Apontamentos das aulas teóricas de Álgebra Linear

Instituto Superior Técnico
Departamento de Matemática

Engenharia Aeroespacial

1º Semestre 2014/2015

Prof. Paulo Pinto
http://www.math.tecnico.ulisboa.pt/~ppinto

Conteúdo

1 Matrizes e sistemas lineares **1**
1.1 Álgebra das Matrizes . . . 1
1.2 Operações elementares. Característica . . . 4
1.3 Sistemas de Equações Lineares . . . 8
1.4 Cálculo da matriz inversa . . . 13

2 Determinantes **15**

3 Espaços Lineares (ou Vectoriais) **20**
3.1 Subespaços lineares – p. ex.: núcleo, espaço colunas e linhas de uma matriz . 22
3.2 Independência linear . . . 27
3.3 Bases e dimensão de Espaços Lineares . . . 29
3.4 Coordenadas de um vector numa base . . . 36
3.5 Matriz mudança de base . . . 37

4 Valores Próprios, Vectores Próprios e diagonalização de Matrizes **38**

5 Transformações Lineares **45**
5.1 Representação matricial de uma transformação linear . . . 48
5.2 Transformações injectivas, sobrejectiva e bijectivas – equações lineares . . . 52
5.3 Valores e vectores próprios de transformações lineares . . . 56

6 Produtos Internos **57**
6.1 Bases ortogonais . . . 62
6.2 Complementos e projecções ortogonais . . . 65
6.3 Diagonalização ortogonal de matrizes simétricas . . . 68

7 Algumas Aplicações **76**
7.1 Formas quadráticas . . . 76
7.2 Mínimos quadrados . . . 77
7.3 Equações diferenciais ordinárias . . . 78
7.3.1 Um processo de difusão . . . 80
7.4 Genes ligados ao sexo . . . 81
7.5 Redes e grafos . . . 82

1 Matrizes e sistemas lineares

1.1 Álgebra das Matrizes

As matrizes[1] são uma ferramenta fundamental no estudo de álgebra linear.

Definição 1.1 Uma **matriz** A, do tipo $m \times n$ (m por n), é uma tabela de mn números dispostos em m linhas e n colunas:

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

A matriz **linha** i de A é:

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix},$$

para cada $i = 1, ..., m$. A matriz **coluna** j de A é:

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

para cada $j = 1, ..., n$. Usa-se também a notação $A = [a_{ij}]_{m \times n}$ na qual a_{ij} é a entrada (i, j) da matriz A.

Se $m = n$, diz-se que A é uma **matriz quadrada** do tipo $n \times n$ e as entradas a_{11}, a_{22}, ..., a_{nn} formam a chamada **diagonal principal** de A. Seja $\mathcal{M}_{m \times n}(\mathbb{R})$ o conjunto de todas as matrizes do tipo $m \times n$ com entradas em $\mathbb{R}$. Mais geralmente, sendo $\mathbb{K}$ um conjunto, $\mathcal{M}_{m \times n}(\mathbb{K})$ designa o conjunto de todas as matrizes do tipo $m \times n$ com entradas em $\mathbb{K}$.

Exemplo 1.2 As matrizes

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix},\ B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & -2 & 0 \end{bmatrix},\ C = \begin{bmatrix} 0 & 0 & 7 \end{bmatrix} \text{ e } D = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$$

são dos seguintes tipos: A é 2×2, B é 2×4, C é 1×3, A é 4×1. Tem-se, por exemplo, $a_{21} = -2$, $b_{13} = 3$, $c_{12} = 0$ e $d_{41} = 1$.

Dizemos que as matriz $A = [a_{ij}]$ do tipo $m \times n$ e a matriz $B = [b_{ij}]$ do tipo $p \times q$ são iguais se $m = p$, $n = q$ e $a_{ij} = b_{ij}$ para todos i, j.

Definição 1.3 1. *A* **soma matricial** *de uma matriz $A = [a_{ij}]$ do tipo $m \times n$ com outra matriz $B = [b_{ij}]$ do tipo $m \times n$ é a matriz $A + B$ do mesmo tipo cuja entrada (i, j) é $a_{ij} + b_{ij}$.*

2. *O produto de uma matriz $A = [a_{ij}]$ do tipo $m \times n$ por* **escalar** *λ é a matriz $\lambda A = [\lambda a_{ij}]$.*

[1] O termo *Matriz* foi utilizado pela primeira vez por James Sylvester (1814-1897) em 1850

3. *O* **produto matricial** *$A = [a_{ij}]$ do tipo $m \times p$ com outra matriz $B = [b_{ij}]$ do tipo $p \times n$ é uma matriz $C = c_{ij}$ do tipo $m \times n$, designada por AB, cuja entrada (i, j) é dada por*

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + ... + a_{ip}b_{pm} = \sum_{k=1}^{p} a_{ik}b_{kj}.$$

4. *A* **transposta** *da matriz $A = [a_{ij}]$ de tipo $m \times n$ é a matriz $A^T = [a_{ji}]$ de tipo $n \times m$.*

Exemplo 1.4 Sejam

$$A = \begin{bmatrix} 1 & 4 & -1 \\ -3 & 2 & 6 \end{bmatrix}, B = \begin{bmatrix} 0 & -3 & 2 \\ 4 & -1 & -5 \end{bmatrix}, C = \begin{bmatrix} -1 \\ -1/2 \\ 2 \end{bmatrix} \text{ e } D = \begin{bmatrix} -2 & \sqrt{3} \end{bmatrix}.$$

Tem-se $A+B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ e não é possível somar C com D. Temos $-2A = \begin{bmatrix} -2 & -8 & 2 \\ 6 & -4 & -12 \end{bmatrix}$ e

$$A^T = \begin{bmatrix} 1 & -3 \\ 4 & 2 \\ -1 & 6 \end{bmatrix} \quad \text{e} \quad C^T = \begin{bmatrix} -1 & -\frac{1}{2} & 2 \end{bmatrix}.$$

Não é possível efectuar, por exemplo, AB. Os produtos AC e CD são possíveis e tem-se:

$$AC = \begin{bmatrix} -5 \\ 14 \end{bmatrix} \quad \text{e} \quad CD = \begin{bmatrix} 2 & -\sqrt{3} \\ 1 & -\sqrt{3}/2 \\ -4 & 2\sqrt{3} \end{bmatrix}.$$

Observação 1.5 O produto de matrizes não é comutativo. Por exemplo, para

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \text{ tem-se } AB = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ e } BA = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Logo $AB \neq BA$.

Teorema 1.6 Sejam A, B, C e D matrizes de tipos apropriados, α e β escalares. São válidas as seguintes propriedades para as operações matriciais.

1. (Comutatividade da soma) $A + B = B + A$.
2. (Associatividade da soma) $A + (B + C) = (A + B) + C$.
3. (Elemento neutro da soma) Existe uma única matriz $\mathbf{0}$ do tipo $m \times n$ tal que $A + \mathbf{0} = A$, para toda a matriz A do tipo $m \times n$. À matriz $\mathbf{0}$, cujas entradas são todas iguais a zero, chama-se **matriz nula**.
4. (Simétrico) Para cada matriz A existe uma única matriz B tal que $A + B = \mathbf{0}$. Esta matriz B denota-se por $-A$.
5. (Associatividade do produto por escalares) $\alpha(\beta A) = (\alpha\beta) A$.
6. (Distributividade) $(\alpha + \beta) A = \alpha A + \beta A$.
7. (Distributividade) $\alpha(A + B) = \alpha A + \alpha B$.

8. (Associatividade do produto de matrizes) $A(BC) = (AB)C$.

9. (Distributividade) $A(B+C) = AB + AC$ e $(B+C)D = BD + CD$.

10. $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

11. $\left(A^T\right)^T = A$.

12. $(A+B)^T = A^T + B^T$.

13. $(\alpha A)^T = \alpha A^T$.

14. $(AB)^T = B^T A^T$.

15. $(A_1 A_2 ... A_n)^T = A_n^T ... A_2^T A_1^T$, com A_1, A_2, ..., A_n matrizes de tipos apropriados.

16. À matriz, do tipo $n \times n$,
$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$
chama-se **matriz identidade** (de ordem n) e é tal que
$$AI = A \quad \text{e} \quad IB = B,$$
para todas as matrizes $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{n\times m}$.

Definição 1.7 Seja A uma matriz quadrada $n \times n$. A matriz A é invertível se existir uma matriz B tal que
$$AB = BA = I.$$

Teorema 1.8 Caso exista, a inversa de uma matriz quadrada é única, que designamos por A^{-1}.

A matriz nula não é invertível e a matriz identidade é invertível, tendo-se $I^{-1} = I$.

Teorema 1.9 Sejam A e B matrizes invertíveis e α escalar não nulo. Então AB, A^T e αA também são invertíveis e
$$(AB)^{-1} = B^{-1}A^{-1}, \qquad (\alpha A)^{-1} = \frac{1}{\alpha}\, A^{-1}, \qquad (A^T)^{-1} = (A^{-1})^T.$$

Seja A uma matriz quadrada e $k \in \mathbb{N}$. Chamamos potência de expoente k de A, e designa-se por A^k, à matriz $A \cdots A$ multiplicando a mesma matriz k-vezes (por exemplo, $A^2 = AA$, $A^3 = AAA = (AA)A = A(AA)$). Coloca-se $A^0 = I$ se A for não nula. Se A for invertível, então pelo último teorema A^k também é invertível e
$$(A^k)^{-1} = A^{-1} \cdots A^{-1}$$
onde A^{-1} aparece k-vezes nesta multiplicação. Portanto podemos definir A^{-k} com sendo $(A^k)^{-1}$.

Definição 1.10 Seja $A = [a_{ij}] \in \mathcal{M}_{n\times n}(\mathbb{R})$.

1. A é **simétrica** se $A = A^T$, isto é, se $a_{ij} = a_{ji}$, para $i, j = 1, ..., n$.

2. A é **anti-simétrica** se $A = -A^T$, isto é, se $a_{ij} = -a_{ji}$, para $i, j = 1, ..., n$.

3. A é **ortogonal** se A for invertível e $A^T = A^{-1}$.

1.2 Operações elementares. Característica

Seja A uma matriz $n \times m$. Operações elementares sobre as linhas de A cada uma das seguinte tipo de operação:

1. permutação (i.e. troca) de duas linhas
2. multiplicação de uma linha por um escalar não nulo
3. adição de uma linha a uma outra.

Combinando as operações 2) e 3) obtém-se a chamada operação de Jacobi, que consiste em adicionar a uma linha uma outra previamente multiplicada por um escalar. Por abuso de forma, usaremos a operaçãode Jacobi como sendo uma operação elementar (substituindo a 3) acima descrita. Vamos adoptar as seguintes notações para as transformações elementares sobre as linhas de uma matriz:

1. $\mathbf{L_i} \leftrightarrow \mathbf{L_j}$, para representar que se efectuou a troca das linhas L_i e L_j
2. $\boldsymbol{\alpha}\mathbf{L_i} \rightarrow \mathbf{L_i}$, para representar que a linha L_i foi multiplicada pelo escalar $\alpha \neq 0$.
3. $\mathbf{kL_j} + \mathbf{L_i} \rightarrow \mathbf{L_i}$, para representar que a nova linha L_i é obtida somando à linha L_i a linha L_j previamente multiplicada por um escalar k.

Se a matriz A foi transformada na matriz B do uma operação elementar, então usamos a seguinte notação (respectivamente)

$$A \underset{L_i \leftrightarrow L_j}{\longrightarrow} B, \qquad A \underset{\alpha L_i \rightarrow L_i}{\longrightarrow} B, \qquad A \underset{L_i + kL_j \rightarrow L_i}{\longrightarrow} B.$$

Dizemos então que a matriz A foi condensada na matriz U se U for obtida por sucessiva aplicação de operações elementares de tal forma que a matriz U está em **escada de linhas**: isto é, por baixo do primeiro elemento não nulo de cada linha (e na mesma coluna) todos os elementos são nulos. Chama-se **método de eliminação de Gauss** a este algoritmo de condensação da matriz A.

À primeira entrada não nula em cada linha da uma matriz U em escada de linhas chamamos **pivô**.

Este processo de condensação aplica-se naturalmente a qualquer matriz A (nõ necessariamente quadrada).

Exemplo 1.11 Vamos transformar em escada de linhas a seguinte matriz A:

$$A = \begin{bmatrix} 0 & 0 & 3 & -9 \\ 5 & 15 & -10 & 40 \\ 1 & 3 & -1 & 5 \end{bmatrix} \underset{L_1 \leftrightarrow L_3}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 5 & 15 & -10 & 40 \\ 0 & 0 & 3 & -9 \end{bmatrix} \underset{\frac{1}{5}L_2 \rightarrow L_2}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 1 & 3 & -2 & 8 \\ 0 & 0 & 3 & -9 \end{bmatrix}$$

$$\underset{-L_1 + L_2 \rightarrow L_2}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 3 & -9 \end{bmatrix} \underset{3L_2 + L_3 \rightarrow L_3}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U.$$

Em particular, conclui-se que $\mathrm{car}(A) = 2$, pois a matriz U tem 2 linhas não nulas e as entradas $(1,1)$ e $(2,3)$ são os pivôs de U.

Definição 1.12 Uma **matriz elementar** do tipo $n \times n$ é uma matriz obtida da matriz identidade I através de uma única operação elementar.

(i) A matriz P_{ij}, chamada **matriz de permutação**, é a matriz elementar obtida por troca da linha i com a linha j da matriz I. Tem-se:

$$P_{ij} = \left[\begin{array}{ccccccccccc}
1 & 0 & \cdots & & & & & & & \cdots & 0 \\
0 & \ddots & \ddots & & & & & & & & \vdots \\
\vdots & \ddots & 1 & & & & & & & & \\
 & & & 0 & & & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & & & 0 & & & \\
 & & & & & & & & 1 & \ddots & \vdots \\
\vdots & & & & & & & & \ddots & \ddots & 0 \\
0 & \cdots & & & & & & & \cdots & 0 & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\
\end{array} .$$

(ii) A matriz $E_i(\alpha)$ é a matriz elementar obtida da matriz I através do produto do escalar $\alpha \neq 0$ pela linha i da matriz I. Tem-se:

$$E_i(\alpha) = \left[\begin{array}{ccccccc}
1 & 0 & \cdots & & & \cdots & 0 \\
0 & \ddots & \ddots & & & & \vdots \\
\vdots & \ddots & 1 & & & & \\
 & & & \alpha & & & \\
 & & & & 1 & \ddots & \vdots \\
\vdots & & & & \ddots & \ddots & 0 \\
0 & \cdots & & & \cdots & 0 & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \\ \leftarrow i \\ \\ \\ \\
\end{array} .$$

(iii) A matriz $E_{ij}(\alpha)$ é a matriz elementar obtida da matriz I por soma da linha j com um múltiplo α da linha i. Tem-se:

$$E_{ij}(\alpha) = \left[\begin{array}{ccccccc}
1 & 0 & \cdots & & & \cdots & 0 \\
0 & \ddots & \ddots & & & & \vdots \\
\vdots & \ddots & 1 & & & & \\
 & & & \ddots & & & \\
 & & \alpha & & 1 & \ddots & \vdots \\
\vdots & & & & \ddots & \ddots & 0 \\
0 & \cdots & & & \cdots & 0 & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\
\end{array} .$$

Exemplo 1.13 As matrizes elementares do tipo 2×2 são:

$$P_{12} = P_{21} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E_1(\alpha) = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}, \quad E_2(\alpha) = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix},$$

com $\alpha \neq 0$,

$$E_{12}(\alpha) = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \quad \text{e} \quad E_{21}(\alpha) = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}.$$

Teorema 1.14 Sejam E uma matriz elementar do tipo $m \times m$ e A uma matriz qualquer do tipo $m \times n$. Então, EA é a matriz obtida de A através da mesma operação elementar que originou E. Isto é, aplicar uma operação elementar a uma matriz corresponde a multiplicar essa matriz à esquerda por uma matriz elementar.

Exemplo 1.15 Consideremos a matriz

$$A = \begin{bmatrix} 0 & 0 & 3 & -9 \\ 5 & 15 & -10 & 40 \\ 1 & 3 & -1 & 5 \end{bmatrix}.$$

A operação elementar:

$$\begin{bmatrix} 0 & 0 & 3 & -9 \\ 5 & 15 & -10 & 40 \\ 1 & 3 & -1 & 5 \end{bmatrix} \underset{L_1 \leftrightarrow L_3}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 5 & 15 & -10 & 40 \\ 0 & 0 & 3 & -9 \end{bmatrix},$$

corresponde à seguinte multiplicação (à esquerda):

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 3 & -9 \\ 5 & 15 & -10 & 40 \\ 1 & 3 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 & 5 \\ 5 & 15 & -10 & 40 \\ 0 & 0 & 3 & -9 \end{bmatrix}.$$

A operação elementar:

$$\begin{bmatrix} 1 & 3 & -1 & 5 \\ 5 & 15 & -10 & 40 \\ 0 & 0 & 3 & -9 \end{bmatrix} \underset{\frac{1}{5}L_2 \to L_2}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 1 & 3 & -2 & 8 \\ 0 & 0 & 3 & -9 \end{bmatrix},$$

corresponde à seguinte multiplicação (à esquerda):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 5 & 15 & -10 & 40 \\ 0 & 0 & 3 & -9 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 & 5 \\ 1 & 3 & -2 & 8 \\ 0 & 0 & 3 & -9 \end{bmatrix}.$$

A operação elementar:

$$\begin{bmatrix} 1 & 3 & -1 & 5 \\ 1 & 3 & -2 & 8 \\ 0 & 0 & 3 & -9 \end{bmatrix} \underset{-L_1 + L_2 \to L_2}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 3 & -9 \end{bmatrix},$$

corresponde à seguinte multiplicação (à esquerda):

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 1 & 3 & -2 & 8 \\ 0 & 0 & 3 & -9 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 3 & -9 \end{bmatrix}.$$

Finalmente, a operação elementar:

$$\begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 3 & -9 \end{bmatrix} \underset{3L_2 + L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

corresponde à seguinte multiplicação (à esquerda):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 3 & -9 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Tem-se então:

$$E_{23}\,(3)\,E_{12}\,(-1)\,E_2\left(\frac{1}{5}\right)P_{13}\begin{bmatrix} 0 & 0 & 3 & -9 \\ 5 & 15 & -10 & 40 \\ 1 & 3 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note-se que as matrizes elementares são invertíveis, sendo

$$P_{ij}^{-1} = P_{ij}, \quad E_i(\alpha)^{-1} = E_i(\frac{1}{\alpha}), \quad \text{e } E_{ij}(\alpha)^{-1} = E_{ij}(-\alpha).$$

Uma matriz diz-se **matriz triangular superior** (triangular inferior) se as entradas por baixo (por cima, respectivamente) da diagonal principal são todas nulas.

Seja A uma matriz quadrada e considere-se a sua transformação em matriz em escada de linhas A_k (que é triangular superior) usando matrizes elementares $E_1, ..., E_k$:

$$A \underset{E_1}{\longrightarrow} A_1 \underset{E_2}{\longrightarrow} A_2 \underset{E_3}{\longrightarrow} .... \underset{E_t}{\longrightarrow} A_k.$$

Por aplicação repetida do Teorema 1.14 temos

$$(E_t...E_2E_1)A = A_k, \tag{1}$$

e obtém-se

$$A = (E_k.....E_2E_1)^{-1}A_k = (E_1^{-1}...E_k^{-1})A_k.$$

Se as matrizes elementares $E_1^{-1}, ...E_k^{-1}$ não são de permutação, então a matriz $L := E_1^{-1}...E_k^{-1}$ é uma matriz triangular inferior (o produto de matrizes triangulares inferiores é uma matriz triangular inferior) e $U := A_k$ é uma matriz triangular superior. Temos assim a decomposição $A = LU$ da matriz inicial A como produto de duas matrizes triangulares.

Todavia há matrizes para as quais não é possível efectuar a decomposição $A = LU$. Nestes casos, efectuamos todas as trocas de linhas no início da condensação obtendo uma matriz de permutação P (produto de matrizes elementares associadas a troca de linhas). Como as trocas foram todas realizadas no início, podemos então transformar a matriz PA numa matriz em escada de linhas sem recorrer a mais trocas de linhas.

Teorema 1.16 (Factorização triangular). Seja A uma matriz do tipo $n \times n$. Então ou A admite a factorização $A = LU$ ou existe uma matriz de permutação P tal que PA admite a factorização $PA = LU$, onde L e U são respectivamente uma matriz triangular inferior e uma matriz triangular superior.

Veremos aplicações desta factorização, p.ex, na Secção 1.3.

Exemplo 1.17 Seja $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 5 \end{bmatrix}$. Tem-se:

$$E_{23}(1)E_{13}(-2)E_{12}(-2)A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix}.$$

Logo,

$$A = \left(E_{12}(-2)\right)^{-1}\left(E_{13}(-2)\right)^{-1}\left(E_{23}(1)\right)^{-1}\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix}.$$

Isto é,

$$A = E_{12}(2)E_{13}(2)E_{23}(-1)\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix},$$

ou ainda,

$$A = LU,$$

com

$$L \;=\; E_{12}(2)E_{13}(2)E_{23}(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix}.$$

1.3 Sistemas de Equações Lineares

O estudo dos sistemas lineares remonta aos Matemáticos da Babilónia (c.2000 a.C.) e da China (c.200 a.C.) e tomou a sua forma actual no século XIX, destacando-se os trabalhos de Arthur Cayley (1821-1895) e James Sylvester.

Definição 1.18 Uma **equação linear** com n incógnitas $x_1, x_2, ..., x_n$ é uma equação da forma

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b,$$

em que $a_1, a_2, ..., a_n$ e b são escalares (reais ou complexos).

Definição 1.19 Um **sistema de** m **equações lineares** com n incógnitas é um conjunto de equações lineares da forma

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

em que a_{ij} e b_k são escalares (reais ou complexos), para $i, k = 1, ..., m$ e $j = 1, ..., n$.

Sempre que os a_{ij} e b_k forem todos coeficientes reais, as incógnitas (ou variáveis) $x_1, ..., x_n$ do sistema (2) também se consideram variáveis reais (por defeito). Usando o produto de matrizes definido na Secção 1.1, o sistema linear (2) pode ser escrito como uma equação matricial

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

em que

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

A matriz A é designada por matriz dos coeficientes das incógnitas, $\mathtt{x}$ matriz-coluna das incógnitas e b matriz-coluna dos termos independentes dos sistema linear (2). O sistema anterior também pode ser representado por

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

dizendo-se que esta matriz é a **matriz aumentada** do sistema.

Dizemos que $s = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{bmatrix}$ é uma solução do sistema linear (2) se todas as equações de (2) forem satisfeitas, quando substituímos $x_1 = s_1, ..., x_n = s_n$, portanto

$$\begin{cases} a_{11}s_1 + a_{12}s_2 + ... + a_{1n}s_n = b_1 \\ a_{21}s_1 + a_{22}s_2 + ... + a_{2n}s_n = b_2 \\ \cdots \\ a_{m1}s_1 + a_{m2}s_2 + ... + a_{mn}s_n = b_m \end{cases} . \tag{3}$$

Na forma matricial, todavia equivalente, s é solução do sistema linear (2) se $As = b$. Ao conjunto $\mathcal{S}$ de todos as soluções de (2) damos o nome de conjunto-solução ou solução geral do sistema, i.e.

$$\mathcal{S} = \{s = (s_1, ..., s_n) \in \mathbb{R}^n : A \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{bmatrix} = b\}.$$

Pretendemos uma descrição (algébrica e geométrica) do conjunto-solução. O **método de eliminação de Gauss** (ver Secção 1.2) permite obter essa descrição, que consiste em aplicar as operações elementares na matriz aumentada, obtendo sistemas lineares equivalentes ao inicial (i.e. com o mesmo conjunto-solução) mas de resolução simplificada. Facilmente se prova que se a matriz aumentada $[A_1|b_1]$ é obtida da matriz aumentada $[A|b]$ usando uma operação elementar, então os 2 sistemas lineares associados têm o mesmo conjunto-solução.

Uma matriz em escada por linhas diz-se na **forma reduzida** se

- está em escada por linhas,
- todos os pivôs são iguais a 1 e
- cada pivô é a única entrada não nula da coluna respectiva.

Por exemplo, a matriz $\begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 0 & 1 & 2/5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ está em escada por linhas mas não está em escada reduzida. A matriz $\begin{bmatrix} 1 & 0 & \mathbf{0} & 4 \\ 0 & 0 & 1 & 2/5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ está na forma reduzida em escada por linhas.

É claro que podemos transformar uma qualquer matriz A numa matriz U na forma reduzida em escada reduzida. Em seguida apresentamos uma primeira consequência da invariância do conjunto solução por operações elementares na matriz aumentada do sistema de equações lineares.

Teorema 1.20 *A forma reduzida em escada por linhas de uma matriz é única.*

Dem.: Sejam U e U' duas matrizes em escada reduzidas associadas à mesma matriz A. Pretende-se demonstrar que $U = U'$. Com vista um absurdo, vamos assumir que $U \neq U'$. Escolha-se a coluna mais à esquerda na qual U e U' são diferentes. Em seguida, escolham-se todas as colunas à esquerda desta que contêm os pivôs (isto é, as colunas com povô à esquerda da que já foi escolhida em cima). Obtêm assim duas matrizes $\tilde{U}$ e $\tilde{U}'$:

$$\tilde{U} = \left[\begin{array}{c|c} I_n & b \\ \hline & 0 \\ \mathbf{0} & 0 \\ & \vdots \end{array}\right] \quad \text{ou} \quad \tilde{U} = \left[\begin{array}{c|c} I_n & 0 \\ \hline & 1 \\ \mathbf{0} & 0 \\ & \vdots \end{array}\right]$$

e

$$\tilde{U}' = \left[\begin{array}{c|c} I_n & b' \\ \hline & 0 \\ \mathbf{0} & 0 \\ & \vdots \end{array}\right] \quad \text{ou} \quad \tilde{U}' = \left[\begin{array}{c|c} I_n & 0 \\ \hline & 1 \\ \mathbf{0} & 0 \\ & \vdots \end{array}\right].$$

Note que U pode ser transformada em U', por operações elementares pois ambas foram obtidas usando operações elementares aplicadas à mesma matriz inicial A. Logo $\tilde{U}$ pode ser transformada na matriz $\tilde{U}'$, pois a remoção de colunas não altera a sucessão de operações elementares a ser usada. Olhemos, agora, para $\tilde{U}$ e $\tilde{U}'$ como sendo matrizes aumentadas de sistema lineares (que têm o mesmo conjunto solução pelo que acabámos de dizer). Ou os sistemas são possíveis, e nesse caso $b = b'$ ou os sistemas são impossíveis. Em qualquer caso, podemos concluir que $\tilde{U} = \tilde{U}'$, o que é absurdo porque por cosntrução a última colunas de $\tilde{U}$ não é igual à última coluna de $\tilde{U}'$. QED.

Chama-se **característica** de A, e designa-se por $\mathrm{car}(A)$, ao número de linhas não nulas da (única) matriz reduzida em escada por linhas U associada à matriz inicial A (usando operações elementares).

Exemplo 1.21 O sistema linear

$$\begin{cases} x + z = 3 \\ x + 2y + 2z = 6 \\ 3y + 3z = 6 \end{cases}$$

na forma matricial é

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \\ 0 & 3 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix}.$$

Consideremos então a matriz aumentada e o consequente método de eliminação de Gauss:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 1 & 2 & 2 & 6 \\ 0 & 3 & 3 & 6 \end{array}\right] \underset{-L_1+L_2 \to L_2}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 2 & 1 & 3 \\ 0 & 3 & 3 & 6 \end{array}\right] \underset{-\frac{3}{2}L_2+L_3 \to L_3}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & \frac{3}{2} & \frac{3}{2} \end{array}\right].$$

Logo o sistema linear inicial é equivalente ao sistema

$$\begin{cases} x+z=3 \\ \\ 2y+z=3 \\ \\ \frac{3}{2}z=\frac{3}{2} \end{cases} \Leftrightarrow \begin{cases} x=2 \\ \\ y=1 \\ \\ z=1. \end{cases}$$

Exemplo 1.22 O sistema linear

$$\begin{cases} 3z-9w=6 \\ \\ 5x+15y-10z+40w=-45 \\ \\ x+3y-z+5w=-7 \end{cases}$$

é equivalente a

$$\begin{bmatrix} 0 & 0 & 3 & -9 \\ 5 & 15 & -10 & 40 \\ 1 & 3 & -1 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 6 \\ -45 \\ -7 \end{bmatrix}.$$

Consideremos então a matriz aumentada e o consequente método de eliminação de Gauss:

$$\begin{bmatrix} 0 & 0 & 3 & -9 & | & 6 \\ 5 & 15 & -10 & 40 & | & -45 \\ 1 & 3 & -1 & 5 & | & -7 \end{bmatrix} \underset{\substack{L_1 \leftrightarrow L_3 \\ \frac{1}{5}L_2 \to L_2}}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 & | & -7 \\ 1 & 3 & -2 & 8 & | & -9 \\ 0 & 0 & 3 & -9 & | & 6 \end{bmatrix} \underset{-L_1+L_2 \to L_2}{\longrightarrow}$$

$$\longrightarrow \begin{bmatrix} 1 & 3 & -1 & 5 & | & -7 \\ 0 & 0 & -1 & 3 & | & -2 \\ 0 & 0 & 3 & -9 & | & 6 \end{bmatrix} \underset{3L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 3 & -1 & 5 & | & -7 \\ 0 & 0 & -1 & 3 & | & -2 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

Logo,

$$\begin{cases} x+3y-z+5w=-7 \\ \\ -z+3w=-2 \end{cases} \Leftrightarrow \begin{cases} x=-3y-2w-5 \\ \\ z=3w+2. \end{cases}$$

As incógnitas y e w são livres (isto é podem tomar valores arbitrários) e as incógnitas x e z são não livres. A solução geral do sistema é:

$$X = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -3y-2w-5 \\ y \\ 3w+2 \\ w \end{bmatrix},$$

para quaisquer $y, w \in \mathbb{R}$, isto é, o conjunto solução é dado por:

$$S = \{(-3y-2w-5, y, 3w+2, w) : y, w \in \mathbb{R}\}.$$

Neste exemplo o sistema tem **infinitas soluções**

Dado um sistema linear na forma matricial $A\mathtt{x} = b$, importa investigar o seu conjunto solução.

