

---

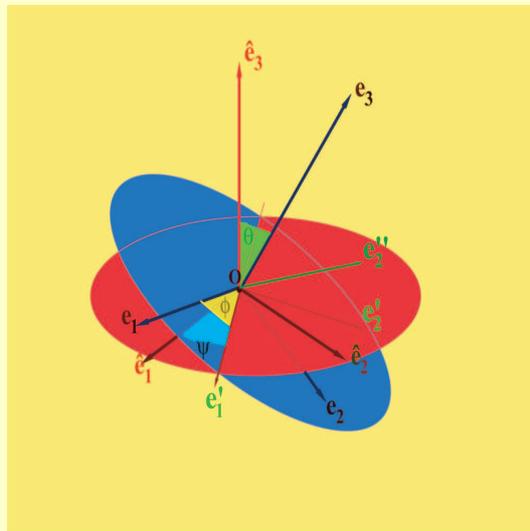
---

Dep. Matemática Pura. FCUP

---

---

## ÁLGEBRA LINEAR II



Resumo das aulas teóricas e práticas

1.º ano da licenciatura em Matemática

Ano lectivo de 2005/06

João Nuno Tavares

---

# ÍNDICE:

<b>1</b>	<b>Determinantes. Produtos vectorial e misto (ou triplo) em <math>\mathbb{R}^3</math></b>	<b>2</b>
1.1	Determinantes . . . . .	2
1.2	Produto vectorial. Produto misto (ou triplo) em $\mathbb{R}^3$ . Interpretação geométrica do determinante . . . . .	6
1.3	Interpretação geométrica do $\det \mathbf{A}$ . . . . .	8
<b>2</b>	<b>Espaços vectoriais com produto interno</b>	<b>10</b>
2.1	Espaços Euclidianos reais . . . . .	10
2.2	Espaços Hermitianos (ou Unitários) complexos . . . . .	13
2.3	Norma . . . . .	14
2.4	Ortogonalidade . . . . .	15
2.5	Aplicações à geometria . . . . .	17
2.6	Bases ortonormadas num espaço vectorial com produto interno . . . . .	20
2.7	Método de ortogonalização de Gram-Schmidt . . . . .	20
2.8	Decomposição ortogonal. Teorema da aproximação óptima . . . . .	22
2.9	Aplicações. Mínimos quadrados . . . . .	27
2.10	Exercícios . . . . .	31
<b>3</b>	<b>Subespaços invariantes. Subespaços próprios. Valores próprios</b>	<b>36</b>
3.1	Conjugação . . . . .	36
3.2	Subespaços invariantes . . . . .	38
3.3	Valores e vectores próprios de um operador linear. Operadores diagonalizáveis . . . . .	39
3.4	Cálculo de valores e vectores próprios . . . . .	40
3.5	Sistemas dinâmicos lineares discretos . . . . .	43
3.6	Exercícios . . . . .	49
<b>4</b>	<b>Transformações ortogonais e unitárias</b>	<b>51</b>
4.1	Transformações ortogonais e unitárias. Exemplos . . . . .	51
4.2	Isometrias em $\mathbb{R}^2$ . Os grupos $\mathcal{O}(2)$ e $\mathcal{SO}(2)$ . . . . .	56

4.3	Isometrias em $\mathbb{R}^3$ . Rotações. Ângulos de Euler. Os grupos $\mathcal{O}(3)$ e $\mathcal{SO}(3)$ . . . . .	58
4.4	Transformações unitárias em $\mathbb{C}^2$ . Os grupos $U(2)$ e $SU(2)$ . . . . .	62
4.5	Exercícios . . . . .	63
<b>5</b>	<b>Operadores auto-adjuntos (simétricos e hermitianos). Teorema espectral</b>	<b>65</b>
5.1	Operadores auto-adjuntos (simétricos e hermitianos) . . . . .	65
5.2	Teorema espectral para operadores auto-adjuntos . . . . .	67
5.3	Diagonalização de formas quadráticas reais . . . . .	69
5.4	Diagonalização simultânea de duas formas quadráticas reais . . . . .	73
5.5	Exercícios . . . . .	77
<b>6</b>	<b>Cônicas e quádricas afins</b>	<b>78</b>
6.1	Parábola, Elipse e Hipérbole . . . . .	78
6.2	Quádricas . . . . .	80
6.3	Cônicas e quádricas afins . . . . .	83
6.4	Redução à forma canônica da equação geral de uma cônica . . . . .	88
<b>7</b>	<b>Quaterniões e Rotações</b>	<b>95</b>

## Referências

1. T.M. Apostol: “**Calculus**, vol.1 e vol.2”. Xerox College Publishing International Text-book series, 1969.
2. Postnikov M.: “**Leçons de Géométrie, vol.1 e 2**”. Éditions MIR, Moscou,1981.
3. Banchoff T., Wermer J.. “**Linear Algebra through Geometry**”. UTM, Springer-Verlag, New York, 1983.
4. Smith L.: “**Linear Algebra**”. UTM, Springer-Verlag, New York, 1978.
5. Curtis C.W.: “**Linear Algebra, An Introductory Approach**”. UTM, Springer-Verlag, New York, 1974.
6. Lipschutz S.: “**Linear Algebra**”. Schaum’s Outline Series. McGraw-Hill Book Company,1968.
7. Hernández E.: “**Álgebra y Geometría**”(2.<sup>a</sup> edición). Addison-Wesley/Universidad Autónoma de Madrid, 1994.

# Capítulo 1

## Determinantes. Produtos vectorial e misto (ou triplo) em $\mathbb{R}^3$

### 1.1 Determinantes

► **1.1 Matrizes  $2 \times 2$  ...** Dada uma matriz  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ , com entradas em  $\mathbb{k}$ , definimos o seu **determinante**  $\det A$ , como sendo o escalar:

$$\det A = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \in \mathbb{k} \quad (1.1.1)$$

Representemos por  $A_1 = \begin{bmatrix} a \\ c \end{bmatrix}$  e  $A_2 = \begin{bmatrix} b \\ d \end{bmatrix}$  as colunas da matriz  $A$ , de tal forma que:

$$\det A = \det [A_1 \quad A_2] = ad - bc \quad (1.1.2)$$

Um cálculo directo mostra que:

$$\det [A_1 \quad A_2] \neq 0 \quad \text{sse } A_1, A_2 \text{ são linearmente independentes} \quad (1.1.3)$$

$$\det [A_1 \quad A_2] = -\det [A_2 \quad A_1] \quad (1.1.4)$$

$$\det [A_1 + A'_1 \quad A_2] = \det [A_1 \quad A_2] + \det [A'_1 \quad A_2] \quad (1.1.5)$$

$$\det [A_1 \quad A_2 + A'_2] = \det [A_1 \quad A_2] + \det [A_1 \quad A'_2] \quad (1.1.6)$$

$$\begin{aligned} \det [\lambda A_1 \quad A_2] &= \lambda \det [A_1 \quad A_2] \\ &= \det [A_1 \quad \lambda A_2] \quad \lambda \in \mathbb{k} \end{aligned} \quad (1.1.7)$$

e ainda que:

$$\det \mathbf{I} = 1 \quad (1.1.8)$$

$$\det (AB) = \det A \det B \quad (1.1.9)$$

$$\det (A^{-1}) = (\det A)^{-1} \quad \forall A \in GL(2; \mathbb{k}) \quad (1.1.10)$$

$$\det (P^{-1}AP) = \det A \quad \forall P \in GL(2; \mathbb{k}) \quad (1.1.11)$$

$$\det (A) = \det (A^t) \quad (1.1.12)$$

onde  $A^t$  é a **transposta** de  $A$ .

Além disso é possível provar que para uma matriz  $A \in \mathcal{M}_{2,2}(\mathbb{K})$ :

$$A \text{ é inversível se e só se } \det A \neq 0$$

e, nesse caso:

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad (1.1.13)$$

Finalmente, se  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  é um operador linear num espaço vectorial  $\mathcal{V}$  de dimensão 2, sobre  $\mathbb{K}$ , define-se o respectivo **determinante**  $\det \mathbf{L}$ , como sendo o determinante da matriz de  $L$ , relativamente a uma qualquer base de  $\mathcal{V}$ . Por (1.1.10), esta definição não depende da base escolhida. Veremos em breve uma interpretação geométrica para  $\det \mathbf{L}$ , no caso real.

► **1.2 Matrizes  $3 \times 3$  ...** ... Dada uma matriz  $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$ , com entradas em  $\mathbb{K}$ , definimos o seu **determinante**  $\det A$ , como sendo o escalar:

$$\begin{aligned} \det A &= \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} \\ &= a \det \begin{bmatrix} e & f \\ h & k \end{bmatrix} - b \det \begin{bmatrix} d & f \\ g & k \end{bmatrix} + c \det \begin{bmatrix} d & e \\ g & h \end{bmatrix} \in \mathbb{K} \end{aligned} \quad (1.1.14)$$

Representemos por:

$$A_1 = \begin{bmatrix} a \\ d \\ g \end{bmatrix}, \quad A_2 = \begin{bmatrix} b \\ e \\ h \end{bmatrix} \quad e \quad A_3 = \begin{bmatrix} c \\ f \\ k \end{bmatrix}$$

as colunas da matriz  $A$ , de tal forma que:

$$\det A = \det [A_1 \ A_2 \ A_3] \quad (1.1.15)$$

É possível mostrar as seguintes propriedades do  $\det$  :

- (i).  $\det [A_1 \ A_2 \ A_3] \neq 0$  sse  $A_1, A_2, A_3$  são linearmente independentes.
- (ii).  $\det [A_1 \ A_2 \ A_3]$  muda de sinal, sempre que se permuta um par de colunas.
- (iii).

$$\det [A_1 + A'_1 \ A_2 \ A_3] = \det [A_1 \ A_2 \ A_3] + \det [A'_1 \ A_2 \ A_3] \quad (1.1.16)$$

$$\det [A_1 \ A_2 + A'_2 \ A_3] = \det [A_1 \ A_2 \ A_3] + \det [A_1 \ A'_2 \ A_3] \quad (1.1.17)$$

$$\det [A_1 \ A_2 \ A_3 + A'_3] = \det [A_1 \ A_2 \ A_3] + \det [A_1 \ A_2 \ A'_3] \quad (1.1.18)$$

$$\begin{aligned} \det [\lambda A_1 \ A_2 \ A_3] &= \lambda \det [A_1 \ A_2 \ A_3] \\ &= \det [A_1 \ \lambda A_2 \ A_3] \\ &= \det [A_1 \ A_2 \ \lambda A_3] \quad \lambda \in \mathbb{K} \end{aligned} \quad (1.1.19)$$

e ainda que:

(iv).

$$\det \mathbf{I} = 1 \quad (1.1.20)$$

$$\det (AB) = \det A \det B \quad (1.1.21)$$

$$\det (A^{-1}) = (\det A)^{-1} \quad \forall A \in GL(3; \mathbb{k}) \quad (1.1.22)$$

$$\det (P^{-1} A P) = \det A \quad \forall P \in GL(3; \mathbb{k}) \quad (1.1.23)$$

$$\det (A) = \det (A^t) \quad (1.1.24)$$

onde  $A^t$  é a **transposta** de  $A$ .

(v). Além disso é possível provar que para uma matriz  $A \in \mathcal{M}_{3,3}(\mathbb{k})$ :

$A$  é inversível se e só se  $\det A \neq 0$

Nesse caso, a inversa de  $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$  pode ser calculada da seguinte forma: em primeiro lugar definimos a chamada **matriz adjunta** de  $A$ ,  $\text{adj } A$ , através de:

$$\text{adj } A = \begin{bmatrix} \begin{vmatrix} e & f \\ h & k \end{vmatrix} & -\begin{vmatrix} d & f \\ g & k \end{vmatrix} & \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ -\begin{vmatrix} b & c \\ h & k \end{vmatrix} & \begin{vmatrix} a & c \\ g & k \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} \\ \begin{vmatrix} b & c \\ e & f \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{bmatrix}^t \quad (1.1.25)$$

Esta matriz é pois obtida substituindo cada entrada de  $A$  pelo determinante  $2 \times 2$ , chamado o **cofactor** dessa entrada, obtido por remoção da linha e coluna que contêm essa entrada, afectado de um sinal  $+$  ou  $-$ , como está indicado. Finalmente:

$$A^{-1} = \frac{1}{\det A} \text{adj } A \quad (1.1.26)$$

Se  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  é um operador linear num espaço vectorial de dimensão 3, sobre  $\mathbb{k}$ , define-se o respectivo **determinante**  $\det \mathbf{L}$ , como sendo o determinante da matriz de  $L$ , relativamente a uma qualquer base de  $\mathcal{V}$ . Por (1.1.23), esta definição não depende da base escolhida. Veremos em breve uma interpretação geométrica para  $\det \mathbf{L}$ , no caso real.

► **1.3 Matrizes  $n \times n$  ...** A generalização da função determinante para matrizes quaisquer  $n \times n$ , está contida no teorema seguinte, cuja demonstração omitimos.

Dada uma matriz  $A \in \mathcal{M}_n(\mathbb{k})$ , digamos  $A = [A_j^i]$ , representemos por  $A_1, A_2, \dots, A_n$  as respectivas  $n$  colunas. A matriz  $A$  será escrita na forma:

$$A = [A_1 \ A_2 \ \dots \ A_n]$$

► **1.4 Teorema** ... Existe uma única aplicação  $\det$  :

$$\begin{aligned} \det : \mathcal{M}_n(\mathbb{k}) &\longrightarrow \mathbb{k} \\ A &\longmapsto \det A \end{aligned} \quad (1.1.27)$$

que satisfaz as seguintes três propriedades:

- $\det [A_1 \cdots A_i \cdots A_j \cdots A_n] = -\det [A_1 \cdots A_j \cdots A_i \cdots A_n]$
- $\det [A_1 \cdots A_i + \lambda A'_i \cdots A_n] = \det [A_1 \cdots A_i \cdots A_n] + \lambda \det [A_1 \cdots A'_i \cdots A_n]$
- $\det \mathbf{1}_n = 1$ , onde  $\mathbf{1}_n$  é a matriz identidade  $n \times n$ .

Esta função determinante verifica, além disso, as seguintes propriedades:

1.  $\det (AB) = \det A \det B$
2.  $\det (A^t) = \det A$
3.  $\det A \neq 0$  se e só se  $A$  for inversível.
4. Se  $A$  for inversível, então  $\det (A^{-1}) = (\det A)^{-1}$
5.  $\det (P^{-1}AP) = \det A$ , se  $P$  for inversível.
6. Se  $\tilde{A}$  se obter a partir de  $A$ , usando as transformações elementares sobre  $A$  então:
  - $\det \tilde{A} = \lambda \det A$ , se  $\tilde{A}$  se obter a partir de  $A$  multiplicando uma linha (ou uma coluna) por  $\lambda \in \mathbb{k}$ .
  - $\det \tilde{A} = -\det A$ , se  $\tilde{A}$  se obter a partir de  $A$  permutando duas linhas (ou duas colunas).
  - $\det \tilde{A} = \det A$ , se  $\tilde{A}$  se obter a partir de  $A$  substituindo uma linha (respectivamente, uma coluna) pela que se obter somando a essa linha (respectivamente, coluna) um múltiplo escalar de uma outra.
7. O  $\det A$  pode ser obtido pela seguinte **regra de Laplace**: fixamos uma qualquer linha  $i$  da matriz  $A = [A_j^i]$  e “desenvolvemos segundo esta linha”:

$$\det A = \sum_{j=1}^n (-1)^{i+j} A_j^i \det \hat{A}_j^i \quad (1.1.28)$$

onde  $\hat{A}_j^i$  representa a matriz  $(n-1) \times (n-1)$  que se obter a partir de  $A$ , omitindo a linha  $i$  e a coluna  $j$ .

■

Estas propriedades serão usadas sistematicamente no cálculo prático de determinantes. Finalmente, se  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  é um operador linear num espaço vectorial de dimensão  $n$ , sobre  $\mathbb{k}$ , define-se o respectivo **determinante**  $\det \mathbf{L}$ , como sendo o determinante da matriz de  $L$ , relativamente a uma qualquer base de  $\mathcal{V}$ . Como  $\det (P^{-1}LP) = \det L$ , se  $P$  for inversível, esta definição não depende da base escolhida.

## 1.2 Produto vectorial. Produto misto (ou triplo) em $\mathbb{R}^3$ . Interpretação geométrica do determinante

► **1.5 Produto vectorial em  $\mathbb{R}^3$**  ... Começemos por recordar o que é o **produto vectorial** de dois vectores em  $\mathbb{R}^3$ . Dados dois vectores  $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$  e  $\mathbf{x}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$ , em  $\mathbb{R}^3$ , define-se o **produto vectorial**  $\mathbf{x} \times \mathbf{x}'$ , de  $\mathbf{x}$  por  $\mathbf{x}'$ , como sendo o seguinte vector de  $\mathbb{R}^3$ :

$$\mathbf{x} \times \mathbf{x}' \stackrel{\text{def}}{=} (yz' - y'z)\mathbf{i} + (zx' - z'x)\mathbf{j} + (xy' - x'y)\mathbf{k} \quad (1.2.1)$$

O produto vectorial  $\mathbf{x} \times \mathbf{y}$ , pode ser obtido desenvolvendo segundo a primeira linha, o determinante formal:

$$\mathbf{x} \times \mathbf{y} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ x' & y' & z' \end{bmatrix}$$

► **1.6 Propriedades** ... A seguir indicam-se as propriedades mais importantes deste produto vectorial, todas elas de demonstração simples (que deve ser feita como exercício).

- O produto vectorial é bilinear:

$$\begin{aligned} (\mathbf{x} + \mathbf{y}) \times \mathbf{z} &= \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z} \\ \mathbf{x} \times (\mathbf{y} + \mathbf{z}) &= \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z} \\ \lambda \mathbf{x} \times \mathbf{y} &= \mathbf{x} \times \lambda \mathbf{y} = \lambda(\mathbf{x} \times \mathbf{y}), \quad \lambda \in \mathbb{R}, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3 \end{aligned} \quad (1.2.2)$$

- O produto vectorial é antissimétrico:

$$\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x} \quad (1.2.3)$$

- Além disso, se  $\mathbf{x} \in \mathbb{R}^3$  e  $\mathbf{y} \in \mathbb{R}^3$ , são ambos não nulos, então:

1.  $\mathbf{x} \times \mathbf{y}$  é perpendicular a  $\mathbf{x}$  e a  $\mathbf{y}$ , i.e.:

$$(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x} = 0 = (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{y} \quad (1.2.4)$$

Se  $\mathbf{x}$  e  $\mathbf{y}$  são linearmente independentes,  $\mathbf{x} \times \mathbf{y}$  é perpendicular ao plano gerado por  $\mathbf{x}$  e  $\mathbf{y}$ .

- 2.

$$\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\|\|\mathbf{y}\|\sin \theta \quad (1.2.5)$$

onde  $\theta$  é o ângulo entre  $\mathbf{x}$  e  $\mathbf{y}$ . Portanto,  $\|\mathbf{x} \times \mathbf{y}\|$  é igual à área do paralelogramo cujos lados adjacentes são  $\mathbf{x}$  e  $\mathbf{y}$ .

3.  $\mathbf{x} \times \mathbf{y} = \mathbf{0} \Leftrightarrow \mathbf{x}$  e  $\mathbf{y}$  são linearmente dependentes.

4. O produto vectorial não é associativo. De facto:

$$(\mathbf{x} \times \mathbf{y}) \times \mathbf{z} = (\mathbf{x} \cdot \mathbf{z})\mathbf{y} - (\mathbf{y} \cdot \mathbf{z})\mathbf{x} \quad (1.2.6)$$

enquanto que:

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) = (\mathbf{x} \cdot \mathbf{z})\mathbf{y} - (\mathbf{x} \cdot \mathbf{y})\mathbf{z} \quad (1.2.7)$$

Em particular, se consideramos o paralelogramo de lados adjacentes  $\mathbf{x} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$  e  $\mathbf{x}' = \begin{bmatrix} x' \\ y' \\ 0 \end{bmatrix}$ , contido no plano  $z = 0$ , vemos que a respectiva área é dada por:

$$\begin{aligned} \|\mathbf{x} \times \mathbf{x}'\| &= \left| \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & 0 \\ x' & y' & 0 \end{bmatrix} \right| \\ &= \left| \det \begin{bmatrix} x & y \\ x' & y' \end{bmatrix} \right| \\ &= xy' - x'y \\ &= \text{área do paralelogramo gerado por } \mathbf{x} \text{ e } \mathbf{x}' \end{aligned} \quad (1.2.8)$$

Uma equação (cartesiana) para o plano vectorial  $\text{span}_{\mathbb{R}}\{\mathbf{u}, \mathbf{v}\}$ , gerado por dois vectores  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3 - \{\mathbf{0}\}$ , linearmente independentes, é:

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{x} = 0 \quad (1.2.9)$$

► **1.7 Produto misto (ou triplo) em  $\mathbb{R}^3$  ...** Definamos agora, ainda em  $\mathbb{R}^3$ , o chamado produto misto (ou triplo).

Dados três vectores  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  em  $\mathbb{R}^3$ , define-se o **produto misto (ou triplo)**  $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$ , de  $\mathbf{x}, \mathbf{y}$  e  $\mathbf{z}$  (por esta ordem), através de:

$$[\mathbf{x}, \mathbf{y}, \mathbf{z}] \equiv \mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) \quad (1.2.10)$$

É fácil ver que  $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$  é dado por:

$$\begin{aligned} [\mathbf{x}, \mathbf{y}, \mathbf{z}] &= \det [\mathbf{x} \ \mathbf{y} \ \mathbf{z}] \\ &= \det \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} \end{aligned} \quad (1.2.11)$$

► **1.8 Propriedades ...** Eis algumas propriedades do produto triplo:

- São válidas as igualdades seguintes, que se deduzem das propriedades sobre determinantes:

$$\begin{aligned} [\mathbf{x}, \mathbf{y}, \mathbf{z}] &= [\mathbf{y}, \mathbf{z}, \mathbf{x}] = [\mathbf{z}, \mathbf{x}, \mathbf{y}] = -[\mathbf{y}, \mathbf{x}, \mathbf{z}] \\ &= -[\mathbf{x}, \mathbf{z}, \mathbf{y}] = -[\mathbf{z}, \mathbf{y}, \mathbf{x}] \end{aligned} \quad (1.2.12)$$

- O volume  $\text{vol}(\mathbf{x}, \mathbf{y}, \mathbf{z})$ , do paralelepípedo de lados adjacentes  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ , é igual ao módulo do produto misto:

$$\text{vol}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = |[\mathbf{x}, \mathbf{y}, \mathbf{z}]| \quad (1.2.13)$$

Com efeito, o volume de um paralelepípedo é igual ao produto da área da base pela sua altura. A base é o paralelogramo de lados adjacentes  $\mathbf{x}$  e  $\mathbf{y}$ , e por isso, a sua área é  $\|\mathbf{x} \times \mathbf{y}\|$ . A altura é igual à norma da projecção de  $\mathbf{z}$  sobre um vector perpendicular à base. Mas  $\mathbf{x} \times \mathbf{y}$  é perpendicular à base, e, portanto, a projecção de  $\mathbf{z}$  sobre  $\mathbf{x} \times \mathbf{y}$ , é igual a:

$$\frac{\mathbf{z} \cdot (\mathbf{x} \times \mathbf{y})}{\|\mathbf{x} \times \mathbf{y}\|_2} (\mathbf{x} \times \mathbf{y}) \quad (1.2.14)$$

donde se deduz facilmente o resultado.

Quando  $\mathbf{x}_1, \mathbf{x}_2$  e  $\mathbf{x}_3$  são linearmente independentes, de tal forma que:

$$\det[\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3] \neq 0$$

dizemos que a **base ordenada**  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$  é **positiva** se  $\det[\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3] > 0$ , e **negativa** se  $\det[\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3] < 0$ .

### 1.3 Interpretação geométrica do det A

Consideremos agora uma aplicação linear  $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ . A imagem do cubo  $\mathcal{Q} \subset \mathbb{R}^3$ , gerado pelos vectores da base canónica (que é positiva)  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ :

$$\mathcal{Q} = \{a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3 : 0 \leq a, b, c \leq 1\}$$

é o paralelepípedo  $\mathbf{A}(\mathcal{Q})$ , de lados adjacentes  $\mathbf{A}(\mathbf{e}_1)$ ,  $\mathbf{A}(\mathbf{e}_2)$  e  $\mathbf{A}(\mathbf{e}_3)$ .

Pondo  $\mathbf{A}(\mathbf{e}_1) = a_1^1\mathbf{e}_1 + a_1^2\mathbf{e}_2 + a_1^3\mathbf{e}_3 = \begin{bmatrix} a_1^1 \\ a_1^2 \\ a_1^3 \end{bmatrix}$ ,  $\mathbf{A}(\mathbf{e}_2) = a_2^1\mathbf{e}_1 + a_2^2\mathbf{e}_2 + a_2^3\mathbf{e}_3 = \begin{bmatrix} a_2^1 \\ a_2^2 \\ a_2^3 \end{bmatrix}$ , e

$\mathbf{A}(\mathbf{e}_3) = a_3^1\mathbf{e}_1 + a_3^2\mathbf{e}_2 + a_3^3\mathbf{e}_3 = \begin{bmatrix} a_3^1 \\ a_3^2 \\ a_3^3 \end{bmatrix}$  sabemos que o volume deste paralelepípedo é igual a:

$$\begin{aligned} \text{vol} \mathbf{A}(\mathcal{Q}) &= |[\mathbf{A}(\mathbf{e}_1), \mathbf{A}(\mathbf{e}_2), \mathbf{A}(\mathbf{e}_3)]| \\ &= |\det[\mathbf{A}(\mathbf{e}_1) \ \mathbf{A}(\mathbf{e}_2) \ \mathbf{A}(\mathbf{e}_3)]| \\ &= \left| \det \begin{bmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{bmatrix} \right| \\ &= |\det \mathbf{A}| \end{aligned} \quad (1.3.1)$$

Portanto:

$$\text{vol} \mathbf{A}(\mathcal{Q}) = |\det \mathbf{A}| \quad (1.3.2)$$

Mais geralmente, se  $\mathcal{P}$  é um paralelepípedo gerado pelos vectores  $\mathbf{x}, \mathbf{y}$  e  $\mathbf{z}$ , então a imagem  $\mathbf{A}(\mathcal{P})$  é o paralelepípedo gerado por  $\mathbf{A}(\mathbf{x}), \mathbf{A}(\mathbf{y})$  e  $\mathbf{A}(\mathbf{z})$ , e é fácil provar que o volume dessa imagem é igual a:

$$\begin{aligned} \text{vol} \mathbf{A}(\mathcal{P}) &= |[\mathbf{A}(\mathbf{x}), \mathbf{A}(\mathbf{y}), \mathbf{A}(\mathbf{z})]| \\ &= |\det[\mathbf{A}(\mathbf{x}) \ \mathbf{A}(\mathbf{y}) \ \mathbf{A}(\mathbf{z})]| \\ &= |\det \mathbf{A}| \text{vol}(\mathcal{P}) \end{aligned} \quad (1.3.3)$$

Em particular, se os vectores  $\mathbf{x}$ ,  $\mathbf{y}$  e  $\mathbf{z}$  são linearmente independentes, de tal forma que  $\text{vol } \mathcal{P} \neq 0$ , então:

$$|\det \mathbf{A}| = \frac{\text{vol } \mathbf{A}(\mathcal{P})}{\text{vol } \mathcal{P}} \quad (1.3.4)$$

Diz-se que uma aplicação linear inversível  $\mathbf{A} : \mathbb{R}_3 \rightarrow \mathbb{R}_3$  **preserva a orientação** (ou é positiva) se  $\det \mathbf{A} > 0$ , e que **inverte a orientação** (ou é negativa) se  $\det \mathbf{A} < 0$ .

## Capítulo 2

# Espaços vectoriais com produto interno

### 2.1 Espaços Euclidianos reais

► **2.1 Definição** ... Seja  $\mathcal{V}$  um espaço vectorial real. Um **produto interno** em  $\mathcal{V}$  é, por definição, uma aplicação:

$$\begin{aligned} \langle | \rangle : \mathcal{V} \times \mathcal{V} &\longrightarrow \mathbb{R} \\ (\mathbf{u}, \mathbf{v}) &\longmapsto \langle \mathbf{u} | \mathbf{v} \rangle \end{aligned} \quad (2.1.1)$$

que satisfaz as três propriedades seguintes:

[PI1]. **é uma forma bilinear:**

$$\begin{aligned} \langle (\mathbf{u} + \mathbf{v}) | \mathbf{w} \rangle &= \langle \mathbf{u} | \mathbf{w} \rangle + \langle \mathbf{v} | \mathbf{w} \rangle \\ \langle \mathbf{u} | (\mathbf{v} + \mathbf{w}) \rangle &= \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{u} | \mathbf{w} \rangle \\ \langle \lambda \mathbf{u} | \mathbf{v} \rangle &= \langle \mathbf{u} | \lambda \mathbf{v} \rangle = \lambda \langle \mathbf{u} | \mathbf{v} \rangle \end{aligned} \quad (2.1.2)$$

[PI2]. **é uma forma simétrica:**

$$\langle \mathbf{u} | \mathbf{v} \rangle = \langle \mathbf{v} | \mathbf{u} \rangle \quad (2.1.3)$$

[PI3]. **é não degenerada:**

$$\langle \mathbf{u} | \mathbf{v} \rangle = 0 \quad \forall \mathbf{v} \in \mathcal{V} \quad \Rightarrow \quad \mathbf{u} = \mathbf{0} \quad (2.1.4)$$

$\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}, \forall \lambda \in \mathbb{R}$ . Um produto interno diz-se um **produto interno Euclidiano**, se satisfaz além disso a seguinte propriedade:

[PI4]. **é uma forma definida positiva:**

$$\langle \mathbf{u} | \mathbf{u} \rangle \geq 0, \quad \forall \mathbf{u} \in \mathcal{V} \quad (2.1.5)$$

Um espaço vectorial real, munido de um produto interno Euclidiano chama-se um **espaço Euclidiano**. Outras notações muito comuns para  $\langle \mathbf{u} | \mathbf{v} \rangle$  são por exemplo  $\langle \mathbf{u}, \mathbf{v} \rangle$ ,  $\beta(\mathbf{u}, \mathbf{v})$ ,  $g(\mathbf{u}, \mathbf{v})$ ,  $\mathbf{u} \cdot \mathbf{v}$  ou ainda  $\mathbf{u} | \mathbf{v}$ .

► **2.2 Exemplo [Produto interno Euclidiano usual em  $\mathbb{R}^n$ ]** ... Dados dois vectores  $\mathbf{x} = [x_i]$  e  $\mathbf{y} = [y_i]$ , em  $\mathbb{R}^n$ , define-se o respectivo **produto interno (Euclidiano)**, como sendo o escalar  $\mathbf{x} \cdot \mathbf{y} \in \mathbb{R}$ , dado por:

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &\stackrel{\text{def}}{=} \sum_{i=1}^n x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\ &= \mathbf{x}^t \mathbf{y} \quad \text{em notação matricial} \end{aligned} \quad (2.1.6)$$

O espaço vectorial  $\mathbb{R}^n$ , munido deste produto interno Euclidiano, diz-se o **espaço Euclidiano** usual e nota-se por  $\mathbb{E}^n$ .

► **2.3 Exemplo [Produto interno  $L^2$  em  $C^o([a, b], \mathbb{R})$ ]** ... Consideremos o espaço vectorial real constituído pelas funções contínuas reais, definidas no intervalo  $[a, b] \subset \mathbb{R}$ . Dadas duas funções  $f, g \in C^o([a, b], \mathbb{R})$ , define-se o respectivo **produto interno  $L^2$** , como sendo o escalar  $\langle f|g \rangle \in \mathbb{R}$ , dado por:

$$\langle f|g \rangle \stackrel{\text{def}}{=} \int_a^b f(t)g(t) dt \quad (2.1.7)$$

► **2.4 Exemplo [Produto interno de Minkowski em  $\mathbb{R}^4$ ]** ... Dados dois vectores  $\mathbf{x} =$

$\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}$  e  $\mathbf{y} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix}$ , em  $\mathbb{R}^4$ , define-se o respectivo **produto interno de Minkowski**, como sendo o escalar  $\mathbf{x} \cdot \mathbf{y} \in \mathbb{R}$ , dado por:

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= -x_0 y_0 + x_1 y_1 + x_2 y_2 + x_3 y_3 \\ &= [-x_0 \ x_1 \ x_2 \ x_3] \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \\ &= \mathbf{x}^t \eta \mathbf{y} \end{aligned} \quad (2.1.8)$$

onde  $\eta$  representa a matriz simétrica:

$$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.1.9)$$

O produto interno de Minkowski **não** é definido positivo, isto é, **não** é verdade que  $\mathbf{x} \cdot \mathbf{x} \geq 0$ ,  $\forall \mathbf{x} \in \mathbb{R}^4$ . Com efeito, por exemplo o vector  $\mathbf{e}_0 = (1, 0, 0, 0)$ , satisfaz  $\mathbf{e}_0 \cdot \mathbf{e}_0 = -1$ . Note no entanto que a restrição do produto escalar de Minkowski ao hiperplano  $\{0\} \times \mathbb{R}^3 = \{\mathbf{x} = (x^\alpha) \in \mathbb{R}^4 : x^0 = 0\} \cong \mathbb{R}^3$ , é um produto interno euclidiano, portanto em particular definido positivo.

► **2.5 Expressões matriciais** ... Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço vectorial real, de dimensão  $n$ , com um produto interno Euclidiano.

Seja  $\mathcal{C} = [\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n]$  uma base qualquer para  $\mathcal{V}$ , escrita como um vector-linha com entradas vectoriais  $\mathbf{e}_i$ . Se  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$  podemos escrever:

$$\begin{aligned} \mathbf{v} &= \sum_i v^i \mathbf{e}_i \\ &= [\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n] \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix} \\ &= \mathcal{C}[\mathbf{v}]_{\mathcal{C}} \end{aligned} \tag{2.1.10}$$

onde  $[\mathbf{v}]_{\mathcal{C}} = \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}$  é o vector-coluna das componentes do vector  $\mathbf{v}$  na base  $\mathcal{C}$ . Analogamente:

$$\mathbf{u} = \sum_i u^i \mathbf{e}_i = \mathcal{C}[\mathbf{u}]_{\mathcal{C}}$$

Calculemos agora o produto interno  $\langle \mathbf{u} | \mathbf{v} \rangle$ :

$$\begin{aligned} \langle \mathbf{u} | \mathbf{v} \rangle &= \left\langle \sum_i u^i \mathbf{e}_i \middle| \sum_j v^j \mathbf{e}_j \right\rangle \\ &= \sum_{i,j} u^i v^j \langle \mathbf{e}_i | \mathbf{e}_j \rangle \\ &= \sum_{i,j} g_{ij} u^i v^j \\ &= [\mathbf{u}]_{\mathcal{C}}^T G_{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}} \end{aligned} \tag{2.1.11}$$

onde definimos a chamada **matriz de Gram**,  $G_{\mathcal{C}} = [g_{ij}]$ , do produto interno  $\langle | \rangle$ , na base  $\mathcal{C}$  através de:

$$g_{ij} \stackrel{\text{def}}{=} \langle \mathbf{e}_i | \mathbf{e}_j \rangle \tag{2.1.12}$$

Como  $\langle \mathbf{u} | \mathbf{v} \rangle = \langle \mathbf{v} | \mathbf{u} \rangle$ , deduzimos que a matriz de Gram  $G_{\mathcal{C}}$  é simétrica, isto é:

$$G_{\mathcal{C}}^T = G_{\mathcal{C}}$$

Como  $\langle \mathbf{v} | \mathbf{v} \rangle > 0, \forall \mathbf{v} \neq \mathbf{0} \in \mathcal{V}$  deduzimos que a matriz de Gram  $G_{\mathcal{C}}$  é definida positiva, isto é:

$$[\mathbf{v}]_{\mathcal{C}}^T G_{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}} = \sum_{i,j} g_{ij} v^i v^j > 0, \quad \forall v^i \text{ não simultaneamente nulos}$$

É possível provar os critérios seguintes (necessários e suficientes) para decidir quando uma matriz simétrica  $G = [g_{ij}]$  é definida positiva:

$n = 2$

$$g_{ij} > 0, \quad \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} > 0$$

$n = 3$

$$g_{ij} > 0, \quad \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} > 0$$

## 2.2 Espaços Hermitianos (ou Unitários) complexos

► **2.6 Definição** ... Seja  $\mathcal{V}$  um espaço vectorial complexo. Um **produto interno Hermitiano** em  $\mathcal{V}$  é, por definição, uma aplicação:

$$\begin{aligned} \langle | \rangle : \mathcal{V} \times \mathcal{V} &\longrightarrow \mathbb{C} \\ (\mathbf{u}, \mathbf{v}) &\longmapsto \langle \mathbf{u} | \mathbf{v} \rangle \end{aligned} \quad (2.2.1)$$

que satisfaz as propriedades seguintes:

[PH1]. **é uma forma sesquilinear**, isto é, é linear na primeira variável e semi-linear na segunda variável <sup>1</sup>:

$$\begin{aligned} \langle (\mathbf{u} + \mathbf{v}) | \mathbf{w} \rangle &= \langle \mathbf{u} | \mathbf{w} \rangle + \langle \mathbf{v} | \mathbf{w} \rangle \\ \langle \mathbf{u} | (\mathbf{v} + \mathbf{w}) \rangle &= \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{u} | \mathbf{w} \rangle \end{aligned} \quad (2.2.2)$$

$$\begin{aligned} \langle \lambda \mathbf{u} | \mathbf{v} \rangle &= \lambda \langle \mathbf{u} | \mathbf{v} \rangle \\ \langle \mathbf{u} | \lambda \mathbf{v} \rangle &= \bar{\lambda} \langle \mathbf{u} | \mathbf{v} \rangle \end{aligned} \quad (2.2.3)$$

[PH2]. **é uma forma Hermitiana**:

$$\langle \mathbf{u} | \mathbf{v} \rangle = \overline{\langle \mathbf{v} | \mathbf{u} \rangle} \quad (2.2.4)$$

[PH3]. **é não degenerada**:

$$\langle \mathbf{u} | \mathbf{v} \rangle = 0 \quad \forall \mathbf{v} \in \mathcal{V} \quad \Rightarrow \quad \mathbf{u} = \mathbf{0} \quad (2.2.5)$$

[PH4]. **é definida positiva**:

$$\langle \mathbf{u} | \mathbf{u} \rangle \geq 0 \quad (2.2.6)$$

$\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}, \forall \lambda \in \mathbb{C}$ .

