. *Pelo Teorema 3.2.45, conclui-se que a sequência* (e_1, w_1, w_2) *é linearmente independente e, portanto,* (e_1, w_1, w_2) *é uma base de* $\mathbb{R}^3$.

Teorema 3.2.56. *Sejam* V *um subespaço vetorial de* $\mathbb{R}^n$ *e* $(v_1, ..., v_r)$ *uma base de* V. *Então qualquer base de* V *tem exatamente* r *vetores.*

Demonstração. O resultado é imediato a partir do Teorema 3.2.43. □

O resultado anterior fundamenta a definição que se segue.

Definição 3.2.57. *Seja* V *um subespaço vetorial de* $\mathbb{R}^n$. *Chama-se* ***dimensão*** *de* V, *e representa-se por* $\dim V$, *ao número de elementos de uma sua qualquer base. Por convenção, diz-se ainda que* $\dim \{0_{\mathbb{R}^n}\} = 0$.

Exemplo 3.2.58. *Para* $n \in \mathbb{N}$, $\dim \mathbb{R}^n = n$.

Teorema 3.2.59. *Seja* V *um subespaço vetorial de* $\mathbb{R}^n$ *de dimensão* $r \geq 1$. *Então:*

1. *se* $v_1, ..., v_p$ *são* p *vetores de* V *com* $p > r$, *então* $(v_1, ..., v_p)$ *é linearmente dependente;*

2. *se* $(v_1, \ldots, v_r)$ *é uma sequência de vetores de* V *linearmente independente, então* $(v_1, \ldots, v_r)$ *é uma base de* V;

3. *se* $v_1, \ldots, v_r$ *são* r *vetores de* V, *distintos dois a dois, e* $\{v_1, \ldots, v_r\}$ *é um conjunto gerador de* V, *então* $(v_1, \ldots, v_r)$ *é uma base de* V.

Demonstração. O resultado *1.* é imediato a partir do Teorema 3.2.43, a alínea *2.* resulta dos teoremas 3.2.56 e 3.2.54, e *3.* resulta dos teoremas 3.2.56 e 3.2.51. □

Teorema 3.2.60. *Sejam* V *e* W *subespaços vetoriais de* $\mathbb{R}^n$ *tais que* $W \subseteq V$. *Então*

1. $\dim W \leq \dim V$;

2. *se* $\dim W = \dim V$, *então* $W = V$.

Demonstração. 1. A prova é feita com base no Teorema 3.2.43.
2. Resulta do teorema anterior. □

Teorema 3.2.61. *Seja* V *um subespaço vetorial de* $\mathbb{R}^n$. *Então existe um suplementar de* V *relativamente a* $\mathbb{R}^n$.

Demonstração. Seja V um subespaço vetorial de $\mathbb{R}^n$.
Se $V = \{0_{\mathbb{R}^n}\}$, então $\mathbb{R}^n$ é um suplementar de V relativamente a $\mathbb{R}^n$.
Se $V = \mathbb{R}^n$, então $\{0_{\mathbb{R}^n}\}$ é um suplementar de V relativamente a $\mathbb{R}^n$.
Se $V \neq \{0_{\mathbb{R}^n}\}$ e $V \neq \mathbb{R}^n$, seja $r = \dim V$ e $(v_1, \ldots, v_r)$ uma base de V. Uma vez que $V \neq \mathbb{R}^n$, temos $r < n$. Por outro lado, como $(v_1, \ldots, v_r)$ é linearmente independente, segue, pelo Teorema 3.2.54, que existe uma base de $\mathbb{R}^n$ da qual fazem parte os vetores $v_1, \ldots, v_r$. Suponha-se, sem perda de generalidade, que $(v_1, \ldots, v_r, u_1, \ldots, u_{n-r})$ é essa base. Seja $U =< u_1, \ldots, u_{n-r} >$. Fica ao cuidado do leitor a verificação de que $\mathbb{R}^n = V \oplus U$. □

Exemplo 3.2.62. *Consideremos, no espaço vetorial real* $\mathbb{R}^3$*, o subespaço* $V =< (0,1,-1), (1,-1,-1) >$. *A sequência* $((0,1,-1), (1,-1,-1))$ *é linearmente independente e* $((1,0,0), (0,1,-1), (1,-1,-1))$ *é uma base de* $\mathbb{R}^3$ *que inclui os vetores* $(0,1,-1)$, $(1,-1,-1)$ *(ver exemplo 3.2.55). Por conseguinte,* $U =< (1,0,0) >$ *é um espaço suplementar de* V *relativamente a* $\mathbb{R}^3$.

Teorema 3.2.63. *Sejam* V *e* W *subespaços vetoriais de* $\mathbb{R}^n$*. Então*

$$\mathrm{dim(V + W) = dimV + dimW - dim(V \cap W)}.$$

Demonstração. Sejam V e W subespaços vetoriais de $\mathbb{R}^n$.
Se $V = \{0_{\mathbb{R}^n}\}$ ou $W = \{0_{\mathbb{R}^n}\}$, o resultado é imediato.
Se $V \neq \{0_{\mathbb{R}^n}\}$ e $W \neq \{0_{\mathbb{R}^n}\}$, admitamos que $\mathrm{dimV = t} \geq 1$ e $\mathrm{dimW = k} \geq 1$. Sejam $(v_1, \ldots, v_t)$ uma base de V e $(w_1, \ldots, w_k)$ uma base de W. Então

$$V + W =< v_1, \ldots, v_t > + < w_1, \ldots, w_k >=< v_1, \ldots, v_t, w_1, \ldots, w_k > .$$

Consideremos, agora, dois casos:
1) $V \cap W = \{0_{\mathbb{R}^n}\}$: Neste caso, prova-se que $(v_1, \ldots, v_t, w_1, \ldots, w_k)$ é linearmente independente. De facto, dados $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_t \in \mathbb{R}$,

$$\begin{array}{rl}
& \beta_1 v_1 + \cdots + \beta_t v_t + \alpha_1 w_1 + \cdots \alpha_k w_k = 0_{\mathbb{R}^n} \\
\Rightarrow & \alpha_1 w_1 + \cdots + \alpha_k w_k = -\beta_1 v_1 - \cdots - \beta_t v_t \\
\Rightarrow & \alpha_1 w_1 + \cdots + \alpha_k w_k \in V \cap W \ \text{ e } \ -\beta_1 v_1 - \cdots - \beta_t v_t \in V \cap W \\
\Rightarrow & \alpha_1 w_1 + \cdots + \alpha_k w_k = 0_{\mathbb{R}^n} \ \text{ e } \ -\beta_1 v_1 - \cdots - \beta_t v_t = 0_{\mathbb{R}^n} \\
\Rightarrow & \alpha_1 = \cdots = \alpha_k = 0 \ \text{ e } \ -\beta_1 = \cdots = \beta_t = 0.
\end{array}$$

Assim, $(v_1, \ldots, v_t, w_1, \ldots, w_k)$ é uma base de $V + W$, e tem-se

$$\dim(V + W) = t + k = \dim V + \mathrm{dimW - dim(V \cap W)}.$$

2) Caso $V \cap W \neq \{0_{\mathbb{R}^n}\}$: Sejam $p = \mathrm{dim(V \cap W)}$ e $(u_1, \ldots, u_p)$ uma base de $V \cap W$. Uma vez que $(u_1, \ldots, u_p)$ é linearmente independente, existe uma base de W da qual fazem parte estes vetores; consideremos, sem perda de generalidade, que $(u_1, \ldots, u_p, w_1, \ldots, w_{k-p})$ é essa base. De modo análogo, existe uma base de V da qual fazem parte os vetores $u_1, \ldots, u_p$; consideremos, sem perda de generalidade, que $(u_1, \ldots, u_p, v_1, \ldots, v_{t-p})$ é essa base. Então

$$\begin{array}{rcl}
V + W & = & < u_1, \ldots, u_p, v_1, \ldots, v_{t-p} > + < u_1, \ldots, u_p, w_1, \ldots, w_{k-p} > \\
& = & < u_1, \ldots, u_p, w_1, \ldots, w_{k-p}, v_1, \ldots, v_{t-p} >
\end{array}$$

Vamos, agora, verificar que a sequência $(u_1, \ldots, u_p, w_1, \ldots, w_{k-p}, v_1, \ldots, v_{t-p})$ é linearmente independente. De facto, dados $\lambda_1, \ldots, \lambda_p, \alpha_1, \ldots, \alpha_{k-p}, \beta_1, \ldots, \beta_{t-p} \in \mathbb{R}$,

$$\begin{array}{rl} & \lambda_1 u_1 + \cdots + \lambda_p u_p + \alpha_1 w_1 + \cdots \alpha_{k-p} w_{k-p} + \beta_1 v_1 + \cdots + \beta_{t-p} v_{t-p} = 0_{\mathbb{R}^n} \\ \Rightarrow & \lambda_1 u_1 + \cdots + \lambda_p u_p + \alpha_1 w_1 + \cdots \alpha_{k-p} w_{k-p} = -\beta_1 v_1 - \cdots - \beta_{t-p} v_{t-p}. \end{array}$$

Então $-\beta_1 v_1 - \cdots - \beta_{t-p} v_{t-p} \in V \cap W$. Logo, existem $\lambda'_1, \ldots, \lambda'_p \in \mathbb{R}$ tais que

$$-\beta_1 v_1 - \cdots - \beta_{t-p} v_{t-p} = \lambda'_1 u_1 + \cdots + \lambda'_p u_n,$$

donde

$$\lambda'_1 u_1 + \ldots + \lambda'_p u_n + \beta_1 v_1 + \ldots + \beta_{t-p} v_{t-p} = 0_{\mathbb{R}^n}$$

e, como $(u_1, \ldots, u_p, v_1, \ldots, v_{t-p})$ é base de V, temos

$$\lambda'_1 = \ldots = \lambda'_p = \beta_1 = \ldots = \beta_{t-p} = 0.$$

De modo análogo, prova-se que $\lambda'_1 = \ldots = \lambda'_p = \alpha_1 = \ldots = \alpha_{k-p} = 0$. Uma vez que $\alpha_1 = \ldots = \alpha_{k-p} = \beta_1 = \ldots = \beta_{t-p} = 0$ e $(u_1, \ldots, u_p)$ é linearmente independente, tem-se também $\lambda_1 = \ldots = \lambda_p = 0$. Logo, a sequência

$$(u_1, \ldots, u_p, w_1, \ldots, w_{k-p}, v_1, \ldots, v_{t-p})$$

é linearmente independente. Portanto, $(u_1, \ldots, u_p, w_1, \ldots, w_{k-p}, v_1, \ldots, v_{t-p})$ é uma base de $V + W$, e tem-se

$$\begin{array}{rcl} \dim(V+W) & = & p + (k-p) + (t-p) \\ & = & k + t - p \\ & = & \dim W + \dim V - \dim(W \cap V). \end{array}$$

□

3.3 Dos espaços vetoriais $\mathbb{R}^n$ aos espaços vetoriais abstratos

As operações de adição de vetores de $\mathbb{R}^n$ e de multiplicação de um escalar por um vetor de $\mathbb{R}^n$, tal como definidas na secção anterior, satisfazem um conjunto de propriedades fundamentais: a adição é comutativa e associativa, existe um vetor nulo que atua como elemento neutro para a adição de vetores, cada vetor tem um oposto aditivo, a multiplicação por escalares distribui-se em relação à adição, entre outras. Estas propriedades não dependem da dimensão do espaço considerado, nem da interpretação geométrica, mas apenas da forma como as operações estão definidas. Por essa razão, é natural considerar mais abstratamente quaisquer conjuntos de objetos em que se possam definir operações de adição e multiplicação por escalares que satisfaçam as mesmas regras.

Definição 3.3.1. *Sejam V um conjunto não vazio, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ e*

$$\begin{array}{rcl} \widetilde{+} : V \times V & \to & V \\ (x,y) & \mapsto & x \widetilde{+} y \end{array}, \qquad \begin{array}{rcl} \widetilde{\cdot} : \mathbb{K} \times V & \to & V \\ (\alpha, x) & \mapsto & \alpha \widetilde{\cdot} x \end{array}$$

funções. Diz-se que $(V, \widetilde{+}, \widetilde{\cdot}, \mathbb{K})$ é um ***espaço vetorial sobre*** $\mathbb{K}$ *ou que V* ***juntamente com as aplicações*** $\widetilde{+}$ ***e*** $\widetilde{\cdot}$ ***é um espaço vetorial sobre*** $\mathbb{K}$ *se são satisfeitas as condições seguintes:*

(1) $\forall_{x,y\in V} \quad x\tilde{+}y = y\tilde{+}x$;

(2) $\forall_{x,y,z\in V} \quad x\tilde{+}\left(y\tilde{+}z\right) = \left(x\tilde{+}y\right)\tilde{+}z$;

(3) $\exists_{0_V\in V}\,\forall_{x\in V} \quad x\tilde{+}0_V = x = 0_V\tilde{+}x$;

(4) $\forall_{x\in V}\,\exists_{x'\in V} \quad x\tilde{+}x' = 0_V = x'\tilde{+}x$;

(5) $\forall_{x,y\in V}\,\forall_{\alpha\in\mathbb{K}} \quad \alpha\tilde{\cdot}\left(x\tilde{+}y\right) = \alpha\tilde{\cdot}x\tilde{+}\alpha\tilde{\cdot}y$;

(6) $\forall_{x\in V}\,\forall_{\alpha,\beta\in\mathbb{K}} \quad (\alpha+\beta)\,\tilde{\cdot}x = \alpha\tilde{\cdot}x\tilde{+}\beta\tilde{\cdot}x$;

(7) $\forall_{x\in V}\,\forall_{\alpha,\beta\in\mathbb{K}} \quad (\alpha\cdot\beta)\,\tilde{\cdot}x = \alpha\tilde{\cdot}\left(\beta\tilde{\cdot}x\right)$;

(8) $\forall_{x\in V} \quad 1\tilde{\cdot}x = x$.

Notação e terminologia:

- Seja $(V,\tilde{+},\tilde{\cdot},\mathbb{K})$ um espaço vetorial sobre $\mathbb{K}$. Para simplificar a linguagem, em vez de dizermos que um conjunto V juntamente com as aplicações $\tilde{+}$ e $\tilde{\cdot}$ é um espaço vetorial sobre $\mathbb{K}$, dizemos apenas que V é um espaço vetorial sobre $\mathbb{K}$ (subentendendo as operações envolvidas).

- Um espaço vetorial sobre $\mathbb{R}$ diz-se um ***espaço vetorial real*** e a um espaço vetorial sobre $\mathbb{C}$ dá-se a designação de ***espaço vetorial complexo***.

- Aos elementos de V dá-se o nome de ***vetores*** e aos elementos de $\mathbb{K}$ o de ***escalares***. O elemento 0_V indicado na propriedade (V_3) da definição de espaço vetorial é único. Ao elemento 0_V dá-se a designação de ***vetor nulo*** e ao zero de $\mathbb{K}$ damos o nome de ***escalar nulo***. Desde que não exista ambiguidade podemos representar tanto o vetor nulo como o escalar nulo por 0.

- A operação $\tilde{+}$ designa-se por ***adição de vetores*** e a operação $\tilde{\cdot}$ por ***multiplicação de um escalar por um vetor***. Simplificamos também a notação, escrevendo $+$ quer se trate da adição em $\mathbb{K}$ quer se trate da adição de vetores e escrevemos $\cdot$ quer seja a multiplicação em $\mathbb{K}$ quer o produto de um escalar por um vetor.

- Para cada $x\in V$ e para cada $\alpha\in\mathbb{K}$, é usual escrever αx para representar $\alpha\cdot x$.

- Para cada $x\in V$, ao elemento x' determinado na condição (4), chama-se ***simétrico de*** x e representa-se por $-x$.

- Dados $x,y\in V$, representa-se por $x-y$ o elemento $x+(-y)$.

Exemplo 3.3.2. *O espaço vetorial real* $(\mathbb{R}^n,+,\cdot,\mathbb{R})$.

Exemplo 3.3.3. *O conjunto* $\mathbb{R}_2[x] = \{ax^2+bx+c : a,b,c\in\mathbb{R}\}$ *dos polinómios, na indeterminada* x *e com coeficientes reais, que têm grau menor ou igual a 2, algebrizado com as operações* $+ : \mathbb{R}_2[x]\times\mathbb{R}_2[x]\to\mathbb{R}_2[x]$ *e* $\cdot : \mathbb{K}\times\mathbb{R}_2[x]\to\mathbb{R}_2[x]$*, definidas, respetivamente, por*

- $(a_1x^2 + b_1x + c_1) + (a_2x^2 + b_2x_2 + c_2) = (a_1 + a_2)x^2 + (b_1 + b_2)x + (c_1 + c_2)$, *para quaisquer* $a_1, a_2, b_1, b_2, c_1, c_2 \in \mathbb{R}$,
- $\alpha \cdot (ax^2 + bx + c) = (\alpha a)x^2 + (\alpha b)x + \alpha c$, *para quaisquer* $a, b, c \in \mathbb{R}$ *e* $\alpha \in \mathbb{R}$.

é um espaço vetorial real.

Exemplo 3.3.4. *O conjunto* $\mathbb{R}[x]$ *de todos os polinómios na ideterminada* x *e de coeficientes reais, com a adição usual de polinómios e a multiplicação de um número real por um polinómio, é um espaço vetorial real.*

Exemplo 3.3.5. *Sejam* $m, n \in I\!N$. *O conjunto* $\mathcal{M}_{m\times n}(\mathbb{R})$, *das matrizes reais de ordem* $m \times n$, *algebrizado com a adição de matrizes e a multiplicação de um real por uma matriz, é um espaço vetorial real.*

No contexto de $\mathbb{R}^n$, estudámos noções tais como subespaço vetorial, conjunto gerador de um subespaço, sequências de vetores linearmente independentes, etc. Na apresentação de tais conceitos e dos teoremas e demonstrações a eles associados, só utilizámos as propriedades da adição de vetores e da mutiplicação de escalares por vetores, não sendo feita referência à natureza dos elementos de $\mathbb{R}^n$. Tal sugere que no estudo dos espaços vetoriais arbitrários podem ser definidos conceitos análogos. Porém, como o estudo de espaços vetoriais arbitrários não faz parte do âmbito do programa deste curso, ficará ao cuidado do leitor um estudo mais geral sobre espaços vetoriais.

3.4 Relação entre $\mathbb{R}^n$ e os espaços vetoriais de matrizes

Nesta secção, destacamos a relação existente entre os espaços vetoriais $\mathbb{R}^n$ e os espaços vetoriais de matrizes.

Do que foi estudado no primeiro capítulo, é simples concluir que o quádruplo $(\mathcal{M}_{m\times n}(\mathbb{R}), +, \cdot, \mathbb{R})$, onde $+$ representa a adição de matrizes e $\cdot$ a multiplicação de um escalar por uma matriz, é um espaço vetorial sobre $\mathbb{R}$.

Em particular, $\mathcal{M}_{n\times 1}(\mathbb{R})$ e $\mathcal{M}_{1\times n}(\mathbb{R})$ são espaços vetoriais. Estes espaços são naturalmente identificados com $\mathbb{R}^n$. De facto, um vetor de $\mathbb{R}^n$ pode ser representado de diferentes formas: como matriz coluna, matriz linha ou n-uplo ordenado. Todas essas representações descrevem a mesma estrutura algébrica, pois as operações de adição e multiplicação por escalar atuam componente a componente.
Por exemplo, na forma de matriz coluna:

$$u + v = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix},$$

sendo o raciocínio análogo para matrizes linha ou n-uplos.
Dessa forma, identificamos os espaços $\mathcal{M}_{n\times 1}(\mathbb{R})$, $\mathcal{M}_{1\times n}(\mathbb{R})$ e $\mathbb{R}^n$, escolhendo a forma mais conveniente conforme o contexto. Ao longo deste texto, usaremos a notação $\mathbb{R}^n$ para qualquer uma dessas representações, salvo quando a forma exata do vetor for relevante, como no produto de matrizes.

Dada uma matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, também podemos associar cada linha i de A ao vetor $(a_{i1}, \ldots, a_{in}) \in \mathbb{R}^n$ e cada coluna j ao vetor $(a_{1j}, \ldots, a_{mj}) \in \mathbb{R}^m$.
O subespaço de $\mathbb{R}^n$ gerado pelas linhas de A é chamado **espaço das linhas** de A e denotado por $\mathcal{L}(A)$. O subespaço de $\mathbb{R}^m$ gerado pelas colunas de A é chamado **espaço das colunas** de A e denotado por $\mathcal{C}(A)$.
As dimensões de $\mathcal{L}(A)$ e $\mathcal{C}(A)$ são, respectivamente, chamadas **característica linha** e **característica coluna** de A, representadas por $car_l(A)$ e $car_c(A)$.

Pelo Teorema 3.2.59 é simples perceber que a característica linha de uma matriz A é igual ao número máximo de linhas de A que são linearmente independentes e, analogamente, a característica coluna de A é igual ao número máximo de colunas de A que são linearmente independentes.

Exemplo 3.4.1. *Se*

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad e \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

tem-se

$$\mathcal{L}(A) =< (2,0,0), (0,0,1), (0,0,0) >, \quad \mathcal{L}(B) =< (2,0,0), (0,1,0), (0,0,0) >$$

e

$$\mathcal{C}(A) =< (2,0,0), (0,0,0), (0,1,0) >, \quad \mathcal{C}(B) =< (2,0,0), (0,1,0), (0,0,0) > .$$

A respeito destes espaços vetoriais facilmente se verifica que $\mathcal{L}(A) \neq \mathcal{L}(B)$, $\mathcal{C}(A) = \mathcal{C}(B)$, $car_l(A) = car_c(A)$ e $car_l(B) = car_c(B)$.

Observação: Para qualquer matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, tem-se $\mathcal{C}(A) = \mathcal{L}(A^T)$.

Teorema 3.4.2. *Sejam $m, n \in \mathbb{N}$ e $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$. Se B é uma matriz obtida de A por meio de uma operação elementar sobre linhas, então $car_l(A) = car_l(B)$.*

Demonstração. Se B é uma matriz obtida de A por meio de uma operação elementar sobre linhas, então do Teorema 3.2.33 segue que $\mathcal{L}(A) = \mathcal{L}(B)$. Logo, $car_l(A) = car_l(B)$. $\square$

Teorema 3.4.3. *Sejam $m, n \in \mathbb{N}$ e $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$. Se B é uma matriz obtida de A por meio de uma operação elementar sobre linhas, então $car_c(A) = car_c(B)$.*

Demonstração. Sejam $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$ tais que B é obtida de A por meio de uma operação elementar sobre linhas.
Pretendemos mostrar que $\dim \mathcal{C}(A) = \dim \mathcal{C}(B)$, ou seja, que o número máximo de colunas de A linearmente independentes é igual ao número máximo de colunas de B linearmente independentes. Para tal, basta mostrar que, para qualquer $k \in \{1, \ldots, n\}$, qualquer sequência $(A_{j_1}, A_{j_2}, \ldots, A_{j_k})$ com k colunas de A e quaisquer $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$,

$$\alpha_1 A_{j_1} + \alpha_2 A_{j_2} + \ldots \alpha_k A_{j_k} = \mathbf{0}$$

se e só se

$$\alpha_1 B_{j_1} + \alpha_2 B_{j_2} + \ldots \alpha_k B_{j_k} = \mathbf{0},$$

onde A_{j_i} e B_{j_i} representam a coluna j_i de A e de B, respetivamente. Ora, como $B = EA$ para alguma matriz elementar E, tem-se

$$[B_{j_1} B_{j_2} \ldots B_{j_k}] = E[A_{j_1} A_{j_2} \ldots A_{j_k}],$$

onde $[A_{j_1} A_{j_2} \ldots A_{j_k}]$ representa a matriz cuja coluna i é a coluna A_{j_i} de A e $[B_{j_1} B_{j_2} \ldots B_{j_k}]$ é a matriz construída de modo análogo a partir da matriz B. Por conseguinte, e tendo em conta que E é invertível, segue que

$$\alpha_1 A_{j_1} + \alpha_2 A_{j_2} + \ldots \alpha_k A_{j_k} = 0$$

se e só se

$$[A_{j_1} A_{j_2} \ldots A_{j_k}] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \mathbf{0}$$

se e só se

$$E[A_{j_1} A_{j_2} \ldots A_{j_k}] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \mathbf{0}$$

se e só se

$$[B_{j_1} B_{j_2} \ldots B_{j_k}] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \mathbf{0}$$

se e só se

$$\alpha_1 B_{j_1} + \alpha_2 B_{j_2} + \ldots \alpha_k B_{j_k} = 0.$$

Do que acabámos de provar segue que o número máximo de colunas de A linearmente independentes é igual ao número máximo de colunas de B linearmente independentes e, por conseguinte, $car_c(A) = car_c(B)$. □

Exemplo 3.4.4. *Sejam*

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \quad e \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

A matriz B é equivalente por linhas à matriz A, uma vez que

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \xrightarrow[l_2 \to l_2 - 2l_1]{} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Tem-se $\mathcal{L}(A) = \mathcal{L}(B)$ e, portanto, $car_l(A) = car_l(B)$. Embora $\mathcal{C}(A) \neq \mathcal{C}(B)$ (pois $(1,2) \in \mathcal{C}(A)$, mas $(1,2) \notin \mathcal{C}(B)$), também temos $car_c(A) = car_c(B)$.

Como vamos verificar nos resultados seguintes, a noção de caracteristica linha e de caracaterística coluna estão relacionadas com a noção de característica de uma matriz.

Teorema 3.4.5. *Sejam $m, n \in \mathbb{N}$. Se $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ é uma matriz em forma de escada, então a sua característica linha e a sua característica coluna são iguais e coincidem com a característica de A, isto é,*

$$car_l(A) = car_c(A) = car(A).$$

Demonstração. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ uma matriz em escada. Então temos

$$A = \begin{bmatrix} 0 & \dots & a_{1j_1} & \dots & a_{1j_2} & \dots & a_{1j_3} & \dots & a_{1j_m} & \dots & a_{1n} \\ 0 & \dots & 0 & \dots & a_{2j_2} & \dots & a_{2j_3} & \dots & a_{2j_m} & \dots & a_{2n} \\ 0 & \dots & 0 & \dots & 0 & \dots & a_{3j_3} & \dots & a_{3j_m} & \dots & a_{3n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & a_{rj_r} & \dots & a_{rn} \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ . & . & . & . & . & . & . & . & . & . & . \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{bmatrix}$$

onde:

- para todo $i \in \{1, \dots, r\}$, $a_{ij_i} \neq 0$,
- para todo $r < i \leq n$, a linha i é nula,
- para todo $j \in \{1, \dots, n\} \setminus (\{j_1, j_2, \dots, j_r\} \cup \{j_{r+1}, \dots, n\})$, a coluna j é nula.

Pretendemos mostrar que $car(A) = car_l(A) = car_c(A)$. Para tal, começamos por observar que $car(A) = r$, uma vez que A é uma matriz em escada com r linhas não nulas.

Como vamos verificar de seguida, também temos $car_l(A) = r$. De facto, se representarmos por l_i o vetor de $\mathbb{R}^n$ que representa a linha i de A, temos

$$\mathcal{L}(A) =< l_1, l_2, \dots, l_r, \dots, l_m >=< l_1, l_2, l_3, \dots, l_r >,$$

uma vez que que, que para todo $i > r$, $l_i = 0_{\mathbb{R}^n}$. Por outro lado, é simples verificar que a sequência $(l_1, l_2, l_3, \dots, l_r)$ é linearmente independente, pois, para quaisquer $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{R}$,

$$\alpha_1 l_1 + \alpha_2 l_2 + \alpha_3 l_3 \dots + \alpha_r l_r = 0_{\mathbb{R}^n}$$

se e só se

$$\begin{array}{rlcccccccccccc}
 & \alpha_1(0, & \ldots, & a_{1j_1}, & \ldots, & a_{1j_2}, & \ldots, & a_{1j_3}, & \ldots, & a_{1j_m}, & \ldots, & a_{1n}) \\
+ & \alpha_2(0, & \ldots, & 0, & \ldots, & a_{2j_2}, & \ldots, & a_{2j_3}, & \ldots, & a_{2j_m}, & \ldots, & a_{2n}) \\
+ & \alpha_3(0, & \ldots, & 0, & \ldots, & 0, & \ldots, & a_{3j_3}, & \ldots, & a_{3j_m}, & \ldots, & a_{3n}) \\
+ & \ldots \\
+ & \alpha_r(0, & \ldots, & 0, & \ldots, & 0, & \ldots, & 0, & \ldots, & a_{rj_r}, & \ldots, & a_{rn}) \\
= & (0, & \ldots, & 0, & \ldots, & 0, & \ldots, & 0, & \ldots, & 0, & \ldots, & 0)
\end{array}$$

se e só se

$$\begin{array}{rl}
 & (0, \ldots, \alpha_1 a_{1j_1}, \ldots, \sum_{i=1}^{2} \alpha_i a_{ij_2}, \ldots, \sum_{i=1}^{3} \alpha_i a_{ij_3}, \ldots, \sum_{i=1}^{r} \alpha_i a_{ij_r}, \ldots, \sum_{i=1}^{r} \alpha_i a_{ij_n}) \\
= & (0, \ldots, 0, \ldots, 0, \ldots, 0, \ldots, 0, \ldots, 0)
\end{array}$$

se e só se

$$\alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_r = 0.$$

Logo, $(l_1, l_2, l_3, \ldots, l_r)$ é uma base de $\mathcal{L}(A)$ e, por conseguinte, $dim\, \mathcal{L}(A) = r$, isto é, $car_l(A) = r$. Assim, $car_l(A) = car(A)$.

Facilmente também provamos que $car_c(A) = r$. Com efeito, se representarmos por c_j o vetor de $\mathbb{R}^m$ que representa a coluna j da matriz A, é imediato que

$$\mathcal{C}(A) =< c_1, c_2, c_3, \ldots, c_n >=< c_{j_1}, c_{j_2}, c_{j_3}, \ldots, c_{j_r}, c_{j_{r+1}}, \ldots, c_n >,$$

uma vez que, para todo $j \in \{1, \ldots, n\} \setminus (\{j_1, j_2, j_3, \ldots, j_r\} \cup \{j_{r+1}, \ldots, n\})$, $c_j = 0_{\mathbb{R}^m}$. Além disso, verifica-se que, para todo $k \geq r + 1$, c_k é combinação linear de $c_{j_1}, c_{j_2}, c_{j_3}, \ldots, c_{j_r}$, ou seja, existem $\alpha_1, \alpha_2, \alpha_3 \ldots \alpha_r \in \mathbb{R}$ tais que

$$\alpha_1 c_{j_1} + \alpha_2 c_{j_2} + \alpha_3 c_{j_3} + \ldots + \alpha_r c_{j_r} = c_k.$$

De facto, se representarmos por A' a matriz que tem as colunas $j_1, j_2, j_3, \ldots, j_r$ de A e por A_k a matriz coluna com a coluna k de A, i.e.,

$$A' = \begin{bmatrix} a_{1j_1} & a_{1j_2} & a_{1j_3} & a_{1j_r} \\ 0 & a_{2j_2} & a_{2j_3} & a_{2j_r} \\ 0 & 0 & a_{3j_3} & a_{3j_r} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_{rj_r} \\ 0 & 0 & 0 & 0 \\ . & . & . & . \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad e \quad A_k = \begin{bmatrix} a_{1k} \\ a_{2k} \\ a_{3k} \\ \vdots \\ a_{rk} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

c_k é combinação linear de $c_{j_1}, c_{j_2}, \ldots, c_{j_r}$ se e só se o sistema

$$A' \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_r \end{bmatrix} = A_k$$

é possivel. Ora, para todo $i \in \{1, \ldots, r\}$, $a_{ij_i} \neq 0$, pelo que $car(A') = r = car([A'|A_k])$ e, portanto, o sistema anterior é possivel.
Então, considerando que, para todo $k \geq r+1$, c_k é combinação linear de $c_{j_1}, c_{j_2}, c_{j_3}, \ldots, c_{j_r}$, temos

$$\mathcal{C}(A) =< c_{j_1}, c_{j_2}, \ldots, c_{j_r}, c_{j_{r+1}}, \ldots, c_n >=< c_{j_1}, c_{j_2}, \ldots, c_{j_r} > .$$

Por último, e de forma semelhante ao que foi feito no caso das linhas, prova-se que a sequência $(c_{j_1}, c_{j_2}, \ldots, c_{j_r})$ é linearmente independente. Logo, $(c_{j_1}, c_{j_2}, \ldots, c_{j_r})$ é uma base de $\mathcal{C}(A)$, pelo que $dim\, \mathcal{C}(A) = r$. Assim, $car_c(A) = car(A)$. □

O resultado anterior pode ser generalizado para qualquer matriz.

Teorema 3.4.6. *Sejam $m, n \in \mathbb{N}$. Para qualquer matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, a sua característica linha e a sua característica coluna são iguais e coincidem com a característica de A, isto é,*

$$car_l(A) = car_c(A) = car(A).$$

Demonstração. Por definição de característica de uma matriz, tem-se $car(A) = car(U)$ onde U é uma matriz em escada obtida de A por meio de operações elementares sobre linhas. Por outro lado, pela Proposição 3.4.5, tem-se $car(U) = car_l(U)$. Finalmente, da Proposição 3.4.2 segue que $car_l(A) = car_l(U)$. Logo $car(A) = car_l(A)$.
De forma simples também provamos que $car_c(A) = car(A)$. Com efeito, pela Proposição 3.4.5 sabemos que $car_c(U) = car(U)$ e pela Proposição 3.4.3 tem-se $car_c(A) = car_c(U)$. Logo, como $car(A) = car(U)$, temos $car_c(A) = car(A)$. □

Teorema 3.4.7. *Sejam $m, n \in \mathbb{N}$. Para qualquer matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, tem-se*

$$car(A) = car(A^T).$$

Demonstração. O resultado segue de imediato, uma vez que

$$car(A) = car_l(A) = car_c(A) = car_l(A^T) = car(A^T).$$

□

4. Aplicações Lineares

4.1 Definições e propriedades

Uma aplicação linear é uma função entre espaços vetoriais que preserva a estrutura desses espaços. Embora este conceito possa ser formulado em geral para aplicações entre quaisquer espaços vetoriais, neste capítulo restringimos o estudo a aplicações lineares definidas entre subespaços de $\mathbb{R}^n$ e subespaços de $\mathbb{R}^m$.