Notamos que se $\mathtt{x}_0$ e $\mathtt{x}_1$ forem soluções de $A\mathtt{x} = b$ tais que $\mathtt{x}_0 \neq \mathtt{x}_1$, então $\mathtt{x}_0 + \lambda(\mathtt{x}_1 - \mathtt{x}_0)$ também é solução de $A\mathtt{x} = b$, para cada escalar λ. Portanto se um sistema linear tiver duas soluções distintas então tem infinitas soluções.

Podemos assim classificar (quando ao tipo solução) os sistemas lineares da seguinte forma:

1. Impossíveis (os que têm o conjunto-solução vazio)
2. Possíveis:
   (a) Determinados (os que têm uma única solução)
   (b) Indeterminados (os que têm um número infinito de soluções)

Em particular, podemos concluir que não há sistemas lineares com precisamente 2 soluções, por exemplo.

Observação 1.23 Seja $[A \mid b]$ a matriz aumentada associada a um sistema linear com n incógnitas.

1. Se car $A =$ car $[A \mid b] = n$ então o sistema é **possível e determinado** (tem uma única solução).
2. Se car $A =$ car $[A \mid b] < n$ então o sistema é **possível e indeterminado** (tem um n^o infinito de soluções).
3. Se car $A <$ car $[A \mid b]$ então o sistema é **impossível** (não tem solução).
4. Podemos escolher como **incógnitas livres** (podem tomar valores arbitrários) do sistema aquelas que correspondem às colunas, que não contenham pivôs, da matriz em escada de linhas obtida de A através de operações elementares.
5. As **incógnitas não livres** do sistema são aquelas que correspondem às colunas, que contenham pivôs, da matriz em escada de linhas obtidas de A através de operações elementares.
6. car $A =$ n ^{o} de linhas não nulas da matriz em escada de linhas obtidas de $A =$ n^o de pivôs = n^o de incógnitas não livres.

Teorema 1.24 *Se A for uma matriz quadrada e invertível, então o sistema $A\mathtt{x} = b$ é possível e determinado. Além disso a única solução é $x = A^{-1}b$.*

Na Secção 1.2 desenvolvemos um método para factorizar uma matriz quadrada A na forma $A = LU$, com L uma matriz triangular inferior e U matriz triangular superior. Ora, o sistema $A\mathtt{x} = b$ pode ser escrito da seguinte forma

$$LU\mathtt{x} = b.$$

Em seguida definimos uma nova variável $\mathtt{y}$ tal que $U\mathtt{x} = \mathtt{y}$. Resolvemos o sistema $L\mathtt{y} = b$ e finalmente determinamos os valores de $\mathtt{x}$ usando a equação $U\mathtt{x} = \mathtt{y}$.

Embora a decomposição LU converta o problema de resolver um único sistema $A\mathtt{x} = b$ nos problemas de resolver dois sistemas $L\mathtt{y} = b$ e $U\mathtt{x} = \mathtt{y}$, esses sistemas são de resolução imediata porque as matrizes dos coeficientes das incógnitas são triangulares. Uma outra vantagem nas aplicações é que esta decomposição só usa a matriz A e não a matriz b, pelo que uma vez conhecida essa decomposição, podemos utilizá-la para para resolver $A\mathtt{x} = b$ para vários valores de b.

Se não existir a decomposição $A = LU$, então sabemos que existe uma matriz de permutação P tal que podemos efectuar a decompisição $PA = LU$. Pelo que a resolução do sistema $A\mathtt{x} = b$ é análoga ao anterior uma vez que as soluções de $PA\mathtt{x} = Pb$ são as mesmas do sistema $A\mathtt{x} = b$.

Sistemas homogéneos

Ao sistema $A\mathtt{x} = \mathbf{0}$ chamamos sistema homogéneo, onde a matriz-coluna dos termos independentes é a matriz-coluna nula. Note-se que o sistema homogéneo $A\mathtt{x} = \mathbf{0}$ é sempre possível, uma vez que $A\mathbf{0} = \mathbf{0}$. O próximo resultado indica-nos que o estudo de sistema lineares possíveis reduzem-se ao estudo dos sistemas homogéneos associados, sabendo uma solução particular.

Teorema 1.25 Seja $\mathcal{S}$ o conjunto solução do sistema $A\mathtt{x} = b$ e $\mathtt{x}_1 \in \mathcal{S}$. Seja ainda $\mathcal{S}_0$ o conjunto-solução dos sistema homogéneo associado $A\mathtt{x} = 0$. Então temos

$$\mathcal{S} = \mathtt{x}_1 + \mathcal{S}_0.$$

A prova deste Teorema é simples: Dado $\mathtt{x}_0$ solução de $A\mathtt{x} = 0$ facilmente se conclui que $\mathtt{x}_1 + \mathtt{x}_0$ é solução de $A\mathtt{x} = b$. Mais, se $\mathtt{x}_2$ é solução de $A\mathtt{x} = b$ então $(\mathtt{x}_2 - \mathtt{x}_1)$ é solução do sistema $A\mathtt{x} = 0$.

No Exemplo 1.21 o conjunto-solução obtido foi

$$\mathcal{S} = \{(-3y - 2w - 5, y, 3w + 2, w) : y, w \in \mathbb{R}\}.$$

Cada solução neste sistema pode ser escrito da seguinte forma (separando a parte que tem as variáveis livres):

$$(-3y - 2w - 5, y, 3w + 2, w) = (-5, 0, 2, 0) + (-3y - 2w, y, 3w, w).$$

Facilmente se verifica que $\{(-3y - 2w, y, 3w, w) : y, w \in \mathbb{R}\}$ é o conjunto-solução $\mathcal{S}_0$ do sistema homogéneo e que $(-5, 0, 2, 0)$ é uma solução particular de $A\mathtt{x} = b$ e portanto

$$\mathcal{S} = (-5, 0, 2, 0) + \mathcal{S}_0.$$

Já vimos que o vector nulo é solução do sistema homogéneo. O próximo resultado dá-nos uma ideia da estrutura crucial que está presente em Álgebra Linear.

Teorema 1.26 Se $\mathtt{x}_0$ e $\mathtt{x}_1$ são soluções do sistema homogéneo $A\mathtt{x} = 0$ e λ escalar, então $\mathtt{x}_0 + \mathtt{x}_1$ e $\lambda\mathtt{x}_0$ também são soluções do mesmo sistema homogéneo.

1.4 Cálculo da matriz inversa

Nesta Secção, vamos fornecer dois algoritmos para determinar a inversa (ver Definição 1.7).

1^0 algoritmo para a determinar a inversa de uma matriz

Na equação (1), se a matriz em escada de linhas A_k tem uma linha L_i toda nula, então A_t não é invertível uma vez que a linha L_i de A_kB é sempre nula, e portanto $A_kB \neq I$. Se A_t tiver todas as linhas não nulas, então com A_k está em escada de linhas e a matriz A_k é quadrada conclui-se que as entradas de diagonal de A_k são todas não nulas. Pelo que podemos prosseguir com as operações elementares a partir da matriz A_k por forma a transforma-la na matriz identidade I:

$$A \underset{E_1}{\longrightarrow} A_1 \underset{E_2}{\longrightarrow} A_2 \underset{E_3}{\longrightarrow} .... \underset{E_k}{\longrightarrow} A_k \underset{E_{k+1}}{\longrightarrow} A_{k+1} \underset{E_{k+2}}{\longrightarrow} ... \underset{E_s}{\longrightarrow} I.$$

Concluimos que

$$(E_s...E_k.....E_2E_1)A = I,$$

em particular concluimos que A é invertível e que

Teorema 1.27 *1.* $A^{-1} = E_s ... E_k ..... E_2 E_1$.

2. $car(A) = n$ *se e só se* A *é invertível.*

2^o algoritmo para calcular a inversa: método de Gauss-Jordan[2]

Note-se que se A for invertível, aplicando o método de eliminação de Gauss, podemos transformar matriz aumentada $[A|b]$ numa matriz do tipo $[I|c]$, em que a matriz coluna c será naturalmente a solução de $Ax = b$.

Se a matriz A do tipo $n \times n$. Pretendemos determinar uma matriz X tal que $AX = I$. Ora esta equação matricial pode ser resolvida através da resolução de n sistemas lineares:

$$A\mathtt{x}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; A\mathtt{x}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, ..., \; A\mathtt{x}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

em que $\mathtt{x}_1$ é a primeira coluna de X, ..., $\mathtt{x}_n$ é a n-ésima coluna de X. Estes n sistemas podem ser resolvidos em simultânea considerando a matriz aumentada $[A|I]$. pelo que foi exposto em cima, podemos transformar esta matriz aumentada numa matriz $[I|C]$, caso A seja invertível. Desta forma obtém-se a inversa de A como sendo $A^{-1} = C$.

Este algoritmo para obter a inversa de uma matriz chama-se **método de Gauss-Jordan** e consiste na continuação do método de eliminação de Gauss agora aplicado a

$$[\text{ matriz triangular superior } | \ast]$$

efectuando-se as eliminaç ões de baixo para cima de modo a obter-se $[I \mid A^{-1}]$.

Exemplo 1.28 **(i)** Seja $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 5 \end{bmatrix}$. Tem-se

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 4 & 0 & 1 & 0 \\ 2 & 3 & 5 & 0 & 0 & 1 \end{array}\right] \underset{\substack{-2L_1+L_2 \longrightarrow L_2 \\ -2L_1+L_3 \longrightarrow L_3}}{\longrightarrow} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & -2 & 1 & 0 \\ 0 & 1 & 3 & -2 & 0 & 1 \end{array}\right] \underset{L_2+L_3 \longrightarrow L_3}{\longrightarrow}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 5 & -4 & 1 & 1 \end{array}\right] \underset{\frac{1}{5}L_3 \longrightarrow L_3}{\longrightarrow} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -4/5 & 1/5 & 1/5 \end{array}\right] \underset{\substack{-2L_3+L_2 \longrightarrow L_2 \\ -L_3+L_1 \longrightarrow L_1}}{\longrightarrow}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 9/5 & -1/5 & -1/5 \\ 0 & -1 & 0 & -2/5 & 3/5 & -2/5 \\ 0 & 0 & 1 & -4/5 & 1/5 & 1/5 \end{array}\right] \underset{L_2+L_1 \longrightarrow L_1}{\longrightarrow}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 7/5 & 2/5 & -3/5 \\ 0 & -1 & 0 & -2/5 & 3/5 & -2/5 \\ 0 & 0 & 1 & -4/5 & 1/5 & 1/5 \end{array}\right] \underset{-L_2 \longrightarrow L_2}{\longrightarrow}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 7/5 & 2/5 & -3/5 \\ 0 & 1 & 0 & 2/5 & -3/5 & 2/5 \\ 0 & 0 & 1 & -4/5 & 1/5 & 1/5 \end{array}\right].$$

[2] O 3^o algoritmo para calcular a inversa de uma matriz invertível será dado no Capítulo 2

Portanto A é invertível e $A^{-1} = \begin{bmatrix} 7/5 & 2/5 & -3/5 \\ 2/5 & -3/5 & 2/5 \\ -4/5 & 1/5 & 1/5 \end{bmatrix}$.

(ii) Seja $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$. Tem-se

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \underset{-L_1+L_2 \longrightarrow L_2}{\longrightarrow} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -1 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \underset{L_2+L_3 \longrightarrow L_3}{\longrightarrow}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right].$$

Logo, A não é invertível e como tal não é invertível.

2 Determinantes

Definição *2.1* Dados os números naturais $1, 2, ..., n$ chama-se **permutação** desses n números a qualquer lista em em que os mesmos sejam apresentados por ordem arbitrária.

Definição *2.2* Seja $(i_1 i_2 ... i_n)$ uma permutação dos números naturais $1, 2, ..., n$. Diz-se que um par $(i_j i_k)$ é uma **inversão** quando $(j-k)(i_j - i_k) < 0$ (isto é, quando i_j e i_k aparecerem na permutação por ordem decrescente).

Definição *2.3* Uma permutação $(i_1 i_2 ... i_n)$ diz-se **par** (**ímpar**) quando o n^o máximo de inversões incluídas fôr par (ímpar).

Exemplo *2.4* A permutação (21453) é ímpar pois contem as inversões $(21), (43)$ e (53).

Definição *2.5* Seja A matriz $n \times n$. Chama-se **determinante**[3] de A, e escreve-se $|A|$ ou $\det(A)$, o número que se obtém do seguinte modo:

(i) Formam-se todos os produtos possíveis de n factores em que intervenha um elemento de cada linha e, simultaneamente, um elemento de cada coluna de A.

(ii) Afecta-se cada produto do sinal $+$ ou do sinal $-$ conforme as permutações (dos números naturais $1, 2, ..., n$) que figuram nos índices de linha e de coluna tenham a mesma paridade ou não.

(iii) Somam-se as parcelas obtidas.

Em resumo:

$$\det(A) = |A| = \sum_{\substack{(j_1 j_2 ... j_n) \\ \text{permutação de } 1,2,...,n}} (-1)^{\sigma} a_{1j_1} a_{2j_2} ... a_{nj_n},$$

em que

$$\sigma = \begin{cases} 0 & \text{se } (j_1 j_2 ... j_n) \text{ é par} \\ \\ 1 & \text{se } (j_1 j_2 ... j_n) \text{ é ímpar.} \end{cases}$$

[3]O Determinante de uma matriz foi pela primeira vez considerado por Talakazu Seki 1642–1708

Observação ***2.6*** Podemos ainda escrever de modo equivalente:

$$|A| = \sum_{\substack{(i_1 i_2 ... i_n) \\ \text{permutação de } 1,2,...,n}} (-1)^{\sigma} a_{i_1 1} a_{i_2 2} ... a_{i_n n},$$

em que

$$\sigma = \begin{cases} 0 & \text{se } (i_1 i_2 ... i_n) \text{ é par} \\ 1 & \text{se } (i_1 i_2 ... i_n) \text{ é ímpar.} \end{cases}$$

Teorema ***2.7*** Seja A matriz 2×2. Então

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

(ii) Seja A matriz 3×3. Então

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Exemplo ***2.8*** (i)

$$\begin{vmatrix} 1 & -1 \\ 2 & -2 \end{vmatrix} = 1(-2) - (-1)2 = 0.$$

(ii)

$$\begin{vmatrix} 1 & 2 & 1 \\ 3 & -1 & 2 \\ 2 & 1 & -3 \end{vmatrix} = 1(-1)(-3) + 3 + 8 - 1(-1)2 - 6(-3) - 2 = 32.$$

Observação ***2.9*** i) A área do paralelogramo definido pelos vectores $(a_1, b_1), (a_2, b_2)$ de $\mathbb{R}^2$ é

$$A = |\det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}|.$$

ii) O volume do paralelipípedo definido pelos vectores $(a_1, b_1, c_1), (a_2, b_2, c_2), (a_3, b_3, c_3)$ de $\mathbb{R}^3$ é dado por

$$A = |\det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}|.$$

Observação ***2.10*** Se A é do tipo $n \times n$ então $|A|$ tem $\mathbf{n}!$ parcelas, pelo que p.ex. se aplicarmos a definição de determinante a uma matriz 4×4, teremos $4! = 24$ parcelas. Em seguida vamos estudar métodos mais expeditos para o cálculo do determinante de uma matriz evitando o cálculo através desta definição.

Podemos facilmente calcular o determinante de cada matriz elementar:

$$\det(P_{ij}) = -1 \text{ (se} i \neq \text{j)}, \det(\mathrm{E_i}(\alpha)) = \alpha \text{ e } \det(\mathrm{E_{ij}}(\alpha)) = 1.$$

Teorema ***2.11*** Sejam A, B matrizes $n \times n$. Seja $\lambda \in \mathbb{R}$.

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

(ii) Se A fôr uma matriz triangular superior ou triangular inferior então $\det A =$ produto dos elementos da diagonal principal de A.

(iii) Se A tiver uma linha nula então $\det A = 0$.

(iv) Se B fôr obtida de A multiplicando uma linha de A por um número real λ então $\det B = \lambda \det A$.

(v) Se B fôr obtida de A somando a uma linha de A um múltiplo real λ de uma outra linha de A então $\det B = \det A$.

(vi) Se duas linhas de A forem iguais então $\det A = 0$.

(vii) Se B fôr obtida de A trocando duas linhas de A então $\det B = -\det A$.

(viii) $\det\left(A^T\right) = \det A$.

(ix) Se A fôr invertível $\det(A^{-1}) = \dfrac{1}{\det A}$.

(x) $\det(\lambda A) = \lambda^n \det A$.

(xi) $\det(AB) = 0 \Rightarrow \det A = 0 \text{ ou } \det B = 0$.

(xii) $\det(AB) = \det(BA)$.

(xiii) $\det(A) \neq 0$ se e só se A invertível.

Observação ***2.12*** **(i)** Para estabelecer parte (i) do teorema 2.11, sugere-se provar esse resultado para matrizes elementares.
(ii) Em geral, $\det(A + B) \neq \det(A) + \det(B)$.

Definição ***2.13*** Seja $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$, com $n > 1$. Seja A_{ij} a matriz do tipo $(n-1) \times (n-1)$ que se obtém de A suprimindo a linha i e a coluna j de A. Chama-se a A_{ij} o **menor-**ij da matriz A.

Teorema ***2.14*** (Fórmula de Laplace[4].) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, com $n > 1$. Tem-se

$$\det A = \sum_{j=1}^{n} a_{ij}(-1)^{i+j} \det A_{ij}.$$

Observação ***2.15*** Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, com $n > 1$. Tem-se

$$\det A = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} \det A_{ij}.$$

Exemplo ***2.16***

$$\begin{vmatrix} 1 & 0 & -2 & 3 \\ 2 & 1 & -1 & 4 \\ 0 & -1 & 0 & -2 \\ 1 & 0 & -2 & -3 \end{vmatrix} = (-1)(-1)^{3+2} \begin{vmatrix} 1 & -2 & 3 \\ 2 & -1 & 4 \\ 1 & -2 & -3 \end{vmatrix} + (-2)(-1)^{3+4} \begin{vmatrix} 1 & 0 & -2 \\ 2 & 1 & -1 \\ 1 & 0 & -2 \end{vmatrix} =$$

$$= (-1)(-3) + (-2)4 + 2(-2)3 - (-1)3 - (-2)2(-3) - 4(-2) + 2\left[(-2) - (-2)\right] = -18$$

Definição ***2.17*** Seja $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$, com $n > 1$. Seja $a'_{ij} = (-1)^{i+j} \det A_{ij}$ onde A_{ij} é o menor-ij da matriz A. Chama-se a a'_{ij} o **cofactor-**ij da matriz A e à matriz $\operatorname{cof} A = (a'_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$, com $n > 1$, a matriz dos cofactores de A.

Teorema ***2.18*** Para qualquer matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, com $n > 1$, tem-se

$$A\,(\operatorname{cof} A)^T = (\det A)\,I.$$

Se $\det A \neq 0$ então A é invertível e

$$A^{-1} = \frac{1}{\det A}\,(\operatorname{cof} A)^T.$$

Exemplo ***2.19*** Seja $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R})$ tal que $\det A \neq 0$. Então A é invertível e

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Note que $ad - bc = \det A$.

(ii) Podemos usar o teorema 2.18 para calcular não só a inversa de uma matriz (não singular) mas também entradas concretas dessa inversa. Seja

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

A entrada $(2,3)$ da matriz A^{-1} é dada por

$$(A^{-1})_{23} = \frac{1}{\det A}\left((\operatorname{cof} A)^T\right)_{23} = \frac{1}{\det A}\left((-1)^{3+2} \det A_{32}\right) = \frac{1}{-3}\left(-\det\left(\begin{bmatrix} 1 & 0 \\ 4 & 6 \end{bmatrix}\right)\right) = 2.$$

[4] Pierre-Simon Laplace 1749–1827

Exemplo *2.20* Seja $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Então

$$\operatorname{cof}(A) = \begin{bmatrix} a_{22}a_{33} - a_{23}a_{32} & -(a_{21}a_{33} - a_{23}a_{31}) & a_{21}a_{32} - a_{22}a_{31} \\ -(a_{12}a_{33} - a_{13}a_{32}) & a_{11}a_{33} - a_{13}a_{31} & -(a_{11}a_{32} - a_{12}a_{31}) \\ a_{12}a_{23} - a_{13}a_{22} & -(a_{11}a_{23} - a_{13}a_{21}) & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix}.$$

Apelando aos teoremas 2.7 e 2.18, podemos escrever a inversa de cada matriz invertível 3×3.

Teorema *2.21* (Regra de Cramer[5].) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é invertível. Então a única solução do sistema de equações lineares $AX = B$ é dada por

$$X = A^{-1}B = \frac{1}{\det A}\left(\operatorname{cof} A\right)^T B.$$

Isto é, sendo $X = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T$ e $B = \begin{bmatrix} b_1 & \ldots & b_n \end{bmatrix}^T$ tem-se

$$x_j = \frac{1}{\det A}\sum_{k=1}^{n} a'_{kj}b_k = \frac{\det B_j}{\det A},$$

onde B_j é a matriz obtida de A substituindo a coluna j de A pela matriz coluna B dos termos independentes.

Exemplo *2.22* O sistema de equações lineares

$$\begin{cases} 2x + y = 8 \\ -x + 2y + 4z = 7 \\ -x + z = 1 \end{cases}$$

pode ser resolvido usando a regra de Cramer:

$$x = \frac{\begin{vmatrix} 8 & 1 & 0 \\ 7 & 2 & 4 \\ 1 & 0 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 0 \\ -1 & 2 & 4 \\ -1 & 0 & 1 \end{vmatrix}} = 13, \quad y = \frac{\begin{vmatrix} 2 & 8 & 0 \\ -1 & 7 & 4 \\ -1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 0 \\ -1 & 2 & 4 \\ -1 & 0 & 1 \end{vmatrix}} = -18 \quad \text{e} \quad z = \frac{\begin{vmatrix} 2 & 1 & 8 \\ -1 & 2 & 7 \\ -1 & 0 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 0 \\ -1 & 2 & 4 \\ -1 & 0 & 1 \end{vmatrix}} = 14.$$

[5]Gabriel Cramer 1704–1752

3 Espaços Lineares (ou Vectoriais)

No final do século XIX e no começo do século XX tornou-se claro – graças a Grassmann[6], Peano[7] e a Weyl[8] – que o desenvolvimento axiomático da geometria Euclideana podia ser feito apelando a estruturas matemáticas — Espaços Vectoriais e Euclidianos — que desempenham um papel determinante noutras áreas da matemática e de outras ciências. O estudo das estruturas matemáticas independente quer dos contextos que lhes deram origem quer dos contextos em que aplicam constitui uma das ideias mais ricas da matemática do século XX e é indissociável da matemática Emmy Noether[9]. A Álgebra linear é basicamente o estuda dessas estruturas.

Definição 3.1 Um conjunto não vazio V é um **espaço linear** (real) se existirem duas operações associadas a V, uma soma de elementos de V e um produto de escalares (números reais) por elementos de V, com as seguintes propriedades:

(a) (Fecho da soma). Para quaisquer $u, v \in V$ tem-se $u + v \in V$.

(b) (Fecho do produto por escalares). Para quaisquer $\alpha \in \mathbb{R}$ e $u \in V$ tem-se $\alpha u \in V$.

(c) (Comutatividade da soma). Para quaisquer $u, v \in V$, $u + v = v + u$.

(d) (Associatividade da soma). Para quaisquer $u, v, w \in V$, $u + (v + w) = (u + v) + w$.

(e) (Elemento neutro da soma). Existe um elemento de V designado por $\mathbf{0}$ tal que, para qualquer $u \in V$, $u + \mathbf{0} = u$.

(f) (Simétrico). Para cada (qualquer) $u \in V$ existe $v \in V$ tal que $u + v = \mathbf{0}$. A v chama-se o **simétrico** de u e denota-se por $-u$.

(g) (Associatividade do produto por escalares). Para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u \in V$, $\alpha (\beta u) = (\alpha\beta) u$.

(h) (Distributividade em relação à soma de vectores). Para quaisquer $\alpha \in \mathbb{R}$ e $u, v \in V$, $\alpha (u + v) = \alpha u + \alpha v$.

(i) (Distributividade em relação à soma de escalares). Para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u \in V$, $(\alpha + \beta) u = \alpha u + \beta u$.

(j) Para qualquer $u \in V$, $1u = u$.

Observação 3.2 Aos elementos de V chamaremos vectores.

[6] Hermann Grassmann 1809–1877
[7] Giuseppe Peano 1858–1932
[8] Hermanm Weyl 1885–1955
[9] Emmy Noether 1882–1935

Exemplo 3.3 Exemplos de espaços lineares:

(i) $\mathbb{R}^n$, com as operações usuais:

$$(u_1, u_2, ..., u_n) + (v_1, v_2, ..., v_n) = (u_1 + v_1, u_2 + v_2, ..., u_n + v_n),$$

$$\alpha(u_1, u_2, ..., u_n) = (\alpha u_1, \alpha u_2, ..., \alpha u_n).$$

(ii) $\mathcal{M}_{m\times n}(\mathbb{R})$ (conjunto de todas as matrizes reais do tipo $m \times n$), com as operações (usuais): $A + B$ e αA.

(iii) O conjunto de todas as funções reais de variável real definidas num conjunto não vazio $S \subseteq \mathbb{R}$, com as operações usuais:

$$(f + g)(x) = f(x) + g(x),$$

$$(\alpha f)(x) = \alpha f(x).$$

(iv) O conjunto $\mathcal{P}$ de todos os polinómios reais, com as operações usuais.

(v) O conjunto $\mathcal{P}_n$ (por vezes designado por P_n) de todos os polinómios reais de grau menor ou igual a n (incluindo o polinómio nulo), com as operações usuais.

Observação 3.4 Um mesmo conjunto pode servir para formar espaços lineares diferentes:

(i) O conjunto dos números reais $\mathbb{R}$, com a soma definida por

$$u \boxplus v = u + v + 1,$$

e o produto por escalares definido por

$$\alpha \cdot u = \alpha u + \alpha - 1,$$

é um espaço linear. (Neste caso o elemento neutro é -1.)

(ii) O conjunto dos números reais maiores do que zero, com a soma definida por

$$u \boxplus v = uv,$$

e o produto por escalares definido por

$$\alpha \cdot u = u^\alpha,$$

é um espaço linear. (Neste caso o elemento neutro é 1.)

Observação 3.5 Alterações nos conjuntos considerados anteriormente podem resultar em conjuntos que não são espaços lineares.

(i) O conjunto $\{(x, y) \in \mathbb{R}^2 : x \geq 0 \text{ e } y \geq 0\}$, com as operações usuais, não é um espaço linear. Por exemplo, os simétricos não estão no conjunto.

(ii) O conjunto $V = \{a_0 + a_1 t + ... + a_n t^n : a_0, a_1, ..., a_n \in \mathbb{R} \text{ e } a_n \neq 0\}$, com as operações usuais, não é um espaço linear. Por exemplo:

$$t^n, -t^n + t \in V, \quad \text{mas} \;\; t^n + (-t^n + t) = t \notin V.$$

(iii) O conjunto $U = \{f : \mathbb{R} \longrightarrow \mathbb{R} \text{ tais que } f(1) = 2\}$, com as operações usuais, não é um espaço linear. Por exemplo, se $f_1, f_2 \in U$,

$$(f_1 + f_2)(1) = f_1(1) + f_2(1) = 2 + 2 = 4 \neq 2.$$

Logo, $f_1 + f_2 \notin U$.

3.1 Subespaços lineares – p. ex.: núcleo, espaço colunas e linhas de uma matriz

Definição 3.6 Seja V um espaço linear. Diz-se que U é um **subespaço** de V se U é um subconjunto de V e se U, com as operações de V, fôr um espaço linear.

Observação 3.7 No entanto, para mostrar que um certo conjunto $S \subset V$ é um subespaço do espaço linear V, não será necessário verificar os 10 axiomas da definição 3.1, como se pode ver no seguinte teorema.

Teorema 3.8 Um subconjunto não vazio U de um espaço linear V é um subespaço linear de V se e só se:

(i) Para quaisquer $u, v \in U$ tem-se $u + v \in U$.

(ii) Para quaisquer $\alpha \in \mathbb{R}$ e $u \in U$ tem-se $\alpha u \in U$.

Exemplo 3.9 Exemplos de subespaços:

(i) Os únicos subespaços do espaço linear $\mathbb{R}$, com as operações usuais, são $\{0\}$ e $\mathbb{R}$.

(ii) Os subespaços do espaço linear $\mathbb{R}^3$, com as operações usuais, são: $\{(0, 0, 0)\}$, $\mathbb{R}^3$, todas as rectas que passam pela origem e todos os planos que passam pela origem.

(iii) O conjunto de todas as matrizes (reais) triangulares superiores (do tipo $n \times n$) é um subespaço do espaço linear $\mathcal{M}_{n \times n}(\mathbb{R})$, com as operações usuais.

(iv) O conjunto de todas as funções reais definidas e contínuas em $I \subset \mathbb{R}$ (I é um intervalo) é um subespaço do espaço linear de todas as funções $f : I \longrightarrow \mathbb{R}$, com as operações usuais.

(v) Seja A uma matriz (real) do tipo $m \times n$. O conjunto

$$\mathcal{C}(A) = \{b \in \mathbb{R}^m : Au = b \text{ tem pelo menos uma solução } u\}$$

é um subespaço do espaço linear $\mathbb{R}^m$, com as operações usuais, ao qual se dá o nome de **espaço das colunas** de A.