Um espaço vectorial complexo, munido de um produto interno Hermitiano chama-se um **espaço Hermitiano** ou um **espaço unitário**.

► **2.7 Exemplo** [**Produto interno Hermitiano usual em  $\mathbb{C}^n$** ] ... Dados dois vectores  $\mathbf{z} = [z_i]$  e  $\mathbf{w} = [w_i]$ , em  $\mathbb{C}^n$ , define-se o respectivo **produto interno (Hermitiano)**, como sendo o escalar  $\langle \mathbf{z} | \mathbf{w} \rangle \in \mathbb{C}$ , dado por:

$$\begin{aligned} \langle \mathbf{z} | \mathbf{w} \rangle &\stackrel{\text{def}}{=} \sum_{i=1}^n z_i \bar{w}_i = z_1 \bar{w}_1 + z_2 \bar{w}_2 + \cdots + z_n \bar{w}_n \\ &= [z_1 \quad z_2 \quad \cdots \quad z_n] \begin{bmatrix} \bar{w}_1 \\ \bar{w}_2 \\ \vdots \\ \bar{w}_n \end{bmatrix} \\ &= \mathbf{z}^t \bar{\mathbf{w}} \quad \text{em notação matricial} \end{aligned} \quad (2.2.7)$$

O espaço vectorial  $\mathbb{C}^n$ , munido deste produto interno Euclideano, diz-se o **espaço unitário usual** e nota-se por  $\mathbf{U}^n$ .

<sup>1</sup>em Física, nomeadamente em Mecânica Quântica, é usual considerar outra convenção - linearidade na segunda variável e semi-linearidade na primeira variável!

► **2.8 Exemplo** [Produto interno  $L^2$  em  $C^o([a, b], \mathbb{C})$ ] ... Consideremos o espaço vectorial real constituído pelas funções contínuas complexas, definidas no intervalo  $[a, b] \subset \mathbb{R}$ . Dadas duas funções  $f, g \in C^o([a, b], \mathbb{C})$ , define-se o respectivo **produto interno**  $L^2$ , como sendo o escalar  $\langle f|g \rangle \in \mathbb{C}$ , dado por:

$$\langle f|g \rangle \stackrel{\text{def}}{=} \int_a^b f(t)\overline{g(t)} dt \quad (2.2.8)$$

## 2.3 Norma

► **2.9 Definição** [Norma] ... Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço com um produto interno (*Euclideo se  $\mathcal{V}$  é real ou Hermitiano se  $\mathcal{V}$  é complexo*). Define-se a **norma**  $\|\mathbf{v}\|$ , de um vector  $\mathbf{v} \in \mathcal{V}$ , através da fórmula:

$$\|\mathbf{v}\| \stackrel{\text{def}}{=} \sqrt{\langle \mathbf{v}|\mathbf{v} \rangle} \quad (2.3.1)$$

► **2.10** A norma verifica as propriedades seguintes:

[N1]. é **positiva e não degenerada**:

$$\|\mathbf{v}\| \geq 0 \quad \text{e} \quad \|\mathbf{v}\| = 0 \quad \text{sse} \quad \mathbf{v} = \mathbf{0} \quad (2.3.2)$$

[N2]. é **homogénea (positiva)**:

$$\|\lambda\mathbf{v}\| = |\lambda| \|\mathbf{v}\| \quad (2.3.3)$$

[N3]. **satisfaz a “desigualdade triangular” seguinte**:

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\| \quad (2.3.4)$$

$\forall \mathbf{v}, \mathbf{w} \in \mathcal{V}, \forall \lambda \in \mathbb{k} = \mathbb{R} \text{ ou } \mathbb{C}$ .

Todas as propriedades são de demonstração imediata com excepção da desigualdade triangular, que resulta da seguinte proposição:

► **2.11 Proposição** [Desigualdade de Cauchy-Schwarz] ...

$$|\langle \mathbf{v}|\mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|, \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V} \quad (2.3.5)$$

**Dem.:** Se  $\mathbf{w} = \mathbf{0}$  a desigualdade é trivial. Se  $\mathbf{w} \neq \mathbf{0}$  consideremos o vector:

$$\mathbf{u} = \mathbf{v} - \frac{\langle \mathbf{v}|\mathbf{w} \rangle}{\|\mathbf{w}\|^2} \mathbf{w}$$

de tal forma que  $\langle \mathbf{u}|\mathbf{w} \rangle = 0$ . Temos então que:

$$\begin{aligned} 0 \leq \|\mathbf{u}\|^2 &= \left\langle \left( \mathbf{v} - \frac{\langle \mathbf{v}|\mathbf{w} \rangle}{\|\mathbf{w}\|^2} \mathbf{w} \right) \middle| \left( \mathbf{v} - \frac{\langle \mathbf{v}|\mathbf{w} \rangle}{\|\mathbf{w}\|^2} \mathbf{w} \right) \right\rangle \\ &= \langle \mathbf{v}|\mathbf{v} \rangle - \frac{\langle \mathbf{v}|\mathbf{w} \rangle \langle \mathbf{w}|\mathbf{v} \rangle}{\|\mathbf{w}\|^2} \\ &= \|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v}|\mathbf{w} \rangle|^2}{\|\mathbf{w}\|^2} \end{aligned} \quad (2.3.6)$$

o que demonstra a desigualdade.

► **2.12** Demonstremos agora a desigualdade triangular (2.3.4):

$$\begin{aligned}
 \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v} | \mathbf{u} + \mathbf{v} \rangle \\
 &= \langle \mathbf{u} | \mathbf{u} \rangle + \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{v} | \mathbf{u} \rangle + \langle \mathbf{v} | \mathbf{v} \rangle \\
 &= \|\mathbf{u}\|^2 + \langle \mathbf{u} | \mathbf{v} \rangle + \overline{\langle \mathbf{u} | \mathbf{v} \rangle} + \|\mathbf{v}\|^2 \\
 &= \|\mathbf{u}\|^2 + 2\operatorname{Re} \langle \mathbf{u} | \mathbf{v} \rangle + \|\mathbf{v}\|^2 \\
 &\leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u} | \mathbf{v} \rangle| + \|\mathbf{v}\|^2 \\
 &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2, \quad \text{pela desigualdade de Cauchy-Schwarz (2.3.5)} \\
 &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2
 \end{aligned}$$

e portanto  $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ , como se pretendia.

► **2.13 Exemplos** ... (i). No espaço Euclidiano  $\mathbb{E}^n$ , a norma de um vector  $\mathbf{x} = (x_i) \in \mathbb{R}^n$  é dada pelo teorema de Pitágoras:

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^t \mathbf{x}} = \left[ \sum_{i=1}^n (x_i)^2 \right]^{1/2} \quad (2.3.7)$$

(ii). No espaço Unitário  $\mathbf{U}^n$ , a norma de um vector  $\mathbf{z} = (z_i) \in \mathbb{C}^n$  é dada por:

$$\|\mathbf{z}\| = \sqrt{\mathbf{z}^t \bar{\mathbf{z}}} = \left[ \sum_{i=1}^n |z_i|^2 \right]^{1/2} \quad (2.3.8)$$

(iii). No espaço Unitário  $C^o([a, b], \mathbb{C})$ , munido do produto interno  $L^2$ , dado por (2.2.8):  $\langle f | g \rangle \stackrel{\text{def}}{=} \int_a^b f(t) \overline{g(t)} dt$ , a norma de uma função  $f \in C^o([a, b], \mathbb{C})$  é dada por:

$$\|f\| = \sqrt{\langle f | f \rangle} = \left[ \int_a^b |f(t)|^2 dt \right]^{1/2} \quad (2.3.9)$$

Neste exemplo, a desigualdade de Cauchy-Schwarz toma o aspecto:

$$\left| \int_a^b f(t) \overline{g(t)} dt \right| \leq \left[ \int_a^b |f(t)|^2 dt \right]^{1/2} \left[ \int_a^b |g(t)|^2 dt \right]^{1/2} \quad (2.3.10)$$

enquanto que a desigualdade triangular tem o aspecto seguinte:

$$\left[ \int_a^b |f(t) + g(t)|^2 dt \right]^{1/2} \leq \left[ \int_a^b |f(t)|^2 dt \right]^{1/2} + \left[ \int_a^b |g(t)|^2 dt \right]^{1/2} \quad (2.3.11)$$

## 2.4 Ortogonalidade

► **2.14 Definição** ... Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço com um produto interno (*Euclidiano se  $\mathcal{V}$  é real ou Hermitiano se  $\mathcal{V}$  é complexo*). Dois vectores  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$  dizem-se **ortogonais** se:

$$\langle \mathbf{u} | \mathbf{v} \rangle = 0 \quad (2.4.1)$$

► **2.15 Ângulo não orientado** ... Suponhamos agora que  $(\mathcal{V}, \langle | \rangle)$  é um espaço real Euclidiano. Dados dois vectores não nulos  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ , deduzimos da desigualdade de Cauchy-Schwarz que:

$$-1 \leq \frac{\langle \mathbf{u} | \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1 \quad (2.4.2)$$

o que permite definir o **ângulo (não orientado)**  $\theta = \theta(\mathbf{u}, \mathbf{v}) \in [0, \pi]$ , entre os referidos vectores não nulos  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ , como sendo o único  $\theta \in [0, \pi]$ , tal que:

$$\cos \theta = \frac{\langle \mathbf{u} | \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \in [-1, 1] \quad (2.4.3)$$

Portanto:

$$\langle \mathbf{u} | \mathbf{v} \rangle = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta(\mathbf{u}, \mathbf{v}) \quad (2.4.4)$$

Como vimos antes, dois vectores  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$  dizem-se **ortogonais** se  $\langle \mathbf{u} | \mathbf{v} \rangle = 0$ . Se ambos são não nulos isto significa que o ângulo  $\theta(\mathbf{u}, \mathbf{v})$  é igual a  $\pi/2$ .

► **2.16 Definição [Ortogonal de um subconjunto]** ... Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço com um produto interno (*Euclidiano se  $\mathcal{V}$  é real ou Hermitiano se  $\mathcal{V}$  é complexo*). Se  $S$  é um subconjunto não vazio de  $\mathcal{V}$ , define-se o **ortogonal de  $S$**  como sendo o subconjunto  $S^\perp$  de  $\mathcal{V}$  constituído por todos os vectores que são ortogonais a todos os vectores de  $S$ :

$$S^\perp \stackrel{\text{def}}{=} \{ \mathbf{u} \in \mathcal{V} : \langle \mathbf{u} | \mathbf{s} \rangle = 0, \forall \mathbf{s} \in S \} \quad (2.4.5)$$

Vamos verificar que  $S^\perp$  é um subespaço de  $\mathcal{V}$ . De facto, se  $\mathbf{u}, \mathbf{v} \in S^\perp$ , então  $\langle \mathbf{u} | \mathbf{s} \rangle = 0$  e  $\langle \mathbf{v} | \mathbf{s} \rangle = 0, \forall \mathbf{s} \in S$  e portanto  $\langle \mathbf{u} + \mathbf{v} | \mathbf{s} \rangle = \langle \mathbf{u} | \mathbf{s} \rangle + \langle \mathbf{v} | \mathbf{s} \rangle = 0, \forall \mathbf{s} \in S$ , i.e.,  $\mathbf{u} + \mathbf{v} \in S^\perp$ . Anàlogamente  $\lambda \mathbf{u} \in S^\perp, \forall \lambda \in \mathbb{k}$ , se  $\mathbf{u} \in S^\perp$ .

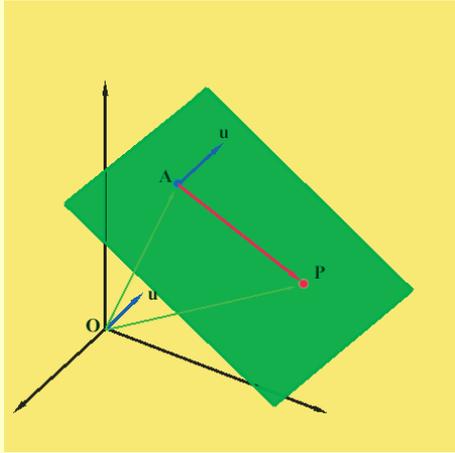
► **2.17 Hiperplanos vectoriais** ... No espaço Euclidiano  $\mathbb{E}^n$ , dado um vector não nulo  $\mathbf{u} \in \mathbb{R}^n - \{0\}$ , o conjunto dos vectores  $\mathbf{x} \in \mathbb{E}^n$  que são ortogonais a  $\mathbf{u}$ :

$$\{ \mathbf{x} \in \mathbb{E}^n : \mathbf{x} \cdot \mathbf{u} = 0 \} \quad (2.4.6)$$

formam um subespaço em  $\mathbb{E}^n$ , que se diz o **hiperplano (vectorial) ortogonal** a  $\mathbf{u}$ . Se  $\mathbf{x} = (x_i)$  é um ponto genérico desse hiperplano, e se  $\mathbf{u} = (u_i)$ , a equação  $\mathbf{x} \cdot \mathbf{u} = 0$ , é equivalente à seguinte **equação cartesiana**:

$$\sum_i u_i x_i = u_1 x_1 + u_2 x_2 + \cdots + u_n x_n = 0 \quad (2.4.7)$$

► **2.18 Hiperplanos afins em  $\mathbb{E}^n$**  ...



No espaço Euclidiano  $\mathbb{E}^n$ , com a estrutura afim canônica, dado um ponto  $A$  e um vetor não nulo  $\mathbf{u} \in \mathbb{R}^n - \{\mathbf{0}\}$ , o conjunto dos pontos  $P \in \mathbb{E}^n$  tais que  $\overrightarrow{AP} = P - A$  é ortogonal a  $\mathbf{u}$ :

$$\{P \in \mathbb{E}^n : \overrightarrow{AP} \cdot \mathbf{u} = 0\} \quad (2.4.8)$$

diz o **hiperplano (afim) ortogonal** a  $\mathbf{u}$ , que passa em  $A$ . Se  $\overrightarrow{OA} = (a_i)$ ,  $\mathbf{u} = (u_i)$  e se  $\overrightarrow{OP} = (x_i)$  é um ponto genérico desse hiperplano, a equação  $\overrightarrow{AP} \cdot \mathbf{u} = 0$ , é equivalente a:

$$0 = (\overrightarrow{OP} - \overrightarrow{OA}) \cdot \mathbf{u} = \overrightarrow{OP} \cdot \mathbf{u} - \overrightarrow{OA} \cdot \mathbf{u} = \sum_i u_i x_i - \sum_i a_i u_i$$

e portanto à seguinte **equação cartesiana**:

$$\sum_i u_i x_i = u_1 x_1 + u_2 x_2 + \dots + u_n x_n = c \quad (2.4.9)$$

onde  $c = \overrightarrow{OA} \cdot \mathbf{u} = \sum_i a_i u_i$ .

► **2.19 Teorema [Pitágoras]** ... *Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço com um produto interno (Euclidiano se  $\mathcal{V}$  é real ou Hermitiano se  $\mathcal{V}$  é complexo), e  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$  dois vectores ortogonais. Então:*

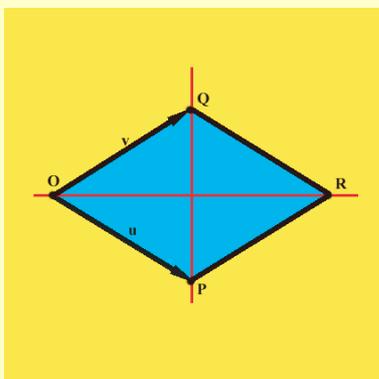
$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \quad (2.4.10)$$

**Dem.:**

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v} | \mathbf{u} + \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{v} | \mathbf{u} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \end{aligned} \quad (2.4.11)$$

## 2.5 Aplicações à geometria

► **2.20 Exemplo** ... As diagonais de um losango intersectam-se perpendicularmente.

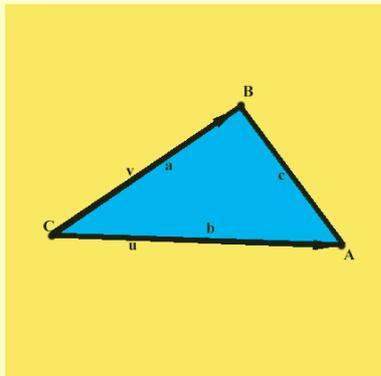


**Dem.:** Como  $OQRP$  é um losango,  $\|\mathbf{u}\| = \|\mathbf{v}\|$ . Pretende-se provar que  $\overrightarrow{QP} \perp \overrightarrow{OR}$ , isto é que,  $(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = 0$ . Mas:

$$(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2 = 0$$

► **2.21 Exemplo** [Lei dos cossenos] ... Num triângulo plano  $\triangle(ABC)$ , onde  $a = \overline{BC}$ , etc. tem-se que:

$$c^2 = a^2 + b^2 - 2ab \cos C$$



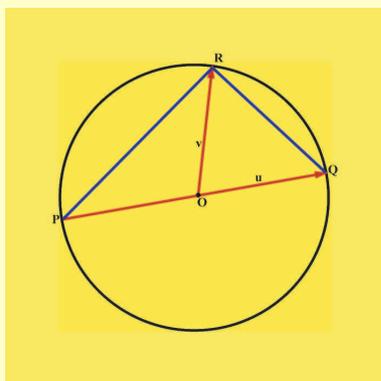
**Dem.:** Escolhamos um referencial com origem em  $C$ , e ponhamos  $\mathbf{u} = \overrightarrow{CA}$  e  $\mathbf{v} = \overrightarrow{CB}$ . Então  $\overrightarrow{AB} = \mathbf{v} - \mathbf{u}$ , e daí que:

$$\|\overrightarrow{AB}\|^2 = \|\mathbf{v} - \mathbf{u}\|^2 = \|\mathbf{v}\|^2 - 2\mathbf{u} \cdot \mathbf{v} + \|\mathbf{u}\|^2$$

ou, com as notações referidas:

$$c^2 = a^2 + b^2 - 2ab \cos C$$

► **2.22 Exemplo** ... Se  $R$  é um ponto sobre um círculo de diâmetro  $POQ$ , mostre que  $PR \perp QR$ .



**Dem.:** Seja  $\mathbf{u} = \overrightarrow{OQ}$ ,  $\mathbf{v} = \overrightarrow{OR}$ . Então

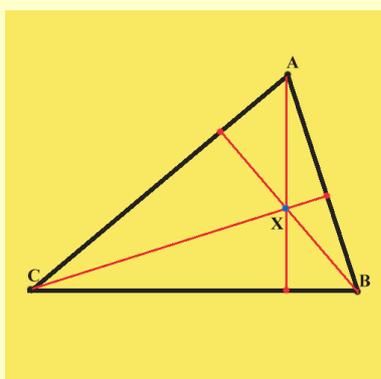
$$\overrightarrow{PR} = \overrightarrow{OR} - \overrightarrow{OP} = \mathbf{u} + \mathbf{v}$$

$$\overrightarrow{QR} = \overrightarrow{OR} - \overrightarrow{OQ} = \mathbf{v} - \mathbf{u}$$

Sabe-se que  $\|\mathbf{u}\| = \|\mathbf{v}\|$  e portanto:

$$\overrightarrow{PR} \cdot \overrightarrow{QR} = (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{v} - \mathbf{u}) = \|\mathbf{v}\|^2 - \|\mathbf{u}\|^2 = 0$$

► **2.23 Exemplo** ... As alturas de um triângulo intersectam-se num único ponto (chamado o **ortocentro** do triângulo).



**Dem.:** Pretende-se encontrar um ponto  $X$  tal que:

$$\overrightarrow{AX} \cdot \overrightarrow{BC} = 0, \quad \overrightarrow{BX} \cdot \overrightarrow{CA} = 0, \quad \overrightarrow{CX} \cdot \overrightarrow{AB} = 0$$

Identificando um ponto  $P$  com o seu vector de posição  $\overrightarrow{OP}$ , relativamente a uma origem fixa  $O$  no plano, é fácil verificar a identidade seguinte:

$$(X - A) \cdot (C - B) + (X - B) \cdot (A - C) + (X - C) \cdot (B - A) = 0 \quad (2.5.1)$$

Seja  $X$  o ponto de intersecção de duas das alturas, digamos, das alturas partindo de  $A$  e de  $B$ . Temos então que, lembrando que  $\overrightarrow{AX} = X - A$ , etc:

$$(X - A) \cdot (C - B) = 0 \quad (2.5.2)$$

$$(X - B) \cdot (A - C) = 0 \quad (2.5.3)$$

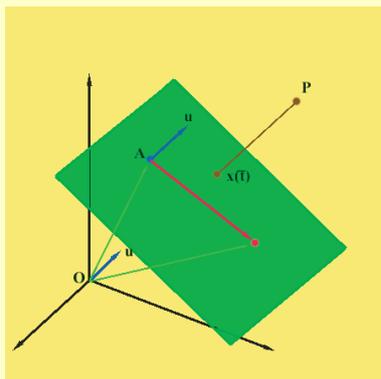
Subtraindo (2.5.2) e (2.5.3) de (2.5.1), obtemos:

$$(X - C) \cdot (B - A) = 0$$

como se pretendia.

► **2.24 Exemplo** ... Dados dois pontos distintos  $A \neq B$  no plano, mostrar que o lugar geométrico dos pontos  $P$  cuja distância a  $A$  é o dobro da distância a  $B$  é um círculo.

► **2.25 Exemplo** ... Calcular a distância entre um ponto  $P$  e um hiperplano afim em  $\mathbb{E}^n$ .



**Res...** Suponhamos que esse hiperplano é perpendicular ao vector  $\mathbf{u} \neq \mathbf{0}$  e passa num ponto  $\mathbf{a}$  e, portanto, tem equação:

$$(\mathbf{x} - \mathbf{a}) \cdot \mathbf{u} = 0$$

ou

$$\mathbf{x} \cdot \mathbf{u} + c = 0, \quad c = -\mathbf{a} \cdot \mathbf{u}$$

A recta que passa em  $P \simeq \overrightarrow{OP} = \mathbf{p}$  e tem a direcção do vector  $\mathbf{u}$ , tem equação:

$$\mathbf{x}(t) = \mathbf{p} + t\mathbf{u}$$

O ponto desta recta que pertence ao plano referido, corresponde ao valor do parâmetro  $\bar{t}$  que verifica:

$$0 = \mathbf{x}(\bar{t}) \cdot \mathbf{u} + c = (\mathbf{p} + \bar{t}\mathbf{u}) \cdot \mathbf{u} + c = \mathbf{p} \cdot \mathbf{u} + \bar{t}\|\mathbf{u}\|^2 + c \Rightarrow \bar{t} = -\frac{\mathbf{p} \cdot \mathbf{u} + c}{\|\mathbf{u}\|^2}$$

A distância entre um ponto  $P \simeq \mathbf{p}$  e o hiperplano afim é pois dada por:

$$d = \|\mathbf{p} - \mathbf{x}(\bar{t})\| = \left\| \mathbf{p} - \mathbf{p} + \frac{\mathbf{p} \cdot \mathbf{u} + c}{\|\mathbf{u}\|^2} \mathbf{u} \right\| = \frac{|\mathbf{p} \cdot \mathbf{u} + c|}{\|\mathbf{u}\|}$$

Assim por exemplo:

- No plano, a distância entre um ponto  $P = (\alpha, \beta)$  e a recta afim  $ax + by + c = 0$  é:

$$d = \frac{|\mathbf{p} \cdot \mathbf{u} + c|}{\|\mathbf{u}\|} = \frac{|(\alpha, \beta) \cdot (a, b) + c|}{\|(a, b)\|} = \frac{|a\alpha + b\beta + c|}{(a^2 + b^2)^{1/2}}$$

- No espaço, a distância entre um ponto  $P = (\alpha, \beta, \gamma)$  e o plano afim  $ex + fy + gz + h = 0$  é:

$$d = \frac{|\mathbf{p} \cdot \mathbf{u} + c|}{\|\mathbf{u}\|} = \frac{|(\alpha, \beta, \gamma) \cdot (e, f, g) + h|}{\|(e, f, g)\|} = \frac{|e\alpha + f\beta + g\gamma + h|}{(e^2 + f^2 + g^2)^{1/2}}$$

► **2.26 Exemplo** ... Calcular a distância entre um ponto  $P$  e uma recta afim em  $\mathbb{E}^3$ , quando:

1. essa recta é definida parametricamente.
2. essa recta é definida como intersecção de dois planos afins.

## 2.6 Bases ortonormadas num espaço vectorial com produto interno

► **2.27 Definição [Base ortonormada]** ... Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço vectorial de dimensão  $n$  com um produto interno (*Euclideano se  $\mathcal{V}$  é real ou Hermitiano se  $\mathcal{V}$  é complexo*).

Uma base  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  diz-se uma **base ortonormada** para  $\mathcal{V}$  se:

$$\langle \mathbf{e}_i | \mathbf{e}_j \rangle = \delta_{ij} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \quad (2.6.1)$$

► **2.28 Proposição** ... *Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço vectorial de dimensão  $n$  com um produto interno (Euclideano se  $\mathcal{V}$  é real ou Hermitiano se  $\mathcal{V}$  é complexo) e  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  uma base ortonormada para  $\mathcal{V}$ . Então  $\mathbf{v} \in \mathcal{V}$ :*

$$\mathbf{v} = \sum_{i=1}^n \langle \mathbf{v} | \mathbf{e}_i \rangle \mathbf{e}_i \quad (2.6.2)$$

e:

$$\|\mathbf{v}\|^2 = \sum_{i=1}^n |\langle \mathbf{v} | \mathbf{e}_i \rangle|^2 \quad (2.6.3)$$

**Dem.:** Cálculo directo.

## 2.7 Método de ortogonalização de Gram-Schmidt

► **2.29 Ortogonalização de Gram-Schmidt** ... Dada uma base qualquer  $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ , para  $\mathcal{V}$ , é possível construir, a partir dela, uma base ortogonal  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ , para  $\mathcal{V}$ :

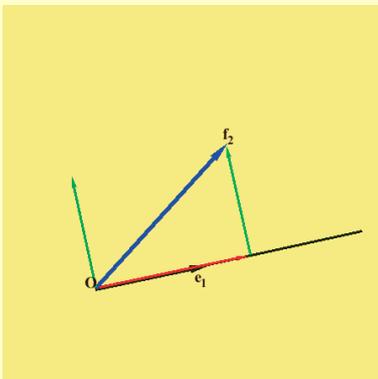
$$\langle \mathbf{e}_i | \mathbf{e}_j \rangle = 0, \quad i \neq j$$

através do chamado **processo de ortogonalização de Gram-Schmidt**, que passamos a descrever:

[1.] Em primeiro lugar põmos:

$$\mathbf{e}_1 = \mathbf{f}_1 \quad (2.7.1)$$

[2.]



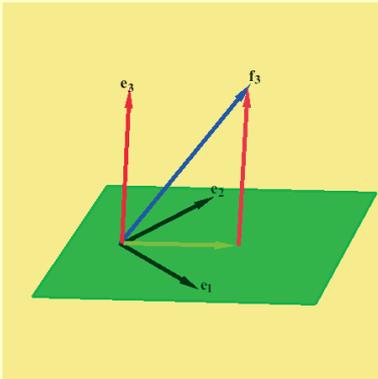
Em segundo lugar, começamos por calcular a chamada **projecção ortogonal de  $\mathbf{f}_2$**  sobre a recta gerada por  $\mathbf{f}_1 = \mathbf{e}_1$ . Esta projecção ortogonal, por estar na recta gerada por  $\mathbf{f}_1 = \mathbf{e}_1$ , vai ser um vector do tipo  $\lambda \mathbf{e}_1$ , onde  $\lambda \in \mathbb{K}$  é calculado pela condição de que  $\langle \mathbf{f}_2 - \lambda \mathbf{e}_1 | \mathbf{e}_1 \rangle = 0$ . Obtemos então:

$$\lambda = \frac{\langle \mathbf{f}_2 | \mathbf{e}_1 \rangle}{\|\mathbf{e}_1\|^2}$$

Pômos agora  $\mathbf{e}_2$  igual a:

$$\mathbf{e}_2 = \mathbf{f}_2 - \frac{\langle \mathbf{f}_2 | \mathbf{e}_1 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 \quad (2.7.2)$$

[3.]



Em terceiro lugar, começamos por calcular a chamada **projecção ortogonal de  $\mathbf{f}_3$**  sobre o plano gerado por  $\{\mathbf{f}_1, \mathbf{f}_2\}$ , que é também o plano gerado por  $\{\mathbf{e}_1, \mathbf{e}_2\}$ . Esta projecção ortogonal, por estar no plano gerado por  $\{\mathbf{e}_1, \mathbf{e}_2\}$ , vai ser um vector do tipo  $\lambda\mathbf{e}_1 + \eta\mathbf{e}_2$ , onde  $\lambda, \eta \in \mathbb{K}$  são calculados pela condição de que  $\langle \mathbf{f}_3 - (\lambda\mathbf{e}_1 + \eta\mathbf{e}_2) | \mathbf{e}_1 \rangle = 0$  e  $\langle \mathbf{f}_3 - (\lambda\mathbf{e}_1 + \eta\mathbf{e}_2) | \mathbf{e}_2 \rangle = 0$ . Fazendo os cálculos, atendendo a que  $\mathbf{e}_1 \perp \mathbf{e}_2$ , obtemos:

$$\lambda = \frac{\langle \mathbf{f}_3 | \mathbf{e}_1 \rangle}{\|\mathbf{e}_1\|^2}, \quad \eta = \frac{\langle \mathbf{f}_3 | \mathbf{e}_2 \rangle}{\|\mathbf{e}_2\|^2}$$

Portanto a projecção ortogonal de  $\mathbf{f}_3$  sobre o plano gerado por  $\{\mathbf{e}_1, \mathbf{e}_2\}$  é dada por:

$$\frac{\langle \mathbf{f}_3 | \mathbf{e}_1 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 + \frac{\langle \mathbf{f}_3 | \mathbf{e}_2 \rangle}{\|\mathbf{e}_2\|^2} \mathbf{e}_2$$

Pômos agora  $\mathbf{e}_3$  igual a:

$$\mathbf{e}_3 = \mathbf{f}_3 - \frac{\langle \mathbf{f}_3 | \mathbf{e}_1 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 - \frac{\langle \mathbf{f}_3 | \mathbf{e}_2 \rangle}{\|\mathbf{e}_2\|^2} \mathbf{e}_2 \quad (2.7.3)$$

[k.] o processo decorre agora indutivamente: se supômos já construídos os vectores ortogonais  $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ , de tal forma que:

$$\text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_k\} = \text{span}\{\mathbf{f}_1, \dots, \mathbf{f}_k\}$$

o vector  $\mathbf{e}_{k+1}$  será construído da seguinte forma - começamos por calcular a chamada **projecção ortogonal de  $\mathbf{f}_{k+1}$**  sobre o subespaço gerado por  $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ . Esta projecção ortogonal é dada por:

$$\sum_{i=1}^k \frac{\langle \mathbf{f}_{k+1} | \mathbf{e}_i \rangle}{\|\mathbf{e}_i\|^2} \mathbf{e}_i$$

Pômos agora  $\mathbf{e}_{k+1}$  igual a:

$$\mathbf{e}_{k+1} = \mathbf{f}_{k+1} - \sum_{i=1}^k \frac{\langle \mathbf{f}_{k+1} | \mathbf{e}_i \rangle}{\|\mathbf{e}_i\|^2} \mathbf{e}_i \quad (2.7.4)$$

É claro que a base ortogonal assim obtida, pode ser transformada numa base ortonormada, normalizando os vectores  $\mathbf{e}_i$ , isto é, dividindo cada um deles pela respectiva norma.

► **2.30 Polinómios de Legendre** ... Consideremos o espaço vectorial  $\mathcal{V}$  constituído por todas as funções polinomiais de grau  $\leq n$ , definidas no intervalo  $[-1, 1]$ , munido do produto interno  $L^2$ :

$$\langle p | q \rangle = \int_{-1}^1 p(t)q(t) dt$$

Uma base para  $\mathcal{V}$  é  $\{1, t, t^2, \dots, t^n\}$ . Quando aplicamos o processo de ortogonalização de Gram-Schmidt a esta base obtemos os chamados **polinômios de Legendre**  $\{\psi_0, \psi_1, \psi_2, \dots, \psi_n\}$ . Vejamos como. Em primeiro lugar põmos:

$$\psi_0(t) = 1$$

Depois põmos:

$$\begin{aligned} \psi_1 &= t - \frac{\langle t|1 \rangle}{\|1\|^2} 1 \\ &= t - \frac{\int_{-1}^1 t dt}{\|\int_{-1}^1 1^2 dt\|^2} 1 \\ &= t \end{aligned} \tag{2.7.5}$$

Em seguida:

$$\begin{aligned} \psi_2 &= t^2 - \frac{\langle t^2|1 \rangle}{\|1\|^2} 1 - \frac{\langle t^2|t \rangle}{\|t\|^2} t \\ &= t^2 - \frac{\int_{-1}^1 t^2 dt}{\|\int_{-1}^1 1^2 dt\|^2} 1 - \frac{\int_{-1}^1 t^3 dt}{\|\int_{-1}^1 t^2 dt\|^2} t \\ &= t^2 - \frac{1}{3} \end{aligned} \tag{2.7.6}$$

e procedendo da mesma forma:

$$\begin{aligned} \psi_3 &= t^3 - \frac{3}{5}t \\ \psi_4 &= t^4 - \frac{6}{7}t^2 + \frac{3}{35} \\ &\vdots \end{aligned} \tag{2.7.7}$$

Quando normalizamos estes polinômios obtemos os chamados **polinômios de Legendre normalizados**  $\{\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_n\}$ :

$$\begin{aligned} \varphi_0 &= \sqrt{\frac{1}{2}} \\ \varphi_1 &= \sqrt{\frac{3}{2}}t \\ \varphi_2 &= \frac{1}{2}\sqrt{\frac{5}{2}}(3t^2 - 1) \\ \varphi_3 &= \frac{1}{2}\sqrt{\frac{7}{2}}(5t^3 - 3t) \\ &\vdots \end{aligned} \tag{2.7.8}$$

## 2.8 Decomposição ortogonal. Teorema da aproximação óptima

► **2.31 Teorema [Decomposição ortogonal]** ... Consideremos um espaço vectorial com um produto interno  $(\mathcal{V}, \langle | \rangle)$  (Euclideano se  $\mathcal{V}$  é real ou Hermitiano se  $\mathcal{V}$  é complexo), e seja  $\mathcal{S}$  um

subespaço de dimensão finita. Então:

$$\mathcal{V} = \mathcal{S} \oplus \mathcal{S}^\perp \quad (2.8.1)$$

isto é, qualquer vector  $\mathbf{v} \in \mathcal{V}$  pode ser representado de maneira única como uma soma de dois vectores:

$$\mathbf{v} = \mathbf{s} + (\mathbf{v} - \mathbf{s}), \quad \text{onde } \mathbf{s} \in \mathcal{S} \text{ e } \mathbf{v} - \mathbf{s} \in \mathcal{S}^\perp \quad (2.8.2)$$

Além disso:

$$\|\mathbf{v}\|^2 = \|\mathbf{s}\|^2 + \|\mathbf{v} - \mathbf{s}\|^2 \quad (2.8.3)$$

**Dem.:** Como  $\mathcal{S}$  tem dimensão finita, existe uma base ortonormada  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$  para  $\mathcal{S}$ , onde  $m = \dim \mathcal{S}$ . Dado um vector qualquer  $\mathbf{v} \in \mathcal{V}$ , definamos:

$$\mathbf{s} \stackrel{\text{def}}{=} \sum_{i=1}^m \langle \mathbf{v} | \mathbf{e}_i \rangle \mathbf{e}_i \quad (2.8.4)$$

É claro que  $\mathbf{s} \in \mathcal{S}$ . Por outro lado, como:

$$\langle \mathbf{v} - \mathbf{s} | \mathbf{e}_j \rangle = \langle \mathbf{v} | \mathbf{e}_j \rangle - \langle \mathbf{s} | \mathbf{e}_j \rangle = \langle \mathbf{v} | \mathbf{e}_j \rangle - \langle \mathbf{v} | \mathbf{e}_j \rangle = 0, \quad j = 1, \dots, m$$

o que significa que  $\mathbf{v} - \mathbf{s}$  está em  $\mathcal{S}^\perp$ . Obtemos portanto a decomposição (2.8.2).