Definição 4.1.1. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$. Uma aplicação $f : V \to V'$ diz-se uma **aplicação linear** (ou **transformação linear** ou **homomorfismo**) de V em V' se*

i) $\forall_{x,y\in V}\ f(x+y) = f(x) + f(y)$;

ii) $\forall_{x\in V}\ \forall_{\lambda\in\mathbb{R}}\ f(\lambda x) = \lambda f(x)$.

O conjunto de todas as aplicações lineares de V em V' é representado por $\mathcal{L}(V, V')$.

Da definição anterior resulta facilmente a seguinte caracterização para as aplicações lineares.

Teorema 4.1.2. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$. Então uma aplicação $f : V \to V'$ é uma aplicação linear se e só se*

$$\forall_{x,y\in V}\,\forall_{\alpha,\beta\in\mathbb{R}}\ f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

Demonstração. Exercício. □

Exemplo 4.1.3. *Consideremos os espaços vetoriais reais $\mathbb{R}^2$ e $\mathbb{R}^3$. A aplicação $f : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $f(a, b) = (2a, a - b, a + 3b)$, para todo $(a, b) \in \mathbb{R}^2$, é uma aplicação linear de $\mathbb{R}^2$ em $\mathbb{R}^3$. De facto, para quaisquer $(a, b), (a', b') \in \mathbb{R}^2$ e qualquer $\lambda \in \mathbb{R}$, tem-se*

$$\begin{aligned} f((a,b)+(a',b')) &= f(a+a', b+b') \\ &= (2(a+a'), (a+a')-(b+b'), (a+a')+3(b+b')) \\ &= (2a+2a', (a-b)+(a'-b'), (a+3b)+(a'+3b')) \\ &= (2a, a-b, a+3b) + (2a', a'-b', a'+3b') \\ &= f(a,b) + f(a',b') \end{aligned}$$

$$\begin{array}{rcl} f(\lambda(a,b)) & = & f(\lambda a, \lambda b) \\ & = & (2(\lambda a), \lambda a - \lambda b, \lambda a + 3(\lambda b)) \\ & = & (\lambda(2a), \lambda(a-b), \lambda(a+3b)) \\ & = & \lambda(2a, a-b, a+3b) \\ & = & \lambda f(a,b) \end{array}$$

Exemplo 4.1.4. *A aplicação $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(a,b) = 2a+3$, para todo $(a,b) \in \mathbb{R}^2$, não é uma aplicação linear de $\mathbb{R}^2$ em $\mathbb{R}$. Dados, por exemplo, $x = (1,0)$ e $y = (2,1) \in \mathbb{R}^2$, tem-se*

$$f((1,0)+(2,1)) = f(3,1) = 9 \text{ e } f(1,0)+f(2,1) = (2+3)+(4+3) = 12,$$

i.e., existem $x, y \in \mathbb{R}^2$ tais que $f(x+y) \neq f(x)+f(y)$.

Exemplo 4.1.5. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$. As aplicações*

$$\begin{array}{rccc} 0_{\mathcal{L}(V,V')} : & V & \to & V' \\ & x & \mapsto & 0_{\mathbb{R}^m} \end{array} \quad e \quad \begin{array}{rccc} id_V : & V & \to & V \\ & x & \mapsto & x \end{array}$$

são aplicações lineares designadas, respectivamente, por ***aplicação linear nula*** *de V em V' e* ***aplicação identidade*** *em V.*

Teorema 4.1.6. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$ e $f \in \mathcal{L}(V,V')$. Então*

1. *$f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$;*
2. *Para todo $x \in V$, $f(-x) = -f(x)$;*
3. *Para todos $x_1, \ldots, x_k \in V$, $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$,*

$$f(\alpha_1 x_1 + \ldots + \alpha_k x_k) = \alpha_1 f(x_1) + \ldots + \alpha_k f(x_k).$$

Demonstração. 1. $f(0_{\mathbb{R}^n}) = f(0 \cdot 0_{\mathbb{R}^n}) = 0 \cdot f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$.
2. Para todo $x \in V$, tem-se $f(-x) = f((-1)x) = (-1)f(x) = -f(x)$.
3. A prova é feita por indução em k. □

Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$. O resultado seguinte estabelece que uma aplicação linear $f : V \to V'$ fica completamente determinada se conhecermos a imagem, por f, dos vetores de uma base de V.

Teorema 4.1.7 (Teorema da Extensão Linear)**.** *Sejam V um subespaço vetorial de $\mathbb{R}^n$, $(v_1, \ldots, v_r)$ uma base de V, V' um subespaço vetorial de $\mathbb{R}^m$ e $v'_1, \ldots, v'_r \in V'$. Então existe uma e uma só aplicação linear f de V em V' tal que $f(v_i) = v'_i$, para todo $i \in \{1, \ldots, r\}$.*

Demonstração. Sejam $(v_1, \dots v_r)$ uma base de V e $v'_1, \dots, v'_r \in V'$. Cada vetor v de V escreve-se de modo único único como combinação linear dos vetores $v_1, \dots, v_r$, i.e., v pode escrever-se na forma

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r,$$

sendo os escalares univocamente determinados. Pode, então, definir-se uma correspondência de V em V', associando a cada vetor $v \in V$ o vetor $z = f(v) \in V'$, definido do seguinte modo,

$$z = \lambda_1 v'_1 + \dots + \lambda_n v'_r.$$

Como $z \in V'$ e é bem determinado, a correspondência f assim definida é uma aplicação.

Vamos agora provar que f é uma aplicação linear. Sejam $x, y \in V$ e $\alpha, \beta \in \mathbb{R}$. Então existem $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_r \in \mathbb{R}$ tais que

$$x = \alpha_1 v_1 + \dots + \alpha_r v_r \ \text{ e } \ y = \beta_1 v_1 + \dots + \beta_r v_r.$$

Logo,

$$\begin{aligned} f(\alpha x + \beta y) &= f((\alpha\alpha_1 + \beta\beta_1)v_1 + \dots + (\alpha\alpha_r + \beta\beta_r)v_r) \\ &= (\alpha\alpha_1 + \beta\beta_1)v'_1 + \dots + (\alpha\alpha_r + \beta\beta_r)v'_r \\ &= \alpha(\alpha_1 v'_1 + \dots + \alpha_r v'_r) + \beta(\beta_1 v'_1 + \dots + \beta_r v'_r) \\ &= \alpha f(x) + \beta f(y). \end{aligned}$$

Portanto, f é uma aplicação linear.

É também simples verificar que $f(v_i) = v'_i$, para todo $i \in \{1, \dots, r\}$. De facto,

$$\begin{aligned} f(v_i) &= f(0v_1 + \dots + 0v_{i-1} + 1v_i + 0v_{i+1} + \dots + 0v_r) \\ &= 0v'_1 + \dots + 0v'_{i-1} + 1v'_i + 0v'_{i+1} + \dots + 0v'_r \\ &= v'_i \end{aligned}$$

Falta mostrar que f, assim definida, é a única aplicação linear que satisfaz as condições requeridas. Suponhamos que $g : V \to V'$ é uma aplicação linear tal que $g(v_i) = v'_i$, para todo $i \in \{1, \dots, r\}$. Então, para quaisquer $\alpha_1, \dots, \alpha_r \in \mathbb{R}$, tem-se

$$\begin{aligned} g(\alpha_1 v_1 + \dots + \alpha_r v_r) &= \alpha_1 g(v_1) + \dots + \alpha_r g(v_r) \\ &= \alpha_1 f(v_1) + \dots + \alpha_r f(v_r) \\ &= f(\alpha_1 v_1 + \dots + \alpha_r v_r). \end{aligned}$$

Logo $g = f$.

Existe, portanto, uma única aplicação linear $f : V \to V'$ tal que $f(v_i) = v'_i$, para todo $i \in \{1, \dots, r\}$. □

Exemplo 4.1.8. *Consideremos os espaços vetoriais reais* $\mathbb{R}^3$ *e* $\mathbb{R}^2$ *e a base* $((1,1,1), (1,1,0), (1,0,0))$ *de* $\mathbb{R}^3$. *Pelo teorema anterior, existe uma e uma só aplicaçao linear* f *de* $\mathbb{R}^3$ *em* $\mathbb{R}^2$ *tal que*

$$f(1,1,1) = (2,1), \quad f(1,1,0) = (-4,1), \quad f(1,0,0) = (0,3).$$

Com base nas imagens dos vetores da base de $\mathbb{R}^3$, *facilmente determinamos* $f(2,0,1)$. *De facto, como*

$$(2,0,1) = 1(1,1,1) + (-1)(1,1,0) + 2(1,0,0)$$

tem-se

$$\begin{aligned} f(2,0,1) &= 1f(1,1,1) + (-1)f(1,1,0) + 2f(1,0,0) \\ &= 1(2,1) + (-1)(-4,1) + 2(0,3) = (6,6). \end{aligned}$$

Conhecidas as imagens dos vetores de uma base de $\mathbb{R}^3$*, pode-se determinar a imagem de qualquer vetor de* $\mathbb{R}^3$*. Dado* $(a,b,c) \in \mathbb{R}^3$*, escrevemos* (a,b,c) *como combinação linear dos vetores* $(1,1,1)$*,* $(1,1,0)$*,* $(1,0,0)$ *e determinamos* $f(a,b,c)$*. Uma vez que*

$$(a,b,c) = c(1,1,1) + (b-c)(1,1,0) + (a-b)(1,0,0),$$

segue que

$$\begin{aligned} f(a,b,c) &= cf(1,1,1) + (b-c)f(1,1,0) + (a-b)f(1,0,0) \\ &= c(2,1) + (b-c)(-4,1) + (a-b)(0,3) \\ &= (-4b+6c, 3a-2b). \end{aligned}$$

4.2 Operações com aplicações lineares

A partir de aplicações lineares dadas podem construir-se outras. Estudamos seguidamente algumas operações envolvendo aplicações lineares.

Começamos por definir as operações de *adição de aplicações lineares* e de *multiplicação de um escalar por uma aplicação linear*, as quais permitem dar a estrutura de espaço vetorial ao conjunto $\mathcal{L}(V,V')$ de todas aplicações lineares de um subespaço vetorial V de $\mathbb{R}^n$ num subespaço vetorial V' de $\mathbb{R}^m$.

Definição 4.2.1. *Sejam* V *um subespaço vetorial de* $\mathbb{R}^n$ *e* V' *um subespaço vetorial de* $\mathbb{R}^m$*,* $f,g \in \mathcal{L}(V,V')$ *e* $\lambda \in \mathbb{R}$*. Designa-se por:*

- ***soma de*** f ***e*** g *a aplicação* $f+g : V \to V'$ *definida por* $(f+g)(x) = f(x)+g(x)$*, para todo* $x \in V$*.*
- ***produto de*** λ ***por*** f *a aplicação* $\lambda f : V \to V'$ *definida por* $(\lambda f)(x) = \lambda f(x)$*, para todo* $x \in V$*.*

Nas condições da definição anterior, é óbvio que $f+g$ e λf são aplicações e é simples provar que estas aplicações também são aplicações lineares.

Teorema 4.2.2. *Sejam* V *um subespaço vetorial de* $\mathbb{R}^n$ *e* V' *um subespaço vetorial de* $\mathbb{R}^m$*,* $f,g \in \mathcal{L}(V,V')$ *e* $\lambda \in \mathbb{R}$*. Então* $f+g \in \mathcal{L}(V,V')$ *e* $\lambda f \in \mathcal{L}(V,V')$*.*

Demonstração. As aplicações $f+g$ e λf são aplicações lineares de V em V'. De facto, para quaisquer $x,y \in V$ e para quaisquer $\alpha,\beta \in \mathbb{R}$, temos

$$\begin{aligned} (f+g)(\alpha x+\beta y) &= f(\alpha x+\beta y) + g(\alpha x+\beta y) && \text{(definição de } f+g\text{)} \\ &= (\alpha f(x)+\beta f(y)) + (\alpha g(x)+\beta g(y)) && \text{(Teorema 6.1.17)} \\ &= (\alpha f(x)+\alpha g(x)) + (\beta f(y)+\beta g(y)) && \text{(propriedades de } \mathbb{R}^m\text{)} \\ &= \alpha(f(x)+g(x)) + \beta(f(y)+g(y)) && \text{(propriedades de } \mathbb{R}^m\text{)} \\ &= \alpha((f+g)(x)) + \beta((f+g)(y)) && \text{(definição de } f+g\text{)} \end{aligned}$$

o que permite concluir que $f+g$ é aplicação linear de V em V'.

Relativamente a λf tem-se o seguinte

$$\begin{array}{rcll}(\lambda f)(\alpha x+\beta y) & = & \lambda(f(\alpha x+\beta y)) & \text{(definição de } \lambda f\\ & = & \lambda(\alpha f(x)+\beta f(y)) & \text{(Teorema 6.1.17)}\\ & = & \lambda(\alpha f(x))+\lambda(\beta f(y)) & \text{(propriedades de } \mathbb{R}^m)\\ & = & \alpha(\lambda(f(x)))+\beta(\lambda f(y)) & \text{(propriedades de } \mathbb{R}^m)\\ & = & \alpha((\lambda f)(x))+\beta((\lambda f)(y)) & \text{(definição de } \lambda f)\end{array}$$

e, portanto, λf é uma aplicação linear. □

Definição 4.2.3. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, V'' um subespaço vetorial de $\mathbb{R}^p$, $f\in\mathcal{L}(V,V')$ e $g\in\mathcal{L}(V',V'')$. Designa-se por **composta de** g **com** f a aplicação $g\circ f: V\to V''$ definida por*

$$(g\circ f)(x)=g(f(x)),$$

para todo $x\in V$.

Teorema 4.2.4. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, V'' um subespaço vetorial de $\mathbb{R}^p$, $f\in\mathcal{L}(V,V')$ e $g\in\mathcal{L}(V',V'')$. Então $g\circ f\in\mathcal{L}(V,V'')$.*

Demonstração. Considerando que f é uma aplicação de V em V' e g é uma aplicação de V' em V'', então, por definição de composição de funções, $g\circ f$ é uma aplicação de V em V''. Além disso,

$$\begin{array}{rcll}\forall_{x,y\in V}\ (g\circ f)(x+y) & = & g(f(x+y)) & \text{(definição de } g\circ f)\\ & = & g(f(x)+f(y)) & (f \text{ é aplicação linear})\\ & = & g(f(x))+g(f(y)) & (g \text{ é aplicação linear})\\ & = & (g\circ f)(x)+(g\circ f)(y) & \text{(definição de } g\circ f)\end{array}$$

e

$$\begin{array}{rcll}\forall_{x\in V}\,\forall_{\lambda\in\mathbb{R}}\ (g\circ f)(\lambda x) & = & g(f(\lambda x)) & \text{(definição de } g\circ f)\\ & = & g(\lambda f(x)) & (f \text{ é aplicação linear})\\ & = & \lambda(g(f(x))) & (g \text{ é aplicação linear})\\ & = & \lambda(g\circ f)(x) & \text{(definição de } g\circ f).\end{array}$$

Portanto, $g\circ f$ é uma aplicação linear. □

Teorema 4.2.5. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, V'' um subespaço vetorial de $\mathbb{R}^p$, $f,g\in\mathcal{L}(V,V')$, $h,k\in\mathcal{L}(V',V'')$ e $\lambda\in\mathbb{R}$. Então são válidas as seguintes propriedades:*

1. *$h\circ(f+g)=h\circ f+h\circ g$;*
2. *$(h+k)\circ f=h\circ f+k\circ f$;*
3. *$\lambda(h\circ f)=(\lambda h)\circ f=h\circ(\lambda f)$.*

Demonstração. Exercício. □

4.3 Núcleo e espaço imagem de uma aplicação linear

O estudo dos conceitos de núcleo e espaço imagem tem interesse na sistematização do estudo de problemas que envolvem aplicações lineares.

No sentido de definirmos estes conceitos, começamos por recordar algumas noções e notações de teoria de conjuntos. Dados conjuntos A e B, C um subconjunto de A, D um subconjunto de B e f uma aplicação de A em B, designa-se por:

- ***imagem de*** C ***por*** f o conjunto
$$f(C) = \{f(x) : x \in C\} = \{y \in B : (\exists x \in C)\ y = f(x)\};$$
- ***imagem inversa de*** D ***por*** f o conjunto
$$f^{\leftarrow}(D) = \{x \in A : f(x) \in D\}.$$
Se $D = \{y\}$, é usual representar $f^{\leftarrow}(D)$ por $f^{\leftarrow}(y)$.

Teorema 4.3.1. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$ e $f \in \mathcal{L}(V, V')$. Então*

1. *$f(V)$ é um subespaço vetorial de $\mathbb{R}^m$.*
2. *$f^{\leftarrow}(V')$ é um subespaço vetorial de $\mathbb{R}^n$.*

Demonstração. Prova-se a propriedade *2.*, ficando a prova de *1.* como exercício.
Por definição de $f^{\leftarrow}(V')$, tem-se $f^{\leftarrow}(V') \subseteq \mathbb{R}^n$. O conjunto $f^{\leftarrow}(V')$ não é vazio, pois $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$ e $0_{\mathbb{R}^m} \in V'$; portanto, $0_{\mathbb{R}^n} \in f^{\leftarrow}(V')$.
Para quaisquer $x, y \in f^{\leftarrow}(V')$, tem-se $x, y \in V$ e $f(x), f(y) \in V'$. Logo, para qualquer $\alpha \in \mathbb{R}$, $x + y \in V$ e $\alpha x \in V$, pois V é subespaço vetorial de $\mathbb{R}^n$, e
$$f(x + y) = f(x) + f(y) \in V' \quad \text{e} \quad f(\alpha x) = \alpha f(x) \in V'$$
uma vez que V' é subespaço vetorial de $\mathbb{R}^m$. Logo, $x + y \in f^{\leftarrow}(V')$ e $\alpha x \in f^{\leftarrow}(V')$.
Portanto, $f^{\leftarrow}(V')$ é um subespaço vetorial de $\mathbb{R}^n$. □

Teorema 4.3.2. *Sejam V um subespaço de $\mathbb{R}^n$, $v_1, \ldots v_r \in V$, V' um subespaço vetorial de $\mathbb{R}^m$ e $f \in \mathcal{L}(V, V')$. Se $V =< v_1, \ldots, v_r >$, então $f(V) =< f(v_1), \ldots, f(v_r) >$.*

Demonstração. Admita-se que $V =< v_1, \ldots, v_r >$. Então $v_1, \ldots, v_r \in V$ e, portanto, $f(v_1), \ldots, f(v_r) \in f(V)$. Logo, como $f(V)$ é um subespaço de $\mathbb{R}^n$ e $< f(v_1), \ldots, f(v_r) >$ é o menor subespaço de $\mathbb{R}^n$ que contém $\{f(v_1), \ldots, f(v_r)\}$, tem-se $< f(v_1), \ldots, f(v_r) > \subseteq f(V)$. Por outro lado,
$$\begin{aligned}
y \in f(V) &\Rightarrow \exists x \in V : y = f(x) \\
&\Rightarrow \exists \alpha_1, \ldots, \alpha_r \in \mathbb{K} : x = \alpha_1 v_1 + \ldots + \alpha_r v_r \text{ e } y = f(x) \\
&\Rightarrow \exists \alpha_1, \ldots, \alpha_r \in \mathbb{K} : y = f(\alpha_1 v_1 + \ldots + \alpha_r v_r) \\
&\Rightarrow \exists \alpha_1, \ldots, \alpha_r \in \mathbb{K} : y = \alpha_1 f(v_1) + \ldots + \alpha_r f(v_r) \\
&\Rightarrow y \in < f(v_1), \ldots, f(v_r) >,
\end{aligned}$$
e, portanto, $f(V) \subseteq < f(v_1), \ldots, f(v_r) >$. Logo $f(V) =< f(v_1), \ldots, f(v_r) >$. □

Definição 4.3.3. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$ e $f \in \mathcal{L}(V, V')$.*

- *Chama-se* ***núcleo*** *de f, e representa-se por* $\mathrm{Nuc}f$*, ao conjunto*

$$\mathrm{Nuc}f = f^{\leftarrow}(\{0_{\mathbb{R}^m}\}) = \{x \in V : f(x) = 0_{\mathbb{R}^m}\}.$$

- *Chama-se* ***espaço imagem*** *de f, e representa-se por* $\mathrm{Im}f$ *ou $f(V)$, ao conjunto imagem de V por f.*

Exemplo 4.3.4. *Consideremos os espaços vetoriais reais $\mathbb{R}^3$ e $\mathbb{R}^2$ e a aplicação linear $f : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $f(a, b, c) = (a, a + b + c)$, para todo $(a, b, c) \in \mathbb{R}^3$. Então*

$$\begin{aligned}
\mathrm{Nuc}f &= \{(a, b, c) \in \mathbb{R}^3 : f(a, b, c) = (0, 0)\} \\
&= \{(a, b, c) \in \mathbb{R}^3 : (a, a + b + c) = (0, 0)\} \\
&= \{(a, b, c) \in \mathbb{R}^3 : a = 0, a + b + c = 0\} \\
&= \{(a, b, c) \in \mathbb{R}^3 : a = 0, b = -c\} \\
&= \{(0, -c, c) \in \mathbb{R}^3 : a, c \in \mathbb{R}\} \\
&= < (0, -1, 1) >
\end{aligned}$$

e

$$\begin{aligned}
\mathrm{Im}f &= \{f(a, b, c) \in \mathbb{R}^2 : (a, b, c) \in \mathbb{R}^3\} \\
&= \{(a, a + b + c) \in \mathbb{R}^2 : a, b, c \in \mathbb{R}\} = \mathbb{R}^2.
\end{aligned}$$

Do Teorema 4.3.1 é imediato o resultado seguinte.

Teorema 4.3.5. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$ e $f \in \mathcal{L}(V, V')$. Então*

1. $\mathrm{Nuc}f$ *é um subespaço vetorial de $\mathbb{R}^n$.*
2. $\mathrm{Im}f$ *é um subespaço vetorial de $\mathbb{R}^m$.* □

Definição 4.3.6. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$ e $f \in \mathcal{L}(V, V')$. As dimensões de* $\mathrm{Nuc}f$ *e de* $\mathrm{Im}f$ *são designadas, respectivamente, por* ***nulidade*** *de f e* ***característica*** *de f. A nulidade de f representa-se por n_f e a característica de f por c_f.*

O resultado seguinte permite relacionar a nulidade e a característica de uma aplicação linear.

Teorema 4.3.7 (Teorema da Dimensão)**.** *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$ e $f \in \mathcal{L}(V, V')$. Então* $\dim V = \dim \mathrm{Nuc}f + \dim \mathrm{Im}f$.

Demonstração. Sejam $(v_1, \ldots, v_r)$ uma base de $\mathrm{Nuc}f$ e $(w_1, \ldots, w_s)$ uma base de $\mathrm{Im}f$. Como $w_1, \ldots, w_s \in \mathrm{Im}f$, podemos escolher $u_1, \ldots, u_s \in V$ tais que $f(u_i) = w_i$, para $i = 1, \ldots, s$. Seguidamente, mostramos que $(v_1, \ldots, v_r, u_1, \ldots, u_s)$ é uma base de V, o que prova o resultado pretendido.

Seja $x \in V$. Como $f(x) \in \mathrm{Im}f$, existem $\alpha_1, \dots, \alpha_s \in \mathbb{R}$ tais que $f(x) = \sum_{i=1}^{s} \alpha_i w_i$, donde segue que

$$f(x) = \sum_{i=1}^{s} \alpha_i f(u_i) = \sum_{i=1}^{s} f(\alpha_i u_i) = f\left(\sum_{i=1}^{s} \alpha_i u_i\right).$$

Logo,

$$x - \sum_{i=1}^{s} \alpha_i u_i \in \mathrm{Nuc}f$$

e, portanto, existem $\beta_1, \dots, \beta_r \in \mathbb{R}$ tais que

$$x - \sum_{i=1}^{s} \alpha_i u_i = \sum_{i=1}^{r} \beta_i v_i.$$

Da igualdade anterior segue que

$$x = \sum_{i=1}^{r} \beta_i v_i + \sum_{i=1}^{s} \alpha_i u_i$$

Assim, todo o vetor x de V é combinação linear dos vetores $< v_1, \dots, v_r, u_1, \dots, u_s >$ e, portanto $V =< v_1, \dots, v_r, u_1, \dots, u_s >$.
Mostremos, agora, que a sequência $(v_1, \dots, v_r, u_1, \dots, u_s)$ é linearmente independente. De facto, se admitirmos que existem $\alpha_1, \dots, \alpha_s, \beta_1, \dots, \beta_r \in \mathbb{R}$ tais que

$$\sum_{i=1}^{r} \beta_i v_i + \sum_{i=1}^{s} \alpha_i u_i = 0_{\mathbb{R}^n}, \quad (1)$$

então

$$\sum_{i=1}^{r} \beta_i v_i = -\sum_{i=1}^{s} \alpha_i u_i \in \mathrm{Nuc}f$$

e, consequentemente,

$$\sum_{i=1}^{s} \alpha_i w_i = \sum_{i=1}^{s} \alpha_i f(u_i) = f\left(\sum_{i=1}^{s} \alpha_i u_i\right) = 0_{\mathbb{R}^n},$$

donde $\alpha_1 = \cdots = \alpha_s = 0$, pois $(w_1, \dots, w_s)$ é uma base de $\mathrm{Im}f$. Assim, de (1) resulta $\sum_{i=1}^{r} \beta_i v_i = 0_{\mathbb{R}^n}$ e, por conseguinte, $\beta_1 = \cdots = \beta_r = 0$, pois $(v_1, \dots, v_r)$ é uma base de $\mathrm{Nuc}f$. Portanto, a sequência de vetores $(v_1, \dots, v_r, u_1, \dots, u_s)$ é linearmente independente.
Provou-se que $(v_1, \dots, v_r, u_1, \dots, u_s)$ é uma base de V e, portanto,

$$\dim V = r + s = \dim \mathrm{Nuc}f + \dim \mathrm{Im}f. \qquad \square$$

De acordo com a notação introduzida anteriormente e nas condições deste último teorema, tem-se $\dim V = n_f + c_f$.

4.4 Aplicações lineares especiais

Algumas aplicações lineares tomam designações especiais atendendo às suas propriedades enquanto aplicações.

Definição 4.4.1. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$. Uma aplicação linear $f : V \to V'$ diz-se:*

- *um **monomorfismo** se f é injetiva;*
- *um **epimorfismo** se f é sobrejetiva;*
- *um **isomorfismo** se f é bijetiva;*
- *um **endomorfismo** se $V = V'$;*
- *um **automorfismo** se f é um endomorfismo e é bijetiva.*

As aplicações lineares sobrejetivas e as aplicações lineares injetivas podem ser caracterizadas através do seu espaço imagem e do seu núcleo.

Teorema 4.4.2. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$ e $f : V \to V'$ uma aplicação linear. Então*

1. *f é injetiva se e só se $\mathrm{Nuc}f = \{0_{\mathbb{R}^n}\}$.*
2. *f é sobrejetiva se e só se $\mathrm{Im}f = V'$.*

Demonstração. 1. Seja $f \in \mathcal{L}(V, V')$. É claro que $0_{\mathbb{R}^n} \in \mathrm{Nuc}f$, pois $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$. Admitindo que f é injetiva, então, para qualquer $x \in V$, tem-se

$$x \in \mathrm{Nuc}f \quad \Rightarrow \quad f(x) = 0_{\mathbb{R}^m} \Rightarrow f(x) = f(0_{\mathbb{R}^n}) \Rightarrow x = 0_{\mathbb{R}^n}.$$

Logo, $\mathrm{Nuc}f \subseteq \{0_{\mathbb{R}^n}\}$. Assim, $\mathrm{Nuc}f = \{0_{\mathbb{R}^n}\}$.
Reciprocamente, admitamos que $\mathrm{Nuc}f = \{0_{\mathbb{R}^n}\}$. Então, para quaisquer $x, y \in V$,

$$\begin{aligned} f(x) = f(y) \quad &\Rightarrow \quad f(x) - f(y) = 0_{\mathbb{R}^m} \Rightarrow f(x-y) = 0_{\mathbb{R}^m} \Rightarrow x - y \in \mathrm{Nuc}f \\ &\Rightarrow \quad x - y = 0_{\mathbb{R}^n} \Rightarrow x = y. \end{aligned}$$

Logo, f é injetiva.
2. Imediato pela definição de função sobrejetiva. □

Teorema 4.4.3. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, $f : V \to V'$ uma aplicação linear, $r \in \mathbb{N}$ e $v_1, \ldots, v_r \in V$.*

1. *Se a sequência $(f(v_1), \ldots, f(v_r))$ é linearmente independente, então a sequência $(v_1, \ldots, v_r)$ é linearmente independente.*
2. *Se f é injetiva e a sequência $(v_1, \ldots, v_r)$ é linearmente independente, então $(f(v_1), \ldots, f(v_r))$ é linearmente independente.*
3. *Se f é sobrejetiva e $\{v_1, \ldots, v_r\}$ é um conjunto gerador de V, então $\{f(v_1), \ldots, f(v_r)\}$ é um conjunto gerador de V'.*

Demonstração. 1. Suponhamos que a sequência $(f(v_1), \ldots, f(v_r))$ é linearmente independente. Então, para quaisquer $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$,

$$\begin{aligned}\alpha_1 v_1 + \cdots + \alpha_r v_r = 0_{\mathbb{R}^n} &\Rightarrow f(\alpha_1 v_1 + \ldots + \alpha_r v_r) = f(0_{\mathbb{R}^n})\\ &\Rightarrow \alpha_1 f(v_1) + \cdots + \alpha_r f(v_r) = 0_{\mathbb{R}^m}\\ &\Rightarrow \alpha_1 = \cdots = \alpha_r = 0.\end{aligned}$$

Logo $(v_1, \ldots, v_r)$ é linearmente independente.
2. Admitamos que f é injetiva e que a sequência $(v_1, \ldots, v_r)$ é linearmente independente. Então, recorrendo ao teorema anterior, prova-se que, para quaisquer $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$,

$$\begin{aligned}\alpha_1 f(v_1) + \cdots + \alpha_r f(v_r) = 0_{\mathbb{R}^m} &\Rightarrow f(\alpha_1 v_1 + \cdots + \alpha_r v_r) = f(0_{\mathbb{R}^n})\\ &\Rightarrow \alpha_1 v_1 + \cdots + \alpha_r v_r = 0_{\mathbb{R}^n}\\ &\Rightarrow \alpha_1 = \cdots = \alpha_r = 0.\end{aligned}$$

Logo a sequência $(f(v_1), \ldots, f(v_r))$ é linearmente independente.
3. Imediato pelo Teorema 4.3.2. □

Teorema 4.4.4. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, $f \in \mathcal{L}(V, V')$ e $(v_1, \ldots, v_r)$ uma base de V. Então*

1. *f é injetiva se e só se a sequência $(f(v_1), \ldots, f(v_r))$ é linearmente independente.*

2. *f é sobrejetiva se e só se $\{f(v_1), \ldots, f(v_r)\}$ é um conjunto gerador de V'.*

3. *f é bijetiva se e só se $(f(v_1), \ldots, f(v_r))$ é uma base de V'.*

Demonstração. Seja $(v_1, \ldots, v_r)$ uma base de V. Então a sequência $(v_1, \ldots, v_r)$ é linearmente independente e $\{v_1, \ldots, v_r\}$ é um conjunto gerador de V.

1. Como a sequência $(v_1, \ldots, v_r)$ é linearmente independente e f é injetiva, pelo teorema anterior segue que a sequência $(f(v_1), \ldots, f(v_r))$ também é linearmente independente.
Reciprocamente, suponhamos que a sequência $(f(v_1), \ldots, f(v_r))$ é linearmente independente. Para provar que f é injetiva é suficiente mostrar que $\mathrm{Nuc}f = \{0_{\mathbb{R}^n}\}$. Uma vez que $\{0_{\mathbb{R}^n}\} \subseteq \mathrm{Nuc}f$, resta provar que $\mathrm{Nuc}f \subseteq \{0_{\mathbb{R}^n}\}$. Seja $x \in \mathrm{Nuc}f$. Tem-se $f(x) = 0_{\mathbb{R}^m}$ e, como $\{v_1, \ldots, v_r\}$ é um conjunto gerador de V, existem $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$ tais que $x = \alpha_1 v_1 + \ldots + \alpha_r v_r$. Então $f(\alpha_1 v_1 + \ldots + \alpha_r v_r) = 0_{\mathbb{R}^m}$ e, uma vez que f é aplicação linear, segue que $\alpha_1 f(v_1) + \ldots + \alpha_r f(v_r) = 0_{\mathbb{R}^m}$. Por conseguinte, como $(f(v_1), \ldots, f(v_r))$ é linearmente independente, resulta que $\alpha_1 = \ldots = \alpha_r = 0$; logo $x = 0v_1 + \ldots + 0v_r = 0_{\mathbb{R}^n}$. Portanto, $\mathrm{Nuc}f \subseteq \{0_{\mathbb{R}^n}\}$.

2. Uma vez que $\{v_1, \ldots, v_r\}$ é um conjunto gerador de V, do teorema anterior segue que $\{f(v_1), \ldots, f(v_r)\}$ é um conjunto gerador de V'. A implicação recíproca é imediata a partir do Teorema 4.3.2. De facto, se $\{f(v_1), \ldots, f(v_r)\}$ é um conjunto gerador de V', tem-se $V' =< f(v_1), \ldots, f(v_r) >= \mathrm{Im}f$ e, portanto, f é sobrejetiva.