(vi) Seja A uma matriz (real) do tipo $m \times n$. O conjunto

$$\mathcal{N}(A) = \{u \in \mathbb{R}^n : Au = \mathbf{0}\}$$

é um subespaço do espaço linear $\mathbb{R}^n$, com as operações usuais, ao qual se dá o nome de **espaço nulo ou núcleo** de A. (Confronte com o teorema 1.26)

Observação 3.10 **(i)** Se A é invertível então $\mathcal{N}(A) = \{\mathbf{0}\}$.

(ii) Se $\mathcal{N}(A) = \{\mathbf{0}\}$ então A é invertível.

(iii) Poderemos obter subespaços de um espaço linear através de combinações lineares de vectores desse espaço.

Definição 3.11 Seja S um subconjunto não vazio de um espaço linear V. Diz-se que um vector u é **combinação linear** finita dos elementos de S, se existir um n^o finito de elementos de S, $u_1, ..., u_k$, e de escalares $\lambda_1, ..., \lambda_k$ tais que

$$u = \lambda_1 u_1 + ... + \lambda_k u_k = \sum_{i=1}^{k} \lambda_i u_i.$$

Ao cojunto de todas as combinações lineares finitas de elementos de S chama-se **expansão linear** de S e designa-se por $L(S)$. Se S é o conjunto vazio $\varnothing$, escreve-se $L(\varnothing) = \{0\}$.

Teorema 3.12 Seja S um subconjunto não vazio de um espaço linear V. A expansão linear $L(S)$ de S é o menor subespaço de V que contém S. Deste modo, a $L(S)$ também se chama o **subespaço gerado** por S, e diz-se que S **gera** $L(S)$.

Observação 3.13 Seja S e T dois subconjuntos não vazios de um espaço linear V, com $S \subset T$. Se $L(S) = V$ então $L(T) = V$.

Exemplo 3.14 (i) O espaço linear $\mathbb{R}^2$ é gerado por qualquer dos seguintes conjuntos de vectores:

$$\{(1,0),(0,1)\}, \quad \{(1,2),(-1,11)\} \quad \text{e} \quad \{(23,8),(6,14)\}.$$

(ii) O subespaço $\{(x,y) \in \mathbb{R}^2 : y = 2x\}$ do espaço linear $\mathbb{R}^2$ é gerado por qualquer dos seguintes conjuntos de vectores:

$$\{(1,2)\}, \quad \{(-2,-4)\} \quad \text{e} \quad \{(77,154)\}.$$

(iii) O espaço linear P_n de todos os polinómios de grau menor ou igual a n, é gerado por qualquer dos seguintes conjuntos de vectores:

$$\{1,t,t^2,...,t^n\}, \quad \{1,1+t,(1+t)^2,...,(1+t)^n\} \quad \text{e} \quad \{1,\frac{t}{1!},\frac{t^2}{2!},...,\frac{t^n}{n!}\}.$$

(iv) O espaço linear P de todos os polinómios, é gerado pelo conjunto infinito de vectores:

$$\{1,t,t^2,...\}.$$

(v) O espaço linear V de todas as funções $f : \mathbb{R} \to \mathbb{R}$ diferenciáveis tais que $f'(x) = af(x)$ é gerado pela função $f_1(x) = e^{ax}$, i.e. $V = L(\{f_1\})$.

(vi) Seja A uma matriz (real) do tipo $m \times n$. O espaço das colunas de A,

$$\mathcal{C}(A) = \{b \in \mathbb{R}^m : Au = b \text{ tem pelo menos uma solução } u\},$$

é o subespaço (do espaço linear $\mathbb{R}^m$) gerado pelas colunas de A, uma vez que:

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = u_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + u_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + ... + u_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

(vii) Seja A uma matriz (real) do tipo $m \times n$. Ao subespaço linear de $\mathbb{R}^n$ gerado pelas linhas de A dá-se o nome de **espaço das linhas** de A e designa-se por $\mathcal{L}(A)$.

(viii) Sejam

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -3 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & 2 \\ 2 & -4 \\ -2 & 4 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}.$$

Tem-se

$$\mathcal{C}(A) = \{(0,0)\}, \quad \mathcal{N}(A) = \mathbb{R}^3 \quad \text{e} \quad \mathcal{L}(A) = \{(0,0,0)\}.$$

$$\mathcal{C}(B) = L\left(\{(1,0,0),(1,7,0)\}\right), \quad \mathcal{N}(B) = L\left(\{(3,1,0)\}\right) \quad \text{e} \quad \mathcal{L}(B) = L\left(\{(1,-3,1),(0,0,7)\}\right).$$

$$\mathcal{C}(C) = L\left(\{(-1,2,-2)\}\right), \quad \mathcal{N}(C) = L\left(\{(2,1)\}\right) \quad \text{e} \quad \mathcal{L}(C) = L\left(\{(-1,2)\}\right).$$

$$\mathcal{C}(D) = L\left(\{(2,0),(0,-1)\}\right), \quad \mathcal{N}(D) = \{(0,0)\} \quad \text{e} \quad \mathcal{L}(D) = L\left(\{(2,0),(0,-1)\}\right).$$

(ix) Seja $U = \{A \in \mathcal{M}_{3\times 2}(\mathbb{R}) : a_{12} = a_{21} = a_{32} = 0 \text{ e } a_{11} + 2a_{31} = 0\}$. Tem-se, para $A \in U$,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} -2a_{31} & 0 \\ 0 & a_{22} \\ a_{31} & 0 \end{bmatrix} = a_{31} \begin{bmatrix} -2 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

com $a_{31}, a_{22} \in \mathbb{R}$. Logo,

$$U = L\left(\left\{\begin{bmatrix} -2 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\right\}\right).$$

(x) Seja $U = \{p(t) = a_0 + a_1 t + a_2 t^2 \in P_2 : p(1) = p(0)\}$. Tem-se, para $p(t) \in U$,

$$p(1) = p(0) \iff a_0 + a_1 + a_2 = a_0 \iff a_1 + a_2 = 0 \iff a_1 = -a_2.$$

Logo,

$$p(t) = a_0 - a_2 t + a_2 t^2 = a_0 1 + a_2\left(-t + t^2\right),$$

com $a_0, a_2 \in \mathbb{R}$. Assim,

$$U = L\left(\left\{1, -t + t^2\right\}\right).$$

Teorema 3.15 Se U e V são subespaços do espaço linear W, então:

(i) O conjunto $U \cap V$ é um subespaço linear de W.

(ii) O conjunto $U + V = \{u + v : u \in U \text{ e } v \in V\}$ é um subespaço de W. É o menor subespaço de W que contém $U \cup V$. O conjunto $U \cup V$ em geral não é um subespaço. Tem-se $U + V = L(U \cup V)$.

Se $U \cap V = \{0\}$ então dizemos que U e V estão em soma directa e escrevemos $U \oplus V$ para designar $U + V$.

Exemplo 3.16 (i) Em $\mathbb{R}^3$, considere os subespaços:

$$U = \{(x, y, z) \in \mathbb{R}^3 : x + y - 2z = 0\} \quad \text{e} \quad V = L\left(\{(1, 1, -1), (1, 2, 1)\}\right).$$

Seja $v \in V$, então

$$v = \alpha(1, 1, -1) + \beta(1, 2, 1) = (\alpha + \beta, \alpha + 2\beta, -\alpha + \beta),$$

com $\alpha, \beta \in \mathbb{R}$. Para que v esteja também em U é preciso que:

$$(\alpha + \beta) + (\alpha + 2\beta) - 2\left(-\alpha + \beta\right) = 0.$$

A última equação é equivalente a $4\alpha + \beta = 0 \iff \beta = -4\alpha$. Logo,

$$U \cap V = \{(-3\alpha, -7\alpha, -5\alpha) : \alpha \in \mathbb{R}\} = \{\alpha(-3, -7, -5) : \alpha \in \mathbb{R}\} = L\left(\{(3, 7, 5)\}\right).$$

(ii) Em $\mathbb{R}^3$, considere os subespaços:

$$U = L\left(\{(1, -1, 1), (1, 2, 2)\}\right) \quad \text{e} \quad V = L\left(\{(2, 1, 1), (-1, 1, 3)\}\right).$$

Seja $v \in U$, então

$$v = \alpha(1, -1, 1) + \beta(1, 2, 2) = (\alpha + \beta, -\alpha + 2\beta, \alpha + 2\beta),$$

com $\alpha, \beta \in \mathbb{R}$. Para que v esteja também em V é preciso que:

$$\begin{aligned}(\alpha+\beta, -\alpha+2\beta, \alpha+2\beta) &= \lambda(2,1,1)+\mu(-1,1,3) = \\ &= (2\lambda-\mu, \lambda+\mu, \lambda+3\mu)\,,\end{aligned}$$

com $\lambda, \mu \in \mathbb{R}$. Deste modo,

$$\begin{cases} \alpha+\beta = 2\lambda-\mu \\ \\ -\alpha+2\beta = \lambda+\mu \\ \\ \alpha+2\beta = \lambda+3\mu. \end{cases}$$

Considerando a matriz aumentada tem-se

$$\left[\begin{array}{ccc} 1 & 1 & | & 2\lambda-\mu \\ -1 & 2 & | & \lambda+\mu \\ 1 & 2 & | & \lambda+3\mu \end{array}\right] \underset{\substack{L_1+L_2\to L_2 \\ -L_1+L_3\to L_3}}{\longrightarrow} \left[\begin{array}{ccc} 1 & 1 & | & 2\lambda-\mu \\ 0 & 3 & | & 3\lambda \\ 0 & 1 & | & -\lambda+4\mu \end{array}\right] \underset{-\frac{1}{3}L_2+L_3\to L_3}{\longrightarrow} \left[\begin{array}{ccc} 1 & 1 & | & 2\lambda-\mu \\ 0 & 3 & | & 3\lambda \\ 0 & 0 & | & -2\lambda+4\mu \end{array}\right]$$

Logo,

$$\begin{cases} \alpha+\beta = 2\lambda-\mu \\ \\ \beta = \lambda \\ \\ 0 = -2\lambda+4\mu. \end{cases} \Longleftrightarrow \begin{cases} \alpha = \mu \\ \\ \beta = 2\mu \\ \\ \lambda = 2\mu. \end{cases}$$

Assim,

$$\alpha(1,-1,1)+\beta(1,2,2) = \mu(1,-1,1)+2\mu(1,2,2) = (3\mu,3\mu,5\mu) = \mu(3,3,5).$$

Logo,

$$U\cap V = \{(3\mu,3\mu,5\mu) : \mu\in\mathbb{R}\} = \{\mu(3,3,5) : \mu\in\mathbb{R}\} = L\left(\{(3,3,5)\}\right).$$

Observação 3.17 Neste exemplo (ii), os subespaços U e V poderiam ter sido apresentados inicialmente na forma:

$$U = \{(x,y,z)\in\mathbb{R}^3 : 4x+y-3z=0\} \quad \text{e} \quad V = \{(x,y,z)\in\mathbb{R}^3 : 2x-7y+3z=0\},$$

uma vez que

$$U = \{(x,y,z)\in\mathbb{R}^3 : 4x+y-3z=0\} = L\left(\{(1,-4,0),(0,3,1)\}\right) = L\left(\{(1,-1,1),(1,2,2)\}\right)$$

e

$$V = \{(x,y,z)\in\mathbb{R}^3 : 2x-7y+3z=0\} = L\left(\{(7,2,0),(-3,0,2)\}\right) = L\left(\{(2,1,1),(-1,1,3)\}\right).$$

(iii) Sejam $W = \mathcal{M}_{n\times n}(\mathbb{R})$, U o subespaço (de W) das matrizes triangulares superiores, V o subespaço (de W) das matrizes triangulares inferiores. Então

$$U+V=W \quad \text{e} \quad U\cap V = \text{ subespaço das matrizes diagonais.}$$

(iv) Sejam $W=\mathbb{R}^2$, $U = L(\{(1,0)\})$ e $V = L(\{(0,1)\})$. O conjunto

$$U\cup V = \{(x,y)\in\mathbb{R}^2 : x=0 \vee y=0\}$$

não é um espaço linear:

$$\underbrace{(1,0)}_{\in U} + \underbrace{(0,1)}_{\in V} = (1,1) \notin U\cup V$$

Teorema 3.18 Se U e V subespaços do espaço linear W, então $U \cup V$ é subespaço de W se e só se $U \subset V$ ou $V \subset U$.

Teorema 3.19 Sejam W_1 e W_2 subespaços de um espaço linear V tais que

$$W_1 \cap W_2 = \{\mathbf{0}\}.$$

Se $V = W_1 + W_2$ então todo o vector $v \in V$ pode ser escrito de modo único na forma

$$v = w_1 + w_2$$

com $w_1 \in W_1$ e $w_2 \in W_2$. Neste caso escreve-se $V = W_1 \oplus W_2$ e diz-se que V é a **soma directa** dos espaços W_1 e W_2.

Teorema 3.20 O espaço das linhas $\mathcal{L}(A)$ e o núcleo $\mathcal{N}(A)$ de uma matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ mantêm-se invariantes por aplicação do método de eliminação de Gauss. Isto é, sendo A' a matriz em escada que se obtem de A por aplicação desse método, tem-se

$$\mathcal{L}(A) = \mathcal{L}(A') \text{ e } \mathcal{N}(A) = \mathcal{N}(A').$$

Observação 3.21 Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Se A' fôr a matriz em escada que se obtém de A por aplicação do método de eliminação de Gauss, tem-se

$$\mathcal{C}(A) \neq \mathcal{C}(A').$$

Teorema 3.22 Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Tem-se

$$\mathcal{C}(A) = \mathcal{L}(A^T) \text{ e } \mathcal{L}(A) \cap \mathcal{N}(A) = \{\mathbf{0}\}.$$

3.2 Independência linear

Definição 3.23 Seja V um espaço linear. Seja $S = \{v_1, v_2, ..., v_k\} \subset V$. Diz-se que o conjunto S é **linearmente dependente** se e só se algum dos vectores de S se escrever como combinação linear dos restantes, isto é, se e só se existir algum $i \in \{1, 2, ..., k\}$ e escalares $\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_k \in \mathbb{R}$ tais que

$$v_i = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + ... + \lambda_k v_k.$$

Definição 3.24 Seja V um espaço linear. Seja $S = \{v_1, v_2, ..., v_k\} \subset V$. Diz-se que o conjunto S é **linearmente independente** se e só se nenhum dos vectores de S se puder escrever como combinação linear dos restantes, isto é, se e só a única solução do sistema homogéneo

$$\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k = \mathbf{0}$$

fôr a solução trivial, ou seja, $\lambda_1 = \lambda_2 = ... = \lambda_k = 0$.

Se $V = \mathbb{R}^n$, sendo A a matriz cujas colunas sã os vectores de S, entaão S é **linearmente independente** se e só se $\mathcal{N}(A) = \{\mathbf{0}\}$ se e só se $\operatorname{car}(A) = k$.

Teorema 3.25 Seja A' uma matriz em escada de linhas.

(i) As colunas de A' que contêm pivôs são linearmente independentes.

(ii) As linhas não nulas de A' são linearmente independentes.

(iii) O n^o de linhas independentes e o n^o de colunas independentes (de A') são ambos iguais à característica de A'.

Observação 3.26 **(i)** Assim, atendendo ao teorema anterior, a independência linear de $S = \{v_1, v_2, ..., v_k\} \subset \mathbb{R}^n$ pode ser decidida aplicando o método de eliminação à matriz A cujas colunas são os vectores de S, de modo a colocá-la em escada de linhas. Sendo A' essa matriz em escada, tem-se pelo teorema 3.20

$$\mathcal{N}(A) = \mathcal{N}(A') \quad (*).$$

Uma vez que as colunas de A' que contêm pivôs são linearmente independentes então, devido a (*), as colunas de A nas posições correspondentes também serão linearmente independentes.

(ii) Em $\mathbb{R}$, quaisquer dois vectores são linearmente dependentes.

(iii) Em $\mathbb{R}^2$, dois vectores são linearmente independentes se não forem colineares.

(iv) Em $\mathbb{R}^3$, três vectores são linearmente independentes se não forem coplanares.

(v) Qualquer conjunto que contenha o vector nulo (elemento neutro) é linearmente dependente. Em particular, o conjunto $\{\mathbf{0}\}$, formado apenas pelo vector nulo, é linearmente dependente.

(vi) O conjunto vazio $\varnothing$ é linearmente independente.

Teorema 3.27 Sejam S_1 e S_2 dois subconjuntos finitos de um espaço linear, tais que $S_1 \subset S_2$.

(i) Se S_1 é linearmente dependente então S_2 também é linearmente dependente.

(ii) Se S_2 é linearmente independente então S_1 também é linearmente independente.

Observação 3.28 Sejam S_1 e S_2 dois subconjuntos finitos de um espaço linear, tais que $S_1 \subset S_2$.

(i) Se S_2 fôr linearmente dependente então S_1 tanto pode ser linearmente dependente como linearmente independente.

(ii) Se S_1 fôr linearmente independente então S_2 tanto pode ser linearmente dependente como linearmente independente.

Exemplo 3.29 Seja $S = \{(1,0,2),(2,0,4),(0,1,2)\}$. Tem-se

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 2 & 4 & 2 \end{bmatrix} \underset{-2L_1+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \underset{-2L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = A'.$$

Logo, como apenas existem dois pivôs e portanto uma variável livre, as três colunas de A são linearmente dependentes, isto é, o conjunto S é linearmente dependente. O subconjunto de S:

$$\{(1,0,2),(2,0,4)\}$$

também é linearmente dependente. No entanto, uma vez que a 1^a e 3^a colunas de A são independentes pois correspondem às colunas da matriz em escada A' que contêm os pivôs, o subconjunto de S:

$$\{(1,0,2),(0,1,2)\}$$

é linearmente independente.

3.3 Bases e dimensão de Espaços Lineares

Definição 3.30 Chama-se **base** de um espaço linear V a qualquer subconjunto $\mathcal{S}$ de V que verifique as duas condições:

(i) $\mathcal{S}$ gera V, isto é, $L(\mathcal{S}) = V$.

(ii) $\mathcal{S}$ é linearmente independente.

O seguinte resuldado foi provado por George Hamel[10].

Teorema 3.31 Qualquer espaço linear $V \neq \{\mathbf{0}\}$ tem pelo menos uma base.[11]

Observação 3.32 Qualquer espaço linear $V \neq \{\mathbf{0}\}$ tem um n^o infinito de bases. Por exemplo, se $\mathcal{S} = \{u_1, ..., u_k\}$ fôr uma base de V então para cada $\alpha \neq 0$ o conjunto $\{\alpha u_1, ..., \alpha u_k\}$ é também uma base de V.

Teorema 3.33 Todas as bases de um espaço linear $V \neq \{\mathbf{0}\}$ têm o mesmo n^o de vectores.

Dem.: Sejam $\{u_1, ..., u_m\}$ e $\{v_1, ..., v_n\}$ bases de V. Então $m = n$. Suponhamos que $n < m$ e cheguemos a uma contradição.

Como $u_1, ..., u_m \in V$ e $\{v_1, ..., v_n\}$ é uma base de V, existem escalares a_{ij}:

$$\begin{matrix} u_1 = a_{11}v_1 + a_{12}v_2 + ... + a_{1n}v_n \\ \vdots \\ u_m = a_{m1}v_1 + a_{m2}v_2 + ... + a_{mn}v_n \end{matrix} \quad , \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathcal{M}_{m\times n}$$

[10] George Hamel 1877—1954

[11] A prova deste teorema é dificíl e nela intervêm de maneira crucial o facto de os escalares envolvidos serem $\mathbb{R}$ ou $\mathbb{C}$, isto é, os escalares têm que ser um *corpo*, portanto têm que conter fracções (o que não sucede com os naturais ou inteiros).

O sistema homogéneo $A^T\mathbf{x} = \mathbf{0}$ é indeterminado porque $n < m$.

Seja $(\alpha_1, ..., \alpha_m)$ uma solução <u>não nula</u> desse sistema:

$$\begin{cases} a_{11}\alpha_1 + a_{21}\alpha_2 + ... + a_{m1}\alpha_m = 0 \\ a_{12}\alpha_1 + a_{22}\alpha_2 + ... + a_{m2}\alpha_n = 0 \\ \vdots \\ a_{1n}\alpha_1 + a_{2n}\alpha_2 + ... + a_{mn}\alpha_m = 0 \end{cases}.$$

Tem-se:

$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m =$

$$= \alpha_1(a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n) + \cdots + \alpha_m(a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n) =$$

$$(a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{m1}\alpha_m)v_1 + \cdots + (a_{1n}\alpha_1 + a_{2n}\alpha_2 + ... + a_{mn}\alpha_m)v_n =$$

$$= 0v_1 + \cdots + 0v_n = \mathbf{0},$$

isto é, os vectores $u_1, ..., u_m$ são <u>linearmente dependentes</u>, o que contradiz a hipótese de $\{u_1, ..., u_m\}$ ser uma base!!

O caso $m < n$ é análogo ao caso $n < m$. Conclusão: $n = m$. QED.

Definição 3.34 Chama-se **dimensão** de um espaço linear $V \neq \{\mathbf{0}\}$ ao n^o de vectores de uma base qualquer de V, e escreve-se $\dim V$. Se $V = \{\mathbf{0}\}$ então $\dim V = 0$ uma vez que o conjunto vazio $\varnothing$ é base de $\{\mathbf{0}\}$. Um espaço linear terá dimensão finita se uma sua base tiver um n^o finito de vectores.

Exemplo 3.35 (i) O conjunto $\{1\}$ é uma base de $\mathbb{R}$, chamada base canónica ou natural de $\mathbb{R}$. Logo,

$$\dim \mathbb{R} = 1.$$

(ii) O conjunto $\{(1,0),(0,1)\}$ é uma base de $\mathbb{R}^2$, chamada base canónica ou natural de $\mathbb{R}^2$. Logo,

$$\dim \mathbb{R}^2 = 2.$$

(iii) O conjunto $\{(1,0,0),(0,1,0),(0,0,1)\}$ é uma base de $\mathbb{R}^3$, chamada base canónica ou natural de $\mathbb{R}^3$. Logo,

$$\dim \mathbb{R}^3 = 3.$$

(iv) O conjunto

$$\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$$

é uma base de $\mathcal{M}_{2\times 3}(\mathbb{R})$, chamada base canónica ou natural de $\mathcal{M}_{2\times 3}(\mathbb{R})$. Logo,

$$\dim \mathcal{M}_{2\times 3}(\mathbb{R}) = 6.$$

(v) Tem-se

$$\dim \mathbb{R}^n = n \quad \text{e} \quad \dim \mathcal{M}_{m\times n}(\mathbb{R}) = mn.$$

(vi) O conjunto $\{1, t, t^2, ..., t^n\}$ é uma base de $\mathcal{P}_n$ (espaço linear de todos os polinómios reais de grau menor ou igual a n, incluindo o polinómio nulo), chamada base canónica ou natural de $\mathcal{P}_n$. Logo,

$$\dim \mathcal{P}_n = n + 1.$$

(vii) O conjunto $\{1, t, t^2, ...\}$ é uma base de $\mathcal{P}$ (espaço linear de todos os polinómios reais), chamada base canónica ou natural de $\mathcal{P}$. Logo,

$$\dim \mathcal{P} = \infty.$$

Exemplo 3.36 O conjunto dos números complexos $E = \mathbb{C}$ é um espaço linear tanto sobre $\mathbb{R}$ como sobre $\mathbb{C}$. Mais, $\dim_{\mathbb{C}}(E) = 1$ e $\{1\}$ é uma base; $\dim_{\mathbb{R}}(E) = 2$ e $\{1, i\}$ é uma base.

Observação 3.37 Chama-se **nulidade** à dimensão do núcleo ou espaço nulo de uma matriz A e escreve-se $\operatorname{nul} A$.

Teorema 3.38 Seja A uma matriz do tipo $m \times n$.

(i) Tem-se
$$\operatorname{car}(A) = \operatorname{car}(A^T), \qquad \dim \mathcal{C}(A) = \dim \mathcal{L}(A) = \operatorname{car} A.$$

(ii) Tem-se
$$\operatorname{car} A + \operatorname{nul} A = n.$$

Dem.: Provemos que $\dim(\mathcal{C}(A)) = \dim(\mathcal{L}(A))$. Podemos transformar a matriz A numa matriz em escada de linhas U. Seja $k = \operatorname{car}(A)$, $R_1, ...R_k$ as linhas não nulas de U e $L_1, ..., L_m$ as linhas de A. Como $\mathcal{L}(A) = \mathcal{L}(U)$, existem escalares c_{ij} tais que

$$L_1 = c_{11}R_1 + ... + c_{1k}R_k,$$
$$\vdots$$
$$L_m = c_{m1}R_1 + ... + c_{mk}R_k.$$

Para $i = 1, ..., m$, sejam a_{ij} e r_{ij} as componentes j das linhas L_i e R_i respectivamente. Assim tem-se,

$$a_{1j} = c_{11}r_{1j} + ... + c_{1k}r_{kj},$$
$$\vdots$$
$$a_{mj} = c_{m1}r_{1j} + ... + c_{mk}r_{kj},$$

ou matricialmente
$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} = r_{1j} \begin{bmatrix} c_{11} \\ \vdots \\ c_{m1} \end{bmatrix} + ... + r_{kj} \begin{bmatrix} c_{1k} \\ \vdots \\ c_{mk} \end{bmatrix}.$$

Como $\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$ é a coluna j de A, a última igualdade prova que $\dim\mathcal{C}(A) \leq \dim\mathcal{L}(A)$, uma vez que cada vector coluna de A é combinação linear de k vectores e $\dim\mathcal{L}(A) = \dim\mathcal{L}(U) = k$.

Aplicando esta desigualdade à matriz A^T obtém-se $\dim\mathcal{C}(A^T) \leq \dim\mathcal{L}(A^T)$, i.e. $\dim\mathcal{L}(A) \leq \dim\mathcal{C}(A)$. Portanto $\dim\mathcal{C}(A) = \dim\mathcal{L}(A)$. Fica assim demonstrado que $\dim(\mathcal{C}(A)) = \dim(\mathcal{L}(A))$. O resto da prova é agora fácil. QED.

Teorema 3.39 Sejam W_1 e W_2 dois subespaços de dimensão finita de um espaço linear V. Então,
$$\dim (W_1 + W_2) = \dim W_1 + \dim W_2 - \dim (W_1 \cap W_2).$$

Teorema 3.40 Sejam V um espaço linear de dimensão finita e W um subespaço de V.

(i) Seja $S = \{u_1, ..., u_k\} \subset V$. Se S é linearmente independente então S será um subconjunto de uma base de V e ter-se-á $\dim V \geq k$.

(ii) Se $\dim V = n$, então quaisquer m vectores de V, com $m > n$, são linearmente dependentes.

(iii) Se $\dim V = n$, então nenhum conjunto com m vectores de V, em que $m < n$, pode gerar V.

(iv) O subespaço W tem dimensão finita e $\dim W \leq \dim V$.

(v) Se $\dim W = \dim V$, então $W = V$.

(vi) Se $\dim V = n$, então quaisquer n vectores de V linearmente independentes constituem uma base de V.

(vii) Se $\dim V = n$, então quaisquer n vectores geradores de V constituem uma base de V.

Observação 3.41 O n^o de elementos de uma base de um espaço linear é igual ao n^o mínimo de vectores possam constituir um conjunto gerador desse espaço e é também igual ao n^o máximo de vectores que possam constituir um conjunto linearmente independente nesse espaço.

Exemplo 3.42 Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Como $\mathcal{L}(A)$ e $\mathcal{N}(A)$ são subespaços de $\mathbb{R}^n$ então

$$\mathcal{L}(A) + \mathcal{N}(A) = L\left(\mathcal{L}(A) \cup \mathcal{N}(A)\right)$$

é também um subepaço de $\mathbb{R}^n$. Por outro lado, atendendo a que

$$\mathcal{L}(A) \cap \mathcal{N}(A) = \{\mathbf{0}\}$$

(teorema 3.22), tem-se

$$\dim\left(\mathcal{L}(A) \cap \mathcal{N}(A)\right) = 0.$$

Assim,

$$\begin{aligned} \dim\left(\mathcal{L}(A) + \mathcal{N}(A)\right) &= \dim \mathcal{L}(A) + \dim \mathcal{N}(A) - \dim\left(\mathcal{L}(A) \cap \mathcal{N}(A)\right) = \\ &= \operatorname{car} A + \operatorname{nul} A - 0 = \\ &= n. \end{aligned}$$

Logo, pelo teorema 3.40 (v), tem-se

$$\mathbb{R}^n = \mathcal{L}(A) \oplus \mathcal{N}(A),$$

isto é $\mathbb{R}^n = \mathcal{L}(A) + \mathcal{N}(A)$ e $\mathcal{L}(A) \cap \mathcal{N}(A) = \{0\}$.

Exemplo 3.43 (i) Os seguintes conjuntos são **todos** os subespaços de $\mathbb{R}$:

$$\{0\} \text{ e } \mathbb{R}.$$

(ii) Os seguintes conjuntos são **todos** os subespaços de $\mathbb{R}^2$:

$$\{(0,0)\}, \text{ todas as rectas que contêm a origem e } \mathbb{R}^2.$$

(iii) Os seguintes conjuntos são **todos** os subespaços de $\mathbb{R}^3$:

$\{(0,0,0)\}$, todas as rectas que contêm a origem, todos os planos que contêm a origem e $\mathbb{R}^3$.

Teorema 3.44 O método de eliminação de Gauss permite determinar a dimensão e uma base quer para o espaço das linhas $\mathcal{L}(A)$ quer para o espaço das colunas $\mathcal{C}(A)$ de uma matriz A. Seja A' a matriz em escada que se obtém de A por aplicação do método de eliminação de Gauss. Então,

(i) Uma base para $\mathcal{L}(A)$ será formada pelas linhas não nulas de A'.

(ii) Uma base para $\mathcal{C}(A)$ será formada pelas colunas de A que correspondem às posições das colunas de A' que contêm os pivôs.