Mostremos agora que esta decomposição é única. Isto é equivalente a provar, como já sabemos, que  $\mathcal{S} \cap \mathcal{S}^\perp = \{\mathbf{0}\}$ . Suponhamos então que  $\mathbf{0} \neq \mathbf{u} \in \mathcal{S} \cap \mathcal{S}^\perp$ . Então, por definição de  $\mathcal{S}^\perp$ , e como  $\mathbf{u} \in \mathcal{S}$ ,  $\mathbf{u}$  é ortogonal a todo o vector de  $\mathcal{S}$ . Em particular é ortogonal a si próprio, isto é,  $0 = \langle \mathbf{u} | \mathbf{u} \rangle = \|\mathbf{u}\|^2$ , o que implica que  $\mathbf{u} = \mathbf{0}$ .

Finalmente (2.8.3) deduz-se do Teorema de Pitágoras (ver o teorema 2.19).

► **2.32 Projectores ...** Consideremos de novo um espaço vectorial com um produto interno  $(\mathcal{V}, \langle | \rangle)$  (Euclideano se  $\mathcal{V}$  é real ou Hermitiano se  $\mathcal{V}$  é complexo), e suponhamos que  $\mathcal{S}$  é um subespaço de dimensão finita em  $\mathcal{V}$ . Então, como  $\mathcal{V} = \mathcal{S} \oplus \mathcal{S}^\perp$ , podemos ainda definir uma aplicação linear:

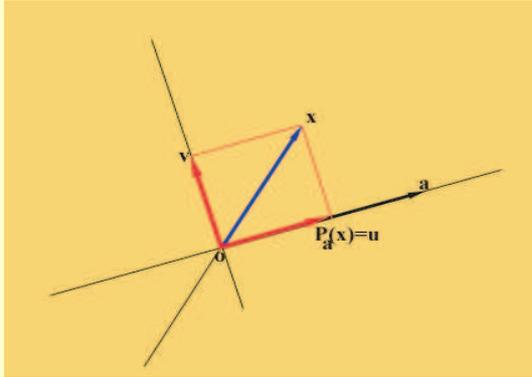
$$\mathbf{P}_{\mathcal{S}} : \mathcal{V} \longrightarrow \mathcal{V} \quad (2.8.5)$$

chamada a **projecção ortogonal sobre  $\mathcal{S}$**  da seguinte forma. Por definição de soma directa, todo o vector  $\mathbf{v} \in \mathcal{V}$  admite uma decomposição única da forma:  $\mathbf{v} = \mathbf{s} + (\mathbf{v} - \mathbf{s})$ , onde  $\mathbf{s} \in \mathcal{S}$  e  $\mathbf{v} - \mathbf{s} \in \mathcal{S}^\perp$ . Pômos então  $\mathbf{P}_{\mathcal{S}}(\mathbf{v}) = \mathbf{s}$ . É fácil ver que  $\mathbf{P}_{\mathcal{S}}$  verifica as propriedades seguintes:

- $\text{im } \mathbf{P}_{\mathcal{S}} = \mathcal{S}$
- $\text{ker } \mathbf{P}_{\mathcal{S}} = \mathcal{S}^\perp$
- $\mathbf{P}_{\mathcal{S}}^2 = \mathbf{P}_{\mathcal{S}}$
- $\|\mathbf{P}_{\mathcal{S}}(\mathbf{v})\| \leq \|\mathbf{v}\|, \quad \forall \mathbf{v} \in \mathcal{V}$
- Se  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$  é uma base ortonormada para  $\mathcal{S}$ , então:

$$\mathbf{P}_{\mathcal{S}}(\mathbf{v}) = \sum_{i=1}^m \langle \mathbf{v} | \mathbf{e}_i \rangle \mathbf{e}_i \quad (2.8.6)$$

► 2.33 Exemplo [Projectção ortogonal sobre uma recta, em  $\mathbb{E}^3$ ] ...



Sejam  $\mathbf{a} \neq \mathbf{0}$  e  $\mathbf{x}$  dois vectores em  $\mathbb{R}^3$ , com  $\mathbf{a}$  não nulo. Então existe um único vector  $\mathbf{u}$ , na recta gerada por  $\mathbf{a}$ , e um único vector  $\mathbf{v}$ , ortogonal a  $\mathbf{a}$ , tais que  $\mathbf{x} = \mathbf{u} + \mathbf{v}$ . O vector  $\mathbf{u}$ , notado por  $\mathbf{P}_a(\mathbf{x})$ , diz-se a **projectção ortogonal** de  $\mathbf{x}$  sobre a recta gerada por  $\mathbf{a}$ , e é dado por:

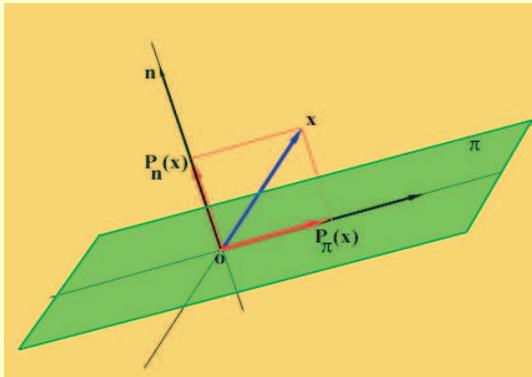
$$\mathbf{P}_a(\mathbf{x}) = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} \quad (2.8.7)$$

A aplicação  $\mathbf{P}_a : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  definida por (4.1.12), é linear. Note que  $\mathbf{P}_a^2 = \mathbf{P}_a$ . Por outro lado, se considerarmos um qualquer vector  $\mathbf{b} \neq \mathbf{0}$  ortogonal a  $\mathbf{a}$  (i.e.:  $\mathbf{a} \cdot \mathbf{b} = 0$ ), vemos que  $\mathbf{P}_a(\mathbf{b}) = \mathbf{0}$  e portanto:

$$\ker \mathbf{P}_a = \text{span}\{\mathbf{b}\} = \{\mathbf{b} \in \mathbb{R}^3 : \mathbf{b} \cdot \mathbf{a} = 0\} = \mathbf{a}^\perp$$

é o plano vectorial ortogonal a  $\mathbf{a}$ .

► 2.34 Exemplo [Projectção ortogonal sobre um plano vectorial, em  $\mathbb{E}^3$ ] ...



Consideremos um plano vectorial ortogonal a um vector  $\mathbf{n} \in \mathbb{R}^3 - \{\mathbf{0}\}$  (se esse plano é gerado por dois vectores  $\mathbf{u}, \mathbf{v}$  linearmente independentes, podemos tomar  $\mathbf{n} = \mathbf{u} \times \mathbf{v}$ ). Notemos esse plano por  $\pi = \mathbf{n}^\perp$ . Dado um vector  $\mathbf{x} \in \mathbb{R}^3$ , ao vector:

$$\mathbf{P}_{\mathbf{n}^\perp} \equiv \mathbf{x} - \mathbf{P}_n(\mathbf{x})$$

chamamos a **projectção ortogonal** de  $\mathbf{x}$  sobre o plano vectorial ortogonal a  $\mathbf{n}$ .

De acordo com (4.1.12), temos que:

$$\begin{aligned} \mathbf{P}_{\mathbf{n}^\perp} &\equiv \mathbf{x} - \mathbf{P}_n(\mathbf{x}) \\ &= \mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n} \end{aligned} \quad (2.8.8)$$

A aplicação  $\mathbf{P}_{\mathbf{n}^\perp} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  definida por (4.1.13), é linear. Note que  $\mathbf{P}_{\mathbf{n}^\perp}^2 = \mathbf{P}_{\mathbf{n}^\perp}$ . Se  $\mathbf{x} \cdot \mathbf{n} = 0$ , i.e., se  $\mathbf{x}$  é ortogonal a  $\mathbf{n}$ , então  $\mathbf{P}_{\mathbf{n}^\perp}(\mathbf{x}) = \mathbf{x}$ , enquanto que, por outro lado,  $\mathbf{P}_{\mathbf{n}^\perp}(\mathbf{n}) = \mathbf{0}$ . Portanto vemos que:

$$\ker \mathbf{P}_{\mathbf{n}^\perp} = \text{span}\{\mathbf{n}\}$$

e:

$$\mathbf{P}_{\mathbf{n}^\perp}(\mathbf{x}) = \mathbf{x} \quad \forall \mathbf{x} \in \mathbf{n}^\perp$$

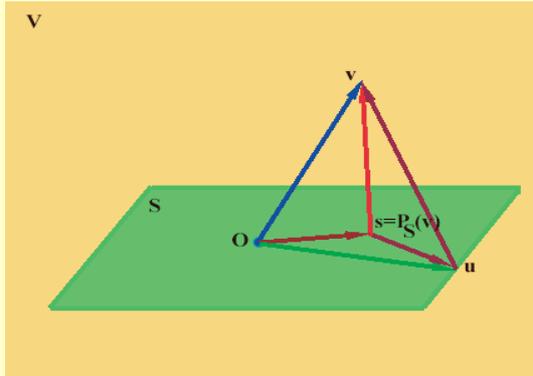
► **2.35 Teorema [da aproximação óptima]** ... Consideremos um espaço vectorial com um produto interno  $(\mathcal{V}, \langle \cdot | \cdot \rangle)$  (Euclidiano se  $\mathcal{V}$  é real ou Hermitiano se  $\mathcal{V}$  é complexo), e seja  $\mathcal{S}$  um subespaço de dimensão finita. Dado um vector  $\mathbf{v} \in \mathcal{V}$ , a projecção ortogonal de  $\mathbf{v}$  sobre  $\mathcal{S}$ :

$$\mathbf{s} = \mathbf{P}_{\mathcal{S}}(\mathbf{v}) \in \mathcal{S}$$

é o vector de  $\mathcal{S}$  que está mais perto de  $\mathbf{v}$ , isto é:

$$\|\mathbf{v} - \mathbf{P}_{\mathcal{S}}(\mathbf{v})\| \leq \|\mathbf{v} - \mathbf{u}\|, \quad \forall \mathbf{u} \in \mathcal{S} \quad (2.8.9)$$

e  $\|\mathbf{v} - \mathbf{P}_{\mathcal{S}}(\mathbf{v})\| = \|\mathbf{v} - \mathbf{u}\|$ , com  $\mathbf{u} \in \mathcal{S}$  se e só se  $\mathbf{u} = \mathbf{P}_{\mathcal{S}}(\mathbf{v})$ .



**Dem.:** Por (2.8.2), temos que  $\mathbf{v} = \mathbf{s} + (\mathbf{v} - \mathbf{s})$ , onde  $\mathbf{s} = \mathbf{P}_{\mathcal{S}}(\mathbf{v}) \in \mathcal{S}$  e  $\mathbf{v} - \mathbf{s} \in \mathcal{S}^{\perp}$ . Como  $\forall \mathbf{u} \in \mathcal{S}$  se tem:

$$\mathbf{v} - \mathbf{u} = \underbrace{(\mathbf{s} - \mathbf{u})}_{\in \mathcal{S}} + \underbrace{(\mathbf{v} - \mathbf{s})}_{\in \mathcal{S}^{\perp}}$$

esta é a decomposição ortogonal de  $\mathbf{v} - \mathbf{u}$ . Pelo teorema de Pitágoras:

$$\|\mathbf{v} - \mathbf{u}\|^2 = \|\mathbf{s} - \mathbf{u}\|^2 + \|\mathbf{v} - \mathbf{s}\|^2 \geq \|\mathbf{v} - \mathbf{s}\|^2$$

sendo a igualdade válida sse  $\|\mathbf{s} - \mathbf{u}\|^2 = 0$ , isto é, sse  $\mathbf{s} = \mathbf{u}$ .

► **2.36 Exemplo (Aproximação de funções contínuas em  $[0, 2\pi]$  por polinómios trigonométricos)** ... Seja  $\mathcal{V} = C^o([0, 2\pi]; \mathbb{R})$  o espaço das funções reais contínuas definidas em  $[0, 2\pi]$ , munido do produto  $L^2$ :

$$\langle f | g \rangle = \int_0^{2\pi} f(t)g(t) dt$$

e  $\mathcal{S}_n$  o subespaço de dimensão  $2n + 1$  seguinte:

$$\mathcal{S}_n = \text{span}_{\mathbb{R}} \left\{ \varphi_0(t) = \frac{1}{\sqrt{2}}, \varphi_{2k-1}(t) = \frac{\cos kt}{\sqrt{\pi}}, \varphi_{2k}(t) = \frac{\sin kt}{\sqrt{\pi}} : k = 1, \dots, n \right\} \quad (2.8.10)$$

As  $2n + 1$  funções  $\{\varphi_0, \varphi_1, \dots, \varphi_{2n-1}, \varphi_{2n}\}$ , chamadas **polinómios trigonométricos**, formam uma base ortonormada para  $\mathcal{S}$  (mostrar isto<sup>2</sup>).

Se  $f \in C^o([0, 2\pi]; \mathbb{R})$ , representemos por  $\mathcal{F}_n(f)$  a projecção ortogonal de  $f$  sobre  $\mathcal{S}_n$ . De acordo com a fórmula da projecção ortogonal (2.8.6), temos que:

$$\mathcal{F}_n(f) = \sum_{k=0}^{2n} \langle f | \varphi_k \rangle \varphi_k \quad (2.8.11)$$

<sup>2</sup>Usar as relações trigonométricas seguintes:

$$\begin{aligned} \cos A \cos B &= \frac{1}{2} \{ \cos(A - B) + \cos(A + B) \} \\ \sin A \sin B &= \frac{1}{2} \{ \cos(A - B) - \cos(A + B) \} \\ \sin A \cos B &= \frac{1}{2} \{ \sin(A - B) + \sin(A + B) \} \end{aligned}$$

onde:

$$\langle f|\varphi_k \rangle = \int_0^{2\pi} f(t)\varphi_k(t) dt \quad (2.8.12)$$

são os chamados **coeficientes de Fourier** de  $f$ . Usando a definição das funções  $\varphi_k$ , podemos escrever as fórmulas anteriores na forma:

$$\mathcal{F}_n(f) = \frac{1}{2}a_0 + \sum_{k=1}^n (a_k \cos kt + b_k \sin kt) \quad (2.8.13)$$

onde os coeficientes de Fourier são dados por:

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt dt \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt dt \end{aligned} \quad (2.8.14)$$

para  $k = 0, 1, 2, \dots, n$ . O teorema da aproximação óptima diz-nos que o polinómio trigonométrico  $\mathcal{F}_n(f) \in \mathcal{S}_n$ , dado por (2.8.13), aproxima  $f$  melhor que qualquer outro polinómio trigonométrico em  $\mathcal{S}_n$ , no sentido em que  $\|f - \mathcal{F}_n(f)\|$  é o menor possível.

► **2.37 Exemplo (Aproximação de funções contínuas em  $[-1, 1]$  por polinómios de grau  $\leq n$ )** ... Seja  $\mathcal{V} = C^0([-1, 1]; \mathbb{R})$  o espaço das funções reais contínuas definidas em  $[-1, 1]$ , munido do produto  $L^2$ :

$$\langle f|g \rangle = \int_{-1}^1 f(t)g(t) dt$$

e  $\mathcal{S}_n$  o subespaço de dimensão  $n + 1$  gerado pelos polinómios de Legendre normalizados, introduzidos no exemplo 2.30:

$$\mathcal{S}_n = \text{span}_{\mathbb{R}} \{\varphi_0, \varphi_1, \dots, \varphi_n\} \quad (2.8.15)$$

É claro que  $\mathcal{S}$  é o subespaço constituído por todas as funções polinomiais de grau  $\leq n$ , definidas no intervalo  $[-1, 1]$ .  $f \in C^0([-1, 1]; \mathbb{R})$ , representemos por  $\mathbf{P}_n(f)$  a projecção ortogonal de  $f$  sobre  $\mathcal{S}_n$ . De acordo com a fórmula da projecção ortogonal (2.8.6), temos que:

$$\mathbf{P}_n(f) = \sum_{k=0}^n \langle f|\varphi_k \rangle \varphi_k, \quad \text{onde } \langle f|\varphi_k \rangle = \int_{-1}^1 f(t)\varphi_k(t) dt \quad (2.8.16)$$

que é o polinómio de grau  $\leq n$ , para o qual  $\|f - \mathbf{P}_n(f)\|$  é o menor possível. Por exemplo, se  $f(t) = \sin \pi t$ , os coeficientes  $\langle f|\varphi_k \rangle$  são dados por:

$$\langle f|\varphi_k \rangle = \int_{-1}^1 \sin \pi t \varphi_k(t) dt$$

Em particular,  $\langle f|\varphi_0 \rangle = 0$  E.

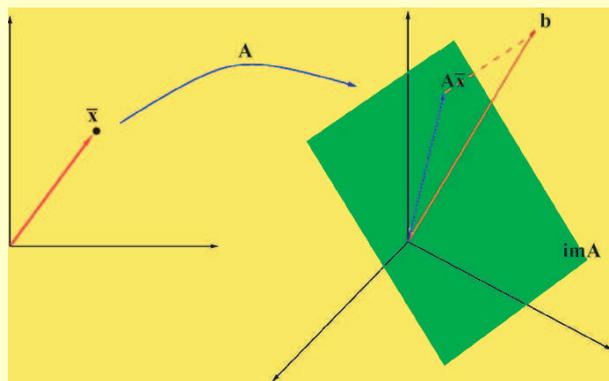
$$\langle f|\varphi_1 \rangle = \int_{-1}^1 \sqrt{\frac{3}{2}} t \sin \pi t dt = \sqrt{\frac{3}{2}} \frac{2}{\pi}$$

## 2.9 Aplicações. Mínimos quadrados

► **2.38 Solução dos mínimos quadrados ...** Seja:

$$\mathbf{Ax} = \mathbf{b} \quad (2.9.1)$$

um sistema de equações lineares, não homogéneo, escrito em forma matricial.  $\mathbf{A}$  é uma matriz  $m \times n$ ,  $\mathbf{x} \in \mathbb{R}^n$  e  $\mathbf{b} \in \mathbb{R}^m$  é um vector fixo.



Uma “solução” dos mínimos quadrados do sistema (2.9.1) é, por definição, um vector  $\hat{\mathbf{x}}$ , que satisfaz:

$$\|\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}\| \text{ é mínimo} \quad (2.9.2)$$

Interpretando  $\mathbf{A}$  como a matriz de uma aplicação linear  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , relativamente às bases canónicas de cada um destes espaços, vemos que o significado de uma “solução” dos mínimos quadrados é o seguinte: é um vector  $\hat{\mathbf{x}} \in \mathbb{R}^n$  cuja imagem está mais próxima de  $\mathbf{b}$ .

► **2.39** Quando  $\ker A = \{\mathbf{0}\}$  a “solução”  $\hat{\mathbf{x}}$  é única. Quando  $\mathbf{b} \in \text{im } A$ ,  $\hat{\mathbf{x}}$  é uma solução exacta do sistema. Quando  $\mathbf{b} \notin \text{im } A$ , e  $\ker A = \{\mathbf{0}\}$  a “solução”  $\hat{\mathbf{x}}$  é dada por:

$$\hat{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{P}_{\text{im } A}(\mathbf{b}) \quad (2.9.3)$$

Isto é, para calcular a “solução” dos mínimos quadrados do sistema (2.9.1) procede-se da seguinte forma:

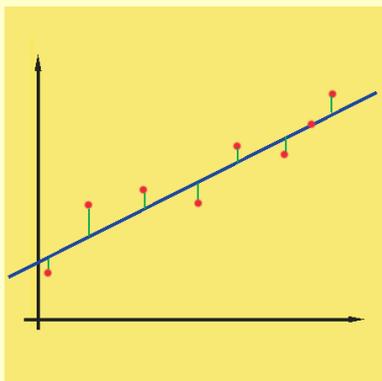
1. Calcula-se a projecção ortogonal  $\hat{\mathbf{y}} \in \text{im } A$ , de  $\mathbf{b}$  sobre a imagem de  $A$ . Pelo teorema da aproximação óptima, este será o vector da imagem de  $A$ , que melhor aproxima  $\mathbf{b}$ .
2. Calcula-se  $\hat{\mathbf{x}}$  tal que  $\mathbf{A}\hat{\mathbf{x}} = \hat{\mathbf{y}}$

► **2.40 Exemplo ...** Calcular a “solução” dos mínimos quadrados do sistema:

$$\begin{cases} x + 2y & = & 1 \\ 3x - y + z & = & 0 \\ -x + 2y + z & = & -1 \\ x - y - 2z & = & 2 \\ 2x + y - z & = & 2 \end{cases} \quad (2.9.4)$$

e o erro correspondente.

► **2.41 Aproximação de dados por uma recta pelo método dos mínimos quadrados ...**



Suponhamos que se fazem  $n$  medições de uma certa grandeza  $y$ , em  $n$  instantes  $t_i$ ,  $i = 1, \dots, n$ , obtendo os resultados:

$$\begin{array}{|c|c|c|c|c|} \hline t_1 & t_2 & t_3 & \cdots & t_n \\ \hline y_1 & y_2 & y_3 & \cdots & y_n \\ \hline \end{array} \quad (2.9.5)$$

Representemos os  $n$  pontos  $(t_i, y_i)$  no plano em  $\mathbb{R}_{t,y}^2$ , e suponhamos que se pretende calcular uma recta do tipo:

$$y = \alpha t + \beta \quad (2.9.6)$$

que melhor ajuste esses dados. Em que sentido deve ser entendido este “melhor” ajustamento?

Para cada  $t_i$ , o erro  $e_i$  entre o valor medido  $y_i$  e o valor estimado a partir da recta referida (supondo que ela está já calculada) é igual a:

$$e_i = y_i - (\alpha t_i + \beta), \quad i = 1, 2, \dots, n$$

Em forma matricial:

$$\mathbf{e} = \mathbf{y} - \mathbf{A}\mathbf{x} \quad (2.9.7)$$

onde:

$$\mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} t_1 & 1 \\ t_2 & 2 \\ \vdots & \vdots \\ t_n & n \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$\mathbf{e}$  é o chamado **vector de erro** e  $\mathbf{y}$  o **vector dos dados**. Os coeficientes  $\alpha, \beta$  - as incógnitas do problema - são as componentes do vector  $\mathbf{x}$ .

Se os dados se ajustassem exactamente,  $y_i = \alpha t_i + \beta$ , os erros seriam todos nulos  $e_i = 0$ , e poderíamos resolver o sistema  $\mathbf{A}\mathbf{x} = \mathbf{y}$ . Por outras palavras, os dados estarão todos numa linha recta sse  $\mathbf{y} \in \text{im } \mathbf{A}$ . Se eles não forem colineares então devemos procurar a recta para a qual o erro total:

$$\|\mathbf{e}\| = (e_1^2 + \cdots + e_n^2)^{1/2}$$

seja mínimo.

Em linguagem vectorial, procuramos pois o vector  $\mathbf{x} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$  que minimiza a norma Euclídeana do vector erro:

$$\|\mathbf{e}\| = \|\mathbf{A}\mathbf{x} - \mathbf{y}\|$$

que é exactamente a situação que caracteriza a procura da solução dos mínimos quadrados para o sistema  $\mathbf{A}\mathbf{x} = \mathbf{y}$ , que foi explicada no ponto anterior.

► **2.42 Exemplo** ... Calcular a recta de aproximação dos mínimos quadrados para os dados seguintes:

$$\begin{array}{|c|c|c|c|c|} \hline t_i & 0 & 1 & 3 & 6 \\ \hline y_i & 2 & 3 & 7 & 12 \\ \hline \end{array} \quad (2.9.8)$$

**Solução:**  $y = 12/7(1 + t)$ .

► **2.43 Exemplo ...** Considere a aplicação linear  $\mathbf{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  definida por:

$$\mathbf{A}(x, y) = (x + y, x - y, x)$$

a.) Calcule o ortogonal da imagem de  $\mathbf{A}$  em  $\mathbb{R}^3$ , com a estrutura Euclideana usual.

b.) Calcule a “solução” dos mínimos quadrados do sistema:

$$\begin{cases} x + y = 1 \\ x - y = 1 \\ x = 0 \end{cases}$$

Calcule o erro associado a essa solução e explique qual o seu significado geométrico (da solução e do seu erro).

**Resolução ...**

a.) A imagem de  $\mathbf{A}$  é constituída por todos os vectores  $(X, Y, Z) \in \mathbb{R}^3$  tais que:

$$(X, Y, Z) = \mathbf{A}(x, y) = (x + y, x - y, x)$$

para algum vector  $(x, y) \in \mathbb{R}^2$ . A questão é pois: quais os vectores  $(X, Y, Z) \in \mathbb{R}^3$  para os quais existe  $(x, y)$  tal que:

$$\begin{cases} x + y = X \\ x - y = Y \\ x = Z \end{cases} ?$$

Resolvendo o sistema em ordem a  $x, y$  (com  $X, Y, Z$  como parâmetros), vem que:

$$\begin{cases} x = Z \\ y = X - Z \\ 0 = X + Y - 2Z \end{cases}$$

Portanto a imagem de  $\mathbf{A}$  é o plano  $X + Y - 2Z = 0$  em  $\mathbb{R}^3$ . O seu ortogonal é a recta gerada pelo vector  $\mathbf{n} = (1, 1, -2)$ .

b.) Por definição (e pelo teorema da aproximação óptima), a “solução” dos mínimos quadrados é a solução do sistema:

$$\mathbf{A}\mathbf{x} = \mathbf{P}_{\text{im } \mathbf{A}}(\mathbf{b})$$

onde  $\mathbf{P}_{\text{im } \mathbf{A}}(\mathbf{b})$  é a projecção ortogonal do vector  $\mathbf{b} = (1, 1, 0)$  sobre o plano imagem de  $\mathbf{A}$ :  $X + Y - 2Z = 0$ .

Essa projecção pode ser calculada pela seguinte fórmula:

$$\mathbf{P}_{\text{im } \mathbf{A}}(1, 1, 0) = (1, 1, 0) - \frac{(1, 1, 0) \cdot (1, 1, -2)}{\|(1, 1, -2)\|^2} (1, 1, -2) = \frac{2}{3}(1, 1, 1)$$

Logo a solução procurada é a solução do sistema:

$$\begin{cases} x + y = 2/3 \\ x - y = 2/3 \\ x = 2/3 \end{cases}$$

que é:

$$x = 2/3, \quad y = 0$$

O erro associado é, por definição, igual à distância entre o ponto  $(1, 1, 0)$  e a  $\mathbf{P}_{\text{im } \mathbf{A}}(\mathbf{b})$ :

$$e = \|(1, 1, 0) - \frac{2}{3}(1, 1, 1)\| = \sqrt{6}/3$$

► **2.44 Exemplo** ... Considere o espaço vectorial  $\mathbb{R}_3[t]$  das funções polinomiais  $p(t)$ , de grau  $\leq 3$ , de coeficientes reais, munido do produto interno:

$$\langle p(t)|q(t) \rangle = \int_0^{+1} p(t)q(t) dt$$

a.) Mostre que:

$$S = \{p(t) \in \mathbb{R}_3[t] : p(t) = p(-t)\}$$

é um subespaço vectorial. Calcule  $\dim S$  e determine uma base ortonormada para  $S$ .

b.) Calcule o polinómio de  $S$  que está mais próximo do polinómio  $p(t) = t$ .

c.) Calcule o ortogonal de  $\mathcal{T} = \text{span}\{1\}$  em  $\mathbb{R}_3[t]$ .

d.) Calcule o núcleo e a imagem da aplicação linear:

$$\begin{aligned} \mathbf{T} : \mathbb{R}_3[t] &\longrightarrow \mathbb{R}_3[t] \\ p(t) &\longmapsto \mathbf{T}[p(t)] = p''(t) - 2tp'(t) \end{aligned}$$

**Resolução** ...

a.) Se  $p, q \in S$  então  $(p + q)(t) = p(t) + q(t) = p(-t) + q(-t) = (p + q)(-t)$  e portanto  $p + q \in S$ . Se  $p \in S$  e  $\lambda \in \mathbb{R}$  então  $(\lambda p)(t) = \lambda p(t) = \lambda p(-t) = \lambda p(-t)$  e portanto  $\lambda p \in S$ .

Se  $p(t) = a + bt + ct^2 + dt^3 \in S$  então  $a + bt + ct^2 + dt^3 = p(t) = p(-t) = a - bt + ct^2 - dt^3$ , isto é,  $2bt + 2dt^3 = 0$  e portanto  $b = d = 0$ . Logo:

$$\begin{aligned} S &= \{p(t) = a + bt + ct^2 + dt^3 \in \mathbb{R}_3[t] : b = d = 0\} \\ &= \{p(t) = a + ct^2 \in \mathbb{R}_3[t] : a, c \in \mathbb{R}\} \\ &= \text{span}\{1, t^2\} \end{aligned}$$

e  $\dim S = 2$ . Os polinómios  $p(t) \equiv 1$  e  $q(t) = t^2$  constituem uma base para  $S$ .

Uma base ortonormada obtém-se pelo processo de Gram-Schmidt.  $\|1\|^2 = \int_0^1 1 dt = 1$  e  $t^2 - \frac{\langle t^2|1 \rangle}{\|1\|^2} 1 = t^2 - \int_0^1 t^2 dt = t^2 - 1/3$ . Além disso  $\|t^2 - 1/3\|^2 = \int_0^1 (t^2 - 1/3)^2 dt = 4/45$ . Logo os polinómios  $1$  e  $(3\sqrt{5}/2)(t^2 - 1/3)$  constituem uma base ortonormada para  $S$ .

b.) Pelo teorema da aproximação óptima esse polinómio é dado pela projecção ortogonal de  $t$  sobre  $S$ :

$$\begin{aligned} \mathbf{P}_S(t) &= \langle t|1 \rangle 1 + \langle t|(3\sqrt{5}/2)(t^2 - 1/3) \rangle (3\sqrt{5}/2)(t^2 - 1/3) \\ &= \int_0^1 t dt + (45/4) \left( \int_0^1 t(t^2 - 1/3) dt \right) (t^2 - 1/3) \\ &= 1/2 + (45/48)(t^2 - 1/3) \end{aligned}$$

c.) Um polinómio  $p(t) = a + bt + ct^2 + dt^3 \in \mathbb{R}_3[t]$  estará em  $\mathcal{T}^\perp$  sse  $\langle (a + bt + ct^2 + dt^3) | 1 \rangle = 0$  isto é, sse  $a + b/2 + c/3 + d/4 = 0$ . Portanto:

$$\mathcal{T}^\perp = \{p(t) = a + bt + ct^2 + dt^3 \in \mathbb{R}_3[t] : a + b/2 + c/3 + d/4 = 0\}$$

que é um hiperplano em  $\mathbb{R}_3[t]$ .

d.) Um polinómio  $p(t) = a + bt + ct^2 + dt^3 \in \mathbb{R}_3[t]$  estará em  $\ker \mathbf{T}$  sse:

$$\begin{aligned} 0 &= \mathbf{T}[p(t)] = p''(t) - 2tp'(t) \\ &= (a + bt + ct^2 + dt^3)'' - 2t(a + bt + ct^2 + dt^3)' \\ &= (2c + 6dt) - 2t(b + 2ct + 3dt^2) \\ &= 2c + (6d - 2b)t - 4ct^2 - 6dt^3 \end{aligned}$$

donde  $2c = 0, 6d - 2b = 0, 4c = 0, 6d = 0$ , isto é,  $b = c = d = 0$ . Portanto o  $\ker \mathbf{T}$  é constituído pelos polinómios  $p(t) = a + bt + ct^2 + dt^3 \in \mathbb{R}_3[t]$  tais que  $b = c = d = 0$ , isto é,  $\ker \mathbf{T} = \{a : a \in \mathbb{R}\} = \text{span}\{1\}$ .

$\text{im } \mathbf{T}$  é constituída pelos polinómios  $P(t) = A + Bt + Ct^2 + Dt^3 \in \mathbb{R}_3[t]$  tais que:

$$\mathbf{T}(a + bt + ct^2 + dt^3) = A + Bt + Ct^2 + Dt^3$$

para algum polinómio  $p(t) = a + bt + ct^2 + dt^3 \in \mathbb{R}_3[t]$ . Como  $\mathbf{T}[p(t)] = 2c + (6d - 2b)t - 4ct^2 - 6dt^3$ , vem que:

$$2c + (6d - 2b)t - 4ct^2 - 6dt^3 = A + Bt + Ct^2 + Dt^3$$

isto é:

$$\left\{ \begin{array}{l} 2c = A \\ -2b + 6d = B \\ -4c = C \\ -6d = D \end{array} \right. \Rightarrow \left\{ \begin{array}{l} -2b + 6d = B \\ 2c = A \\ -6d = D \\ 0 = 2A + C \end{array} \right. \Rightarrow$$

e portanto  $\text{im } \mathbf{T} = \{P(t) = A + Bt + Ct^2 + Dt^3 \in \mathbb{R}_3[t] : 2A + C = 0\}$ .

## 2.10 Exercícios

▷ **Exercício 2.1** ... Verifique quais das seguintes funções são produtos internos Euclidianos em  $\mathbb{R}^2$  ou  $\mathbb{R}^3$ :

- $\langle \mathbf{u}, \mathbf{v} \rangle = x^1 y^1 - x^1 y^2 - x^2 y^1 + 3x^2 y^2$ , sabendo que  $\mathbf{u} = (x^1, x^2)$ , e  $\mathbf{v} = (y^1, y^2)$ .
- $\langle \mathbf{u}, \mathbf{v} \rangle = x^1 y^1 + x^1 y^2 - 2x^2 y^1 + 3x^2 y^2$ , sabendo que  $\mathbf{u} = (x^1, x^2)$ , e  $\mathbf{v} = (y^1, y^2)$ .
- $\langle \mathbf{u}, \mathbf{v} \rangle = 6x^1 y^1 + 2x^2 y^2$ , sabendo que  $\mathbf{u} = (x^1, x^2)$ , e  $\mathbf{v} = (y^1, y^2)$ .
- $\langle \mathbf{u}, \mathbf{v} \rangle = x^1 y^1 + 3x^2 y^2 + 4x^3 y^3$ , sabendo que  $\mathbf{u} = (x^1, x^2, x^3)$ , e  $\mathbf{v} = (y^1, y^2, y^3)$ .
- $\langle \mathbf{u}, \mathbf{v} \rangle = x^1 y^1 + 3x^2 y^2 + 4x^3 y^3 - x^1 y^2 - y^1 x^2$ , sabendo que  $\mathbf{u} = (x^1, x^2, x^3)$ , e  $\mathbf{v} = (y^1, y^2, y^3)$ .

▷ **Exercício 2.2** ... Calcule em cada caso  $\langle \mathbf{u}, \mathbf{v} \rangle$  usando o produto interno Euclidiano usual e o produto interno definido em 2.1-a). Depois, calcule  $\|\mathbf{u}\|$  e  $\|\mathbf{v}\|$  recorrendo também a cada um desses dois produtos internos.

- $\mathbf{u} = (1, 1)$ ,  $\mathbf{v} = (-1, 1)$ ;
- $\mathbf{u} = (1, 0)$ ,  $\mathbf{v} = (1, 2)$ ;
- $\mathbf{u} = (2, 1)$ ,  $\mathbf{v} = (4, -1)$ ;

▷ **Exercício 2.3** ... Calcule em cada caso  $\langle \mathbf{u}, \mathbf{v} \rangle$  usando o produto interno euclidiano usual e o produto interno definido em 2.1-d). Depois, calcule  $\|\mathbf{u}\|$  e  $\|\mathbf{v}\|$  recorrendo também a cada um destes dois produtos internos.

a)  $\mathbf{u} = (1, 1, 1), \mathbf{v} = (-1, 1, 2);$

b)  $\mathbf{u} = (1, 0, -1), \mathbf{v} = (3, -1, 2);$

c)  $\mathbf{u} = (0, 0, 1), \mathbf{v} = (-1, 4, 6);$

▷ **Exercício 2.4** ... Determine todos os valores reais de  $k$  para os quais  $\langle \mathbf{u}, \mathbf{v} \rangle$  é um produto interno Euclidiano em  $\mathbb{R}^2$  :

$$\langle \mathbf{u}, \mathbf{v} \rangle = x^1 y^1 - 3x^1 y^2 - 3x^2 y^1 + kx^2 y^2$$

▷ **Exercício 2.5** ... Determine todos os valores reais de  $a, b, c, d$  para os quais  $\langle \mathbf{u}, \mathbf{v} \rangle$  é um produto interno Euclidiano em  $\mathbb{R}^2$  :

$$\langle \mathbf{u}, \mathbf{v} \rangle = ax^1 y^1 + bx^1 y^2 + cx^2 y^1 + dx^2 y^2$$

▷ **Exercício 2.6** ... Sejam,  $\mathbf{u} = (z^1, z^2)$  e  $\mathbf{v} = (w^1, w^2)$  elementos de  $\mathbb{C}^2$ . Verifique que a função que se segue é um produto interno Hermitiano em  $\mathbb{C}^2$  :

$$f(\mathbf{u}, \mathbf{v}) = z^1 \overline{w^1} + (1+i)z^1 \overline{w^2} + (1-i)z^2 \overline{w^1} + 3z^2 \overline{w^2}$$

Calcule a norma de  $\mathbf{v} = (1-2i, 2+3i)$  usando o produto interno Hermitiano usual e depois o produto interno definido neste exercício.

▷ **Exercício 2.7** ... Em cada caso, determine o cos do ângulo  $\theta$  entre os vectores  $\mathbf{u}$  e  $\mathbf{v}$  :

a)  $\mathbf{u} = (1, -3, 2), \mathbf{v} = (2, 1, 5)$  em  $\mathbb{R}^3$ , usando o produto interno euclidiano usual e o produto interno definido em 2.1-d).

b)  $\mathbf{u} = 2t - 1, \mathbf{v} = t^2$  em  $\mathbb{R}[t]$ , usando o produto interno Euclidiano definido no exercício 2.14.