3. Imediato das alíneas anteriores. □

Corolário 4.4.5. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$ e $f \in \mathcal{L}(V, V')$. Se $\dim V = \dim V'$, então f é injetiva se e só se f é sobrejetiva.* □

Definição 4.4.6. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$. Diz-se que V é* **isomorfo** *a V', e escreve-se $V \cong V'$, se existe um isomorfismo de V em V'.*

Exemplo 4.4.7. *Seja V um subespaço vetorial de $\mathbb{R}^n$. Então $V \cong V$, uma vez que $id_V : V \to V$ é um isomorfismo.*

Exemplo 4.4.8. *Sejam V e V' os subespaços vetoriais de $\mathbb{R}^3$ e $\mathbb{R}^4$, respetivamente, definidos por*

$$V =< (1,0,0), (0,1,0) > \quad e \quad V' =< (0,1,0,0), (0,1,0,1) > .$$

A aplicação $f : V \to V'$ definida por

$$f(1,0,0) = (0,1,0,0) \text{ e } f(0,1,0) = (0,1,0,1)$$

é um isomorfismo.

Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$. O resultado seguinte permite concluir que se $V \cong V'$, então também se tem $V' \cong V$. Diz-se, então, que os espaços V e V' são isomorfos.

Teorema 4.4.9. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$. Se $f : V \to V'$ é um isomorfismo, então f^{-1} é um isomorfismo.*

Demonstração. Seja $f : V \to V'$ um isomorfismo. Então f é bijetiva e, portanto, existe uma e uma só aplicação $f^{-1} : V' \to V$ tal que $f^{-1} \circ f = id_{\mathbb{R}^n}$ e $f \circ f^{-1} = id_{\mathbb{R}^m}$. Para cada $a \in \mathbb{R}^m$, tem-se $a = id_{\mathbb{R}^m}(a) = (f \circ f^{-1})(a) = f(f^{-1}(a))$. Logo, para quaisquer $x, y \in \mathbb{R}^m$ e qualquer $\alpha \in \mathbb{R}$, tem-se

$$\begin{array}{rcl} f^{-1}(x+y) & = & f^{-1}(f(f^{-1}(x)) + f(f^{-1}(y))) \\ & = & f^{-1}(f(f^{-1}(x) + f^{-1}(y))) \\ & = & (f^{-1} \circ f)(f^{-1}(x) + f^{-1}(y)) \\ & = & id_{\mathbb{R}^n}(f^{-1}(x) + f^{-1}(y)) \\ & = & f^{-1}(x) + f^{-1}(y), \end{array}$$

$$\begin{array}{rcl} f^{-1}(\alpha x) & = & f^{-1}(\alpha f(f^{-1}(x))) \\ & = & f^{-1}(f(\alpha f^{-1}(x))) \\ & = & (f^{-1} \circ f)(\alpha f^{-1}(x)) \\ & = & id_{\mathbb{R}^n}(\alpha f^{-1}(x)) \\ & = & \alpha f^{-1}(x). \end{array}$$

Então f^{-1} é aplicação linear. Uma vez que f^{-1} também é bijetiva, conclui-se que $f^{-1} : V \to V'$ é um isomorfismo. □

Teorema 4.4.10. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$. Então $V \cong V'$ se e só se* $\dim V = \dim V'$.

Demonstração. Se $V = \{0_{\mathbb{R}^n}\}$ ou $V' = \{0_{\mathbb{R}^m}\}$, o resultado é óbvio.
Consideremos, agora o caso em que $V \neq \{0_{\mathbb{R}^n}\}$ e $V' \neq \{0_{\mathbb{R}^m}\}$.
Se V e V' são espaços vetoriais isomorfos, então existe um isomorfismo $f : V \to V'$. Seja $(v_1, \ldots, v_r)$ uma base de V. Atendendo ao Teorema 4.4.4 segue que $(f(v_1), \ldots, f(v_r))$ é uma base de V'. Logo $\dim V = \dim V'$.
Reciprocamente, suponhamos que $\dim V = r = \dim V'$, $r \in \mathbb{N}$. Sejam $(v_1, \ldots, v_r)$ e $(u_1, \ldots, u_r)$ bases de V e V', respetivamente. Pelo Teorema 4.1.7 sabemos que existe uma aplicação linear $f : V \to V'$ tal que $f(v_i) = u_i$, para todo $i \in \{1, \ldots, r\}$. Então, considerando que $(f(v_1), \ldots, f(v_r))$ é uma base de V', pelo Teorema 4.4.4 podemos concluir que f é um isomorfismo. □

4.5 Representação matricial de uma aplicação linear

Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$. Seguidamente iremos verificar que, se fixarmos uma base do espaço vetorial V e uma base do espaço vetorial V', toda aplicação linear de V em V' pode ser representada, em relação a essas bases, por uma única matriz, que a caracteriza completamente.
Se V é um espaço vetorial de dimensão $r \geq 1$, então, pelo Teorema da Extensão Linear, uma aplicação linear f de V em V' fica completamente definida pelas imagens dos vetores de uma base de V, i.e., sendo $(v_1, \ldots, v_r)$ uma base de V, uma aplicação linear $f : V \to V'$ fica completamente determinada por $f(v_1), \ldots, f(v_r)$. Por outro lado, se V' tem dimensão finita $p \geq 1$ e $(v'_1, \ldots, v'_p)$ é uma base de V', cada vetor de V' escreve-se, de modo único, como combinação linear de $v'_1, \ldots, v'_p$. Em particular, cada vetor $f(v_j)$, $j = 1, \ldots, r$, fica perfeitamente determinado se forem conhecidas as suas coordenadas em relação à base $(v'_1, \ldots, v'_p)$. De facto, se $(a_{1j}, \ldots, a_{pj}) \in \mathbb{R}^n$ é a sequência de coordenadas de $f(v_j)$ na base $(v'_1, \ldots, v'_p)$, então

$$f(v_j) = a_{1j}v'_1 + \ldots + a_{pj}v'_p.$$

Assim, fixadas as bases $(v_1, \ldots, v_r)$ de V e $({v_1}', \ldots, {v_p}')$ de V', a aplicação linear f fica completamente caracterizada pelos escalares a_{ij}, $i \in \{1, \ldots, p\}$, $j \in \{1, \ldots r\}$, e podemos associar a f a matriz

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \ldots & a_{pr} \end{bmatrix}$$

que caracteriza completamente f.

Definição 4.5.1. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$, $\mathcal{B} = (v_1, \ldots, v_r)$ uma base de V, $\mathcal{B}' = (v'_1, \ldots, v'_p)$ uma base de V' e $f \in \mathcal{L}(V, V')$. Designa-se por* ***matriz da aplicação linear f em relação às bases $\mathcal{B}$ e $\mathcal{B}'$****, e representa-se por $M(f; \mathcal{B}, \mathcal{B}')$ ou $M(f; (v_j), (v'_i))$, a matriz $A = [a_{ij}] \in \mathcal{M}_{p \times r}(\mathbb{K})$ tal que $f(v_j) = \sum_{i=1}^{p} a_{ij}v'_i$, $j \in \{1, \ldots, r\}$.*

Terminologia: Sejam V um subespaço vetorial de $\mathbb{R}^n$ e $\mathcal{B}$ uma base de V. Se f é endomorfismo de V, à matriz de f em relação a $\mathcal{B}$ e a $\mathcal{B}$ chama-se apenas matriz de f em relação a $\mathcal{B}$.

Teorema 4.5.2. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, $\mathcal{B} = (v_1, \ldots, v_r)$ uma base de V e $\mathcal{B}' = (v_1', \ldots, v_p')$ uma base de V'. Então a aplicação*

$$\begin{array}{rccc} \Phi : & \mathcal{L}(V, V') & \to & \mathcal{M}_{p\times r}(\mathbb{K}) \\ & f & \mapsto & M(f; \mathcal{B}, \mathcal{B}') \end{array}$$

é bijetiva.

Observação: Sendo V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, $(v_1, \ldots, v_r)$ uma base de V e $(v_1', \ldots, v_p')$ uma base de V', do teorema anterior resulta que toda a matriz $A \in \mathcal{M}_{p\times r}(\mathbb{K})$ é matriz em relação às bases $(v_1, \ldots, v_r)$ e $(v_1', \ldots, v_p')$ de uma e uma só aplicação linear $f : V \to V'$.

Exemplo 4.5.3. *Seja $f : \mathbb{R}^3 \to \mathbb{R}^2$ a aplicação definida por*

$$f(x, y, z) = (2x - 3y + z, 3x - 2y),$$

para todo $(x, y, z) \in \mathbb{R}^3$. Consideremos as bases

$$\mathcal{B}_1 = ((1, 0, 0), (1, 1, 0), (1, 1, 1)) \text{ de } \mathbb{R}^3 \quad \text{e} \quad \mathcal{B}_1{}' = ((1, 1), (1, -1) \text{ de } \mathbb{R}^2.$$

Tem-se

$$\begin{array}{l} f(1, 0, 0) = (2, 3) = \frac{5}{2}(1, 1) - \frac{1}{2}(1, -1), \\ f(1, 1, 0) = (-1, 1) = 0(1, 1) + (-1)(1, -1), \\ f(1, 1, 1) = (0, 1) = \frac{1}{2}(1, 1) - \frac{1}{2}(1, -1), \end{array}$$

Portanto, $M(f; \mathcal{B}_1, \mathcal{B}_1{}') = \begin{bmatrix} \frac{5}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & -1 & -\frac{1}{2} \end{bmatrix} \in \mathcal{M}_{2\times 3}(\mathbb{R})$.

Exemplo 4.5.4. *Sendo V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, $\mathcal{B} = (v_1, \ldots, v_r)$ uma base de V, $\mathcal{B}' = (v_1', \ldots, v_p')$ uma base de V', tem-se $M(0_{\mathcal{L}(V,V')}; \mathcal{B}, \mathcal{B}') = 0_{p\times r}$.*

Exemplo 4.5.5. *Sendo V um subespaço vetorial de $\mathbb{R}^n$ e $\mathcal{B} = (v_1, \ldots, v_r)$ uma base de V, tem-se $M(id_V; \mathcal{B}, \mathcal{B}) = I_r$.*

O resultado seguinte permite determinar a característica de uma aplicação linear, isto é, a dimensão do seu espaço imagem, através da característica de uma qualquer matriz da aplicação linear.

Teorema 4.5.6. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, $(v_1, \ldots, v_r)$ uma base de V, $(v_1', \ldots, v_p')$ uma base de V', $f \in \mathcal{L}(V, V')$ e*

$$A = M(f; (v_j), (v_i')).$$

Então $c_f = car(A)$.

Demonstração. Sejam $A = [a_{ij}] = M(f;(v_j),(v'_i))$, $(e_1,\ldots,e_p)$ a base canónica de $\mathbb{R}^p$ e $\psi : \mathbb{R}^p \to V'$ a aplicação linear definida por $\psi(e_i) = v'_i$, para todo $i \in \{1,\ldots,p\}$. Então ψ é um isomorfismo (Teorema 4.4.4) e, para todo $(x_1,\ldots,x_p) \in \mathbb{R}^p$, tem-se $\psi(x_1,\ldots,x_p) = \sum_{i=1}^p x_i v'_i$. Logo, para qualquer subespaço F de $\mathbb{R}^p$, tem-se $\dim F = \dim \psi(F)$ e, por conseguinte,

$$\begin{aligned}
car(A) = car_c(A) &= \dim < (a_{11},a_{21},\ldots,a_{p1}),\ldots,(a_{1r},a_{2r},\ldots,a_{pr}) > \\
&= \dim \psi(< (a_{11},a_{21},\ldots,a_{p1}),\ldots,(a_{1r},a_{2r},\ldots,a_{pr}) >) \\
&= \dim(< \psi(a_{11},a_{21},\ldots,a_{p1}),\ldots,\psi(a_{1r},a_{2r},\ldots,a_{pr}) >) \\
&= \dim < \textstyle\sum_{i=1}^p a_{i1}v'_i,\ldots,\sum_{i=1}^p a_{ir}v'_i > \\
&= \dim < f(v_1),\ldots,f(v_r) > \\
&= \dim \mathrm{Im} f = c_f.
\end{aligned}$$

□

Sendo V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, $f \in \mathcal{L}(V,V')$, $\mathcal{B}$ uma base de V e $\mathcal{B}'$ uma base de V', vejamos como utilizar a matriz $M(f;\mathcal{B},\mathcal{B}')$ para determinar a imagem por f de qualquer vetor $v \in V$.

Definição 4.5.7. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e $\mathcal{B} = (v_1,\ldots,v_r)$ uma base de V. Dado $v = \alpha_1 v_1 + \ldots + \alpha_r v_r \in V$, designa-se por* ***vetor coluna de*** *v na base $\mathcal{B}$, e representa-se por $[v]_{(v_j)}$ ou por $[v]_{\mathcal{B}}$, a matriz*

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} \in \mathcal{M}_{r\times 1}(\mathbb{R})$$

das coordenadas de v relativamente à base $\mathcal{B}$.

Teorema 4.5.8. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, $\mathcal{B} = (v_1,\ldots,v_r)$ uma base de V e $\mathcal{B}' = (v'_1,\ldots,v'_p)$ uma base de V'. Sejam $f \in \mathcal{L}(V,V')$, $A = M(f;\mathcal{B},\mathcal{B}')$ e $v \in V$.*
Se $(\alpha_1,\ldots,\alpha_r)$ é a sequência de coordenadas de v na base $\mathcal{B}$, então a sequência de coordenadas de $f(v)$ na base $\mathcal{B}'$ é $(\beta_1,\ldots,\beta_p)$ tal que

$$A\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}.$$

Demonstração. Sejam $A = [a_{ij}] = M(f;\mathcal{B},\mathcal{B}')$ e $v = \sum_{j=1}^r \alpha_j v_j$. O vetor coluna de v na base $\mathcal{B}$ é

$$[v]_{(v_j)} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} \in \mathcal{M}_{r\times 1}(\mathbb{R})$$

e tem-se

$$A[v]_{(v_j)} = \begin{bmatrix} \sum_{j=1}^r a_{1j}\alpha_j \\ \vdots \\ \sum_{j=1}^r a_{pj}\alpha_j \end{bmatrix} \in \mathcal{M}_{p\times 1}(\mathbb{R}).$$

O vetor $A[v]_{(v_j)}$ é o vetor coluna de $f(v)$ na base $(v'_1, \ldots, v'_p)$, uma vez que

$$\begin{aligned}
f(v) &= f(\textstyle\sum_{j=1}^r \alpha_j v_j) = \sum_{j=1}^r \alpha_j f(v_j) = \sum_{j=1}^r \alpha_j (\sum_{i=1}^p a_{ij} v'_i) \\
&= \textstyle\sum_{j=1}^r (\sum_{i=1}^p \alpha_j (a_{ij} v'_i)) = \sum_{j=1}^r (\sum_{i=1}^p (\alpha_j a_{ij}) v'_i) \\
&= \textstyle\sum_{i=1}^p (\sum_{j=1}^r (\alpha_j a_{ij}) v'_i) = \sum_{i=1}^p (\sum_{j=1}^r \alpha_j a_{ij}) v'_i \\
&= \textstyle\sum_{i=1}^p (\sum_{j=1}^r a_{ij} \alpha_j) v'_i.
\end{aligned}$$

□

Observação: De acordo com o teorema anterior, podemos determinar a imagem por f de um vetor v multiplicando a matriz $\mathcal{M}(f; \mathcal{B}, \mathcal{B}')$ pelo vetor coluna de v relativamente à base $\mathcal{B}$. O resultado deste produto é o vetor coluna de $f(v)$ relativamente à base $\mathcal{B}'$.

Exemplo 4.5.9. *Consideremos as bases*

$$\begin{aligned}
\mathcal{B} &= ((1,1),(1,0)) \text{ de } \mathbb{R}^2, \\
\mathcal{B}' &= ((1,1,1),(1,1,0),(1,0,0)) \text{ de } \mathbb{R}^3,
\end{aligned}$$

e seja $f : \mathbb{R}^2 \to \mathbb{R}^3$ *a aplicação linear tal que* $M(f; \mathcal{B}, \mathcal{B}') = \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix}$.
Vamos determinar $f(2,-1)$ *seguindo o processo descrito anteriormente. Para tal, comecemos por escrever* $(2,-1)$ *como combinação linear dos vetores da base* $\mathcal{B}$. *Tem-se* $(2,-1) = (-1)(1,1) + 3(1,0)$, *pelo que o vetor coluna de* $(2,-1)$ *na base* $\mathcal{B}$ *é* $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$. *Logo*

$$A \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \\ 6 \end{bmatrix}$$

é o vetor coluna de $f(2,-1)$ *na base* $\mathcal{B}'$. *Por conseguinte,*

$$f(2,-1) = -2(1,1,1) + 4(1,1,0) + 6(1,0,0) = (8,2,-2).$$

Estudámos anteriormente operações entre aplicações lineares pelo que é natural questionar quais serão as matrizes das aplicações lineares resultantes destas operações.

Teorema 4.5.10. *Sejam* V *um subespaço vetorial de* $\mathbb{R}^n$, V' *um subespaço vetorial de* $\mathbb{R}^m$, V'' *um subespaço vetorial de* $\mathbb{R}^p$, $\mathcal{B} = (v_1, \ldots, v_r)$ *uma base de* V, $\mathcal{B}' = (v'_1, \ldots, v'_s)$ *uma base de* V' *e* $\mathcal{B}'' = (v''_1, \ldots, v''_p)$ *uma base de* V''. *Então*

1. $\forall f, g \in \mathcal{L}(V, V')$, $M(f+g; \mathcal{B}, \mathcal{B}') = M(f; \mathcal{B}, \mathcal{B}') + M(g; \mathcal{B}, \mathcal{B}')$;

2. $\forall f \in \mathcal{L}(V, V')$, $\forall \alpha \in \mathbb{K}$, $M(\alpha f; \mathcal{B}, \mathcal{B}') = \alpha M(f; \mathcal{B}, \mathcal{B}')$;

3. $\forall f \in \mathcal{L}(V, V')$, $\forall g \in \mathcal{L}(V', V'')$,
$$M(g \circ f; \mathcal{B}, \mathcal{B}'') = M(g; \mathcal{B}', \mathcal{B}'') M(f; \mathcal{B}, \mathcal{B}').$$

Demonstração. 1. Sejam $A = [a_{ij}] = M(f; \mathcal{B}, \mathcal{B}')$ e $B = [b_{ij}] = M(g; \mathcal{B}, \mathcal{B}')$. Tem--se $A, B \in \mathcal{M}_{s\times r}(\mathbb{R})$, pelo que $A + B \in \mathcal{M}_{s\times r}(\mathbb{K})$. Como $f, g \in \mathcal{L}(V, V')$, também se tem $f + g \in \mathcal{L}(V, V')$ e $M(f + g; \mathcal{B}, \mathcal{B}') \in \mathcal{M}_{s\times r}(\mathbb{R})$. Logo as matrizes $A + B$ e $M(f + g; \mathcal{B}, \mathcal{B}')$ são do mesmo tipo. Além disso, para cada $j \in \{1, \dots, r\}$, tem-se

$$\begin{aligned}(f+g)(v_j) = f(v_j) + g(v_j) &= \textstyle\sum_{i=1}^s a_{ij}v_i' + \sum_{i=1}^s b_{ij}v_i' \\ &= \textstyle\sum_{i=1}^s a_{ij}v_i' + b_{ij}v_i' \\ &= \textstyle\sum_{i=1}^s (a_{ij} + b_{ij})v_i'.\end{aligned}$$

Logo, $M(f + g; \mathcal{B}, \mathcal{B}') = [a_{ij} + b_{ij}]_{s\times r} = M(f; \mathcal{B}, \mathcal{B}') + M(g; \mathcal{B}, \mathcal{B}')$.

2. Sejam $\alpha \in \mathbb{R}$ e $A = [a_{ij}] = M(f; \mathcal{B}; \mathcal{B}')$. Tem-se $A \in \mathcal{M}_{s\times r}(\mathbb{R})$, pelo que $\alpha A \in \mathcal{M}_{s\times r}(\mathbb{R})$. Como $f \in \mathcal{L}(V, V')$, também se tem $\alpha f \in \mathcal{L}(V, V')$ e $M(\alpha f; \mathcal{B}, \mathcal{B}') \in \mathcal{M}_{s\times r}(\mathbb{K})$. Logo as matrizes αA e $M(\alpha f; \mathcal{B}, \mathcal{B}')$ são do mesmo tipo. Para cada $j \in \{1, \dots, r\}$, tem-se

$$(\alpha f)(v_j) = \alpha f(v_j) = \alpha \sum_{i=1}^s a_{ij}v_i' = \sum_{i=1}^s \alpha(a_{ij}v_i') = \sum_{i=1}^s (\alpha a_{ij})v_i'.$$

Logo, $M(\alpha f; \mathcal{B}, \mathcal{B}') = \alpha[a_{ij}]_{s\times r} = \alpha M(f; \mathcal{B}, \mathcal{B}')$.

3. Sejam $f \in \mathcal{L}(V, V')$, $g \in \mathcal{L}(V', V'')$, $B = [b_{ij}] = M(f; \mathcal{B}, \mathcal{B}')$ e $A = [a_{ki}] = M(g; \mathcal{B}', \mathcal{B}'')$. Então $A \in \mathcal{M}_{p\times s}(\mathbb{R})$ e $B \in \mathcal{M}_{s\times r}(\mathbb{R})$, pelo que $AB \in \mathcal{M}_{p\times r}(\mathbb{R})$. Como $f \in \mathcal{L}(V, V')$ e $g \in \mathcal{L}(V', V'')$, então $g \circ f \in \mathcal{L}(V, V'')$ e $M(g \circ f; \mathcal{B}, \mathcal{B}'') \in \mathcal{M}_{p\times r}(\mathbb{K})$. Assim, as matrizes AB e $M(g \circ f; \mathcal{B}, \mathcal{B}'')$ são do mesmo tipo. Para cada $j \in \{1, \dots, r\}$, tem-se

$$\begin{aligned}(g\circ f)(v_j) &= g(f(v_j)) = g(\textstyle\sum_{i=1}^s b_{ij}v_i') = \sum_{i=1}^s b_{ij}g(v_i') \\ &= \textstyle\sum_{i=1}^s b_{ij}(\sum_{k=1}^p a_{ki}v_k'') = \sum_{i=1}^s \sum_{k=1}^p b_{ij}(a_{ki}v_k'') \\ &= \textstyle\sum_{i=1}^s \sum_{k=1}^p (b_{ij}a_{ki})v_k'' = \sum_{k=1}^p \sum_{i=1}^s (b_{ij}a_{ki})v_k'' \\ &= \textstyle\sum_{k=1}^p (\sum_{i=1}^s b_{ij}a_{ki})v_k'' = \sum_{k=1}^p (\sum_{i=1}^s a_{ki}b_{ij})v_k''.\end{aligned}$$

Logo, $M(g \circ f; \mathcal{B}, \mathcal{B}'') = AB = M(g; \mathcal{B}', \mathcal{B}'')M(f; \mathcal{B}, \mathcal{B}')$. □

O resultado seguinte permite determinar a aplicação inversa de um isomorfismo f recorrendo à representação matricial de f.

Teorema 4.5.11. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e V' um subespaço vetorial de $\mathbb{R}^m$ tais que $\dim V = \dim V' = r$. Sejam $\mathcal{B}$ uma base de V, $\mathcal{B}'$ uma base de V', $f \in \mathcal{L}(V, V')$ e $A = M(f; \mathcal{B}, \mathcal{B}')$. Então*

1) A é invertível se e só se f é um isomorfismo.

2) Se f é um isomorfismo, então $A^{-1} = M(f^{-1}; \mathcal{B}', \mathcal{B})$.

Demonstração. 1) Sejam $f \in \mathcal{L}(V, V')$ e $A = M(f; \mathcal{B}, \mathcal{B}')$.
Admitamos que A é invertível. Então existe $A^{-1} \in \mathcal{M}_r(\mathbb{R})$ tal que

$$AA^{-1} = I_r = A^{-1}A.$$

Como $f : V \to V'$ é uma aplicação linear, resta mostrar que f é bijetiva, ou seja, que existe uma aplicação $g : V' \to V$ tal que $f \circ g = \mathrm{id}_{V'}$ e $g \circ f = \mathrm{id}_V$.
Seja $g : V' \to V$ a aplicação linear tal que $A^{-1} = M(g; \mathcal{B}', \mathcal{B})$. Atendendo ao Teorema 4.5.10, tem-se

$$M(f \circ g; \mathcal{B}', \mathcal{B}') = M(f; \mathcal{B}, \mathcal{B}')M(g; \mathcal{B}', \mathcal{B}) = AA^{-1} = I_r = M(id_{V'}; \mathcal{B}', \mathcal{B}');$$

logo $f \circ g = id_{V'}$. Também se tem

$$M(g \circ f; \mathcal{B}, \mathcal{B}) = M(g; \mathcal{B}', \mathcal{B})M(f; \mathcal{B}, \mathcal{B}') = A^{-1}A = \mathrm{I}_r = M(id_V; \mathcal{B}, \mathcal{B}),$$

e, portanto, $g \circ f = id_V$.
Assim, $f : V \to V'$ é uma aplicação linear bijetiva; i.e., f é um isomorfismo de V em V'.
Reciprocamente, se $f : V \to V'$ é um isomorfismo, a aplicação $f^{-1} : V' \to V$ também é um isomorfismo. Seja $B = M(f^{-1}; \mathcal{B}', \mathcal{B})$. Atendendo ao Teorema 4.5.10, tem-se

$$\begin{aligned} AB &= M(f; \mathcal{B}, \mathcal{B}')M(f^{-1}; \mathcal{B}', \mathcal{B}) = M(f \circ f^{-1}; \mathcal{B}', \mathcal{B}') = M(id_{V'}; \mathcal{B}', \mathcal{B}') = I_r, \\ BA &= M(f^{-1}; \mathcal{B}', \mathcal{B})M(f; \mathcal{B}, \mathcal{B}') = M(f^{-1} \circ f; \mathcal{B}, \mathcal{B}) = M(id_{V'}; \mathcal{B}, \mathcal{B}) = I_r. \end{aligned}$$

Assim, $AB = I_r = BA$ e, portanto, A é invertível.
2) Exercício. □

Exemplo 4.5.12. *Seja V um subespaço vetorial de $\mathbb{R}^3$ com base $\mathcal{B} = (v_1, v_2)$. Seja $f : V \to V$ a aplicação linear tal que $M(f; \mathcal{B}, \mathcal{B}) = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}$. A matriz $M(f; \mathcal{B}, \mathcal{B})$ é invertivel, logo f é um isomorfismo. Pretendemos determinar $f^{-1}(av_1 + bv_2)$, para quaisquer $a, b \in \mathbb{R}$. Tem-se*

$$M(f^{-1}; \mathcal{B}, \mathcal{B}) = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{bmatrix}.$$

Como

$$\begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \frac{1}{3}a + \frac{2}{3}b \\ b \end{bmatrix},$$

então

$$f^{-1}(av_1 + bv_2) = \left(\frac{1}{3}a + \frac{2}{3}b\right) v_1 + bv_2.$$

Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$, $f \in \mathcal{L}(V, V')$, $\mathcal{B}_1$ e $\mathcal{B}_1'$ bases de V, $\mathcal{B}_2$ e $\mathcal{B}_2'$ bases de V'. A respeito da representação matricial de aplicações lineares, iremos verificar que as matrizes $M(f; \mathcal{B}_1, \mathcal{B}_2)$ e $M(f; \mathcal{B}_1', B_2')$, que em geral são distintas, estão relacionadas. Para estabelecer a relação existente entre matrizes de uma mesma aplicação linear relativas a bases distintas, introduzimos a noção de matriz de *mudança de base*.

Definição 4.5.13. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e $\mathcal{B}$ e $\mathcal{B}'$ bases de V. Dá-se o nome de* **matriz de mudança de base de** *$\mathcal{B}$ para $\mathcal{B}'$ à matriz $M(id_V; \mathcal{B}, \mathcal{B}')$.*

Teorema 4.5.14. *Sejam* V *um subespaço vetorial de* $\mathbb{R}^n$, $\mathcal{B}$ *e* $\mathcal{B}'$ *bases de* V *e* $A = M(id_V; \mathcal{B}, \mathcal{B}')$. *Então* A *é invertível e*

$$A^{-1} = M(id_V; \mathcal{B}', \mathcal{B}).$$

Demonstração. Imediato pelo teorema anterior. □

Teorema 4.5.15. *Sejam* V *um subespaço vetorial de* $\mathbb{R}^n$, *e* $\mathcal{B} = (v_1, \dots, v_r)$ *e* $\mathcal{B}' = (v_1', \dots, v_r')$ *bases de* V. *Se* $v = \sum_{j=1}^{r} \alpha_j v_j$ *e* $v = \sum_{i=1}^{r} \alpha_i' v_j'$, *tem-se*

$$M(\mathrm{id}_V; \mathcal{B}, \mathcal{B}') \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} = \begin{bmatrix} \alpha_1' \\ \vdots \\ \alpha_r' \end{bmatrix}.$$

Demonstração. Considerando que $v = id_V(v)$, o resultado é imediato pelo Teorema 4.5.8. □

Do teorema anterior conclui-se que conhecida a expressão de v como combinação linear dos vetores de uma base $\mathcal{B}$, a matriz de mudança de base de $\mathcal{B}$ para $\mathcal{B}'$ permite escrever v como combinação linear de $\mathcal{B}'$.

Teorema 4.5.16. *Sejam* V *um subespaço vetorial de* $\mathbb{R}^n$, V' *um subespaço vetorial de* $\mathbb{R}^m$, $\mathcal{B}_1, \mathcal{B}_2$ *bases de* V, $\mathcal{B}_1', \mathcal{B}_2'$ *bases de* V' *e* $f \in \mathcal{L}(V, V')$. *Então*

$$M(f; \mathcal{B}_2, \mathcal{B}_2') = M(\mathrm{id}_V; \mathcal{B}_1', \mathcal{B}_2') M(f; \mathcal{B}_1, \mathcal{B}_1') M(\mathrm{id}_V; \mathcal{B}_2, \mathcal{B}_1).$$

Demonstração. Tem-se $f = id_{V'} \circ (f \circ id_V)$. Logo, pelo Teorema 4.5.10,

$$\begin{aligned} M(f; \mathcal{B}_2, \mathcal{B}_2') &= M(id_{V'} \circ (f \circ id_V); \mathcal{B}_2, \mathcal{B}_2') \\ &= M(id_{V'}; \mathcal{B}_1', \mathcal{B}_2') M(f \circ id_V; \mathcal{B}_2, \mathcal{B}_1') \\ &= M(id_{V'}; \mathcal{B}_1', \mathcal{B}_2') M(f; \mathcal{B}_1, \mathcal{B}_1') M(id_V; \mathcal{B}_2, \mathcal{B}_1). \end{aligned}$$

□

Exemplo 4.5.17. *Consideremos os espaços vetoriais reais* $\mathbb{R}^3$ *e* $\mathbb{R}^2$. *Sejam* $\mathcal{B}_1$ *e* $\mathcal{B}_1'$ *as bases de* $\mathbb{R}^3$ *a seguir indicadas*

$$\mathcal{B}_1 = ((1,0,0), (0,1,0), (0,0,1)), \quad \mathcal{B}_1' = ((0,0,1), (0,1,1), (1,1,1))$$

e $\mathcal{B}_2$, $\mathcal{B}_2'$ *as bases de* $\mathbb{R}^2$ *definidas por*

$$\mathcal{B}_2 = ((1,1), (0,1)), \quad \mathcal{B}_2' = ((1,0), (1,2)).$$

Seja $f : \mathbb{R}^3 \to \mathbb{R}^2$ *a aplicação linear definida por*

$$f(a,b,c) = (2a + b, c - b),$$

para todo $(a,b,c) \in \mathbb{R}^3$. *Tem-se*

$$\begin{aligned} f(0,0,1) &= (0,1) &= 0(1,1) + 1(0,1), \\ f(0,1,1) &= (1,0) &= 1(1,1) + (-1)(0,1), \\ f(1,1,1) &= (3,0) &= 3(1,1) + (-3)(0,1). \end{aligned}$$

Logo $M(f; \mathcal{B}_1', \mathcal{B}_2) = \begin{bmatrix} 0 & 1 & 3 \\ 1 & -1 & -3 \end{bmatrix} = A.$

Vamos determinar $B = M(f;\mathcal{B}_1,\mathcal{B}'_2)$. *Pelo Teorema 4.5.16, tem-se*

$$B = M(id_{\mathbb{R}^2};,\mathcal{B}_2,\mathcal{B}'_2)AM(id_{\mathbb{R}^3};\mathcal{B}_1,\mathcal{B}'_1),$$

pelo que comecemos por determinar $M(id_{\mathbb{R}^3};\mathcal{B}_1,\mathcal{B}'_1)$ *e* $M(id_{\mathbb{R}^2};\mathcal{B}_2,\mathcal{B}'_2)$. *Tem-se*

$$\begin{array}{rcccl} id_{\mathbb{R}^3}(1,0,0) & = & (1,0,0) & = & 0(0,0,1)+(-1)(0,1,1)+1(1,1,1) \\ id_{\mathbb{R}^3}(0,1,0) & = & (0,1,0) & = & (-1)(0,0,1)+1(0,1,1)+0(1,1,1) \\ id_{\mathbb{R}^3}(0,0,1) & = & (0,0,1) & = & 1(0,0,1)+0(0,1,1)+0(1,1,1) \end{array}$$

e

$$\begin{array}{rcccl} id_{\mathbb{R}^2}(1,1) & = & (1,1) & = & \frac{1}{2}(1,0)+\frac{1}{2}(1,2) \\ id_{\mathbb{R}^2}(0,1) & = & (0,1) & = & -\frac{1}{2}(1,0)+\frac{1}{2}(1,2). \end{array}$$

Assim, $M(id_{\mathbb{R}^3};\mathcal{B}_1,\mathcal{B}'_1) = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ *e* $M(id_{\mathbb{R}^2};\mathcal{B}_2,\mathcal{B}'_2) = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

Logo,

$$\begin{aligned} B = M(f;\mathcal{B}_1,\mathcal{B}'_2) &= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & 3 \\ 1 & -1 & -3 \end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{2} & 1 & 3 \\ \frac{1}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & \frac{3}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}. \end{aligned}$$

Sejam V um subespaço vetorial de $\mathbb{R}^n$, V' um subespaço vetorial de $\mathbb{R}^m$ tais que $\dim V = r$ e $\dim V' = p$, $\mathcal{B}_1$, $\mathcal{B}_2$ bases de V, $\mathcal{B}'_1$, $\mathcal{B}'_2$ bases de V' e $f \in \mathcal{L}(V,V')$. Se $A = M(f;\mathcal{B}_1,\mathcal{B}'_1)$ e $B = M(f;\mathcal{B}_2,\mathcal{B}'_2)$, então, pelo teorema anterior, existem matrizes invertíveis $P \in \mathcal{M}_p(\mathbb{K})$ e $Q \in \mathcal{M}_r(\mathbb{K})$ tais que $B = PAQ$.