Dem.: A prova de (i) é fácil. Provemos a parte (ii). Considere uma sequência de operações elementares que transforma a matriz numa matriz em escada de linhas U:

$$A \to \cdots \to U \qquad (*)$$

Seja k=car(A). Claro que $\dim(\mathcal{C}(A)) = k$ e portanto k vectores linearmente independentes em $\mathcal{C}(A)$ formam uma base de $\mathcal{C}(A)$.

Seja A' a matriz que se obtém de A removendo as colunas de A que correspondem às colunas de U sem pivô. (Analogamente, podemos definir a matriz U'). Assim, A' e U' são matrizes $m \times k$. Além disso, podemos usar a mesma sequência de operações de (*) e concluir que $A' \to U'$ e car$(A') = k$, isto é, o conjunto de todas as k colunas de A' é linearmente independente. Q.E.D.

Exemplo 3.45 Seja

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

Tem-se

$$A = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 2 & 3 & 3 \\ -6 & -3 & 1 & 1 \end{bmatrix} \underset{\substack{-2L_1+L_2 \to L_2 \\ 3L_1+L_3 \to L_3}}{\longrightarrow} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 4 & 4 \end{bmatrix} \underset{-4L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = A'.$$

Logo, $\{(2,1,1,1),(0,0,1,1)\}$ é uma base de $\mathcal{L}(A)$ e $\{(2,4,-6),(1,3,1)\}$ é uma base de $\mathcal{C}(A)$. Assim,

$$\dim \mathcal{L}(A) = 2 = \dim \mathcal{C}(A)$$

e

$$\mathcal{L}(A) = L\left(\{(2,1,1,1),(0,0,1,1)\}\right), \ \mathcal{C}(A) = L\left(\{(2,4,-6),(1,3,1)\}\right).$$

Por outro lado,

$$\begin{aligned}
\mathcal{N}(A') &= \left\{ (x,y,z,w) \in \mathbb{R}^4 : A' \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\} = \\
&= \{(x,-2x,-w,w) : x,w \in \mathbb{R}\} = \\
&= L\{(1,-2,0,0),(0,0,-1,1)\}.
\end{aligned}$$

Como o conjunto $\{(1,-2,0,0),(0,0,-1,1)\}$ é linearmente independente e gera $\mathcal{N}(A')$ então é uma base de $\mathcal{N}(A')$. Finalmente, uma vez que $\mathcal{N}(A) = \mathcal{N}(A')$, o conjunto

$$\{(1,-2,0,0),(0,0,-1,1)\}$$

é uma base de $\mathcal{N}(A)$ e portanto $\dim \mathcal{N}(A) = 2$, com

$$\mathcal{N}(A) = L\{(1,-2,0,0),(0,0,-1,1)\}.$$

Exemplo 3.46 Seja $S = \{1,2,-1),(2,1,1),(-1,-2,1),(0,1,0)\} \subset \mathbb{R}^3$. Determinemos uma base para $L(S)$.

Considere a seguinte matriz cujas colunas são os vectores de S:

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

Tem-se

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & -2 & 1 \\ -1 & 1 & 1 & 0 \end{bmatrix} \underset{\substack{-2L_1+L_2 \to L_2 \\ L_1+L_3 \to L_3}}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & -3 & 0 & 1 \\ 0 & 3 & 0 & 0 \end{bmatrix} \underset{L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & -3 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Logo, $S' = \{1,2,-1),(2,1,1),(0,1,0)\}$ é uma base de $L(S)$. Como $\dim \mathbb{R}^3 = 3$, então tem-se mesmo: $L(S) = \mathbb{R}^3$ e S' é uma base de $\mathbb{R}^3$.

Resolução alternativa: Considere a seguinte matriz cujas linhas são os vectores de S:

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

Tem-se

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ -1 & -2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underset{\substack{-2L_1+L_2 \to L_2 \\ L_1+L_3 \to L_3}}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \underset{L_3 \leftrightarrow L_4}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \underset{\frac{1}{3}L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Logo, $S' = \{1,2,-1),(0,-3,3),(0,0,1)\}$ é uma base de $L(S)$. Como $\dim \mathbb{R}^3 = 3$, então tem-se mesmo: $L(S) = \mathbb{R}^3$ e S' é uma base de $\mathbb{R}^3$.

Exemplo 3.47 Seja $S_{a,b} = \{1, 0, 1), (0, 1, a), (1, 1, b), (1, 1, 1)\} \subset \mathbb{R}^3$. Determinemos os valores dos parâmetros a e b para os quais $S_{a,b}$ não gere $\mathbb{R}^3$.

Considere a seguinte matriz cujas colunas são os vectores de S:

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & a & b & 1 \end{bmatrix}.$$

Tem-se

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & a & b & 1 \end{bmatrix} \underset{-L_1+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & a & b-1 & 0 \end{bmatrix} \underset{-aL_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & b-a-1 & -a \end{bmatrix}.$$

Logo, $S_{a,b}$ não gera $\mathbb{R}^3$ se e só se $b - a - 1 = 0$ e $-a = 0$, isto é, se e só se $a = 0$ e $b = 1$.

Teorema 3.48 **(i)** Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. As colunas de A geram $\mathbb{R}^m$ se e só se $\operatorname{car} A = m$.

(ii) Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. As colunas de A são linearmente independentes se e só se $\operatorname{car} A = n$.

(iii) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. A matriz A é invertível se e só se as colunas de A (ou as linhas de A) formarem uma base de $\mathbb{R}^n$. No caso de A ser invertível tem-se

$$\mathcal{C}(A) = \mathcal{L}(A) = \mathbb{R}^n.$$

Observação 3.49 Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e considere o sistema de equações lineares $Au = b$.

(i) O sistema $Au = b$ é **impossível** (não tem solução) se e só se $b \notin \mathcal{C}(A)$, isto é, se e só se $\operatorname{car} A < \operatorname{car}[A \mid b]$.

(ii) O sistema $Au = b$ é **possível e indeterminado** (tem um n^o infinito de soluções) se e só se $b \in \mathcal{C}(A)$ e as colunas de A forem linearmente dependentes, isto é, se e só se $\operatorname{car} A = \operatorname{car}[A \mid b] < n$, isto é, se e só se $\operatorname{car} A = \operatorname{car}[A \mid b]$ e $\operatorname{nul} A \neq 0$.

(iii) O sistema $Au = b$ é **possível e determinado** (tem uma única solução) se e só se $b \in \mathcal{C}(A)$ e as colunas de A forem linearmente independentes, isto é, se e só se $\operatorname{car} A = \operatorname{car}[A \mid b] = n$, isto é, se e só se $\operatorname{car} A = \operatorname{car}[A \mid b]$ e $\operatorname{nul} A = 0$.

Observação 3.50 Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e considere o sistema de equações lineares $Au = b$.

(i) Existência de solução: Se $m \leq n$ então o sistema $Au = b$ tem pelo menos uma solução u para cada $b \in \mathbb{R}^m$ se e só se $\operatorname{car} A = m$.

(ii) Unicidade de solução: Se $m \geq n$ então o sistema $Au = b$ tem no máximo uma solução u para cada $b \in \mathbb{R}^m$ se e só se $\operatorname{car} A = n$, isto é, se e só se $\operatorname{nul} A = 0$.

(iii) Existência e unicidade de solução: Se $m = n$ então o sistema $Au = b$ tem solução única u para cada $b \in \mathbb{R}^m$ se e só se A fôr invertível.

Teorema 3.51 Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. As seguintes afirmações são equivalentes.

(i) A é não singular.

(ii) A é invertível.

(iii) $\mathcal{N}(A) = \{\mathbf{0}\}$.

(iv) $\operatorname{nul} A = 0$.

(v) $Au = \mathbf{0}$ tem apenas a solução trivial $u = \mathbf{0}$.

(vi) $Au = b$ tem solução única u para cada $b \in \mathbb{R}^n$.

(vii) A característica de A é máxima, isto é, $\operatorname{car} A = n$.

(viii) As colunas de A geram $\mathbb{R}^n$.

(ix) As colunas de A são independentes.

(x) As linhas de A geram $\mathbb{R}^n$.

(xi) As linhas de A são independentes.

3.4 Coordenadas de um vector numa base

Definição 3.52 Seja $\mathcal{B} = \{v_1, v_2, ..., v_k\}$ uma base ordenada de um espaço linear V e seja u um vector de V. Chamam-se **coordenadas** do vector u na base ordenada $\mathcal{B}$ aos escalares $\lambda_1, \lambda_2, ..., \lambda_k$ da combinação linear:

$$u = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k.$$

Designamos por $u_{\mathcal{S}}$ as coordenadas de u na base $\mathcal{S}$, i.e. $u_{\mathcal{B}} = (\lambda_1, \cdots, \lambda_k)$.

Teorema 3.53 Seja V um espaço linear.

(i) Um conjunto $\mathcal{B}$ de vectores não nulos de V é uma base de V se e só se todo o vector de V puder ser escrito de modo único como combinação linear dos vectores de $\mathcal{B}$.

(ii) Se $\dim V = n$, então dados $u, w \in V$ e $\mathcal{S} = \{v_1, v_2, \ldots, v_n\}$ uma base ordenada de V, tem-se $u = w$ se e só se as coordenadas de u e de w na base $\mathcal{B}$ forem iguais.

Exemplo 3.54 (i) Sejam $B_1 = \{e_1, e_2\}$ e $B_2 = \{v_1, v_2\}$ duas bases de $\mathbb{R}^2$, onde $e_1 = (1,0), e_2 = (0,1), v_1 = (1,1)$ e $v_2 = (1,-1)$. Seja ainda $u = (11,3)$.
Então $u_{B_1} = (11,3)$ enquanto $u_{B_2} = (7,4)$.

(ii) Seja V o subespaço de $\mathbb{R}^3$ gerado pelos vectores $v_1 = (1,1,1)$ e $v_2 = (1,0,1)$. Claro que $B = \{v_1, v_2\}$ é uma base de F, uma vez que os vectores v_1, v_2 são linearmente independentes. Sendo $u = (3,1,3) \in F$, então as coordenadas u_B de u na base B são $u_B = (1,2)$, uma vez que $\lambda_1 = 1, \lambda_2 = 2$ é a única solução de $u = \lambda_1 v_1 + \lambda_2 v_2$.

(iii) Considerando a mesma base B de ii), sabendo que as coordenadas de um vector u na base B são $u_B = (2,1)$, então o vector $u = 2v_1 + 1v_2 = (3,2,3)$.

3.5 Matriz mudança de base

Teorema 3.55 Sejam $\mathcal{B}_1 = \{v_1, v_2, \ldots, v_n\}$ e $\mathcal{B}_2 = \{w_1, w_2, \ldots, w_n\}$ duas bases ordenadas de V. Seja $S_{\mathcal{S}_1 \to \mathcal{B}_2}$ a matriz cujas colunas são as coordenadas dos vectores de $\mathcal{B}_1$ em relação à base $\mathcal{B}_2$. Isto é,

$$S_{\mathcal{B}_1 \to \mathcal{B}_2} = (s_{ij})_{n \times n} \quad \text{com} \quad v_j = \sum_{i=1}^{n} s_{ij} w_i \quad \text{para todo o } j = 1, ..., n.$$

A matriz $S_{\mathcal{B}_1 \to \mathcal{B}_2}$ é invertível e chama-se **matriz de mudança de base** (da base $\mathcal{B}_1$ para $\mathcal{B}_2$). Assim, se tivermos

$$u = \sum_{i=1}^{n} \lambda_i v_i,$$

isto é, se $(\lambda_1, ..., \lambda_n)$ forem as coordenadas do vector u na base $\mathcal{B}_1$ então as coordenadas $(\mu_1, ..., \mu_n)$ de u na base $\mathcal{B}_2$ são dadas por

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = S_{\mathcal{B}_1 \to \mathcal{B}_2} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

Dem. Tem-se

$$u = \sum_{i=1}^{n} \mu_i w_i = \sum_{j=1}^{n} \lambda_j v_j = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{n} s_{ij} w_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} s_{ij} \lambda_j \right) w_i.$$

Atendendo ao teorema 3.53 (i), as coordenadas de um vector u numa base são únicas. Logo,

$$\mu_i = \left(\sum_{j=1}^{n} s_{ij} \lambda_j \right),$$

para todo o $i = 1, ..., n$. Isto é,

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = S_{\mathcal{B}_1 \to \mathcal{B}_2} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

Observação 3.56 Tem-se

$$S_{\mathcal{B}_2 \to \mathcal{B}_1} = \left(S_{\mathcal{B}_1 \to \mathcal{B}_2} \right)^{-1}.$$

Exemplo 3.57 Seja $\mathcal{B}_c = \{(1,0), (0,1)\}$ a base canónica de $\mathbb{R}^2$ e $\mathcal{B} = \{(1,2), (2,1)\}$ outra base ordenada de $\mathbb{R}^2$. Sejam $(2,3)$ as coordenadas de um vector u na base canónica $\mathcal{B}_c$ e determinemos as coordenadas de u na base $\mathcal{B}$ usando a matriz de mudança de base $S_{\mathcal{B}_c \to \mathcal{B}}$. Tem-se

$$S_{\mathcal{B}_c \to \mathcal{B}} = \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix},$$

uma vez que

$$(1,0) = -\frac{1}{3}(1,2) + \frac{2}{3}(2,1) \quad \text{e} \quad (0,1) = \frac{2}{3}(1,2) - \frac{1}{3}(2,1).$$

Logo, as coordenadas de u na base $\mathcal{B}$ são dadas por

$$S_{\mathcal{B}_c \to \mathcal{B}} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4/3 \\ 1/3 \end{bmatrix}.$$

Logo, $(4/3, 1/3)$ são as coordenadas de $(2, 3)$ na base ordenada $\mathcal{B}$, isto é

$$(2,3) = \frac{4}{3}(1,2) + \frac{1}{3}(2,1).$$

4 Valores Próprios, Vectores Próprios e diagonalização de Matrizes

Neste Capítulo, discutiremos equaçõs lineares da forma $A\mathbf{x} = \mathbf{x}$ e, mais geralmente, equações da forma $A\mathbf{x} = \lambda\mathbf{x}$, onde λ é um escalar. Estas equações aparecem numa variedade de aplicações importantes e constituem um tema recorrente ao longo da disciplina.

Definição 4.1 Seja A uma matriz $n \times n$. Chama-se a

$$p(\lambda) = \det(A - \lambda I),$$

o **polinómio característico** da matriz A.

Teorema 4.2 (**Teorema Fundamental da Álgebra$^{(*)}$**). Seja $p(\lambda)$ um polinómio de grau n com coeficientes em $\mathbb{C}$. Então $p(\lambda)$ tem n raízes, eventualmente repetidas.

Observação 4.3 **(i)** Note-se que $p(\lambda) = \det(A - \lambda I)$ é de facto um polinómio, de grau n, e o coeficiente do termo de grau n é $(-1)^n$ e o termo constante é $p(0) = \det A$.
(ii) Um polinómio com coeficientes em $\mathbb{R}$ nõ tem necessariamente todas as suas raízes (zeros) em $\mathbb{R}$ — exemplo: $p(\lambda) = 1 + \lambda^2$.

Definição 4.4 Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Chama-se **valor próprio** de A a qualquer escalar λ tal que $A - \lambda I$ seja singular (não invertível), isto é, tal que $\det(A - \lambda I) = 0$. Chama-se **vector próprio** de A, associado ao valor próprio λ de A, a qualquer vector não nulo v que verifique

$$(A - \lambda I)v = \mathbf{0}.$$

Observação 4.5 Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. O escalar $\lambda = 0$ é valor próprio de A se e só se A fôr singular. Isto é, a matriz A é invertível se e só se 0 não fôr valor próprio de A.

Definição 4.6 Sejam $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$. As matrizes A e B dizem-se **semelhantes** se existir uma matriz S invertível tal que

$$B = SAS^{-1}$$

Teorema 4.7 Sejam $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$. Se A e B forem semelhantes então A e B têm o mesmo polinómio característico. Em particular, se A e B forem semelhantes então A e B têm os mesmos valores próprios.

Dem. Tem-se

$$\begin{aligned}
\det(B-\lambda I) &= \det(SAS^{-1}-\lambda I) = \\
&= \det(SAS^{-1}-\lambda SS^{-1}) = \\
&= \det(S(A-\lambda I)S^{-1}) = \\
&= \det S \det(A-\lambda I)\det\left(S^{-1}\right) = \\
&= \det S \det(A-\lambda I)\frac{1}{\det S} = \\
&= \det(A-\lambda I).
\end{aligned}$$

Teorema 4.8 Seja A matriz $n \times n$. Sejam $\lambda_1, ..., \lambda_k$ valores próprios distintos de A e e $u_1, ..., u_k$ vectores próprios associados a cada um destes valores próprios, respectivamente. Então $u_1, ..., u_k$ são vectores linearmente independentes.

Definição 4.9 Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Se existir uma matriz P invertível tal que

$$D = PAP^{-1},$$

com D **matriz diagonal**, então diz-se que A é uma **matriz diagonalizável** e que S (matriz de mudança de base, ver mais sobre este assunto na Secção 6.1) é a **matriz diagonalizante**.

Teorema 4.10 Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. A matriz A é diagonalizável se e só se existir uma base de $\mathbb{R}^n$ constituída por vectores próprios de A. Neste caso, as entradas da diagonal principal dessa matriz diagonal serão os valores próprios associados aos vectores próprios da base de $\mathbb{R}^n$ pela ordem da mesma. Sendo $\lambda_1, ..., \lambda_n$ os valores próprios de A então a matriz diagonal é:

$$D = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & & 0 \\ 0 & ... & 0 & \lambda_n \end{bmatrix}.$$

Construindo uma base para cada espaço proóprio E_{λ_1}, $E_{\lambda_2}, ...$ obtém-se uma base de $\mathbb{R}^n$ juntando os vectores próprios de cada base de cada espaço próprio. Se A for diagonalizável, então temos uma base B_{vp} de $\mathbb{R}^n$ formada por vectores próprios de A. A coluna j da matriz P^{-1} é o vector número j de B_{vp} colocado em coluna.

O mesmo se aplica em $\mathbb{C}^n$.

Em particular, se A tiver n valores próprios distintos $\lambda_1, ..., \lambda_n$ então a matriz A é diagonalizável.

Observação 4.11 [12] Seja A a matriz $n \times n$.

(1) Seja $p(\lambda)$ o polinómio característico de A. Para cada raiz λ_1 de $p(\lambda)$, a sua multiplicidade enquanto raiz do polinómio chama-se **mutliplicidade algébrica** de λ_1 e denota-se por $ma(\lambda_1)$. Mais precisamente, λ_0 tem tem multiplicidade algébrica m quando

$$p(\lambda) = (\lambda - \lambda_1)^m q(\lambda)$$

e $q(\lambda_1) \neq 0$.

(2) Chamamos espectro de A ao conjunto de todos os valores próprios da matriz A, e denota-se por σ_A ou $\sigma(A)$.

(3) À dimensão de $\mathcal{N}(A - \lambda_1 I)$ chama-se **multiplicidade geométrica** e designa-se por $mg(\lambda_1)$.

Teorema 4.12 Seja λ_i valor próprio de uma matriz A matriz $n \times n$.

Então $0 \neq \mathrm{mg}(\lambda_i) \leq \mathrm{ma}(\lambda_i)$.

Dem.: Seja k a multiplicidade geométrica de λ_i (i.e. $\mathrm{mg}(\lambda_i)$=dim E_{λ_i}=dim $\mathcal{N}(A - \lambda_i I)$). Como $\dim(E_{\lambda_i}) \neq 0$, $\mathrm{mg}(\lambda_i) \neq 0$.

Seja $\{v_1, ..., v_k\}$ uma base de E_{λ_i}. Juntemos $n-k$ vectores de modo a obter n vectores linearmente independentes. Seja B a matriz cujas colunas são esses vectores. Então $B = [W \ Z]$ onde W é uma matriz $n \times k$ cujas colunas são $v_1, ..., v_k$. A matriz B é invertível, e de $B^{-1}B = I$ temos que $B^{-1}W = \begin{bmatrix} I_k \\ 0 \end{bmatrix}$. Como $Av_j = \lambda_i v_j$ para cada $j \in \{1, ..., k\}$, tem-se

$$AB = [Av_1 \ .... \ Av_k \ AZ] = [\lambda_1 v_1 \ .... \ \lambda_k v_k \ AZ] = [\lambda_j W \ AZ],$$

donde

$$B^{-1}AB = B^{-1}[\lambda_j W \ AZ] = [\lambda_j B^{-1}W \ B^{-1}AZ] = \begin{bmatrix} \lambda_j I_k & * \\ 0 & * \end{bmatrix}.$$

Calculando o polinómio característico desta última matriz temos que $(\lambda - \lambda_i)^k$ é um factor desse polinómio, e portanto $B^{-1}AB$ tem λ_r como valor próprio com multiplicidade algébrica pelo menos igual a k. Como A e $B^{-1}AB$ têm o mesmo polinómio característico, podemos concluir que $\mathrm{mg}(\lambda_i) \leq \mathrm{ma}(\lambda_i)$. Q.E.D.

A matriz $A \in \mathcal{M}_{n\times n}$ é diagonalizável se e só se

$$\sum_{\lambda \text{ valores próprios}} \dim\mathcal{N}(A - \lambda I) = \dim(V).$$

Ou seja, existe uma base de V na qual a representação matricial de T é uma matriz diagonal sse

$$\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k} = n,$$

onde $\lambda_1, \cdots, \lambda_k$ $(k \leq n)$ são os valores próprios de T.

Observação 4.13 i) Claro que qualquer matriz diagonal D é automaticamente diagonalizável, pondo $S = I$.
ii) Na Secção 6.3 pode encontrar mais matrizes diagonalizáveis.

[12] Para algumas aplicações, ver secções 7.3, 7.3.1, 7.4, 7.5

Exemplo 4.14 (i) Uma matriz com valores próprios distintos.

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

O polinómio característico é dado por

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1-\lambda & 5 & -1 \\ 0 & -2-\lambda & 1 \\ -4 & 0 & 3-\lambda \end{vmatrix} = \\ &= (1-\lambda)(-2-\lambda)(3-\lambda) - 20 + 4(2+\lambda) = \\ &= (1-\lambda)(-2-\lambda)(3-\lambda) + 4\lambda - 12 = \\ &= (3-\lambda)\left[(\lambda-1)(\lambda+2) - 4\right] = \\ &= (3-\lambda)\left(\lambda^2 + \lambda - 6\right) = \\ &= (3-\lambda)(\lambda-2)(\lambda+3). \end{aligned}$$

Os valores próprios de A são os valores de λ para os quais $\det(A - \lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 3, \quad \lambda_2 = 2 \quad \text{e} \quad \lambda_3 = -3.$$

Os vectores próprios de A associados ao valor próprio λ são os vectores não nulos $u \in \mathbb{R}^3$ para os quais

$$(A - \lambda I)\, u = 0,$$

isto é, são os vectores não nulos de $\mathcal{N}(A - \lambda I)$.

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$. Tem-se

$$\mathcal{N}(A - \lambda_1 I) = \mathcal{N}\left(\begin{bmatrix} -2 & 5 & -1 \\ 0 & -5 & 1 \\ -4 & 0 & 0 \end{bmatrix}\right) = L(\{(0,1,5)\}).$$

Logo, o subespaço próprio E_{λ_1} é dado por

$$E_{\lambda_1} = \mathcal{N}(A - \lambda_1 I) = L(\{(0,1,5)\}).$$

Os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$ são

$$u = (0, s, 5s), \ \text{ com } \ s \in \mathbb{R} \backslash \{0\}.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$. Tem-se

$$\mathcal{N}(A - \lambda_2 I) = \mathcal{N}\left(\begin{bmatrix} -1 & 5 & -1 \\ 0 & -4 & 1 \\ -4 & 0 & 1 \end{bmatrix}\right) = L(\{(1,1,4)\}).$$

Logo, o subespaço próprio E_{λ_2} é dado por

$$E_{\lambda_2} = \mathcal{N}(A - \lambda_2 I) = L(\{(1,1,4)\}).$$

Os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$ são

$$u = (s, s, 4s), \ \text{ com } \ s \in \mathbb{R} \backslash \{0\}.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_3 = -3$. Tem-se

$$\mathcal{N}(A - \lambda_3 I) = \mathcal{N}\left(\begin{bmatrix} 4 & 5 & -1 \\ 0 & 1 & 1 \\ -4 & 0 & 6 \end{bmatrix}\right) = L(\{(3, -2, 2)\}).$$

Logo, o subespaço próprio E_{λ_3} é dado por

$$E_{\lambda_3} = \mathcal{N}(A - \lambda_3 I) = L(\{(3, -2, 2)\}).$$

Os vectores próprios de A associados ao valor próprio $\lambda_3 = -3$ são

$$u = (3s, -2s, 2s), \text{ com } s \in \mathbb{R} \setminus \{0\}.$$

Atendendo a que os valores próprios de A são distintos, pelo teorema 4.8, os vectores próprios de A associados a esses valores próprios são linearmente independentes. Como $\dim \mathbb{R}^3 = 3$, então 3 vectores em $\mathbb{R}^3$ linearmente independentes formarão desde logo uma base de $\mathbb{R}^3$. Logo, o conjunto

$$\mathcal{S} = \{(0, 1, 5), (1, 1, 4), (3, -2, 2)\}$$

é uma base de $\mathbb{R}^3$. Deste modo, temos uma base de $\mathbb{R}^3$ formada só por vectores próprios de A. Logo, a matriz A é diagonalizável, isto é, existe uma matriz invertível S diagonalizante tal que a matriz PAP^{-1} é diagonal, tendo-se

$$D = PAP^{-1} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix},$$

com

$$P^{-1} = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 1 & -2 \\ 5 & 4 & 2 \end{bmatrix}.$$

Note que cada coluna de S^{-1} é formada pelo vector próprio associado ao valor próprio respectivo e na posição respectiva.

(ii) Uma matriz com valores próprios repetidos mas diagonalizável.

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

O polinómio característico é dado por

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 2 \\ 3 & 3 & 4-\lambda \end{vmatrix} = \\ &= (2-\lambda)(3-\lambda)(4-\lambda) + 6 + 6 - 3(3-\lambda) - 6(2-\lambda) - 2(4-\lambda) = \\ &= -\lambda^3 + 9\lambda^2 - 15\lambda + 7 = \\ &= -(\lambda - 1)(\lambda - 1)(\lambda - 7). \end{aligned}$$

Os valores próprios de A são os valores de λ para os quais $\det(A - \lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 1 \quad \text{e} \quad \lambda_2 = 7.$$

Os vectores próprios de A associados ao valor próprio λ são os vectores não nulos $u \in \mathbb{R}^3$ para os quais

$$(A - \lambda I)\, u = 0,$$

isto é, são os vectores não nulos de $\mathcal{N}(A - \lambda I)$.

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_1 = 1$. Tem-se

$$\mathcal{N}(A - \lambda_1 I) = \mathcal{N}\left(\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}\right) = L\left(\{(-1,1,0),(-1,0,1)\}\right).$$

Logo, o subespaço próprio E_{λ_1} é dado por

$$E_{\lambda_1} = \mathcal{N}(A - \lambda_1 I) = L\left(\{(-1,1,0),(-1,0,1)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_1 = 1$ são

$$u = (-s-t, s, t)\,, \text{ com } s,t \in \mathbb{R}, \text{ não simultâneamente nulos.}$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_2 = 7$. Tem-se

$$\mathcal{N}(A - \lambda_2 I) = \mathcal{N}\left(\begin{bmatrix} -5 & 1 & 1 \\ 2 & -4 & 2 \\ 3 & 3 & -3 \end{bmatrix}\right) = L\left(\{(1,2,3)\}\right).$$

Logo, o subespaço próprio E_{λ_2} é dado por

$$E_{\lambda_2} = \mathcal{N}(A - \lambda_2 I) = L\left(\{(1,2,3)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_2 = 7$ são

$$u = (s, 2s, 3s)\,, \text{ com } s \in \mathbb{R}\backslash\{0\}\,.$$

Atendendo a que

$$\dim E_{\lambda_1} + \dim E_{\lambda_2} = 3,$$

podemos ter a seguinte base de $\mathbb{R}^3$ formada só por vectores próprios de A

$$\mathcal{S} = \{(-1,1,0),(-1,0,1),(1,2,3)\}\,.$$

Logo, a matriz A é diagonalizável, isto é, existe uma matriz invertível P diagonalizante tal que a matriz PAP^{-1} é diagonal, tendo-se

$$D = PAP^{-1} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix},$$

com

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

Note que cada coluna de P^{-1} é formada pelo vector próprio associado ao valor próprio respectivo e na posição respectiva.

(iii) Uma matriz com valores próprios repetidos e não diagonalizável.

$$A = \begin{bmatrix} 7 & 5 & -1 \\ 0 & -2 & 1 \\ 20 & 0 & 3 \end{bmatrix}$$

O polinómio característico é dado por

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} 7-\lambda & 5 & -1 \\ 0 & -2-\lambda & 1 \\ 20 & 0 & 3-\lambda \end{vmatrix} = \\
&= (7-\lambda)(-2-\lambda)(3-\lambda) + 100 - 20(2+\lambda) = \\
&= (3-\lambda)\left[(7-\lambda)(-2-\lambda) + 20\right] = \\
&= (3-\lambda)\left(\lambda^2 - 5\lambda + 6\right) = \\
&= (3-\lambda)(\lambda-3)(\lambda-2).
\end{aligned}$$

Os valores próprios de A são os valores de λ para os quais $\det(A-\lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 3 \quad \text{e} \quad \lambda_2 = 2.$$

Os vectores próprios de A associados ao valor próprio λ são os vectores não nulos $u \in \mathbb{R}^3$ para os quais

$$(A - \lambda I)\, u = 0,$$

isto é, são os vectores não nulos de $\mathcal{N}(A - \lambda I)$.