▷ **Exercício 2.8** ... No espaço linear  $\mathbb{R}[t]$  verifique se  $\langle f, g \rangle$  é um produto interno.

a)  $\langle f, g \rangle = f(1)g(1)$

b)  $\langle f, g \rangle = \left| \int_0^1 f(t)g(t) dt \right|$

c)  $\langle f, g \rangle = \int_0^1 f'(t)g'(t) dt$

d)  $\langle f, g \rangle = \left( \int_0^1 f(t) dt \right) \left( \int_0^1 g(t) dt \right)$

▷ **Exercício 2.9** ... No espaço vectorial real das funções contínuas em  $[-1, 1]$ , seja  $\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt$ . Considere as três funções  $u_1, u_2, u_3$  dadas por:

$$u_1(t) = 1, \quad u_2(t) = t, \quad u_3(t) = 1 + t.$$

Mostre que duas delas são ortogonais, duas fazem um ângulo de  $\frac{\pi}{3}$  entre si e as outras duas fazem um ângulo de  $\frac{\pi}{6}$  entre si.

▷ **Exercício 2.10** ... Prove cada uma das afirmações das alíneas seguintes e interprete-as geometricamente no caso do produto interno usual em  $\mathbb{R}^2$  ou  $\mathbb{R}^3$ .

- a)  $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \Leftrightarrow \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ .  
 b)  $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \Leftrightarrow \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2$ .  
 c)  $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \Leftrightarrow \|\mathbf{x} + c\mathbf{y}\| \geq \|\mathbf{x}\|$  para todo o real  $c$ .  
 d)  $\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = 0 \Leftrightarrow \|\mathbf{x}\| = \|\mathbf{y}\|$ .

▷ **Exercício 2.11** ... Calcule o ângulo que o vector  $(1, 1, \dots, 1)$  de  $\mathbb{R}^n$  faz com os vectores coordenados unitários de  $\mathbb{R}^n$ .

▷ **Exercício 2.12** ... Como se sabe, num espaço Euclidiano real com produto interno  $\langle \mathbf{x}, \mathbf{y} \rangle$  fica definida uma norma por  $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{\frac{1}{2}}$ . Dê uma fórmula para obter o produto interno  $\langle \mathbf{x}, \mathbf{y} \rangle$  a partir de normas de vectores apropriados.

▷ **Exercício 2.13** ... Seja  $V$  um espaço linear real normado e designe-se a norma de  $\mathbf{x} \in V$  por  $\|\mathbf{x}\|$ . Prove que se a norma se pode obter de um produto interno na forma  $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{y} \rangle^{\frac{1}{2}}$  então:

$$\|\mathbf{x} - \mathbf{y}\|^2 + \|\mathbf{x} + \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$$

Esta identidade é conhecida por **lei do paralelogramo**. Verifique que corresponde a afirmar que para um paralelogramo a soma dos quadrados dos comprimentos dos lados é igual à soma dos quadrados dos comprimentos das diagonais.

▷ **Exercício 2.14** ... Considere o espaço vectorial real  $\mathbb{R}[t]$  no qual está definido o seguinte produto interno:  $\langle f, g \rangle = \int_0^1 f(t)g(t) dt$ . Seja  $f(t) = t + 2$  e  $g(t) = t^2 - 2t - 3$ . Determine :

- a)  $\langle f, g \rangle$     b)  $\|f\|$     c) Um vector unitário com a direcção de  $g$ .

▷ **Exercício 2.15** ... Seja  $E$  um espaço vectorial no qual está definido um produto escalar. Mostre que :

$$a) \|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2 \quad b) \langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4} \|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4} \|\mathbf{u} - \mathbf{v}\|^2$$

▷ **Exercício 2.16** ... Em cada um dos casos, determine uma base ortonormada do subespaço de  $\mathbb{R}^3$  gerado pelos seguintes vectores:

- a)  $\mathbf{x}_1 = (1, 1, 1)$ ,     $\mathbf{x}_2 = (1, 0, 1)$ ,     $\mathbf{x}_3 = (3, 2, 3)$ .  
 b)  $\mathbf{x}_1 = (1, 1, 1)$ ,     $\mathbf{x}_2 = (-1, 1, -1)$ ,     $\mathbf{x}_3 = (1, 0, 1)$ .

▷ **Exercício 2.17** ... Em cada um dos casos, determine uma base ortonormada do subespaço de  $\mathbb{R}^4$  gerado pelos seguintes vectores:

- a)  $\mathbf{x}_1 = (1, 1, 0, 0)$ ,     $\mathbf{x}_2 = (0, 1, 1, 0)$ ,     $\mathbf{x}_3 = (0, 0, 1, 1)$ ,     $\mathbf{x}_4 = (1, 0, 0, 1)$ .  
 b)  $\mathbf{x}_1 = (1, 1, 0, 1)$ ,     $\mathbf{x}_2 = (1, 0, 2, 1)$ ,     $\mathbf{x}_3 = (1, 2, -2, 1)$ .

▷ **Exercício 2.18** ... No espaço vectorial real  $\mathbb{R}[t]$ , com o produto interno  $\langle x, y \rangle = \int_0^1 x(t)y(t) dt$ , mostre que as funções que se seguem formam uma base ortonormada do subespaço por elas gerado:

$$y_1(t) = 1, \quad y_2(t) = \sqrt{3}(2t - 1), \quad y_3(t) = \sqrt{5}(6t^2 - 6t + 1).$$

▷ **Exercício 2.19** ... Seja  $S$  um subespaço de um espaço vectorial  $V$ . Mostre que o  $S^\perp$  é o conjunto dos vectores ortogonais a todos os vectores de uma base de  $S$ .

▷ **Exercício 2.20** ... Seja  $W$  o subespaço de  $\mathbb{R}^5$  gerado pelos vectores  $\mathbf{u} = (1, 2, 3, -1, 2)$  e  $\mathbf{v} = (2, 4, 7, 2, -1)$ . Determine uma base do complemento ortogonal  $W^\perp$  de  $W$ .

▷ **Exercício 2.21** ... Determine uma base do subespaço  $W$  de  $\mathbb{R}^4$  ortogonal a  $\mathbf{u}_1 = (1, -2, 3, 4)$  e  $\mathbf{u}_2 = (3, -5, 7, 8)$ .

▷ **Exercício 2.22** ... Considere o espaço vectorial real  $\mathbb{R}_2[t]$  no qual está definido o produto interno  $\langle f, g \rangle = \int_0^1 f(t)g(t) dt$ .

a) Determine uma base do subespaço  $W$  ortogonal a  $h(t) = 2t + 1$ .

b) Aplique o método de ortogonalização de Gram-Schmidt à base  $(1, t, t^2)$  para obter uma base ortonormada  $(u_1(t), u_2(t), u_3(t))$  de  $\mathbb{R}_2[X]$ .

▷ **Exercício 2.23** ... Seja  $V$  o espaço linear das matrizes  $2 \times 2$  de componentes reais, com as operações usuais. Prove que fica definido um produto interno em  $V$  por:

$$\langle A, B \rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22} \text{ onde } A = (a_{ij}) \text{ e } B = (b_{ij}).$$

Calcule a matriz da forma  $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ , com  $a, b \in \mathbb{R}$ , mais próxima da matriz  $A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$ .

▷ **Exercício 2.24** ... Considere o subespaço  $S$  de  $\mathbb{R}^3$  gerado pelos vectores  $(1, 0, 0)$  e  $(0, 1, 0)$ .

a) Verifique que fica definido em  $\mathbb{R}^3$  um produto interno por:

$$\langle x, y \rangle = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 + x_3y_3, \text{ onde } x = (x_1, x_2, x_3) \text{ e } y = (y_1, y_2, y_3).$$

b) Determine uma base ortonormal para o subespaço  $S$ , com este produto interno.

c) Determine o elemento de  $S$  mais próximo do ponto  $(0, 0, 1)$ , usando o produto interno de a).

d) Calcule um vector diferente de zero e ortogonal a  $S$  usando o produto interno de a).

▷ **Exercício 2.25** ... No espaço vectorial real das funções contínuas definidas em  $[0, 2]$ , com o produto interno  $\langle f, g \rangle = \int_0^2 f(x)g(x) dx$ , seja  $f(x) = \exp(x)$ . Mostre que, o polinómio constante  $g$ , mais próximo de  $f$  é  $g = \frac{1}{2}(\exp(2) - 1)$ . Calcule  $\|g - f\|^2$ .

▷ **Exercício 2.26** ... Usando os produtos internos usuais em  $\mathbb{R}^2$  e  $\mathbb{R}^3$ , calcule em cada caso a projecção ortogonal  $\mathbf{P}_{\mathbf{u}}(\mathbf{v})$ , de  $\mathbf{v}$  sobre a recta gerada por  $\mathbf{u}$ :

a)  $\mathbf{u}=(1,1)$ ,  $\mathbf{v}=(2,3)$ ;

b)  $\mathbf{u}=(4,3)$ ,  $\mathbf{v}=(0,1)$ ;

c)  $\mathbf{u}=(1,1,1)$ ,  $\mathbf{v}=(1,-1,0)$ ;

d)  $\mathbf{u}=(1,0,0)$ ,  $\mathbf{v}=(0,1,2)$ .

▷ **Exercício 2.27** ... Determine as projecções ortogonais seguintes:

a)  $\mathbf{v} = (1, -1, 2)$ ,  $\mathbf{w} = (0, 1, 1)$  sobre  $F = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$  usando o produto interno Euclidiano usual de  $\mathbb{R}^3$ .

b)  $\mathbf{v} = 2t - 1$ ,  $\mathbf{w} = t^2$  sobre  $\mathbb{R}_1[t]$  usando o produto interno  $L^2$ .

## Capítulo 3

# Subespaços invariantes. Subespaços próprios. Valores próprios

### 3.1 Conjugação

► **3.1 Mudança de base** ... Suponhamos que  $\mathcal{V}$  é um espaço vectorial e que:

$$\mathcal{C} = [ \mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n ]$$

é uma base qualquer, escrita como um vector-linha com entradas vectoriais  $\mathbf{e}_i$ . Se  $\mathbf{v} \in \mathcal{V}$  é um vector qualquer em  $\mathcal{V}$ , designemos por  $v^i$  as suas componentes na base  $\mathcal{C}$ , isto é:

$$\begin{aligned} \mathbf{v} &= \sum_i v^i \mathbf{e}_i \\ &= [ \mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n ] \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix} \\ &= \mathcal{C}[\mathbf{v}]_{\mathcal{C}} \end{aligned} \tag{3.1.1}$$

Suponhamos agora que mudamos de base:

$$\mathcal{C} \longrightarrow \mathcal{C}P = \widehat{\mathcal{C}} = [ \widehat{\mathbf{e}}_1 \quad \widehat{\mathbf{e}}_2 \quad \cdots \quad \widehat{\mathbf{e}}_n ] \tag{3.1.2}$$

que escrevemos na forma matricial seguinte:

$$[ \widehat{\mathbf{e}}_1 \quad \widehat{\mathbf{e}}_2 \quad \cdots \quad \widehat{\mathbf{e}}_n ] = [ \mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n ] \begin{bmatrix} P_1^1 & P_2^1 & \cdots & P_n^1 \\ P_1^2 & P_2^2 & \cdots & P_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ P_1^n & P_2^n & \cdots & P_n^n \end{bmatrix} \tag{3.1.3}$$

ou muito simplesmente:

$$\widehat{\mathcal{C}} = \mathcal{C}P$$

Se  $\widehat{v}^i$  são as componentes do mesmo vector  $\mathbf{v}$  na base  $\widehat{\mathcal{C}}$ , isto é, se:

$$\begin{aligned} \mathbf{v} &= \sum_i \widehat{v}^i \widehat{\mathbf{e}}_i \\ &= \widehat{\mathcal{C}}[\mathbf{v}]_{\widehat{\mathcal{C}}} \end{aligned} \tag{3.1.4}$$

então vem que:

$$\mathcal{C}[\mathbf{v}]_{\mathcal{C}} = \mathbf{v} = \widehat{\mathcal{C}}[\mathbf{v}]_{\widehat{\mathcal{C}}} = \mathcal{C}P[\mathbf{v}]_{\mathcal{C}P}$$

donde se conclui que:

$$\mathcal{C} \longrightarrow \mathcal{C}P \quad \Rightarrow \quad [\mathbf{v}]_{\mathcal{C}P} = P^{-1}[\mathbf{v}]_{\mathcal{C}} \quad (3.1.5)$$

► **3.2** Suponhamos agora que  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  é um operador linear, cuja matriz relativamente à base  $\mathcal{C} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ , para  $\mathcal{V}$ , é:

$$[\mathbf{L}]_{\mathcal{C}} = [L_j^i] \quad (3.1.6)$$

Recorde que isto significa que:

$$\mathbf{L}(\mathbf{e}_j) = \sum_i L_j^i \mathbf{e}_i$$

Portanto, se  $\mathbf{v} = \mathcal{C}[\mathbf{v}]_{\mathcal{C}} \in \mathcal{V}$ , isto é, se o vector das coordenadas de  $\mathbf{v}$ , relativamente à base  $\mathcal{C}$  é:

$$[\mathbf{v}]_{\mathcal{C}} = \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix}$$

então:

$$\mathbf{L}(\mathbf{v}) = \mathbf{L}(v^j \mathbf{e}_j) = v^j \mathbf{L}(\mathbf{e}_j) = v^j (L_j^i \mathbf{e}_i) = (L_j^i v^j) \mathbf{e}_i$$

isto é, o vector das coordenadas de  $\mathbf{L}(\mathbf{v})$ , relativamente à base  $\mathcal{C}$ , é obtido multiplicando a matriz  $[\mathbf{L}]_{\mathcal{C}}$  pelo vector-coluna  $[\mathbf{v}]_{\mathcal{C}}$ :

$$[\mathbf{L}\mathbf{v}]_{\mathcal{C}} = [\mathbf{L}]_{\mathcal{C}}[\mathbf{v}]_{\mathcal{C}} \quad (3.1.7)$$

► **3.3 Conjugação ...** Suponhamos agora que escolhemos uma nova base para  $\mathcal{V}$ :

$$\widehat{\mathcal{C}} = \mathcal{C}P$$

**Como muda a representação matricial de  $\mathbf{L}$ ? Isto é, se a matriz de  $\mathbf{L}$  nesta nova base é  $\widehat{L}_j^i$ , como é que esta matriz se relaciona com a matriz  $L_j^i$ ?**

Para responder a esta questão, consideremos um vector arbitrário  $\mathbf{v} \in \mathcal{V}$ . Podemos então escrever:

$$\mathbf{v} = \mathcal{C}[\mathbf{v}]_{\mathcal{C}} = (\mathcal{C}P)[\mathbf{v}]_{\mathcal{C}P} \quad \Rightarrow \quad [\mathbf{v}]_{\mathcal{C}P} = P^{-1}[\mathbf{v}]_{\mathcal{C}}$$

Portanto:

- por um lado:

$$\mathbf{L}(\mathbf{v}) = \mathcal{C}[\mathbf{L}(\mathbf{v})]_{\mathcal{C}} = \mathcal{C}[\mathbf{L}]_{\mathcal{C}}[\mathbf{v}]_{\mathcal{C}} \quad (3.1.8)$$

- e, por outro lado:

$$\begin{aligned} \mathbf{L}(\mathbf{v}) &= (\mathcal{C}P)[\mathbf{L}(\mathbf{v})]_{\mathcal{C}P} \\ &= (\mathcal{C}P)[\mathbf{L}]_{\mathcal{C}P}[\mathbf{v}]_{\mathcal{C}P} \\ &= (\mathcal{C}P)[\mathbf{L}]_{\mathcal{C}P}P^{-1}[\mathbf{v}]_{\mathcal{C}} \end{aligned} \quad (3.1.9)$$

Comparando (3.1.8) com (3.1.9), vem que:

$$\mathcal{C}[\mathbf{L}]_{\mathcal{C}}[\mathbf{v}]_{\mathcal{C}} = (\mathcal{C}P)[\mathbf{L}]_{\mathcal{C}P}P^{-1}[\mathbf{v}]_{\mathcal{C}} \Rightarrow [\mathbf{L}]_{\mathcal{C}}[\mathbf{v}]_{\mathcal{C}} = P[\mathbf{L}]_{\mathcal{C}P}P^{-1}[\mathbf{v}]_{\mathcal{C}}$$

e como esta igualdade é válida  $\forall \mathbf{v}$ , temos que:

$$[\mathbf{L}]_{\mathcal{C}P} = P^{-1}[\mathbf{L}]_{\mathcal{C}}P \quad (3.1.10)$$

Concluindo:

**Se  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  é um operador linear num espaço vectorial de dimensão finita, então a representação matricial de  $\mathbf{L}$  varia, com a escolha da base, numa classe de conjugação de matrizes:**

$$\mathcal{C} \rightarrow \mathcal{C}P \Rightarrow [\mathbf{L}]_{\mathcal{C}P} = P^{-1}[\mathbf{L}]_{\mathcal{C}}P \quad (3.1.11)$$

► **3.4** Esta possibilidade de variar a representação matricial de  $\mathbf{L}$ , variando a base, conduz-nos naturalmente ao seguinte problema:

**Como escolher a base de  $\mathcal{V}$  de tal forma que a representação matricial de  $\mathbf{L}$  seja o mais “simples” possível? Mais formalmente - se  $L = [\mathbf{L}]_{\mathcal{C}}$  é a representação matricial de  $\mathbf{L}$  numa certa base  $\mathcal{C}$ , como seleccionar na classe de conjugação de  $L$ :**

$$\{P^{-1}LP : P \in Gl(n)\}$$

**o representante mais “simples” possível?**

► **3.5** Uma solução intuitiva para este problema consiste, grosso modo, em decompôr o espaço vectorial  $\mathcal{V}$  em “blocos simples” onde a acção de  $\mathbf{L}$  seja fácil de descrever. Os conceitos que intervêm nesta discussão são os seguintes:

- subespaços invariantes, em particular, subespaços próprios (e valores próprios associados)
- decomposição de  $\mathcal{V}$  como soma directa de subespaços invariantes
- estrutura da restrição de  $\mathbf{L}$  a cada subespaço invariante

Vamos de seguida discutir estes conceitos e posteriormente, no capítulo 8, vamos dar uma solução do problema anterior para uma classe muito importante de operadores - a classe de operadores hermíticos em espaços unitários (em particular, os operadores simétricos em espaços Euclidianos).

## 3.2 Subespaços invariantes

► **3.6 Definição** ... Seja  $\mathcal{V}$  um espaço vectorial e  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  um operador linear. Um subespaço  $\mathcal{S} \subseteq \mathcal{V}$  diz-se um **subespaço invariante do operador  $\mathbf{L}$**  se:

$$\mathbf{L}(\mathcal{S}) \subseteq \mathcal{S} \quad (3.2.1)$$

Um subespaço invariante de dimensão um diz-se um **subespaço próprio do operador  $\mathbf{L}$** .

► **3.7 Teorema** ... Seja  $\mathcal{V}$  um espaço vectorial e  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  um operador linear. Então  $\mathcal{V}$ ,  $\{\mathbf{0}\}$ ,  $\ker \mathbf{L}$  e  $\text{im } \mathbf{L}$  são subespaços invariantes do operador  $\mathbf{L}$ .

**Dem.:** Basta aplicar directamente as definições.

► **3.8 Teorema** ... Seja  $\mathcal{V}$  um espaço vectorial de dimensão finita  $n$ , e  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  um operador linear.

1. Suponhamos que  $\mathcal{S}$  é um subespaço invariante de dimensão  $k \leq n$ . Então existe uma representação matricial de  $\mathbf{L}$  da forma:

$$L = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \quad (3.2.2)$$

onde  $A$  é uma matriz  $k \times k$ ,  $B$  uma matriz  $k \times (n - k)$  e  $D$  uma matriz  $(n - k) \times (n - k)$ .

2. Suponhamos que  $\mathcal{S}$  e  $\mathcal{T}$  são subespaços invariantes de dimensão  $k$  e  $n - k$ , respectivamente, tais que:

$$\mathcal{V} = \mathcal{S} \oplus \mathcal{T}$$

Então existe uma representação matricial de  $\mathbf{L}$  da forma:

$$L = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} \quad (3.2.3)$$

onde  $A$  é uma matriz  $k \times k$  e  $D$  uma matriz  $(n - k) \times (n - k)$ .

**Dem.:** 1. Seja  $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$  uma base para  $\mathcal{S}$ , e completemos essa base a uma base  $\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$  de  $\mathcal{V}$  (isto é possível, pelo teorema da base incompleta). É claro que o subespaço  $\mathcal{T} = \text{span}\{\mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$  não é, em geral, um subespaço invariante de  $\mathbf{L}$ , embora  $\mathcal{V} = \mathcal{S} \oplus \mathcal{T}$ . De qualquer forma, podemos sempre pôr:

$$\begin{aligned} \mathbf{L}(\mathbf{e}_i) &= \sum_{j=1}^k A_i^j \mathbf{e}_j + \sum_{\beta=k+1}^n C_i^\beta \mathbf{e}_\beta, \quad i = 1, \dots, k \\ \mathbf{L}(\mathbf{e}_\alpha) &= \sum_{j=1}^k B_\alpha^j \mathbf{e}_j + \sum_{\beta=k+1}^n D_\alpha^\beta \mathbf{e}_\beta, \quad \alpha = k + 1, \dots, n \end{aligned}$$

Mas como, por hipótese,  $\mathbf{L}(\mathcal{S}) \subseteq \mathcal{S}$ , temos que  $C_i^\beta = 0$ ,  $\forall i, \beta$ , e portanto a representação matricial de  $\mathbf{L}$ , na base indicada, é:

$$L = \begin{bmatrix} A_i^j & B_\alpha^j \\ 0 & D_\alpha^\beta \end{bmatrix}$$

2. Análogo.

### 3.3 Valores e vectores próprios de um operador linear. Operadores diagonalizáveis

► **3.9** Suponhamos que  $\mathcal{S} \subseteq \mathcal{V}$  é um subespaço próprio do operador  $\mathbf{L}$ , isto é,  $\mathcal{S}$  é um subespaço invariante de dimensão um. Como  $\dim \mathcal{S} = 1$ ,  $\mathcal{S}$  é gerado por um qualquer dos seus vectores não nulos. Suponhamos que  $\mathbf{v} \in \mathcal{S} - \{\mathbf{0}\}$ . Então, como  $\dim \mathcal{S} = 1$ , tem-se que:

$$\mathbf{L}(\mathbf{v}) = \lambda \mathbf{v} \quad (3.3.1)$$

para algum escalar  $\lambda \in \mathbb{K}$ .

► **3.10 Definições** ...  $\lambda \in \mathbb{k}$  diz-se um **valor próprio de  $\mathbf{L}$**  se existir um vector não nulo  $\mathbf{v} \neq \mathbf{0}$ , em  $\mathcal{V}$ , tal que:

$$\mathbf{L}(\mathbf{v}) = \lambda \mathbf{v} \quad (3.3.2)$$

Neste caso,  $\mathbf{v}$  diz-se um **vector próprio pertencente ao valor próprio  $\lambda$** . Ao subespaço gerado por todos os vectores próprios, associados ao valor próprio  $\lambda$ , chama-se o **espaço próprio de  $\mathbf{L}$ , associado ao valor próprio  $\lambda$**  e nota-se usualmente por  $\mathcal{E}_{\mathbf{L}}(\lambda)$ , ou simplesmente por  $\mathcal{E}(\lambda)$ . Portanto:

$$\mathcal{E}(\lambda) = \mathcal{E}_{\mathbf{L}}(\lambda) \stackrel{\text{def}}{=} \{\mathbf{v} \in \mathcal{V} : \mathbf{L}(\mathbf{v}) = \lambda \mathbf{v}\} \quad (3.3.3)$$

À dimensão  $\dim \mathcal{E}(\lambda)$  chama-se a **multiplicidade geométrica** do valor próprio  $\lambda$ . O valor próprio  $\lambda$  diz-se **degenerado** quando  $\dim \mathcal{E}(\lambda) \geq 2$ .

► **3.11 Teorema** ... *Suponhamos que  $\mathbf{u}, \mathbf{v} \in \mathcal{V} - \{\mathbf{0}\}$  são vectores próprios pertencentes respectivamente aos valores próprios distintos  $\lambda, \eta \in \mathbb{k}$ , de um operador linear  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$ . Então  $\mathbf{u}$  e  $\mathbf{v}$  são linearmente independentes.*

**Dem.:** De facto, se por exemplo  $\mathbf{v} = r\mathbf{u}$ , para algum  $r \in \mathbb{k} - \{0\}$ , então viria que:

$$\eta r\mathbf{u} = \eta \mathbf{v} = \mathbf{L}(\mathbf{v}) = \mathbf{L}(r\mathbf{u}) = r\mathbf{L}(\mathbf{u}) = r\lambda\mathbf{u}$$

e portanto:

$$r(\lambda - \eta)\mathbf{u} = \mathbf{0}$$

o que implica, uma vez que  $\lambda \neq \eta$  e  $r \neq 0$ , que  $\mathbf{u} = \mathbf{0}$ , o que é absurdo.

► **3.12 Definição [Operador diagonalizável]** ... Um operador linear  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  diz-se **diagonalizável** se qualquer das seguintes condições equivalentes se verifica:

- Existe uma base de  $\mathcal{V}$ , relativamente à qual a matriz de  $\mathbf{L}$  é uma matriz diagonal.
- $\mathcal{V}$  decompõe-se numa soma directa de subespaços próprios (*subespaços invariantes de dimensão um*) de  $\mathbf{L}$ .

## 3.4 Cálculo de valores e vectores próprios

► **3.13** Suponhamos que  $\lambda \in \mathbb{k}$  é um valor próprio do operador  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  e que  $\mathcal{E}(\lambda)$  é espaço próprio associado. Como já vimos, a restrição de  $\mathbf{L}$  a  $\mathcal{E}(\lambda)$  é uma **homotetia** de razão  $\lambda$  (eventualmente  $\lambda$  pode ser 0), isto é:

$$\mathbf{L}(\mathbf{v}) = \lambda \mathbf{v} \quad \forall \mathbf{v} \in \mathcal{E}(\lambda)$$

Em particular, se  $\lambda = 0$  é valor próprio de  $\mathbf{L}$ , isto significa que o núcleo de  $\mathbf{L}$ :

$$\ker \mathbf{L} = \mathcal{E}(0)$$

não se reduz ao vector nulo  $\mathbf{0}$ , e portanto  $\mathbf{L}$  é não inversível (por outras palavras,  $\mathbf{L}$  é **singular**), ou de forma equivalente,  $\det \mathbf{L} = 0$ .

Quando  $\lambda \neq 0$ , dizer que  $\lambda$  é valor próprio de  $\mathbf{L}$ , é equivalente a dizer que 0 é valor próprio de  $\mathbf{L} - \lambda \text{Id}$ , o que, pelo parágrafo anterior, é equivalente a dizer que  $\mathbf{L} - \lambda \text{Id}$  é singular, ou ainda que:

$$\det(\mathbf{L} - \lambda \text{Id}) = 0 \quad (3.4.1)$$

► **3.14 Definição** ... O polinómio:

$$p(t) = \det(\mathbf{L} - t\text{Id}) \quad (3.4.2)$$

diz-se o **polinómio característico** de  $\mathbf{L}$ .

Portanto as raízes em  $\mathbb{k}$  da chamada **equação característica** de  $\mathbf{L}$ :

$$p(t) = \det(\mathbf{L} - t\text{Id}) = 0 \quad (3.4.3)$$

(se existirem), são exactamente os valores próprios de  $\mathbf{L}$  em  $\mathbb{k}$ .

► **3.15** ... Para calcular o polinómio característico de  $\mathbf{L}$ , usamos uma representação matricial qualquer  $L$  do operador  $\mathbf{L}$ , e pômos  $p(t) = \det(L - t\text{Id})$ . Note que o polinómio característico não depende da representação matricial de  $\mathbf{L}$ . De facto, qualquer outra representação matricial de  $\mathbf{L}$ , é do tipo  $PLP^{-1}$ , onde  $P$  é uma matriz inversível, e tem-se que:

$$\begin{aligned} \det(PLP^{-1} - t\text{Id}) &= \det(PLP^{-1} - tPP^{-1}) = \det(P(L - t\text{Id})P^{-1}) \\ &= \det(L - t\text{Id}) = p(t) \end{aligned}$$

► **3.16 Exemplo [Cálculo de valores próprios]** ... Calcule os valores e vectores próprios (reais) do operador linear  $\mathbf{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , cuja matriz na base canónica de  $\mathbb{R}^2$  :

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

A equação característica de  $\mathbf{A}$  é:

$$\begin{aligned} p(t) &= \det(A - t\text{Id}) \\ &= \det \begin{bmatrix} 3-t & 4 \\ 4 & -3-t \end{bmatrix} \\ &= t^2 - 25 = 0 \end{aligned} \quad (3.4.4)$$

cujas raízes reais (os valores próprios reais de  $\mathbf{A}$ ) são  $\lambda_1 = 5$  e  $\lambda_2 = -5$ .

Para calcular os vectores próprios  $\mathbf{x} = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}$ , pertencentes ao valor próprio  $\lambda = 5$ , devemos resolver o sistema:

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

isto é:

$$\begin{cases} -2x^1 + 4x^2 = 0 \\ 4x^1 - 8x^2 = 0 \end{cases}$$

cuja solução geral é:

$$\begin{cases} x^1 = 2s \\ x^2 = s \end{cases} \quad s \in \mathbb{R}$$

Portanto os vectores próprios de  $\mathbf{A}$ , pertencentes ao valor próprio  $\lambda_1 = 5$ , são da forma:

$$s \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad s \in \mathbb{R} - \{0\}$$

Por outras palavras, o espaço próprio  $\mathcal{E}(5)$  é:

$$\mathcal{E}(5) = \text{span} \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

Procedendo da mesma forma relativamente ao outro valor próprio  $\lambda_2 = -5$ , podemos calcular que os vectores próprios de  $\mathbf{A}$ , pertencentes ao valor próprio  $\lambda_2 = -5$ , são da forma:

$$s \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad s \in \mathbb{R} - \{0\}$$

Note que neste exemplo os vectores próprios  $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$  e  $\mathbf{u}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$  formam uma base  $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$  de  $\mathbb{R}^2$  relativamente à qual a matriz de  $\mathbf{A}$  é diagonal:

$$[\mathbf{A}]_{\mathcal{B}} = \begin{bmatrix} 5 & 0 \\ 0 & -5 \end{bmatrix}$$

portanto  $\mathbf{A}$  é um operador diagonalizável.

► **3.17 Exemplo [Cálculo de valores próprios]** ... Calcule os valores e vectores próprios (reais) do operador linear  $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , cuja matriz na base canónica de  $\mathbb{R}^3$  é:

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

A equação característica de  $\mathbf{A}$  é:

$$\begin{aligned} p(t) &= \det(A - t \text{Id}) \\ &= \det \begin{bmatrix} 1-t & 0 & 0 \\ -5 & 2-t & 0 \\ 2 & 3 & 7-t \end{bmatrix} \\ &= (1-t)(2-t)(7-t) = 0 \end{aligned} \tag{3.4.5}$$

cujas raízes reais (os valores próprios reais de  $\mathbf{A}$ ) são  $\lambda_1 = 1$ ,  $\lambda_2 = 2$  e  $\lambda_3 = 7$ . Para calcular os vectores próprios  $\mathbf{x} = \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}$ , pertencentes ao valor próprio  $\lambda_2 = 2$ , devemos resolver o sistema:

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

isto é:

$$\begin{cases} -x^1 & = 0 \\ -5x^1 & = 0 \\ 2x^1 + 3x^2 + 5x^3 & = 0 \end{cases}$$

cuja solução geral é:

$$\begin{cases} x^1 & = 0 \\ x^2 & = -\frac{5}{3}s \\ x^3 & = s \end{cases} \quad s \in \mathbb{R}$$

Portanto os vectores próprios de  $\mathbf{A}$ , pertencentes ao valor próprio  $\lambda_2 = 2$ , são da forma:

$$s \begin{bmatrix} 0 \\ -\frac{5}{3} \\ 1 \end{bmatrix} \quad s \in \mathbb{R} - \{0\}$$

Procedendo da mesma forma relativamente aos outros valores próprios  $a_1 = 1$  e  $a_3 = 7$ , podemos calcular os correspondentes vectores próprios.

### Notas ...

1. Note que o polinómio característico  $p(t) = \det(\mathbf{L} - t\text{Id})$ , de um operador linear  $\mathbf{L} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , é sempre um polinómio do 3.º grau, do tipo:

$$p(t) = -t^3 + bt^2 + ct + d \quad b, c, d \in \mathbb{R}$$

e por isso admite sempre uma raiz real  $\lambda \in \mathbb{R}$  (eventualmente nula). Se  $\lambda \neq 0$ , concluímos portanto que, neste caso, existe sempre um subespaço próprio invariante  $\mathcal{E}(\lambda) \subseteq \mathbb{R}^3$ , de dimensão superior ou igual a 1.

2. Todo o operador linear  $\mathbf{L} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  tem quando muito 3 valores próprios distintos. Se  $\mathbf{L}$  tem exactamente 3 valores próprios distintos, então os correspondentes vectores próprios formam uma base de  $\mathbb{R}^3$ , e a matriz de  $\mathbf{L}$  nessa base, é uma matriz diagonal cujas entradas da diagonal principal, são esses valores próprios.

## 3.5 Sistemas dinâmicos lineares discretos

► **3.18** Um **sistema dinâmico linear discreto** é um sistema recursivo do tipo:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) \quad (3.5.1)$$

onde  $\mathbf{A}$  é uma matriz  $n \times n$ , e

$$\mathbf{x} : \mathbb{N}_0 \rightarrow \mathbb{R}^n$$

é uma função que a cada "instante de tempo" discreto  $k = 0, 1, 2, \dots$ , associa um vector  $\mathbf{x}(k) \in \mathbb{R}^n$ .

A equação (3.5.1) indica pois a lei de evolução do sistema: conhecido o **valor inicial** do sistema:

$$\mathbf{x}(0) = \mathbf{x}_0 \quad (3.5.2)$$

os valores nos instantes seguintes são calculados sucessivamente através de:

$$\begin{aligned} \mathbf{x}(1) &= \mathbf{A}\mathbf{x}_0 \\ \mathbf{x}(2) &= \mathbf{A}\mathbf{x}(1) = \mathbf{A}^2\mathbf{x}_0 \\ \mathbf{x}(3) &= \mathbf{A}\mathbf{x}(2) = \mathbf{A}^3\mathbf{x}_0 \\ &\vdots \\ \mathbf{x}(k) &= \mathbf{A}\mathbf{x}(k-1) = \mathbf{A}^k\mathbf{x}_0 \\ &\vdots \end{aligned} \quad (3.5.3)$$

► **3.19** Quando a matriz  $\mathbf{A}$  de evolução é diagonalizável, o cálculo explícito da evolução através da equação (3.5.3):

$$\mathbf{x}(k) = \mathbf{A}^k \mathbf{x}(0) \quad (3.5.4)$$

torna-se particularmente simples.

De facto, suponhamos que  $\mathcal{B} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n]$  é uma base de  $\mathbb{R}^n$  constituída por vectores próprios (não necessariamente distintos) da matriz  $\mathbf{A}$ :

$$\mathbf{A} \mathbf{v}_j = \lambda_j \mathbf{v}_j, \quad j = 1, 2, \dots, n \quad (3.5.5)$$

Se  $\mathcal{C} = [\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n]$  é a base canónica de  $\mathbb{R}^n$ , põmos, como habitualmente:

$$\mathcal{B} = \mathcal{C}P \quad \Rightarrow \quad \mathbf{x}_{\mathcal{B}} = \mathbf{x}_{\mathcal{C}}P = P^{-1}\mathbf{x}_{\mathcal{C}} \quad (3.5.6)$$

Portanto, pondo  $\mathbf{x}_{\mathcal{C}}(k) = \mathbf{x}(k)$  em (3.5.4), vem que:

$$\begin{aligned} \mathbf{x}_{\mathcal{B}}(k) &= P^{-1}\mathbf{x}_{\mathcal{C}}(k) \\ &= P^{-1}\mathbf{A}^k\mathbf{x}_{\mathcal{C}}(0) \\ &= P^{-1}\mathbf{A}^kP\mathbf{x}_{\mathcal{B}}(0) \\ &= (P^{-1}\mathbf{A}P)^k\mathbf{x}_{\mathcal{B}}(0) \\ &= (\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n))^k\mathbf{x}_{\mathcal{B}}(0) \\ &= \text{diag}(\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k)\mathbf{x}_{\mathcal{B}}(0) \end{aligned} \quad (3.5.7)$$

Isto é, a  $i$ -componente de  $\mathbf{x}(k)$  na base  $\mathcal{B}$ , que diagonaliza  $\mathbf{A}$ , é obtida muito simplesmente multiplicando a potência de expoente  $k$ , do valor próprio  $\lambda_i$ , pela  $i$ -componente do vector inicial  $\mathbf{x}(0)$  na base  $\mathcal{B}$ :

$$x_{\mathcal{B}}^i(k) = (\lambda_i)^k x_{\mathcal{B}}^i(0) \quad (3.5.8)$$

Note que no membro direito da equação anterior não há soma no índice  $i$ !