Considerando que toda a matriz invertível é igual a um produto de matrizes elementares, do último resultado conclui-se que quaisquer duas matrizes de uma mesma aplicação linear são equivalentes por linhas e por colunas.

5. Álgebra Vetorial

Neste capítulo, iremos estudar conceitos ligados à noção intuitiva de vetor no plano e no espaço, nomeadamente os conceitos de comprimento de um vetor, ângulo entre dois vetores e ortogonalidade de vetores. Em particular, pretendemos dar um sentido em $\mathbb{R}^n$ a estas noções. Para tal, é necessário introduzir a noção de *produto escalar* (também designado por *produto interno euclidiano*), que permitirá definir e estudar estas noções de forma rigorosa.

5.1 Produto escalar, norma e distância em $\mathbb{R}^n$

Para dar sentido, em $\mathbb{R}^2$ ou $\mathbb{R}^3$, às noções de distância e ângulo entre vetores do plano ou do espaço, a estrutura do espaço vetorial real é enriquecida com um novo conceito: o *produto escalar*. Embora o objetivo deste capítulo seja estudar a álgebra vetorial no plano e no espaço, apresentaremos essas noções de forma mais geral.

Definição 5.1.1. *Sejam $n \in \mathbb{N}$ e $u = (a_1, \ldots, a_n), v = (b_1, \ldots, b_n) \in \mathbb{R}^n$. Chama-se **produto escalar** (ou **produto interno euclidiano**) de u e v, e representa-se por $u \mid v$ ao número real*

$$u \mid v = a_1 b_1 + \cdots + a_n b_n.$$

Exemplo 5.1.2. *Dados $u = (1, -2, 3), v = (4, 0, 1) \in \mathbb{R}^3$, tem-se*

$$u \mid v = 1 \times 4 + (-2) \times 0 + 3 \times 1 = 7.$$

A prova das propriedades seguintes, relativas ao produto escalar, é deixada como exercício.

Teorema 5.1.3. *Para quaisquer $u, v, w \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, tem-se:*

1. $u \mid v = v \mid u$;
2. $u \mid (v + w) = u \mid v + u \mid w$;
3. $u \mid \alpha v = \alpha(u \mid v) = (\alpha u) \mid v$;
4. $0_{\mathbb{R}^n} \mid u = 0 = u \mid 0_{\mathbb{R}^n}$;
5. $u \mid u \geq 0$, *e* $u \mid u = 0$ *se e só se* $u = 0_{\mathbb{R}^n}$.

Recorrendo à noção de produto escalar, definimos a noção que corresponde ao conceito intuitivo de comprimento de um vetor.

Definição 5.1.4. *Seja $v \in \mathbb{R}^n$. Chama-se* ***norma de*** *v, e representa-se por $\|v\|$, o número real não negativo $\|v\| = \sqrt{v \mid v}$. Se $\|v\| = 1$, diz-se que x é um* ***vector unitário****.*

Exemplo 5.1.5. *Seja $v = (1, 3, 1) \in \mathbb{R}^3$. Então*

$$\|v\| = \sqrt{v \mid v} = \sqrt{(1,3,1) \mid (1,3,1)} = \sqrt{1 \cdot 1 + 3 \cdot 3 + 1 \cdot 1} = \sqrt{11}.$$

Da definição anterior, seguem de imediato as propriedades seguintes.

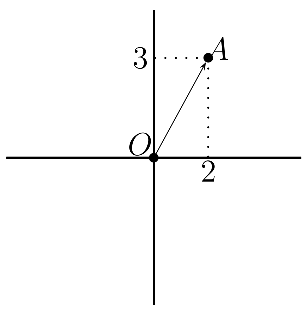
Teorema 5.1.6. *Para quaisquer $v \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, tem-se:*

1. $\|v\| = 0$ *se e só se* $v = 0_{\mathbb{R}^n}$*;*
2. $\|\alpha v\| = |\alpha| \|v\|$.

Definição 5.1.7. *Sejam $u, v \in \mathbb{R}^n$. Designa-se por* ***distância*** *entre u e v, e representa-se por $d(u, v)$, o real $\|v - u\|$.*

Exemplo 5.1.8. *Considerem-se, no plano, os pontos O e A, com coordenadas $u = (0, 0)$ e $v = (2, 3)$, respetivamente. A distância do ponto O ao ponto A é*

$$\|v - u\| = \|(2, 3) - (0, 0)\| = \sqrt{2^2 + 3^2} = \sqrt{13}.$$



Teorema 5.1.9. *Para quaisquer $u, v \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, tem-se:*

1. $d(u, v) \geq 0$*;*
2. $d(u, v) = 0$ *se e só se* $u = v$*;*
3. $d(u, v) = d(v, u)$.

Demonstração. Exercício. □

A definição de norma a partir da noção de produto interno permite estabelecer uma relação entre o produto interno de dois vectores e as suas normas.

Teorema 5.1.10 (Desigualdade de Cauchy-Schwarz)**.** *Para quaisquer $u, v \in \mathbb{R}^n$, tem-se:*

1. $|u \mid v| \leq \|u\|\|v\|$*;*

2. $|u \mid v| = \|u\|\|v\|$ *se e só se a sequência* (u, v) *é linearmente dependente.*

Demonstração. 1. Dados $u, v \in \mathbb{R}^n$, sejam $\alpha = v \mid v$ e $\beta = -(u \mid v)$. Então

$$\begin{aligned}(\alpha u + \beta v) \mid (\alpha u + \beta v) &= \alpha^2(u \mid u) + 2\alpha\beta(u \mid v) + \beta^2(v \mid v) \\ &= (v \mid v)^2(u \mid u) - 2(v \mid v)(u \mid v)^2 + (u \mid v)^2(v \mid v) \\ &= (v \mid v)^2(u \mid u) - (v \mid v)(u \mid v)^2.\end{aligned}$$

Pelo Teorema 5.1.3, tem-se $(\alpha u + \beta v) \mid (\alpha u + \beta v) \geq 0$, logo

$$(v \mid v)(u \mid v)^2 \leq (v \mid v)^2(u \mid u).$$

Se $v \neq 0_{\mathbb{R}^n}$, tem-se $v \mid v > 0$ e, portanto, $(u \mid v)^2 \leq (v \mid v)(u \mid u) = \|u\|^2\|v\|^2$. Assim, $|u \mid v| \leq \|u\|\|v\|$.
Se $v = 0_{\mathbb{R}^n}$, tem-se $u \mid v = 0$ e $\|v\| = 0$, pelo que $|u \mid v| \leq \|u\|\|v\| = 0$.

2. $\Leftarrow$) Suponha-se que a sequência (u, v) é linearmente dependente. Admita-se, sem perda de generalidade, que $v = \lambda u$, para algum $\lambda \in \mathbb{R}$. Então $|u \mid v| = |u \mid (\lambda u)| = |\lambda(u \mid u)| = |\lambda|\|u\|^2$ e $\|v\| = \|\lambda u\| = |\lambda|\|u\|$. Logo $|u \mid v| = \|u\|\|v\|$.
$\Rightarrow$) Se $|u \mid v| = \|u\|\|v\|$, então $(v \mid v)(u \mid v)^2 = (v \mid v)^2(u \mid u)$. Sendo $\alpha = (v \mid v)$ e $\beta = -(u \mid v)$, tem-se $(\alpha u + \beta v) \mid (\alpha u + \beta v) = (v \mid v)^2(u \mid u) - (v \mid v)(u \mid v)^2 = 0$. Logo $\alpha u + \beta v = 0$ e, portanto, (u, v) é linearmente dependente (pois ou $v = 0_{\mathbb{R}^n}$ ou $\alpha = v \mid v \neq 0$). ☐

Corolário 5.1.11 (Desigualdade Triangular)**.** *Para quaisquer $u, v \in \mathbb{R}^n$, tem-se*

$$\|u + v\| \leq \|u\| + \|v\|.$$

Demonstração. Uma vez que

$$\begin{aligned}\|u + v\|^2 &= (u + v) \mid (u + v) = u \mid u + 2(u \mid v) + v \mid v \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

temos $\|u + v\| \leq \|u\| + \|v\|$. ☐

Pela Desigualdade de Cauchy-Schwarz, temos

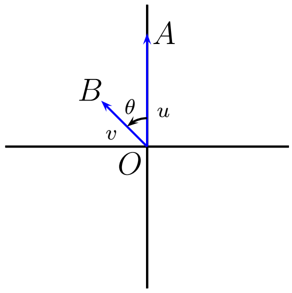
$$-1 \leq \frac{u \mid v}{\|u\|\|v\|} \leq 1$$

para quaisquer $u, v \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$. Assim, sendo a restrição da função cosseno ao intervalo $[0, \pi]$ injetiva e tendo contradomínio $[-1, 1]$, existe um número real $\theta \in [0, \pi]$ tal que $\cos\theta = \frac{u|v}{\|u\|\|v\|}$. Tal motiva a definição seguinte.

Definição 5.1.12. *Sejam $u, v \in \mathbb{R}^n$. Chama-se* ângulo dos vectores *u e v, e representa-se por $\angle(u, v)$, o número real $\theta \in [0, \pi]$ tal que*

$$\cos\theta = \frac{u \mid v}{\|u\|\|v\|}.$$

Exemplo 5.1.13. *No plano, sejam O, A e B os pontos de coordenadas $(0,0)$, $(0,2)$ e $(-1,1)$, respetivamente. Consideremos os vetores $\overrightarrow{OA}$ e $\overrightarrow{OB}$, representados em $\mathbb{R}^2$ por $u=(0,2)$ e $v=(-1,1)$.*



O ângulo entre u e v é o real $\theta \in [0,\pi]$ tal que

$$\cos\theta = \frac{(0,2) \mid (-1,1)}{\|(0,2)\|\|(-1,1)\|} = \frac{0\times(-1)+2\times 1}{\sqrt{0^2+2^2}\sqrt{(-1)^2+1^2}} = \frac{2}{2\sqrt{2}} = \frac{\sqrt{2}}{2}.$$

Logo, $\theta = \pi/4$.

Exemplo 5.1.14. *Sejam $u=(1,-1,0), v=(1,0,1) \in \mathbb{R}^3$. O ângulo entre os vectores u e v é o número real $\theta \in [0,\pi]$ tal que*

$$\cos\theta = \frac{u \mid v}{\|u\|\|v\|} = \frac{1}{\sqrt{2}\sqrt{2}} = \frac{1}{2}.$$

Logo, $\theta = \frac{\pi}{3}$.

5.2 Ortogonalidade

Relacionado com a noção de ângulo de dois vectores temos o conceito de ortogonalidade.

Definição 5.2.1. *Sejam $u,v \in \mathbb{R}^n$. Diz-se que os vetores u e v são* ***ortogonais*** *(ou* ***perpendiculares****), e escreve-se $u \perp v$, se $u \mid v = 0$.*

Observações: Para quaisquer $u,v \in \mathbb{R}^n$:

- se $u \perp v$, então $u \mid v = 0$, pelo que $v \mid u = 0$ e, portanto, também se tem $v \perp u$;
- $v \perp 0_{\mathbb{R}^n}$.

Exemplo 5.2.2. *Sejam $u=(1,3,1), v=(1,0,-1) \in \mathbb{R}^3$. Tem-se*

$$u \mid v = 1\times 1 + 3\times 0 + 1\times(-1) = 0,$$

pelo que os vectores u e v são ortogonais.

Teorema 5.2.3 (Teorema de Pitágoras)**.** *Para quaisquer $u, v \in \mathbb{R}^n$, se u e v são ortogonais, tem-se $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.*

Demonstração. Exercício. □

Definição 5.2.4. *Uma sequência $(v_1, \ldots, v_r)$ de vetores de $\mathbb{R}^n$ diz-se* ***ortogonal*** *se os vetores $v_1, \ldots, v_r$ são ortogonais dois a dois, i.e., se $v_i \mid v_j = 0$, para quaisquer $i, j \in \{1, \ldots, r\}$ tais que $i \neq j$.*
Uma base $(v_1, \ldots, v_r)$ de um subespaço vetorial V de $\mathbb{R}^n$ diz-se uma ***base ortogonal*** *se a sequência $(v_1, \ldots, v_r)$ é ortogonal.*

Observação: Para qualquer $v \in \mathbb{R}^n$, (v) é ortogonal.

Podemos relacionar os conceitos de ortogonalidade e independência linear, através do seguinte resultado.

Teorema 5.2.5. *Para quaisquer $v_1, \ldots, v_r \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, se $(v_1, \ldots, v_r)$ é ortogonal, então $(v_1, \ldots, v_r)$ é linearmente independente.*

Demonstração. Sejam $v_1, \ldots, v_r \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$ tais que $(v_1, \ldots, v_r)$ é ortogonal. Consideremos $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$ tais que $\alpha_1 v_1 + \cdots + \alpha_r v_r = 0_{\mathbb{R}^n}$. Então, para qualquer $i \in \{1, \ldots, r\}$,

$$(\alpha_1 v_1 + \cdots + \alpha_r v_r) \mid v_i = 0_{\mathbb{R}^n} \mid v_i,$$

donde

$$\alpha_1(v_1 \mid v_i) + \cdots + \alpha_i(v_i \mid v_i) + \cdots + \alpha_r(v_r \mid v_i) = 0.$$

Como $(v_1, \ldots, v_r)$ é ortogonal, temos $v_i \mid v_j = 0$, para quaisquer $i, j \in \{1, \ldots, r\}$ tais que $i \neq j$ e, portanto, $\alpha_i(v_i \mid v_j) = 0$. Além disso, dado que $v_i \neq 0_{\mathbb{R}^n}$, tem-se $v_i \mid v_i \neq 0$. Logo, como $\alpha_i(v_i \mid v_i) = 0$, resulta que $\alpha_i = 0$. Portanto, $(v_1, \ldots, v_r)$ é linearmente independente. □

O teorema seguinte descreve um processo, designado por processo de ortogonalização de Gram-Schmidt, para obter uma base ortogonal de qualquer subespaço vetorial de $\mathbb{R}^n$.

Teorema 5.2.6. *[Processo de ortogonalização de Gram-Schmidt] Sejam $(v_1, v_2, ..., v_r)$ uma base de um subespaço vetorial V de $\mathbb{R}^n$ e*

$$\begin{aligned}
u_1 &= v_1, \\
u_2 &= v_2 - \frac{v_2 \mid u_1}{u_1 \mid u_1} u_1, \\
u_3 &= v_3 - \frac{v_3 \mid u_1}{u_1 \mid u_1} u_1 - \frac{v_3 \mid u_2}{u_2 \mid u_2} u_2, \\
&\vdots \\
u_r &= v_r - \frac{v_n \mid u_1}{u_1 \mid u_1} u_1 - \ldots - \frac{v_r \mid u_{r-1}}{u_{r-1} \mid u_{r-1}} u_{n-1}.
\end{aligned}$$

Então $(u_1, u_2, ..., u_r)$ é uma base ortogonal de V.

Demonstração. Sejam $(v_1, v_2, ..., v_r)$ uma base de um subespaço V de $\mathbb{R}^n$ e u_1,u_2,. . .,u_r os vetores indicados. Por indução, prova-se que, para qualquer $p \leq r$, são satisfeitas as propriedades:

i) $(v_1, v_2, ..., v_p)$ é uma sequência ortogonal de vetores não nulos;

ii) $< v_1, v_2, ..., v_p >=< u_1, u_2, ..., u_p >$.

Em particular, $(u_1, u_2, ..., u_r)$ é uma sequência ortogonal de vetores não nulos e, portanto, é uma sequência linearmente independente. Logo, $(u_1, u_2, ..., u_r)$ é uma base de V. □

Exemplo 5.2.7. *Sejam* $v_1 = (1, 0, 1), v_2 = (0, 0, 1), v_3 = (0, 1, 1) \in \mathbb{R}^3$. *A sequência* (v_1, v_2, v_3) *é uma base de* $\mathbb{R}^3$. *Consideremos*

$$\begin{aligned}
u_1 &= v_1 = (1, 0, 1),\\
u_2 &= v_2 - \frac{v_2 \mid u_1}{u_1 \mid u_1}u_1 = (0, 0, 1) - \frac{(0, 0, 1) \mid (1, 0, 1)}{(1, 0, 1) \mid (1, 0, 1)}(1, 0, 1) = \left(-\frac{1}{2}, 0, \frac{1}{2}\right),\\
u_3 &= v_3 - \frac{v_3 \mid u_1}{u_1 \mid u_1}u_1 - \frac{v_3 \mid u_2}{u_2 \mid u_2}u_2\\
&= (0, 1, 1) - \frac{(0, 1, 1) \mid (1, 0, 1)}{(1, 0, 1) \mid (1, 0, 1)}(1, 0, 1) - \frac{(0, 1, 1) \mid (-\frac{1}{2}, 0, \frac{1}{2})}{\left(-\frac{1}{2}, 0, \frac{1}{2}\right) \mid \left(-\frac{1}{2}, 0, \frac{1}{2}\right)}\left(-\frac{1}{2}, 0, \frac{1}{2}\right)\\
&= (0, 1, 0).
\end{aligned}$$

Pelo processo de ortogonalização de Gram-Schmidt podemos afirmar que (u_1, u_2, u_3) *é uma base ortogonal de* $\mathbb{R}^3$.

Definição 5.2.8. *Uma sequência* $(v_1, ..., v_r)$ *de vetores de* $\mathbb{R}^n$ *diz-se* ***ortonormada*** *se é ortogonal e cada um dos vetores* v_i, $i \in \{1, \ldots, r\}$, *é unitário.*
Uma base $(v_1, \ldots, v_r)$ *de um subespaço vetorial* V *de* $\mathbb{R}^n$ *diz-se uma* ***base ortonormada*** *se a sequência* $(v_1, \ldots, v_r)$ *é ortonormada.*

Observação: Para qualquer $v \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, tem-se $\|v\| \neq 0$ e, sendo $u = \frac{1}{\|v\|}v$, segue que $\|u\| = 1$.

Da observação anterior e do Teorema 5.2.6 é imediato o resultado seguinte.

Teorema 5.2.9. *Todo o subespaço vetorial de* $\mathbb{R}^n$ *de dimensão* ≥ 1 *admite uma base ortonormada.*

Exemplo 5.2.10. *Sejam* u_1, u_2, u_3 *os vectores obtidos no exemplo anterior. Então* (w_1, w_2, w_3), *onde*

$$w_1 = \frac{1}{\|u_1\|}\cdot u_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), w_2 = \frac{1}{\|u_2\|}\cdot u_2 = (-1, 0, 1), w_3 = \frac{1}{\|u_3\|}\cdot u_3 = (0, 1, 0),$$

é uma base ortonormada de $\mathbb{R}^3$.

Teorema 5.2.11. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ de dimensão $r \geq 1$ e $(v_1, \ldots, v_r)$ uma base ortogonal de V. Então, para todo $v \in V$, tem-se*

$$v = \frac{v \mid v_1}{\|v_1\|^2} v_1 + \cdots + \frac{v \mid v_r}{\|v_r\|^2} v_r.$$

Em particular, se $(v_1, \ldots, v_r)$ é uma base ortonormada de V, temos

$$v = (v \mid v_1)v_1 + \cdots + (v \mid v_r)v_r.$$

Demonstração. Sejam $(v_1, \ldots, v_r)$ uma base ortogonal de V, $v \in V$ e $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$ tais que $v = \alpha_1 v_1 + \cdots + \alpha_r v_r$. Então, para todo $i \in \{1, \ldots, r\}$, temos

$$\begin{aligned} v \mid v_i &= (\alpha_1 v_1 + \cdots + \alpha_i v_i + \cdots + \alpha_r v_r) \mid v_i \\ &= \alpha_1(v_1 \mid v_i) + \cdots + \alpha_i(v_i \mid v_i) + \cdots + \alpha_r(v_r \mid v_i) \\ &= \alpha_1(0) + \cdots \alpha_i(v_i \mid v_i) + \cdots + \alpha_r(0). \end{aligned}$$

Assim, $\alpha_i = \dfrac{v \mid v_i}{v_i \mid v_i} = \dfrac{v \mid v_i}{\|v_i\|^2}$. No caso em que $(v_1, \ldots, v_r)$ é uma base ortonormada, tem-se $\|v_i\| = 1$ e, portanto, $\alpha_i = v \mid v_i$. □

Dado um subconjunto V de $\mathbb{R}^n$, representa-se por $V^\perp$ o conjunto

$$V^\perp = \{u \in \mathbb{R}^n : u \mid v = 0, \forall v \in V\}.$$

No caso particular em que V é um subespaço vetorial de $\mathbb{R}^n$, prova-se o seguinte.

Teorema 5.2.12. *Seja V um subespaço vetorial de $\mathbb{R}^n$. Então $V^\perp$ é um subespaço vetorial de $\mathbb{R}^n$.*

Demonstração. Seja V um subespaço vetorial de $\mathbb{R}^n$. Pela definição de $V^\perp$, tem-se $V^\perp \subseteq \mathbb{R}^n$. Uma vez que, para todo $v \in V$, $0_{\mathbb{R}^n} \mid v = 0$, então $0_{\mathbb{R}^n} \in V^\perp$. Dados $u, w \in V^\perp$ e $\alpha \in \mathbb{R}$, tem-se, para todo $v \in V$, $u \mid v = 0$, $w \mid v = 0$, pelo que

$$(u + w) \mid v = u \mid v + w \mid v = 0 + 0 = 0,$$

$$(\alpha u) \mid v = \alpha(u \mid v) = 0,$$

e, portanto, $u + w \in V^\perp$, $\alpha u \in V^\perp$. Logo, $V^\perp$ é um subespaço vetorial de $\mathbb{R}^n$. □

Definição 5.2.13. *Seja V um subespaço vetorial de $\mathbb{R}^n$. O subespaço*

$$V^\perp = \{u \in \mathbb{R}^n : u \mid v = 0, \forall v \in V\}$$

*designa-se por **complemento ortogonal de** V.*

Teorema 5.2.14. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão ≥ 1 e $(v_1, ..., v_r)$ uma base de V. Então*

$$V^\perp = \{u \in \mathbb{R}^n : u \mid v_i = 0, \forall i \in \{1, ..., r\}\}.$$

Demonstração. Exercício. □

Exemplo 5.2.15. *Consideremos o subespaço vetorial de* $\mathbb{R}^3$ *definido por* $V =< (-1,1,0),(0,1,1) >$. *A sequência* $((-1,1,0),(0,1,1))$ *é linearmente independente e, portanto, é uma base de* V. *O complemento ortogonal de* V *é*

$$\begin{aligned} V^\perp &= \{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y,z) \,|\, (-1,1,0) = 0, (x,y,z) \,|\, (0,1,1) = 0\} \\ &= \{(x,y,z) \in \mathbb{R}^3 \,|\, -x+y=0, y+z=0\} \\ &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x=y, z=-y\} \\ &= \{(y,y,-y) \in \mathbb{R}^3 \;\, |\, y \in \mathbb{R}\} \\ &= < (1,1,-1) > . \end{aligned}$$

Teorema 5.2.16. *Sejam* U *e* V *subespaços vetoriais de* $\mathbb{R}^n$. *Então*

1. *Se* $U \subseteq V$, *então* $V^\perp \subseteq U^\perp$.
2. $V = (V^\perp)^\perp$.
3. $U^\perp + V^\perp \subseteq (U \cap V)^\perp$.

Demonstração. Apresentamos a prova de *2.*, ficando a prova das restantes propriedades ao cuidado do leitor.

$[V_1 \subseteq (V_1^\perp)^\perp]$ Sejam $v \in V_1$ e $w \in V_1^\perp$. Por definição de $V_1^\perp$, temos $v \mid w = 0$. Portanto, v é ortogonal a todo vetor de $V_1^\perp$, o que significa que $v \in (V_1^\perp)^\perp$. Assim, $V_1 \subseteq (V_1^\perp)^\perp$.

$[(V_1^\perp)^\perp \subseteq V_1]$ Se $V_1 = \{0_{\mathbb{R}^n}\}$, o resultado é imediato.
Caso $\dim V_1 = k \geq 1$, seja $(v_1, \ldots, v_k)$ uma base ortonormada de V_1 e consideremos a sua extensão a uma base de $\mathbb{R}^n$. Seja $(v_1, \ldots, v_k, u_{k+1}, \ldots, u_n)$ a base de $\mathbb{R}^n$ obtida. A partir desta base pode-se obter uma base ortonormada de V_1 que inclui os vetores $v_1, \ldots, v_k$; seja $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ essa base. Pela construção desta última base, os vetores $v_{k+1}, \ldots, v_n$ são ortogonais a todos os vetores de $\{v_1, \ldots, v_k\}$ e, portanto, são ortogonais a todos os elementos de V_1, logo $v_{k+1}, \ldots, v_n \in V_1^\perp$.
Seja $x \in (V_1^\perp)^\perp$. Como $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ é uma base ortonormada de $\mathbb{R}^n$, tem-se

$$x = (x \mid v_1)v_1 + \ldots + (x \mid v_k)v_k + (x \mid v_{k+1})v_{k+1} + \ldots + (x \mid v_n)v_n.$$

Considerando que $x \in (V_1^\perp)^\perp$ segue que $x \mid v_{k+1} = 0$, ... $x \mid v_n = 0$. Logo,

$$x = (x \mid v_1)v_1 + \ldots + (x \mid v_k)v_k$$

e, portanto, $x \in V_1$. Assim, $(V_1^\perp)^\perp \subseteq V_1$.

Logo, $V_1 = (V_1^\perp)^\perp$. □

Teorema 5.2.17. *Seja* V *um subespaço vetorial de* $\mathbb{R}^n$. *Então* $\mathbb{R}^n$ *é soma direta de* V *e* V^T.

Demonstração. Sejam V um subespaço vetorial de $\mathbb{R}^n$ e $V^\perp$ o complemento ortogonal de V. Considerando que V e V^T são subespaços de $\mathbb{R}^n$, é imediato que $\{0_{\mathbb{R}^n}\} \subseteq V \cap V^\perp$. A inclusão contrária também é válida. De facto, se $u \in V \cap V^T$, então $u \mid v = 0$, para todo $v \in V$. Em particular, $u \mid u = 0$, donde segue $u = 0_{\mathbb{R}^n}$ e, portanto, $V \cap V^\perp \subseteq \{0_{\mathbb{R}^n}\}$. Assim, $V \cap V^\perp = \{0_{\mathbb{R}^n}\}$. Logo, a soma $V + V^\perp$ é direta. Mostremos que $\mathbb{R}^n = V + V^\perp$. Se $V = \{0_{\mathbb{R}^n}\}$, então $V^\perp = \mathbb{R}^n$. Consideremos, agora, o caso em que $\dim V = k \geq 1$. Como $V + V^\perp$ é um subespaço de $\mathbb{R}^n$, temos $n \geq \dim(V + V^\perp)$. Sendo $V + V^T$ uma soma direta, segue que $n \geq \dim(V + V^\perp) = \dim V + \dim V^\perp = k + \dim V^\perp$ e, portanto, $\dim V^\perp \leq n - k$. Sejam $(v_1, \ldots, v_k)$ uma base ortogonal de V e $(v_1 \ldots, v_k, u_1, \ldots, u_{n-k})$ uma base de $\mathbb{R}^n$. Utilizando o processo de ortogonalização de Gram-Schmidt é possível obter uma base ortogonal $(v_1, \ldots, v_k, w_1, \ldots, w_{n-k})$ de $\mathbb{R}^n$. Então $w_i \in V^\perp$, para todo $i \in \{1, \ldots, n-k\}$ (pois $w_i \perp v_j$, para todo $j \in \{1, \ldots, k\}$). Como $(w_1, \ldots, w_{n-k})$ é linearmente independente, temos $\dim V^\perp \geq n - k$. Assim, $\dim V^\perp = n - k$. Como $V + V^\perp \leq \mathbb{R}^n$, e $\dim(V + V^\perp) = \dim V + \dim V^\perp = n + (n - k) = n = \dim \mathbb{R}^n$, conclui-se que $\mathbb{R}^n = V + V^\perp$. □

Dos teoremas 3.2.18 e 5.2.17 resulta que todo o vetor u de $\mathbb{R}^n$ escreve-se de modo único na forma $u_1 + u_2$, com $u_1 \in V$ e $u_2 \in V^\perp$.

Definição 5.2.18. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, $u \in \mathbb{R}^n$ e u_1, u_2 os únicos vetores de $\mathbb{R}^n$ tais que $u_1 \in V$, $u_2 \in V^\perp$ e $u = u_1 + u_2$. Ao vetor u_1 dá-se a designação de **projeção ortogonal de** u **sobre** V e representa-se por* $\mathrm{proj}_V u$.
*No caso particular em que $V = < v >$, com $v \neq 0_{\mathbb{R}^n}$, o vetor vetor $\mathrm{proj}_V u$ é designado por **projeção ortogonal de** u **sobre** v e representa-se por* $\mathrm{proj}_v u$.

Observação: Se $V = \{0_{\mathbb{R}^n}\}$ e $u \in \mathbb{R}^n$, tem-se $\mathrm{proj}_V u = 0_{\mathbb{R}^n}$,

Teorema 5.2.19. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ de $\dim \geq 1$, $u \in \mathbb{R}^n$ e $(v_1, \ldots, v_r)$ uma base ortogonal de V. Então*

$$\mathrm{proj}_V u = \frac{u \mid v_1}{\|v_1\|^2} v_1 + \cdots + \frac{u \mid v_r}{\|v_r\|^2} v_r.$$

No caso em que $(v_1, \ldots, v_r)$ é uma base ortonormada, tem-se

$$\mathrm{proj}_V u = (u \mid v_1) v_1 + \cdots + (u \mid v_r) v_r.$$

Demonstração. Sejam $u \in \mathbb{R}^n$ e $(v_1, \ldots, v_r)$ uma base ortogonal de V. Consideremos a expansão da base $(v_1, \ldots, v_r)$ a uma base ortogonal $(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ de $\mathbb{R}^n$. Pelo Teorema 5.2.11, tem-se

$$u = \frac{u \mid v_1}{\|v_1\|^2} v_1 + \cdots + \frac{u \mid v_r}{\|v_r\|^2} v_r + \frac{u \mid v_{r+1}}{\|v_{r+1}\|^2} v_{r+1} + \cdots + \frac{u \mid v_n}{\|v_n\|^2} v_n.$$

Sejam

$$u_1 = \frac{u \mid v_1}{\|v_1\|^2} v_1 + \cdots + \frac{u \mid v_r}{\|v_r\|^2} v_r \quad \text{e} \quad u_2 = \frac{u \mid v_{r+1}}{\|v_{r+1}\|^2} v_{r+1} + \cdots + \frac{u \mid v_n}{\|v_n\|^2} v_n.$$

Claramente, tem-se $u_1 \in V$. Do Teorema 5.2.14 segue que $u_2 \in V^\perp$, pois, para todos $i \in \{1, \ldots, r\}$, $j \in \{r+1, \ldots, n\}$, $v_j \perp v_i$ e, portanto, $u_2 \perp v_i$, para todo $i \in \{1, \ldots, r\}$. Pelos teoremas 3.2.18 e 5.2.17, sabe-se que u_1 e u_2 são os únicos vetores de $\mathbb{R}^n$ tais que $u = u_1 + u_2$. Assim,

$$\text{proj}_V u = u_1 = \frac{u \mid v_1}{\|v_1\|^2} v_1 + \cdots + \frac{u \mid v_r}{\|v_r\|^2} v_r.$$

Se a base $(v_1, \ldots, v_r)$ é ortonormada, então, para todo $i \in \{1, \ldots, r\}$, $\|v_i\|^2 = 1$, pelo que

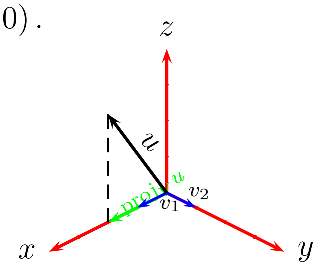
$$\text{proj}_V u = (u \mid v_1) v_1 + \cdots + (u \mid v_r) v_r.$$ □

Corolário 5.2.20. *Sejam u e v vetores de $\mathbb{R}^n$ tais que $v \neq 0_{\mathbb{R}^n}$. Então*

$$\text{proj}_v u = \frac{u \mid v}{\|v\|^2} v.$$

Exemplo 5.2.21. *Sejam $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$ e V o subespaço de $\mathbb{R}^3$ gerado por $\{v_1, v_2\}$. Vamos determinar a projeção ortogonal de $u = (2, 0, 3)$ sobre V. Os vetores v_1 e v_2 são ortogonais entre si e não são nulos. Logo, (v_1, v_2) é linearmente independente e, portanto, é uma base ortogonal de V. Pelo teorema anterior, tem-se*

$$\text{proj}_V u = \frac{u \mid v_1}{\|v_1\|^2} v_1 + \frac{u \mid v_2}{\|v_2\|^2} v_2 = 2v_1 + 0v_2 = (2, 0, 0).$$



Teorema 5.2.22. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e $u \in \mathbb{R}^n$. Então $u - \text{proj}_V u \in V^\perp$.*

Demonstração. Da prova do teorema anterior segue que $u = \text{proj}_V u + u_2$, com $u_2 \in V^\perp$. Logo, $u - \text{proj}_V u = u_2 \in V^\perp$. □

Teorema 5.2.23. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ e $u \in \mathbb{R}^n$. Então o vetor $\text{proj}_V u$ satisfaz a propriedade de minimizar a distância a u de entre os elementos de V, isto é,*

$$\|u - \text{proj}_V u\| = \min\{\|u - v\| : v \in V\}.$$

Demonstração. Sejam $p = \text{proj}_V u$ e $v \in V$ arbitrário. Como $u - p$ é ortogonal a todos os vetores de V, então u é ortogonal a $v - p$. Pelo Teorema de Pitágoras segue que

$$\|u - p\|^2 + \|v - p\|^2 = \|(u - p) - (v - p)\|^2 = \|u - v\|^2.$$

Logo, para qualquer $v \in V$, tem-se

$$\|u - p\| \leq \|u - v\|,$$

e a igualdade verifica-se se e só se $v = p$. □

O teorema seguinte, cuja prova é deixada como exercício, estabelece que a projeção ortogonal sobre um subespaço V de $\mathbb{R}^n$ é um endomorfimso de $\mathbb{R}^n$.