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$. Tem-se

$$\mathcal{N}(A - \lambda_1 I) = \mathcal{N}\left(\begin{bmatrix} 4 & 5 & -1 \\ 0 & -5 & 1 \\ 20 & 0 & 0 \end{bmatrix}\right) = L\left(\{(0,1,5)\}\right).$$

Logo, o subespaço próprio E_{λ_1} é dado por

$$E_{\lambda_1} = \mathcal{N}(A - \lambda_1 I) = L\left(\{(0,1,5)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$ são

$$u = (0, s, 5s), \ \text{ com } \ s \in \mathbb{R} \backslash \{0\}.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$. Tem-se

$$\mathcal{N}(A - \lambda_2 I) = \mathcal{N}\left(\begin{bmatrix} 5 & 5 & -1 \\ 0 & -4 & 1 \\ 20 & 0 & 1 \end{bmatrix}\right) = L\left(\{(1,-5,-20)\}\right).$$

Logo, o subespaço próprio E_{λ_2} é dado por

$$E_{\lambda_2} = \mathcal{N}(A - \lambda_2 I) = L\left(\{(1,-5,-20)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$ são

$$u = (s, -5s, -20s), \ \text{ com } \ s \in \mathbb{R} \backslash \{0\}.$$

Atendendo a que

$$\dim E_{\lambda_1} + \dim E_{\lambda_2} = 2 < 3,$$

não é possível ter uma base de $\mathbb{R}^3$ formada só por vectores próprios de A. Logo, a matriz A não é diagonalizável, isto é, não existe uma matriz invertível P diagonalizante tal que a matriz PAP^{-1} seja diagonal.

(iv) Uma matriz com apenas um valor próprio real.

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

O polinómio característico é dado por

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{vmatrix} = \\ &= \lambda^2 (1-\lambda) + (1-\lambda) = \\ &= (1-\lambda)\left(\lambda^2 + 1\right). \end{aligned}$$

Os valores próprios de A são os valores de λ para os quais $\det(A - \lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 1, \quad \lambda_2 = i \quad \text{e} \quad \lambda_3 = -i.$$

Logo, a matriz A não é diagonalizável numa matriz de entradas reais, isto é, não existe uma matriz invertível P diagonalizante tal que a matriz PAP^{-1} seja diagonal com entradas reais. No entanto e atendendo a que os três valores próprios são distintos, a matriz A é diagonalizável numa matriz de entradas complexas:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}$$

5 Transformações Lineares

Definição 5.1 Sejam U e V espaços lineares. Diz-se que

$$T : U \to V$$

é uma **transformação linear** se e só se verificar as duas condições:

(i) $T(u+v) = T(u) + T(v)$, para todos os $u, v \in U$.
(ii) $T(\lambda u) = \lambda T(u)$, para todos os $u \in U$ e $\lambda \in \mathbb{R}$.

Observação 5.2 Sejam U e V espaços lineares. Sejam $\mathbf{0}$ o vector nulo de U e $\mathbf{0}'$ o vector nulo de V.

(i) Se $T : U \to V$ fôr uma transformação linear então $T(U)$ é um subespaço de V e além disso tem-se $T(\mathbf{0}) = \mathbf{0}'$. Logo, se T não verificar $T(\mathbf{0}) = \mathbf{0}'$ então T não será uma transformação linear.

(ii) $T : U \to V$ é uma transformação linear se e só se

$$T(\lambda u + \mu v) = \lambda T(u) + \mu T(v),$$

para todos os $\lambda, \mu \in \mathbb{R}$ e $u, v \in U$.

(iii) Seja $T : U \to V$ uma transformação linear e seja $\{v_1, v_2, \ldots, v_n\}$ uma base de U. Seja $u \in U$. Logo, existem $\lambda_1, \lambda_2, ..., \lambda_n \in \mathbb{R}$ tais que

$$u = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_n v_n.$$

Tem-se então

$$T(u) = \lambda_1 T(v_1) + \lambda_2 T(v_2) + ... + \lambda_n T(v_n).$$

Exemplo 5.3 Consideremos a base canónica $\{(1,0),(0,1)\}$ de $\mathbb{R}^2$. Seja $T : \mathbb{R}^2 \to \mathbb{R}$ uma transformação linear tal que $T(1,0) = 1$ e $T(0,1) = 1$.

Para qualquer $(x, y) \in \mathbb{R}^2$ tem-se

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

Então,

$$T(x, y) = T(x(1, 0) + y(0, 1)) = xT(1, 0) + yT(0, 1) = x + y.$$

Logo, $T : \mathbb{R}^2 \to \mathbb{R}$ é a transformação linear definida explicitamente por

$$T(x, y) = x + y.$$

Teorema 5.4 Sejam U e V espaços lineares e seja $\{v_1, v_2, \ldots, v_n\}$ uma base de U. Sejam $T_1, T_2 : U \to V$ duas transformações lineares.

$$\text{Se } T_1(v_i) = T_2(v_i) \text{ para todo o } i = 1, \ldots, n, \text{ então } T_1(u) = T_2(u),$$

para todo o $u \in U$, isto é, $T_1 = T_2$.

Exemplo 5.5 Sejam U e V espaços lineares e seja $\mathbf{0}$ o vector nulo de V.

(i) Seja $O : U \to V$ definida por

$$O(u) = \mathbf{0},$$

para todo o $u \in U$. O é uma transformação linear e chama-se **transformação nula**.

(ii) Seja $\lambda \in \mathbb{R}$. Seja $T_\lambda : U \to U$ definida por

$$T_\lambda(u) = \lambda u,$$

para todo o $u \in U$. T_λ é uma transformação linear. Se $\lambda = 1$ então chama-se a T_1 a **transformação identidade** e denota-se por I. Tem-se $I(u) = u$, para todo o $u \in U$.

(iii) Seja

$$\text{tr} : \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathbb{R}$$

definida por

$$\text{tr}(A) = a_{11} + a_{22} + ... + a_{nn} = \sum_{i=1}^{n} a_{ii},$$

para todo o $A = (a_{ij})_{n\times n} \in \mathcal{M}_{n\times n}(\mathbb{R})$. tr (traço) é uma transformação linear.

(iv) Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Seja

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

definida por

$$T(u) = Au,$$

para todo o $u \in \mathbb{R}^n$. T é uma transformação linear.

(v) Seja E o espaço das funções diferenciáveis. Então $T: E \to E$ definida por

$$T(f) = f'$$

é uma transformação linear.

Definição 5.6 Sejam U e V espaços lineares e $T_1, T_2 : U \to V$ transformações lineares. Seja $\lambda \in \mathbb{R}$. Sejam $T_1 + T_2, \lambda T_1 : U \to V$ definidas por

$$(T_1 + T_2)(u) = T_1(u) + T_2(u) \;\text{ e }\; (\lambda T_1)(u) = \lambda T_1(u),$$

para todo o $u \in U$.

Então $T_1 + T_2$ e λT_1 são transformações lineares.

Definição 5.7 Sejam U, V e W espaços lineares e, $T : U \to V$ e $S : V \to W$ transformações lineares. Seja $S \circ T$ (ou ST): $U \to W$ definida por

$$(S \circ T)(u) = S(T(u)),$$

para todo o $u \in U$. $S \circ T$ é uma transformação linear. Chama-se a $S \circ T$ (ou ST) a **composição de** S com T.

Observação 5.8 A composição de transformações lineares é uma transformação linear; no entanto $S \circ T \neq T \circ S$ (em geral).

Teorema 5.9 **(i)** Sejam $T : U \to V$, $S : V \to W$ e $R : W \to X$. Então, tem-se

$$R \circ (S \circ T) = (R \circ S) \circ T.$$

(ii) Sejam $R, S : U \to V$ e $T : V \to W$. Seja $\lambda \in \mathbb{R}$. Então, tem-se

$$T \circ (R + S) = T \circ R + T \circ S \;\text{ e }\; T \circ (\lambda R) = \lambda (T \circ R).$$

Se o contradomínio de Q estiver contido em U então

$$(R + S) \circ Q = R \circ Q + S \circ Q \;\text{ e }\; (\lambda R) \circ Q = \lambda (R \circ Q).$$

Definição 5.10 Define-se

$$T^0 = I \;\text{ e }\; T^k = T \circ T^{k-1}, \text{ para todo o } k = 1, 2, ...$$

Observação 5.11 Tem-se $T^{m+n} = T^m \circ T^n$ para todos os $m, n \in \mathbb{N}$.

5.1 Representação matricial de uma transformação linear

Teorema 5.12 Sejam U e V espaços lineares de dimensões finitas tais que $\dim U = n$ e $\dim V = m$. Sejam $\mathcal{S}_1 = \{u_1, u_2, \ldots, u_n\}$ e $\mathcal{S}_2 = \{v_1, v_2, \ldots, v_m\}$ duas bases ordenadas de U e V respectivamente. Seja $T : U \to V$ uma transformação linear. Considere-se a matriz $A = (a_{ij})_{m \times n} \in \mathcal{M}_{m \times n}(\mathbb{R})$ cuja coluna j, para cada $j = 1, ..., n$, é formada pelas coordenadas de $T(u_j)$ na base $\mathcal{S}_2$. Isto é,

$$T(u_j) = \sum_{i=1}^{m} a_{ij} v_i.$$

Chama-se a esta matriz A a **representação matricial** de T em relação às bases $\mathcal{S}_1$ e $\mathcal{S}_2$ e escreve-se

$$A = M(T; \mathcal{S}_1; \mathcal{S}_2).$$

Além disso, sendo $\alpha_1, \alpha_2, ..., \alpha_n$ as coordenadas de um vector $v \in U$ na base ordenada $\mathcal{S}_1$ então as coordenadas $\beta_1, \beta_2, ..., \beta_m$ de $T(v) \in V$ na base ordenada $\mathcal{S}_2$ são dadas por

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} = M(T; \mathcal{S}_1; \mathcal{S}_2) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Observação 5.13 (a) Seja V um espaço linear de dimensão finita, com $\dim V = n$. Sejam $\mathcal{S}_1 = \{u_1, u_2, \ldots, u_n\}$ e $\mathcal{S}_2 = \{v_1, v_2, \ldots, v_n\}$ duas bases ordenadas de V. A representação matricial da transformação identidade $I : V \to V$ em relação às bases $\mathcal{S}_1$ e $\mathcal{S}_2$ é igual à matriz de mudança da base $\mathcal{S}_1$ para $\mathcal{S}_2$. Isto é,

$$M(I; \mathcal{S}_1; \mathcal{S}_2) = S_{\mathcal{S}_1 \to \mathcal{S}_2}.$$

(b) Quando a base de partida e chegada coincidem $\mathcal{S}_2 = \mathcal{S}_1$, denota-se $M(T; \mathcal{S}_1; \mathcal{S}_2)$ simplesmente por $M(T; \mathcal{S}_1)$.

Teorema 5.14 Sejam $\mathcal{B}_c^n = \{e_1, e_2, \ldots, e_n\}$ e $\mathcal{B}_c^m = \{e'_1, e'_2, \ldots, e'_m\}$ as bases canónicas (ordenadas) de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Considere-se a matriz $A = (a_{ij})_{m\times n} = M(T; \mathcal{B}_c^n; \mathcal{B}_c^m) \in \mathcal{M}_{m\times n}(\mathbb{R})$ cuja coluna j, para cada $j = 1, ..., n$, é formada pelas coordenadas de $T(e_j)$ na base $\mathcal{B}_c^m$. Isto é,

$$T(e_j) = \sum_{i=1}^{m} a_{ij}e'_i = a_{1j}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + ... + a_{mj}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Então, tem-se, para todo o $u \in \mathbb{R}^n$,

$$T(u) = Au.$$

Dem. Seja $u \in \mathbb{R}^n$. Então, existem $\lambda_1, \lambda_2, ..., \lambda_n \in \mathbb{R}$ tais que

$$u = \lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_n e_n = \sum_{j=1}^{n} \lambda_j e_j.$$

Uma vez que, para todo o $j = 1, ..., n$,

$$T(e_j) = \sum_{i=1}^{m} a_{ij}e'_i,$$

tem-se

$$\begin{aligned} T(u) &= T\left(\sum_{j=1}^{n} \lambda_j e_j\right) \underset{\text{T é linear}}{=} \sum_{j=1}^{n} \lambda_j T(e_j) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij}e'_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\lambda_j\right) e'_i = \\ &= \left(\sum_{j=1}^{n} a_{1j}\lambda_j, ..., \sum_{j=1}^{n} a_{mj}\lambda_j\right) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = Au. \end{aligned}$$

Exemplo 5.15 (i) Seja $T : \mathbb{R}^4 \to \mathbb{R}^3$ definida por $T(x, y, z, w) = (3x + y - 2z, 0, x + 4z)$. T é uma transformação linear e a matriz $M(T; \mathcal{B}_c^4; \mathcal{B}_c^3)$ que representa T em relação às bases canónicas (ordenadas) $\mathcal{B}_c^4$ e $\mathcal{B}_c^3$ de $\mathbb{R}^4$ e $\mathbb{R}^3$ respectivamente, é dada por

$$M(T; \mathcal{B}_c^4; \mathcal{B}_c^3) = \begin{bmatrix} 3 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 4 & 0 \end{bmatrix},$$

uma vez que $T(1, 0, 0, 0) = (3, 0, 1)$, $T(0, 1, 0, 0) = (1, 0, 0)$, $T(0, 0, 1, 0) = (-2, 0, 4)$ e $T(0, 0, 0, 1) = (0, 0, 0)$. Tem-se então:

$$T(x, y, z, w) = M(T; \mathcal{B}_c^4; \mathcal{B}_c^3)\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}.$$

(ii) Sejam $\mathcal{S}_1 = \{1, t, t^2\}$ e $\mathcal{S}_2 = \{1, t, t^2, t^3\}$ as bases canónicas (ordenadas) de P_2 e P_3 respectivamente. Seja $D : P_2 \to P_3$ tal que $D(1) = 0$, $D(t) = 1$ e $D(t^2) = 2t$. D é uma transformação linear e a matriz $M(D;\mathcal{S}_1;\mathcal{S}_2)$ que representa D em relação às bases canónicas $\mathcal{S}_1$ e $\mathcal{S}_2$, é dada por

$$M(D;\mathcal{S}_1;\mathcal{S}_2) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Teorema 5.16 Sejam U e V espaços lineares e $T_1, T_2 : U \to V$ transformações lineares. Seja B_1 uma base de U, B_2 uma base de V e λ um escalar. Então Então temos:

$$M(T_1+T_2;B_1;B_2) = M(T_1;B_1;B_2)+M(T_2;B_1;B_2), \qquad M(\lambda T_1;B_1;B_2) = \lambda M(T_1;B_1;B_2).$$

Teorema 5.17 Sejam U, V e W espaços lineares de dimensões finitas. Sejam $\mathcal{S}_1, \mathcal{S}_2$ e $\mathcal{S}_3$ bases de U, V e W respectivamente. Sejam $T : U \to V$ e $S : V, \to W$ transformações lineares. Então, tem-se

$$M(S \circ T;\mathcal{S}_1;\mathcal{S}_3) = M(S;\mathcal{S}_2;\mathcal{S}_3)M(T;\mathcal{S}_1;\mathcal{S}_2).$$

Teorema 5.18 Seja V um espaço linear de dimensão finita. Seja $T : V \to V$ uma transformação linear. Sejam $\mathcal{S}_1$ e $\mathcal{S}_2$ duas bases ordenadas de V. Seja $M(T;\mathcal{S}_1;\mathcal{S}_1)$ a matriz que representa T em relação à base $\mathcal{S}_1$.

Então, a matriz $M(T;\mathcal{S}_2;\mathcal{S}_2)$ que representa T em relação à base $\mathcal{S}_2$, é dada por

$$M(T;\mathcal{S}_2;\mathcal{S}_2) = S_{\mathcal{S}_1\to\mathcal{S}_2} M(T;\mathcal{S}_1;\mathcal{S}_1) \left(S_{\mathcal{S}_1\to\mathcal{S}_2}\right)^{-1},$$

onde $S_{\mathcal{S}_1\to\mathcal{S}_2}$ é a matriz de mudança da base $\mathcal{S}_1$ para $\mathcal{S}_2$.

Além disso,

$$S_{\mathcal{S}_1\to\mathcal{S}_2} M(T;\mathcal{S}_1;\mathcal{S}_1) = M(T;\mathcal{S}_1;\mathcal{S}_2)$$

e

$$M(T;\mathcal{S}_2;\mathcal{S}_2) S_{\mathcal{S}_1\to\mathcal{S}_2} = M(T;\mathcal{S}_1;\mathcal{S}_2).$$

Isto é, o diagrama seguinte é comutativo.

$$\begin{array}{ccc}
(V,\mathcal{S}_1) & \underset{T}{\overset{M(T;\mathcal{S}_1;\mathcal{S}_1)}{\longrightarrow}} & (V,\mathcal{S}_1) \\
S_{\mathcal{S}_1\to\mathcal{S}_2} \downarrow I & & I \downarrow S_{\mathcal{S}_1\to\mathcal{S}_2} \\
(V,\mathcal{S}_2) & \underset{M(T;\mathcal{S}_2;\mathcal{S}_2)}{\overset{T}{\longrightarrow}} & (V,\mathcal{S}_2)
\end{array}$$

Teorema 5.19 (Caso geral.) Sejam U e V dois espaços lineares de dimensões finitas. Seja $T : U \to V$ uma transformação linear. Sejam $\mathcal{S}_1$ e $\mathcal{S}_1'$ duas bases ordenadas de U. Sejam $\mathcal{S}_2$ e $\mathcal{S}_2'$ duas bases ordenadas de V. Seja $M(T;\mathcal{S}_1;\mathcal{S}_2)$ a matriz que representa T em relação às bases $\mathcal{S}_1$ e $\mathcal{S}_2$.

Então, a matriz $M(T;\mathcal{S}_1';\mathcal{S}_2')$ que representa T em relação às bases $\mathcal{S}_1'$ e $\mathcal{S}_2'$, é dada por

$$M(T;\mathcal{S}_1';\mathcal{S}_2') = S_{\mathcal{S}_2\to\mathcal{S}_2'} M(T;\mathcal{S}_1;\mathcal{S}_2) \left(S_{\mathcal{S}_1\to\mathcal{S}_1'}\right)^{-1},$$

onde $S_{\mathcal{S}_2\to\mathcal{S}_2'}$ e $S_{\mathcal{S}_1\to\mathcal{S}_1'}$ são as matrizes de mudança das bases $\mathcal{S}_2$ para $\mathcal{S}_2'$ e de $\mathcal{S}_1$ para $\mathcal{S}_1'$ respectivamente.

Além disso,
$$S_{\mathcal{S}_2\to\mathcal{S}_2'}M(T;\mathcal{S}_1;\mathcal{S}_2)=M(T;\mathcal{S}_1;\mathcal{S}_2')$$
e
$$M(T;\mathcal{S}_1';\mathcal{S}_2')S_{\mathcal{S}_1\to\mathcal{S}_1'}=M(T;\mathcal{S}_1;\mathcal{S}_2').$$
Isto é, o diagrama seguinte é comutativo.
$$\begin{array}{ccc} (U,\mathcal{S}_1) & \overset{M(T;\mathcal{S}_1;\mathcal{S}_2)}{\underset{T}{\longrightarrow}} & (V,\mathcal{S}_2) \\ S_{\mathcal{S}_1\to\mathcal{S}_1'}\downarrow I & & I\downarrow S_{\mathcal{S}_2\to\mathcal{S}_2'} \\ (U,\mathcal{S}_1') & \underset{M(T;\mathcal{S}_1';\mathcal{S}_2')}{\overset{T}{\longrightarrow}} & (V,\mathcal{S}_2') \end{array}$$

Observação 5.20 As demonstrações dos Teoremas 5.18 e 5.19 resultam do Teorema 5.17.

Exemplo 5.21 Seja $T:\mathbb{R}^2\to\mathbb{R}^2$ definida por $T(x,y)=(y,x)$. T é uma transformação linear. A matriz $M(T;\mathcal{B}_c^2;\mathcal{B}_c^2)$ que representa T em relação à base canónica (ordenada) $\mathcal{B}_c^2$ de $\mathbb{R}^2$, é dada por
$$M(T;\mathcal{B}_c^2;\mathcal{B}_c^2)=\begin{bmatrix}0&1\\1&0\end{bmatrix}.$$
Seja $\mathcal{S}=\{(1,1),(-1,1)\}$ uma base ordenada de $\mathbb{R}^2$.

A matriz $M(T;\mathcal{S};\mathcal{S})$ que representa T em relação à base ordenada $\mathcal{S}$ de $\mathbb{R}^2$, é dada por
$$M(T;\mathcal{S};\mathcal{S})=\begin{bmatrix}1&0\\0&-1\end{bmatrix},$$
uma vez que $T(1,1)=(1,1)=1(1,1)+0(-1,1)$ e $T(-1,1)=(1,-1)=0(1,1)+(-1)(-1,1)$.

Vamos agora verificar que se tem
$$M(T;\mathcal{S};\mathcal{S})=S_{\mathcal{B}_c^2\to\mathcal{S}}M(T;\mathcal{B}_c^2;\mathcal{B}_c^2)\left(S_{\mathcal{B}_c^2\to\mathcal{S}}\right)^{-1}.$$
Uma vez que $(0,1)=\frac{1}{2}(1,1)+\frac{1}{2}(-1,1)$ e $(1,0)=\frac{1}{2}(1,1)-\frac{1}{2}(-1,1)$, tem-se então
$$S_{\mathcal{B}_c^2\to\mathcal{S}}=\begin{bmatrix}1/2&1/2\\1/2&-1/2\end{bmatrix}.$$
Logo,
$$\begin{aligned}S_{\mathcal{B}_c^2\to\mathcal{S}}M(T;\mathcal{B}_c^2;\mathcal{B}_c^2)\left(S_{\mathcal{B}_c^2\to\mathcal{S}}\right)^{-1} &= \begin{bmatrix}1/2&1/2\\1/2&-1/2\end{bmatrix}\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}1/2&1/2\\1/2&-1/2\end{bmatrix}^{-1}=\\ &= \begin{bmatrix}1/2&1/2\\-1/2&1/2\end{bmatrix}\begin{bmatrix}1&1\\1&-1\end{bmatrix}=\\ &= \begin{bmatrix}1&0\\0&-1\end{bmatrix}=\\ &= M(T;\mathcal{S};\mathcal{S}).\end{aligned}$$
Isto é,
$$M(T;\mathcal{S};\mathcal{S})=S_{\mathcal{B}_c^2\to\mathcal{S}}M(T;\mathcal{B}_c^2;\mathcal{B}_c^2)\left(S_{\mathcal{B}_c^2\to\mathcal{S}}\right)^{-1}.$$
Além disso,
$$S_{\mathcal{B}_c^2\to\mathcal{S}}M(T;\mathcal{B}_c^2;\mathcal{B}_c^2)=M(T;\mathcal{B}_c^2;\mathcal{S})$$
e
$$M(T;\mathcal{S};\mathcal{S})S_{\mathcal{B}_c^2\to\mathcal{S}}=M(T;\mathcal{B}_c^2;\mathcal{S}).$$

5.2 Transformações injectivas, sobrejectiva e bijectivas – equações lineares

Definição 5.22 **(i)** $T : U \to V$ diz-se **injectiva** se e só se

$$T(u) = T(w) \quad \Rightarrow \quad u = w,$$

para todos os $u, w \in U$, isto é, se e só se

$$u \neq w \quad \Rightarrow \quad T(u) \neq T(w),$$

para todos os $u, w \in U$.

(ii) $T : U \to V$ diz-se **sobrejectiva** se e só se

$$T(U) = V.$$

(iii) $T : U \to V$ diz-se **bijectiva** se e só se fôr injectiva e sobrejectiva.

Definição 5.23 Sejam U e V espaços lineares. Diz-se que U e V são isomorfos se e só se existir um **isomorfismo** entre U e V, isto é, se e só se existir uma transformação linear bijectiva $T : U \to V$.

Teorema 5.24 Sejam U e V dois espaços lineares de dimensões finitas. U e V são isomorfos se e só se $\dim U = \dim V$.

Teorema 5.25 Sejam U e V espaços lineares de dimensões finitas tais que $\dim U = \dim V$. Seja $T : U \to V$ uma transformação linear. Então, T é injectiva se e só se T é sobrejectiva.

Definição 5.26 Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear. Seja $\mathbf{0}$ o vector nulo de V.

(i) Chama-se **contradomínio** ou imagem de T ao conjunto

$$T(U) = \{T(u) : u \in U\},$$

que também se denota por $\mathcal{I}(T)$.

(ii) Chama-se **núcleo** ou espaço nulo de T ao conjunto

$$\mathcal{N}(T) = \{u \in U : T(u) = \mathbf{0}\}.$$

Teorema 5.27 Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear. Então, os conjuntos $\mathcal{N}(T)$ e $\mathcal{I}(T)$ são subespaços de U e V respectivamente.

Exemplo 5.28 Sejam U e V espaços lineares. Sejam $\mathbf{0}$ e $\mathbf{0}'$ os vectores nulos de U e V respectivamente.

(i) Considere a transformação nula $O : U \to V$ definida por

$$O(u) = \mathbf{0}',$$

para todo o $u \in U$. Tem-se

$$\mathcal{N}(O) = U \text{ e } \mathcal{I}(O) = \{\mathbf{0}'\}.$$

(ii) Considere a transformação identidade $I : U \to U$ definida por

$$I(u) = u,$$

para todo o $u \in U$. Tem-se

$$\mathcal{N}(I) = \{\mathbf{0}\} \text{ e } \mathcal{I}(I) = U.$$

Exemplo 5.29 Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Seja

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

definida por

$$T(u) = Au,$$

para todo o $u \in \mathbb{R}^n$. Tem-se

$$\mathcal{N}(T) = \mathcal{N}(A) \text{ e } \mathcal{I}(T) = \mathcal{C}(A).$$

Definição 5.30 Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear.

(i) Chama-se **característica** de T à dimensão de $\mathcal{I}(T)$, isto é,

$$\operatorname{car} T = \dim \mathcal{I}(T).$$

(ii) Chama-se **nulidade** de T à dimensão de $\mathcal{N}(T)$, isto é,

$$\operatorname{nul} T = \dim \mathcal{N}(T).$$

Teorema 5.31 Sejam U um espaço linear de dimensão finita e T uma transformação linear definida em U. Então, o subespaço $\mathcal{I}(T)$ tem dimensão finita e

$$\dim \mathcal{N}(T) + \dim \mathcal{I}(T) = \dim U.$$

Teorema 5.32 Sejam $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Sejam $\mathcal{B}_c^n$ e $\mathcal{B}_c^m$ as bases canónicas (ordenadas) de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente. Seja $A = M(T; \mathcal{B}_c^n; \mathcal{B}_c^m) \in \mathcal{M}_{m\times n}(\mathbb{R})$ a matriz que representa T em relação às bases $\mathcal{B}_c^n$ e $\mathcal{B}_c^m$. Tem-se então:

(i) $\dim \mathcal{N}(T) = \operatorname{nul} A$;

(ii) $\dim \mathcal{I}(T) = \operatorname{car} A$;

(iii) T é injectiva se e só se $\operatorname{nul} A = 0$, isto é, se e só se $\operatorname{car} A = n$;

(iv) T é sobrejectiva se e só se $\operatorname{car} A = m$.

Teorema 5.33 Seja $T : U \to V$ uma transformação linear entre dois espaços lineares U e V de dimensão finita. Seja ainda B_1 uma base de U e B_2 uma base de V e $A = M(T; B_1; B_2)$. Então:

1. $\mathcal{N}(T) = T(U) = \{u \in U : \ u_{B_1} \in \mathcal{N}(A)\}$, onde u_{B_1} designa as coordenadas de u na base B_1.

2. $\mathcal{I}(U) = \{v \in V : \ v_{B_2} \in \mathcal{C}(A)\}$, onde v_{B_2} designa as coordenadas de v na base B_2.

Exemplo 5.34 Considere a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^4$ tal que a sua representação matricial nas bases ordenadas $\mathcal{B}_1 = \{(1,2,0),(3,2,1),(2,1,0)\}$ e
$\mathcal{B}_2$=$\{(1,1,1,1),(0,1,1,1),(0,0,2,3),(0,0,0,4)\}$ de $\mathbb{R}^4$ e $\mathbb{R}^3$, respectivamente é

$$M(T;\mathcal{B}_1;\mathcal{B}_2) = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 11 & 19 \\ 4 & 16 & 28 \end{bmatrix}.$$

(a) Verifique se T é injectiva ou sobrejectiva.
(b) Determine uma base para o núcleo de T.
(c) Determine uma base pata o contradomínio de T.

Resolução: Ora aplicando o método de eliminação de Gauss, temos:

$$A := \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 11 & 19 \\ 4 & 16 & 28 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & 9 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

E portanto $\mathrm{car}(A) = 2$ e $\dim(\mathcal{N}(A)) = 1$. Logo T não é injectiva pois $\mathcal{N}(A) \neq 0$; e também não é sobrejectiva, pois $\dim(\mathcal{I}(T)) = 2 \neq \dim(\mathbb{R}^4)$.

Além disso $\{(1,-2,1\}$ é uma base de $\mathcal{N}(A)$ e $\{(1,2,3,4),(5,6,11,16)\}$ é uma base para $\mathcal{C}(A)$. Como $1(1,2,0) - 2(3,2,1) + 1(2,1,0) = (-3,-1,-2)$ concluimos que $\{(-3,-1,-2)\}$ é uma base de $\mathcal{N}(T)$. Finalmente como

$$1(1,1,1,1) + 2(0,1,1,1) + 3(0,0,2,3) + 4(0,0,0,4) = (1,3,9,28),$$

$$5(1,1,1,1) + 6(0,1,1,1) + 11(0,0,2,3) + 16(0,0,0,4) = (5,11,33,108),$$

concluímos que $\{(1,3,9,28),(5,11,33,108)\}$ é uma base para o contrdomínio de T, isto é $\mathcal{I}(T)$.