Na prática procedemos como segue:

[1]. Escrevemos o vector inicial  $\mathbf{x}(0)$  na base  $\mathcal{B}$ , calculando assim as componentes  $c^i = x_{\mathcal{B}}^i(0)$ :

$$\mathbf{x}(0) = \mathcal{B}\mathbf{x}_{\mathcal{B}}(0) = \sum_i c^i \mathbf{v}_i$$

[2]. Pômos:

$$\mathbf{x}(k) = \mathcal{C}\mathbf{x}_{\mathcal{C}}(k) = \mathcal{B}\mathbf{x}_{\mathcal{B}}(k) = \sum_i (c^i \lambda_i^k) \mathbf{v}_i$$

Concluindo :

$$\mathbf{x}(k) = \sum_i (c^i \lambda_i^k) \mathbf{v}_i, \quad \text{onde} \quad \mathbf{x}(0) = \sum_i c^i \mathbf{v}_i \quad (3.5.9)$$

► **3.20** Números de Fibonacci ... são definidos pela lei recursiva (de segunda ordem) seguinte:

$$x(k+2) = x(k+1) + x(k) \quad (3.5.10)$$

isto é, cada número de Fibonacci é obtido somando os dois anteriores. As condições iniciais são:

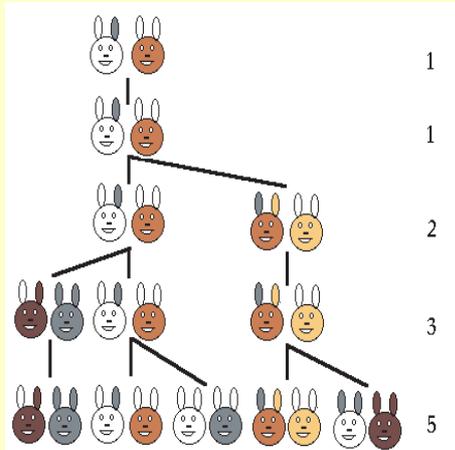
$$x(0) = a, \quad x(1) = b \quad (3.5.11)$$

Por exemplo, para:

$$x(0) = a = 0, x(1) = b = 1 \tag{3.5.12}$$

obtem-se:

$$0 \quad 1 \quad 1 \quad 2 \quad 3 \quad 5 \quad 8 \quad 13 \quad 21 \quad 34 \quad \dots \tag{3.5.13}$$



Foram criados pelo matemático italiano Fibonacci como um modelo simplificado do crescimento de uma população de coelhos. Neste modelo:

$$x(n) = \text{número total de pares de coelhos no ano } n \tag{3.5.14}$$

O processo inicia-se no ano  $n = 0$  com um único par de coelhos jovens. Ao fim de cada ano, cada par dá origem a um novo par de descendentes. No entanto, cada par necessita de um ano para procriar o seu par de descendentes.

► **3.21 Números de Fibonacci. Escrita matricial** ... Definamos, para cada  $k \in \mathbb{N}$ , um vector  $\mathbf{x}(k) \in \mathbb{R}^2$  através de:

$$\mathbf{x}(k) = \begin{bmatrix} x(k) \\ x(k+1) \end{bmatrix} \in \mathbb{R}^2 \tag{3.5.15}$$

Então (3.5.10) pode ser escrita na forma matricial:

$$\begin{bmatrix} x(k+1) \\ x(k+2) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x(k) \\ x(k+1) \end{bmatrix} \tag{3.5.16}$$

isto é:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k), \quad \text{onde } \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \tag{3.5.17}$$

► **3.22 Cálculo explícito dos números de Fibonacci** ... Para calcular a forma explícita dos números de Fibonacci, usamos o método descrito no número 3.19.

Para isso, determinamos os valores e vectores próprios da matriz  $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ . Um cálculo simples mostra que eles são:

$$\begin{aligned} \lambda_1 &= \frac{1 + \sqrt{5}}{2} = 1.618034\dots, & \mathbf{v}_1 &= \begin{bmatrix} \frac{-1 + \sqrt{5}}{2} \\ 1 \end{bmatrix} \\ \lambda_2 &= \frac{1 - \sqrt{5}}{2} = -0.618034\dots, & \mathbf{v}_2 &= \begin{bmatrix} \frac{-1 - \sqrt{5}}{2} \\ 1 \end{bmatrix} \end{aligned} \tag{3.5.18}$$

Escrevemos agora o vector inicial na base  $\mathcal{B}$ :

$$\begin{aligned} \mathbf{x}_{\mathcal{B}}(0) &= P^{-1}\mathbf{x}_{\mathcal{C}}(0) \\ &= \begin{bmatrix} \frac{-1 + \sqrt{5}}{2} & \frac{-1 - \sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} a \\ b \end{bmatrix} \\ &= \begin{bmatrix} \frac{2a + (1 + \sqrt{5})b}{2\sqrt{5}} \\ -\frac{2a + (1 - \sqrt{5})b}{2\sqrt{5}} \end{bmatrix}_{\mathcal{B}} \end{aligned} \tag{3.5.19}$$

isto é:

$$\mathbf{x}(0) = \frac{2a + (1 + \sqrt{5})b}{2\sqrt{5}} \mathbf{v}_1 - \frac{2a + (1 - \sqrt{5})b}{2\sqrt{5}} \mathbf{v}_2 \quad (3.5.20)$$

Usando a fórmula (3.5.9) vem então que:

$$\begin{aligned} \mathbf{x}(k) &= \frac{2a + (1 + \sqrt{5})b}{2\sqrt{5}} \lambda_1^k \mathbf{v}_1 - \frac{2a + (1 - \sqrt{5})b}{2\sqrt{5}} \lambda_2^k \mathbf{v}_2 \\ &= \frac{2a + (1 + \sqrt{5})b}{2\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^k \begin{bmatrix} \frac{-1 + \sqrt{5}}{2} \\ 1 \end{bmatrix} - \frac{2a + (1 - \sqrt{5})b}{2\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^k \begin{bmatrix} \frac{-1 - \sqrt{5}}{2} \\ 1 \end{bmatrix} \end{aligned}$$

donde se deduz que:

$$x(k) = \frac{(-1 + \sqrt{5})a + 2b}{2\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^k + \frac{(1 + \sqrt{5})a - 2b}{2\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^k \quad (3.5.21)$$

► **3.23 Fórmula de Binet** ... Para os valores iniciais  $a = 0$  e  $b = 1$ , obtemos a chamada **fórmula de Binet**:

$$x(k) = \frac{1}{\sqrt{5}} \left[ \left( \frac{1 + \sqrt{5}}{2} \right)^k - \left( \frac{1 - \sqrt{5}}{2} \right)^k \right] \quad (3.5.22)$$

► **3.24 Número de ouro** ... Os valores próprios da matriz  $\mathbf{A}$ , verificam as desigualdades seguintes:

$$0 < |\lambda_2| = \frac{\sqrt{5} - 1}{2} < 1 < \lambda_1 = \frac{1 + \sqrt{5}}{2} \quad (3.5.23)$$

Portanto os termos que envolvem  $\lambda_1^k$  divergem para  $\infty$ , enquanto que os que envolvem  $\lambda_2^k$  convergem para 0.

O valor próprio dominante  $\lambda_1 = \frac{1 + \sqrt{5}}{2} = 1.618034\dots$  é o chamado **número de ouro** (ou **razão de ouro**). Desempenha um papel muito importante em crescimento em espiral em vários fenómenos naturais bem como em certas criações artísticas em arquitectura e pintura.

► **3.25 Exercício** ... Considere a aplicação linear:

$$\begin{aligned} \mathbf{T}: \quad \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x, y, z) &\longmapsto \mathbf{T}(x, y, z) = (4z, x + 2y + z, 2x + 4y - 2z) \end{aligned}$$

a.) Calcular a matriz de  $\mathbf{T}$  relativamente à base canónica de  $\mathbb{R}^3$ . Calcular o núcleo e a imagem de  $\mathbf{T}$ .

b.) Calcular os valores próprios de  $\mathbf{T}$  e, se possível, uma base de  $\mathbb{R}^3$  constituída por vectores próprios de  $\mathbf{T}$ . Calcule a matriz de  $\mathbf{T}$  relativamente a esta nova base.

c.) Usando os resultados das alíneas anteriores, calcule  $\mathbf{T}^3(0, 0, -4)$ , onde  $\mathbf{T}^3 = \mathbf{T} \circ \mathbf{T} \circ \mathbf{T}$ .

Resolução ...

a.) A matriz é  $T = \begin{bmatrix} 0 & 0 & 4 \\ 1 & 2 & 1 \\ 2 & 4 & -2 \end{bmatrix}$ .  $\ker \mathbf{T} = \{(x, y, z) \in \mathbb{R}^3 : \mathbf{T}(x, y, z) = (4z, x + 2y + z, 2x + 4y - 2z) = (0, 0, 0)\}$  o que implica que:

$$\begin{cases} 4z = 0 \\ x + 2y + z = 0 \\ 2x + 4y - 2z = 0 \end{cases} \Rightarrow \begin{cases} z = 0 \\ x + 2y = 0 \\ 2x + 4y = 0 \end{cases} \Rightarrow \begin{cases} z = 0 \\ x + 2y = 0 \end{cases} \Rightarrow \begin{cases} x = -2t \\ y = t \\ z = 0 \end{cases} \quad t \in \mathbb{R}$$

isto é  $\ker \mathbf{T} = \{t(-2, 1, 0) : t \in \mathbb{R}^3\} = \text{span}\{(-2, 1, 0)\}$  que é a recta de  $\mathbb{R}^3$  gerada por  $(-2, 1, 0)$  e de equações cartesianas  $x + 2y = 0$  e  $z = 0$ .

A imagem de  $\mathbf{T}$  é gerada por  $\mathbf{T}(\mathbf{e}_1) = (0, 1, 2)$ ,  $\mathbf{T}(\mathbf{e}_2) = (0, 2, 4)$  e  $\mathbf{T}(\mathbf{e}_3) = (4, 1, -2)$ , isto é:

$$\begin{aligned} \text{im} \mathbf{T} &= \text{span}\{(0, 1, 2), (0, 2, 4), (4, 1, -2)\} \\ &= \{(x, y, z) \in \mathbb{R}^3 : (x, y, z) = a(0, 1, 2) + b(0, 2, 4) + c(4, 1, -2), \quad a, b, c \in \mathbb{R}\} \end{aligned}$$

Portanto:

$$\begin{cases} 4c = x \\ a + 2b + c = y \\ 2a + 4b - 2c = z \end{cases} \Rightarrow \dots \Rightarrow \begin{cases} a + 2b + c = y \\ 4c = 2y - z \\ 0 = x - 2y + z \end{cases}$$

isto é,  $\text{im} \mathbf{T}$  é o plano  $x - 2y + z = 0$  em  $\mathbb{R}^3$ .

b.) A equação característica é  $\det(T - \lambda \text{Id}) = \det \begin{bmatrix} -\lambda & 0 & 4 \\ 1 & 2 - \lambda & 1 \\ 2 & 4 & -2 - \lambda \end{bmatrix} = -\lambda^3 + 16\lambda = 0$ ,  
cujas raízes são  $\lambda = -4, 0, +4$ .

$$\mathcal{E}(\mathbf{T}; -4) = \text{span}\{(1, 0, -1)\}$$

$$\mathcal{E}(\mathbf{T}; 0) = \text{span}\{(-2, 1, 0)\}$$

$$\mathcal{E}(\mathbf{T}; +4) = \text{span}\{(1, 1, 1)\}$$

e os vectores  $\{\mathbf{e}_1 = (1, 0, -1), \mathbf{e}_2 = (-2, 1, 0), \mathbf{e}_3 = (1, 1, 1)\}$  constituem uma base de vectores próprios de  $\mathbf{T}$  que é, por isso, diagonalizável. Nesta base a matriz de  $\mathbf{T}$  é  $\text{diag}(-4, 0, 4)$ .

c.) Calculando as componentes do vector  $(0, 0, -4)$  na base de vectores próprios de  $\mathbf{T}$ , calculada anteriormente, vem que:

$$(0, 0, -4) = a(1, 0, -1) + b(-2, 1, 0) + c(1, 1, 1) = (a - 2b + c, b + c, -a + c)$$

donde se deduz que  $a = -1, b = 1, c = -1$ . Portanto:

$$\begin{aligned} \mathbf{T}^3(0, 0, -4) &= -\mathbf{T}^3(1, 0, -1) + \mathbf{T}^3(-2, 1, 0) - \mathbf{T}^3(1, 1, 1) \\ &= -(-4)^3(1, 0, -1) + 0^3(-2, 1, 0) - 4^3(1, 1, 1) \\ &= (0, -64, -128) \end{aligned}$$

► **3.26 Exercício** ... Considere a aplicação linear  $\mathbf{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definida por:

$$\mathbf{A}(x, y) = (6x - 2y, -2x + 9y)$$

a.) Mostrar que  $\mathbf{A}$  é diagonalizável e calcular uma base ortonormada para  $\mathbb{R}^2$  (com a estrutura Euclideana usual) constituída por vectores próprios de  $\mathbf{A}$ .

b.) Considere as sucessões  $(x_n)$  e  $(y_n)$ , definidas pelas fórmulas de recorrência seguintes:

$$\begin{cases} x_{n+1} = 6x_n - 2y_n \\ y_{n+1} = -2x_n + 9y_n \end{cases}, \quad n \geq 0 \quad \text{e} \quad \begin{cases} x_0 = 1 \\ y_0 = 1 \end{cases}$$

Calcule  $x_n$  e  $y_n$  como funções de  $n$ .

### Resolução ...

a.) A matriz de  $\mathbf{A}$  relativamente à base canónica de  $\mathbb{R}^2$  é a matriz simétrica:

$$A = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}$$

Os valores próprios calculam-se por:

$$\det(A - \lambda \text{Id}) = \det \begin{pmatrix} 6 - \lambda & -2 \\ -2 & 9 - \lambda \end{pmatrix} = (6 - \lambda)(9 - \lambda) - 4 = 0 \Rightarrow \lambda^2 - 15\lambda + 50 = 0 \Rightarrow \lambda = 5, 10$$

Como existem dois ( $= \dim \mathbb{R}^2$ ) valores próprios distintos,  $\mathbf{A}$  é diagonalizável. Os espaços próprios calculam-se da forma habitual e são:

$$\mathcal{E}(5) = \mathbb{R} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{e} \quad \mathcal{E}(10) = \mathbb{R} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Estes espaços são ortogonais (tinham que o ser, pelo teorema espectral!). Uma base ortonormada para  $\mathbb{R}^2$  constituída por vectores próprios de  $\mathbf{A}$  é:

$$\mathcal{B} = \left\{ \mathbf{u}_1 = \frac{(2, 1)}{\sqrt{5}}, \quad \mathbf{u}_2 = \frac{(1, -2)}{\sqrt{5}} \right\}$$

a.) Pondo  $\mathbf{x}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$ , as fórmulas de recorrência dadas escrevem-se na forma vectorial:

$$\mathbf{x}_{n+1} = A\mathbf{x}_n, \quad \mathbf{x}_0 = (1, 1)$$

onde  $A = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}$ . Os cálculos devem ser feitos na base  $\mathcal{B}$  que diagonaliza o operador  $\mathbf{A}$ . Escrevendo o vector  $\mathbf{x}_n$  na base  $\mathcal{B}$ , vem que:

$$\begin{aligned} \mathbf{x}_n &= (\mathbf{x}_n \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{x}_n \cdot \mathbf{u}_2)\mathbf{u}_2 \\ &= \frac{1}{\sqrt{5}}(2x_n + y_n)\mathbf{u}_1 + \frac{1}{\sqrt{5}}(x_n - 2y_n)\mathbf{u}_2 \end{aligned} \quad (3.5.24)$$

isto é, as componentes de  $\mathbf{x}_n$  na base  $\mathcal{B}$  são  $\tilde{x}_n = \frac{2x_n + y_n}{\sqrt{5}}$ ,  $\tilde{y}_n = \frac{x_n - 2y_n}{\sqrt{5}}$ .

Na base  $\mathcal{B}$  as fórmulas de recorrência escrevem-se na forma:

$$\begin{pmatrix} \tilde{x}_{n+1} \\ \tilde{y}_{n+1} \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} \tilde{x}_n \\ \tilde{y}_n \end{pmatrix} = \begin{pmatrix} 5\tilde{x}_n \\ 10\tilde{y}_n \end{pmatrix}$$

Portanto:

$$\begin{pmatrix} \tilde{x}_1 \\ \tilde{y}_1 \end{pmatrix} = \begin{pmatrix} 5\tilde{x}_0 \\ 10\tilde{y}_0 \end{pmatrix}, \quad \begin{pmatrix} \tilde{x}_2 \\ \tilde{y}_2 \end{pmatrix} = \begin{pmatrix} 5\tilde{x}_1 \\ 10\tilde{y}_1 \end{pmatrix} = \begin{pmatrix} 5^2\tilde{x}_0 \\ 10^2\tilde{y}_0 \end{pmatrix}, \quad \dots \quad \begin{pmatrix} \tilde{x}_n \\ \tilde{y}_n \end{pmatrix} = \begin{pmatrix} 5^n\tilde{x}_0 \\ 10^n\tilde{y}_0 \end{pmatrix}$$

Mas  $\tilde{x}_0 = \frac{2x_0+y_0}{\sqrt{5}} = \frac{3}{\sqrt{5}}$ ,  $\tilde{y}_0 = \frac{x_0-2y_0}{\sqrt{5}} = \frac{-1}{\sqrt{5}}$ . Portanto:

$$\begin{cases} \tilde{x}_n = \frac{2x_n+y_n}{\sqrt{5}} = 5^n \frac{3}{\sqrt{5}} \\ \tilde{y}_n = \frac{x_n-2y_n}{\sqrt{5}} = 10^n \frac{-1}{\sqrt{5}} \end{cases}$$

e resolvendo em ordem a  $x_n$  e  $y_n$  obtemos:

$$x_n = 2 \times 5^{n-1}(3 - 2^{n-1}), \quad y_n = 5^{n-1}(3 + 4 \times 2^{n-1})$$

### 3.6 Exercícios

▷ **Exercício 3.1** ... Seja  $f$  um endomorfismo de  $\mathbb{R}_2[X]$  tal que  $X + X^2$  é um vector próprio associado ao valor próprio 2,  $-1 + X$  é um vector próprio associado ao valor próprio 5 e  $X^2$  é um vector próprio associado ao valor próprio -3. Determine  $f(a_0 + a_1X + a_2X^2)$ .

▷ **Exercício 3.2** ... Seja  $f$  um endomorfismo de  $\mathbb{C}_2[X]$  munido da estrutura usual de espaço vectorial complexo. Suponha que :

$1 + iX$  é um vector próprio de valor próprio  $i$ ,

$1 - X$  é um vector próprio de valor próprio 1 e

$X^2$  é um vector próprio de valor próprio  $-1$ .

Calcule  $f(a + bX + cX^2)$ .

▷ **Exercício 3.3** ... Seja  $f$  um automorfismo de um espaço vectorial  $E$ . Qual a relação entre os valores próprios de  $f$  e os valores próprios de  $f^{-1}$ ?

▷ **Exercício 3.4** ... Sejam  $f$  e  $g$  endomorfismos de  $E$ .

a) Mostre que, se  $\mathbf{u}$  é um vector próprio de  $f$ , com valor próprio associado  $\lambda$  então  $\mathbf{u}$  é um vector próprio de  $f \circ f$  com valor próprio associado  $\lambda^2$ .

b) Mostre que, se  $\mathbf{u}$  é um vector próprio de  $f$  e de  $g$ , então  $\mathbf{u}$  é um vector próprio de  $g \circ f$  e de qualquer combinação linear de  $f$  e de  $g$ ,  $af + bg$ .

c) Mostre que, se todos os elementos não nulos de  $E$  são vectores próprios de  $f$ , então  $f$  tem um único valor próprio (e, portanto, existe  $\alpha \in \mathbb{R}$  tal que, para qualquer  $\mathbf{u} \in E$ ,  $f(\mathbf{u}) = \alpha\mathbf{u}$ ).

▷ **Exercício 3.5** ... Seja  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  um endomorfismo tal que  $\{(x, y, z) \in \mathbb{R}^3 : x = y = z\}$  e  $\{(x, y, z) \in \mathbb{R}^3 : x - y + z = 0\}$  são subespaços próprios associados respectivamente aos valores próprios 1 e 2. Determine  $f((x, y, z))$ .

▷ **Exercício 3.6** ... Em cada um dos seguintes casos, determine, se existirem, os valores próprios de  $f$ , os subespaços próprios associados e as respectivas dimensões e diga se  $f$  é diagonalizável; no caso de  $f$  ser diagonalizável, indique uma base do domínio de  $f$  composta por vectores próprios de  $f$  e indique a matriz de  $f$  relativamente a essa base.

- a)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $f(x, y) = (2x - y, y)$ ;    b)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $f(x, y) = (-x, -y)$ ;  
 c)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $f(x, y) = (3x + y, 12x + 2y)$ ;  
 d)  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $f(x, y, z) = (3x + y + z, 3y + z, 3z)$ ;  
 e)  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $f(x, y, z) = (3x + y + z, 3y, 3z)$ ;  
 f)  $f : \mathbb{R}_2[X] \rightarrow \mathbb{R}_2[X]$ ,  $f(P) = P(0) + XP(1) + X^2P(-1)$ ;  
 g)  $f : \mathbb{R}_3[X] \rightarrow \mathbb{R}_3[X]$ ,  $f(P) = P + (X + 1)P'$ ;  
 h)  $f : M_{2,2}(\mathbb{R}) \rightarrow M_{2,2}(\mathbb{R})$ ,  $f \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) = \left( \begin{array}{cc} 3a + 2b + c + d & 2a + 3b + c - d \\ 2c & -c \end{array} \right)$ .  
 i)  $f : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ ,  $f(u, v) = (iu, u + v)$ ;

▷ **Exercício 3.7** ... Calcular fórmulas explícitas para as soluções das seguintes fórmulas recursivas:

$$\text{a). } \begin{cases} x(k+1) = x(k) - 2y(k) \\ y(k+1) = -2x(k) + y(k) \end{cases}, \quad \begin{cases} x(0) = 1 \\ y(0) = 0 \end{cases}$$

$$\text{b). } \begin{cases} x(k+1) = \frac{1}{2}x(k) + y(k) \\ y(k+1) = y(k) - 2z(k) \\ z(k+1) = \frac{1}{3}z(k) \end{cases}, \quad \begin{cases} x(0) = 1 \\ y(0) = -1 \\ z(0) = 1 \end{cases}$$

$$\text{c). } x(k+2) = -x(k+1) + 2x(k), \quad x(0) = 1, \quad x(1) = 2$$

$$\text{d). } x(k+3) = 2x(k+2) + x(k+1) - 2x(k), \quad x(0) = 0, \quad x(1) = 2, \quad x(2) = 3$$

## Capítulo 4

# Transformações ortogonais e unitárias

### 4.1 Transformações ortogonais e unitárias. Exemplos

► **4.1 Definição** ... [Transformações ortogonais] ... Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço Euclidiano de dimensão  $n$ , isto é, um espaço vectorial real com um produto interno Euclidiano. Um operador linear  $\mathbf{A} : \mathcal{V} \rightarrow \mathcal{V}$  diz-se uma **transformação ortogonal** de  $\mathcal{V}$ , se  $\mathbf{A}$  preserva o produto interno  $\langle | \rangle$ , i.e.:

$$\langle \mathbf{A}(\mathbf{v}) | \mathbf{A}(\mathbf{w}) \rangle = \langle \mathbf{v} | \mathbf{w} \rangle \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V} \quad (4.1.1)$$

Se  $A$  é a matriz de uma tal transformação ortogonal, relativamente a uma base ortonormada de  $\mathcal{V}$ , então (4.1.1) escreve-se na seguinte forma matricial:

$$(\mathbf{A}\mathbf{v})^t \mathbf{A}\mathbf{w} = \mathbf{v}^t \mathbf{w} \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V}$$

ou ainda:

$$\mathbf{v}^t \mathbf{A}^t \mathbf{A}\mathbf{w} = \mathbf{v}^t \mathbf{w} = \mathbf{v}^t \mathbf{I} \mathbf{w} \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V}$$

o que significa que a matriz  $A$  é uma **matriz ortogonal**, isto é:

$$\mathbf{A}^t \mathbf{A} = \mathbf{I} \quad (4.1.2)$$

Note ainda que se  $A$  é uma matriz ortogonal então, uma vez que:

$$1 = \det \mathbf{I} = \det (\mathbf{A}\mathbf{A}^t) = \det A \det (\mathbf{A}^t) = (\det A)^2, \quad \text{e} \quad \det A \in \mathbb{R}$$

concluimos que  $\det A = \pm 1$  e, em particular  $A$  é inversível com:

$$\mathbf{A}^{-1} = \mathbf{A}^t$$

O conjunto de todas as matrizes ortogonais  $n \times n$  reais formam um subgrupo de  $Gl(n) = Gl(n; \mathbb{R})$ , que se diz o **grupo ortogonal** em dimensão  $n$  e nota-se por  $\mathcal{O}(n)$ . O conjunto de todas as matrizes ortogonais  $n \times n$  reais, de determinante 1, formam um subgrupo de  $\mathcal{O}(n)$ , que se diz o **grupo ortogonal especial** em dimensão  $n$  e nota-se por  $\mathcal{SO}(n)$ :

$$\begin{aligned} \mathcal{O}(n) &= \{A \in \mathcal{M}_n(\mathbb{R}) : \mathbf{A}^t \mathbf{A} = \mathbf{I}\} \\ \mathcal{SO}(n) &= \{A \in \mathcal{M}_n(\mathbb{R}) : \mathbf{A}^t \mathbf{A} = \mathbf{I}, \quad \text{e} \quad \det A = 1\} \end{aligned} \quad (4.1.3)$$

► **4.2 Definição** ... [Transformações unitárias] ... Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço unitário de dimensão  $n$ , isto é, um espaço vectorial complexo com um produto interno Hermitiano. Um operador linear  $\mathbf{A} : \mathcal{V} \rightarrow \mathcal{V}$  diz-se uma **transformação unitária** de  $\mathcal{V}$ , se  $\mathbf{A}$  preserva o produto interno hermitiano  $\langle | \rangle$ , i.e.:

$$\langle \mathbf{A}(\mathbf{v}) | \mathbf{A}(\mathbf{w}) \rangle = \langle \mathbf{v} | \mathbf{w} \rangle \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V} \quad (4.1.4)$$

Se  $A$  é a matriz de uma tal transformação unitária, relativamente a uma base ortonormada de  $\mathcal{V}$ , então (4.4.1) escreve-se na seguinte forma matricial:

$$(A\mathbf{v})^t \overline{A\mathbf{w}} = \mathbf{v}^t \overline{\mathbf{w}} \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V}$$

ou ainda:

$$\mathbf{v}^t A^t \overline{A\mathbf{w}} = \mathbf{v}^t \overline{\mathbf{w}} = \mathbf{v}^t \mathbf{I} \overline{\mathbf{w}} \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V}$$

o que significa que a matriz  $A$  é uma **matriz unitária**, isto é:

$$A^t \overline{A} = \mathbf{I} \quad (4.1.5)$$

Dada uma matriz  $A$ , define-se a respectiva **matriz adjunta**  $A^\dagger$ , como sendo a conjugada transposta de  $A$ :

$$A^\dagger = \overline{A}^t \quad (4.1.6)$$

Portanto  $A$  é unitária sse:

$$AA^\dagger = \mathbf{I} \quad (4.1.7)$$

Note ainda que, uma vez que:

$$\det(AA^\dagger) = \det(A\overline{A}^t) = \det A \det(\overline{A}^t) = \det A \overline{\det A} = |\det A|^2$$

concluimos que, se  $A$  é unitária, então  $|\det A| = 1$  e, em particular  $A$  é inversível com:

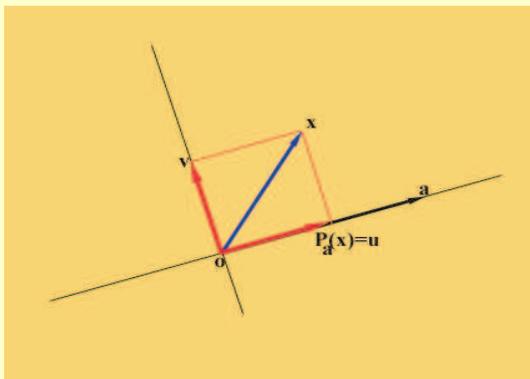
$$A^{-1} = A^\dagger$$

Note que agora  $\det A \in \mathbb{C}$ .

O conjunto de todas as matrizes unitárias  $n \times n$  complexas formam um subgrupo de  $Gl(n; \mathbb{C})$ , que se diz o **grupo unitário** em dimensão  $n$  e nota-se por  $\mathcal{U}(n)$ . O conjunto de todas as matrizes unitárias  $n \times n$  complexas, de determinante 1, formam um subgrupo de  $\mathcal{U}(n)$ , que se diz o **grupo unitário especial** em dimensão  $n$  e nota-se por  $\mathcal{SU}(n)$ :

$$\begin{aligned} \mathcal{U}(n) &= \left\{ A \in \mathcal{M}_n(\mathbb{C}) : A^\dagger A = \mathbf{I} \right\} \\ \mathcal{SU}(n) &= \left\{ A \in \mathcal{M}_n(\mathbb{C}) : A^\dagger A = \mathbf{I}, \text{ e } \det A = 1 \right\} \end{aligned} \quad (4.1.8)$$

► **4.3 Projecção ortogonal sobre uma recta, em  $\mathbb{E}^2$  ...**



Sejam  $\mathbf{a} \neq \mathbf{0}$  e  $\mathbf{x}$  dois vectores em  $\mathbb{E}^2$ . Então existe um único vector  $\mathbf{u}$ , na recta gerada por  $\mathbf{a}$ , e um único vector  $\mathbf{v}$ , ortogonal a  $\mathbf{a}$ , tais que  $\mathbf{x} = \mathbf{u} + \mathbf{v}$ . O vector  $\mathbf{u}$ , notado por  $\mathbf{P}_a(\mathbf{x})$ , diz-se a **projecção ortogonal** de  $\mathbf{x}$  sobre a recta gerada por  $\mathbf{a}$ , e é calculado da seguinte forma.

Uma vez que  $\mathbf{u} = \mathbf{P}_a(\mathbf{x})$  pertence à recta gerada por  $\mathbf{a}$ ,  $\mathbf{u}$  é da forma  $\mathbf{u} = t\mathbf{a}$  para um certo  $t \in \mathbb{R}$ , caracterizado pela condição de que:

$$(\mathbf{x} - t\mathbf{a}) \cdot \mathbf{a} = 0$$

Obtemos então que  $t = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}$  e portanto:

$$\mathbf{P}_a(\mathbf{x}) = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} \quad (4.1.9)$$

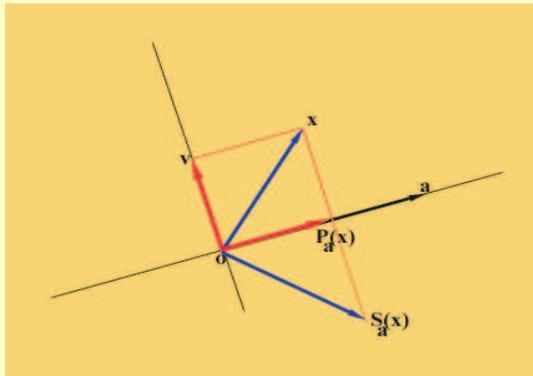
A aplicação  $\mathbf{P}_a : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definida por (4.1.9), é linear e é um projector, isto é,  $\mathbf{P}_a^2 = \mathbf{P}_a$ . Uma vez que  $\mathbf{P}_a(\mathbf{a}) = \mathbf{a}$  vemos que  $\mathbf{a}$  é vector próprio de  $\mathbf{P}_a$ , pertencente ao valor próprio 1. por outro lado, se considerarmos um qualquer vector  $\mathbf{b} \neq \mathbf{0}$  ortogonal a  $\mathbf{a}$  (i.e.:  $\mathbf{a} \cdot \mathbf{b} = 0$ ), vemos que  $\mathbf{P}_a(\mathbf{b}) = \mathbf{0}$  e portanto:

$$\ker \mathbf{P}_a = \text{span}\{\mathbf{b}\} \equiv \{t\mathbf{b} : t \in \mathbb{R}\}$$

A matriz de  $\mathbf{P}_a$  na base  $\{\mathbf{a}, \mathbf{b}\}$  é portanto:

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

#### ► 4.4 Simetria relativamente a uma recta, em $\mathbb{E}^2$ ...



Seja  $\mathbf{a}$  um vector não nulo em  $\mathbb{R}^2$ . A **simetria** relativamente à recta gerada por  $\mathbf{a}$ , é a aplicação linear  $\mathbf{S}_a : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , definida pela condição:

$$\frac{1}{2}(\mathbf{S}_a(\mathbf{x}) + \mathbf{x}) = \mathbf{P}_a(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^2 \quad (4.1.10)$$

isto é, o ponto médio do segmento que une  $\mathbf{x}$  a  $\mathbf{S}_a(\mathbf{x})$  deve ser igual à projecção de  $\mathbf{x}$  sobre a recta gerada por  $\mathbf{a}$ .

Atendendo a (4.1.9), vemos que:

$$\begin{aligned} \mathbf{S}_a(\mathbf{x}) &= 2\mathbf{P}_a(\mathbf{x}) - \mathbf{x} \\ &= 2 \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} - \mathbf{x} \quad \forall \mathbf{x} \in \mathbb{R}^2 \end{aligned} \quad (4.1.11)$$

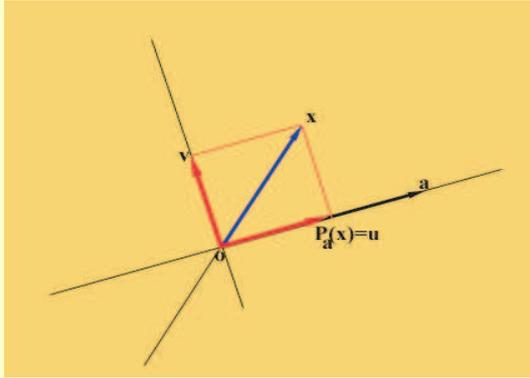
Note que  $\mathbf{S}_a^2 = \text{Id}$ . Uma vez que  $\mathbf{P}_a(\mathbf{a}) = \mathbf{a}$  vemos que  $\mathbf{S}_a \mathbf{a} = \mathbf{a}$ , e portanto  $\mathbf{a}$  é vector próprio de  $\mathbf{S}_a$ , pertencente ao valor próprio 1. Se considerarmos um qualquer vector  $\mathbf{b} \neq \mathbf{0}$  ortogonal a  $\mathbf{a}$  (i.e.:  $\mathbf{a} \cdot \mathbf{b} = 0$ ), vemos que  $\mathbf{P}_a(\mathbf{b}) = \mathbf{0}$  e portanto  $\mathbf{S}_a(\mathbf{b}) = -\mathbf{b}$ .

A matriz de  $\mathbf{S}_a$  na base  $\{\mathbf{a}, \mathbf{b}\}$  é portanto:

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

o que mostra que  $\det \mathbf{S}_a = -1 < 0$ , i.e.,  $\mathbf{S}_a$  inverte orientação (embora preserve o módulo da área de paralelogramos).

► 4.5 Projectção ortogonal sobre uma recta, em  $\mathbb{E}^3$  ...



Sejam  $\mathbf{a} \neq \mathbf{0}$  e  $\mathbf{x}$  dois vectores em  $\mathbb{R}^3$ , com  $\mathbf{a}$  não nulo. Então existe um único vector  $\mathbf{u}$ , na recta gerada por  $\mathbf{a}$ , e um único vector  $\mathbf{v}$ , ortogonal a  $\mathbf{a}$ , tais que  $\mathbf{x} = \mathbf{u} + \mathbf{v}$ . O vector  $\mathbf{u}$ , notado por  $\mathbf{P}_a(\mathbf{x})$ , diz-se a **projectção ortogonal** de  $\mathbf{x}$  sobre a recta gerada por  $\mathbf{a}$ , e é dado por:

$$\mathbf{P}_a(\mathbf{x}) = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} \quad (4.1.12)$$

A aplicação  $\mathbf{P}_a : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  definida por (4.1.12), é linear. Note que  $\mathbf{P}_a^2 = \mathbf{P}_a$ . Uma vez que  $\mathbf{P}_a(\mathbf{a}) = \mathbf{a}$  vemos que  $\mathbf{a}$  é vector próprio de  $\mathbf{P}_a$ , pertencente ao valor próprio 1. Por outro lado, se considerarmos um qualquer vector  $\mathbf{b} \neq \mathbf{0}$  ortogonal a  $\mathbf{a}$  (i.e.:  $\mathbf{a} \cdot \mathbf{b} = 0$ ), vemos que  $\mathbf{P}_a(\mathbf{b}) = \mathbf{0}$  e portanto:

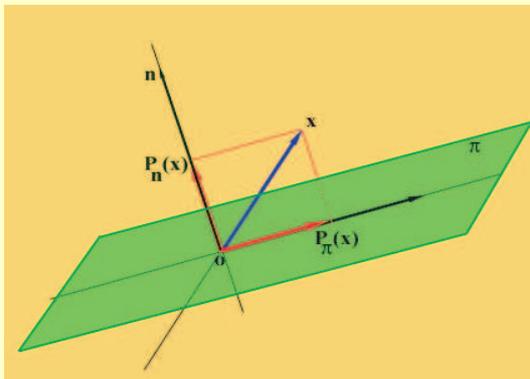
$$\ker \mathbf{P}_a = \text{span}\{\mathbf{b}\} = \{\mathbf{b} \in \mathbb{R}^3 : \mathbf{b} \cdot \mathbf{a} = 0\} = \mathbf{a}^\perp$$

é o plano vectorial ortogonal a  $\mathbf{a}$ .