Teorema 5.2.24. *Sejam V um subespaço vetorial de $\mathbb{R}^n$. Então a correspondência p de $\mathbb{R}^n$ em $\mathbb{R}^n$ definida por $p(u) = \mathrm{proj}_V u$, para todo $u \in \mathbb{R}^n$, é uma aplicação linear e* $\mathrm{Nuc}\, p = V^T$.

6. Determinantes

O determinante de uma matriz quadrada sobre $\mathbb{K}$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, é um elemento de $\mathbb{K}$ calculado a partir dos elementos da matriz e que, entre outras aplicações, pode ser usado na resolução de certos sistemas de equações lineares e para decidir sobre a invertibilidade de uma matriz.

6.1 Definição e algumas propriedades

O determinante de uma matriz pode ser definido de diversas formas. No texto que se segue optamos por apresentar uma definição indutiva deste conceito.

Para uma matriz de ordem 1×1

$$A = [a]$$

é fácil concluir que a matriz é invertível se e só se $a \neq 0$.

Dada uma matriz quadrada de ordem 2×2

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

é também simples concluir em que condições a matriz é invertível; aplicando o método de eliminação de Gauss à matriz A, conclui-se que A é invertível se e só se $ad - bc \neq 0$.

Como iremos ver mais à frente, a qualquer matriz $A \in \mathcal{M}_n(\mathbb{K})$, $n \in \mathbb{N}$, podemos associar um elemento de $\mathbb{K}$ com a propriedade de A ser invertível se e só se esse escalar for não nulo. A este elemento de $\mathbb{K}$ chamaremos o *determinante* de A.

Para matrizes de ordem superior apresentamos uma definição indutiva para o determinante de uma matriz, i.e., define-se o determinante de uma matriz 2×2 em função do determinante de matrizes de ordem 1×1, define-se o determinante de uma matriz 3×3 em função do determinante de matrizes de ordem 2×2, e assim sucessivamente.

No sentido de apresentarmos a definição referida, começamos por introduzir alguma notação.

Dada uma matriz $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{K})$, $n \in \mathbb{N}$, representa-se por $A(i|j)$ a matriz quadrada de ordem $n-1$, obtida de A retirando a linha i e a coluna j.

Exemplo 6.1.1. *Se* $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 3 \\ 1 & -5 & 1 \end{bmatrix}$ *então* $A(2|3) = \begin{bmatrix} 1 & 1 \\ 1 & -5 \end{bmatrix}$.

Definição 6.1.2. *Sejam* $n \in \mathbb{N}$ *e* $A \in \mathcal{M}_n(\mathbb{K})$. *Chama-se* ***determinante de A****, e representa-se por* $\det A$ *ou* $|A|$*, ao elemento de* $\mathbb{K}$ *definido da seguinte forma:*

i) *Se* $n = 1$, *então* $\det A = a_{11}$.

ii) *Se* $n > 1$, *então* $\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A(1|j)$.

Exemplo 6.1.3. *Se* $A = \begin{bmatrix} 2 & 1 \\ 1 & -5 \end{bmatrix}$, *então*

$$\begin{aligned} \det A &= (-1)^{1+1} \times 2 \times \det[-5] + (-1)^{1+2} \times 1 \times \det[1] \\ &= 1 \times 2 \times (-5) + (-1) \times 1 \times 1 \\ &= -10 - 1 \\ &= -11. \end{aligned}$$

Exemplo 6.1.4. *Seja*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R}).$$

Então

$$\begin{aligned} \det A &= (-1)^{1+1} \times a \times \det[d] + (-1)^{1+2} \times b \times \det[c] \\ &= a \times d - b \times c. \end{aligned}$$

Exemplo 6.1.5. *Se* $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 3 \\ 1 & -5 & 1 \end{bmatrix}$, *então*

$$\begin{aligned} \det A &= (-1)^{1+1} \times 1 \times \det \begin{bmatrix} 1 & 3 \\ -5 & 1 \end{bmatrix} \\ &\quad +(-1)^{1+2} \times 1 \times \det \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} \\ &\quad +(-1)^{1+3} \times (-1) \times \det \begin{bmatrix} 2 & 1 \\ 1 & -5 \end{bmatrix} \\ &= 1 \times 1 \times (1 \times 1 - (-5) \times 3) \\ &\quad -1 \times 1 \times (2 \times 1 - 1 \times 3) \\ &\quad +1 \times (-1) \times (2 \times (-5) - 1 \times 1) \\ &= 28. \end{aligned}$$

Exemplo 6.1.6. *Se $A = [a_{ij}]$ é uma matriz real de ordem 3, então*

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21}.$$

Dada uma matriz $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{K})$ e dados $i, j \in \{1, \ldots, n\}$, designa-se por **complemento algébrico** do elemento a_{ij}, e representa-se por $\widehat{a}_{ij}$, o elemento de $\mathbb{K}$ definido por $(-1)^{i+j} \det A(i|j)$.

De acordo com a definição que apresentámos para o determinante de uma matriz $A \in \mathcal{M}_n(\mathbb{K})$, o determinante de A é igual à soma dos elementos da linha 1 multiplicados pelos respetivos complementos algébricos, ou seja,

$$\det A = \sum_{j=1}^{n} a_{1j}\widehat{a}_{1j}.$$

O resultado seguinte, que não será aqui demonstrado, estabelece que se procedermos de forma análoga para uma qualquer linha ou uma qualquer coluna de A obtemos também o determinante de A.

Teorema 6.1.7 (Teorema de Laplace)**.** *Sejam $n \in \mathbb{N}$ tal que $n \geq 2$ e $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{K})$. Então, para qualquer $k \in \{1, 2, ..., n\}$,*

1. $\det A = \sum\limits_{j=1}^{n} (-1)^{k+j} a_{kj} \det A(k|j) = \sum\limits_{j=1}^{n} a_{kj}\widehat{a}_{kj}.$

2. $\det A = \sum\limits_{i=1}^{n} (-1)^{i+k} a_{ik} \det A(i|k) = \sum\limits_{i=1}^{n} a_{ik}\widehat{a}_{ik}.$ □

Observação: A expressão indicada em *1.* designa-se por *desenvolvimento do determinante de A ao longo da linha k de A*; a expressão indicada em *2.* designa-se por *desenvolvimento do determinante de A ao longo da coluna k de A*.

Exemplo 6.1.8. *Seja* $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 0 \\ 1 & -5 & 1 \end{bmatrix}$.

Por definição, temos

$$\begin{aligned} \det A &= 1 \begin{vmatrix} 0 & 0 \\ -5 & 1 \end{vmatrix} - 1 \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} + 1 \begin{vmatrix} 2 & 0 \\ 1 & -5 \end{vmatrix} \\ &= 1\,(0-0) - 1\,(2-0) + 1\,(-10-0) \\ &= -12. \end{aligned}$$

Aplicando o teorema de Laplace, desenvolvendo o determinante ao longo da linha $k = 2$, vem

$$\begin{aligned} \det A &= (-1)^{2+1} \times 2 \times \begin{vmatrix} 1 & -1 \\ -5 & 1 \end{vmatrix} \\ &= -2\,(1+5) \\ &= -12. \end{aligned}$$

Teorema 6.1.9. *Sejam* $n \in \mathbb{N}$ *e* $A \in \mathcal{M}_n(\mathbb{K})$*. Então,* $\det A^T = \det A$.

Demonstração. A prova segue por indução sobre a ordem n da matriz.
Caso base ($n = 1$): Se $A \in \mathcal{M}_1(\mathbb{K})$, temos $A = [a_{11}] = A^T$ e, portanto, $\det A = \det A^T$.
Passo de indução: Dado $p \in \mathbb{N}$, admitamos, por hipótese de indução, que o determinante de qualquer matriz de $\mathcal{M}_p(\mathbb{K})$ é igual ao determinante da sua transposta. Com base nesta hipótese prova-se facilmente que se $A \in \mathcal{M}_{p+1}(\mathbb{K})$, então $\det A^T = \det A$. Com efeito, pelo Teorema de Laplace, e desenvolvendo o determinante da matriz A ao longo de uma linha k, $k \in \{1, 2, \ldots, p+1\}$, temos

$$\det A = \sum_{j=1}^{p+1} (-1)^{k+j} \, a_{kj} \det A(k|j).$$

Considerando que $A(k|j)$ é uma matriz de $\mathcal{M}_p(\mathbb{K})$ tem-se, por hipótese de indução, $\det A(k|j) = \det(A(k|j))^T$, donde resulta que

$$\det A = \sum_{j=1}^{p+1} (-1)^{k+j} \, a_{kj} \det A^T(j|k).$$

Agora, tendo em conta que a_{kj} é o elemento na linha j e coluna k de A^T, conclui-se que $\sum\limits_{j=1}^{p+1} (-1)^{k+j} \, a_{kj} \det A^T(j|k)$ é o determinante de A^T desenvolvido ao longo da coluna k. Por conseguinte, $\det A = \det A^T$.
Do que foi provado conclui-se, pelo Princípio de Indução em $\mathbb{N}$, que, para todo $n \in \mathbb{N}$, se $A \in \mathcal{M}_n(\mathbb{K})$, então $\det A = \det A^T$. ☐

Observação. Do teorema anterior resulta que, dada uma propriedade sobre determinantes expressa em termos de linhas (resp. colunas), podemos sempre enunciar uma propriedade análoga expressa em termos de colunas (resp. linhas).

Teorema 6.1.10. *Sejam* $n \in \mathbb{N}$ *e* $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{K})$ *uma matriz triangular superior (resp. inferior). Então* $\det A = a_{11} \times a_{22} \times \cdots \times a_{nn}$.

Demonstração. A prova é realizada por indução sobre n.

Caso base ($n = 1$): Para $n = 1$, temos $A = [a_{11}]$ e o resultado é imediato.

Passo de indução: Dado $p \in \mathbb{N}$, admitamos, por hipótese de indução, que para qualquer matriz triangular superior B de $\mathcal{M}_p(\mathbb{K})$, $\det B = b_{11} \times b_{22} \times \ldots \times b_{pp}$. Então, para qualquer matriz triangular superior $A \in \mathcal{M}_{p+1}(\mathbb{K})$, prova-se que $\det A = a_{11} \times a_{22} \times \cdots \times a_{p+1p+1}$. De facto, aplicando o Teorema de Laplace ao longo da linha $p+1$ da matriz A e atendendo a que A é uma matriz traingular superior, temos

$$\det A = \sum_{j=1}^{p+1} (-1)^{(p+1)+j} \, a_{p+1j} \det A(p+1|j) = (-1)^{2(p+1)} a_{p+1p+1} \det A(p+1|p+1).$$

Agora, como $A(p{+}1|p{+}1)$ é uma matriz de $\mathcal{M}_p(\mathbb{K})$ e é também uma matriz triangular superior, por hipótese de indução temos $\det A(p+1|p+1) = a_{11} \times a_{22} \times \cdots \times a_{pp}$. Logo, $\det A = a_{11} \times a_{22} \times \cdots \times a_{p+1p+1}$.
Assim, pelo Princípio de Indução em $\mathbb{N}$, concluímos que, para todo $n \in \mathbb{N}$, se $A \in \mathcal{M}_n(\mathbb{K})$ é uma matriz triangular superior, então $\det A = a_{11} \times a_{22} \times \cdots \times a_{nn}$.

Se A é uma matriz triangular inferior, o resultado segue imediatamente tendo em conta a proposição anterior. $\square$

Teorema 6.1.11. *Para $n \in \mathbb{N}$ e $k \in \{1, 2, \ldots, n\}$, tem-se*

$$\det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k-1\,1} & a_{k-2\,2} & \cdots & a_{k-1\,n} \\ a_{k1}+b_{k1} & a_{k2}+b_{k2} & \cdots & a_{kn}+b_{kn} \\ a_{k+1\,1} & a_{k+1\,2} & \cdots & a_{k+1\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} =$$

$$\det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k-1\,1} & a_{k-2\,2} & \cdots & a_{k-1\,n} \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ a_{k+1\,1} & a_{k+1\,2} & \cdots & a_{k+1\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} + \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k-1\,1} & a_{k-2\,2} & \cdots & a_{k-1\,n} \\ b_{k1} & b_{k2} & \cdots & b_{kn} \\ a_{k+1\,1} & a_{k+1\,2} & \cdots & a_{k+1\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Demonstração. Sejam

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k-1\,1} & a_{k-2\,2} & \cdots & a_{k-1\,n} \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ a_{k+1\,1} & a_{k+1\,2} & \cdots & a_{k+1\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \text{ e } B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k-1\,1} & a_{k-2\,2} & \cdots & a_{k-1\,n} \\ b_{k1} & b_{k2} & \cdots & b_{kn} \\ a_{k+1\,1} & a_{k+1\,2} & \cdots & a_{k+1\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Então, se $C = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k-1\,1} & a_{k-2\,2} & \cdots & a_{k-1\,n} \\ a_{k1}+b_{k1} & a_{k2}+b_{k2} & \cdots & a_{kn}+b_{kn} \\ a_{k+1\,1} & a_{k+1\,2} & \cdots & a_{k+1\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, temos

$$\begin{aligned} \det C &= \textstyle\sum_{j=1}^n (-1)^{k+j} (a_{kj} + b_{kj}) \det C(k|j) \\ &= \textstyle\sum_{j=1}^n (-1)^{k+j} a_{kj} \det C(k|j) + \sum_{j=1}^n (-1)^{k+j} b_{kj} \det C(k|j) \\ &= \textstyle\sum_{j=1}^n (-1)^{k+j} a_{kj} \det A(k|j) + \sum_{j=1}^n (-1)^{k+j} b_{kj} \det B(k|j) \\ &= \det A + \det B. \end{aligned}$$

$\square$

Corolário 6.1.12. *Para* $n \in \mathbb{N}$ *e* $k \in \{1, 2, \ldots, n\}$, *tem-se*

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1\,k-1} & a_{1\,k} + b_{1\,k} & a_{1\,k+1} & \cdots & a_{1n} \\ a_{2\,1} & \cdots & a_{2\,k-1} & a_{2\,k} + b_{2\,k} & a_{2\,k+1} & \cdots & a_{k-1\,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n\,k-1} & a_{n\,k} + b_{n\,k} & a_{n\,k+1} & \cdots & a_{nn} \end{bmatrix} =$$

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1\,k-1} & a_{1\,k} & a_{1\,k+1} & \cdots & a_{1n} \\ a_{2\,1} & \cdots & a_{2\,k-1} & a_{2\,k} & a_{2\,k+1} & \cdots & a_{k-1\,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n\,k-1} & a_{n\,k} & a_{n\,k+1} & \cdots & a_{nn} \end{bmatrix}$$

$$+ \det \begin{bmatrix} a_{11} & \cdots & a_{1\,k-1} & b_{1\,k} & a_{1\,k+1} & \cdots & a_{1n} \\ a_{2\,1} & \cdots & a_{2\,k-1} & b_{2\,k} & a_{2\,k+1} & \cdots & a_{k-1\,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n\,k-1} & b_{n\,k} & a_{n\,k+1} & \cdots & a_{nn} \end{bmatrix}.$$

Demonstração. Imediato pelo Teorema 6.1.9. □

Teorema 6.1.13. *Sejam* $n \in \mathbb{N}$ *e* $A \in \mathcal{M}_n(\mathbb{K})$. *Se* B *é uma matriz obtida de* A, *multiplicando uma sua linha por* $\alpha \in \mathbb{K}$. *Então,*

$$\det B = \alpha \det A.$$

Demonstração. Sejam $n \in \mathbb{N}$ e

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k-1\,1} & a_{k-2\,2} & \cdots & a_{k-1\,n} \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ a_{k+1\,1} & a_{k+1\,2} & \cdots & a_{k+1\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \text{ e } B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k-1\,1} & a_{k-2\,2} & \cdots & a_{k-1\,n} \\ \alpha \cdot a_{k1} & \alpha \cdot a_{k2} & \cdots & \alpha \cdot a_{kn} \\ a_{k+1\,1} & a_{k+1\,2} & \cdots & a_{k+1\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Então, desenvolvendo o determinante de B ao longo da sua linha k, tem-se

$$\begin{aligned} \det B &= \textstyle\sum_{j=1}^{n} (-1)^{k+j} (\alpha \cdot a_{kj}) \det B(k|j) \\ &= \alpha \textstyle\sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det A(k|j) \\ &= \alpha \det A. \end{aligned}$$

□

Corolário 6.1.14. *Sejam* $n \in \mathbb{N}$ *e* $A \in \mathcal{M}_n(\mathbb{K})$. *Se* B *é uma matriz obtida de* A, *multiplicando uma sua coluna por* $\alpha \in \mathbb{K}$. *Então,*

$$\det B = \alpha \det A.$$

Demonstração. Imediato pelo Teorema 6.1.9. □

Teorema 6.1.15. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se A tem uma linha ou coluna só com zeros, então* $\det A = 0$.

Demonstração. Se A tem uma linha ou coluna só com zeros, então A tem uma linha ou uma coluna multiplicada por $\alpha = 0$. Logo, o resultado é imediato pelo corolário anterior. $\square$

Teorema 6.1.16. *Sejam $n \in \mathbb{N}$, $A \in \mathcal{M}_n(\mathbb{K})$ e $\alpha \in \mathbb{K}$. Então,*

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

Demonstração. A prova é realizada por indução sobre n.
Caso base ($n = 1$): Se $A \in \mathcal{M}_1(\mathbb{K})$ temos $A = [a_{11}]$, pelo que

$$\det(\alpha A) = \det[\alpha a_{11}] = \alpha a_{11} = \alpha^1 \det A.$$

Passo de indução: Seja $p \in \mathbb{N}$. Admitamos, por hipótese de indução, que, para qualquer matriz $A' \in \mathcal{M}_p(\mathbb{K})$, $\det(\alpha A') = \alpha^p \det A'$. Com base nesta hipótese prova--se que se $A \in \mathcal{M}_{p+1}(\mathbb{K})$, então $\det(\alpha A) = \alpha^{p+1} \det A$. De facto, desenvolvendo o determinante de αA ao longo da linha k, $k \in \{1, 2 \ldots, p+1\}$, tem-se

$$\det(\alpha A) = \sum_{j=1}^{p+1} (-1)^{k+j} (\alpha a_{kj}) \det(\alpha A)(k|j).$$

Mas $(\alpha A)(k|j) = \alpha A(k|j)$, $A(k|j) \in \mathcal{M}_p(\mathbb{K})$ e, por hipótese de indução,

$$\det \alpha A(k|j) = \alpha^p \det A(k|j).$$

Logo,

$$\begin{aligned}
\det(\alpha A) &= \sum_{j=1}^{p+1} (-1)^{k+j} (\alpha a_{kj})(\alpha^p \det A(k|j)) \\
&= \alpha^{p+1}(\sum_{j=1}^{p+1} (-1)^{k+j} a_{kj} \det A(k|j)) \\
&= \alpha^{p+1} \det A.
\end{aligned}$$

Assim, pelo Princípio de Indução em $\mathbb{N}$, conclui-se que, para todo $n \in \mathbb{N}$, se $A \in \mathcal{M}_n(\mathbb{K})$, então $\det(\alpha A) = \alpha^n \det A$. $\square$

Teorema 6.1.17. *Sejam $n \in \mathbb{N}$ tal que $n \geq 2$ e seja $A \in \mathcal{M}_n(\mathbb{K})$. Se B é uma matriz obtida de A trocando duas das suas linhas, então*

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

Demonstração. A prova é realizada por indução sobre a ordem n da matriz.
Caso base ($n = 2$): Para $n = 2$ o resultado é válido, pois

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21} = -(a_{21}a_{12} - a_{22}a_{11}) = -\det \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix}.$$

Passo de indução: Seja $p \in \mathbb{N}$ tal que $p \geq 2$. Admitamos, por hipótese de indução, que se B' é uma matriz obtida de uma matriz $A' \in \mathcal{M}_p(\mathbb{K})$ trocando duas das suas linhas, então $\det B' = -\det A'$. Com base nesta hipótese mostra-se que se B é uma matriz obtida de uma matriz $A \in \mathcal{M}_{p+1}(\mathbb{K})$ trocando duas das suas linhas, então $\det B = -\det A$. Suponhamos que B é a matriz obtida de A trocando as linhas i e j, $i \neq j$, $i, j \in \{1, \ldots, p+1\}$. Como $p + 1 \geq 3$, a matriz A tem uma linha $k \in \{1, \ldots, p+1\}$ tal que $k \neq i$ e $k \neq j$. Então, aplicando o Teorema de Laplace ao longo da linha k da matriz A, temos

$$\det A = \sum_{l=1}^{p+1} (-1)^{k+l} \times a_{kl} \times \det A(k|l).$$

Para quaisquer $k, l \in \{1, \ldots, p+1\}$, a matriz $A(k|l)$ é uma matriz de $\mathcal{M}_p(\mathbb{K})$ e, como $k \neq i$ e $k \neq j$, a matriz $B(k|l)$ é a matriz obtida de $A(k|l)$ trocando as linhas i e j. Por conseguinte, por hipótese de indução, temos $\det B(k|l) = -\det A(k|l)$, donde segue que

$$\sum_{l=1}^{p+1} (-1)^{k+l} \times a_{kl} \times \det A(k|l) = \sum_{l=1}^{p+1} (-1)^{k+l} \times a_{kl} \times (-\det B(k|l)).$$

Agora, como a linha k de A é igual à linha k de B, temos

$$\sum_{l=1}^{p+1} (-1)^{k+l} \times a_{kl} \times (-\det B(k|l)) = \sum_{l=1}^{p+1} (-1)^{k+l} \times b_{kl} \times (-\det B(k|l)) = -\det B.$$

Assim, pelo Princípio de Indução em $\mathbb{N}$, concluímos que, para todo $n \in \mathbb{N}$, se $A \in \mathcal{M}_n(\mathbb{K})$ e B é uma matriz obtida de A trocando duas das suas linhas, então $\det B = -\det A$. □

Corolário 6.1.18. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se B é uma matriz obtida de A trocando duas das suas colunas, então*

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

Demonstração. Resulta do Teorema 6.1.9. □

Corolário 6.1.19. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se A tem duas linhas iguais, então* $\det A = 0$.

Demonstração. Se trocarmos as duas linhas iguais da matriz A, obtemos a mesma matriz A. Mas, pelo Teorema 6.1.17, $\det A = -\det A$, pelo que $\det A = 0$. □

Corolário 6.1.20. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se A tem duas colunas iguais, então* $\det A = 0$.

Teorema 6.1.21. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se B é uma matriz obtida de A, substituindo uma sua linha pela sua soma com um múltiplo de outra linha, então*

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

Demonstração. Sejam $n, k, p \in \mathbb{N}$ tais que $1 \leq k < p \leq n$, $\alpha \in \mathbb{K}$ e

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k\,1} & a_{k\,2} & \cdots & a_{k\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p\,1} & a_{p\,2} & \cdots & a_{p\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \text{ e } B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k\,1} & a_{k\,2} & \cdots & a_{k\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p\,1} + \alpha \cdot a_{k1} & a_{p\,2} + \alpha \cdot a_{k2} & \cdots & a_{p\,n} + \alpha \cdot a_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Então

$$\begin{aligned} \det B &= \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k\,1} & a_{k\,2} & \cdots & a_{k\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p\,1} + \alpha \cdot a_{k1} & a_{p\,2} + \alpha \cdot a_{k2} & \cdots & a_{p\,n} + \alpha \cdot a_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \\ &= \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k\,1} & a_{k\,2} & \cdots & a_{k\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p\,1} & a_{p\,2} & \cdots & a_{p\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} + \alpha \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k\,1} & a_{k\,2} & \cdots & a_{k\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \\ &= \det A + \alpha \cdot 0 = \det A. \end{aligned}$$

$\square$

Corolário 6.1.22. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se B é uma matriz obtida de A, substituindo uma sua coluna pela sua soma com um múltiplo de outra coluna, então*

$$\det B = \det A. \quad \square$$

Considerando algumas das propriedades dos determinantes referidas anteriormente, o cálculo do determinante de uma matriz pode ser realizado recorrendo ao método de eliminação de Gauss.

Teorema 6.1.23. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Se U é uma matriz em escada obtida de A por aplicação do método de eliminação de Gauss, então* $\det A = \det U$ *ou* $\det A = -\det U$.

Demonstração. Verificámos atrás que o determinante de uma matriz A não se altera se substituirmos uma das suas linhas pela sua soma com outra multiplicada por um escalar. Assim, se U é uma matriz em escada obtida de A por aplicação de transformações lineares que sejam troca de linhas (resp. colunas) ou substituição de uma das suas linhas (resp. colunas) pela sua soma com outra linha (resp. coluna) multiplicada por um escalar, tem-se $\det A = (-1)^l \det U$, onde l é o número de trocas de linhas efetuadas até à obtenção da matriz U. Se o número de trocas de linhas realizadas for par, tem-se $\det A = \det U$; se o número de trocas de linhas for ímpar, então $\det A = -\det U$. $\square$

De um modo geral, se $U = [u_{ij}] \in \mathcal{M}_n(\mathbb{K})$ é uma matriz triangular superior (inferior) obtida de uma matriz $A \in \mathcal{M}_n(\mathbb{K})$ efectuando operações elementares sobre as linhas (ou colunas) de A , tem-se

$$\det A = (-1)^l \times \beta \times u_{11} \times u_{22} \times \cdots \times u_{nn},$$

onde l é o número de vezes que trocamos duas linhas ou duas colunas e β é o inverso do produto dos escalares pelos quais multiplicamos as linhas ou colunas.

Exemplo 6.1.24. *Seja* $A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}$. *Então*

$$\begin{vmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{vmatrix} \overset{l_2 \to l_2 - 2l_2}{=} \begin{vmatrix} 1 & 1 & 2 & 2 \\ 0 & 0 & -2 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{vmatrix} \overset{l_2 \leftrightarrow l_3}{=} - \begin{vmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix}$$

$$\overset{l_3 \to \frac{1}{2}l_3}{=} -2 \begin{vmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & -1 & \frac{1}{2} \\ 0 & 0 & 1 & 2 \end{vmatrix} \overset{l_4 \to l_4 + l_3}{=} -2 \begin{vmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & -1 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{5}{2} \end{vmatrix}$$

$$= -2 \times \left(1 \times 1 \times (-1) \times \tfrac{5}{2}\right) = 5.$$

Facilmente se encontram exemplos de matrizes quadradas A e B tais que $\det(A + B) \neq \det A + \det B$. No entanto, como iremos verificar mais à frente, o determinante do produto de matrizes quadradas é sempre igual ao produto dos determinantes das matrizes fatores. No sentido de estabelecermos tal resultado, começamos por provar alguns resultados auxiliares.

Teorema 6.1.25. *Sejam* $n \in \mathbb{N}$, $A \in \mathcal{M}_n(\mathbb{K})$ *e* $E_1, \ldots, E_s$ *matrizes elementares de* $\mathcal{M}_n(\mathbb{K})$. *Então*

$$\det(E_1 \ldots E_s A) = \det(E_1 \ldots E_s) \det A.$$

Demonstração. A prova é feita por indução sobre s.

Caso base ($s = 1$): Pretendemos mostrar que $\det E_1 A = \det E_1 \det A$. Para provar este resultado vamos considerar os 3 casos seguintes:

i) E_1 é a matriz obtida de I_n trocando as linhas i e j, com $i \neq j$;

ii) E_1 é a matriz obtida de I_n multiplicando a linha i por $\alpha \in \mathbb{K} \setminus \{0\}$.

iii) E_1 é a matriz obtida de I_n substituindo a linha i pela sua soma com a linha j multiplicada por $\alpha \in \mathbb{K}$, com $i \neq j$.

Caso i): Pela Proposição 6.1.17 temos $\det E_1 = -\det \mathrm{I}_n = -1$ e, uma vez que $E_1 A$ é a matriz obtida de A trocando as linhas i e j, também pela Proposição 6.1.17, temos $\det(E_1 A) = -\det A$. Logo

$$\det(E_1 A) = -\det A = \det E_1 \det A.$$

Caso ii) Pela proposição 6.1.13 temos $\det E_1 = \alpha \det \mathrm{I}_n = \alpha$ e, uma vez que $E_1 A$ é a matriz obtida de A multiplicando a linha i por α tem-se $\det(E_1 A) = \alpha \det A$. Logo

$$\det(E_1 A) = \alpha \det A = \det E_1 \det A.$$

Caso iii): Pela Proposição 6.1.21 temos $\det E_1 = \det I_n = 1$ e, uma vez que $E_1 A$ é a matriz obtida de A susbstituindo a linha i pela sua soma com a linha j multiplicada por α, tem-se $\det(E_1 A) = \det A$. Logo

$$\det(E_1 A) = \det A = \det E_1 \det A.$$

Passo de indução: Dado $k \in \mathbb{N}$, admitamos, por hipótese de indução, que o resultado é válido para o produto de quaisquer k matrizes elementares de $\mathcal{M}_n(\mathbb{K})$ por qualquer matriz de $\mathcal{M}_n(\mathbb{K})$. Então, considerando que

$$\det(E_1 \dots E_k E_{k+1} A) = \det((E_1 \dots E_k)(E_{k+1} A)),$$

pela hipótese de indução tem-se

$$\det(E_1 \dots E_k E_{k+1} A) = \det(E_1 \dots E_k) \det(E_{k+1} A).$$

Além disso, do que foi provado no caso base temos

$$\det(E_{k+1} A) = \det E_{k+1} \det A.$$

Assim,

$$\begin{aligned} \det(E_1 \dots E_k E_{k+1} A) &= \det(E_1 \dots E_k) \det E_{k+1} \det A \\ &= \det(E_1 \dots E_k E_{k+1}) \det A. \end{aligned}$$

Do que foi provado concluímos, pelo Princípio de Indução em $\mathbb{N}$, que, para qualquer $s \in \mathbb{N}$, para quaisquer matrizes elementares $E_1, \dots E_s \in \mathcal{M}_n(\mathbb{K})$ e para qualquer matriz $A \in \mathcal{M}_n(\mathbb{K})$,

$$\det(E_1 \dots E_s A) = \det(E_1 \dots E_s) \det A.$$

□

Teorema 6.1.26. *Sejam* $n \in \mathbb{N}$ *e* $A, B \in \mathcal{M}_n(\mathbb{K})$*. Então,*

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

Demonstração. Na prova deste resultado consideramos dois casos: i) A é invertível; ii) A não é invertível.

Caso i): Se A é invertível, então A é equivalente por linhas à matriz I_n. Por conseguinte, existem matrizes elementares $E_1, \ldots E_s \in \mathcal{M}_n(\mathbb{K})$ tais que $A = E_1 \ldots E_s I_n$. Logo, pelo resultado anterior, tem-se

$$\det(AB) = \det(E_1 \ldots E_s B) = \det(E_1 \ldots E_s) \det B = \det A \det B.$$

Caso ii): Se A não é invertível, então $car(A) < n$. Logo existem matrizes elementares $E_1, \ldots, E_s \in \mathcal{M}_n(\mathbb{K})$ tais que

$$E_1 \ldots E_s A = A'$$

onde A' é uma matriz triangular superior com um número de linhas não nulas menor do que n, ou seja, A' tem pelo menos uma linha nula.

Da igualdade anterior, e tendo em conta que toda a matriz elementar é invertível, temos

$$A = {E_s}^{-1} \ldots {E_1}^{-1} A'.$$

Então, atendendo a que inversa de uma matriz elementar é também uma matriz elementar, pelo teorema anterior temos

$$\begin{aligned} \det(AB) &= \det(({E_s}^{-1} \ldots {E_1}^{-1} A')B) \\ &= \det(({E_s}^{-1} \ldots {E_1}^{-1})(A'B)) \\ &= \det({E_s}^{-1} \ldots {E_1}^{-1}) \det(A'B). \end{aligned}$$

Daqui segue que $\det(AB) = 0$, pois, como A' tem pelo menos uma linha nula, a matriz $A'B$ também tem pelo menos uma linha nula, pelo que $\det(A'B) = 0$. A matriz $A = {E_s}^{-1} \ldots {E_1}^{-1} A'$ também tem pelo menos uma linha nula e, portanto, $\det A = 0$. Assim,

$$\det(AB) = 0 = 0 \det B = \det A \det B. \qquad \square$$

6.2 Cálculo da inversa a partir da adjunta

No capítulo anterior foram apresentadas várias condições para a caracterização de matrizes invertíveis e apresentou-se um processo para o cálculo da inversa de matrizes invertíveis. Seguidamente apresentamos mais uma caracterização de matrizes invertíveis e um processo para o cálculo da inversa de uma matriz invertível, mas neste caso recorrendo a determinantes.

Definição 6.2.1. *Sejam* $n \in \mathbb{N}$ *e* $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{K})$*. Chama-se* ***matriz adjunta de A****, e representa-se por* $\operatorname{Adj} A$*, à matriz*

$$\operatorname{Adj}A = [\hat{a}_{ij}]^T.$$

Exemplo 6.2.2. *Seja* $A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 5 & 0 \\ 4 & 0 & 7 \end{bmatrix}$.