Definição 5.35 Diz-se que $T : U \to V$ é invertível se existir $S : T(U) \to U$ tal que

$$S \circ T = I_U \ \text{ e } \ T \circ S = I_{T(U)},$$

onde I_U e $I_{T(U)}$ são as funções identidade em U e $T(U)$ respectivamente. Chama-se a S a inversa de T e escreve-se

$$S = T^{-1}.$$

Teorema 5.36 Sejam U e V espaços lineares de dimensões finitas. Seja $T : U \to V$ uma transformação linear. Seja $\mathbf{0}$ o vector nulo de U. As seguintes afirmações são equivalentes.

(i) T é injectiva.

(ii) T é invertível e a inversa $T^{-1} : T(U) \to U$ é linear.

(iii) $\mathcal{N}(T) = \{\mathbf{0}\}$.

(iv) $\dim U = \dim T(U)$.

(v) T transforma vectores linearmente independentes de U em vectores linearmente independentes de V.

(vi) T transforma bases de U em bases de $T(U)$.

Teorema 5.37 Sejam U e V dois espaços lineares de dimensões finitas. Seja $T : U \to V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ duas bases ordenadas de U e V respectivamente. Seja $A = M(T; \mathcal{B}_1; \mathcal{B}_2)$ a matriz que representa T em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$.

Se $V = T(U)$ então T é invertível se e só se A fôr uma matriz quadrada não singular. Tem-se então

$$A^{-1} = M(T^{-1}; \mathcal{B}_2; \mathcal{B}_1),$$

isto é, A^{-1} será a matriz que representa T^{-1} em relação às bases $\mathcal{S}_2$ e $\mathcal{S}_1$.

Teorema 5.38 Seja $B = \{v_1, .., v_n\}$ uma base ordenada de uma espaço linear V de dimensão finita (sobre $\mathbb{R}$). Dado $u \in V$ sejam $(\lambda_1, ..., \lambda_n)$ as coordenadas de u na base B (isto é $u = \lambda_1 v_1 + ... + \lambda_n v_n$). Então

$$T : V \to \mathbb{R}^n$$

tal que $T(U) = (\lambda_1, .., \lambda_n)$ é uma transformação linear bijectiva.

Definição 5.39 Seja $T : U \to V$ uma transformação linear e $b \in V$ fixo. A equação $T(u) = b$ designa-se por equação linear associada a T e a b. Resolver a equação linear $T(u) = b$ é descrever o conjunto de todos of vectores $u \in U$ (caso existam) tais que $T(u) = b$.

Teorema 5.40 Sejam U e V espaços lineares. Seja $T : U \to V$ uma transformação linear. Seja $b \in V$. Então:

(i) Existência de solução: a equação linear $T(u) = b$ tem pelo menos uma solução u se e só se $b \in T(U)$;

(ii) Unicidade de solução: a equação linear $T(u) = b$ tem no máximo uma solução u se e só se T fôr injectiva;

(iii) Existência e unicidade de solução: equação linear $T(u) = b$ tem solução única u se e só se $b \in T(U)$ e T fôr injectiva.

Teorema 5.41 Sejam U e V espaços lineares. Seja $T : U \to V$ uma transformação linear. Seja $b \in V$. A conjunto solução S da equação linear $T(u) = b$ obtém-se somando a uma solução particular u_0 dessa equação linear ao conjunto soluçã S_0 da equação linear homogéneo $T(u) = \mathbf{0}$, isto é

$$S = u_0 + S_0.$$

5.3 Valores e vectores próprios de transformações lineares

Definição 5.42 Seja V espaço linear e $T : V \to V$ uma transformação linear. Diz-se que um escalar λ é um **valor próprio** de T se existir um vector não nulo $u \in V$ tal que

$$T(u) = \lambda u.$$

Aos vectores não nulos u que satisfazem a equação anterior chamam-se **vectores próprios** associados ao valor próprio λ. Dado um valor próprio λ de T, o conjunto

$$E_\lambda = \{u \in V : T(u) = \lambda u\}$$

é um subespaço linear de V. Chama-se a E_λ o **subespaço próprio** de T associado ao valor próprio λ.

Teorema 5.43 Sejam V um espaço linear e $\mathbf{0}$ o vector nulo de V. Seja $T : V \to V$ uma transformação linear.

(i) Um escalar λ é um valor próprio de T se e só se $\mathcal{N}(T - \lambda I) \neq \{\mathbf{0}\}$. Sendo λ um valor próprio de T, o subespaço próprio de T, associado ao valor próprio λ, é dado por

$$E_\lambda = \mathcal{N}(T - \lambda I).$$

(ii) Se o espaço linear V tiver dimensão finita e se $A = M(T; B, B)$ fôr uma matriz que representa T em relação a uma base B de V, então um escalar λ é um valor próprio de T se e só se esse escalar λ fôr solução da equação

$$\det(A - \lambda I) = 0,$$

i.e. λ fôr valor próprio de A.

Teorema 5.44 *Sejam $v_1, ..., v_k$ vectores próprios de uma transformação linear T associados a valores próprios distintos dois a dois. Então $v_1, ..., v_k$ são linearmente independentes.*

Prova: Seja $U = L(\{v_1, ..., v_k\})$, como $v_1, ..., v_k$ são vectores próprios de T, dado $u \in U$ temos $T(u) \in U$, pelo que podemos definir uma transformação linear $T : U \to U$ tal que $u \mapsto T(u)$. Vamos supor que $\dim(U) := m < k$. Por um lado T tem m valores próprios (eventualmente repetidos), por outro lado temos k valores próprios distintos. contradição. Q.E.D.

Observação 5.45 Se $A = M(T; B, B)$ representa uma transformação linear $T : V \to V$ na base B, então T diz-se diagonalizável se A o fôr. Neste caso, sendo B_{vp} a base de V constituída por vectores próprios de T, então:

$$D = PAP^{-1}$$

onde $P^{-1} = S_{B_{vp} \to B}$, e $D = M(T; B_{vp}, B_{vp})$ é a matriz diagonal cujas entradas da diagonal são os valores próprios de A (iguais aos de T).

6 Produtos Internos

Definição 6.1 Sejam V um espaço linear real e $\mathbf{0}$ o vector nulo de V. Chama-se **produto interno** em V à aplicação

$$\begin{aligned} \langle , \rangle : V \times V &\to \mathbb{R} \\ (u,v) &\to \langle u,v \rangle \end{aligned}$$

que verifique as três condições seguintes.

(i) Simetria: para todos os $u, v \in V$

$$\langle u,v \rangle = \langle v,u \rangle .$$

(ii) Linearidade: para todo o $v \in V$ (fixo) a aplicação

$$V \to \mathbb{R}$$

$$u \to \langle u,v \rangle$$

é linear.

(iii) Positividade: para todo o $u \in V$ tal que $u \neq \mathbf{0}$,

$$\langle u,u \rangle > 0.$$

Observação 6.2 Se V é uma espaço linear complexo, então $\langle , \rangle : V \times V \to \mathbb{C}$ é um produto interno se, os axiomas de definição anterior forem satisfeitos, com excepção ao simetria que é substituindo por:

$$\langle u,v \rangle = \overline{\langle v,u \rangle}.$$

onde $\overline{\langle v,u \rangle}$ designa o complexo conjugado de $\langle v,u \rangle$.

Definição 6.3 Chama-se **espaço euclidiano** a um espaço linear com um produto interno.

Observação 6.4 Seja V um espaço euclidiano real. Seja $\mathcal{S} = \{w_1, w_2, ..., w_n\}$ uma base de V. Sejam $u, v \in V$. Sejam

$$\alpha_1, \alpha_2, ..., \alpha_n \;\text{ e }\; \beta_1, \beta_2, ..., \beta_n$$

as coordenadas de u e de v na base $\mathcal{S}$ respectivamente, isto é,

$$u = \alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_n w_n = \sum_{i=1}^{n} \alpha_i w_i \quad \text{e} \quad v = \beta_1 w_1 + \beta_2 w_2 + ... + \beta_n w_n = \sum_{i=1}^{n} \beta_i w_i.$$

Logo,

$$\begin{aligned} \langle u,v \rangle &= \left\langle \sum_{i=1}^{n} \alpha_i w_i, \sum_{i=1}^{n} \beta_i w_i \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \beta_j \langle w_i, w_j \rangle = \\ &= \begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{bmatrix} \begin{bmatrix} \langle w_1,w_1 \rangle & \langle w_1,w_2 \rangle & \ldots & \langle w_1,w_n \rangle \\ \langle w_2,w_1 \rangle & \langle w_2,w_2 \rangle & \ldots & \langle w_2,w_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle w_n,w_1 \rangle & \langle w_n,w_2 \rangle & \ldots & \langle w_n,w_n \rangle \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}. \end{aligned}$$

Isto é, existe uma matriz simétrica e definida positiva (todos os seus valores próprios são positivos):

$$A = \begin{bmatrix} \langle w_1, w_1 \rangle & \langle w_1, w_2 \rangle & \dots & \langle w_1, w_n \rangle \\ \langle w_2, w_1 \rangle & \langle w_2, w_2 \rangle & \dots & \langle w_2, w_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle w_n, w_1 \rangle & \langle w_n, w_2 \rangle & \dots & \langle w_n, w_n \rangle \end{bmatrix} \quad \text{tal que} \quad \langle u, v \rangle = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{bmatrix} A \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

Teorema 6.5 [13] Seja V um espaço linear real com $\dim V = n$. Seja $\{w_1, w_2, ..., w_n\}$ uma base de V. Então, uma aplicação

$$\langle , \rangle : V \times V \to \mathbb{R}$$

é um produto interno (em V) se e só se

$$\langle u, v \rangle = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{bmatrix} A \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix},$$

com

$$u = \alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_n w_n, \quad v = \beta_1 w_1 + \beta_2 w_2 + ... + \beta_n w_n$$

e A é uma matriz simétrica cujos valores próprios são todos positivos. Se a aplicação $\langle , \rangle$ fôr um produto interno tem-se

$$A = \begin{bmatrix} \langle w_1, w_1 \rangle & \langle w_1, w_2 \rangle & \dots & \langle w_1, w_n \rangle \\ \langle w_2, w_1 \rangle & \langle w_2, w_2 \rangle & \dots & \langle w_2, w_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle w_n, w_1 \rangle & \langle w_n, w_2 \rangle & \dots & \langle w_n, w_n \rangle \end{bmatrix}.$$

Observação 6.6 i) No caso complexo, também podemos encontrar uma matriz A com entradas complexas tal que

$$\langle u, v \rangle = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{bmatrix} A \begin{bmatrix} \overline{\beta_1} \\ \overline{\beta_2} \\ \vdots \\ \overline{\beta_n} \end{bmatrix}$$

com os valores próprios de A todos positivos e $A = \overline{A}^T$, onde $\overline{A}$ é a matriz que se obtém de A passando todas as entradas de A ao complexo conjugado.
ii) O sinal dos valores próprios da matriz A é fulcral no estudo da Secção 7.1.

Exemplo 6.7 (i) Seja $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por:

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = \alpha_1 \beta_1 + \alpha_2 \beta_2,$$

[13] A prova deste resultado será feitas nas aulas teóricas quando for lecionado a Secção 6.3

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$. Esta aplicação é um produto interno em $\mathbb{R}^2$ a que se dá o nome de produto interno usual em $\mathbb{R}^2$, uma vez que

$$\langle(\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = \alpha_1\beta_1 + \alpha_2\beta_2 = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} A \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

com

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

A matriz A é simétrica e o único valor próprio de A é $1 > 0$.

(ii) Seja $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por:

$$\langle(\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = -2\alpha_1\beta_1 + 3\alpha_2\beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$. Esta aplicação não é um produto interno em $\mathbb{R}^2$, uma vez que

$$\langle(\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = -2\alpha_1\beta_1 + 3\alpha_2\beta_2 = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} A \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

com

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

A matriz A é simétrica, no entanto, os valores próprios de A: -2 e 3 não são ambos positivos.

Exemplo 6.8 $\mathbb{R}^2$ com um produto interno não usual. Seja $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por:

$$\langle(\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = 2\alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_2\beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$.

É fácil ver que esta aplicação é simétrica e linear em relação a (α_1, α_2) (fixando (β_1, β_2)). Vejamos por exemplo que a condição

$$\langle(\alpha_1, \alpha_2), (\alpha_1, \alpha_2)\rangle > 0, \quad \text{para todo o } (\alpha_1, \alpha_2) \neq (0, 0),$$

é satisfeita.

Atendendo a que

$$\langle(\alpha_1, \alpha_2), (\alpha_1, \alpha_2)\rangle = 2\alpha_1^2 + 2\alpha_1\alpha_2 + \alpha_2^2 = \alpha_1^2 + (\alpha_1 + \alpha_2)^2,$$

tem-se

$$\langle(\alpha_1, \alpha_2), (\alpha_1, \alpha_2)\rangle = 0 \Leftrightarrow (\alpha_1 = 0 \text{ e } \alpha_1 + \alpha_2 = 0) \Leftrightarrow (\alpha_1 = 0 \text{ e } \alpha_2 = 0) \Leftrightarrow (\alpha_1, \alpha_2) = (0, 0).$$

Em alternativa, podemos escrever

$$\langle(\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = 2\alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_2\beta_2 = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} A \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

com

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

A matriz A é simétrica e os valores próprios de A: $\frac{3+\sqrt{5}}{2}$ e $\frac{3-\sqrt{5}}{2}$ são ambos positivos.

Definição 6.9 Sejam V um espaço euclidiano e $\mathbf{0}$ o vector nulo de V. Sejam $u, v \in V$.

(i) Chama-se **norma** de u a:

$$\|u\| = \sqrt{\langle u, u\rangle}.$$

(ii) Chama-se **projecção ortogonal** de v sobre $u \neq \mathbf{0}$ a:

$$\operatorname{proj}_u v = \frac{\langle v, u\rangle}{\|u\|^2} u.$$

(iii) Diz-se que u e v são **ortogonais** se $\langle u, v\rangle = 0$.

(iv) Chama-se **ângulo** entre dois vectores não nulos u e v a:

$$\theta = \arccos \frac{\langle u, v\rangle}{\|u\| \, \|v\|}.$$

Observação 6.10 O ângulo θ entre dois vectores não nulos u e v é $\frac{\pi}{2}$ se e só se u e v são ortogonais.

Teorema 6.11 (**Desigualdade de Cauchy-Schwarz**) Seja $\langle \cdot, \cdot \rangle$ um produto interno num espaço linear V, então temos

$$|\langle u, v\rangle| \leq ||u|| \; ||v||, \qquad \text{para quaisquer } u, v \in V,$$

tendo-se a igualdade se e só se u e v forem linearmente dependentes.

Dem.: 1: Vamos supor que $v \neq 0$ (o caso $v = 0$ é trivial). Seja $\lambda := \frac{\langle u,v\rangle}{||v||^2}$ e observe que $\langle u - \lambda v, v\rangle = 0$, logo

$$\begin{aligned}
0 \leq ||u - \lambda v||^2 &= \langle u - \lambda v, u - \lambda v\rangle = \langle u - \lambda v, u\rangle - \overline{\lambda}\langle u - \lambda v, v\rangle \\
&= \langle u - \lambda v, u\rangle - 0 = \langle u, u\rangle - \lambda\langle v, u\rangle \\
&= \frac{\langle u, u\rangle\langle v, v\rangle - \langle u, v\rangle\langle v, u\rangle}{\langle v, v\rangle} \\
&= \frac{||u||^2 \; ||v||^2 - |\langle u, v\rangle|^2}{||v||^2} \qquad (*).
\end{aligned}$$

Como $0 < ||v||^2$ podemos concluir que $0 \leq ||u||^2 \; ||v||^2 - |\langle u, v\rangle|^2$, como pretendido.

Se u e v são linearmente dependentes então existe α tal que $u = \alpha v$; assim

$$|\langle u, v\rangle| = |\langle \alpha v, v\rangle| = |\alpha| \; \langle v, v\rangle = |\alpha| \; ||v||^2 = |\alpha| \; ||v|| \; ||v|| = ||\alpha v|| \; ||v|| = ||u|| \; ||v||.$$

Reciprocamente, se $|\langle u, v\rangle| = ||u|| \; ||v||$ então por (*) temos $||u - \lambda v||^2 = 0$, i.e. $u - \lambda v = 0$, logo $u = \lambda v$ como pretendido. Q.E.D.

Observação 6.12 (i) **Teorema de Pitágoras.** Sejam $u, v \in \mathbb{R}^2$. Tem-se u e v ortogonais se e só se

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 .$$

Dem.

$$\begin{aligned} \|u - v\|^2 &= \langle u - v, u - v\rangle = \langle u, u\rangle - \langle v, u\rangle - \langle u, v\rangle + \langle v, v\rangle = \\ &= \|u\|^2 - 2\langle u, v\rangle + \|v\|^2 = \|u\|^2 + \|v\|^2 \end{aligned}$$

se e só se

$$\langle u, v\rangle = 0,$$

isto é, se e só se u e v forem ortogonais.

(ii) Em $\mathbb{R}^2$ com o produto interno usual, a desigualdade de Cauchy-Schwarz é dada por

$$|\alpha_1\beta_1 + \alpha_2\beta_2| \leq \sqrt{\alpha_1^2 + \alpha_2^2}\sqrt{\beta_1^2 + \beta_2^2},$$

uma vez que

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = \alpha_1\beta_1 + \alpha_2\beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$.

(iii) Em $\mathbb{R}^n$ com o produto interno usual, a desigualdade de Cauchy-Schwarz é dada por

$$\left|\sum_{i=1}^{n} \alpha_i\beta_i\right| \leq \sqrt{\sum_{i=1}^{n} \alpha_i^2}\sqrt{\sum_{i=1}^{n} \beta_i^2},$$

uma vez que

$$\langle (\alpha_1, ..., \alpha_n), (\beta_1, ..., \beta_n)\rangle = \alpha_1\beta_1 + ... + \alpha_n\beta_n,$$

com $(\alpha_1, ..., \alpha_n), (\beta_1, ..., \beta_n) \in \mathbb{R}^n$.

Teorema 6.13 Sejam V um espaço euclidiano e $\mathbf{0}$ o vector nulo de V. Sejam $u.v \in V$ e $\lambda \in \mathbb{R}$. A norma satisfaz as seguintes propriedades.

(i) Positividade: $\|u\| > 0$ se $u \neq \mathbf{0}$.

(ii) Homogeneidade: $\|\lambda u\| = |\lambda| \|u\|$

(iii) Desigualdade triangular: $\|u + v\| \leq \|u\| + \|v\|$

Observação 6.14 Pode definir-se **norma** num espaço linear V, sem estar associada a qualquer produto interno, como sendo uma aplicação de V em $\mathbb{R}$ que satisfaz as propriedades do teorema anterior. A um espaço linear com uma norma chama-se **espaço normado**.

Observação 6.15 Seja V um espaço euclidiano. Sejam $u, v \in V$. Tem-se

$$\langle u, v\rangle = \frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right).$$

Observação 6.16 Seja V um espaço linear real normado. Sejam $u, v \in V$. Então, a norma pode ser obtida de um produto interno na forma

$$\|u\| = \sqrt{\langle u, u\rangle}$$

se e só se

$$\|u-v\|^2 + \|u+v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Esta última equação é conhecida por **lei do paralelogramo**.

6.1 Bases ortogonais

Definição 6.17 Sejam V um espaço euclidiano e $S \subset V$. Diz-se que S é **ortogonal** se para todos os $u, v \in S$ com $u \neq v$,

$$\langle u, v\rangle = 0.$$

Diz-se que S é **ortonormado** se fôr ortogonal e para todo o $u \in S$,

$$\|u\| = 1.$$

Teorema 6.18 Sejam V um espaço euclidiano e $S \subset V$. Seja $\mathbf{0}$ o vector nulo de V. Se S é ortogonal e $\mathbf{0} \notin S$ então S é linearmente independente. Em particular, se $n = \dim V$ então qualquer conjunto S ortogonal de n vectores não nulos é uma base de V.

Teorema 6.19 Seja V um espaço euclidiano com $\dim V = n$. Seja $\mathcal{S} = \{u_1, ..., u_n\}$ uma base ortogonal de V. Então, as coordenadas de um vector $v \in V$ em relação à base S são dadas por:

$$\alpha_j = \frac{\langle v, u_j\rangle}{\langle u_j, u_j\rangle},$$

com $j = 1, ..., n$. Isto é:

$$v = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle} u_1 + \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle} u_2 + ... + \frac{\langle v, u_n\rangle}{\langle u_n, u_n\rangle} u_n.$$

Teorema 6.20 Seja V um espaço euclidiano real com $\dim V = n$. Seja $\mathcal{S} = \{w_1, ..., w_n\}$ uma base ortonormada de V. Então, para todos os $u, v \in V$,

$$\langle u, v\rangle = \sum_{i=1}^{n} \langle u, w_i\rangle \langle v, w_i\rangle \quad \text{(fórmula de Parseval)} \quad \text{e} \quad \text{tem-se} \quad \|u\| = \sqrt{\sum_{i=1}^{n} \langle u, w_i\rangle^2}.$$

Observação 6.21 Seja V um espaço euclidiano real com $\dim V = n$. Seja $\mathcal{S} = \{w_1, ..., w_n\}$ uma base ortonormada de V. Sejam $u, v \in V$, com

$$u = \alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_n w_n, \quad v = \beta_1 w_1 + \beta_2 w_2 + ... + \beta_n w_n.$$

Então, atendendo ao teorema 6.19, a fórmula de Parseval é dada por:

$$\langle u, v \rangle = \sum_{i=1}^{n} \alpha_i \beta_i = \alpha_1 \beta_1 + \alpha_2 \beta_2 + ... + \alpha_n \beta_n \quad \text{e} \quad \text{tem-se} \quad \|u\| = \sqrt{\sum_{i=1}^{n} \alpha_i^2}.$$

Notação 6.22 Sejam V um espaço euclidiano e $\mathbf{0}$ o vector nulo de V. Para qualquer $v \in V$, com $v \neq \mathbf{0}$, o vector $\frac{1}{\|v\|} v$ será denotado por $\frac{v}{\|v\|}$.

Teorema 6.23 Método de ortogonalização de Gram-Schmidt[14]. Seja V um espaço euclidiano. Considere o conjunto linearmente independente:

$$\{v_1, v_2, ..., v_k\} \subset V.$$

Sejam

$$\begin{aligned}
u_1 &= v_1, \\
u_2 &= v_2 - \operatorname{proj}_{u_1} v_2 = v_2 - \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1, \\
u_3 &= v_3 - \operatorname{proj}_{u_2} v_3 - \operatorname{proj}_{u_1} v_3 = v_3 - \frac{\langle v_3, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 - \frac{\langle v_3, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 \\
&\ \ ... \\
u_k &= v_k - \operatorname{proj}_{u_1} v_k - ... - \operatorname{proj}_{u_{k-1}} v_k.
\end{aligned}$$

Então:

(i) $L(\{u_1, u_2, ..., u_k\}) = L(\{v_1, v_2, ..., v_k\})$

(ii) O conjunto $\{u_1, u_2, ..., u_k\}$ é uma base ortogonal de $L(\{v_1, v_2, ..., v_k\})$.

(iii) O conjunto $\left\{ \frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}, ..., \frac{u_k}{\|u_k\|} \right\}$ é uma base ortonormada de $L(\{v_1, v_2, ..., v_k\})$.

Exemplo 6.24 Considere-se $\mathbb{R}^4$ com o produto interno usual. Seja

$$U = L(\{(1, 1, -1, -1), (1, 2, 3, 4), (2, 1, -6, -7), (1, 3, 7, 9)\}).$$

Determinemos a dimensão de U e uma base ortonormada para U. Tem-se

[14] Jorgen Pedersen Gram 1850–1916. Erhard Schmidt 1876–1959

$$\begin{bmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \\ -1 & 3 & -6 & 7 \\ -1 & 4 & -7 & 9 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 4 & -4 & 8 \\ 0 & 5 & -5 & 10 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Logo, o conjunto $\{v_1, v_2\}$, com $v_1 = (1, 1, -1, -1)$ e $v_2 = (1, 2, 3, 4)$, é uma base de U e como tal $\dim U = 2$.

Sejam

$$u_1 = v_1 \quad \text{e} \quad u_2 = v_2 - \operatorname{proj}_{u_1} v_2.$$

Logo, o conjunto $\{u_1, u_2\}$, com $u_1 = (1, 1, -1, -1)$ e

$$u_2 = (1, 2, 3, 4) - \frac{1 + 2 - 3 - 4}{4}(1, 1, -1, -1) = (2, 3, 2, 3),$$

é uma base ortogonal de U. Uma base ortonormada para U:

$$\left\{ \frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|} \right\} = \left\{ \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2} \right), \left(\frac{\sqrt{26}}{13}, \frac{3\sqrt{26}}{26}, \frac{\sqrt{26}}{13}, \frac{3\sqrt{26}}{26} \right) \right\}$$

Teorema 6.25 Qualquer espaço euclidiano de dimensão finita tem uma base ortonormada.

Teorema 6.26 Seja $\{v_1, v_2, ..., v_n\}$ uma base de $\mathbb{R}^n$. Então, existe um único produto interno em $\mathbb{R}^n$ para o qual esta base é ortonormada.

Exemplo 6.27 Considere em $\mathbb{R}^2$ a base $\mathcal{S} = \{v_1, v_2\}$, com $v_1 = (1, 0)$ e $v_2 = (1, 1)$. Vejamos que existe um e um só produto interno para o qual a base $\mathcal{S}$ é ortonormada.

Seja $\mathcal{B}_c^2 = \{(1, 0), (0, 1)\}$ a base canónica de $\mathbb{R}^2$. Tem-se

$$S_{B_c^2 \to \mathcal{S}} = \left(S_{\mathcal{S} \to B_c^2} \right)^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Sejam $u, v \in \mathbb{R}^2$. Tem-se

$$u = (\alpha_1, \alpha_2) \quad \text{e} \quad v = (\beta_1, \beta_2),$$

onde α_1, α_2 e β_1, β_2 são as coordenadas na base $\mathcal{B}_c^2$ de u e v respectivamente. Seja $S = S_{B_c^2 \to \mathcal{S}}$. Logo, a aplicação $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2$ definida por

$$\langle u, v \rangle = (Su)^T A (Sv), \quad \text{com } A = \begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

é um produto interno e é o único para o qual a base $\mathcal{S}$ é ortonormada. Tem-se então

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = \alpha_1\beta_1 - \alpha_1\beta_2 - \alpha_2\beta_1 + 2\alpha_2\beta_2.$$

É fácil verificar que para este produto interno a base $\mathcal{S}$ é ortonormada:

$$\langle (1, 0), (1, 1) \rangle = 0 \quad \text{e} \quad \langle (1, 0), (1, 0) \rangle = \langle (1, 1), (1, 1) \rangle = 1.$$

6.2 Complementos e projecções ortogonais

Definição 6.28 Sejam V um espaço euclidiano e S um subespaço de V. Diz-se que um elemento de V é **ortogonal a** S se fôr ortogonal a todos os elementos de S. Ao conjunto de todos os elementos ortogonais a S chama-se **complemento ortogonal** de S e designa-se por $S^{\perp}$.

Teorema 6.29 Qualquer que seja o subespaço S de um espaço euclidiano V, também $S^{\perp}$ é um subespaço de V.

Exemplo 6.30 **(i)** Se $S \subset \mathbb{R}^3$ é um plano que passa pela origem, então $S^{\perp}$ é uma recta que passa pela origem e é perpendicular ao plano.

(ii) Se $S \subset \mathbb{R}^3$ é uma recta que passa pela origem, então $S^{\perp}$ é um plano que passa pela origem e é perpendicular à recta.

(iii) Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Então,

$$\mathcal{N}(A) = (\mathcal{L}(A))^{\perp}.$$

Teorema 6.31 Se S é um subespaço de dimensão finita de um espaço euclidiano V, então V é a soma directa de S e $S^{\perp}$, isto é, $V = S \oplus S^{\perp}$. Logo, cada elemento $v \in V$ pode ser escrito de modo único como soma de um elemento de S com um elemento de $S^{\perp}$:

$$v = v_S + v_{S^{\perp}}, \quad \text{com} \quad v_S \in S \quad \text{e} \quad v_{S^{\perp}} \in S^{\perp}.$$

À aplicação $P_S : V \to S$ definida por $P_S(v) = v_S$ chama-se **projecção ortogonal de** V **sobre** S e à aplicação $P_{S^{\perp}} : V \to S^{\perp}$ definida por $P_{S^{\perp}}(v) = v_{S^{\perp}}$ chama-se **projecção ortogonal de** V **sobre** $S^{\perp}$. Tem-se

$$I = P_S + P_{S^{\perp}}.$$

Se $\{v_1, v_2, ..., v_n\}$ é uma base ortonormada de S, então

$$P_S(v) = \sum_{i=1}^{n} \langle v, v_i \rangle \, v_i,$$

para todo o $v \in V$.

Se $\{u_1, u_2, ..., u_k\}$ é uma base ortonormada de $S^{\perp}$, então

$$P_{S^{\perp}}(v) = \sum_{j=1}^{k} \langle v, u_j \rangle \, u_j,$$

para todo o $v \in V$.

As aplicações P_S e $P_{S^{\perp}}$ são transformações lineares de V em V que satisfazem as propriedades:

(i) $P_S(V) = S, \quad P_{S^{\perp}}(V) = S^{\perp}$;

(ii) $(P_S)^2 = P_S, \quad (P_{S^{\perp}})^2 = P_{S^{\perp}}$;

(iii) $\langle P_S(u), v\rangle = \langle u, P_S(v)\rangle, \quad \langle P_{S^{\perp}}(u), v\rangle = \langle u, P_{S^{\perp}}(v)\rangle$, para todos os $u, v \in V$;

(iv) $\|u\|^2 = \|P_S(u)\|^2 + \|P_{S^{\perp}}(u)\|^2$, para todo o $u \in V$ (Teorema de Pitágoras);

Observação 6.32 Seja V um espaço euclidiano de dimensão finita e U é um subespaço linear de V.