A matriz de  $\mathbf{P}_a$  numa base  $\{\mathbf{a}, \mathbf{b}_1, \mathbf{b}_2\}$ , onde  $\mathbf{b}_1, \mathbf{b}_2$  geram o  $\ker \mathbf{P}_a$ , é portanto:

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

► 4.6 Projectção ortogonal sobre um plano vectorial, em  $\mathbb{E}^3$  ...



Consideremos um plano vectorial ortogonal a um vector  $\mathbf{n} \in \mathbb{R}^3 - \{\mathbf{0}\}$  (se esse plano é gerado por dois vectores  $\mathbf{u}, \mathbf{v}$  linearmente independentes, podemos tomar  $\mathbf{n} = \mathbf{u} \times \mathbf{v}$ ). Notemos esse plano por  $\pi = \mathbf{n}^\perp$ . Dado um vector  $\mathbf{x} \in \mathbb{R}^3$ , ao vector:

$$\mathbf{P}_{\mathbf{n}^\perp} \equiv \mathbf{x} - \mathbf{P}_n(\mathbf{x})$$

chamamos a **projectção ortogonal** de  $\mathbf{x}$  sobre o plano vectorial ortogonal a  $\mathbf{n}$ .

De acordo com (4.1.12), temos que:

$$\begin{aligned} \mathbf{P}_{\mathbf{n}^\perp} &\equiv \mathbf{x} - \mathbf{P}_n(\mathbf{x}) \\ &= \mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n} \end{aligned} \quad (4.1.13)$$

A aplicação  $\mathbf{P}_{\mathbf{n}^\perp} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  definida por (4.1.13), é linear. Note que  $\mathbf{P}_{\mathbf{n}^\perp}^2 = \mathbf{P}_{\mathbf{n}^\perp}$ . Se  $\mathbf{x} \cdot \mathbf{n} = 0$ , i.e., se  $\mathbf{x}$  é ortogonal a  $\mathbf{n}$ , então  $\mathbf{P}_{\mathbf{n}^\perp}(\mathbf{x}) = \mathbf{x}$ , enquanto que, por outro lado,  $\mathbf{P}_{\mathbf{n}^\perp}(\mathbf{n}) = \mathbf{0}$ . Portanto vemos que:

$$\ker \mathbf{P}_{\mathbf{n}^\perp} = \text{span}\{\mathbf{n}\}$$

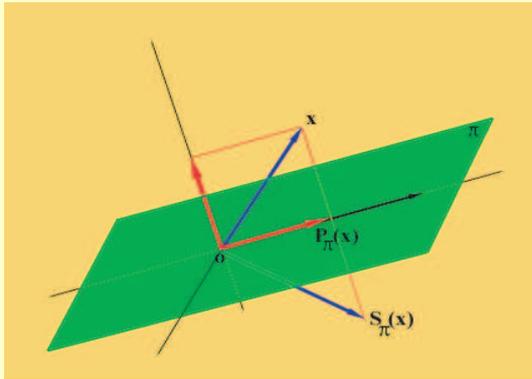
e:

$$\mathbf{P}_{\mathbf{n}^\perp}(\mathbf{x}) = \mathbf{x} \quad \forall \mathbf{x} \in \mathbf{n}^\perp$$

Portanto a matriz de  $\mathbf{P}_{\mathbf{n}^\perp}$  numa base  $\{\mathbf{n}, \mathbf{b}_1, \mathbf{b}_2\}$ , onde  $\mathbf{b}_1, \mathbf{b}_2$  geram o plano  $\mathbf{n}^\perp$ , é:

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

#### ► 4.7 Simetria relativamente a um plano vectorial ...



Consideremos novamente um plano vectorial  $\mathbf{n}^\perp$ , ortogonal a um vector  $\mathbf{n} \in \mathbb{R}^3 - \{\mathbf{0}\}$  (se esse plano é gerado por dois vectores  $\mathbf{u}, \mathbf{v}$  linearmente independentes, podemos tomar  $\mathbf{n} = \mathbf{u} \times \mathbf{v}$ ).

A **simetria** relativamente ao plano vectorial  $\mathbf{n}^\perp$ , é a aplicação linear  $\mathbf{S}_{\mathbf{n}^\perp} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , definida pela condição:

$$\frac{1}{2}(\mathbf{S}_{\mathbf{n}^\perp}(\mathbf{x}) + \mathbf{x}) = \mathbf{P}_{\mathbf{n}^\perp}(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^3 \quad (4.1.14)$$

isto é, o ponto médio do segmento que une  $\mathbf{x}$  a  $\mathbf{S}_{\mathbf{n}^\perp}(\mathbf{x})$  deve ser igual à projecção de  $\mathbf{x}$  sobre o plano vectorial  $\mathbf{n}^\perp$ .

Atendendo a (4.1.13), vemos que:

$$\begin{aligned} \mathbf{S}_{\mathbf{n}^\perp}(\mathbf{x}) &= 2\mathbf{P}_{\mathbf{n}^\perp}(\mathbf{x}) - \mathbf{x} \\ &= 2\left(\mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n}\right) - \mathbf{x} \\ &= \mathbf{x} - 2 \frac{\mathbf{x} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n} \quad \forall \mathbf{x} \in \mathbb{R}^3 \end{aligned} \quad (4.1.15)$$

Note que  $\mathbf{S}_{\mathbf{n}^\perp}^2 = \text{Id}$ . Além disso, é fácil ver que :

$$\mathbf{S}_{\mathbf{n}^\perp}(\mathbf{n}) = -\mathbf{n}$$

o que significa que  $\mathbf{n}$  é vector próprio de  $\mathbf{S}_{\mathbf{n}^\perp}$ , pertencente ao valor próprio  $-1$ , e ainda que:

$$\mathbf{P}_{\mathbf{n}^\perp}(\mathbf{x}) = \mathbf{x} \quad \forall \mathbf{x} \in \mathbf{n}^\perp$$

Portanto a matriz de  $\mathbf{S}_{\mathbf{n}^\perp}$  numa base  $\{\mathbf{n}, \mathbf{b}_1, \mathbf{b}_2\}$ , onde  $\mathbf{b}_1, \mathbf{b}_2$  geram o plano  $\mathbf{n}^\perp$ , é:

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

o que mostra que  $\det \mathbf{S}_{\mathbf{n}^\perp} = -1 < 0$ , i.e.,  $\mathbf{S}_{\mathbf{n}^\perp}$  inverte orientação.

► **4.8 Projecção ortogonal sobre uma recta, em  $\mathbb{C}^2$  ...** Sejam  $\mathbf{w} \neq \mathbf{0}$  e  $\mathbf{z}$  dois vectores em  $\mathbb{C}^2$ . Então existe um único vector  $\mathbf{u}$ , na recta complexa gerada por  $\mathbf{w}$ , e um único vector  $\mathbf{v}$ , ortogonal a  $\mathbf{w}$ , tais que  $\mathbf{z} = \mathbf{u} + \mathbf{v}$ . O vector  $\mathbf{u}$ , notado por  $\mathbf{P}_{\mathbf{w}}(\mathbf{z})$ , diz-se a **projecção ortogonal** de  $\mathbf{z}$  sobre a recta complexa gerada por  $\mathbf{w}$ , e é calculado da seguinte forma: uma vez que  $\mathbf{u} = \mathbf{P}_{\mathbf{w}}(\mathbf{z})$  pertence à recta gerada por  $\mathbf{w}$ ,  $\mathbf{u}$  é da forma  $\mathbf{u} = \alpha\mathbf{w}$  para um certo  $\alpha \in \mathbb{C}$ , caracterizado pela condição de que:

$$\langle \mathbf{z} - \alpha\mathbf{w} | \mathbf{w} \rangle = 0$$

Obtemos que  $\alpha = \frac{\langle \mathbf{z} | \mathbf{w} \rangle}{\|\mathbf{w}\|^2}$  e portanto:

$$\mathbf{P}_{\mathbf{w}}(\mathbf{z}) = \frac{\langle \mathbf{z} | \mathbf{w} \rangle}{\|\mathbf{w}\|^2} \mathbf{w} \quad (4.1.16)$$

► **4.9 Simetria relativamente a uma recta complexa ...** Seja  $\mathbf{w}$  um vector não nulo em  $\mathbb{C}^2$ . A **simetria** relativamente à recta complexa gerada por  $\mathbf{w}$ , é a aplicação linear  $\mathbf{S}_{\mathbf{w}} : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ , definida pela condição:

$$\frac{1}{2}(\mathbf{S}_{\mathbf{w}}(\mathbf{z}) + \mathbf{z}) = \mathbf{P}_{\mathbf{w}}(\mathbf{z}) \quad \forall \mathbf{z} \in \mathbb{C}^2 \quad (4.1.17)$$

isto é, o ponto médio do segmento que une  $\mathbf{z}$  a  $\mathbf{S}_{\mathbf{w}}(\mathbf{z})$  deve ser igual à projecção de  $\mathbf{z}$  sobre a recta gerada por  $\mathbf{w}$ . Atendendo a (4.1.16), vemos que:

$$\begin{aligned} \mathbf{S}_{\mathbf{w}}(\mathbf{z}) &= 2\mathbf{P}_{\mathbf{w}}(\mathbf{z}) - \mathbf{z} \\ &= 2 \frac{\langle \mathbf{z} | \mathbf{w} \rangle}{\|\mathbf{w}\|^2} \mathbf{w} - \mathbf{z} \quad \forall \mathbf{z} \in \mathbb{C}^2 \end{aligned} \quad (4.1.18)$$

## 4.2 Isometrias em $\mathbb{R}^2$ . Os grupos $\mathcal{O}(2)$ e $\mathcal{SO}(2)$

► **4.10** Como já vimos, uma aplicação linear  $\mathbf{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  diz-se uma **transformação ortogonal** ou uma **isometria** de  $\mathbb{R}^2$ , se  $\mathbf{A}$  preserva o produto interno (Euclideano) usual de  $\mathbb{R}^2$ , i.e.:

$$\mathbf{A}(\mathbf{x}) \cdot \mathbf{A}(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y} \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^2 \quad (4.2.1)$$

Esta condição é equivalente a:

$$\|\mathbf{A}(\mathbf{x})\| = \|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^2 \quad (4.2.2)$$

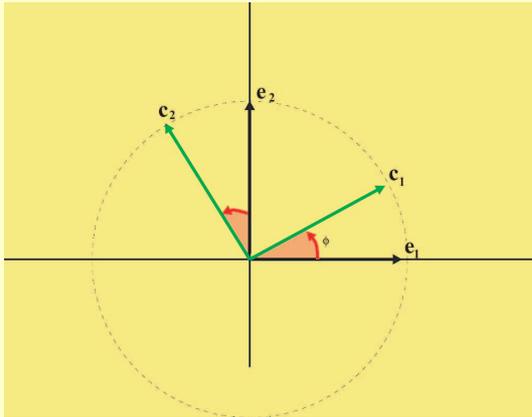
i.e.,  $\mathbf{A}$  preserva os comprimentos dos vectores. Se  $A$  é a matriz de uma tal transformação ortogonal, relativamente a uma qualquer base ortonormada  $\{\mathbf{e}_1, \mathbf{e}_2\}$  de  $\mathbb{R}^2$  (por exemplo, a base canónica),  $A$  é uma matriz ortogonal, isto é,  $A^t A = \mathbf{I}$ . Portanto  $A \in \mathcal{O}(2)$ . Vejamos como é a forma geral de uma tal matriz.

► **4.11** Se  $\mathbf{c}_1 = \mathbf{A}(\mathbf{e}_1)$ ,  $\mathbf{c}_2 = \mathbf{A}(\mathbf{e}_2)$  são as colunas de  $A$ , então:

$$\mathbf{c}_i \cdot \mathbf{c}_j = \delta_{ij}$$

o que significa que  $\mathbf{c}_1$  e  $\mathbf{c}_2$  são ortonormais. Portanto  $\mathbf{A}$  transforma bases ortonormadas em bases ortonormadas, preservando ou invertendo orientação, conforme  $\det \mathbf{A} = +1$  ou  $\det \mathbf{A} = -1$ , respectivamente. Por exemplo, a simetria  $\mathbf{S}_{\mathbf{a}}$ , descrita em (4.1.17), é uma transformação ortogonal com  $\det$  igual a  $-1$ .

## ► 4.12



Como  $\mathbf{c}_1 = \mathbf{A}(\mathbf{e}_1) \equiv \begin{bmatrix} a \\ b \end{bmatrix}$  é um vector de norma 1, sabemos que  $a^2 + b^2 = 1$  e portanto existe um único  $\varphi \in [0, 2\pi[$  tal que  $a = \cos \varphi$  e  $b = \sin \varphi$  ( $\varphi \in [0, 2\pi[$  é o ângulo polar de  $\mathbf{c}_1$ , i.e., o ângulo orientado que  $\mathbf{c}_1$  faz com a parte positiva do eixo dos  $xx$ ):

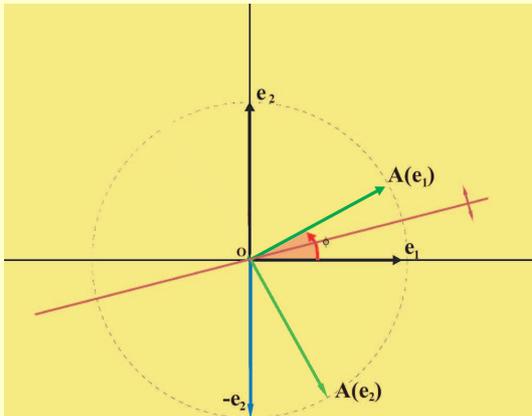
Portanto  $\mathbf{c}_1 = \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix}$ , e como  $\mathbf{c}_2 = \mathbf{A}(\mathbf{e}_2)$  é também um vector unitário e ortogonal a  $\mathbf{c}_1$ , dois casos podem ocorrer:

$$(i). \quad \mathbf{c}_2 = \begin{bmatrix} -\sin \varphi \\ \cos \varphi \end{bmatrix}, \quad \text{ou} \quad (ii). \quad \mathbf{c}_2 = \begin{bmatrix} \sin \varphi \\ -\cos \varphi \end{bmatrix}$$

No primeiro caso, a matriz  $A$  tem a forma:

$$A = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \quad (4.2.3)$$

cujo determinante é 1. Neste caso  $\mathbf{A}$  diz-se uma **rotação de ângulo**  $\varphi$  (no sentido positivo), em torno da origem, e nota-se por  $\mathbf{R}_\varphi$ :



No segundo caso, a matriz  $A$  tem a forma:

$$\begin{aligned} A &= \begin{bmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{bmatrix} \\ &= \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ &= \mathbf{R}_\varphi \mathbf{S}_{\mathbf{e}_1} \end{aligned} \quad (4.2.4)$$

cujo determinante é  $-1$ . Neste caso  $\mathbf{A}$  pode ser interpretada como uma reflexão relativamente ao eixo dos  $xx$  seguida de uma rotação  $\mathbf{R}_\varphi$ .

Essa reflexão fixa  $\mathbf{e}_1$  e transforma  $\mathbf{e}_2$  em  $-\mathbf{e}_2$ . Se então rodamos de ângulo  $\varphi$ , temos que:

$$\begin{aligned} \mathbf{e}_1 &\rightarrow \mathbf{e}_1 \rightarrow \cos \varphi \mathbf{e}_1 + \sin \varphi \mathbf{e}_2 \\ \mathbf{e}_2 &\rightarrow -\mathbf{e}_2 \rightarrow -(-\sin \varphi \mathbf{e}_1 + \cos \varphi \mathbf{e}_2) \end{aligned} \quad (4.2.5)$$

De facto, neste caso  $\mathbf{A}$  representa uma simetria relativamente à recta que faz um ângulo  $\frac{\varphi}{2}$  com a parte positiva do eixo dos  $xx$ :

### 4.3 Isometrias em $\mathbb{R}^3$ . Rotações. Ângulos de Euler. Os grupos $\mathcal{O}(3)$ e $\mathcal{SO}(3)$

► **4.13** Como já vimos, uma aplicação linear  $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  diz-se uma **transformação ortogonal** ou uma **isometria** de  $\mathbb{R}^3$ , se  $\mathbf{A}$  preserva o produto interno (Euclideano) usual de  $\mathbb{R}^3$ , i.e.:

$$\mathbf{A}(\mathbf{x}) \cdot \mathbf{A}(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y} \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^3 \quad (4.3.1)$$

Esta condição é equivalente a:

$$\|\mathbf{A}(\mathbf{x})\| = \|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^3 \quad (4.3.2)$$

i.e.,  $\mathbf{A}$  preserva os comprimentos dos vectores. Se  $A$  é a matriz de uma tal transformação ortogonal, relativamente a uma qualquer base ortonormada  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  de  $\mathbb{R}^3$  (por exemplo, a base canónica),  $A$  é uma matriz ortogonal, isto é,  $A^t A = \mathbf{I}$ . Portanto  $A \in \mathcal{O}(3)$ . Vejamos como é a forma geral de uma tal matriz.

► **4.14** Se  $\mathbf{c}_1 = \mathbf{A}(\mathbf{e}_1)$ ,  $\mathbf{c}_2 = \mathbf{A}(\mathbf{e}_2)$ ,  $\mathbf{c}_3 = \mathbf{A}(\mathbf{e}_3)$  são as colunas de  $A$ , então:

$$\mathbf{c}_i \cdot \mathbf{c}_j = \delta_{ij}$$

o que significa que  $\mathbf{c}_1, \mathbf{c}_2$  e  $\mathbf{c}_3$  são ortonormais. Portanto  $\mathbf{A}$  transforma bases ortonormadas em bases ortonormadas, preservando ou invertendo orientação, conforme  $\det \mathbf{A} = +1$  ou  $\det \mathbf{A} = -1$ , respectivamente. Por exemplo, a simetria  $\mathbf{S}_{\mathbf{n}^\perp}$ , descrita em (4.1.14), é uma transformação ortogonal com  $\det$  igual a  $-1$ .

► **4.15** Como já vimos  $\mathbf{A}$  admite sempre um valor próprio real. De facto, se  $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  é uma isometria então esse valor próprio (real) ou é 1 ou  $-1$ . Com efeito, se  $a \in \mathbb{R}$  é valor próprio de  $\mathbf{A}$ , e  $\mathbf{x}$  é um vector próprio pertencente a  $a$ , temos que:

$$\|\mathbf{x}\| = \|\mathbf{A}(\mathbf{x})\| = \|a\mathbf{x}\| = |a|\|\mathbf{x}\|$$

o que implica que  $|a| = 1$  (uma vez que  $\mathbf{x} \neq \mathbf{0}$ ), i.e.,  $a = \pm 1$ .

Analisemos agora a estrutura das isometrias de  $\mathbb{R}^3$  com determinante igual a 1, isto é, a estrutura das matrizes  $A \in \mathcal{SO}(3)$ . Seja  $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  uma tal isometria, com:

$$\det \mathbf{A} = 1$$

Pelo parágrafo anterior,  $\mathbf{A}$  admite o valor próprio 1 ou  $-1$ . Vamos analisar cada um destes casos:

(i).  $a = 1$  é valor próprio de  $\mathbf{A}$  (e  $\det \mathbf{A} = 1$ ) ... Seja  $\mathbf{u} \neq \mathbf{0}$  um vector próprio de  $\mathbf{A}$ , pertencente ao valor próprio 1:

$$\mathbf{A}(\mathbf{u}) = \mathbf{u}$$

Podemos supôr também que  $\|\mathbf{u}\| = 1$ . Se  $\Pi = \mathbf{u}^\perp$  é o plano ortogonal a  $\mathbf{u}$ , é fácil ver que  $\mathbf{A}$  deixa  $\Pi$  invariante:

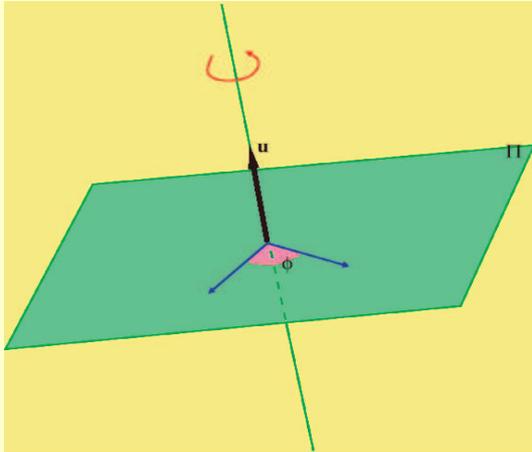
$$\mathbf{A}(\Pi) \subseteq \Pi$$

e que a restrição de  $\mathbf{A}$  a  $\Pi$  é uma isometria de  $\Pi$ . Portanto existe uma base ortonormada  $\{\mathbf{e}, \mathbf{f}\}$  de  $\Pi$ , relativamente à qual a matriz da restrição de  $\mathbf{A}$  a  $\Pi$ , é de um dos seguintes dois tipos:

$$(i1). \quad \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \quad (4.3.3)$$

ou:

$$(i2). \quad \begin{bmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{bmatrix} \quad (4.3.4)$$



A matriz de  $\mathbf{A}$ , relativamente à base ortonormada  $\{\mathbf{u}, \mathbf{e}, \mathbf{f}\}$  de  $\mathbb{R}^3$  é portanto no caso (i1):

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{bmatrix} \quad (4.3.5)$$

que tem de facto determinante 1, e representa uma rotação em torno da recta gerada por  $\mathbf{u} \in \Pi$  (que se diz o **eixo da rotação**), de ângulo  $\varphi$ .

Por outro lado, no caso (i2), a matriz de  $\mathbf{A}$ , relativamente à base ortonormada  $\{\mathbf{u}, \mathbf{e}, \mathbf{f}\}$  de  $\mathbb{R}^3$ , é:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & \sin \varphi \\ 0 & \sin \varphi & -\cos \varphi \end{bmatrix} \quad (4.3.6)$$

que tem determinante  $-1$  e por isso não pode ser a matriz de  $\mathbf{A}$ .

**(i).**  $a = -1$  é valor próprio de  $\mathbf{A}$  (e  $\det \mathbf{A} = 1$ ) ... Seja  $\mathbf{u} \neq \mathbf{0}$  um vector próprio de  $\mathbf{A}$ , pertencente ao valor próprio  $-1$ :

$$\mathbf{A}(\mathbf{u}) = -\mathbf{u}$$

Podemos supôr também que  $\|\mathbf{u}\| = 1$ .

Mais uma vez, se  $\Pi = \mathbf{u}^\perp$  é o plano ortogonal a  $\mathbf{u}$ ,  $\mathbf{A}$  deixa  $\Pi$  invariante:

$$\mathbf{A}(\Pi) \subseteq \Pi$$

e a restrição de  $\mathbf{A}$  a  $\Pi$  é uma isometria de  $\Pi$ . Portanto existe uma base ortonormada  $\{\mathbf{e}, \mathbf{f}\}$  de  $\Pi$ , relativamente à qual a matriz da restrição de  $\mathbf{A}$  a  $\Pi$ , é de um dos seguintes dois tipos:

$$(ii1). \quad \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \quad (4.3.7)$$

ou:

$$(ii2). \quad \begin{bmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{bmatrix} \quad (4.3.8)$$

Como vimos anteriormente, esta é uma matriz de uma simetria relativamente a uma recta no plano  $\Pi$ , e portanto podemos escolher uma base ortonormada  $\{\mathbf{e}', \mathbf{f}'\}$  para  $\Pi$ , relativamente à qual a matriz dessa simetria é:

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

A matriz de  $\mathbf{A}$ , relativamente à base ortonormada  $\{\mathbf{u}, \mathbf{e}, \mathbf{f}\}$  de  $\mathbb{R}^3$  é portanto no caso (ii 1):

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{bmatrix} \quad (4.3.9)$$

que tem determinante  $-1$ , e por isso não pode ser a matriz de  $\mathbf{A}$ .

Finalmente no caso (ii 2), a matriz de  $\mathbf{A}$ , relativamente à base ortonormada  $\{\mathbf{u}, \mathbf{e}', \mathbf{f}'\}$  de  $\mathbb{R}^3$ , é:

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (4.3.10)$$

que tem determinante  $1$ , e representa uma rotação em torno da recta gerada por  $\mathbf{e}' \in \Pi$ , de ângulo  $\pi$ .

► **4.16 Resumindo** ... Uma isometria  $\mathbf{A}$  em  $\mathbb{R}^3$ , com  $\det \mathbf{A} = 1$ , é sempre uma rotação em torno de uma certa recta  $\mathbb{R}\{\mathbf{u}\}$  (o eixo de rotação), e de ângulo  $\varphi$  no sentido directo. Representamos uma tal rotação por  $\mathbf{R}_{(\mathbf{u};\varphi)}$ . As matrizes das rotações em torno dos eixos coordenados de  $\mathbb{R}^3$ , e de ângulo  $\varphi$  no sentido directo, são respectivamente:

$$\mathbf{R}_1(\varphi) = \mathbf{R}_{(\mathbf{e}_1;\varphi)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{bmatrix} \quad (4.3.11)$$

$$\mathbf{R}_2(\varphi) = \mathbf{R}_{(\mathbf{e}_2;\varphi)} = \begin{bmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{bmatrix} \quad (4.3.12)$$

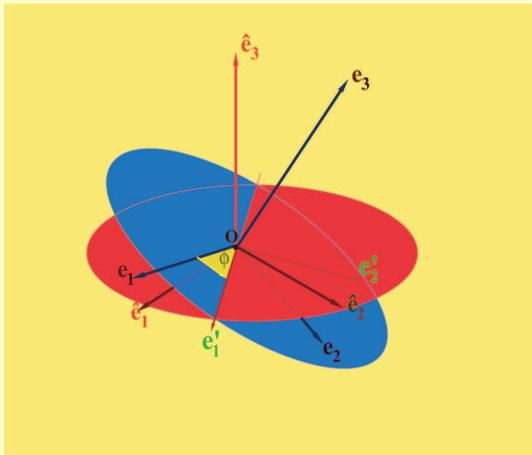
$$\mathbf{R}_3(\varphi) = \mathbf{R}_{(\mathbf{e}_3;\varphi)} = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.3.13)$$

► **4.17 Ângulos de Euler** ... Qualquer rotação pode ser escrita como um produto de rotações dos tipos acima indicados.

Com efeito consideremos uma qualquer rotação  $\mathbf{R} \in \mathcal{SO}(3)$  e duas bases ortonormadas de  $\mathbb{R}^3$ :

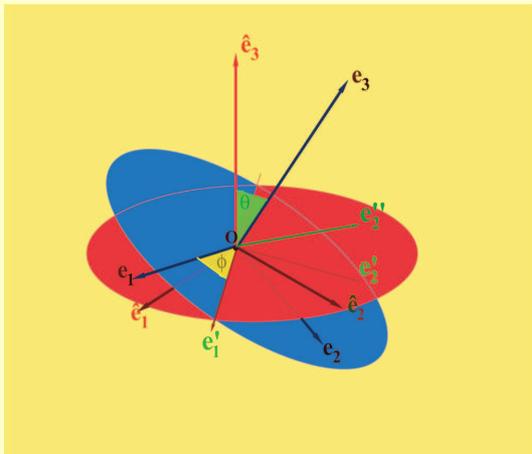
$$\begin{aligned} \mathcal{B} &= \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \\ \widehat{\mathcal{B}} = \mathcal{B}\mathbf{R} &= \{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2, \widehat{\mathbf{e}}_3\} \end{aligned} \quad (4.3.14)$$

com a mesma orientação. A base  $\widehat{\mathcal{B}} = \mathcal{B}\mathbf{R}$  pode ser obtida através das seguintes três fases sucessivas:



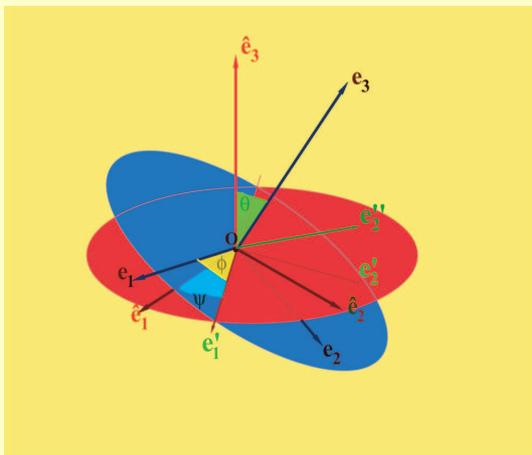
1. Obter uma base ortonormada  $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3 = \mathbf{e}_3\}$ , através de uma rotação  $\mathbf{R}_3(\phi)$ , em torno de  $\mathbf{e}_3$  e de ângulo  $\phi$ , onde  $\phi$  é o ângulo entre  $\mathbf{e}_1$  e a chamada **linha dos nodos** (a recta de intersecção dos planos gerados respectivamente por  $\{\mathbf{e}_1, \mathbf{e}_2\}$  e  $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$ ):

$$\mathcal{B}' = \mathcal{B} \mathbf{R}_3(\phi) \quad (4.3.15)$$



2. Obter uma base ortonormada  $\mathcal{B}'' = \{\mathbf{e}'_1, \mathbf{e}''_2, \hat{\mathbf{e}}_3\}$ , através de uma rotação  $\mathbf{R}_2(\theta)$ , em torno da linha dos nodos, gerada por  $\mathbf{e}'_1$ , e de ângulo  $\theta$ , onde  $\theta$  é o ângulo entre  $\mathbf{e}_3$  e  $\hat{\mathbf{e}}_3$ :

$$\mathcal{B}'' = \mathcal{B}' \mathbf{R}_2(\theta) \quad (4.3.16)$$



3. Finalmente, obter a base ortonormada  $\hat{\mathcal{B}} = \mathcal{B} \mathbf{R} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ , através de uma rotação  $\mathbf{R}_3(\psi)$ , em torno de  $\hat{\mathbf{e}}_3$ , e de ângulo  $\psi$ , onde  $\psi$  é o ângulo entre a linha dos nodos e  $\hat{\mathbf{e}}_1$ :

$$\hat{\mathcal{B}} = \mathcal{B}'' \mathbf{R}_3(\psi) \quad (4.3.17)$$

► 4.18 Portanto:

$$\begin{aligned} \hat{\mathcal{B}} &= \mathcal{B} \mathbf{R} \\ &= \mathcal{B} \mathbf{R}_3(\phi) \mathbf{R}_2(\theta) \mathbf{R}_3(\psi) \end{aligned} \quad (4.3.18)$$

e:

$$\begin{aligned} \mathbf{R} &= \mathbf{R}_3(\phi)\mathbf{R}_2(\theta)\mathbf{R}_3(\varphi) \\ &= \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (4.3.19)$$

Os ângulos  $\phi, \theta, \psi$  chamam-se **ângulos de Euler**.

#### 4.4 Transformações unitárias em $\mathbb{C}^2$ . Os grupos $\mathcal{U}(2)$ e $\mathcal{SU}(2)$

► **4.19** Uma aplicação linear  $\mathbf{A} : \mathbb{C}^2 \rightarrow \mathbb{C}^2$  diz-se uma **transformação unitária** de  $\mathbb{C}^2$ , se  $\mathbf{A}$  preserva o produto interno hermitiano usual de  $\mathbb{C}^2$ , i.e.:

$$\langle \mathbf{A}(\mathbf{z}) | \mathbf{A}(\mathbf{w}) \rangle = \langle \mathbf{z} | \mathbf{w} \rangle \quad \forall \mathbf{z}, \mathbf{w} \in \mathbb{C}^2 \quad (4.4.1)$$

Se  $A$  é a matriz de uma tal transformação unitária, relativamente à base canónica de  $\mathbb{C}^2$ , então (4.4.1) escreve-se na seguinte forma matricial:

$$(A\mathbf{z})^t \overline{A\mathbf{w}} = \mathbf{z}^t \overline{\mathbf{w}} \quad \forall \mathbf{z}, \mathbf{w} \in \mathbb{C}^2$$

ou ainda:

$$\mathbf{z}^t A^t \overline{A\mathbf{w}} = \mathbf{z}^t \overline{\mathbf{w}} = \mathbf{z}^t \mathbf{I} \overline{\mathbf{w}} \quad \forall \mathbf{z}, \mathbf{w} \in \mathbb{C}^2$$

o que significa que a matriz  $A$  é uma **matriz unitária**, i.e.:

$$A^t \overline{A} = \mathbf{I} \quad (4.4.2)$$

Recordemos que, dada uma matriz  $A$ , define-se a respectiva **matriz adjunta**  $A^\dagger$ , como sendo a conjugada transposta de  $A$ :

$$A^\dagger = \overline{A}^t$$

Portanto  $A$  é unitária sse:

$$AA^\dagger = \mathbf{I} \quad (4.4.3)$$

Note ainda que, uma vez que  $\det(AA^\dagger) = \det(A\overline{A}^t) = \det A \det(\overline{A}^t) = \det A \overline{\det A} = |\det A|$ , concluímos que, se  $A$  é unitária, então  $|\det A| = 1$  e, em particular  $A$  é inversível com  $A^{-1} = A^\dagger$ .

► **4.20** O subgrupo de  $\mathcal{U}(2)$  constituído por todas as transformações unitárias de  $\mathbb{C}^2$ , que têm determinante 1 diz-se o **grupo unitário especial** e nota-se por  $\mathcal{SU}(2)$ . Este grupo é isomorfo ao grupo das matrizes unitárias de determinante 1, também notado por  $\mathcal{SU}(2)$ .

Suponhamos que  $A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$  é uma matriz em  $\mathcal{SU}(2)$ , de tal forma que  $A^{-1} = A^\dagger$  e  $\det A = \alpha\delta - \beta\gamma = 1$ . Temos então que:

$$A^{-1} = \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} = A^\dagger = \begin{bmatrix} \overline{\alpha} & \overline{\gamma} \\ \overline{\beta} & \overline{\delta} \end{bmatrix}$$

isto é:  $\delta = \overline{\alpha}$  e  $\gamma = -\overline{\beta}$ . Portanto  $\mathcal{SU}(2)$  é o grupo das matrizes que são da forma:

$$A = \begin{bmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{bmatrix} \quad e \quad \det A = |\alpha|^2 + |\beta|^2 = 1 \quad (4.4.4)$$

## 4.5 Exercícios

▷ **Exercício 4.1** ... Classifique as seguintes isometrias em  $\mathbb{R}^2$  :

a)  $f(x, y) = (\frac{1}{2}x + \frac{\sqrt{3}}{2}y, \frac{\sqrt{3}}{2}x - \frac{1}{2}y)$ .

b)  $f(x, y) = (\frac{1}{2}x + \frac{\sqrt{3}}{2}y, -\frac{\sqrt{3}}{2}x + \frac{1}{2}y)$ .

c)  $f(x, y) = (-\frac{4}{5}x + \frac{3}{5}y, -\frac{3}{5}x - \frac{4}{5}y)$ .

d)  $f(x, y) = (x, y)$ .

e)  $f(x, y) = (-y, x)$ .

▷ **Exercício 4.2** ... Em cada um dos casos que se seguem, determine  $\mathbf{S}_r(x, y)$ ,  ${}_b M_b(\mathbf{S}_r)$  e uma base  $b$  de  $\mathbb{R}^2$  tal que  ${}_b M_b(\mathbf{S}_r) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ .

a)  $r$  é a recta de equação  $y = 2x$ ;

b)  $r$  é a recta de equação  $3x - y = 0$ ;

c)  $r$  é a recta de equação  $y = (tg \frac{\pi}{5})x$ ;

▷ **Exercício 4.3** ... Em cada um dos seguintes casos, mostre que o endomorfismo  $f$  de  $\mathbb{R}^2$  ou  $\mathbb{R}^3$  é uma isometria linear e descreva  $f$  geometricamente (isto é, diga se  $f$  é uma simetria ou uma rotação; no caso de ser uma simetria, diga relativamente a que recta, no caso de ser uma rotação determine o ângulo).

a)  $f(x, y) = (y, x)$ ;

b)  $f(x, y) = (y, -x)$ ;

c)  $f(x, y) = (\frac{\sqrt{2}x - \sqrt{2}y}{2}, \frac{\sqrt{2}x + \sqrt{2}y}{2})$ ;

d)  $f(x, y) = ((-\cos \frac{\pi}{8})x + (\sin \frac{\pi}{8})y, (\sin \frac{\pi}{8})x + (\cos \frac{\pi}{8})y)$ ;

▷ **Exercício 4.4** ... Dado:

a)  $\mathbf{a} = (1, 4, -3)$ , calcule  $\mathbf{P}_a(\mathbf{x})$  sendo  $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$ . Calcule  $\ker \mathbf{P}_a$ . Defina  $\mathbf{S}_a(\mathbf{x})$ .

b)  $\mathbf{a} = (0, 1, 2)$ , calcule  $\mathbf{P}_a(\mathbf{x})$  sendo  $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$ . Calcule  $\ker \mathbf{P}_a$ . Defina  $\mathbf{S}_a(\mathbf{x})$ .

c)  $\mathbf{a} = (1, 1, 1)$ , calcule  $\mathbf{P}_a(\mathbf{x})$  sendo  $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$ . Calcule  $\ker \mathbf{P}_a$ . Defina  $\mathbf{S}_a(\mathbf{x})$ .

d)  $\mathbf{a} = (1, 1)$ , calcule  $\mathbf{P}_a(\mathbf{x})$  sendo  $\mathbf{x} = (x, y) \in \mathbb{R}^2$ . Calcule  $\ker \mathbf{P}_a$ . Defina  $\mathbf{S}_a(\mathbf{x})$ .

e)  $\mathbf{a} = (1, 0)$ , calcule  $\mathbf{P}_a(\mathbf{x})$  sendo  $\mathbf{x} = (x, y) \in \mathbb{R}^2$ . Calcule  $\ker \mathbf{P}_a$ . Defina  $\mathbf{S}_a(\mathbf{x})$ .