Como

$$\hat{a}_{11} = (-1)^2 \begin{vmatrix} 5 & 0 \\ 0 & 7 \end{vmatrix} = 35, \quad \hat{a}_{12} = (-1)^3 \begin{vmatrix} 0 & 0 \\ 4 & 7 \end{vmatrix} = 0, \quad \hat{a}_{13} = (-1)^4 \begin{vmatrix} 0 & 5 \\ 4 & 0 \end{vmatrix} = -20,$$

$$\hat{a}_{21} = (-1)^3 \begin{vmatrix} 0 & 3 \\ 0 & 7 \end{vmatrix} = 0, \quad \hat{a}_{22} = (-1)^4 \begin{vmatrix} 1 & 3 \\ 4 & 7 \end{vmatrix} = -5, \quad \hat{a}_{23} = (-1)^5 \begin{vmatrix} 1 & 0 \\ 4 & 0 \end{vmatrix} = 0,$$

$$\hat{a}_{31} = (-1)^4 \begin{vmatrix} 0 & 3 \\ 5 & 0 \end{vmatrix} = -15, \quad \hat{a}_{32} = (-1)^5 \begin{vmatrix} 1 & 3 \\ 0 & 0 \end{vmatrix} = 0, \quad \hat{a}_{33} = (-1)^6 \begin{vmatrix} 1 & 0 \\ 0 & 5 \end{vmatrix} = 5,$$

temos

$$AdjA = \begin{bmatrix} 35 & 0 & -20 \\ 0 & -5 & 0 \\ -15 & 0 & 5 \end{bmatrix}^T = \begin{bmatrix} 35 & 0 & -15 \\ 0 & -5 & 0 \\ -20 & 0 & 5 \end{bmatrix}.$$

Teorema 6.2.3. *Sejam* $n \in \mathbb{N}$ *e* $A = [a_{ij}]_n$ *uma matriz quadrada de ordem* n *sobre* $\mathbb{K}$. *Então, se* $i \neq j$,

1. $a_{i1}\hat{a}_{j1} + a_{i2}\hat{a}_{j2} + \cdots a_{in}\hat{a}_{jn} = 0$.

2. $a_{1i}\hat{a}_{1j} + a_{2i}\hat{a}_{2j} + \cdots a_{ni}\hat{a}_{nj} = 0$.

Demonstração. 1. Sejam $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{K})$ e A' a matriz que se obtém de A substituindo a linha j pela linha i. Então $\det A' = 0$, uma vez que A' tem duas linhas iguais. Por outro lado, aplicando o Teorema de Laplace ao longo da linha j de A', tem-se

$$\begin{aligned} \det A' &= {a'}_{j1}\hat{a}'_{j1} + {a'}_{j2}\hat{a}'_{j2} + \ldots {a'}_{jn}\hat{a}'_{jn} \\ &= a_{i1}(-1)^{j+1} \det A'(j|1) + a_{i2}(-1)^{j+2} \det A'(j|2) + \ldots + a_{in}(-1)^{j+n} \det A'(j|n) \\ &= a_{i1}(-1)^{j+1} \det A(j|1) + a_{i2}(-1)^{j+2} \det A(j|2) + \ldots + a_{in}(-1)^{j+n} \det A(j|n), \end{aligned}$$

pois, para cada $k \in \{1, \ldots, n\}$, $A'(j|k) = A(j|k)$.
Logo $a_{i1}\hat{a}_{j1} + a_{i2}\hat{a}_{j2} + \cdots a_{in}\hat{a}_{jn} = 0$.
2. A prova é análoga à anterior; basta considerar a matriz A' que se obtém de A substituindo a coluna j pela coluna i e desenvolver o determinante ao longo da coluna j de A'. □

Teorema 6.2.4. *Sejam* $n \in \mathbb{N}$ *e* A *uma matriz quadrada de ordem* n *sobre* $\mathbb{K}$. *Então*

1. *A é invertível se e só se* $\det A \neq 0$.

2. *Se A é invertível, então* $\det A^{-1} = (\det A)^{-1}$.

3. $A\mathrm{Adj}A = (\det A)I_n$.

4. *Se A é invertível, então*
$$A^{-1} = \frac{1}{\det A}\mathrm{Adj}A.$$

Demonstração. 1. Seja $A \in \mathcal{M}_n(\mathbb{K})$ tal que A não é invertível. Então $c(A) < n$. Seja $F \in \mathcal{M}_n(\mathbb{K})$ uma matriz em escada equivalente por linhas à matriz A. Então $A = E_1 \dots E_s F$, onde E_1,..., E_s são matrizes elementares de ordem n. Logo,

$$\det A = \det(E_1 \dots E_s) \det F.$$

Uma vez que $c(A) < n$, a matriz F tem, pelo menos, uma linha nula. Assim, $\det F = 0$ e, portanto $\det A = 0$.

Reciprocamente, suponhamos que $A \in \mathcal{M}_n(\mathbb{K})$ é uma matriz invertível. Então, existe $A^{-1} \in \mathcal{M}_n(\mathbb{K})$ tal que $AA^{-1} = I_n = A^{-1}A$. Logo, como A e A^{-1} são ambas matrizes quadradas de ordem n, tem-se $\det A . \det A^{-1} = \det(AA^{-1}) = \det I_n = 1$, pelo que $\det A \neq 0$.

2. Segue da prova de *1.*

3. Pela definição de produto de matrizes, o elemento na linha i e coluna j da matriz $A\mathrm{Adj}A$ é o elemento

$$a_{i1}\hat{a}_{j1} + a_{i2}\hat{a}_{j2} + \cdots a_{in}\hat{a}_{jn}.$$

Se $i = j$, tem-se

$$a_{i1}\hat{a}_{i1} + a_{i2}\hat{a}_{i2} + \cdots a_{in}\hat{a}_{in} = \det A.$$

Se $i \neq j$,

$$a_{i1}\hat{a}_{j1} + a_{i2}\hat{a}_{j2} + \cdots a_{in}\hat{a}_{jn} = 0.$$

Por conseguinte,

$$A\mathrm{Adj}A = \begin{bmatrix} \det A & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \det A & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det A \end{bmatrix} = (\det A)I_n$$

4. Se A é invertível, então, por *3.*, tem-se $A^{-1}A\mathrm{Adj}A = A^{-1}(\det A)I_n$, pelo que $\mathrm{Adj}A = (\det A)A^{-1}$. Como $\det A \neq 0$, segue que $A^{-1} = \frac{1}{\det A}\mathrm{Adj}A$. □

Exemplo 6.2.5. *Seja* $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$. *Como* $\det A = 1 \times 6 - 3 \times 2 = 0$, *então a matriz* A *não é invertível.*

Exemplo 6.2.6. *Seja*

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

Como $\det A = -25$ *e*

$$\mathrm{Adj}A = \begin{bmatrix} 35 & 0 & -15 \\ 0 & -5 & 0 \\ -20 & 0 & 5 \end{bmatrix},$$

tem-se

$$A^{-1} = \begin{bmatrix} -\frac{35}{25} & 0 & \frac{15}{25} \\ 0 & \frac{5}{25} & 0 \\ \frac{20}{25} & 0 & -\frac{5}{25} \end{bmatrix} = \begin{bmatrix} -\frac{7}{5} & 0 & \frac{3}{5} \\ 0 & \frac{1}{5} & 0 \\ \frac{4}{5} & 0 & -\frac{1}{5} \end{bmatrix}.$$

6.3 Regra de Cramer

Um sistema de n equações lineares em n incógnitas que seja possível e determinado diz-se um *sistema de Cramer*. O resultado seguinte estabelece como calcular a única solução de um sistema de Cramer recorrendo a deteminantes.

Definição 6.3.1. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{K})$. Um sistema de equações lineares (S) com equação matricial $Ax = b$ diz-se um* ***sistema de Cramer*** *se A é uma matriz invertível.*

Teorema 6.3.2. *Sejam $n \in \mathbb{N}$, $A \in \mathcal{M}_n(\mathbb{K})$ e (S) um sistema de equações lineares com equação maticial $Ax = b$. Se A é invertível, então o sistema é possível e determinado e a única solução do sistema (S) é o n-uplo $(\alpha_1, \alpha_2, ..., \alpha_n)$, onde*

$$\alpha_i = \frac{\det\left(A^{(i)}\right)}{\det A}, \qquad i = 1, 2, ..., n,$$

e $A^{(i)}$ é a matriz quadrada de ordem n obtida de A substituíndo a sua coluna i pela coluna b.

Demonstração. Seja $A = [a_{ij}]_n$ uma matriz invertível. Então $c(A) = n$. Como $A \in \mathcal{M}_n(\mathbb{K})$ e $c(A) = c(A|b) = n$, o sistema é possível e determinado.
Sendo $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ a única solução do sistema, tem-se

$$A \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = b.$$

Da igualdade anterior obtem-se

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = A^{-1}b.$$

Considerando que $A^{-1} = \frac{1}{\det A}\mathrm{Adj}A$, segue que

$$\begin{aligned} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} &= \frac{1}{\det A} \begin{bmatrix} \hat{a}_{11} & \hat{a}_{21} & \cdots & \hat{a}_{n1} \\ \hat{a}_{12} & \hat{a}_{22} & \cdots & \hat{a}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{a}_{1n} & \hat{a}_{2n} & \cdots & \hat{a}_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ &= \frac{1}{\det A} \begin{bmatrix} \hat{a}_{11}b_1 + \hat{a}_{21}b_2 + \cdots + \hat{a}_{n1}b_n \\ \hat{a}_{12}b_1 + \hat{a}_{22}b_2 + \cdots + \hat{a}_{n2}b_n \\ \vdots \\ \hat{a}_{1n}b_1 + \hat{a}_{2n}b_2 + \cdots + \hat{a}_{nn}b_n \end{bmatrix}. \end{aligned}$$

Assim, para cada $i = 1, 2, ..., n$,

$$\begin{aligned}\alpha_i &= \frac{1}{\det A}\left(\hat{a}_{1i}b_1 + \hat{a}_{2i}b_2 + \cdots + \hat{a}_{ni}b_n\right)\\ &= \frac{1}{\det A}\begin{vmatrix} a_{11} & \cdots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \cdots & a_{1n}\\ a_{21} & \cdots & a_{2\,i-1} & b_2 & a_{2\,i+1} & \cdots & a_{2n}\\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{n\,i-1} & b_n & a_{n\,i+1} & \cdots & a_{nn}\end{vmatrix}.\end{aligned}$$

□

Exemplo 6.3.3. *Consideremos o sistema de equações lineares a seguir indicado*

$$\begin{cases} x_1 + 2x_2 - 3x_3 + x_4 = -5\\ x_2 + 3x_3 + x_4 = 6\\ 2x_1 + 3x_2 + x_3 + x_4 = 4\\ x_1 + x_3 + x_4 = 1\end{cases}$$

O sistema pode ser representado matricialmente por $Ax = b$, onde

$$A = \begin{bmatrix} 1 & 2 & -3 & 1\\ 0 & 1 & 3 & 1\\ 2 & 3 & 1 & 1\\ 1 & 0 & 1 & 1\end{bmatrix}, \quad b = \begin{bmatrix} -5\\ 6\\ 4\\ 1\end{bmatrix}, \quad x = \begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}.$$

Uma vez que $|A| = 20 \neq 0$, a matriz A é invertível e o sistema indicado é um sistema de Cramer. A única solução deste sistema é $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, onde

$$\alpha_1 = \frac{\begin{vmatrix} -5 & -2 & -3 & 1\\ 6 & 1 & 3 & 1\\ 4 & 3 & 1 & 1\\ 1 & 0 & 1 & 1\end{vmatrix}}{20} = \frac{0}{20} = 0, \qquad \alpha_2 = \frac{\begin{vmatrix} 1 & -5 & -3 & 1\\ 0 & 6 & 3 & 1\\ 2 & 4 & 1 & 1\\ 1 & 1 & 1 & 1\end{vmatrix}}{20} = \frac{20}{20} = 1,$$

$$\alpha_3 = \frac{\begin{vmatrix} 1 & 2 & -5 & 1\\ 0 & 1 & 6 & 1\\ 2 & 3 & 4 & 1\\ 1 & 0 & 1 & 1\end{vmatrix}}{20} = \frac{40}{20} = 2, \qquad \alpha_4 = \frac{\begin{vmatrix} 1 & 2 & -3 & -5\\ 0 & 1 & 3 & 6\\ 2 & 3 & 1 & 4\\ 1 & 0 & 1 & 1\end{vmatrix}}{20} = \frac{-20}{20} = -1.$$

6.4 Determinantes e positividade de matrizes

As matrizes simétricas definidas positivas podem ser caracterizadas recorrendo ao cálculo de determinantes.

No sentido de estabelecermos tal caracterização, começamos por introduzir alguma terminologia.

Definição 6.4.1. *Sejam $A \in \mathcal{M}_n(\mathbb{K})$ e $k \in \{1, \dots, n\}$. Ao determinante de uma submatriz principal de A de ordem k dá-se a designação de* ***menor principal de*** A ***de ordem*** k *e ao determinante de uma submatriz principal primária de A de ordem k dá-se a designação de* ***menor principal primário de*** A ***de ordem*** k.

Exemplo 6.4.2. *Consideremos a matriz* $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$. *Então os menores principais primários de A são os seguintes:*

- *menor principal primário de ordem 1:* $|A_1| = |1| = 1$;
- *menor principal primário de ordem 2:* $|A_2| = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0$;
- *menor principal primário de ordem 3:* $|A_3| = |A| = -1$.

Teorema 6.4.3. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz simétrica. A matriz A é simétrica definida positiva se e só se todos os seus menores principais primários são positivos, ou seja, se e só se, para todo $k \in \{1, \dots, n\}$, $|A_k| > 0$.*

Demonstração. Prova por indução sobre a ordem n da matriz. □

Exemplo 6.4.4. *Consideremos a matriz* $A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \in \mathcal{M}_3(\mathbb{R})$.
Como A é simétrica,

$$|A_1| = \det[2] = 2 > 0,$$

$$|A_2| = \det \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 2 \times 2 - 0 \times 0 = 4 > 0$$

e

$$|A_3| = \det \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} = 4 + 0 + 0 - 2 - 0 - 0 = 2 > 0,$$

a matriz A é simétrica definida positiva.

Corolário 6.4.5. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz simétrica. Se a matriz A é simétrica definida positiva, então A é invertível.*

Teorema 6.4.6. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz simétrica. A matriz A é simétrica semidefinida positiva se e só se todas as submatrizes principais de A têm determinantes não negativos, ou seja, se e só se todos os menores principais de A são positivos ou nulos.*

Exemplo 6.4.7. *Seja*

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathcal{M}_{3\times 3}(\mathbb{R}).$$

Calculando os menores principais de A, obtemos:

- *Menores de ordem* 1*:* $\det([2]) = 2$, $\det([1]) = 1$, $\det([0]) = 0$*;*
- *Menores de ordem* 2*:*

$$\det \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = 1, \quad \det \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = 0, \quad \det \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 0.$$

- *Menor de ordem* 3*:* $\det(A) = 0$.

Todos os menores principais de A são positivos ou nulos, pelo que a matriz é simétrica semidefinida positiva (mas não é simétrica definida positiva, pois o menor principal primário de ordem 3 é nulo).

Exemplo 6.4.8. *Consideremos a matriz*

$$A = \begin{bmatrix} -5 & 3 \\ 3 & 0 \end{bmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R}).$$

Calculando os menores principais de A, obtemos:

- *Menores de ordem* 1*:* $\det([-5]) = -5$, $\det([0]) = 0$*;*
- *Menores de ordem* 2*:* $\det A = (-5) \times 0 - 3 \times 3 = -9$.

A matriz tem menores principais negativos, logo a matriz não é semidefinida positiva. Portanto, a matriz também não é definida positiva.

Todos os resultados anteriores, a respeito de matrizes simétricas definidas positivas, podem também ser estabelecidos para matrizes hermíticas definidas (resp. semidefinidas) positivas.

7. Valores e Vetores Próprios

7.1 Definição e propriedades

Neste capítulo estudamos os conceitos de vetor próprio e valor próprio de aplicações lineares que sejam endomorfismos de um subespaço vectorial de $\mathbb{R}^n$ e estudamos também os conceitos de vetor próprio e de valor próprio de matrizes quadradas reais. O interesse no estudo destes conceitos deve-se às suas diversas aplicações práticas. No contexto do estudo matricial, os valores e vetores próprios aplicam-se no processo de diagonalização de matrizes quadradas.

Definição 7.1.1. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ e $f \in \mathcal{L}(V, V)$. Diz-se que $v \in V$ é* ***vetor próprio*** *de f se são satisfeitas as duas condições seguintes:*

i) $v \neq 0_{\mathbb{R}^n}$;

ii) existe $\lambda \in \mathbb{R}$ *tal que* $f(v) = \lambda v$.

O escalar λ diz-se um ***valor próprio de*** f. *Ao conjunto de valores próprios de f dá-se a designação de* ***espectro de*** f.

O conceito anterior pode ser generalizado a endomorfismos de espaços vetoriais em geral, mas, no âmbito deste curso, restringimos o estudo a endomorfismos de subespaços vetorias de $\mathbb{R}^n$.

Observação: Sejam V um subespaço vectorial de $\mathbb{R}^n$ e $f \in \mathcal{L}(V, V)$.

1) Se $v \in V$ é um vetor próprio de f, então existe um único valor próprio associado a v, i.e., existe um único escalar $\lambda \in \mathbb{R}$ tal que $f(v) = \lambda v$. De facto, se $\lambda_1, \lambda_2 \in \mathbb{R}$ são escalares tais que $f(v) = \lambda_1 v$ e $f(v) = \lambda_2 v$, tem-se $\lambda_1 v - \lambda_2 v = 0_{\mathbb{R}^n}$, pelo que $(\lambda_1 - \lambda_2)v = 0_{\mathbb{R}^n}$ e, como $v \neq 0_{\mathbb{R}^n}$, segue que $\lambda_1 - \lambda_2 = 0$, donde $\lambda_1 = \lambda_2$.
Assim, diz-se que v **é vetor próprio de** f **associado ao valor próprio** λ.

2) Se $\lambda \in \mathbb{R}$ é um valor próprio de f, prova-se que existem vetores próprios distintos associados ao mesmo valor próprio λ. Com efeito, se v é um vetor próprio associado a λ, então, para qualquer $\alpha \in \mathbb{R}$,

$$f(\alpha v) = \alpha f(v) = \alpha(\lambda v) = \lambda(\alpha v).$$

Logo, para qualquer $\alpha \in \mathbb{R} \setminus \{0\}$, αv é vetor próprio de f associado a λ, uma vez que $\alpha v \neq 0_V$ e $f(\alpha v) = \lambda(\alpha v)$.

Exemplo 7.1.2. *Consideremos, no espaço vetorial real* $\mathbb{R}^3$*, o endomorfismo* f *definido por* $f(a,b,c) = (3a, 3b, b)$*, para todo* $(a,b,c) \in \mathbb{R}^3$*, e os vetores* $(0,3,1)$*,* $(1,6,2)$*,* $(2,0,0)$*. Tem-se*

$$\begin{array}{lclcl} f(0,3,1) & = & (0,9,3) & = & 3(0,3,1), \\ f(1,6,2) & = & (3,18,6) & = & 3(1,6,2), \\ f(2,0,0) & = & (6,0,0) & = & 3(2,0,0), \end{array}$$

logo $(0,3,1)$*,* $(1,6,2)$*,* $(2,0,0)$ *são vetores próprios de* f *associados ao valor próprio* 3*. Temos também*

$$\begin{array}{lclcl} f(0,0,5) & = & (0,0,0) & = & 0(0,0,5), \\ f(0,0,-2) & = & (0,0,0) & = & 0(0,0,-2), \end{array}$$

e, portanto, $(0,0,5)$*,* $(0,0,-2)$ *são vetores próprios de* f *associados ao valor próprio* 0.

Exemplo 7.1.3. *Sejam* V *um subespaço vectorial de* $\mathbb{R}^n$*,* (v_1, v_2) *uma base de* V *e* $f : V \to V$ *uma aplicação linear definida por* $f(v_1) = v_1 - v_2$ *e* $f(v_2) = -v_1 + v_2$*. Se considerarmos os vetores* $v_1' = v_1 + v_2$ *e* $v_2' = v_1 - v_2$*, tem-se* $v_1' \neq 0_{\mathbb{R}^n}$*,* $v_2' \neq 0_{\mathbb{R}^n}$*,* $f(v_1') = 0v_1'$ *e* $f(v_2') = 2v_2'$*. Logo,* v_1' *e* v_2' *são vetores próprios de* f *associados aos valores próprios* 0 *e* 2*, respetivamente.*

Exemplo 7.1.4. *Sejam* $\mathcal{B}$ *a base canónica de* $\mathbb{R}^2$ *e* f *o endomorfismo de* $\mathbb{R}^2$ *tal que*

$$M(f; \mathcal{B}, \mathcal{B}) = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}.$$

Então $(0,2)$ *é vetor próprio de* f *associado ao valor próprio* 2*, pois* $(0,2) \neq (0,0)$ *e* $\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$*, pelo que* $f(0,2) = (0,4) = 2(0,2)$*.*

Sendo V um subespaço vectorial de $\mathbb{R}^n$ de dimensão finita $r \geq 1$, $\mathcal{B} = (v_1, \ldots, v_r)$ uma base de V, $f \in \mathcal{L}(V, V)$, $A = M(f; \mathcal{B}, \mathcal{B})$, $v \in V$ e $[v]_{(v_i)}$ o vetor coluna de v de V na base $\mathcal{B}$, tem-se

$$v \neq 0_{\mathbb{R}^n} \text{ se e só se } [v]_{(v_i)} \neq 0_{r \times 1}$$

e

$$f(v) = \lambda v \text{ se e só se } A[v]_{(v_i)} = \lambda [v]_{(v_i)},$$

o que motiva as definições seguintes.

Definição 7.1.5. *Sejam* $r \in \mathbb{N}$ *e* $A \in \mathcal{M}_r(\mathbb{R})$*. Diz-se que* $y \in \mathcal{M}_{r \times 1}(\mathbb{R})$ *é* ***vetor próprio*** *de* A *se são satisfeitas as duas condições seguintes:*

i) $y \neq 0_{r \times 1}$*;*

ii) existe $\lambda \in \mathbb{R}$ *tal que* $Ay = \lambda y$*.*

O escalar λ *diz-se um* ***um valor próprio de*** A*. Ao conjunto de valores próprios de* A *dá-se a designação de* ***espectro de*** A*.*

Os conceitos anteriores estendem-se naturalmente a matrizes complexas. Seja $A \in \mathcal{M}_{r\times r}(\mathbb{C})$. Diz-se que um vetor $y \in \mathcal{M}_{r\times 1}(\mathbb{C})$ é um vetor próprio de A se $y \neq 0_{r\times 1}$ e existe um escalar $\lambda \in \mathbb{C}$ tal que $Ay = \lambda y$. Neste caso, o número λ é denominado valor próprio de A.
Observe-se que toda matriz $A \in \mathcal{M}_{r\times r}(\mathbb{R})$ pode ser vista também como uma matriz complexa. Assim, ao estudar uma matriz quadradra $r \times r$ com entradas reais, é necessário explicitar se a consideramos como elemento de $\mathcal{M}_{r\times r}(\mathbb{R})$, restringindo-nos a valores próprios reais e vetores próprios em $\mathcal{M}_{r\times 1}(\mathbb{R})$, ou como elemento de $\mathcal{M}_{r\times r}(\mathbb{C})$, admitindo também valores próprios complexos e vetores próprios em $\mathcal{M}_{r\times 1}(\mathbb{C})$.
Neste curso, a não ser que algo seja dito em contrário, dada uma matriz quadrada real, estaremos apenas interessados no estudo de valores próprios reais e de vetores próprios de elementos reais.

Exemplo 7.1.6. *Sejam*

$$A = \begin{bmatrix} 1 & -2 \\ \frac{1}{2} & 3 \end{bmatrix}, x = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \text{ e } \lambda = 2.$$

Uma vez que

$$Ax = \begin{bmatrix} 1 & -2 \\ \frac{1}{2} & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \lambda x.$$

concluímos que $\lambda = 2$ é valor próprio de A e $x = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ é um vetor próprio de A associado a $\lambda = 2$.

Observação: 1) De modo análogo ao que sucede com os vetores próprios de um endomorfismo, um vetor próprio de uma matriz $A \in \mathcal{M}_r(\mathbb{R})$ está associado a um único valor próprio de A. De facto, para quaisquer $y \in \mathcal{M}_{r\times 1} \setminus \{0_{r\times 1}\}$ e $\lambda, \mu \in \mathbb{R}$,

$$(Ay = \lambda y \quad \text{e} \quad Ay = \mu y) \Rightarrow \lambda y = \mu y \Rightarrow (\lambda - \mu)\, y = 0_{r\times 1} \Rightarrow \lambda - \mu = 0 \Rightarrow \lambda = \mu.$$

Diz-se, então, que y **é vetor próprio de** A **associado ao valor próprio** λ.

2) A cada valor próprio λ de A está associada uma infinidade de vetores próprios de A. Com efeito, se $y \in \mathcal{M}_{r\times 1} \setminus \{0_{r\times 1}\}$ é um vetor próprio de A, então, para todo $\alpha \in \mathbb{R} \setminus \{0\}$, $x = \alpha y$ é também um vetor próprio de A, uma vez que $x \neq 0_{r\times 1}$ e

$$Ax = A\,(\alpha y) = \alpha\,(Ay) = \alpha\,(\lambda y) = \lambda\,(\alpha y) = \lambda x.$$

Como já foi observado anteriormente, as noções de valor próprio e vetor próprio de endomorfismos e de valor próprio e vetor próprio de matrizes quadradas estão relacionadas.

Teorema 7.1.7. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ de dimensão finita $r \geq 1$, $\mathcal{B} = (v_1, \ldots, v_r)$ uma base de V, $f \in \mathcal{L}(V, V)$ e $A = M(f; \mathcal{B}, \mathcal{B})$. Então*

1) $v \in V$ é vetor próprio de f se e só se o vetor coluna de v na base $\mathcal{B}$ é vetor próprio de A.

2) $\lambda \in \mathbb{R}$ é valor próprio de f se e só se λ é valor próprio de A.

Demonstração. 1) Sejam $v = \sum_{i=1}^{r} \alpha_i v_i \in V$ e $[v]_{v_i} = \left[\begin{array}{c} \alpha_1 \dots \alpha_r \end{array}\right]^T$ o vetor coluna de v na base $\mathcal{B}$.
Se v é vetor próprio de f tem-se $v \neq 0_{\mathbb{R}^n}$ e existe $\lambda \in \mathbb{R}$ tal que $f(v) = \lambda v$. Então $[v]_{(v_i)} \neq 0_{r\times 1}$ e

$$A[v]_{(v_i)} = \begin{bmatrix} \lambda\alpha_1 \\ \vdots \\ \lambda\alpha_r \end{bmatrix} = \lambda[v]_{(v_i)}.$$

Logo, $[v]_{v_i}$ é vetor próprio de A.

Reciprocamente, se admitirmos que $[v]_{(v_i)}$ é vetor próprio de A, temos $[v]_{(v_i)} \neq 0_{r\times 1}$ e $A[v]_{(v_i)} = \lambda[v]_{(v_i)}$, para algum $\lambda \in \mathbb{R}$. Logo $v \neq 0_{\mathbb{R}^n}$ e

$$f(v) = (\lambda\alpha_1)v_1 + \ldots + (\lambda\alpha_r)v_r = \lambda(\alpha_1 v_1 + \ldots + \alpha_r v_r) = \lambda v;$$

portanto v é vetor próprio de f.

2) Imediato a partir de *1)*. □

Exemplo 7.1.8. *Sejam $\mathcal{B}$ a base canónica de $\mathbb{R}^3$ e $f \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^3)$ tal que*

$$M(f; \mathcal{B}, \mathcal{B}) = \begin{bmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{bmatrix}.$$

Sendo $A = M(f; \mathcal{B}, \mathcal{B})$, tem-se

$$A \begin{bmatrix} 3 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} -6 \\ -6 \\ 0 \end{bmatrix} = -2 \begin{bmatrix} 3 \\ 3 \\ 0 \end{bmatrix}.$$

Logo, considerando que $(3, 3, 0) \neq 0_{\mathbb{R}^3}$ e $f(3, 3, 0) = (-6, -6, 0) = -2(3, 3, 0)$, o vetor $(3, 3, 0)$ é vetor próprio de f associado ao valor próprio -2. Atendendo a que

$$A \begin{bmatrix} 0 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 20 \\ 20 \end{bmatrix} = 4 \begin{bmatrix} 0 \\ 5 \\ 5 \end{bmatrix},$$

conclui-se que o vetor $(0, 5, 5)$ também é um vetor próprio de f, pois $(0, 5, 5) \neq 0_{\mathbb{R}^3}$ e temos $f(0, 5, 5) = (0, 20, 20) = 4(0, 5, 5)$; neste caso, $(0, 5, 5)$ é um vetor próprio de f associado ao valor próprio 4.

Teorema 7.1.9. *Sejam V um subespaço vetorial de $\mathbb{R}^n$, $f \in \mathcal{L}(V, V)$ e $\lambda \in \mathbb{R}$. Seja*

$$V_{[f,\lambda]} = \{v \in V : f(v) = \lambda v\} = \mathrm{Nuc}(f - \lambda id_V).$$

Então

1) *$V_{[f,\lambda]}$ é subespaço vetorial de $\mathbb{R}^n$.*

2) *Se $\lambda \in \mathbb{R}$ é um valor próprio de f, os vetores próprios de f associados ao valor próprio λ são os elementos de $V_{[f,\lambda]} \setminus \{0_{\mathbb{R}^n}\}$.*

Demonstração. 1) O resultado é imediato pelo Teorema 4.3.5.

2) Imediato pela definição de $V_{[f,\lambda]}$ e pela definição de vetor próprio de f associado a λ. $\square$

Observação: Sejam V um subespaço vectorial de $\mathbb{R}^n$, $f \in \mathcal{L}(V,V)$ e $\lambda \in \mathbb{R}$. Se λ não é valor próprio de f, tem-se $V_{[f,\lambda]} = \{0_{\mathbb{R}^n}\}$.

Definição 7.1.10. *Sejam V um subespaço vectorial de $\mathbb{R}^n$, $f \in \mathcal{L}(V,V)$ e $\lambda \in \mathbb{R}$ um valor próprio de f. Ao subespaço*

$$V_{[f,\lambda]} = \{v \in V : f(v) = \lambda v\} = \mathrm{Nuc}(f - \lambda id_V)$$

dá-se a designação de ***subespaço próprio de*** f ***associado ao valor próprio*** λ. *À dimensão do subespaço $V_{[f,\lambda]}$ dá-se a designação de* ***multiplicidade geométrica do valor próprio*** λ *e representa-se por* m.g.(λ).

Correspondendo ao Teorema 7.1.9 temos o resultado seguinte.

Teorema 7.1.11. *Sejam $r \in \mathbb{N}$, $A \in \mathcal{M}_r(\mathbb{R})$, $\lambda \in \mathbb{R}$ e*

$$\begin{aligned} M_{[A,\lambda]} &= \{y \in \mathcal{M}_{r\times 1}(\mathbb{R}) : Ay = \lambda y\} \\ &= \{y \in \mathcal{M}_{r\times 1}(\mathbb{R}) : (A - \lambda I_r)y = 0_{r\times 1}\}. \end{aligned}$$

Então

1) *$M_{[A,\lambda]}$ é um subespaço vetorial de $\mathcal{M}_{r\times 1}(\mathbb{R})$.*

2) *Se λ é um valor próprio de A, o conjunto dos vetores próprios de A associados ao valor próprio λ é $M_{[A,\lambda]} \setminus \{0_{r\times 1}\}$.*

Definição 7.1.12. *Sejam $r \in \mathbb{N}$, $A \in \mathcal{M}_r(\mathbb{R})$ e $\lambda \in \mathbb{R}$ um valor próprio de A. Ao subespaço*

$$\begin{aligned} M_{[A,\lambda]} &= \{y \in \mathcal{M}_{r\times 1}(\mathbb{R}) : Ay = \lambda y\} \\ &= \{y \in \mathcal{M}_{r\times 1}(\mathbb{R}) : (A - \lambda I_r)y = 0_{r\times 1}\}. \end{aligned}$$

dá-se a designação de ***subespaço próprio de*** A ***associado ao valor próprio*** λ. *À dimensão do subespaço $M_{[A,\lambda]}$ dá-se a designação de* ***multiplicidade geométrica do valor próprio*** λ *e representa-se por* m.g.(λ).

Atendendo à relação existente entre os vetores próprios de um endomorfismo f de um espaço vetorial de dimensão finita e os vetores próprios de uma matriz de f, é possível estabelecer o resultado seguinte.

Teorema 7.1.13. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $\mathcal{B} = (v_1, \ldots, v_r)$ uma base de V, $f \in \mathcal{L}(V,V)$, λ um valor próprio de f e $A = M(f; \mathcal{B}, \mathcal{B})$. Então existe uma bijeção entre $V_{[f,\lambda]}$ e $M_{[A,\lambda]}$.*

Demonstração. Do Teorema 7.1.7 resulta que a aplicação $\mu : V_{[f,\lambda]} \to M_{[A,\lambda]}$ tal que $\mu(v) = [v]_{\mathcal{B}}$, para todo $v \in V_{[f,\lambda]}$, está bem definida. É simples verificar que μ é uma aplicação bijetiva. ☐

Seguidamente estudamos processos para determinar os valores próprios e os vetores próprios de um endomorfismo dum subespaço vectorial de $\mathbb{R}^n$ de dimensão ≥ 1. Começamos por estudar como determinar os valores próprios de um dado endomorfismo e seguidamente veremos como determinar os vetores próprios associados a uma determinado valor próprio.