(i) $\dim U + \dim U^\perp = \dim V$ onde $U^\perp$ designa o complemento ortogonal de U em V.

(ii) $\left(U^\perp\right)^\perp = U$

(iii) Seja $v \in V$. Se $\{u_1, v_2, ..., u_k\}$ é uma base de U então $v \in U^\perp$ se e só se

$$\langle v, u_1\rangle = \langle v, u_2\rangle = ... = \langle v, u_k\rangle = 0.$$

(iv) Nas condições de (iii), seja $V = \mathbb{R}^n$. Seja A a matriz $k \times n$ cuja linha i é igual ao vector v_i. Então $U = \mathcal{L}(A)$ e $U^\perp = \mathcal{N}(A)$.

Teorema 6.33 Seja S é um subespaço de dimensão finita de um espaço euclidiano V. Seja $v \in V$. Então, existe um **elemento de** S **mais próximo de** v do que qualquer dos outros pontos de S. **Este elemento é a projecção ortogonal** $P_S(v)$ **de** v **sobre** S e tem-se

$$\|v - P_S(v)\| \leq \|v - u\|,$$

para todo o $u \in S$, e a igualdade verifica-se se e só se $u = P_S(v)$.

Definição 6.34 Seja V um espaço euclidiano. Seja S é um subespaço de V com $\dim S = k$. Seja $q \in V$. Chama-se ao conjunto

$$\{q\} + S$$

um k-plano. A **distância** d **de um ponto** $p \in V$ **a um** k**-plano** $\mathcal{P} = \{q\} + S$ é dada por:

$$d(p, \mathcal{P}) = \|P_{S^\perp}(p - q)\|.$$

Observação 6.35 A distância entre dois k-planos paralelos $\mathcal{P}_1 = \{a\} + S$ e $\mathcal{P}_2 = \{b\} + S$ é dada por:

$$d(\mathcal{P}_1, \mathcal{P}_2) = \|P_{S^\perp}(a - b)\|.$$

Exemplo 6.36 Considere-se $\mathbb{R}^3$ com o produto interno usual.

(i) Seja $\mathcal{P}$ o plano (em $\mathbb{R}^3$) que passa pelos pontos: $(1,2,1), (1,0,-1)$ e $(1,1,1)$. Tem-se

$$\mathcal{P} = \{(1,2,1)\} + L(\{(0,-2,-2), (0,-1,0)\})$$

Equação vectorial de $\mathcal{P}$:

$$(x,y,z) = (1,2,1) + \alpha(0,-2,-2) + \beta(0,-1,0),$$

com $\alpha, \beta \in \mathbb{R}$.

Equações paramétricas de $\mathcal{P}$:

$$\begin{cases} x = 1 \\ \\ y = 2 + 2\beta - 2\alpha - \beta \\ \\ z = 1 - 2\alpha \end{cases}$$

com $\alpha, \beta \in \mathbb{R}$.

Equação cartesiana de $\mathcal{P}$:

$$x = 1.$$

Em alternativa, podemos determinar uma **equação cartesiana de** $\mathcal{P}$ do seguinte modo. Atendendo a que

$$\mathcal{P} = \{(1,2,1)\} + L\left(\{(0,-2,-2),(0,-1,0)\}\right),$$

seja

$$S = L\left(\{(0,-2,-2),((0,-1,0)\}\right).$$

Logo,

$$\begin{aligned} S^{\perp} &= \left\{(x,y,z) \in \mathbb{R}^3 : \langle (x,y,z),(0,-2,-2)\rangle = 0 \ \text{ e } \ \langle (x,y,z),(0,-1,0)\rangle = 0\right\} = \\ &= \mathcal{N}\left(\begin{bmatrix} 0 & -2 & -2 \\ 0 & -1 & 0 \end{bmatrix}\right) = L\left(\{(1,0,0)\}\right) \end{aligned}$$

e assim, a equação cartesiana do plano $\mathcal{P}$ que passa pelo ponto $(1,2,1)$ é dada por:

$$\left(\langle (x-1,y-2,z-1),(1,0,0)\rangle = 0\right) \Leftrightarrow$$

ou seja por

$$x = 1.$$

(ii) Determinemos a **equação cartesiana** da recta que passa pelos pontos $(1,1,0)$ e $(1,2,1)$. Tem-se

$$r = \{(1,1,0)\} + L\left(\{(0,1,1)\}\right),$$

uma vez que $(0,1,1) = (1,2,1) - (1,1,0)$. Seja

$$S = L\left(\{(0,1,1)\}\right).$$

Logo,

$$S^{\perp} = \left\{(x,y,z) \in \mathbb{R}^3 : \langle (x,y,z),(0,1,1)\rangle = 0\right\} = \mathcal{N}\left(\begin{bmatrix} 0 & 1 & 1 \end{bmatrix}\right) = L\left(\{(1,0,0),(0,1,-1)\}\right)$$

e assim, a equação cartesiana da recta r é dada por:

$$\left(\langle (x-1,y-1,z),(1,0,0)\rangle = 0 \ \text{ e } \ \langle (x-1,y-1,z),(0,1,-1)\rangle = 0\right) \Leftrightarrow$$

$$\Leftrightarrow \left(1\,(x-1) = 0 \ \text{ e } \ 1\,(y-1) - 1z = 0\right),$$

ou seja por

$$\begin{cases} x = 1 \\ \\ y - z = 1. \end{cases}$$

6.3 Diagonalização ortogonal de matrizes simétricas

Recordamos que $Q \in \mathcal{M}_{n\times n}(\mathbb{R})$ diz-se ortogonal se $Q^TQ = I$.

Observação 6.37 i) Se Q é uma matriz ortogonal, então Q é invertível, $Q^{-1} = Q^T$, $\det(Q) = \pm 1$ e portanto a transposta Q^T também é ortogonal $QQ^T = I$.
ii) Seja $\{v_1, v_2, \cdots, v_n\}$ uma base ortonormada de $\mathbb{R}^n$ quando munido com o produto interno usual. Seja ainda Q a matriz cuja coluna i é o vector v_i (com $i = 1, 2, \cdots, n$). Então Q é ortogonal e podemos resumidamente escrever

$$\begin{bmatrix} | & \vdots & | \\ v_1 & \cdots & v_n \\ | & \vdots & | \end{bmatrix} \text{matriz ortogonal} \iff \{v_1, ...v_n\} \text{ base ortonormada de } \mathbb{R}^n \quad \text{com o p.i. usual.}$$

Definição 6.38 Seja $A = [a_{ij}] \in \mathcal{M}_{n\times n}(\mathbb{C})$. A matriz A^* cuja entrada (i,j) é dada por $\overline{a_{ji}}$ é habitualmente designada pela matriz transconjudada de A; i.e. $A^* = \overline{A}^T$. Dizemos que a matriz A é

1. normal se $AA^* = A^*A$,
2. hermitiana se $A = A^*$,
3. anti-hermitiana se $A = -A^*$,
4. unitária A for invertível e se $A^{-1} = A^*$.

Se A tiver entradas em $\mathbb{R}$, então a matriz A diz-se normal, simétrica, anti-simétrica e ortogonal, respectivamente.

Temos $(A^*)^* = A$ e $(\alpha AB)^* = \overline{\alpha}B^*A^*$ e

- $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ hermitiana ou anti-hermitiana ou unitária $\implies A$ normal.
- $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ simétrica ou anti-simétrica ou ortogonal $\implies A$ normal.
- Note que $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ é normal $AA^* = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} = A^*A$ mas não é unitária nem (anti)hermitiana!

Seja $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Usando a definição de A^* juntamente com o produto interno usual em $\mathbb{K}^n$, dado por

$$\langle u, v\rangle = x_1\overline{y}_1 + ... + x_n\overline{y}_n \quad \text{com } u = (x_1, ..., x_n),\ v = (y_1, ..., y_n),$$

podemos provar o seguinte resultado fulcral.

Teorema 6.39 *$\langle Au, v\rangle = \langle u, A^*v\rangle$ para quaisquer $u, v \in \mathbb{K}^n$.*

Prova Sejam $u = (x_1, ..., x_n)$ e $v = (y_1, ..., y_n)$. Usando a definição de produto interno usual e produto matricial, temos

$$\langle Au, v\rangle = \langle A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \rangle = \sum_i (Au)_i\ \overline{y}_i = \sum_i \sum_j a_{ij}\ x_j\ \overline{y}_i \qquad (*).$$

Analogamente,

$$\langle u, A^* v\rangle = \sum_i x_i \overline{(A^*v)_i} = \sum_i \sum_j x_i \ \overline{(A^*)}_{ij} \ \overline{y}_j = \sum_i \sum_j x_i \ \overline{\overline{a}}_{ji} \ \overline{y}_j = \sum_i \sum_j a_{ji} \ x_i \ \overline{y}_j,$$

que é a expressão em (*), efectuando a troca de i com j. Q.E.D.

Teorema 6.40 *1. λ valor próprio de A se e só se $\overline{\lambda}$ valor próprio de A^* (i.e. $\sigma(A^*)=\overline{\sigma(A)}$).*

2. A normal se e só se $\langle Au, Av\rangle = \langle A^*u, A^*v\rangle$ para quaisquer u, v.

3. A hermitiana se e só se $\langle Au, v\rangle = \langle u, Av\rangle$ para quaisquer u, v.

4. A unitária se e só se $\langle Au, Av\rangle = \langle u, v\rangle$ para quaisquer u, v.

5. A hermitiana, então os valores próprios de A são todos reais.

6. A unitária então os valores próprios de A têm modulo 1, isto é $|\lambda| = 1$ para qualquer valor próprio λ de A.

7. Se A é normal, então $Au = \lambda u$ se e só se se $A^*u = \overline{\lambda}u$.

8. Se A é normal, vectores próprios associados a valores próprios distintos são ortogonais.

Prova: 1. Uma vez que

$$p_{A^*}(\overline{\lambda}) = \det(A^* - \overline{\lambda} I) = \det(\overline{A}^T - \overline{\lambda} I) = \det((\overline{A} - \overline{\lambda} I)^T) = \det(\overline{A} - \overline{\lambda} I) = \overline{\det(A - \lambda I)} = \overline{p_A(\lambda)}$$

temos $p_{A^*}(\overline{\lambda}) = \overline{p_A(\lambda)}$; logo $\overline{\lambda}$ é valor próprio de A^* sse λ é valor próprio de A.
2. Observe que se $AA^* = A^*A$, então pelo Teorema 6.39 temos

$$\langle Au, Av\rangle = \langle u, A^*Av\rangle = \langle u, AA^*v\rangle = \langle A^*u, A^*v\rangle.$$

Reciprocamente, novamente pelo Teorema 6.39 e hipótese temos

$$\langle AA^*u, A^*Au\rangle = \langle A^*u, A^*A^*Au\rangle = \langle Au, AA^*Au\rangle = \langle A^*Au, A^*Au\rangle$$

e analogamente $\langle A^*Au, AA^*u\rangle = \langle AA^*u, AA^*u\rangle$. Assim podemos facilmente calcular e concluir que

$$||(A^*A - AA^*)u||^2 = \big\langle (A^*A - AA^*)u, (A^*A - AA^*)u \big\rangle = 0$$

logo $(A^*A - AA^*)u = 0$ para todo o u, i.e. $A^*A - AA^* = 0$.
3. Trivial pelo Teorema 6.39.
4. Se A é uma matriz unitária então temos

$$\langle Au, Av\rangle = \langle u, A^*Av\rangle = \langle u, A^{-1}Av\rangle = \langle u, v\rangle,$$

como pretendido. Reciprocamente, a hipótese implica que $||Au||^2 = ||u||^2$ para todo o u (fazendo $u = v$), logo $\mathcal{N}(A) = \{0\}$ e portanto A é invertível. Por outro lado,

$$\langle (A^*A - I)u, v\rangle = \langle A^*Au, v\rangle - \langle u, v\rangle = 0,$$

para quaisquer u, v. logo $A^*A - I = 0$.
5. Seja λ um valor próprio de A e u um vector próprio associado: $Au = \lambda u$. Então usando a parte 3) deste Teorema temos

$$\lambda ||u||^2 = \lambda \langle u, u\rangle = \langle \lambda u, u\rangle = \langle Au, u\rangle = \langle u, Au\rangle = \langle u, \lambda u\rangle = \overline{\lambda}\langle u, u\rangle = \overline{\lambda}||u||^2,$$

logo $\lambda = \overline{\lambda}$ (i.e. $\lambda \in \mathbb{R}$).
6. Seja λ um valor próprio de A e u um vector próprio associado: $Au = \lambda u$. Então temos

$$\lambda\overline{\lambda}||u||^2 = \lambda\overline{\lambda}\langle u, u\rangle = \langle \lambda u, \lambda u\rangle = \langle Au, Au\rangle = \langle u, u\rangle = ||u||^2,$$

logo $\lambda\overline{\lambda} = 1$.

7. Sendo A normal, pela parte 2) deste teorema podemos concluir que $||Au|| = ||A^*u||$ para todo o vector u. Mais, podemos verificar que $A - \lambda I$ também é uma matriz normal, para todo o escalar λ. Assim se $Au = \lambda u$ temos

$$0 = ||(A - \lambda I)u||| = ||(A - \lambda I)^*u|| = ||(A^* - \overline{\lambda}I)u||$$

o que prova que u é um vector próprio de A associado ao valor próprio λ sse u é um vector de A^* associado ao valor próprio $\overline{\lambda}$.
8. Sejam u e v vectores próprios de A associados a valores próprios λ_1 e λ_2 distintos, respectivamente. Assim temos $Au = \lambda_1 u$ e $Av = \lambda_2 v$ e portanto,

$$(\lambda_1 - \lambda_2)\langle u, v\rangle = \langle \lambda_1 u, v\rangle - \langle u, \overline{\lambda_2}v\rangle = \langle Au, v\rangle - \langle u, A^*v\rangle = 0,$$

onde usamos as partes 1) e 7) deste teorema assim como o Teorema 6.39. Como $\lambda_1 \neq \lambda_2$ podemos concluir que $\langle u, v\rangle = 0$. Q.E.D.

Como caso particular, temos o seguinte para matrizes reais.

Teorema 6.41 Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ e considere o p.i. usual em $\mathbb{R}^n$.

1. temos $\langle Au, v\rangle = \langle u, A^T v\rangle$ para quaisquer $u, v \in \mathbb{R}^n$,
2. Os valores próprios de uma matriz simétrica são todos reais. Mais, se u for um vector próprio de uma matriz simétrica então $u \in \mathbb{R}^n$.
3. Se A fôr simétrica, então vectores próprios associados a valores próprios diferentes são ortogonais (e portanto linearmente independentes).

Prova 1. É um caso particular do Teorema 6.39 e 3. Segue pela parte 7) do Teorema 6.40.
2. Os valores próprios de uma matriz (real) simétrica são reais resulta, como caso particular, da parte 1) do Teorema 6.40. Ora se $\lambda \in \mathbb{R}$ e A também tem entradas em $\mathbb{R}$, então os vectores em $\mathcal{N}(A - \lambda I)$ são vectores em $\mathbb{R}^n$.

O conceito de matriz transconjugada/transposta indica-nos como construir transformações lineares em espaços euclideanos gerais.

Lemma 6.42 (Teorema de representação de Riesz) *Seja E um espaço euclidiano de dimensão finita sobre $\mathbb{K}$ e $\Phi : E \to \mathbb{K}$ uma transformação linear. Então existe um e um só $u_0 \in E$ tal que $\Phi(u) = \langle u, u_0\rangle$, para todo o $u \in E$.*

Prova: Se $\Phi(u) = 0$ para todo $u \in E$ então podemos escolher $u_0 := 0$. Se $\Phi \neq 0$, seja

$$M := \mathcal{N}(\Phi) = \{u \in E : \Phi(u) = 0\}.$$

Claro que M é um subespaço linear de E e $M \neq E$ pois $\Phi \neq 0$. Seja $y \in M^{\perp}$. Então $\Phi(y) = c \neq 0$ seja $z = \frac{1}{c}y$, pelo que $\Phi(z) = 1$. Para qualquer $u \in E$, temos $\Phi(u)z - u \in M$ logo $\Phi(u)z - u \perp z$, portanto

$$\langle \Phi(u)z - u, z \rangle = 0, \qquad \forall u \in E.$$

Esta equação pode ser escrita como $\Phi(u) = \langle u, \frac{1}{||z||^2} z \rangle$ pelo que basta escolher $u_0 := \frac{1}{||z||^2} z$, o que prova a existência. Quanto à unicidade, sejam u_0 e u_1 tais que $\langle u, u_0 \rangle = \langle u, u_1 \rangle$ para todo $u \in E$. Assim $\langle u, u_0 - u_1 \rangle = 0$ para todo $u \in E$, i.e. $u_0 - u_1 \in E^{\perp} = \{0\}$. Q.E.D.

Seja E for um espaço euclidiano de dimensão finita sobre $\mathbb{K}$ e $T : E \to E$ uma transformação linear.

Teorema 6.43 (Transformação linear adjunta) *Sendo E um espaço euclidiano de dimensão finita, a equação*

$$\langle T(u), v \rangle = \langle u, T^*(v) \rangle, \tag{4}$$

para quaisquer $u, v \in E$ define unicamente uma transformação linear $T^* : E \to E$ (a transformação adjunta de T).

Prova: Para cada $v \in E$ seja $L_v : E \to \mathbb{K}$ definido por $L_v(u) = \langle T(u), v \rangle$. Claro que L_v é uma transformação linear, logo pelo Lema 6.42 existe um único vector (designado por "$T^*(v)$") tal que $L_v(u) = \langle u, T^*(v) \rangle$. Fica assim construída uma função $T^* : E \to E$ obedecendo à Eq. (4). Falta provar que T^* é uma transformação linear. Ora, isto resulta da Eq. (4) porque

$$\langle u, T^*(\lambda v_1 + v_2) \rangle = \langle T(u), \lambda v_1 + v_2 \rangle = \overline{\lambda} \langle T(u), v_1 \rangle$$

$$= \langle T(u), v_2 \rangle = \overline{\lambda} \langle u, T^*(v_1) \rangle + \langle u, T^*(v_2) \rangle = \langle u, \lambda T^*(v_1) + T^*(v_2) \rangle$$

logo $T^*(\lambda v_1 + v_2) = \lambda T^*(v_1) + T^*(v_2)$ para $\lambda \in \mathbb{K}$, $v_1, v_2 \in E$. Q.E.D.

Observação 6.44

1. *Seja $A \in \mathcal{M}_{n \times n}(\mathbb{K})$ e $T : \mathbb{K}^n \to \mathbb{K}^n$ a transformação linear definida por $T(u) = Au$; então $T^*(u) = A^*u$ se o produto interno em $\mathbb{K}^n$ for o produto interno usual. Assim a transformação adjunta T^* é uma generalização de matriz transconjugada.*

2. *Podemos definir o que é uma transformação normal $TT^* = T^*T$, hermíteana..., exactamente como o fizemos no caso das matrizes na Definição 6.38.*

3. *Mais, os resultados do Teorema 6.40 continuam válidos com as mesmas demonstrações para o contexto de transformações adjuntas (substituindo A por T e A^* por T^*) num espaço euclidiano qualquer de dimensão finita!!*[15]

Claro que $(T^*)^* = T$ e se $v_1, ..., v_n$ for base ortonormada de E; é fácil ver que essa transformação T^* é dada por

$$T^*(u) = \sum_{i=1}^{n} \langle u, T(v_i) \rangle \ v_i \tag{5}$$

Teorema 6.45 Seja $T : E \to E$ uma transforma cão linear num espaço euclidiano E de dimensão finita com coeficientes em $\mathbb{K}$. Então as seguintes afirmações são equivalentes:

[15]podemos remover a restrição de E ser de *dimensão finita*, mas para tal precisamos de desenvolver o conceito de continuidade para transformações lineares, e é para essas transformações lineares que o Lema 6.42 e a Eq. (4) continuam válidos!

1. Existe uma base ortonormada de E constituída por vectores próprios de A.

2. T é normal e os seus valores próprios pertencem a $\mathbb{K}$.

Dem.: Prova de 1) $\Rightarrow$ 2). Seja $v_1, ..., v_n$ uma base de E formada por vectores próprios de T. Temos então $T(v_i) = \lambda_i \ v_i$ com $\lambda_i \in \mathbb{K}$. A matriz de T nessa base (ordenada) é portanto a matriz diagonal

$$D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

pelo que o polinómio característico de T é $p(\lambda) = \det(D - \lambda I) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$. Sendo então $\{\lambda_1, .., \lambda_n\}$ os valores próprios de T, que estão em $\mathbb{K}$. Admitindo agora que essa base $v_1, ..., v_n$ é ortonormada, temos que para $u = \xi_1 v_1 + ... + \xi_n v_n$ e $v = \eta_1 v_1 + ... + \eta_n v_n$ (com os coeficientes em $\mathbb{K}$):

$$\begin{aligned} \langle T(u), T(v)\rangle &= \langle \xi_1 T(v_1) + ... + \xi_n T(v_n), \eta_1 T(v_1) + ... + \eta_n T(v_n)\rangle = \\ &= \langle \xi_1 \lambda_1 v_1 + ... \xi_n \lambda_n v_n, \eta_1 \lambda_1 v_1 + ... \eta_n \lambda_n v_n\rangle \\ &= \sum_{i,j=1}^{n} \xi_i \lambda_i \bar{\lambda}_j \bar{\eta}_j \langle v_i, v_j\rangle = \sum_{i=1}^{n} \xi_i \lambda_i \bar{\lambda}_i \bar{\eta}_i. \end{aligned}$$

Por outro lado, usando a equação (5) temos:

$$\begin{aligned} \langle T^*(u), T^*(v)\rangle &= \Big\langle \sum_{i=1}^{n} \langle u, T(v_i)\rangle \ v_i, \sum_{j=1}^{n} \langle v, T(v_j)\rangle \ v_j \Big\rangle \\ &= \sum_{i,j=1}^{n} \langle u, T(v_i)\rangle \overline{\langle v, T(v_j)\rangle} \langle v_i, v_j\rangle = \sum_{i=1}^{n} \langle u, T(v_i)\rangle \overline{\langle v, T(v_i)\rangle} \\ &= \sum_{i=1}^{n} \langle u, \lambda_i v_i\rangle \overline{\langle v, \lambda_i v_i\rangle} = \sum_{i=1}^{n} \lambda_i \overline{\lambda}_i \langle u, v_i\rangle \overline{\langle v, v_i\rangle} \\ &= \sum_{i=1}^{n} \lambda_i \overline{\lambda}_i \langle \xi_1 v_1 + ... + \xi_n v_n, v_i\rangle \overline{\langle \eta_1 v_1 + ... + \eta_n v_n, v_i\rangle} \\ &= \sum_{i=1}^{n} \lambda_i \overline{\lambda}_i \xi_i \overline{\eta_i}. \end{aligned}$$

Ou seja $\langle T(u), T(v)\rangle = \langle T^*(u), T^*(v)\rangle$ e portanto T é normal.

Prova de 2) $\Rightarrow$ 1). Vamos mostrar, utilizando o método de indução, que existe uma base própria de E, relativa a T que é ortonormada. Se $\dim(E) = 1$, não há nada a provar. Suponhamos que o enunciado é válido para $\dim(E) = n - 1 \neq 0$ e vamos provar que o resultado também é válido para $\dim(E) = n$. Seja $\lambda_1 \in \mathbb{K}$ um valor próprio de T e seja $v_1 \neq 0$ tal que $T(v_1) = \lambda_1 v_1$ e ponha-se $w_1 = \frac{v_1}{||v_1||}$; pelo que $||w_1|| = 1$. Seja $F = L(\{v_1\})^{\perp}$ o complemento ortogonal do espaço gerado pelo vector v_1 em E; portanto $\dim(F) = n - 1$. Provamos que $T(F) \subseteq F$ (para tal basta verificar que $\langle T(v), v_1\rangle = 0$ para todo $v \in F$, o que é de facto acontece usando a definições de T^* e de F). Pelo que a restrição $T|_F$ é uma transformação linear de F para F. Ora essa restrição $T|_F$ continua a ser uma transformação linear normal, pelo que pela hipótese de indução, existe uma base ortonormada $w_2, ..., w_n$ de F formada por vectores próprios de $T|_F$. É claro que então $w_1, w_2, ..., w_n$ é uma base ortonormada de E de vectores próprios de T. Q.E.D.

Usando o produto interno usual em $\mathbb{K}^n$ e a transformação definida à custa da matriz $T(u) = Au$, temos a seguinte consequência do Teorema 6.45.

Corolário 6.46

1. **(Matrizes unitariamente diagonalizáveis)** *A matriz normal $\iff$ existe uma base ortonormada de $\mathbb{C}^n$ constituida por vectores próprios de A.*

2. **(Matrizes ortogonalmente diagonalizáveis)** *A real e simétrica $\implies$ os valores próprios de A são reais e existe uma base ortonormada de $\mathbb{R}^n$ constituida por vectores próprios de A.*

Observação 6.47 *Nas condições deste Corolário 6.46:*

1. *A* **unitariamente diagonalizável** $\iff$ *existe uma matriz diagonal D e uma matriz unitária U tais que*
$$D = UAU^*.$$

2. *A ortogonalmente diagonalizável $\iff$ existe uma matriz diagonal real D e uma matriz ortogonal Q tais que*
$$D = QAQ^T.$$

3. *A matriz real tal que $AA^T = A^TA$ então A é unitariamente diagonalizável porque A é matriz normal. No entanto, A poderá não ser ortogonalmente diagonalizável! por exemplo $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ é normal $AA^T = I = A^TA$ no entanto $\sigma_A = \{\pm i\}$ e os vectores próprios são vectores com entradas em $\mathbb{C}$ e não em $\mathbb{R}$.*

Procedimento para diagonalizar ortogonalmente uma matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ simétrica

1. Calcule os valores prṕrios $\lambda_1, ..., \lambda_n$ de A e determine uma base para cada espaço próprio de A.

2. Aplique o processo de ortogonalização de Gram-Schmidt a cada base de cada de espaços próprio de A. Temos assim uma base ortogonal de $\mathbb{R}^n$ constituída por vetores próprios de A. Normalize esta base, construindo assim uma base $\{v_1, ..., v_n\}$ ortonormada de $\mathbb{R}^n$ constituída por vetores próprios.

3. Seja Q^T a matriz cuja coluna i é o vector v_i e D a matriz diagonal cuja entrada (i,i) é o valor próprio λ_i de A associado ao vector próprio v_i, com $i \in \{1, ..., n\}$.

4. A teoria garante que $D = QAQ^T$.

Exemplo 6.48 Seja $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$. Como A é simétrica sabemos que existe uma matriz ortogonal Q e uma matriz diagonal D tais que $D = QAQ^T$. Vamos então construir Q^T, D e naturalmente $Q = (Q^T)^T$.

1) o polinómio característico de A é

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 4-\lambda & 2 & 2 \\ 2 & 4-\lambda & 2 \\ 2 & 2 & 4-\lambda \end{bmatrix} = ... = (\lambda - 2)^2(\lambda - 8),$$

pelo que os valores próprios de A são $\lambda = 2$ (raiz dupla) e $\lambda = 8$ (raiz simples). O espaço próprio associado a $\lambda = 2$ é $E_2 = \mathcal{N}(A - 2I)$ cujos vectores $u_1 = (-1, 1, 0), u_2 = (-1, 0, 1)$ forma uma base de E_2. O espaço próprio associado a $\lambda = 8$ é $E_8 = \mathcal{N}(A - 8I)$ e o vector $u_3 = (1, 1, 1)$.
2) Aplicando o processo de Gram-Schmidt às bases $\{u_1, u_2\}$ e $\{u_3\}$ e depois normalizando, obtém-se:

$$v_1 = \big(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\big), \quad v_2 = \big(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\big), \quad v_3 = \big(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\big).$$

3) Então temos $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$, $Q^T = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$ e $Q = (Q^T)^T$ e 4) sem fazer cálculos $D = QAQ^T$.

Teorema 6.49 *Seja $B = \{v_1, ..., v_n\}$ uma base ordenada num espaço linear real E, A matriz $n \times n$ real e u_B as coordenadas do vector u na base B. Então $\langle u, v\rangle := [-\ u_B\ -]A \begin{bmatrix} | \\ v_B \\ | \end{bmatrix}$ define um produto interno em E se e só se*

1. *$A = A^T$ (matriz simétrica) e*
2. *$\sigma_A \subseteq \mathbb{R}^+$ (os valores próprios de A são todos estritamente positivos).*

Prova: i) O axioma da linearidade verifica-se sem nenhuma restrição; nomeadamente

$$\langle \lambda u + v, w\rangle = [-\ (\lambda u + vu)_B\ -]A \begin{bmatrix} | \\ w_B \\ | \end{bmatrix} = \lambda[-\ u_B\ -]A \begin{bmatrix} | \\ w_B \\ | \end{bmatrix} + [-\ v_B\ -]A \begin{bmatrix} | \\ w_B \\ | \end{bmatrix} = \lambda\langle u, w\rangle + \langle v, w\rangle.$$

ii) Vejamos o axioma da simetria. Ora $[-\ u_B\ -]A \begin{bmatrix} | \\ v_B \\ | \end{bmatrix}$ é uma matriz 1×1 logo simétrica, portanto

$$\langle u, v\rangle = [-\ u_B\ -]A \begin{bmatrix} | \\ v_B \\ | \end{bmatrix} = \Big([-\ u_B\ -]A \begin{bmatrix} | \\ v_B \\ | \end{bmatrix}\Big)^T = [-\ v_B\ -]A^T \begin{bmatrix} | \\ u_B \\ | \end{bmatrix},$$

Portanto $\langle u, v\rangle = \langle v, u\rangle$ sse $A = A^T$.
ii) Vejamos o axioma da positividade ($\langle u, u\rangle > 0,\ \forall u \neq 0$). Note que

$$\langle u, u\rangle = [-\ u_B\ -]A \begin{bmatrix} | \\ u_B \\ | \end{bmatrix} = [-\ u_B\ -]\ \frac{A + A^T}{2}\ \begin{bmatrix} | \\ u_B \\ | \end{bmatrix}$$

e $\frac{A+A^T}{2}$ é sempre matriz simétrica. Assim, podemos assumir que A é simétrica. Pelo Corolário 6.46 a matriz A é ortogonalmente diagonalizável: $D = QAQ^T$ com $D = \text{diag}(\lambda_1, ..., \lambda_n)$ matriz diagonal e Q matriz ortogonal. Dado $u \in E$ seja cada $x = u_B$ e $y = Qx$. Assim

$$\langle u, u\rangle = [-\ x\ -\ ]A \begin{bmatrix} | \\ x \\ | \end{bmatrix} = [-\ x\ -]Q^T DQ \begin{bmatrix} | \\ x \\ | \end{bmatrix} = [-\ y\ -]D \begin{bmatrix} | \\ y \\ | \end{bmatrix} = \sum_{i=1}^{n} \lambda_i y_i^2.$$

Como $\sum_{i=1}^{n} \lambda_i y_i^2 > 0$ para todo $y = (y_1, ..., y_n)$ sse $\lambda_1 > 0, ..., \lambda_n > 0$ podemos concluir que o axioma da positividade é equivalente à positividade de todos os valores próprios da matriz simétrica A. Q.E.D.