▷ **Exercício 4.5** ... Defina a simetria relativamente à recta  $2x - y = 0$  em  $\mathbb{R}^2$ .

▷ **Exercício 4.6** ... Em cada uma das alíneas que se seguem, calcule  $\mathbf{P}_\pi(\mathbf{x})$  e  $\ker \mathbf{P}_\pi$ , em  $\mathbb{R}^3$  sendo  $\pi$  cada um dos planos que se seguem. Calcule também em cada caso, os valores próprios e os vectores próprios de  $\mathbf{P}_\pi$ . Finalmente, defina Defina  $\mathbf{S}_\pi(\mathbf{x})$ .

a)  $2x - y + 3z = 0$ ;

b)  $x + y + z = 0$ ;

c)  $3x + y + 2z = 0$ .

▷ **Exercício 4.7** ... As matrizes que se seguem, representam rotações em  $\mathbb{R}^3$  relativamente à base canónica. Mostre que são matrizes ortogonais de determinante igual a 1. Calcule o eixo e o ângulo de rotação:

$$\text{a) } A = \begin{bmatrix} 0 & 1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{bmatrix}; \quad \text{b) } A = \begin{bmatrix} 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}; \quad \text{c) } A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

## Capítulo 5

# Operadores auto-adjuntos (simétricos e hermitianos). Teorema espectral

### 5.1 Operadores auto-adjuntos (simétricos e hermitianos)

► **5.1** Como já vimos numa secção anterior, se  $\mathbf{L} : \mathcal{V} \rightarrow \mathcal{V}$  é um operador linear num espaço vectorial de dimensão finita, então a representação matricial de  $\mathbf{L}$  varia com a escolha da base numa classe de conjugação de matrizes:

$$\mathcal{C} \rightarrow \mathcal{C}P \quad \Rightarrow \quad [\mathbf{L}]_{\mathcal{C}} \rightarrow [\mathbf{L}]_{\mathcal{C}P} = P^{-1} [\mathbf{L}]_{\mathcal{C}} P \quad (5.1.1)$$

Esta possibilidade de variar a representação matricial de  $\mathbf{L}$ , variando a base, conduz-nos naturalmente ao seguinte problema:

**Como escolher a base de  $\mathcal{V}$  de tal forma que a representação matricial de  $\mathbf{L}$  seja o mais “simples” possível? Mais formalmente - se  $[\mathbf{L}]_{\mathcal{C}}$  é a representação matricial de  $\mathbf{L}$  numa certa base  $\mathcal{C}$ , como seleccionar na classe de conjugação de  $L$ :**

$$\{[\mathbf{L}]_{\mathcal{C}P} = P^{-1} [\mathbf{L}]_{\mathcal{C}} P : P \in Gl(n)\}$$

**o representante mais “simples” possível ?**

► **5.2** Suponhamos agora que  $\mathcal{V}$  é um espaço vectorial com um produto interno  $\langle | \rangle$  (como sempre, Euclideano se  $\mathcal{V}$  é real, ou Hermitiano, se  $\mathcal{V}$  for complexo). É claro que nestes espaços, a classe de todas as bases ortonormadas desempenha um papel central.

► **5.3** Suponhamos que  $\mathcal{C}$  e  $\widehat{\mathcal{C}} = \mathcal{C}P$  são duas bases ortonormadas em  $\mathcal{V}$ . Então a matriz  $P$  é:

- uma matriz ortogonal,  $P \in \mathcal{O}(n)$ , se  $\mathcal{V}$  é Euclideano.
- uma matriz unitária,  $P \in \mathcal{U}(n)$ , se  $\mathcal{V}$  é Hermitiano.

De facto, se  $\mathcal{C} = \{\mathbf{e}_i\}$  e  $\widehat{\mathcal{C}} = \{\widehat{\mathbf{e}}_j\}$ , com  $\langle \mathbf{e}_i | \mathbf{e}_j \rangle = \delta_{ij}$  e análogamente  $\langle \widehat{\mathbf{e}}_\ell | \widehat{\mathbf{e}}_k \rangle = \delta_{\ell k}$ , então, como:

$$\widehat{\mathbf{e}}_i = \mathbf{e}_\ell P_i^\ell$$

vem que (supondo que  $\mathcal{V}$  é Hermitiano):

$$\begin{aligned} \delta_{ij} &= \langle \widehat{\mathbf{e}}_i | \widehat{\mathbf{e}}_j \rangle \\ &= \langle \mathbf{e}_\ell P_i^\ell | \mathbf{e}_k P_j^k \rangle \\ &= P_i^\ell \overline{P_j^k} \langle \mathbf{e}_\ell | \mathbf{e}_k \rangle \\ &= P_i^\ell \overline{P_j^k} \delta_{\ell k} \\ &= \sum_k P_i^k \overline{P_j^k} \\ &= (P^t)_k^i \overline{P_j^k} \quad \Rightarrow \quad P^t \overline{P} = \text{Id} \end{aligned} \quad (5.1.2)$$

o que mostra que  $P$  é unitária:  $P^\dagger P = \text{Id}$ . No caso Euclideano, a demonstração é análoga e, neste caso,  $P$  é ortogonal:  $P^t P = \text{Id}$ .

► **5.4** Portanto, quando  $\mathcal{V}$  é um espaço vectorial com um produto interno, a pergunta anterior deve ser reformulada da seguinte forma:

**Como escolher a base ortonormada de  $\mathcal{V}$  de tal forma que a representação matricial de  $L$  seja o mais “simples” possível? Mais formalmente - se  $[L]_{\mathcal{C}}$  é a representação matricial de  $L$  numa certa base ortonormada  $\mathcal{C}$ , como seleccionar na classe de conjugação de  $[L]_{\mathcal{C}}$ :**

$$\{[L]_{\mathcal{C}} P = P^{-1} [L]_{\mathcal{C}} P : P \in \mathcal{U}(n)\}$$

**o representante mais “simples” possível? (no caso Euclideano,  $\mathcal{U}(n)$  será substituído por  $\mathcal{O}(n)$ , é claro!)**

► **5.5 Definição** ... Seja  $(\mathcal{V}, \langle | \rangle)$  um espaço com um produto interno (*Euclideano se  $\mathcal{V}$  é real, ou Hermitiano, se  $\mathcal{V}$  for complexo*). Um operador linear  $\mathbf{S} : \mathcal{V} \rightarrow \mathcal{V}$ , diz-se **auto-adjunto** se  $\mathbf{S}$  satisfaz a condição:

$$\langle \mathbf{S}(\mathbf{v}) | \mathbf{w} \rangle = \langle \mathbf{v} | \mathbf{S}(\mathbf{w}) \rangle \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V} \quad (5.1.3)$$

No caso Euclideano  $\mathbf{S}$  diz-se um **operador simétrico**, enquanto que no caso Hermitiano,  $\mathbf{S}$  diz-se um **operador Hermitiano**.

► **5.6 Teorema** ... A matriz  $S = [S_j^i]$  de um operador auto-adjunto  $\mathbf{S} : \mathcal{V} \rightarrow \mathcal{V}$ , num espaço com um produto interno  $(\mathcal{V}, \langle | \rangle)$ , relativamente a uma base ortonormada  $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  de  $\mathcal{V}$ , é:

- uma **matriz simétrica**,  $S = S^t$ , no caso Euclideano.
- uma **matriz Hermitiana**,  $S = S^\dagger$ , no caso Hermitiano<sup>1</sup>.

<sup>1</sup>Se  $U(\epsilon)$  é uma curva de matrizes unitárias, tais que:

$$U(0) = \text{Id}, \quad \text{e} \quad U'(0) = iH$$

então:

$$U(\epsilon)^t \overline{U(\epsilon)} = \text{Id} \quad \Rightarrow \quad U'(0)^t + \overline{U'(0)} = 0 \quad \Rightarrow \quad iH^t - i\overline{H} = 0 \quad \Rightarrow \quad H^t = \overline{H}$$

isto é,  $H$  é Hermitiana

**Dem.:** De facto (no caso Hermitiano), se  $\mathbf{S}(\mathbf{e}_j) = S_j^k \mathbf{e}_k$ , então:

$$\langle \mathbf{e}_i | \mathbf{S}(\mathbf{e}_j) \rangle = \langle \mathbf{e}_i | S_j^k \mathbf{e}_k \rangle = \overline{S_j^k} \langle \mathbf{e}_i | \mathbf{e}_k \rangle = \overline{S_j^k} \delta_{ik} = \overline{S_j^i}$$

enquanto que, por outro lado, atendendo a (5.1.3):

$$\langle \mathbf{e}_i | \mathbf{S}(\mathbf{e}_j) \rangle = \langle \mathbf{S}(\mathbf{e}_i) | \mathbf{e}_j \rangle = S_i^k \langle \mathbf{e}_k | \mathbf{e}_j \rangle = S_i^k \delta_{kj} = S_i^j = (S^t)^j_i$$

Portanto  $S^t = \overline{S}$ , ou ainda  $S^\dagger = S$ . O caso Euclideano é análogo.

► **5.7 Teorema ...** *Seja  $\mathbf{S} : \mathcal{V} \rightarrow \mathcal{V}$ , um operador auto-adjunto num espaço com um produto interno  $(\mathcal{V}, \langle | \rangle)$ . Então:*

- *Se  $\mathbf{S}$  tem um valor próprio, esse valor próprio é real.*
- *Suponhamos que  $\mathbf{v}$  e  $\mathbf{w}$  são vectores próprios, pertencentes respectivamente aos valores próprios distintos  $\lambda$  e  $\eta$ , de  $\mathbf{S}$ . Então  $\mathbf{v}$  e  $\mathbf{w}$  são ortogonais:  $\langle \mathbf{v} | \mathbf{w} \rangle = 0$ .*

**Dem.:** 1. Seja  $\mathbf{v} \in \mathcal{V} - \{0\}$ , um vector próprio pertencente ao valor próprio  $\lambda$ :

$$\mathbf{S}(\mathbf{v}) = \lambda \mathbf{v} \tag{5.1.4}$$

Usando o produto interno  $\langle | \rangle$ , podemos exprimir o valor próprio  $\lambda$ , na forma:

$$\lambda = \frac{\langle \mathbf{S}\mathbf{v} | \mathbf{v} \rangle}{\|\mathbf{v}\|^2} \tag{5.1.5}$$

onde  $\mathbf{v}$  é um vector próprio pertencente ao valor próprio  $\lambda$ . De facto:

$$\mathbf{S}(\mathbf{v}) = \lambda \mathbf{v} \quad \Rightarrow \quad \langle \mathbf{S}\mathbf{v} | \mathbf{v} \rangle = \langle \lambda \mathbf{v} | \mathbf{v} \rangle = \lambda \|\mathbf{v}\|^2$$

o que implica (5.1.5), já que  $\mathbf{v} \neq 0$ . Portanto se  $\mathbf{S}$  é auto-adjunto temos que:

$$\lambda = \frac{\langle \mathbf{S}(\mathbf{v}) | \mathbf{v} \rangle}{\|\mathbf{v}\|^2} = \frac{\langle \mathbf{v} | \mathbf{S}(\mathbf{v}) \rangle}{\|\mathbf{v}\|^2} = \overline{\lambda}$$

isto é  $\lambda \in \mathbb{R}$ .

2. Por hipótese,  $\mathbf{S}(\mathbf{v}) = \lambda \mathbf{v}$  e  $\mathbf{S}(\mathbf{w}) = \eta \mathbf{w}$ . Por 1. sabemos já que  $\lambda, \eta \in \mathbb{R}$ . Temos então sucessivamente que (no caso Hermitiano):

$$\lambda \langle \mathbf{v} | \mathbf{w} \rangle = \langle \lambda \mathbf{v} | \mathbf{w} \rangle = \langle \mathbf{S}\mathbf{v} | \mathbf{w} \rangle = \langle \mathbf{v} | \mathbf{S}\mathbf{w} \rangle = \langle \mathbf{v} | \eta \mathbf{w} \rangle = \eta \langle \mathbf{v} | \mathbf{w} \rangle = \eta \langle \mathbf{v} | \mathbf{w} \rangle$$

o que implica que  $(\lambda - \eta) \langle \mathbf{v} | \mathbf{w} \rangle = 0$ , e portanto  $\langle \mathbf{v} | \mathbf{w} \rangle = 0$ , já que  $\lambda \neq \eta$ . O caso Euclideano é análogo.

## 5.2 Teorema espectral para operadores auto-adjuntos

► **5.8** Notemos que um operador linear real pode não ter valores próprios reais (por exemplo, uma rotação em  $\mathbb{R}^2$ ). No entanto, é possível provar que todo o operador auto-adjunto tem pelo menos um valor próprio que, pela proposição anterior, é real.

O facto de maior interesse sobre operadores auto-adjuntos em espaços com produto interno de dimensão finita, é que eles podem ser diagonalizados por conjugação pelo grupo ortogonal  $\mathcal{O}(n)$  (no caso Euclideano, isto é, quando  $\mathbf{S}$  é operador simétrico) ou pelo grupo unitário  $\mathcal{U}(n)$  (no caso Hermitiano, isto é, quando  $\mathbf{S}$  é operador Hermitiano). Mais precisamente, é válido o seguinte teorema fundamental.

► **5.9 Teorema ...** [Teorema espectral para operadores auto-adjuntos em espaços com produto interno de dimensão finita] ...

Seja  $\mathbf{S} : \mathcal{V} \rightarrow \mathcal{V}$ , um operador auto-adjunto num espaço com produto interno  $(\mathcal{V}, \langle | \rangle)$ , de dimensão finita  $n$ .

Então existe uma base ortonormada  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ , para  $\mathcal{V}$ , constituída por vectores próprios de  $\mathbf{S}$ .

A matriz de  $\mathbf{S}$  nessa base é portanto a matriz diagonal  $\mathbf{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ , onde  $\lambda_k$  é o valor próprio correspondente ao vector próprio  $\mathbf{u}_k$ , para  $(k = 1, \dots, n)$ .

**Dem.:** A demonstração far-se-á por indução sobre a dimensão  $n$ . Se  $n = 1$ , o resultado é trivial. Suponhamos que ele é válido, para todo o espaço vectorial com produto interno, com  $\dim \leq n - 1$ .

Como se referiu acima,  $\mathbf{S}$  admite sempre um valor próprio (real)  $\lambda_1$ . Seja  $\mathbf{u}_1 \neq \mathbf{0}$  um vector próprio pertencente ao valor próprio  $\lambda_1$ :  $\mathbf{S}(\mathbf{u}_1) = \lambda_1 \mathbf{u}_1$ . Podemos supôr que  $\|\mathbf{u}_1\| = 1$ . Seja  $\mathcal{S}$  o subespaço ortogonal a  $\mathbf{u}_1$ , de tal forma que:

$$\mathcal{V} = \mathbb{R} \mathbf{u}_1 \oplus \mathcal{S} \quad (5.2.1)$$

Então  $\mathbf{S}$  deixa  $\mathcal{S}$  invariante:  $\mathbf{S}(\mathcal{S}) \subseteq \mathcal{S}$  (porquê?). Além disso,  $\mathcal{S}$  é um espaço vectorial com um produto interno, de dimensão  $n - 1$ , e  $\mathbf{S}|_{\mathcal{S}}$  é auto-adjunto. Resta aplicar a hipótese de indução para concluir a prova.

► **5.10 Exemplo ...** Seja  $\mathbf{S}$  o operador simétrico em  $\mathbb{R}^3$ , cuja matriz na base canónica de  $\mathbb{R}^3$  é (a matriz simétrica):

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

A equação característica é:

$$p(t) = \det(S - t\text{Id}) = \begin{vmatrix} 1-t & 0 & 0 \\ 0 & 1-t & 2 \\ 0 & 2 & 1-t \end{vmatrix} = 0$$

isto é:

$$(1-t)[(1-t)^2 - 4] = 0$$

Os valores próprios de  $S$ , são portanto  $t = 1, -1, 3$ . Calculemos uma base ortonormada de vectores próprios. Para isso substituímos sucessivamente  $t$  por  $1, -1$  e  $3$ , na equação matricial seguinte:

$$\begin{bmatrix} 1-t & 0 & 0 \\ 0 & 1-t & 2 \\ 0 & 2 & 1-t \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Resolvendo os correspondentes sistemas de equações, e tendo o cuidado de normalizar os vectores

próprios para que eles tenham norma 1, obtemos a base seguinte:

$$\begin{aligned} \mathbf{u}_1 &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} && \text{pertencente ao valor próprio } \lambda = 1 \\ \mathbf{u}_2 &= \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} && \text{pertencente ao valor próprio } \lambda = -1 \\ \mathbf{u}_3 &= \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} && \text{pertencente ao valor próprio } \lambda = 3 \end{aligned}$$

Designando por  $\mathcal{C} = [\mathbf{i} \ \mathbf{j} \ \mathbf{k}]$  a base canónica de  $\mathbb{R}^3$  e por  $\mathcal{B} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3]$ , a base constituída pelos vectores próprios de  $\mathbf{S}$ , atrás calculados, e pondo:

$$\mathcal{B} = \mathcal{C}P$$

vemos que a matriz  $P$  (que é ortogonal - ( $P^{-1} = P^{tr}$  - como vimos), é dada por:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Podemos verificar directamente que:

$$P^t S P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

## 5.3 Diagonalização de formas quadráticas reais

► **5.11** Suponhamos agora que  $\mathcal{V}$  é um espaço vectorial real de dimensão  $n$ , com um produto interno Euclideano  $\langle \cdot, \cdot \rangle$ , e que:

$$\beta : \mathcal{V} \times \mathcal{V} \longrightarrow \mathbb{R} \quad (5.3.1)$$

é uma forma bilinear simétrica em  $\mathcal{V}$ . A **forma quadrática** associada a  $\beta$  é, por definição, a função  $Q = Q_\beta : \mathcal{V} \rightarrow \mathbb{R}$  dada por:

$$Q(\mathbf{v}) = \beta(\mathbf{v}, \mathbf{v}), \quad \mathbf{v} \in \mathcal{V} \quad (5.3.2)$$

► **5.12** Seja  $\mathcal{C} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  uma base para  $\mathcal{V}$ . Por definição, a **matriz de Gram** de  $\beta$  na base  $\mathcal{C}$ , é a matriz simétrica  $[\beta]_{\mathcal{C}} = [\beta_{ij}]$ , dada por:

$$\beta_{ij} \stackrel{\text{def}}{=} \beta(\mathbf{e}_i, \mathbf{e}_j), \quad i, j = 1, \dots, n \quad (5.3.3)$$

Se  $\mathbf{v} = x^i \mathbf{e}_i$ , então:

$$\begin{aligned} Q(\mathbf{v}) &= Q(x^i \mathbf{e}_i) \\ &\stackrel{\text{def}}{=} Q(x^1, \dots, x^n) \\ &= \beta(x^i \mathbf{e}_i, x^j \mathbf{e}_j) \\ &= \sum_{ij} \beta_{ij} x^i x^j \\ &= [\mathbf{v}]_{\mathcal{C}}^t [\beta]_{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}}, \quad \text{em notação matricial} \end{aligned} \quad (5.3.4)$$

► **5.13** Se mudarmos a base  $\mathcal{C}$ , para uma nova base  $\mathcal{C}P$ :

$$\mathcal{C} \longrightarrow \mathcal{C}P$$

sabemos já que as coordenadas de um vector  $\mathbf{v}$  mudam de acordo com a fórmula:

$$\mathcal{C} \longrightarrow \mathcal{C}P \quad \Longrightarrow \quad [\mathbf{v}]_{\mathcal{C}P} = P^{-1}[\mathbf{v}]_{\mathcal{C}}$$

**Qual é a matriz de Gram de  $\beta$  na base  $\mathcal{C}P$ ?**

Por um lado:

$$\begin{aligned} Q(\mathbf{v}) &= [\mathbf{v}]_{\mathcal{C}}^t [\beta]_{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}} \\ &= (P[\mathbf{v}]_{\mathcal{C}P})^t [\beta]_{\mathcal{C}} P[\mathbf{v}]_{\mathcal{C}P} \\ &= [\mathbf{v}]_{\mathcal{C}P}^t P^t [\beta]_{\mathcal{C}} P [\mathbf{v}]_{\mathcal{C}P} \end{aligned} \quad (5.3.5)$$

e, por outro lado:

$$Q(\mathbf{v}) = [\mathbf{v}]_{\mathcal{C}P}^t [\beta]_{\mathcal{C}P} [\mathbf{v}]_{\mathcal{C}P}$$

Comparando as duas expressões, concluímos que:

$$\boxed{\mathcal{C} \longrightarrow \mathcal{C}P \quad \Longrightarrow \quad [\beta]_{\mathcal{C}P} = P^t [\beta]_{\mathcal{C}} P} \quad (5.3.6)$$

► **5.14** À forma bilinear simétrica  $\beta$ , podemos associar um operador simétrico  $\mathbf{S} = \mathbf{S}_\beta : \mathcal{V} \rightarrow \mathcal{V}$ , tal que:

$$\beta(\mathbf{u}, \mathbf{v}) = \langle \mathbf{S}(\mathbf{u}) | \mathbf{v} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in \mathcal{V} \quad (5.3.7)$$

De facto, se  $\mathbf{u} \in \mathcal{V}$ , a fórmula (5.3.7) define  $\mathbf{S}(\mathbf{u})$  como sendo o único vector de  $\mathcal{V}$  tal que  $\langle \mathbf{S}(\mathbf{u}) | \mathbf{v} \rangle = \beta(\mathbf{u}, \mathbf{v})$ ,  $\forall \mathbf{v} \in \mathcal{V}$ . Não há ambiguidade nesta definição uma vez que o produto interno  $\langle | \rangle$  é não degenerado. Além disso:

$$\langle \mathbf{S}(\mathbf{u}) | \mathbf{v} \rangle = \beta(\mathbf{u}, \mathbf{v}) = \beta(\mathbf{v}, \mathbf{u}) = \langle \mathbf{S}(\mathbf{v}) | \mathbf{u} \rangle = \langle \mathbf{u} | \mathbf{S}(\mathbf{v}) \rangle$$

e portanto  $\mathbf{S}$  é um operador simétrico.

É fácil ver que a matriz de  $\mathbf{S}$ , relativamente à base  $\mathcal{C}$ , é a matriz de Gram  $[\beta]_{\mathcal{C}}$ . Pelo teorema espectral da secção anterior, podemos encontrar uma base ortonormada  $\mathcal{B} = \mathcal{C}P = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ , de  $\mathcal{V}$ , constituída por vectores próprios de  $\mathbf{S}$ , e relativamente à qual a matriz de  $\mathbf{S}$  é a matriz diagonal:

$$[\beta]_{\mathcal{C}P} = D = \mathbf{diag}[\lambda_1 \ \lambda_2 \ \dots \ \lambda_n]$$

onde  $\lambda_k$  é o valor próprio correspondente ao vector próprio  $\mathbf{u}_k$ , para  $(k = 1, \dots, n)$ .

► **5.15** Atendendo a (5.3.6), vemos que:

$$\begin{aligned} Q(\mathbf{v}) &= [\mathbf{v}]_{\mathcal{C}P}^t [\beta]_{\mathcal{C}P} [\mathbf{v}]_{\mathcal{C}P} \\ &= [\mathbf{v}]_{\mathcal{C}P}^t \mathbf{diag}[\lambda_1 \ \lambda_2 \ \dots \ \lambda_n] [\mathbf{v}]_{\mathcal{C}P} \end{aligned} \quad (5.3.8)$$

Pondo  $\mathbf{v} = x^i \mathbf{e}_i = y^j \mathbf{u}_j$ , isto é:

$$[\mathbf{v}]_{\mathcal{C}} = [x^i], \quad [\mathbf{v}]_{\mathcal{C}P} = [y^j]$$

concluimos que:

$$\begin{aligned}
 Q(\mathbf{v}) &= Q(x^i \mathbf{e}_i) \\
 &\stackrel{\text{def}}{=} Q(x^1, \dots, x^n) \\
 &= [\mathbf{v}]_{\mathcal{C}}^t [\beta]_{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}} \\
 &= Q(y^j \mathbf{u}_j) \\
 &\stackrel{\text{def}}{=} Q(y^1, \dots, y^n) \\
 &= [\mathbf{v}]_{\mathcal{C}P}^t [\beta]_{\mathcal{C}P} [\mathbf{v}]_{\mathcal{C}P} \\
 &= [\mathbf{v}]_{\mathcal{C}P}^t \mathbf{diag}[\lambda_1 \ \lambda_2 \ \dots \ \lambda_n] [\mathbf{v}]_{\mathcal{C}P} \\
 &= \sum_i \lambda_i (y^i)^2
 \end{aligned} \tag{5.3.9}$$

Portanto, a forma quadrática associada a  $\beta$ , que nas  $x$ -coordenadas (relativamente à base  $\mathcal{C}$ ) foi escrita na forma (ver (5.3.4)):

$$Q(x^1, \dots, x^n) = \sum_{ij} b_{ij} x^i x^j$$

escreve-se agora, nas  $y$ -coordenadas (relativamente à base  $\mathcal{B} = \mathcal{C}P$ , que diagonaliza  $\mathbf{S}$ ), na forma:

$$Q(y^1, \dots, y^n) = \sum_i \lambda_i (y^i)^2$$

► **5.16 Exemplo ...** Continuando o exemplo da secção anterior, consideremos a forma quadrática associada ao endomorfismo simétrico aí referido:

$$\begin{aligned}
 q(x^1, x^2, x^3) &= [x^1 \ x^2 \ x^3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} \\
 &= (x^1)^2 + (x^2)^2 + (x^3)^2 + 4x^2x^3
 \end{aligned}$$

Se designamos por  $\begin{bmatrix} y^1 \\ y^2 \\ y^3 \end{bmatrix}$  as coordenadas de um vector  $\mathbf{v}$ , na base  $\mathcal{B}$ , então, se as coordenadas

desse mesmo vector, na base  $\mathcal{C}$ , são  $\begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}$ , vem que:

$$\begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = P \begin{bmatrix} y^1 \\ y^2 \\ y^3 \end{bmatrix}, \quad \text{onde} \quad P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

isto é:

$$\begin{aligned}
 x^1 &= y^1 \\
 x^2 &= \frac{1}{\sqrt{2}}y^2 + \frac{1}{\sqrt{2}}y^3 \\
 x^3 &= -\frac{1}{\sqrt{2}}y^2 + \frac{1}{\sqrt{2}}y^3
 \end{aligned}$$

e nas novas coordenadas  $(y^i)$ ,  $q$  escreve-se na forma:

$$q(y^1, y^2, y^3) = (y^1)^2 - (y^2)^2 + 3(y^3)^2$$

como aliás pode ser verificado directamente.

► **5.17 Definição** ... Uma forma quadrática em  $\mathbb{R}^3$ ,  $Q(\mathbf{x}) = \mathbf{S}\mathbf{x} \cdot \mathbf{x}$ , diz-se:

- **definida positiva**, se  $Q(\mathbf{x}) > 0$ ,  $\forall \mathbf{x} \neq \mathbf{0}$ .
- **definida negativa**, se  $Q(\mathbf{x}) < 0$ ,  $\forall \mathbf{x} \neq \mathbf{0}$ .
- **indefinida**, se  $Q$  toma valores positivos e negativos.

A proposição seguinte é consequência imediata da possibilidade de reduzir uma forma quadrática à forma diagonal.

► **5.18 Teorema** ... Uma forma quadrática em  $\mathbb{R}^3$ ,  $Q(\mathbf{x}) = \mathbf{S}\mathbf{x} \cdot \mathbf{x}$ , é:

- **definida positiva**, se todos os valores próprios de  $\mathbf{S}$  são estritamente positivos.
- **definida negativa**, se todos os valores próprios de  $\mathbf{S}$  são estritamente negativos.
- **indefinida**, se os valores próprios de  $\mathbf{S}$  são alguns positivos e alguns negativos (eventualmente nulos).

Finalizamos esta secção com a seguinte propriedade extremal dos valores próprios de uma matriz simétrica (ou da forma quadrática associada)<sup>2</sup>:

► **5.19 Teorema** ... Seja  $\mathbf{S} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  um endomorfismo simétrico de  $\mathbb{R}^n$ , e  $Q : \mathbb{R}^n \rightarrow \mathbb{R}$  a forma quadrática associada a  $\mathbf{S}$ , definida por  $Q(\mathbf{x}) = \mathbf{x}^t \mathbf{S} \mathbf{x}$ , onde  $S$  é a matriz de  $\mathbf{S}$  relativamente à base canónica de  $\mathbb{R}^n$ .

A base ortonormada  $\{\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_n\}$ , de  $\mathbb{R}^n$ , constituída por vectores próprios de  $\mathbf{S}$  ( $\mathbf{S}(\mathbf{u}_k) = \lambda_k \mathbf{u}_k$ ,  $k = 1, \dots, n$ ), e relativamente à qual a matriz de  $\mathbf{S}$  é a matriz diagonal:

$$D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

pode ser escolhida de tal forma que, para cada  $k = 1, \dots, n$ ,  $\lambda_k = Q(\mathbf{u}_k)$  é o valor máximo de  $Q$ , restrita à esfera unitária no subespaço de  $\mathbb{R}^n$ , perpendicular aos vectores  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$ .

\* **Dem.:** Com efeito, escolhamos  $\mathbf{u}_1$  como sendo um máximo condicionado da restrição de  $Q$ , à esfera  $S_1 \equiv \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|^2 = 1\}$  (isto é sempre possível...). Consideremos o subespaço de  $\mathbb{R}^n$ , perpendicular a  $\mathbf{u}_1$ :

$$V(\mathbf{u}_1) \equiv \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{u}_1 = 0\}$$

e escolhamos  $\mathbf{u}_2$  como sendo um máximo condicionado da restrição de  $Q$ , à esfera  $S_2 \equiv \{\mathbf{x} \in V(\mathbf{u}_1) : \|\mathbf{x}\|^2 = 1\}$  (isto é sempre possível...). Consideremos de seguida, o subespaço de  $\mathbb{R}^n$ , perpendicular a  $\mathbf{u}_1$  e a  $\mathbf{u}_2$ :

$$V(\mathbf{u}_1, \mathbf{u}_2) \equiv \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{u}_1 = 0 = \mathbf{x} \cdot \mathbf{u}_2\}$$

<sup>2</sup>a demonstração que damos, usa o método dos multiplicadores de Lagrange, para o cálculo de extremos condicionados (curso de Cálculo).

e escolhamos  $\mathbf{u}_3$  como sendo um máximo condicionado da restrição de  $Q$ , à esfera  $S_3 \equiv \{\mathbf{x} \in V(\mathbf{u}_1, \mathbf{u}_2) : \|\mathbf{x}\|^2 = 1\}$  (isto é sempre possível...).

Procedendo sucessivamente desta forma, conseguimos  $n$  vectores  $\mathbf{u}_1, \dots, \mathbf{u}_n$  que são evidentemente ortonormais. Resta provar que eles são vectores próprios de  $\mathbf{S}$ . Como por construção,  $Q$  tem um máximo condicionado em  $\mathbf{u}_1$ , quando restrita à esfera  $S_1$ , existe um multiplicador de Lagrange  $\lambda_1$ , tal que:

$$\nabla Q(\mathbf{u}_1) = \lambda_1 \nabla g(\mathbf{u}_1) \quad (5.3.10)$$

onde  $g(\mathbf{x}) = \|\mathbf{x}\|^2 - 1$ . Mas o gradiente de  $Q$  é dado por  $\nabla Q(\mathbf{x}) = 2\mathbf{S}(\mathbf{x})$ , e em particular  $\nabla g(\mathbf{x}) = 2\mathbf{x}$ . Portanto a condição (5.3.10) é equivalente a:

$$\mathbf{S}(\mathbf{u}_1) = \lambda_1 \mathbf{u}_1$$

o que significa exactamente que  $\mathbf{u}_1$  é vector próprio associado ao valor próprio  $\lambda_1$ . O mesmo argumento pode ser utilizado sucessivamente, para concluir que  $\mathbf{u}_k$  é vector próprio de  $\mathbf{S}$ .

A forma quadrática associada a  $\mathbf{S}$  pode então ser escrita na forma diagonal:

$$Q(\mathbf{x}) = Q(y^1, \dots, y^n) = \lambda_1(y^1)^2 + \lambda_2(y^2)^2 + \dots + \lambda_n(y^n)^2 \quad (5.3.11)$$

e é claro que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ .

## 5.4 Diagonalização simultânea de duas formas quadráticas reais

► **5.20 Coordenadas normais e modos normais de vibração ...** Suponhamos que a **energia cinética** de um sistema mecânico com  $n$  graus de liberdade, é dada pela forma quadrática real definida-positiva:

$$\begin{aligned} T &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n g_{ij} \dot{x}^i \dot{x}^j \\ &= \frac{1}{2} g_{ij} \dot{x}^i \dot{x}^j, && \text{usando a notação de Einstein} \\ &= \frac{1}{2} \langle \dot{\mathbf{x}} | G \dot{\mathbf{x}} \rangle, && \text{usando notação vectorial} \\ &= \frac{1}{2} \dot{\mathbf{x}}^t G \dot{\mathbf{x}}, && \text{usando notação matricial} \end{aligned} \quad (5.4.1)$$

onde  $\dot{\phantom{x}}$  representa derivada em ordem ao tempo  $t$ , enquanto que a **energia potencial** é dada pela forma quadrática real:

$$\begin{aligned} V &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n b_{ij} x^i x^j \\ &= \frac{1}{2} b_{ij} x^i x^j, && \text{usando a notação de Einstein} \\ &= \frac{1}{2} \langle \mathbf{x} | B \mathbf{x} \rangle, && \text{usando notação vectorial} \\ &= \frac{1}{2} \mathbf{x}^t B \mathbf{x}, && \text{usando notação matricial} \end{aligned} \quad (5.4.2)$$

Vamos mostrar que é possível introduzir novas coordenadas  $y^1, \dots, y^n$  relativamente às quais as novas expressões de  $T$  e  $V$  são:

$$\begin{aligned} 2T &= (\dot{y}^1)^2 + (\dot{y}^2)^2 + \dots + (\dot{y}^n)^2 \\ 2V &= \lambda_1(y^1)^2 + \lambda_2(y^2)^2 + \dots + \lambda_n(y^n)^2 \end{aligned} \quad (5.4.3)$$

Estas novas coordenadas dizem-se **coordenadas normais** e os números  $\lambda_1, \dots, \lambda_n$  dizem-se os **modos normais** (de vibração).

► **5.21 Teorema** ... Seja  $\mathcal{V}$  um espaço vectorial real de dimensão  $n$ , e  $g$  e  $\beta$  duas formas bilineares simétricas em  $\mathcal{V}$ . Suponhamos além disso que  $g$  é não degenerada definida positiva (e portanto define um produto interno em  $\mathcal{V}$ ).

Então existe uma base  $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ , em  $\mathcal{V}$ , relativamente à qual a matriz de Gram de  $g$  é a matriz identidade e a matriz de Gram de  $\beta$  é uma matriz diagonal real:

$$g(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}, \quad \beta(\mathbf{u}_i, \mathbf{u}_j) = \mathbf{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\} \quad \lambda_i \in \mathbb{R} \quad (5.4.4)$$

**Dem.:** Por hipótese,  $g$  é não degenerada definida positiva e portanto define um produto interno em  $\mathcal{V}$ , que representamos por  $\langle | \rangle = g$ , como habitualmente.

Como vimos anteriormente, num espaço vectorial real com produto interno Euclideano  $(\mathcal{V}, \langle | \rangle)$ , a cada forma bilinear simétrica  $\beta$ , podemos associar um operador simétrico  $\mathbf{S}_\beta : \mathcal{V} \rightarrow \mathcal{V}$ , tal que:

$$\beta(\mathbf{u}, \mathbf{v}) = \langle \mathbf{S}_\beta(\mathbf{u}) | \mathbf{v} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in \mathcal{V} \quad (5.4.5)$$

Associemos então a  $\beta$  um operador auto-adjunto  $\mathbf{S}_\beta$ , de acordo com a fórmula (5.4.5):

$$\beta(\mathbf{u}, \mathbf{v}) = \langle \mathbf{S}_\beta(\mathbf{u}) | \mathbf{v} \rangle = g(\mathbf{S}_\beta(\mathbf{u}), \mathbf{v}), \quad \forall \mathbf{u}, \mathbf{v} \in \mathcal{V}$$

Basta calcular a base  $g$ -ortonormada que diagonaliza o operador  $\mathbf{S}_\beta$ . É essa a base pretendida.