Teorema 7.1.14. *Sejam V um subespaço vetorial de $\mathbb{R}^n$ de dimensão finita $r \geq 1$, $\mathcal{B}$ uma base de V, $f \in \mathcal{L}(V, V)$, $A = M(f; \mathcal{B}, \mathcal{B})$ e $\lambda \in \mathbb{R}$. Então λ é valor próprio de f se e só se*

$$|A - \lambda I_r| = 0.$$

Demonstração. Pelo Teorema 7.1.7, λ é valor próprio de f se e só se λ é um valor próprio de A. Por definição, λ é valor próprio de A se e só se existe $y \in \mathcal{M}_{r\times 1}(\mathbb{R}) \setminus \{0_{r\times 1}\}$ tal que $Ay = \lambda y$. Então, considerando que

$$Ay = \lambda y \Leftrightarrow Ay - \lambda y = 0_{r\times 1} \Leftrightarrow Ay - \lambda I_r y = 0_{n\times 1} \Leftrightarrow (A - \lambda I_r)y = 0_{r\times 1},$$

λ é valor próprio de f se e se só o sistema $(A - \lambda I_r)x = 0_{r\times 1}$ é indeterminado, ou seja, se e só se $|A - \lambda I_r| = 0$. ☐

Teorema 7.1.15. *Sejam $r \in \mathbb{N}$ e $A \in \mathcal{M}_r(\mathbb{R})$. Então λ é valor próprio de A se e só se*

$$|A - \lambda \mathrm{I}_r| = 0.$$

Demonstração. Sejam $\mathcal{B}$ uma base do espaço vetorial $\mathbb{R}^r$ e f o endomorfismo de $\mathbb{R}^r$ tal que

$$M(f; \mathcal{B}, \mathcal{B}) = A.$$

Então, pelos teoremas 7.1.7 e 7.1.14, o resultado é imediato. ☐

O Teorema 7.1.15 motiva a definição seguinte.

Definição 7.1.16. *Sejam $r \in \mathbb{N}$ e $A = [a_{ij}] \in \mathcal{M}_r(\mathbb{R})$. Chama-se* **polinómio característico de** *A, e representa-se por $p_A(x)$, o polinómio na variável x e com coeficientes em $\mathbb{R}$, definido por*

$$p_A(x) = |A - x\mathrm{I}_r| = \begin{vmatrix} a_{11} - x & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} - x & \dots & a_{2r} \\ \vdots & \vdots & \vdots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rr} - x \end{vmatrix}.$$

Chama-se ***raiz característica de*** *A a qualquer raiz do polinómio $p_A(x)$ em $\mathbb{R}$.*

Exemplo 7.1.17. *Seja* $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{bmatrix} \in \mathcal{M}_3(\mathbb{R})$. *O polinómio característico de* A *é*

$$p_A(x) = |A - xI_3| = \begin{vmatrix} 1-x & 2 & 1 \\ 2 & -x & -2 \\ -1 & 2 & 3-x \end{vmatrix} = -x^3 + 4x^2 - 4x = -x(x-2)^2.$$

As raízes características de A *são 0 e 2.*

A respeito do polinómio característico de uma matriz prova-se o resultado seguinte.

Teorema 7.1.18. *Sejam* $r \in \mathbb{N}$, $A = [a_{ij}] \in \mathcal{M}_r(\mathbb{R})$ *e* $p_A(x)$ *o polinómio característico de* A. *Então* $p_A(x)$ *tem grau igual a* r *e se* $p_A = a_r x^r + a_{r-1}x^{r-1} + \ldots + a_0$, *tem-se* $a_r = (-1)^r$, $a_{r-1} = -\sum_{i=1}^{r} a_{ii}$ *e* $a_0 = |A|$.

Demonstração. A prova pode ser realizada por indução sobre r. □

Teorema 7.1.19. *Sejam* $r \in \mathbb{N}$, $A \in \mathcal{M}_r(\mathbb{R})$ *e* $\lambda \in \mathbb{R}$. *São equivalentes as afirmações seguintes:*

1) λ *é valor próprio de* A.

2) λ *é um zero do polinómio característico de* A, *isto é,* $p_A(\lambda) = |A - \lambda I_r| = 0$.

Demonstração. O resultado é imediato pelo Teorema 7.1.15. □

De acordo com o Teorema Fundamental da Álgebra, qualquer equação na variável x da forma

$$a_r x^r + a_{r-1}x^{r-1} + \ldots + a_1 x + a_0 = 0$$

com $a_r \neq 0$, $r \geq 1$, e $a_i \in \mathbb{C}$, $i = 0, 1, \ldots, r$ tem exactamente r raízes em $\mathbb{C}$.

Assim, considerando que os valores próprios de uma matriz $A \in \mathcal{M}_r(\mathbb{R})$ são as raízes características de A, é válido o resultado seguinte.

Teorema 7.1.20. *Sejam* $r \in \mathbb{N}$ *e* $A \in \mathcal{M}_r(\mathbb{R})$. *Então* A *tem, no máximo,* r *valores próprios distintos.*

Demonstração. Resulta do Teorema 7.1.19 e do Teorema Fundamental da Álgebra. □

Definição 7.1.21. *Sejam* $r \in \mathbb{N}$ *e* $A \in \mathcal{M}_r(\mathbb{R})$. *Se* λ *é valor próprio de* A, *designa-se por* ***multiplicidade algébrica de*** λ, *e representa-se por* m.a.(λ), *a multiplicidade de* λ *como raiz do polinómio* $p_A(x)$. *Se* m.a.$(\lambda)=k$, *diz-se que* λ *tem multiplicidade algébrica* k; *no caso particular de* $k = 1$, *diz-se que* λ *é valor próprio simples.*

Atendendo ao Teorema 7.1.7, o problema da determinação dos valores próprios de um endomorfismo f de um subespaço vectorial V de $\mathbb{R}^n$ pode ser reduzido ao problema da determinação dos valores próprios de uma matriz A onde A é a matriz de f em relação a uma dada base de V. Porém, considerando que a matriz de um dado endomorfismo depende da base fixada para a definir, coloca-se a questão se matrizes do mesmo endomorfismo relativas a bases distintas poderão ter polinómios característicos distintos. Seguidamente prova-se que todas as matrizes de um mesmo endomorfismo f têm o mesmo polinómio característico.

Teorema 7.1.22. *Sejam $r \in \mathbb{N}$ e $A, B \in \mathcal{M}_r(\mathbb{R})$.*
Se A e B são semelhantes, então têm o mesmo polinómio característico.

Demonstração. Sejam $r \in \mathbb{N}$ e $A, B \in \mathcal{M}_r(\mathbb{R})$ matrizes semelhantes. Então $B = PAP^{-1}$, para alguma matriz invertível $P \in \mathcal{M}_r(\mathbb{R})$. Logo, atendendo às propriedades de determinantes, tem-se

$$\begin{aligned} p_B(x) &= |B - x\mathrm{I}_r| \\ &= |PAP^{-1} - P(x\mathrm{I}_r)P^{-1}| \\ &= |P(A - x\mathrm{I}_r)P^{-1}| \\ &= |P||A - x\mathrm{I}_r||P^{-1}| \\ &= |A - x\mathrm{I}_r||P||P^{-1}| \\ &= |A - x\mathrm{I}_r| \\ &= p_A(x). \end{aligned}$$

□

Corolário 7.1.23. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $f \in \mathcal{L}(V, V)$, $\mathcal{B}_1$, $\mathcal{B}_2$ bases de V, $A = M(f; \mathcal{B}_1, \mathcal{B}_1)$ e $B = M(f; \mathcal{B}_2, \mathcal{B}_2)$. Então $p_A(x) = p_B(x)$.*

Demonstração. Atendendo aos teoremas 4.5.14 e 4.5.16, A e B são matrizes semelhantes e o resultado segue do teorema anterior. □

Definição 7.1.24. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$ e $f \in \mathcal{L}(V, V)$. Chama-se* ***polinómio característico*** *de f, e representa-se por $p_f(x)$, o polinómio característico de qualquer matriz de f em relação a uma base de V.*

Teorema 7.1.25. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $f \in \mathcal{L}(V, V)$ e $\lambda \in \mathbb{R}$. São equivalentes as afirmações seguintes*

1) λ é valor próprio de f.

2) λ é zero do polinómio característico de f.

Demonstração. Imediato pelo Teorema 7.1.14. □

Teorema 7.1.26. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de $r \geq 1$ e $f \in \mathcal{L}(V, V)$. Então f tem, no máximo, r valores próprios distintos.*

Demonstração. Resulta dos teoremas 7.1.25 e 7.1.20. □

Definição 7.1.27. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $f \in \mathcal{L}(V,V)$. Se λ é valor próprio de f, designa-se por* **multiplicidade algébrica de** λ, *e representa-se por* m.a.(λ), *a multiplicidade de λ como raiz do polinómio $p_f(x)$. Se* m.a.(λ)=k, *diz-se que λ tem multiplicidade algébrica k; no caso particular de $k = 1$, diz-se que λ é valor próprio simples.*

Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão ≥ 1 e $f \in \mathcal{L}(V,V)$. Uma vez conhecidos os valores próprios de f podemos também determinar os vetores próprios de f associados aos valores próprios recorrendo a matrizes.

Teorema 7.1.28. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $\mathcal{B} = (v_1, \ldots, v_r)$ uma base de V, $f \in \mathcal{L}(V,V)$, $A = M(f;\mathcal{B},\mathcal{B})$ e $\lambda \in \mathbb{R}$. Para cada $v \in V$, seja $[v]_{(v_i)}$ o vetor coluna de v na base $\mathcal{B}$. Então*

$$V_{[f,\lambda]} = \{v \in V : (A - \lambda I_r)[v]_{(v_i)} = 0_{r\times 1}\}.$$

Demonstração. Temos $V_\lambda = \mathrm{Nuc}(f - \lambda \mathrm{id}_V)$. Por outro lado, sendo $A = M(f;\mathcal{B},\mathcal{B})$, tem-se $M(f - \lambda\mathrm{id}_V;\mathcal{B},\mathcal{B}) = A - \lambda \mathrm{I}_r$. Assim, uma vez que, para cada $v \in V$,

$$v \in \mathrm{Nuc}(f - \lambda\mathrm{id}_V) \Leftrightarrow (f - \lambda\mathrm{id}_V)v = 0_{\mathbb{R}^n} \Leftrightarrow (A - \lambda I_r)[v]_{(v_i)} = 0_{n\times 1},$$

segue que

$$V_{[f,\lambda]} = \{v \in V : (A - \lambda I_r)[v]_{(v_i)} = 0_{r\times 1}\}. \qquad \square$$

Nas condições do teorema anterior, os vetores próprios de f associados a λ são os vetores de V cujos vetores coluna na base $\mathcal{B}$ são as soluções não nulas do sistema homogéneo $(A - \lambda I_r)x = 0_{r\times 1}$.

Exemplo 7.1.29. *Consideremos o espaço vetorial real $\mathbb{R}^3$ e a base $\mathcal{B} = (v_1, v_2, v_3)$ de $\mathbb{R}^3$, onde $v_1 = (1,0,0)$,$v_2 = (1,1,0)$, $v_3 = (1,1,1)$. Seja $f : \mathbb{R}^3 \to \mathbb{R}^3$ o endomorfismo definido por*

$$f(v_1) = (0,-2,-1), \quad f(v_2) = (0,0,-1), \quad f(v_3) = (0,0,1).$$

Recorrendo à matriz de f em relação à base $\mathcal{B}$, vamos determinar os valores próprios e os vetores próprios de f. Sendo $A = M(f;\mathcal{B},\mathcal{B})$, tem-se

$$A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

Dado $\lambda \in \mathbb{R}$, λ é valor próprio de f se e só se $|A - \lambda \mathrm{I}_3| = 0$. Uma vez que

$$\begin{aligned} |A - \lambda\mathrm{I}_3| &= \begin{vmatrix} 2-\lambda & 0 & 0 \\ -1 & 1-\lambda & -1 \\ -1 & -1 & 1-\lambda \end{vmatrix} \\ &= (2-\lambda)\begin{vmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{vmatrix} \\ &= (2-\lambda)[(1-\lambda)^2 - 1] \end{aligned}$$

então $|A - \lambda\mathrm{I}_3| = 0$ se e só se $\lambda = 2$ ou $\lambda = 0$, i.e., os valores próprios de f são 0 *e* 2.

Relativamente ao subespaço próprio de f associado ao valor próprio 0, tem-se o seguinte:

$$\begin{aligned}
\mathbb{R}^3_{[f,0]} &= \mathrm{Nuc}(f - 0\mathrm{id}_{\mathbb{R}^3}) \\
&= \left\{ \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \in \mathbb{R}^3 : M(f - 0\mathrm{id}_{\mathbb{R}^3}; \mathcal{B}, \mathcal{B}) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0 \right\} \\
&= \left\{ \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \in \mathbb{R}^3 : (A - 0I_3) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0 \right\}.
\end{aligned}$$

Resolvendo o sistema $(A - 0\mathrm{I}_3)x = 0$, i.e., $Ax = 0$,

$$\begin{aligned}
[A|0] &= \left[\begin{array}{rrr|r} 2 & 0 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ -1 & -1 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right] \\
&\longrightarrow \left[\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]
\end{aligned}$$

temos

$$\begin{aligned}
\mathbb{R}^3_{[f,0]} &= \{\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \in \mathbb{R}^3 : \alpha_1 = 0, \alpha_2 - \alpha_3 = 0\} \\
&= \{\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \in \mathbb{R}^3 : \alpha_1 = 0, \alpha_2 = \alpha_3\} \\
&= \{0 v_1 + \alpha_2 v_2 + \alpha_2 v_3 \in \mathbb{R}^3 : \alpha_2 \in \mathbb{R}\} \\
&= \{\alpha_2 (v_2 + v_3) \in \mathbb{R}^3 : \alpha_2 \in \mathbb{R}\} \\
&= \{\alpha_2 (2,2,1) \in \mathbb{R}^3 : \alpha_2 \in \mathbb{R}\} \\
&= < (2,2,1) > .
\end{aligned}$$

Os vetores próprios de f associados ao valor próprio 0 são os elementos de $\mathbb{R}^3_{[f,0]} \setminus \{0_{\mathbb{R}^3}\}$.

Vamos, agora, determinar o subespaço próprio associado ao valor próprio 2:

$$\begin{aligned}
\mathbb{R}^3_{[f,2]} &= \mathrm{Nuc}(f - 2\mathrm{id}_{\mathbb{R}^3}) \\
&= \left\{ \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \in \mathbb{R}^3 : M(f - 2\mathrm{id}_{\mathbb{R}^3}; \mathcal{B}, \mathcal{B}) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0 \right\} \\
&= \left\{ \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \in \mathbb{R}^3 : (A - 2\mathrm{I}_3) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0 \right\}.
\end{aligned}$$

Resolvendo o sistema $(A - 2\mathrm{I}_3)x = 0$,

$$[A - 2\mathrm{I}_3|0] = \left[\begin{array}{rrr|r} 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 \\ -1 & -1 & -1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

temos

$$\begin{aligned}\mathbb{R}^3_{[f,2]} &= \{\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \in \mathbb{R}^3 : \alpha_1 + \alpha_2 + \alpha_3 = 0\} \\ &= \{\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \in \mathbb{R}^3 : \alpha_1 = -\alpha_2 - \alpha_3\} \\ &= \{(-\alpha_2 - \alpha_3) v_1 + \alpha_2 v_2 + \alpha_3 v_3 \in \mathbb{R}^3 : \alpha_2, \alpha_3 \in \mathbb{R}\} \\ &= \{\alpha_2(-v_1 + v_2) + \alpha_3(-v_1 + v_3)\mathbb{R}^3 : \alpha_2, \alpha_3 \in \mathbb{R}\} \\ &= \{\alpha_2((0,1,0)) + \alpha_3((0,1,1))\mathbb{R}^3 : \alpha_2, \alpha_3 \in \mathbb{R}\} \\ &= < (0,1,0), (0,1,1) > .\end{aligned}$$

Os vetores próprios de f associados ao valor próprio 2 são os elementos de $\mathbb{R}^3_{[f,2]} \setminus \{0_{\mathbb{R}^3}\}$.

Embora o resultado anterior nos dê um processo para determinar o subespaço próprio de um endomorfismo associado a um determinado valor próprio, o resultado seguinte permite-nos determinar a multiplicidade geométrica de um valor próprio sem determinar o subespaço próprio que lhe está associado.

Teorema 7.1.30. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $f \in \mathcal{L}(V,V)$, $\mathcal{B}$ uma base de V, $A = M(f;\mathcal{B},\mathcal{B})$ e $\lambda \in \mathbb{R}$ um valor próprio de f. Então*

$$\text{m.g.}(\lambda) = r - car(A - \lambda I_r).$$

Demonstração. Tem-se $V_{[f,\lambda]} = \text{Nuc}(f - \lambda id_V)$. Então, considerando que

$$\dim V = \dim \text{Nuc}(f - \lambda id_V) + \dim Im(f - \lambda id_V)$$

e

$$\dim Im(f - \lambda id_V) = car(A - \lambda I_r),$$

segue que

$$m.g.(\lambda) = \dim V_{[f,\lambda]} = r - car(A - \lambda I_r).$$

□

Um resultado similar ao anterior pode ser estabelecido para matrizes quadradas.

Corolário 7.1.31. *Sejam $r \in \mathbb{N}$ e $A \in \mathcal{M}_r(\mathbb{R})$ e $\lambda \in \mathbb{R}$ um valor próprio de A. Então*

$$\text{m.g.}(\lambda) = r - car(A - \lambda I_r).$$

Demonstração. Sejam $\mathcal{B}$ uma base de $\mathbb{R}^r$ e f o endomorfismo de $\mathbb{R}^r$ tal que $A = M(f;\mathcal{B},\mathcal{B})$. Pelo Teorema 7.1.7, λ é valor próprio de A se e só se é valor próprio de f. Então, considerando que, pelo Teorema 7.1.13, $\dim(\mathbb{R}^r)_{[f,\lambda]} = \dim M_{[A,\lambda]}$, o resultado segue do Teorema 7.1.30. □

Conhecida a multiplicidade algébrica de cada valor próprio de um endomorfismo f, podemos também obter alguma informação sobre a multiplicidade geométrica do valor próprio, uma vez que estas mutliplicidades estão relacionadas da forma estabelecida no resultado seguinte.

Teorema 7.1.32. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$ e $f \in \mathcal{L}(V,V)$. Se $\lambda \in \mathbb{R}$ é valor próprio de f, então*

$$\mathrm{m.g}(\lambda) \leq \mathrm{m.a}(\lambda).$$

Demonstração. Seja λ um valor próprio de f e seja $k = \mathrm{m.g.}(\lambda)$, i.e., $k = \dim V_{[f,\lambda]}$. Sejam $(w_1, \dots, w_k)$ uma base de $V_{[f,\lambda]}$ e $\mathcal{B} = (w_1, \dots, w_k, u_1, \dots, u_{r-k})$ uma base de V. Então

$$A = M(f; \mathcal{B}, \mathcal{B}) = \begin{bmatrix} B & C \\ \mathbf{0} & D \end{bmatrix}$$

onde $B = \mathrm{diag}(\lambda, \dots, \lambda) \in \mathcal{M}_k(\mathbb{R})$, $\mathbf{0} \in \mathcal{M}_{(r-k)\times k}(\mathbb{R})$, $C \in \mathcal{M}_{k\times(r-k)}(\mathbb{R})$ e $D \in \mathcal{M}_{(r-k)}(\mathbb{R})$. Tendo em conta o Teorema de Laplace, segue que

$$\begin{aligned} p_f(x) &= |A - x\mathrm{I}_r| = \begin{vmatrix} B - x\mathrm{I}_k & C \\ 0 & D - x\mathrm{I}_{r-k} \end{vmatrix} = \begin{vmatrix} \lambda - x & \dots & 0 \\ \vdots & \dots & \vdots \\ \mathbf{0} & \dots & \lambda - x \end{vmatrix} |D - x\mathrm{I}_{r-k}| \\ &= (-1)^k (x-\lambda)^k |D - x\mathrm{I}_{r-k}| \end{aligned}$$

Portanto, $\mathrm{m.a.}(\lambda) \geq k = \mathrm{m.g.}(\lambda)$. □

Corolário 7.1.33. *Sejam $r \in \mathbb{N}$ e $A \in \mathcal{M}_r(\mathbb{R})$.*
Se $\lambda \in \mathbb{R}$ é valor próprio de A, então $\mathrm{m.g.}(\lambda) \leq \mathrm{m.a.}(\lambda)$.

Demonstração. O resultado é imediato, considerando o Teorema 7.1.7 e aplicando o teorema anterior ao endomorfismo f de $\mathbb{R}^r$ cuja matriz em relação a uma base $\mathcal{B}$ de $\mathbb{R}^r$ é A. □

Terminamos esta secção com o estudo de mais algumas propriedades a respeito de valores próprios de endomorfismos de espaços vetoriais de dimensão finita e de valores próprios de matrizes quadradas.

Teorema 7.1.34. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $f \in \mathcal{L}(V,V)$ e $\lambda_1, \dots, \lambda_p$ valores próprios de f, distintos dois a dois. Se, para cada $i \in \{1, \dots, p\}$, $(v_{i1}, \dots, v_{is_i})$ é uma sequência linearmente independente formada por vetores próprios de f associados ao valor próprio λ_i, então a sequência*

$$(v_{11}, \dots, v_{1s_1}, \dots, v_{p1}, \dots v_{ps_p})$$

é linearmente independente.

Demonstração. Dado $p \in \mathbb{N}$, representemos por $P(p)$ a afirmação seguinte: *Se $\lambda_1, \dots, \lambda_p$ são valores próprios de f, distintos dois a dois e, para cada $i \in \{1, \dots, p\}$, $(v_{i1}, \dots, v_{is_i})$ é uma sequência linearmente independente formada por vetores próprios de f associados ao valor próprio λ_i, então a sequência*

$$(v_{11}, \dots, v_{1s_1}, \dots, v_{p1}, \dots v_{ps_p})$$

é linearmente independente.

Por indução matemática, mostremos que, para todo $p \in \mathbb{N}$, $P(p)$ é verdadeira.

Se $p = 1$, $P(p)$ é verdadeira, uma vez que $(v_{11}, \ldots, v_{1s_1})$ é linearmente independente.

No sentido de provar o passo indutivo, admitamos, que $P(k)$ é verdadeira. Com base nesta hipótese prova-se que $P(k+1)$ é verdadeira, ou seja, mostra-se que se $\lambda_1, \ldots, \lambda_k, \lambda_{k+1}$ são vetores próprios de f, distintos dois a dois e, para cada $i \in \{1, \ldots k+1\}$, $(v_{i1}, \ldots, v_{is_i})$ é uma sequência linearmente independente formada por vetores próprios de f associados ao valor próprio λ_i, então a sequência

$$(v_{11}, \ldots, v_{1s_1}, \ldots, v_{k1}, \ldots v_{ks_k}, v_{k+1\,1}, \ldots v_{k+1\,s_{k+1}})$$

é linearmente independente.
De facto, se consideramos escalares $\alpha_{ij} \in \mathbb{R}$, $i = 1, \ldots, k+1$, $j = 1, \ldots, s_i$, tais que

$$\sum_{i=1}^{k+1} \alpha_{i1} v_{i1} + \ldots + \alpha_{is_i} v_{is_i} = 0_{\mathbb{R}^n} \quad (1)$$

e aplicarmos f a ambos os lados da igualdade (1), obtemos

$$\sum_{i=1}^{k+1} \alpha_{i1} \lambda_i v_{i1} + \ldots + \alpha_{is_i} \lambda_i v_{is_i} = 0_{\mathbb{R}^n} \quad (2)$$

Por conseguinte, de (1) e (2), considerando $(1) - \lambda_{k+1}(2)$, segue que

$$\sum_{i=1}^{k} (\lambda_i - \lambda_{k+1}) \alpha_{i1} v_{i1} + \ldots + \alpha_{is_i} (\lambda_i - \lambda_{k+1}) v_{is_i} = 0_{\mathbb{R}^n}.$$

Como, por hipótese de indução, a sequência

$$(v_{11}, \ldots, v_{1s_1}, \ldots, v_{k1}, \ldots v_{ks_k})$$

é linearmente independente, resulta que, para todo $i \in \{1, \ldots, k\}$, $\alpha_i = 0$. Logo, por (1), obtem-se $\alpha_{k+1\,1} v_{k+1\,1} + \ldots + \alpha_{k+1\,s_{k+1}} v_{k+1 s_{k+1}} = 0_{\mathbb{R}^n}$ e, uma vez que a sequência $(v_{k+1\,1}, \ldots v_{k+1\,s_{k+1}})$ é linearmente independente, conclui-se que $\alpha_{k+1} = 0$. Por conseguinte, a sequência

$$(v_{11}, \ldots, v_{1s_1}, \ldots, v_{k1}, \ldots v_{ks_k}, v_{k+1\,1}, \ldots v_{k+1\,s_{k+1}})$$

é linearmente independente. □

Corolário 7.1.35. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$ e $f \in \mathcal{L}(V, V)$. Se $v_1, \ldots, v_m$ são vetores próprios de f associados, respetivamente, a valores próprios $\lambda_1, \ldots, \lambda_m$ distintos dois a dois, então a sequência $(v_1, \ldots, v_m)$ é linearmente independente.*

Demonstração. Imediato pelo teorema anterior, uma vez que cada vetor próprio v de f é não nulo e, portanto, (v) é linearmente independente. □

Teorema 7.1.36. *Sejam $r \in \mathbb{N}$, $A \in \mathcal{M}_r(\mathbb{R})$ e $\lambda_1, \ldots, \lambda_p$ valores próprios de A, distintos dois a dois. Se, para cada $i \in \{1, \ldots, p\}$, $(y_{i1}, \ldots, y_{is_i})$ é uma sequência linearmente independente formada por vetores próprios de A associados ao valor próprio λ_i, então a sequência*

$$(y_{11}, \ldots, y_{1s_1}, \ldots, y_{p1}, \ldots y_{ps_p})$$

é linearmente independente.

Demonstração. Sejam $\mathcal{B}$ uma base de $\mathbb{R}^r$ e $f \in \mathcal{L}(\mathbb{R}^r, \mathbb{R}^r)$ tal que $A = M(f; \mathcal{B}, \mathcal{B})$. Para cada $i \in \{1, \ldots p\}$ e para cada $j \in \{1, \ldots, s_i\}$, seja (v_{ij}) o vetor de $\mathbb{R}^r$ tal que $[v_{ij}]_{\mathcal{B}} = y_{ij}$. Se $(y_{i1}, \ldots, y_{is_i})$ é uma sequência linearmente independente, a sequência $(v_{i1}, \ldots, v_{is_i})$ de vetores de $\mathbb{R}^r$ também é linearmente independente. Logo, pelo Teorema 7.1.7, $(v_{i1}, \ldots, v_{is_i})$ é uma sequência linearmente independente formada por vetores próprios de f. Então, pelo teorema anterior,

$$(v_{11}, \ldots, v_{1s_1}, \ldots, v_{k1}, \ldots, v_{ks_k}, v_{k+1\,1}, \ldots v_{k+1\,s_{k+1}})$$

é linearmente independente e, por conseguinte, a sequência

$$(y_{11}, \ldots, y_{1s_1}, \ldots, y_{k1}, \ldots, y_{ks_k}, y_{k+1\,1}, \ldots y_{k+1\,s_{k+1}})$$

é linearmente independente. □

Corolário 7.1.37. *Sejam $r \in \mathbb{N}$, $A \in \mathcal{M}_r(\mathbb{R})$. Se $y_1, \ldots, y_m$ são vetores próprios de A associados, respetivamente, a valores próprios $\lambda_1, \ldots, \lambda_m$ distintos dois a dois, então a sequência $(y_1, \ldots, y_m)$ é linearmente independente.*

Teorema 7.1.38. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $\mathcal{B}$ uma base de V, $f \in \mathcal{L}(V, V)$, $A = M(f; \mathcal{B}, \mathcal{B})$ e $\lambda \in \mathbb{R}$.*
São equivalentes as afirmações seguintes:

1) *λ é valor próprio de f.*

2) *$f - \lambda \mathrm{id}_V$ não é automorfismo de V.*

3) *$A - \lambda \mathrm{I}_r$ não é invertível.*

Demonstração. Sejam $f \in \mathcal{L}(V, V)$, $\mathcal{B}$ uma base de V, $A = M(f; \mathcal{B}, \mathcal{B})$ e $\lambda \in \mathbb{R}$.

1) $\Rightarrow$ 2) Se λ é valor próprio de f, tem-se $V_{[f,\lambda]} \neq \{0_{\mathbb{R}^n}\}$, ou seja $\mathrm{Nuc}(f - \lambda \mathrm{id}_V) \neq \{0_{\mathbb{R}^n}\}$. Logo, $f - \lambda \mathrm{id}_V$ não é injetiva e, portanto, não é um automorfismo de V.

2) $\Rightarrow$ 1) Uma vez que V tem dimensão finita, se $f - \lambda \mathrm{id}_V$ não é automorfismo de V, então $f - \lambda \mathrm{id}_V$ não é injetiva. Logo, $\mathrm{Nuc}(f - \lambda \mathrm{id}_V) \neq \{0_{\mathbb{R}^n}\}$, i.e., $V_{[f,\lambda]} \neq \{0_{\mathbb{R}^n}\}$. Portanto, λ é valor próprio de f.

2) $\Leftrightarrow$ 3) Uma vez que $A - \lambda \mathrm{I}_r = M(f - \lambda \mathrm{id}_V; \mathcal{B}, \mathcal{B})$, o resultado é imediato pelo Teorema 4.5.12. □

Corolário 7.1.39. *Sejam $r \in \mathbb{N}$, $A \in \mathcal{M}_r(\mathbb{R})$ e $\lambda \in \mathbb{R}$.*
São equivalentes as afirmações seguintes:

1) λ é valor próprio de A.

2) $A - \lambda \mathrm{I}_r$ não é invertível.

Demonstração. Considerando $\mathcal{B}$ uma base de $\mathbb{R}^r$ e $A = M(f; \mathcal{B}, \mathcal{B})$, o resultado é imediato pelos teoremas 7.1.7 e 7.1.38. □

Teorema 7.1.40. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $f \in \mathcal{L}(V, V)$, $\lambda \in \mathbb{R}$ um valor próprio de f e $v \in V$ um vetor próprio de f associado ao valor próprio λ. Então:*

1) Para qualquer $\alpha \in \mathbb{R}$, $\alpha\lambda$ é valor próprio de αf e v é vetor próprio de αf associado a $\alpha\lambda$;

2) Para qualquer $\alpha \in \mathbb{R}$, $\lambda - \alpha$ é valor próprio de $f - \alpha id_V$ e v é vetor próprio de $f - \alpha id_V$ associado a $\lambda - \alpha$;

3) λ^2 é valor próprio de f^2 e v é vetor próprio de f^2 associado a λ^2;

4) f é um automorfismo se e só se $\lambda \neq 0$. Se f é um automorfismo, então λ^{-1} é valor próprio de f^{-1} e v é vetor próprio de f^{-1} associado a λ^{-1}.

Demonstração. 1) Considerando que $v \in V$ é um vetor próprio de f associado ao valor próprio λ segue que

$$(\alpha f)(v) = \alpha(f(v)) = \alpha(\lambda v) = (\alpha\lambda)v,$$

e, portanto, $\alpha\lambda$ é valor próprio de αf e v é vetor próprio de αf associado a $\alpha\lambda$.

2) Sendo $v \in V$ um vetor próprio de f associado ao valor próprio λ, temos

$$(f - \alpha id_V)(v) = f(v) - (\alpha id_v)(v) = f(v) - \alpha(id_V(v)) = \lambda v - \alpha v = (\lambda - \alpha)(v).$$

Logo, $\lambda - \alpha$ é valor próprio de $f - \alpha id_V$ e v é vetor próprio de $f - \alpha id_V$ associado a $\lambda - \alpha$.

3) Atendendo a que $v \in V$ é um vetor próprio de f associado ao valor próprio λ, temos,

$$f^2(v) = f(f(v)) = f(\lambda v) = \lambda f(v) = \lambda(\lambda v) = \lambda^2 v.$$

Assim, $v \in V$ é um vetor próprio de f associado ao valor próprio λ.

4) Pelo Teorema 7.1.38, é imediato que f é um automorfismo se e só se $\lambda \neq 0$. Assumindo que f é um automorfismo, prova-se que λ^{-1} é valor próprio de f^{-1} e v é vetor próprio de f^{-1} associado a λ^{-1}. No sentido de fazer esta prova, consideremos $\mathcal{B}$ uma base de V e $A = M(f; \mathcal{B}, \mathcal{B})$. Uma vez que f é um automorfismo, pelo Teorema 4.5.14, A é invertível. Então, considerando que $\lambda \neq 0$, temos

$$\begin{aligned} A[v]_{\mathcal{B}} = \lambda[v]_{\mathcal{B}} &\implies A^{-1}\left(A[v]_{\mathcal{B}}\right) = A^{-1}\left(\lambda[v]_{\mathcal{B}}\right) \\ &\iff \left(A^{-1}A\right)[v]_{\mathcal{B}} = \lambda\left(A^{-1}[v]_{\mathcal{B}}\right) \\ &\iff I_r[v]_{\mathcal{B}} = \lambda\left(A^{-1}[v]_{\mathcal{B}}\right) \\ &\iff [v]_{\mathcal{B}} = \lambda\left(A^{-1}[v]_{\mathcal{B}}\right) \\ &\iff \lambda^{-1}[v]_{\mathcal{B}} = A^{-1}[v]_{\mathcal{B}}, \end{aligned}$$

e, portanto, $[v]_{\mathcal{B}}$ é vetor próprio de A^{-1} associado ao valor próprio λ^{-1}. Assim, considerando o Teorema 7.1.7 e que $M(f^{-1};\mathcal{B},\mathcal{B}) = A^{-1}$, concluimos que λ^{-1} é valor próprio de f^{-1} e v é vetor próprio de f^{-1} associado a λ^{-1}. $\square$

O resultado correspondente ao teorema anterior para matrizes quadradas estabelece o seguinte.