O resultado análogo a este teorema para produtos internos em espaços lineares sobre $\mathbb{C}$ é óbvio: $A = A^*$ e $\sigma_A \subseteq \mathbb{R}^+$.

Exemplo 6.50 Vamos provar que

$$\langle (x_1, x_2 x_3), (y_1, y_2, y_3)\rangle = 4x_1y_1+2x_1y_2+2x_2y_1+2x_1y_3+2x_3y_1+4x_2y_2+2x_2y_3+2x_3y_2+4x_3y_3,$$

define um produto interno em $\mathbb{R}^3$. Ora

$$\langle (x_1, x_2 x_3), (y_1, y_2, y_3)\rangle = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

como a matriz é simétrica, falta somente verificar que os valores próprios de A são estritamente positivos (pelo Teorema 6.49)! Todavia pelo Exemplo 6.48: $\sigma_A = \{2, 2, 8\} \subseteq \mathbb{R}^+$. Portanto (*) define um produto interno em $\mathbb{R}^3$.

Produto Externo e Misto

Sejam $u = (a_1, a_2, a_3), v = (b_1, b_2, b_3)$ vectores de $\mathbb{R}^3$ e $\{e_1, e_2, e_3\}$ a base canónica de $\mathbb{R}^3$. O **produto externo** entre u e v é um vector de $\mathbb{R}^3$, que designamos por $u \times v$ e é definido como:

$$u \times v = \det \begin{bmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = \det \begin{bmatrix} a_2 & a_3 \\ b_2 & b_3 \end{bmatrix} e_1 - \det \begin{bmatrix} a_1 & a_3 \\ b_1 & b_3 \end{bmatrix} e_2 + \det \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} e_3 =$$

$$(a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1).$$

Produto misto é $\langle u, v \times w\rangle = \det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}$.

Teorema 6.51
a) $u \times v = -v \times u$ e $u \times u = 0$,

b) Se u, v são ortogonais e não nulos , então $\{u, v, u \times v\}$ é uma base ortogonal de $\mathbb{R}^3$,

c) $u \times v$ é ortogonal a u e a v,

d) $||u \times v|| = ||u||\ ||v|| \sin(\theta)$ onde θ é o ângulo entre u e v,

e) $||u \times v||$ é a área do paralelogramo definido por u e v,

f) O valor absoluto $|\langle u, v \times w\rangle|$ de $\langle u, v \times w\rangle$ é o volume do paralelipípedo formado pelos vectores u, v e w.

g) $\langle u, u \times v\rangle = \langle u, v \times u\rangle = 0, \qquad \langle u, v \times w\rangle = \langle u \times v, w\rangle$.

h) $V = |\langle u, v \times w\rangle|$ é o volume do paralelipípedo formado pelos vectores u, v e w. Note que

$$V = \underbrace{||u \times v||}_{\text{área da face determinada por u e v}} \quad \underbrace{||w||\ |\cos(\theta)|}_{\text{altura}}.$$

7 Algumas Aplicações

7.1 Formas quadráticas

Formas quadráticas é uma função $Q : \mathbb{R}^n \to \mathbb{R}$ que pode ser escrita na forma

$$Q(u) = \sum_{i,j=1}^{n} a_{ij} x_i x_j, \quad \text{com } u = (x_1, ..., x_n),\ a_{ij} \in \mathbb{R}. \tag{6}$$

Classificação das formas quadráticas Seja Q forma quadrática; Q é

- definida positiva se $Q(u) > 0$, $\forall u \in \mathbb{R}^n$, u$\neq 0$,
- definida negativa se $Q(u) < 0$, $\forall u \in \mathbb{R}^n$, u$\neq 0$,
- semidefinida positiva se $Q(u) \geq 0$, $\forall u \in \mathbb{R}^n$,
- semidefinida negativa se $Q(u) \leq 0$, $\forall u \in \mathbb{R}^n$,
- indefinida se existem u e v tais que $Q(u) > 0$ e $Q(v) < 0$.

A equação (6) pode ser escrita na forma $Q(u) = uAu^T$, com $A = [a_{ij}]$; mas podemos também escrever $Q(u) = u\ \frac{A+A^T}{2}\ u^T$ com a vantagem de $\frac{A+A^T}{2}$ ser uma matriz simétrica.

Exemplo: $Q : \mathbb{R}^2 \to \mathbb{R}$ tal que $Q(x_1, x_2) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_2x_1 + a_{22}x_2^2$. Temos

$$Q(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & \frac{a_{12}+a_{21}}{2} \\ \frac{a_{12}+a_{21}}{2} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Teorema 7.1 Seja $Q(u) = uAu^T$ forma quadrática com A simétrica. Então:

- Q definida positiva se e só se todos os valores próprios de A forem positivos.
- Q definida negativa se e só se todos os valores próprios de A forem negativos.
- Q semidefinida positiva se e só se todos os valores prprios de A forem não negativos.
- Q semidefinida negativa se e só se todos os valores prprios de A forem não positivos.
- Q indefinida se e só se A tiver pelo menos um valor próprio positivo e outro negativo.

Supondo que A é uma matriz real e simétrica, então $Q(u) = uAu^T$ é uma forma quadrática definida positiva se e só $\langle u, v\rangle = uAv^T$ define um produto interno em $\mathbb{R}^n$.

Exemplo 7.2 • Seja $Q(x_1, x_2) = 2x_1^2 + 4x_1x_2 + 2x_2^2$. Então $A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$, cujos valores próprios são $\lambda_1 = 0$ e $\lambda_2 = 4$. Assim, Q é uma forma quadrática semidefinida positiva.

- A forma quadrática Q definida usando a matriz da métrica A de um p.i. $\langle u, v\rangle = uAv^T$ é uma forma quadrática definida positiva, pois $Q(u) = \langle u, u\rangle$.

7.2 Mínimos quadrados

Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e $b \in \mathcal{M}_{m\times 1}(\mathbb{R})$. O sistema linear $A\mathtt{x} = b$ é impossível se e só se $b \notin \mathcal{C}(A)$ (i.e. $S_{A\mathtt{x}=b} = \emptyset$).
Vamos procurar vectores $\widehat{\mathtt{x}}$ que tornem mínima a distância entre $A\widehat{\mathtt{x}}$ e b, isto é $||A\widehat{\mathtt{x}} - b|| = \min_{\mathtt{x}}\{||A\mathtt{x} - b||\}$. Dizemos que tal $\widehat{\mathtt{x}}$ é uma solução de mínimos quadrados associado aos sistema linear $A\mathtt{x} = b$.
Assim, $||A\widehat{\mathtt{x}} - b|| \leq ||A\mathtt{x} - b||$ para todo $\mathtt{x}$; $A\widehat{\mathtt{x}} - b$ o vector erro e $||A\widehat{\mathtt{x}} - b||$ erro de mínimos quadrados.
Claro que $A\mathtt{x} \in \mathcal{C}(A)$ para todo o $\mathtt{x}$, pelo que $||A\mathtt{x} - b||$ é minimizado se

$$A\mathtt{x} = \mathrm{proj}_{\mathcal{C}}(b), \tag{7}$$

onde $\mathrm{proj}_{\mathcal{C}}(b)$ designa a projecção ortogonal de b sobre $\mathcal{C}(A)$. Temos $A\mathtt{x} = \mathrm{proj}_{\mathcal{C}}(b)$ é sempre um sistema possível e as suas soluções são as soluções de mínimos quadrados do sistema inicial $A\mathtt{x} = b$.

Teorema 7.3 • $\widehat{\mathtt{x}}$ solução de mínimos quadrados de $A\mathtt{x} = b$ sse $\widehat{\mathtt{x}}$ é solução do sistema linear $A\mathtt{x} = \mathrm{proj}_{\mathcal{C}}(b)$.

- Existe uma única solução de mńimos quadrados do sistema $A\mathtt{x} = b$ sse $\mathrm{car}(A) = n$.

• Como resolver o sistema linear (7)?

Podemos usar a decomposição $b = \mathrm{proj}_{\mathcal{C}(A)}(b) + \mathrm{proj}_{\mathcal{C}(A)^\perp}(b)$ (note que $\mathcal{C}(A)^\perp = \mathcal{L}_{A^T}^\perp = \mathcal{N}(A^T)$) e concluir que

Teorema 7.4 $\widehat{\mathtt{x}}$ uma solução do sistema linear $A\mathtt{x} = \mathrm{proj}_{\mathcal{C}}(b)$ sse $\widehat{\mathtt{x}}$ é uma solução do sistema linear $(A^TA)\widehat{\mathtt{x}} = A^Tb$.

A equação $(A^TA)\widehat{\mathtt{x}} = A^Tb$ é designada por equação normal.

Teorema 7.5 • $\mathcal{N}(A) = \mathcal{N}(A^TA)$.

- $S_{A^TA\widehat{\mathtt{x}}=A^Tb} \neq \emptyset, \quad S_{A\mathtt{x}=b} \subset S_{A^TA\widehat{\mathtt{x}}=A^Tb}$.

- Se $S_{A\mathtt{x}=b} \neq \emptyset$, então $S_{A\mathtt{x}=b} = S_{A^TA\widehat{\mathtt{x}}=A^Tb}$.

- Se $\mathrm{car}(A) = n$, $\widehat{\mathtt{x}} = (A^TA)^{-1}A^Tb$ é a única solução da equação normal $A^TA\widehat{\mathtt{x}} = A^Tb$.

Exemplo: Sejam $A=\begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix}$, $b=\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. O sistema linear $A\mathtt{x} = b$ é impossível. Por outro lado $\text{car}(A) \neq 2$ pelo que a solução de mínimos quadrados não é única. Podemos verificar isso mesmo, determinando o conjunto solução de $A^TA\widehat{\mathtt{x}} = A^Tb$. Calculando temos $A^TA = \begin{bmatrix} -3 & -6 \\ -6 & -12 \end{bmatrix}$ e $A^Tb = \begin{bmatrix} -3 \\ -6 \end{bmatrix}$, pelo que o conjunto solução de $A^TA\widehat{\mathtt{x}} = A^Tb$ é $\{(x,y) \in \mathbb{R}^2 : x + 3y = 1\}$ (o conjunto solução de mínimos quadrádros de $A\mathtt{x} = b$).

• **Ajusto de curvas a uma tabela**

Pretende-se encontrar uma função $y = f(x)$ que se ajuste a um conjunto de dados experimentais (p.e. em $\mathbb{R}^2$)

$$P_1 = (x_1, y_1),\ P_2 = (x_2, y_2), ..., P_n = (x_n, y_n)$$

da melhor maneira possível.

Modelo Linear: Seja $\mathcal{R}$ a recta $y = \alpha + \beta x$

Para $P_i \in \mathcal{R}$ temos o sistema linear $\begin{cases} \alpha + \beta x_1 = y_1 \\ \alpha + \beta x_2 = y_2 \\ \quad \vdots \\ \alpha + \beta x_n = y_n \end{cases}$ nas variáveis α, β, para o qual $A=\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$, $b=\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$.

Se $P_i \notin \mathcal{R}$ para algum i, então o sistema linear é impossível. Nesse caso, procuramos a recta que melhor se aproxima dos pontos, cuja solução é

$$\widehat{\mathtt{x}} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = (A^TA)^{-1}A^Tb.$$

Exemplo: Sejam $P_1 = (1, 3/2), P_2 = (2, 1/2), P_3 = (3, 3)$. Assim $A=\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$, $b=\begin{bmatrix} 3/2 \\ 1/2 \\ 3 \end{bmatrix}$, cuja solução é $(A^TA)^{-1}A^Tb = \begin{bmatrix} 1/6 \\ 3/4 \end{bmatrix}$ e a recta pretendida é: $y = \frac{1}{6} + \frac{3}{4}x$.

Modelo quadrático: $y = \alpha + \beta x + \gamma x^2$,

originando o sistema $\begin{bmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ nas variáveis α, β, γ.

7.3 Equações diferenciais ordinárias

• Se $f : \mathbb{R} \to \mathbb{R}$ é solução da equação diferencial $f'(t) = \lambda f(t)$ (com λ escalar fixo), então existe um escalar c tal que $f(t) = c\ e^{\lambda t}$.

• Considere funções $x_1(t), x_2(t), \cdots, x_n(t)$ diferenciáveis na variável real t. O sistema da forma

$$\left\{\begin{array}{c} a_{11}x_1(t)+a_{12}x_2(t)+...+a_{1n}x_n(t)=x_1'(t)\\ a_{21}x_1(t)+a_{22}x_2(t)+...+a_{2n}x_n(t)=x_2'(t)\\ \cdots\\ a_{m1}x_1(t)+a_{m2}x_2(t)+...+a_{mn}x_n(t)=x_m'(t)\end{array}\right. \tag{8}$$

chama-se sistema linear de equações diferenciais de primeira ordem, em que a_{ij} é uma constante e $x_i'(t)$ designa a derivada de $x_i(t)$ $(i=1,...,m,\ j=1,...,n)$.

O sistema (8) pode escrever-se na forma matricial: $x'(t)=Ax(t)$ onde $A=[a_{ij}]\in\mathcal{M}_{n\times n}(\mathbb{R})$,

$$x(t)=\begin{bmatrix} x_1(t)\\ x_2(t)\\ \vdots\\ x_n(t)\end{bmatrix},\qquad x'(t)=\begin{bmatrix} x_1'(t)\\ x_2'(t)\\ \vdots\\ x_n'(t)\end{bmatrix}.$$

• **Resolução de $x'=\mathbf{A}x$ com A diagonalizável**

Se a matriz $A=[a_{ij}]\in\mathcal{M}_{n\times n}(\mathbb{R})$ é diagonalizável, para resolver $x'(t)=Ax(t)$ em primeiro lugar encontra-se uma matriz mudança de base

$$S=S_{Bc\to B_{vp}},\qquad S^{-1}=S_{B_{vp}\ \to Bc}$$

onde $B_{vp}=\{v_1,v_2,\cdots,v_n\}$ é uma base de $\mathbb{R}^n$ formada por vectores próprios de A tal que o valor próprio associado a v_i é λ_i, $i=1,2,\cdots,n$, Bc é a base canónica de $\mathbb{R}^n$ e matriz diagonal

$$D=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

(formada pelos valores próprios de A) tais que $D=SAS^{-1}$. Depois, usa-se a mudança de variável $Sx=y$ e e transforma-se o sistema $x'=Ax$ no sistema $y'(t)=Dy(t)$ com as funções separadas, cuja solução geral é $y(t)=\begin{bmatrix} c_1e^{\lambda_1 t}\\ c_2e^{\lambda_2 t}\\ \vdots\\ c_ne^{\lambda_n t}\end{bmatrix}$ onde $\lambda_1,\cdots,\lambda_n$ são os valores próprios de A e $c_1,\cdots,c_n$ são constantes. Finalmente, a solução geral do sistema inicial $x'(t)=Ax(t)$ é

$$x(t)=S^{-1}y(t)=\begin{bmatrix} | & \vdots & |\\ v_1 & \cdots & v_n\\ | & \vdots & |\end{bmatrix}\begin{bmatrix} c_1e^{\lambda_1 t}\\ c_2e^{\lambda_2 t}\\ \vdots\\ c_ne^{\lambda_n t}\end{bmatrix}$$

porque $x'(t)=Ax(t)\Longleftrightarrow x'(t)=S^{-1}DSx(t)\Longleftrightarrow Sx'(t)=DSx(t)\Longleftrightarrow y'(t)=Dy(t)$.

Exemplo: Vamos determinar a solução geral do seguinte sistema de equações diferenciais:

$$\left\{\begin{array}{l} 2x_1(t)+x_2(t)=x_1'(t)\\ -2x_1(t)+5x_2(t)=x_2'(t)\end{array}\right. \tag{9}$$

Claro que $A=\begin{bmatrix} 2 & 1\\ -2 & 5\end{bmatrix}$, cujas valores próprios são $\lambda_1=3$ e $\lambda_2=4$, pelo que A é diagonalizável, $\{(1,1)\}$ é uma base para o espaço próprio para E_{λ_1} e $\{(1,2)\}$ é uma base para o espaço próprio para E_{λ_1}. Assim,

$$D=\begin{bmatrix} 3 & 0\\ 0 & 4\end{bmatrix},\qquad S^{-1}=\begin{bmatrix} 1 & 1\\ 1 & 2\end{bmatrix}$$

e portanto a solução geral do sistema de equações diferenciais (9) é

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = S^{-1} \begin{bmatrix} c_1 e^{2t} \\ c_2 e^{4t} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 e^{2t} \\ c_2 e^{4t} \end{bmatrix} = \begin{bmatrix} c_1 e^{2t} + c_2 e^{4t} \\ c_1 e^{2t} + 2c_2 e^{4t} \end{bmatrix}.$$

Vamos calcular a única solução de (9) sujeita às condições iniciais $x_1(0) = 1,\ x_2(0) = -1$. Ora $(x_1(0), x_2(0)) = (c_1 + c_2, c_1 + 2c_2)$, pelo que $c_1 = 3c_2 = -2$ e a única soluçã de (9) é $(x_1(t), x_2(t)) = (3e^{2t} - 2e^{4t}, 3e^{2t} - 4e^{4t})$.

7.3.1 Um processo de difusão

Considere 2 células adjacentes separadas por uma membrana permeável e suponha que um fluído passa da 1^a célula para a 2^a a uma taxa (em mililitros por minutos) numericamente igual a 3 vezes o volume (em mililitros) do fluído na 1^a célula. Em seguida, passa da 2^a célula para a 1^a a uma taxa numericamente igual a 2 vezes o volume do fluído na 2^a célula. Vamos representar por $x_1(t)$ e $x_2(t)$ os volumes do fluído na 1^a e 2^a células, respectivamente, no instante t. Suponhamos que, inicialmente i.e. $t = 0$, a primeira célula tem 40 ml de fluído, enquanto que 2^a tem 5 ml.

Vamos determinar o volume de fluído em cada célula no instante t.

Solução A variação e volume de fluído em cada célula é a diferença entre a quantidade que entra e a quantidade que sai. Como nenhum fluído entra na primeira célula, temos:

$$\frac{dx_1(t)}{dt} = -3x_1(t),$$

onde o sinal de menos indica que o fluído sai da célula. O fluxo $3x_1(t)$ sai da 1^a célula e entra na 2^a. O fluxo que sai da 2^a célula é de $2x_2(t)$. Logo a variação no volume na 2^a célula é dada por

$$\frac{dx_2(t)}{dt} = 3x_1(t) - 2x_2(t).$$

Obtém-se assim o seguinte sistema de equações diferenciais de 1^a ordem:

$$\begin{cases} -3x_1(t) = x_1'(t) \\ 3x_1(t) - 2x_2(t) = x_2'(t) \end{cases},$$

que pode ser escrito na forma matricial como: $\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$

Os valores próprios da matriz $A = \begin{bmatrix} -3 & 0 \\ 3 & -2 \end{bmatrix}$ são $\lambda_1 = -3$ e $\lambda_2 = -2$. A matriz A é uma diagonalizável onde $\{(1, -3)\}$ é uma base para o espaço próprio E_{λ_1}, enquanto que $\{(0, 1)\}$ é uma base para o espaço próprio E_{λ_2}. Portanto a solução geral do sistema de equações diferenciais acima descrito é:

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} k_1 e^{-3t} \\ k_2 e^{-2t} \end{bmatrix} = k_1 \begin{bmatrix} 1 \\ -3 \end{bmatrix} e^{-3t} + k_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-2t}.$$

Usando as condições iniciais $x_1(0) = 40$ e $x_2(0) = 5$ concluímos que

$$k_1 = 40,\ -3k_1 + k_2 = 5,\ \text{pelo que } k_1 = 40 \text{ e } k_2 = 125.$$

Portanto, o volume de fluido em cada célula no instante t é dado por:

$$x_1(t) = 40e^{-3t}, \qquad x_2(t) = -120e^{-3t} + 125e^{-2t}.$$

7.4 Genes ligados ao sexo

A cegueira para as cores, ou daltonismo, é uma alteração hereditária cujo mecanismo de transmissão só foi compreendido em 1910 (após os estudos de hereditariedade ligado so sexo em diversos animais: aves, borboletas e drasófilas). Sabe-se actualmente que os genes relacionados com determinação deste carácter encontra-se no cromossoma X e que o gene para a visão normal é dominante sobre o alelo que determina o daltonismo.

Desta forma, compreende-se que a transmissão desta característica obedeça às seguintes regras:

- do casamento de um homem daltónico com uma mulher normal, resultem filhas normais e filhos daltónico;
- do casamento de uma mulher daltónica com um homem normal, resultem filhas normais e filhos daltónicos;
- as filhas de pai daltónico são sempre portadoras do daltonismo apesar de fenotipicamente normais.

Este tipo de herança resulta do facto de o homem receber o cromossoma X da mãe e nunca o transmitir aos filhos homens. Por outro lado, as mulheres herdam um cromossoma X da mãe e outro cromossoma X do pai.

Pelo que para encontrar um modelo matemático que descreva o daltonismo numa população, é necessário dividir a população em duas classes, homens e mulheres. Seja $x_m^{(0)}$ a proporção de genes para o daltonismo na população masculina e seja $x_f^{(0)}$ a propoção feminina. Como os homens recebem um cromossoma X da mãe e nenhum do pai, a proporção $x_m^{(1)}$ de homens daltónicos na próxima geração será a mesma que a proporção de genes recessivos na geração actual das mulheres. Como as mulheres recebem um cromossoma X da mãe e outro do pai, a proporção $x_f^{(1)}$ de genes recessivos na próxima geração de mulheres será a média entre $x_m^{(0)}$ e $x_f^{(0)}$. Assim, temos:

$$x_f^{(0)} = x_m^{(1)}, \qquad \frac{1}{2}x_m^{(0)} + \frac{1}{2}x_f^{(0)} = x_f^{(1)}.$$

Se $x_f^{(0)} = x_m^{(0)}$ então a proporção vai manter-se na próxima geração. Vamos então supor que $x_f^{(0)} \neq x_m^{(0)}$ e escrever o sistema na forma matricial

$$\begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_m^{(0)} \\ x_f^{(0)} \end{bmatrix} = \begin{bmatrix} x_m^{(1)} \\ x_f^{(1)} \end{bmatrix}.$$

Vamos designar por A a matriz dos coeficientes do sistema e por $x^{(n)} = \begin{bmatrix} x_m^{(n)} \\ x_f^{(n)} \end{bmatrix}$ a proporção de genes para nas populações masculinas e femininas da $(n+1)$-ésima geração. Então:

$$x^{(n)} = A^n x^{(0)}.$$

Para calcular A^n vamos provar que a matriz A é diagonalizável e construir matriz mudança de base S e matriz diagonal D, tais que $D = SAS^{-1}$. Logo $A = S^{-1}DS$ e portanto $A^n = S^{-1}D^nS$. Ora $\lambda_1 = 1$ e $\lambda_2 = -1/2$ são os valores próprios de A, pelo que $D = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}$.

Além disso $(1,1)$ é vector próprio associado a λ_1 e o $(-2,1)$ é vector próprio associado a λ_2, pelo que

$$S^{-1} = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}, \quad S = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{bmatrix}.$$

Logo,

$$x^{(n)} = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}^n \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} x_m^{(0)} \\ x_f^{(0)} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1-(-\frac{1}{2})^{n-1} & 2+(-\frac{1}{2})^{n-1} \\ 1-(-\frac{1}{2})^{n} & 2+(-\frac{1}{2})^{n} \end{bmatrix} \begin{bmatrix} x_m^{(0)} \\ x_f^{(0)} \end{bmatrix};$$

assim,

$$\lim_{n\to\infty} x^{(n)} = \frac{1}{3} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_m^{(0)} \\ x_f^{(0)} \end{bmatrix} = \begin{bmatrix} \frac{x_m^{(0)}+2x_f^{(0)}}{3} \\ \frac{x_m^{(0)}+2x_f^{(0)}}{3} \end{bmatrix}.$$

Conclusão: as propoções de genes para o daltonismo nas populações masculina e feminina vão tender para o mesmo valor quando o número de geraçõs cresce: se a proporção de homens daltónicos for $p \leq 1$ e se durante um certo número de gerações nenhuma pessoa de fora entrou na população, justifica-se então supor que a proporção de daltonismo na população feminina também é p.

Ora como o daltonismo é recessivo, esperaríamos que a propoção de mulheres daltónicas fosse da ordem p^2, o que este modelo matemático não confirma!

7.5 Redes e grafos

A teoria de grafos é uma das áreas importantes da matemática aplicada. É usada para modelar problemas em praticamente todas as ciências aplicadas. A teoria de grafos é particularmente útil em aplicações envolvendo redes de comunicação.

Um grafo (não orientado) G é definido como um conjunto de pontos chamados vértices junto com um conjunto pares não ordenados de vértices chamados de arestas. Obviamente que podemos representar o grafo G geometricamente onde cada vértice V_i corresponde a nós numa rede de comunicação. Os segmentos de recta unindo os vértices correspondem às arestas. Numa rede, cada aresta representa um elo de comunicação directo entre dois nós da rede. Uma rede de comunicação verdadeira pode envolver um grande número de vértices e arestas, pelo que uma representaçãp gráfica da rede seria muito confusa. Uma alternativa é usar uma representação matricial para a rede. Se o grafo contém um total de n vértices, então a matriz $A = [a_{ij}] \in \mathcal{M}_{n\times n}$ de adjacência do grafo é definida da seguinte maneira:

$$a_{ij} = \begin{cases} 1 & \text{se } \{v_i, v_j\} \text{ é uma aresta de } G \\ 0 & \text{caso contrário} \end{cases},$$

Observe que a matriz A é simétrica $A = A^T$, por definição. Podemos pensar num caminho no grafo G como uma sequência de arestas unindo vértices. Dizemos que o caminho tem comprimento k se o caminho for a sequência de k arestas em G. (Incluir um grafo para ilustrar o texto)

Problema: determinar os caminhos de comprimento k.

Teorema 7.6 Seja A matriz de adjacência de um grafo G e $a_{ij}^{(k)}$ a entrada (i,j) da matriz A^k. Então $a_{ij}^{(k)}$ é o número de caminhos de comprimento k do vértice v_i a v_j.

Demontração: Aplicar indução matemática em k. No caso $k = 1$ segue da definição de matriz de adjacência que a_{ij} é o número de caminhos de comprimento 1 entre os vértices v_i e v_j.

Vamos agora supor que $a_{ij}^{(k)}$ é o número de caminhos de comprimento k entre os vértices v_i e v_j. Queremos provar que $a_{ij}^{(k+1)}$ é o número de caminhos de comprimento $k+1$ entre os vértices v_i e v_j. Ora se existe uma aresta entre v_l e v_j, então $a_{il}^{(k)} a_{lj} = a_{il}$ é o número de caminhos de comprimento $k+1$ entre v_i e v_j da forma

$$v_i \to \cdots \to v_l \to v_j.$$

Temos então que o número total de caminhos de comprimento $k+1$ entre v_i e v_j é dado por

$$a_{i1}^{(k)} a_{1j} + a_{i2}^{(k)} a_{2j} + \cdots a_{in}^{(k)} a_{nj}.$$

Mas isto é por definição de produto matricial a entra (i, j) de A^{k+1}, c.q.d.

Como A é uma matriz simétrica, A é diagonalizável, pelo que os valores próprios fornecem a diagonal da matriz diagonal D, a determinação de bases para os espaços próprios fornecem as colunas para a matriz S^{-1}, pelo que $S = (S^{-1})^{-1}$. Mais $D = SAS^{-1}$, donde

$$A^k = S^{-1} D^k S.$$

Uma vez que A é simétrica podemos escolher as bases dos espaços próprios de tal forma que a matriz S seja ortogonal $S^{-1} = S^T$ (ver aulas teóricas anteriores).

Note que no mesmo gráfo não orientado G podemos definir outra matriz $A' = [a'_{ij}]$ como sendo

$$a'_{ij} = \begin{cases} s & \text{se } \{v_i, v_j\} \text{ estão ligados por } s \text{ arestas} \\ 0 & \text{caso contrário} \end{cases},$$

Problema: verifique a validade do teorema anterior!

Exemplo 7.7 a) Esboce o grafo cuja matriz de adjacência é $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

b) Determine uma matriz ortogonal Q tal que QAQ^T seja uma matriz diagonal.

c) Calcule o número de caminhos de comprimento 10 entre dois vértices diferentes (à sua escolha) do grafo de a).

Grafos orientados

Refaça a secção anterior para grafos *orientados*. Conhecem-se aplicações destes grafos à Sociologia, Telecomunicações etc.

Note que nestes grafos, em geral, a matriz que lhe está associada não é simétrica uma vez que, p.ex., pode haver uma aresta do vértice v_i para o vértice v_j, mas não haver nenhuma aresta de v_j para v_i.