Corolário 7.1.41. *Sejam $r \in \mathbb{N}$, $A \in \mathcal{M}_r(\mathbb{R})$, $\lambda \in \mathbb{R}$ um valor próprio de A, e $y \in \mathcal{M}_{r\times 1}(\mathbb{R})$ um vetor próprio de A associado ao valor próprio λ. Então:*

1) Para qualquer $\alpha \in \mathbb{R}$, $\alpha\lambda$ é valor próprio de αA e y é vetor próprio de αA associado a $\alpha\lambda$;

2) Para qualquer $\alpha \in \mathbb{R}$, $\lambda - \alpha$ é valor próprio de $A - \alpha I_r$ e y é vetor próprio de $A - \alpha I_r$ associado a $\lambda - \alpha$;

3) λ^2 é valor próprio de A^2 e y é vetor próprio de A^2 associado a λ^2;

4) A é invertível se e só se $\lambda \neq 0$. Se A é invertível, então λ^{-1} é valor próprio de A^{-1} e y é vetor próprio de A^{-1} associado a λ^{-1}.

Demonstração. Considerando $\mathcal{B}$ uma base de $\mathbb{R}^r$ e $A = M(f;\mathcal{B},\mathcal{B})$, o resultado é imediato pelos teoremas 7.1.7 e 7.1.40. $\square$

Teorema 7.1.42. *Sejam $A \in \mathcal{M}_r(\mathbb{R})$ e $\lambda \in \mathbb{R}$. Então, λ é valor próprio de A se e só se λ é valor próprio de A^T.*

Demonstração. Imediato tendo em conta que $(A - \lambda I_r)^T = A^T - \lambda I_r$ e

$$\det(A - \lambda I_r)^T = \det(A - \lambda I_r)\,.$$

$\square$

Teorema 7.1.43. *Seja A uma matriz triangular de ordem r. Então os valores próprios de A são os elementos da sua diagonal principal.*

Demonstração. Se $A = [a_{ij}]_r$ é uma matriz triangular, então $A - \lambda I_r$ é também uma matriz triangular, pelo que

$$\det(A - \lambda I_r) = (a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{rr} - \lambda)\,.$$

Logo, os valores próprios de A são os elementos $a_{11}, a_{22}, ..., a_{rr}$. $\square$

7.2 Diagonalização

As matrizes diagonais, para muitos propósitos, são os tipos mais simples de matrizes com as quais podemos trabalhar. No que respeita ao estudo de endomorfismos já vimos que um endomorfismo f de um subespaço vetorial V de $\mathbb{R}^n$ pode ser estudado através de qualquer matriz que o represente, mas há vantagem em considerar matrizes diagonais. Nesta secção, determinamos condições sob as quais um endomorfismo pode ser representado por uma matriz diagonal.

Definição 7.2.1. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$ e $f \in \mathcal{L}(V, V)$. Diz-se que f é **diagonalizável** se existe uma base de V em relação à qual a matriz de f é diagonal.*

Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$. Dado um endomorfismo f de V, é evidente que f é representável por uma matriz diagonal

$$\begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda_r \end{bmatrix}$$

se e só se existe uma base $(v_1, v_2, \ldots, v_r)$ de V tal que $f(v_i) = \lambda_i v_i$, para qualquer $i \in \{1, \ldots, r\}$, ou seja, se e só se existir uma base de V formada por vetores próprios de f. Podemos, então, estabelecer o resultado seguinte.

Na sequência do observado anteriormente e considerando o Teorema 4.5.17., podemos estabelecer o resultado seguinte.

Teorema 7.2.2. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$ e $f \in \mathcal{L}(V, V)$. Então f é diagonalizável se e só se existe uma base de $\mathcal{B}'$ de V formada por vetores próprios de f.*

Corolário 7.2.3. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$ e $f \in \mathcal{L}(V, V)$. Então f é diagonalizável se e só se f tem r vetores próprios linearmente independentes.*

Observação: Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$ e $f \in \mathcal{L}(V, V)$. Se f é diagonalizável, então, pelo Teorema 7.2.2, existe uma base de $\mathcal{B}'$ de V formada por vetores próprios de f. Neste caso, $M(f; \mathcal{B}', \mathcal{B}')$ é uma matriz diagonal e os seus elementos principais são valores próprios de f. Além disso, para qualquer base $\mathcal{B}$ de V, tem-se

$$M(f; \mathcal{B}, \mathcal{B}) = M(id_V; \mathcal{B}', \mathcal{B}) M(f; \mathcal{B}', \mathcal{B}') M(id_V; \mathcal{B}', \mathcal{B})^{-1}.$$

Exemplo 7.2.4. *Consideremos o espaço vetorial real $\mathbb{R}^3$ e a base $\mathcal{B} = (v_1, v_2, v_3)$ de $\mathbb{R}^3$, onde $v_1 = (1, 0, 0)$,$v_2 = (1, 1, 0)$, $v_3 = (1, 1, 1)$. Seja $f : \mathbb{R}^3 \to \mathbb{R}^3$ o endomorfismo definido no exemplo 7.1.29 por*

$$f(v_1) = (0, -2, -1), \quad f(v_2) = (0, 0, -1), \quad f(v_3) = (0, 0, 1).$$

Este endomorfismo é diagonalizável, uma vez que $\mathcal{B}' = ((2, 2, 1), (0, 1, 0), (0, 1, 1))$ é uma base de $\mathbb{R}^3$ formada por vetores próprios de f; $(2, 2, 1)$ é um vetor próprio de f associado ao valor próprio 0 e $(0, 1, 0)$, $(0, 1, 1)$ são vetores próprios de f associados ao valor próprio 2. Logo $M(f; \mathcal{B}', \mathcal{B}') = \mathrm{diag}(0, 2, 2)$.

Do Corolário 7.1.35 resulta uma condição suficiente para que um endomorfismo f seja diagonalizável.

Teorema 7.2.5. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$ e $f \in \mathcal{L}(V, V)$. Se f admite r valores próprios distintos, então f é diagonalizável.*

Demonstração. Suponha-se que f admite r valores próprios $\lambda_1, \ldots, \lambda_r$ distintos dois a dois. Para $i \in \{1, \ldots, r\}$, seja $v_i \in V$ um vetor próprio de f associado a λ_i. Pelo Teorema 7.1.35, a sequência de vetores $(v_1, \ldots, v_r)$ é linearmente independente. Logo $(v_1, \ldots, v_r)$ é uma base de V, pois $\dim V = r$. Assim, V admite uma base formada por vetores próprios de f e, portanto, f é diagonalizável. □

Note-se que a condição estabelecida no teorema anterior é apenas uma condição suficiente, não sendo necessária.

Exemplo 7.2.6. *No exemplo anterior tem-se um endomorfismo f dum espaço vetorial de dimensão 3 com apenas dois valores próprios distintos e, no entanto, f é diagonalizável.*

Seguidamente estudamos mais algumas condições que permitem a caracterização de endomorfismos diagonalizáveis, sendo estas condições estabelecidas com base nas multiplicidades geométricas e algébricas dos valores próprios de um endomorfismo.

Teorema 7.2.7. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $f \in \mathcal{L}(V, V)$ e $\lambda_1, \ldots, \lambda_p$ os valores próprios de f, distintos dois a dois. Então f é diagonalizável se e só se $\sum_{i=1}^{p} m.g.(\lambda_i) = r$.*

Demonstração. Sejam $\lambda_1, \ldots, \lambda_p$ os valores próprios de f (distintos dois a dois) e, para cada $i \in \{1, \ldots, p\}$, sejam $s_i = \text{m.g.}(\lambda_i)$ e $\mathcal{B}_i = (v_{i1}, \ldots, v_{is_i})$ uma base de $V_{[f,\lambda_i]}$. Pelo Teorema 7.1.34, a sequência

$$\mathcal{B} = (v_{11}, \ldots, v_{1s_1}, \ldots, v_{p1}, \ldots, v_{ps_p})$$

é linearmente independente. Então, se admitirmos que $\sum_{i=1}^{p} \text{m.g.}(\lambda_i) = r$, a sequência $\mathcal{B}$ é uma base de V, pois $\dim V = r$. Logo, pelo Teorema 7.2.2, f é diagonalizável. Reciprocamente, admitamos que f é diagonalizável. Então, pelo Teorema 7.2.2, existe uma base de V formada por vetores próprios de f. Por conseguinte, $\sum_{i=1}^{p} \text{m.g.}(\lambda_i) \geq r$. Por outro lado, como para cada $i \in \{1 \ldots p\}$, $\text{m.g.}(\lambda_i) \leq \text{m.a.}(\lambda_i)$ e $\sum_{i=1}^{p} \text{m.a.}(\lambda_i) \leq \text{grau}(p_f) = r$, conclui-se que $\sum_{i=1}^{p} \text{m.g.}(\lambda_i) \leq r$. Logo, $\sum_{i=1}^{p} \text{m.g.}(\lambda_i) = r$. □

Corolário 7.2.8. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $f \in \mathcal{L}(V, V)$ e $\lambda_1, \ldots, \lambda_p$ os valores próprios de f, distintos dois a dois. Então f é diagonalizável se e só se $\sum_{i=1}^{p} m.a.(\lambda_i) = r$ e, para cada $i \in \{1, \ldots, r\}$,* $\text{m.g.}(\lambda_i) = \text{m.a.}(\lambda_i)$.

Demonstração. Sejam $\lambda_1, \ldots, \lambda_p$ os valores próprios de f, distintos dois a dois.
Admitamos que f é diagonalizável. Então, pelo teorema anterior, segue que $\sum_{i=1}^{p} \text{m.g.}(\lambda_i) = r$. Agora, como $\sum_{i=1}^{p} \text{m.a.}(\lambda_i) \leq \text{grau}(p_f) = r$ e, para cada $i \in \{1, \ldots, p\}$, temos $\text{m.g.}(\lambda_i) \leq \text{m.a.}(\lambda_i)$, resulta que $\sum_{i=1}^{p} \text{m.a.}(\lambda_i) = r$ e $\text{m.g.}(\lambda_i) = \text{m.a.}(\lambda_i)$.
Reciprocamente, admitamos que $\sum_{i=1}^{p} \text{m.a.}(\lambda_i) = r$ e que, para cada $i \in \{1, \ldots, r\}$, $\text{m.g.}(\lambda_i) = \text{m.a.}(\lambda_i)$. Então, $\sum_{i=1}^{p} \text{m.g.}(\lambda_i) = r$ e, pelo teorema anterior, f é diagonalizável. □

Considerando que toda a matriz quadrada é matriz de um enfomorfismo de um subespaço vetorial V de $\mathbb{R}^n$, em relação a uma certa base de V, a definição de endomorfismo diagonalizável motiva a definição seguinte.

Definição 7.2.9. *Sejam $r \geq 1$ e $A \in \mathcal{M}_r(\mathbb{R})$. Diz-se que A é* ***diagonalizável*** *se A é semelhante a uma matriz diagonal.*

Teorema 7.2.10. *Sejam V um subespaço vectorial de $\mathbb{R}^n$ de dimensão $r \geq 1$, $\mathcal{B}$ uma base de V, $f \in \mathcal{L}(V, V)$ e $A = M(f; \mathcal{B}, \mathcal{B})$. Então f é diagonalizável se e só se A é diagonalizável.*

Demonstração. Atendendo a que matrizes do endomorfismo f em relação a bases diferentes são semelhantes, é imediato que f é diagonalizável se e só se A é diagonalizável. □

Teorema 7.2.11. *Sejam $r \in \mathbb{N}$ e $A \in \mathcal{M}_r(\mathbb{R})$. Então A é diagonalizável se e só se existe uma base de $\mathcal{M}_{r\times 1}(\mathbb{R})$ formada por vetores próprios de A.*

Demonstração. Sejam $\mathcal{B}$ uma base de $\mathbb{R}^r$ e f o endomorfismo de $\mathbb{R}^r$ cuja matriz em relação à base $\mathcal{B}$ é A. Então o resultado é imediato pelos teoremas 7.1.7 e 7.1.19 e considerando que uma sequência $(v_1, \ldots, v_r)$ de vetores de V é uma base de V se e só se $([v_1]_{\mathcal{B}}, \ldots, [v_r]_{\mathcal{B}})$ é uma base de $\mathcal{M}_{r\times 1}(\mathbb{R})$. □

Corolário 7.2.12. *Sejam $r \in \mathbb{N}$ e $A \in \mathcal{M}_r(\mathbb{R})$. Então A é diagonalizável se e só se A tem r vetores próprios linearmente independentes.*

Observação: Sejam $r \in \mathbb{N}$ e $A \in \mathcal{M}_r(\mathbb{R})$. Se A é diagonalizável, então existe uma base de $\mathcal{M}_r(\mathbb{R})$ formada por vetores próprios de A. Se $p_1, \ldots, p_r \in \mathcal{M}_{r\times 1}(\mathbb{R})$ são vetores próprios de A, associados aos valores próprios $\lambda_1, \ldots, \lambda_r$, respetivamente, e $(p_1, \ldots, p_r)$ é uma base de $\mathcal{M}_{r\times 1}(\mathbb{R})$, tem-se

$$A = P \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda_r \end{bmatrix} P^{-1}$$

onde, para cada $j \in \{1, \ldots, r\}$, a coluna j de P é p_j.

De facto, se $\mathcal{B} = (v_1, \ldots, v_r)$ é uma base de $\mathbb{R}^r$ e f é o endomorfismo de $\mathbb{R}^r$ cuja matriz em relação à base $\mathcal{B}$ é A, então f é diagonalizável. Além disso, se, para cada $j \in \{1, \ldots, r\}$, p_j é um vetor próprio de A associado ao valor próprio λ_j, então o vetor $v'_j \in V$, cujo vetor coluna relativamente à base $\mathcal{B}$ é p_j, é um vetor próprio de f associado ao valor próprio λ_j. Assim, sendo $(p_1, \ldots, p_r)$ uma base de $\mathcal{M}_{r\times 1}(\mathbb{R})$ formada por vetores próprios de f associados aos valores próprios $\lambda_1, \ldots, \lambda_r$, a sequência $\mathcal{B}' = (v'_1, \ldots, v'_r)$ é uma base de V formada por vetores próprios de f e tem-se

$$A = M(f; \mathcal{B}, \mathcal{B}) = M(id_{\mathbb{R}^r}; \mathcal{B}', \mathcal{B})M(f; \mathcal{B}', \mathcal{B}')M(id_{\mathbb{R}^r}; \mathcal{B}', \mathcal{B})^{-1},$$

onde

$$M(f; \mathcal{B}', \mathcal{B}') = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda_r \end{bmatrix}$$

e $M(id_{\mathbb{R}^r}; \mathcal{B}', \mathcal{B})$ é uma invertível tal que a coluna j é p_j, $j \in \{1, \ldots, r\}$.

Teorema 7.2.13. *Sejam $r \in \mathbb{N}$ e $A \in \mathcal{M}_r(\mathbb{R})$.*
Se A admite r valores próprios distintos, então A é diagonalizável.

Demonstração. Sejam $\mathcal{B}$ uma base de $\mathbb{R}^r$ e f o endomorfismo de $\mathbb{R}^r$ cuja matriz em relação à base $\mathcal{B}$ é A. Como A e f têm os mesmos valores próprios, e A é diagonalizável se e só se f é diagonalizável, do Teorema 7.2.5 obtemos o resultado enunciado. $\square$

Teorema 7.2.14. *Sejam $r \in \mathbb{N}$ e $A \in \mathcal{M}_r(\mathbb{R})$ e $\lambda_1, \ldots, \lambda_p$ os valores próprios de A, distintos dois a dois. Então A é diagonalizável se e só se $\sum_{i=1}^{p} m.g.(\lambda_i) = r$.*

Demonstração. Sejam $\mathcal{B}$ uma base de $\mathbb{R}^r$ e f o endomorfismo de $\mathbb{R}^r$ cuja matriz em relação à base $\mathcal{B}$ é A. Como A e f têm os mesmos valores próprios, e A é diagonalizável se e só se f é diagonalizável, do Teorema 7.2.7 obtemos o resultado enunciado. $\square$

Corolário 7.2.15. *Sejam $r \in \mathbb{N}$ e $A \in \mathcal{M}_r(\mathbb{R})$ e $\lambda_1, \ldots, \lambda_p$ os valores próprios de A, distintos dois a dois. Então A é diagonalizável se e só se $\sum_{i=1}^{p} m.a.(\lambda_i) = r$ e, para cada $i \in \{1, \ldots, r\}$, $\mathrm{m.g.}(\lambda_i) = \mathrm{m.a.}(\lambda_i)$.*

Demonstração. Sejam $\mathcal{B}$ uma base de $\mathbb{R}^r$ e f o endomorfismo de $\mathbb{R}^r$ cuja matriz em relação à base $\mathcal{B}$ é A. Como A e f têm os mesmos valores próprios, e A é diagonalizável se e só se f é diagonalizável, do Teorema 7.2.8 obtemos o resultado enunciado. $\square$

Exemplo 7.2.16. *Sejam $\mathcal{B}$ a base canónica de $\mathbb{R}^3$ e $f \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^3)$ tal que*

$$M(f; \mathcal{B}, \mathcal{B}) = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 2 & 1 & 1 \end{bmatrix}.$$

Tem-se

$$p_f(x) = \begin{vmatrix} 2-x & -1 & 0 \\ -1 & 2-x & 0 \\ 2 & 1 & 1-x \end{vmatrix} = (1-x)\begin{vmatrix} 2-x & -1 \\ -1 & 2-x \end{vmatrix} = (-1)(x-1)^2(x-3).$$

Logo o polinómio característico de f decompõe-se em fatores lineares sobre $\mathbb{R}$, e os valores próprios de f são 1 e 3.

Como m.a(3)=1, também m.g.(3)=1.

Tem-se m.a.(1)=2 e m.g.(1)=$\dim\mathbb{R}^3_{[f,1]} = 3 - car(A - 1\mathrm{I}_3)$*. Aplicando o método de eliminação de Gauss à matriz $A - 1\mathrm{I}_3$*

$$A - \mathrm{I}_3 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

concluímos que $car(A - 1\mathrm{I}_3) = 2$, pelo que m.g.(1)=3-2=1.

Como $\dim\mathbb{R}^3 = 3$ *e* $\mathrm{m.g.}(1) + \mathrm{m.g.}(3) \neq 3$*, o endomorfismo f não é diagonalizável.*

Exemplo 7.2.17. *Sejam V um subespaço vetorial de $\mathbb{R}^5$ de dimensão 4, $\mathcal{B} = (v_1, v_2, v_3, v_4)$ uma base de V e $f \in \mathcal{L}(V, V)$ tal que*

$$M(f;\mathcal{B},\mathcal{B}) = \begin{bmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 2 & -2 & 0 & -2 \\ 1 & -1 & 0 & -1 \end{bmatrix} = A.$$

Tem-se

$$\begin{aligned} p_f(x) &= \begin{vmatrix} -1-x & 1 & 0 & 1 \\ 0 & -x & 0 & 0 \\ 2 & -2 & -x & -2 \\ 1 & -1 & 0 & -1-x \end{vmatrix} = (-x)\begin{vmatrix} -1-x & 1 & 1 \\ 0 & -x & 0 \\ 1 & -1 & -1-x \end{vmatrix} \\ &= (-x)(-x)\begin{vmatrix} -1-x & 1 \\ 1 & -1-x \end{vmatrix} = x^3(x+2). \end{aligned}$$

Logo, os valores próprios de f são 0 e -2.

Como $\mathrm{m.a.}(-2) = 1$*, também* $\mathrm{m.g.}(-2) = 1$.

Tem-se $\mathrm{m.a.}(0) = 3$ *e* $\mathrm{m.g.}(0) = \dim V_{[f,0]} = 4 - car(A - 0\mathrm{I}_4) = 4 - car(A)$*. Uma vez que*

$$\begin{bmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 2 & -2 & 0 & -2 \\ 1 & -1 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

segue que $\mathrm{m.g.}(0) = 4 - 1 = 3 = \mathrm{m.a.}(0)$.

Então, como $\dim\mathbb{R}^3 = 3$ *e* $m.g.(0) + m.g.(-2) = 4$*, f é diagonalizável.*

No sentido de determinarmos uma base $\mathcal{B}'$ de V tal que $M(f;\mathcal{B}',\mathcal{B}')$ seja diagonal, comecemos por determinar os subespaços próprios de f associados aos valores próprios 0 e -2.

Calculando o subespaço próprio de f associado ao valor próprio 0, temos

$$\begin{aligned}
V_{[f,0]} &= \{v \in V \mid f(v) = 0 \cdot v\} \\
&= \{v \in V \mid f(v) - 0 \cdot v = 0_{\mathbb{R}^n}\} \\
&= \{v \in V \mid (f - 0id_V)(v) = 0_{\mathbb{R}^n}\} \\
&= \left\{ \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 \in V \mid (A - 0I_4) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\} \\
&= \{\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 \in V \mid \alpha_1 = \alpha_2 + \alpha_4\} \\
&= \{(\alpha_2 + \alpha_4) v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 \in V \mid \alpha_2, \alpha_4 \in \mathbb{R}\} \\
&= \{\alpha_2 (v_1 + v_2) + \alpha_3 v_3 + \alpha_4 (v_1 + v_4) \in V \mid \alpha_2, \alpha_3 \alpha_4 \in \mathbb{R}\} \\
&= \ < v_1 + v_2, v_3, v_1 + v_4 > .
\end{aligned}$$

A sequência de vetores $(v_1 + v_2, v_3, v_1 + v_4)$ é linearmente independente, pois, para quaisquer $\alpha, \beta, \gamma \in \mathbb{R}$,

$$\begin{aligned}
& \alpha(v_1 + v_2) + \beta v_3 + \gamma(v_1 + v_4) = 0_{\mathbb{R}^n} \\
\Rightarrow\ & (\alpha + \gamma) v_1 + \alpha v_2 + \beta v_3 + \gamma v_4 = 0_{\mathbb{R}^n} \\
\Rightarrow\ & \alpha + \gamma = 0 \text{ e } \alpha = 0 \text{ e } \beta = 0 \text{ e } \gamma = 0 \quad \textit{\scriptsize ((v_1, v_2, v_3, v_4) é linearmente independente)} \\
\Rightarrow\ & \alpha = \beta = \gamma = 0
\end{aligned}$$

Assim, $(v_1 + v_2, v_3, v_1 + v_4)$ é uma base de $V_{[f,0]}$ formada por vetores próprios de f assciados ao valor próprio 0.

Determinemos o subespaço próprio de f associado ao valor próprio -2.

Tem-se

$$\begin{aligned}
V_{[f,-2]} &= \{v \in V \mid f(v) = -2v\} \\
&= \{v \in V \mid f(v) + 2v = 0_{\mathbb{R}^n}\} \\
&= \{v \in V \mid (f + 2id_V)(v) = 0_{\mathbb{R}^n}\} \\
&= \left\{ \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 \in V \mid (A + 2I_4) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\} \\
&= \{\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 \in V \mid \alpha_1 = -\alpha_4, \alpha_2 = 0, \alpha_3 = 2\alpha_4\} \\
&= \{(-\alpha_4) v_1 + 0 v_2 + 2\alpha_4 v_3 + \alpha_4 v_4 \in V \mid \alpha_4 \in \mathbb{R}\} \\
&= \{\alpha_4 (-v_1 + 2v_3 + v_4) \in V \mid \alpha_4 \in \mathbb{R}\} \\
&= \ < -v_1 + 2v_3 + v_4 > .
\end{aligned}$$

A sequência $(-v_1 + 2v_3 + v_4)$ é linearmente independente, pois para qualquer $\alpha \in \mathbb{R}$,

$$\alpha(-v_1 + 2v_3 + v_4) = 0 \Rightarrow (-\alpha) v_1 + \alpha v_3 + \alpha v_4 = 0 \Rightarrow \alpha = 0.$$

Logo $(-v_1 + 2v_3 + v_4)$ é uma base de $V_{[f,-2]}$ formada por um vetor próprio de f associado ao valor próprio -2.

Por último, verifica-se que a sequência

$$(-v_1 + 2v_3 + v_4, v_1 + v_2, v_3, v_1 + v_4)$$

é linearmente independente, pois, para quaisquer $\alpha, \beta, \gamma, \delta \in \mathbb{R}$,

$$\begin{aligned}
& \alpha(-v_1 + 2v_3 + v_4) + \beta(v_1 + v_2) + \gamma v_3 + \delta(v_1 + v_4) = 0_{\mathbb{R}^n} \\
\Rightarrow\ & (-\alpha + \beta + \gamma)v_1 + \beta v_2 + (\gamma + 2\alpha)v_3 + (\alpha + \delta)v_4 = 0_{\mathbb{R}^n} \\
\Rightarrow\ & -\alpha + \beta + \gamma = 0 \ e\ \beta = 0 \ e\ \gamma + 2\alpha = 0 \ e\ \alpha + \delta = 0 \quad {\scriptstyle (\,(v_1, v_2, v_3, v_4) \text{ é linearmente independente})} \\
\Rightarrow\ & \alpha = \beta = \gamma = \delta = 0.
\end{aligned}$$

Então, como $\dim V = 4$*, a sequência anterior é uma base de* V *formada por vetores próprios de* f*; designemos esta base por* $\mathcal{B}'$*. Uma vez que* $-v_1 + 2v_3 + v_4$ *é um vetor próprio de* f *associado ao valor próprio* -2 *e que* $v_1 + v_2$*,* v_3*,* $v_1 + v_4$ *são vetores próprios de* f *associados ao valor próprio* 0*, tem-se*

$$M(f; \mathcal{B}', \mathcal{B}') = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Desta forma, provámos também que a matriz $M(f; \mathcal{B}, \mathcal{B})$ *é diagonalizável, pois é semelhante à matriz* $M(f; \mathcal{B}', \mathcal{B}')$*. Note-se que*

$$M(f; \mathcal{B}, \mathcal{B}) = PM(f; \mathcal{B}', \mathcal{B}')P^{-1},$$

onde $P = M(id_V; \mathcal{B}', \mathcal{B})$*.*

Diagonalização de matrizes simétricas

Apresentamos seguidamente alguns resultados sobre a diagonalização de matrizes simétricas.

Para simplificar a escrita, no texto que se segue identificamos uma matriz $[a_{11}] \in \mathcal{M}_{11}(\mathbb{R})$ com o elemento a_{11}, sempre que nos referimos a operações e afirmações envolvendo o único elemento da matriz.

Teorema 7.2.18. *Sejam* $n \in \mathbb{N}$ *e* $A \in \mathcal{M}_n(\mathbb{R})$ *uma matriz simétrica. Então o polinómio característico de* A*,* $p_A(x)$*, tem* n *raízes em* $\mathbb{R}$*, isto é,* A *tem* n *valores próprios (não necessariamente distintos).*

Demonstração. Consideremos A como uma matriz de $\mathcal{M}_n(\mathbb{C})$. Como $p_A(x)$ tem grau n, pelo Teorema Fundamental da Álgebra, possui n raízes em $\mathbb{C}$, ou seja, a matriz A (enquanto matriz de $\mathcal{M}_n(\mathbb{C})$) tem n valores próprios em $\mathbb{C}$. Resta mostrar que tais valores próprios são reais.
Seja $\lambda \in \mathbb{C}$ um valor próprio de A, e $v \in \mathcal{M}_{n\times 1}(\mathbb{C}) \setminus \{0_{n\times 1}\}$ tal que $Av = \lambda v$. Multiplicando esta igualdade à esquerda por v^*, obtém-se:

$$v^*Av = \lambda v^*v.$$

Note-se que $v^*v = \sum_{i=1}^n |v_i|^2 > 0$ é real e positivo para $v \neq 0$. Além disso, como A é real e simétrica, tem-se $A^* = A^T = A$, logo:

$$\overline{v^*Av} = v^*Av,$$

isto é, v^*Av é real. Assim,

$$\lambda = \frac{v^*Av}{v^*v} \in \mathbb{R}.$$

Logo, todo valor próprio de uma matriz real e simétrica é real. □

Teorema 7.2.19 (Teorema de Schur)**.** *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{R})$. Se A tem n valores próprios (não necessariamente distintos), então existe Q ortogonal tal que Q^TAQ é triangular superior.*

Demonstração. A demonstração é feita por indução em n.

Base da indução: Se $n = 1$, o resultado é trivial, pois qualquer matriz 1×1 é triangular superior. Neste caso, basta tomar $Q = [1]$ que é ortogonal.

Hipótese de indução: Admitamos que para toda matriz $A \in \mathcal{M}_{n-1}(\mathbb{R})$ com $n-1$ valores próprios reais existe uma matriz ortogonal P tal que P^TAP é triangular superior.

Seja $A \in \mathcal{M}_n(\mathbb{R})$ com n valores próprios. Então existe pelo menos um valor próprio real λ_1 e um vetor não nulo v_1 tal que $Av_1 = \lambda_1 v_1$.

Seja $u_1 = v_1/\|v_1\|$. Como $u_1 \neq 0_{\mathbb{R}^n}$, podemos estender (u_1) a uma base ortonormal $(u_1, u_2, \ldots, u_n)$ de $\mathbb{R}^n$ pelo processo de Gram-Schmidt.

Seja

$$W = [\, u_1 \mid u_2 \mid \cdots \mid u_n \,].$$

a matriz ortogonal cujas colunas são os vetores u_1, u_2, ..., u_n. Então

$$AW = [\, Au_1 \mid Au_2 \mid \cdots \mid Au_n \,] = [\, \lambda_1 u_1 \mid \lambda_2 u_2 \mid \cdots \mid \lambda_n u_n \,]$$

pelo que

$$W^TAW = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix},$$

onde $A_1 \in \mathcal{M}_{n-1}(\mathbb{R})$. A matriz A_1 tem $n-1$ valores próprios reais e, por hipótese de indução, existe uma matriz ortogonal Q_1 tal que $Q_1^TA_1Q_1$ é triangular superior.

Consideremos, agora,

$$Q = W \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix}.$$

Então Q é ortogonal, pois é o produto de matrizes ortogonais, e tem-se

$$\begin{aligned} Q^TAQ &= \begin{bmatrix} 1 & 0 \\ 0 & Q_1^T \end{bmatrix} W^TAW \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & Q_1^T \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & * \\ 0 & Q_1^TA_1Q_1 \end{bmatrix}, \end{aligned}$$

sendo esta matriz triangular superior.

Assim, por indução, conclui-se que o resultado é válido para todo $n \in \mathbb{N}$. □

Teorema 7.2.20. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz simétrica. Então A é diagonalizável com uma matriz diagonalizante ortogonal, isto é, existe Q ortogonal tal que $Q^T AQ$ é diagonal.*

Demonstração. Pelo Teorema 7.2.19, existe $Q \in \mathcal{M}_n(\mathbb{R})$ ortogonal tal que $Q^T AQ = W$, onde W é triangular superior. Transpondo ambos os membros, e considerando que $A^T = A$, obtemos $Q^T AQ = W^T$. Logo $W^T = W$. Como W é triangular superior, então W é necessariamente uma matriz diagonal. □

Se A for uma matriz complexa, têm-se resultados similares aos anteriores, sendo as respetivas provas análogas às anteriores.

Teorema 7.2.21 (Teorema de Schur para matrizes complexas)**.** *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{C})$. Então existe Q unitária tal que $Q^T AQ$ é triangular superior.*

Teorema 7.2.22. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{C})$ uma matriz hermítica. Então A é diagonalizável com uma matriz diagonalizante unitária, isto é, existe Q unitária tal que $Q^T AQ$ é diagonal.*

7.3 Propriedades espectrais das matrizes definidas positivas

As matrizes simétricas definidas positivas, bem como as matrizes semi-definidas positivas, podem ser classificadas em função dos seus valores próprios.

Teorema 7.3.1. *Sejam $n \in \mathbb{N}$ e $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz simétrica. Então:*

1. *A matriz A é simétrica definida positiva se e só se todos os seus valores próprios são positivos.*
2. *A matriz A é simétrica semidefinida positiva se e só se todos os seus valores próprios são maiores ou iguais a zero.*

Demonstração. 1. Sejam $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz simétrica definida positiva e $\lambda \in \mathbb{R}$ um valor próprio de A. Seja v um vetor próprio de A associado a λ. Então $v \neq 0_{n\times 1}$ e tem-se

$$v^T Av = v^T(\lambda v) = \lambda v^T v.$$

Como A é definida positiva e $v \neq 0_{n\times 1}$, $v^T Av > 0$ e $v^T v > 0$. Logo $\lambda v^T v > 0$ e, portanto, $\lambda > 0$.

Reciprocamente, admitamos que todos os valores próprios de A são positivos. Pretendemos provar que, para todo $x \in \mathcal{M}_{n\times 1}(\mathbb{R}) \setminus \{0_{n\times 1}\}$, $x^T Ax > 0$. Como A é simétrica, então A é diagonalizável e existe uma matriz ortogonal tal que

$Q^TAQ = D$, onde $D = diag(\lambda_1, \ldots, \lambda_n)$. Como $Q^TQ = I_n = QQ^T$, então da igualdade $Q^TAQ = D$ obtem-se $A = QDQ^T$. Dado $x \in \mathcal{M}_{n\times 1}(\mathbb{R}) \setminus \{0_{n\times 1}\}$ e definindo $y = Q^Tx$, tem-se

$$x^TAx = x^TQDQ^Tx = y^TDy = \sum_{i=1}^{n} \lambda_i y_i^2,$$

onde $\lambda_i > 0$, para todo $i \in \{1 \ldots, n\}$.
Como $x \neq 0_{n\times 1}$, então $y = Q^Tx \neq 0$, pois Q^T é invertível. Assim, tem-se $y_i \neq 0$, para algum $i \in \{1, \ldots, n\}$. Por conseguinte,

$$x^TAx = \sum_{i=1}^{n} \lambda_i y_i^2 > 0,$$

pois cada termo é não negativo e pelo menos um é positivo.
Logo, A é definida positiva.

2. Exercício. $\square$

O resultado anterior também é válido para matrizes hermíticas definidas (resp. semidefinidas) positivas.