► **5.22** Portanto, se  $\mathbf{u} = y^i \mathbf{u}_i \in \mathcal{V}$  e se  $Q_g$  e  $Q_\beta$  são as formas quadráticas, associadas respectivamente a  $g$  e a  $\beta$ , então:

$$Q_g(\mathbf{u}) = \sum_i (y^i)^2, \quad Q_\beta(\mathbf{u}) = \sum_i \lambda_i (y^i)^2, \quad \lambda_i \in \mathbb{R} \quad (5.4.6)$$

► **5.23** Na prática as formas bilineares simétricas  $g$  e  $\beta$ , em  $\mathcal{V}$ , são dadas pelas suas matrizes de Gram, relativamente a uma certa base  $\mathcal{C} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  de  $\mathcal{V}$ , digamos:

$$G = [g]_{\mathcal{C}} = [g_{ij}] = [g(\mathbf{e}_i, \mathbf{e}_j)], \quad \text{e} \quad B = [\beta]_{\mathcal{C}} = [b_{ij}] = [\beta(\mathbf{e}_i, \mathbf{e}_j)] \quad (5.4.7)$$

de tal forma que:

$$Q_g(\mathbf{v}) = [\mathbf{v}]_{\mathcal{C}}^t [g]_{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}} = \sum_{ij} g_{ij} x^i x^j, \quad \text{e} \quad Q_\beta(\mathbf{v}) = [\mathbf{v}]_{\mathcal{C}}^t [\beta]_{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}} = \sum_{ij} b_{ij} x^i x^j$$

se  $[\mathbf{v}]_{\mathcal{C}} = [x^i]$ .

Se  $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  é a base  $g$ -ortonormada referida no teorema, e se  $\mathcal{C}P = \mathcal{B}$ , sabemos que as matrizes de Gram de  $g$  e  $\beta$ , na nova base  $\mathcal{B} = \mathcal{C}P$  são dadas por:

$$\mathcal{C} \rightarrow \mathcal{C}P = \mathcal{B} \Rightarrow \text{Id} = [g]_{\mathcal{B}} = [g]_{\mathcal{C}P} = P^t [g]_{\mathcal{C}P} = P^t G P \quad (5.4.8)$$

e analogamente:

$$\mathcal{C} \rightarrow \mathcal{C}P = \mathcal{B} \Rightarrow \mathbf{diag}\{\lambda_1, \dots, \lambda_n\} = [\beta]_{\mathcal{B}} = [\beta]_{\mathcal{C}P} = P^t [\beta]_{\mathcal{C}P} = P^t B P \quad (5.4.9)$$

Os modos normais são pois os valores próprios da matriz  $D = \mathbf{diag}\{\lambda_1, \dots, \lambda_n\}$ , isto é, as raízes (reais) do polinómio  $p(\lambda) = \det(D - \lambda \text{Id})$ . Mas:

$$\begin{aligned} \det(D - \lambda \text{Id}) &= \det(P^t B P - \lambda P^t G P) \\ &= \det(P^t (B - \lambda G) P) \\ &= \det(P^t) \det(B - \lambda G) \det P \end{aligned}$$

vemos que:

$$\det(D - t \text{Id}) = 0 \quad \text{se e só se} \quad \det(B - t G) = 0 \quad (5.4.10)$$

uma vez que  $\det P \neq 0$ . Portanto os modos normais podem ser calculados como as soluções da equação:

$$\det(B - t G) = 0 \quad (5.4.11)$$

o que simplifica drásticamente os cálculos.

► **5.24 Concluindo:** Se as formas bilineares simétricas  $g$  e  $\beta$ , em  $\mathcal{V}$ , são dadas pelas suas matrizes de Gram, relativamente a uma certa base  $\mathcal{C} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  de  $\mathcal{V}$ , digamos:

$$G = [g]_{\mathcal{C}} = [g(\mathbf{e}_i, \mathbf{e}_j)], \quad \text{e} \quad B = [\beta]_{\mathcal{C}} = [\beta(\mathbf{e}_i, \mathbf{e}_j)] \quad (5.4.12)$$

então para diagonalizar simultaneamente  $g$  e  $\beta$ :

- calculamos os modos normais  $\lambda_1, \dots, \lambda_n$ , resolvendo a equação em  $\lambda$ :

$$\det(B - \lambda G) = 0 \quad (5.4.13)$$

- para cada modo normal  $\lambda_k$ , calculamos as soluções  $\mathbf{x} = (x^i)$  da equação:

$$[B - \lambda_k G] \mathbf{x} = 0 \quad (5.4.14)$$

A base que diagonaliza simultaneamente as duas formas bilineares é constituída pelos vectores da forma  $\mathbf{u} = \frac{\sum_i x^i \mathbf{e}_i}{\|\sum_i x^i \mathbf{e}_i\|_g}$ , onde  $\|\cdot\|_g$  é a  $g$ -norma.

► **5.25 Exemplo ...** Fazemos a diagonalização simultânea das formas quadráticas seguintes:

$$g(x, y) = x^2 - 2xy + 4y^2 \quad \text{e} \quad \beta(x, y) = -4xy$$

É fácil ver que  $g$  é não degenerada definida positiva. As matrizes de  $g$  e  $\beta$ , relativamente à base canónica de  $\mathbb{R}^2$ , são, respectivamente:

$$G = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 0 & -2 \\ -2 & 0 \end{bmatrix}$$

Os modos normais são as soluções da equação:

$$\begin{aligned} \det(B - \lambda G) &= \det\left(\begin{bmatrix} 0 & -2 \\ -2 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}\right) \\ &= \det\begin{bmatrix} -\lambda & -2 + \lambda \\ -2 + \lambda & -4\lambda \end{bmatrix} \\ &= 3\lambda^2 - 4\lambda + 4 \end{aligned} \tag{5.4.15}$$

cujas soluções são  $\lambda_1 = -2$  e  $\lambda_2 = 2/3$ .

Vamos agora calcular as soluções  $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$  da equação:

$$(B - \lambda G)(\mathbf{x}) = 0$$

onde  $\lambda$  é modo normal.

Se  $\lambda_1 = -2$  vem que:

$$\begin{bmatrix} 2 & -4 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow 2x - 4y = 0 \quad \text{donde} \quad \mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Como:

$$\|\mathbf{x}_1\|_g^2 = g(\mathbf{x}_1, \mathbf{x}_1) = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 4$$

tomamos:

$$\mathbf{u}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|_g} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}$$

Se  $\lambda_2 = 2/3$  vem que:

$$\begin{bmatrix} -2/3 & -4/3 \\ -4/3 & 8/3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow -2x - 4y = 0 \quad \text{donde} \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Como:

$$\|\mathbf{x}_2\|_g^2 = g(\mathbf{x}_2, \mathbf{x}_2) = \begin{bmatrix} 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 12$$

tomamos:

$$\mathbf{u}_2 = \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|_g} = \begin{bmatrix} 1/\sqrt{3} \\ -1/2\sqrt{3} \end{bmatrix}$$

Portanto, na base  $\{\mathbf{u}_1, \mathbf{u}_2\}$ , se  $\mathbf{u} = x'\mathbf{u}_1 + y'\mathbf{u}_2$ , então:

$$\begin{aligned} Q_g(\mathbf{u}) &\stackrel{\text{def}}{=} Q_g(x', y') = (x')^2 + (y')^2 \\ Q_\beta(\mathbf{u}) &\stackrel{\text{def}}{=} Q_\beta(x', y') = -2(x')^2 + \frac{2}{3}(y')^2 \end{aligned}$$

e os modos normais são  $-2$  e  $\frac{2}{3}$ .

## 5.5 Exercícios

▷ **Exercício 5.1** ... Em cada uma das alíneas que se seguem, determine:

- I) Uma matriz simétrica  $A$  que represente a forma quadrática que se segue;
- II) Os valores próprios de  $A$ ;
- III) Uma base ortonormal de vectores próprios;
- IV) Uma matriz ortogonal diagonalizante  $C$ ;
- V) Diagonalize a forma quadrática.

a)  $q(x_1, x_2) = 4x_1^2 + 4x_1x_2 + x_2^2$ ;

b)  $q(x_1, x_2) = x_1x_2$ ;

c)  $q(x_1, x_2) = x_1^2 + 2x_1x_2 - x_2^2$ ;

d)  $q(x_1, x_2) = 34x_1^2 - 24x_1x_2 + 41x_2^2$ ;

e)  $q(x_1, x_2, x_3) = x_1^2 + x_1x_2 + x_2x_3 + x_1x_3$ ;

f)  $q(x_1, x_2, x_3) = 2x_1^2 + x_2^2 - x_3^2 + 4x_1x_3$ ;

g)  $q(x_1, x_2, x_3) = 3x_1^2 + 4x_1x_2 + 4x_2x_3 + 8x_1x_3 + 3x_3^2$ .

▷ **Exercício 5.2** ... Em cada uma das alíneas que se seguem, faça a diagonalização simultânea das formas quadráticas seguintes:

a)  $\varphi(x_1, x_2) = x_1^2 - 2x_1x_2 + 4x_2^2$  e  $\beta(x_1, x_2) = -4x_1x_2$ ;

b)  $\varphi(x_1, x_2) = x_1^2 + 16x_1x_2 + 56x_2^2$  e  $\beta(x_1, x_2) = x_1^2 + 10x_1x_2 + 26x_2^2$ ;

c)  $\varphi(x_1, x_2, x_3, x_4) = \frac{1}{4}x_1^2 + 2x_2x_4 + x_2^2 + x_3^2 + x_4^2$  e  $\beta(x_1, x_2, x_3, x_4) = x_1x_2 + 2x_2x_4 - 2x_2x_3 + x_1x_3 + 2x_4^2$ .

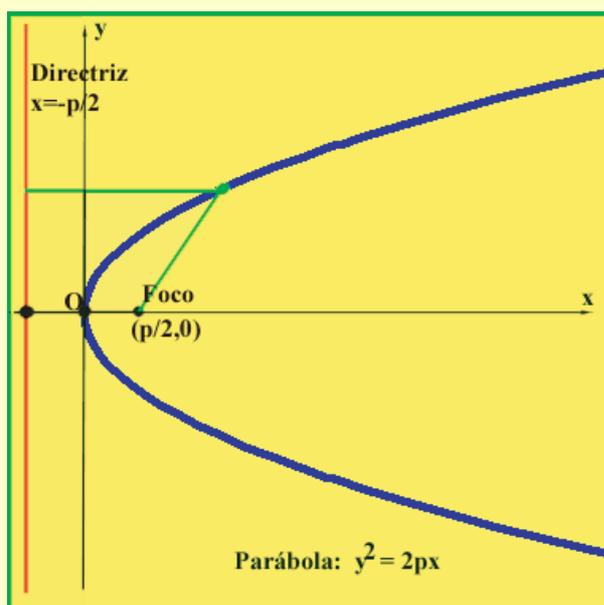
## Capítulo 6

# Cônicas e quádricas afins

### 6.1 Parábola, Elipse e Hipérbole



#### ► 6.1 Parábola



Uma parábola é uma curva em  $\mathbb{E}^2$  cuja equação, em coordenadas cartesianas  $(x, y)$  usuais, é:

$$y^2 = 2px, \quad p > 0 \quad (6.1.1)$$

Os seus elementos principais são:

- O **parâmetro**  $p > 0$
- A **distância focal**  $p/2$
- O **foco**  $F = (p/2, 0)$
- A **directriz** - a recta de equação:

$$x = -p/2$$

► **6.2 Propriedade focal** ... A parábola é o lugar geométrico dos pontos  $P(x, y)$  equidistantes do foco  $F(p/2, 0)$  e da directrix  $x = -p/2$ :

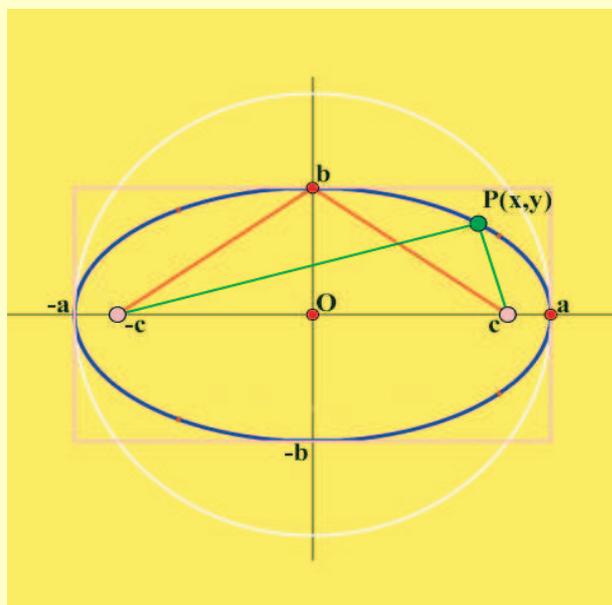
$$d(P, F) = d(P, d)$$

Com efeito:

$$\begin{aligned} d(P, F)^2 &= \|(x, y) - (p/2, 0)\|^2 \\ &= (x - p/2)^2 + y^2 \\ &= (x - p/2)^2 + 2px = (x + p/2)^2 \\ &= d(P, d)^2 \end{aligned}$$

► **6.3 Elipse** ... Uma elipse é uma curva em  $\mathbb{E}^2$  cuja equação, em coordenadas cartesianas  $(x, y)$  usuais, é:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad a \geq b > 0 \quad (6.1.2)$$



Os seus elementos principais são:

- O **semi-eixo maior**  $a > 0$
- O **semi-eixo menor**  $b > 0$
- A **distância focal**  $2c = 2\sqrt{a^2 - b^2}$
- A **excentricidade**  $e = c/a = \sqrt{1 - (b/a)^2}$
- O **parâmetro**  $p = b^2/a$
- Os **focos**  $(\pm c, 0)$
- Os **vértices**  $(\pm a, 0)$  e  $(0, \pm b)$
- As **directrizes** - as rectas de equação:

$$x = \pm a/e$$

► **6.4 Propriedade focal I** ... A elipse é o lugar geométrico dos pontos  $P(x, y)$  cuja soma das distâncias aos focos é constante e igual a  $2a$ :

$$d(P, F_1) + d(P, F_2) \equiv 2a$$

► **6.5 Propriedade focal II** ... A elipse é o lugar geométrico dos pontos  $P(x, y)$  cuja razão das distâncias a um dos focos e à directriz correspondente é constante e igual a  $e$ :

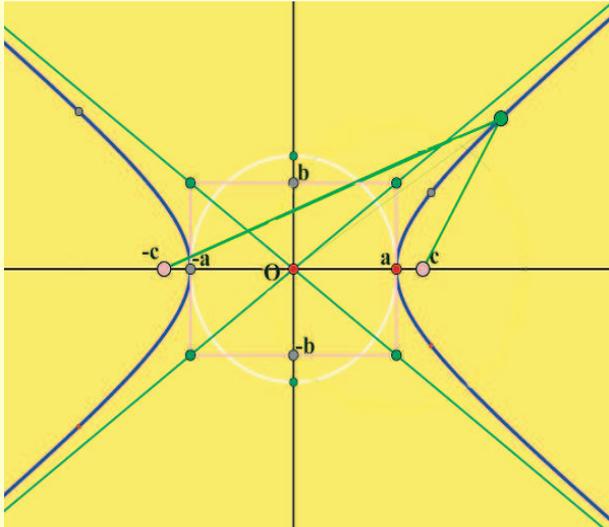
$$\frac{d(P, F_1)}{d(P, d_1)} = e = \frac{d(P, F_2)}{d(P, d_2)}$$

Esta propriedade é análoga à propriedade correspondente para a parábola, se considerarmos a parábola como uma elipse de excentricidade  $e = 1$ .

► **6.6 Hipérbole** ... Uma hipérbole é uma curva em  $\mathbb{E}^2$  cuja equação, em coordenadas cartesianas  $(x, y)$  usuais, é:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad a >, b > 0 \quad (6.1.3)$$

Quando  $a = b$  a hipérbole diz-se equilátera.



Os seus elementos principais são:

- O **semi-eixo real**  $a > 0$
- O **semi-eixo imaginário**  $b > 0$
- A **distância focal**  $2c = 2\sqrt{a^2 + b^2}$
- A **excentricidade**  $e = c/a = \sqrt{1 + (b/a)^2}$ . Claro que  $1 < e < \infty$ .
- O **parâmetro**  $p = b^2/a$
- Os **focos**  $(\pm c, 0)$
- Os **vértices**  $(\pm a, 0)$
- As **directrizes** - as rectas de equação:
 
$$x = \pm a/e$$
- As **assíntotas** - as rectas de equação:
 
$$x = \pm b/a$$

► **6.7 Propriedade focal I** ... A hipérbole é o lugar geométrico dos pontos  $P(x, y)$  cuja diferença das distâncias aos focos é, em valor absoluto, constante e igual a  $2a$ :

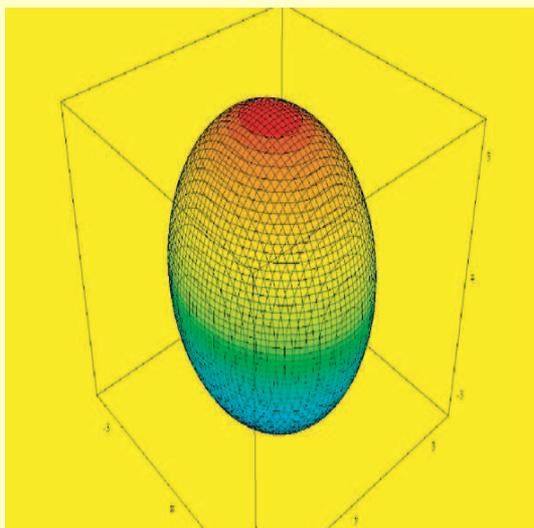
$$|d(P, F_1) - d(P, F_2)| \equiv 2a$$

► **6.8 Propriedade focal II** ... A hipérbole é o lugar geométrico dos pontos  $P(x, y)$  cuja razão das distâncias a um dos focos e à directriz correspondente é constante e igual a  $e$ :

$$\frac{d(P, F_1)}{d(P, d_1)} = e = \frac{d(P, F_2)}{d(P, d_2)}$$

## 6.2 Quádricas

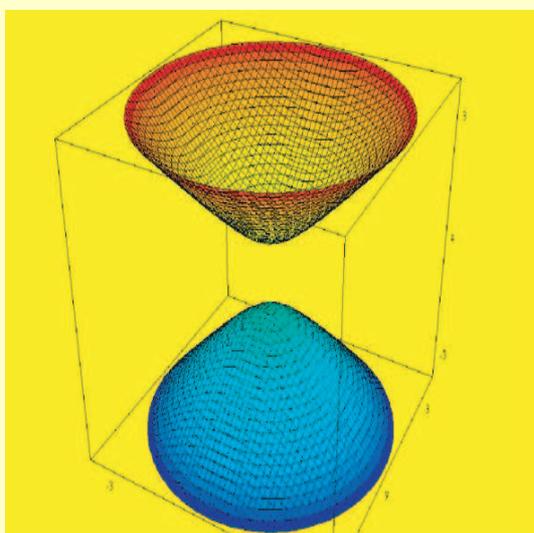
► **6.9 Elipsóides** ...



São superfícies em  $\mathbb{E}^3$  definidas por uma equação do tipo:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad a \geq b \geq c > 0 \quad (6.2.1)$$

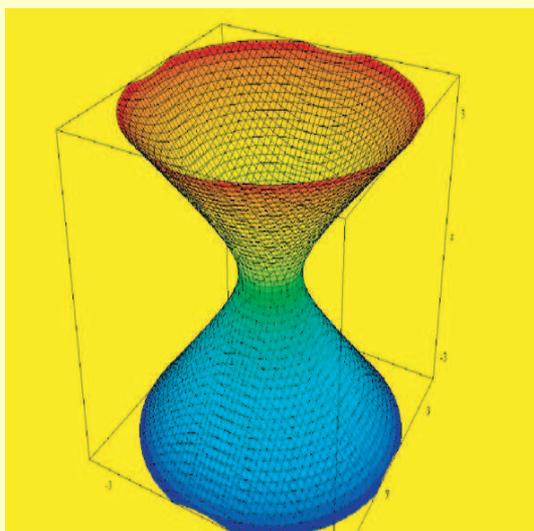
► **6.10** Hiperbolóides de duas folhas ...



São superfícies em  $\mathbb{E}^3$  definidas por uma equação do tipo:

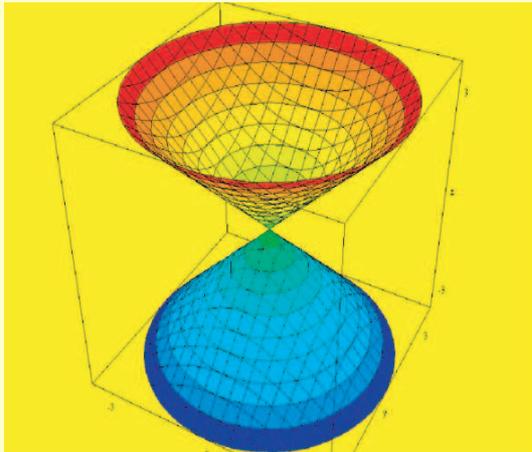
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1, \quad a \geq b > 0, \quad c > 0 \quad (6.2.2)$$

► **6.11** Hiperbolóides de uma folha ...



São superfícies em  $\mathbb{E}^3$  definidas por uma equação do tipo:

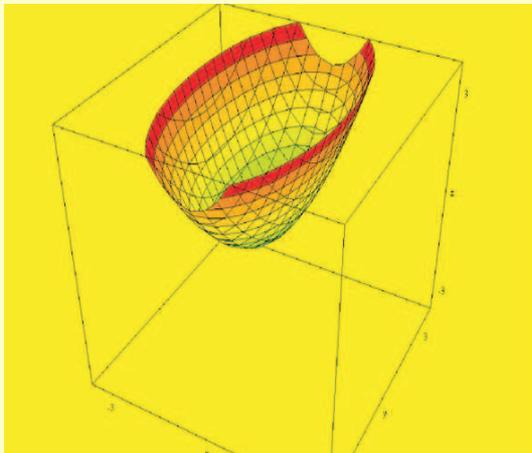
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \quad a \geq b > 0, \quad c > 0 \quad (6.2.3)$$

► 6.12 Cones ...

São superfícies em  $\mathbb{E}^3$  definidas por uma equação do tipo:

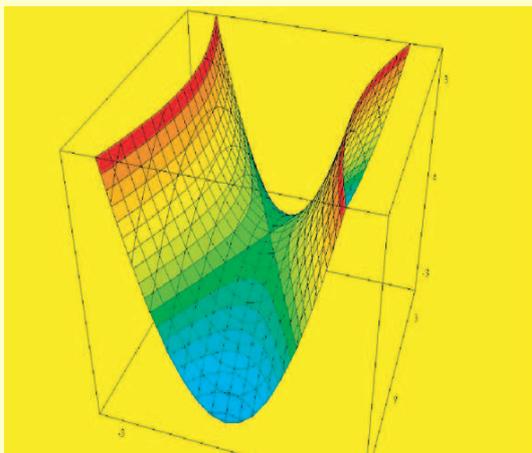
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0, \quad a \geq b > 0, \quad c > 0 \quad (6.2.4)$$

com  $1/a^2 + 1/b^2 + 1/c^2 = 1$ .

► 6.13 Parabolóide elíptico ...

São superfícies em  $\mathbb{E}^3$  definidas por uma equação do tipo:

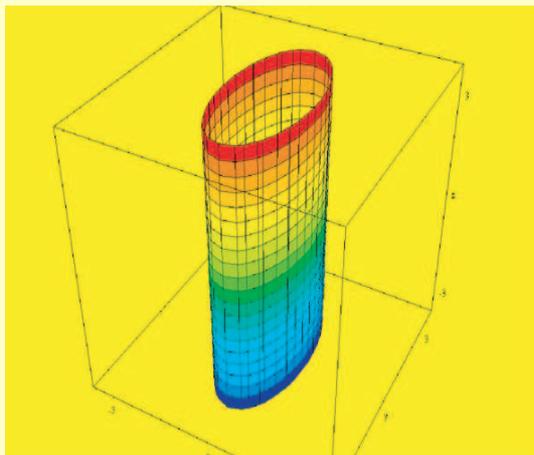
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 2z, \quad a \geq b > 0 \quad (6.2.5)$$

► 6.14 Parabolóide hiperbólico ...

São superfícies em  $\mathbb{E}^3$  definidas por uma equação do tipo:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2z, \quad a > 0, \quad b > 0 \quad (6.2.6)$$

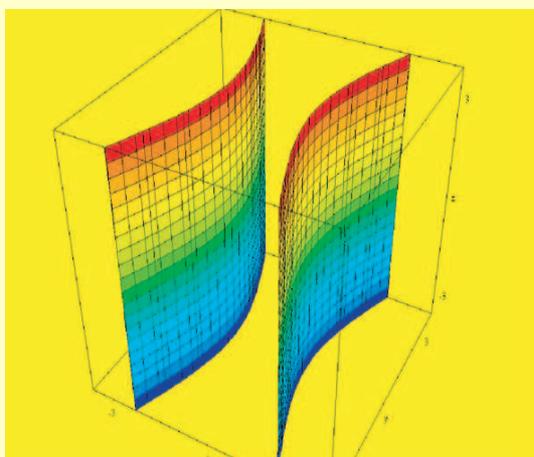
► 6.15 Cilindro elíptico ...



São superfícies em  $\mathbb{E}^3$  definidas por uma equação do tipo:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad a \geq b > 0 \quad (6.2.7)$$

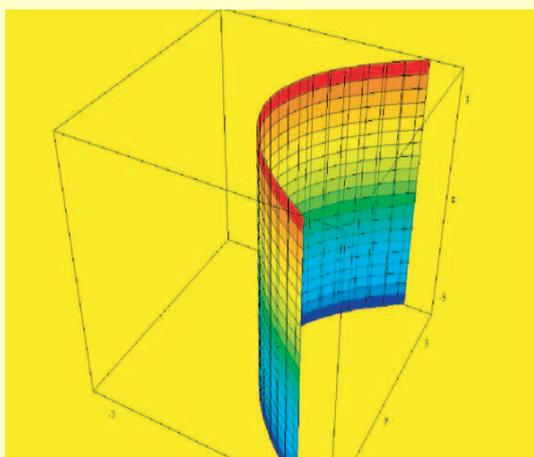
► 6.16 Cilindro hiperbólico ...



São superfícies em  $\mathbb{E}^3$  definidas por uma equação do tipo:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad a > 0, \quad b > 0 \quad (6.2.8)$$

► 6.17 Cilindro parabólico ...



São superfícies em  $\mathbb{E}^3$  definidas por uma equação do tipo:

$$y^2 = 2px, \quad p > 0 \quad (6.2.9)$$

## 6.3 Cónicas e quádricas afins

► 6.18 Cónica afim ... Consideremos o plano  $\mathbb{E}^2$  com a sua estrutura afim e Euclideana usuais. Fixemos um referencial afim ortonormado  $\mathcal{R} = \{O; \mathbf{e}_1, \mathbf{e}_2\}$ .

Um ponto  $P$  em  $\mathbb{E}^2$  será identificado com o seu vector de posição  $\mathbf{x} = \overrightarrow{OP} \in \mathbb{R}^2$ . Uma cónica afim em  $\mathbb{E}^2$  é o conjunto dos pontos  $P$  cujas coordenadas,  $x$  e  $y$ , relativas ao referencial  $\mathcal{R}$ :

$$\overrightarrow{OP} = x\mathbf{e}_1 + y\mathbf{e}_2$$

satisfazem a equação:

$$Q(x, y) = ax^2 + by^2 + 2cxy + 2dx + 2ey + f = 0 \quad (6.3.1)$$

onde  $a, b, c, d, e, f \in \mathbb{R}$  com  $a, b, c$  não simultâneamente nulos.

► **6.19 Quádrica afim** ... Consideremos o espaço  $\mathbb{E}^3$  com a sua estrutura afim e Euclideana usuais. Fixemos um referencial afim ortonormado  $\mathcal{R} = \{O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ .

Um ponto  $P$  em  $\mathbb{E}^3$  será identificado com o seu vector de posição  $\mathbf{x} = \overrightarrow{OP} \in \mathbb{R}^3$ . Uma quádrica afim em  $\mathbb{E}^3$  é o conjunto dos pontos  $P$  cujas coordenadas,  $x, y$  e  $z$ , relativas ao referencial  $\mathcal{R}$ :

$$\mathbf{x} = \overrightarrow{OP} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$$

satisfazem a equação:

$$Q(x, y, z) = ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + 2gx + 2hy + 2kz + l = 0 \quad (6.3.2)$$

onde  $a, b, c, d, \dots \in \mathbb{R}$  com  $a, b, c, d, e, f$  não simultâneamente nulos.

► **6.20 Expressões matriciais. Cónicas afins** ... Podemos escrever a fórmula (6.3.1) em forma matricial:

$$\begin{aligned} Q(x, y) &= [x \ y] \begin{bmatrix} a & c \\ c & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + 2 [d \ e] \begin{bmatrix} x \\ y \end{bmatrix} + f \\ &= \mathbf{x}^t \mathbf{A} \mathbf{x} + 2\mathbf{x}^t \mathbf{b} + f \end{aligned} \quad (6.3.3)$$

ou ainda na forma:

$$\begin{aligned} Q(x, y) &= [x \ y \ 1] \left[ \begin{array}{cc|c} a & c & d \\ c & b & e \\ \hline d & e & f \end{array} \right] \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ &= [\mathbf{x} \ | \ 1]^t \left[ \begin{array}{c|c} \mathbf{A} & \mathbf{b} \\ \hline \mathbf{b}^t & f \end{array} \right] \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \\ &= [\mathbf{x} \ | \ 1]^t \mathbf{B} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \end{aligned} \quad (6.3.4)$$

► **6.21 Expressões matriciais. Quádricas** ... Analogamente podemos escrever a fórmula (6.3.2) em forma matricial:

$$\begin{aligned} Q(x, y, z) &= [x \ y \ z] \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + 2 [g \ h \ k] \begin{bmatrix} x \\ y \\ z \end{bmatrix} + l \\ &= \mathbf{x}^t \mathbf{A} \mathbf{x} + 2\mathbf{x}^t \mathbf{b} + f \end{aligned} \quad (6.3.5)$$

ou ainda na forma:

$$\begin{aligned}
 Q(x, y, z) &= \begin{bmatrix} x & y & z & 1 \end{bmatrix} \left[ \begin{array}{ccc|c} a & d & e & g \\ d & b & f & h \\ e & f & c & k \\ \hline g & h & k & l \end{array} \right] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \\
 &= \begin{bmatrix} \mathbf{x} & | & 1 \end{bmatrix}^t \left[ \begin{array}{c|c} \mathbf{A} & \mathbf{b} \\ \hline \mathbf{b}^t & f \end{array} \right] \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \\
 &= \begin{bmatrix} \mathbf{x} & | & 1 \end{bmatrix}^t \mathbf{B} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \tag{6.3.6}
 \end{aligned}$$

► **6.22 Efeito de uma translacção** ... Estudemos como muda a expressão (6.3.1) quando optamos por um outro referencial  $\mathcal{R}' = \{O'; \mathbf{e}_1, \mathbf{e}_2\}$ , com uma nova origem  $O'$ . Como:

$$\overrightarrow{OP} = \overrightarrow{OO'} + \overrightarrow{O'P} \tag{6.3.7}$$

Pondo:

$$\begin{aligned}
 \overrightarrow{OP} &= x\mathbf{e}_1 + y\mathbf{e}_2 \\
 \overrightarrow{OO'} &= x_o\mathbf{e}_1 + y_o\mathbf{e}_2 \\
 \overrightarrow{O'P} &= x'\mathbf{e}_1 + y'\mathbf{e}_2 \tag{6.3.8}
 \end{aligned}$$

vem que:

$$\begin{cases} x = x_o + x' \\ y = y_o + y' \end{cases} \tag{6.3.9}$$

e substituindo em (6.3.1), obtemos:

$$\begin{aligned}
 Q(x', y') &= Q(x = x_o + x', y = y_o + y') \\
 &= a(x_o + x')^2 + b(y_o + y')^2 + 2c(x_o + x')(y_o + y') + 2d(x_o + x') + 2e(y_o + y') + f \\
 &= a(x')^2 + b(y')^2 + 2cx'y' + 2(ax_o + cy_o + d)x' + 2(by_o + cx_o + e)y' + Q(x_o, y_o) \tag{6.3.10}
 \end{aligned}$$

Quando escrevemos  $Q$  na forma (6.3.1), mas agora nas coordenadas  $x', y'$ :

$$Q(x', y') = a'x'^2 + b'y'^2 + 2c'x'y' + 2d'x' + 2e'y' + f' = 0 \tag{6.3.11}$$

e comparamos com a expressão (6.3.10), obtemos:

$$\begin{aligned}
 a' &= a \\
 b' &= b \\
 c' &= c \\
 d' &= d + ax_o + cy_o \\
 e' &= e + by_o + cx_o \\
 f &= Q(x_o, y_o) \tag{6.3.12}
 \end{aligned}$$

isto é, os termos quadráticos mantêm-se inalterados, mas os lineares alteram-se como é natural. Em particular, o determinante:

$$\delta = \begin{vmatrix} a & c \\ c & b \end{vmatrix} = ab - c^2 \tag{6.3.13}$$

mantem-se inalterado.

► **6.23 Efeito de uma translacção. Escrita matricial** ... Os cálculos do número anterior podem ser feitos em forma matricial o que permite uma generalização imediata para o caso das quádricas afins. De facto, pondo:

$$\mathbf{x} = \mathbf{x}_o + \mathbf{x}'$$

onde  $\mathbf{x} = \overrightarrow{OP}$ ,  $\mathbf{x}_o = \overrightarrow{OO'}$  e  $\mathbf{x}' = \overrightarrow{O'P}$ , e substituindo em (6.3.3) ou (6.3.5), vem que:

$$\begin{aligned} Q(\mathbf{x}') &= Q(\mathbf{x}_o + \mathbf{x}') = (\mathbf{x}_o + \mathbf{x}')^t \mathbf{A} (\mathbf{x}_o + \mathbf{x}') + 2(\mathbf{x}_o + \mathbf{x}')^t \mathbf{b} + C \\ &= \mathbf{x}'^t \mathbf{A} \mathbf{x}' + \mathbf{x}_o^t \mathbf{A} \mathbf{x}' + \mathbf{x}'^t \mathbf{A} \mathbf{x}_o + \mathbf{x}_o^t \mathbf{A} \mathbf{x}_o + 2\mathbf{x}_o^t \mathbf{b} + 2\mathbf{x}'^t \mathbf{b} + C \\ &= \mathbf{x}'^t \mathbf{A} \mathbf{x}' + (\mathbf{x}_o^t \mathbf{A} \mathbf{x}')^t + \mathbf{x}'^t \mathbf{A} \mathbf{x}_o + \mathbf{x}_o^t \mathbf{A} \mathbf{x}_o + 2\mathbf{x}_o^t \mathbf{b} + 2\mathbf{x}'^t \mathbf{b} + C \\ &= \mathbf{x}'^t \mathbf{A} \mathbf{x}' + 2\mathbf{x}'^t (\mathbf{A} \mathbf{x}_o + \mathbf{b}) + \mathbf{x}_o^t \mathbf{A} \mathbf{x}_o + 2\mathbf{x}_o^t \mathbf{b} + C \end{aligned} \quad (6.3.14)$$

Escrevendo  $Q(\mathbf{x}') \stackrel{\text{def}}{=} Q(\mathbf{x}) = Q(\mathbf{x}_o + \mathbf{x}')$  na forma (6.3.5):

$$Q(\mathbf{x}') = \mathbf{x}'^t \mathbf{A}' \mathbf{x}' + 2\mathbf{x}'^t \mathbf{b}' + C' \quad (6.3.15)$$

e comparando com (6.3.14), vem que:

$$\begin{aligned} \mathbf{A}' &= \mathbf{A} \\ \mathbf{b}' &= \mathbf{A} \mathbf{x}_o + \mathbf{b} \\ C' &= Q(\mathbf{x}_o) \end{aligned} \quad (6.3.16)$$

► **6.24** Mas podemos ainda escrever a translacção  $\mathbf{x} = \mathbf{x}_o + \mathbf{x}'$  na seguinte forma matricial:

$$\begin{aligned} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} &= \begin{bmatrix} \text{Id} & \mathbf{x}_o \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}' \\ 1 \end{bmatrix} \\ &= \mathbf{P} \begin{bmatrix} \mathbf{x}' \\ 1 \end{bmatrix} \end{aligned} \quad (6.3.17)$$

Substituindo directamente em (6.3.4) ou (6.3.6) vem que:

$$Q(\mathbf{x}') = \begin{bmatrix} \mathbf{x}' & 1 \end{bmatrix} \mathbf{P}^t \mathbf{B} \mathbf{P} \begin{bmatrix} \mathbf{x}' \\ 1 \end{bmatrix} \quad (6.3.18)$$

onde  $\mathbf{P} = \begin{bmatrix} \text{Id} & \mathbf{x}_o \\ \mathbf{0} & 1 \end{bmatrix}$ . De facto:

$$\mathbf{B}' = \begin{bmatrix} \mathbf{A}' & \mathbf{b}' \\ \mathbf{b}'^t & C' \end{bmatrix} = \mathbf{P}^t \mathbf{B} \mathbf{P} = \begin{bmatrix} \text{Id} & \mathbf{0} \\ \mathbf{x}_o^t & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^t & C \end{bmatrix} \begin{bmatrix} \text{Id} & \mathbf{x}_o \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{A} \mathbf{x}_o + \mathbf{b} \\ \mathbf{x}_o^t \mathbf{A} + \mathbf{b}^t & Q(\mathbf{x}_o) \end{bmatrix}$$

donde se deduz mais uma vez que:

$$\begin{aligned} \mathbf{A}' &= \mathbf{A} \\ \mathbf{b}' &= \mathbf{A} \mathbf{x}_o + \mathbf{b} \\ C' &= Q(\mathbf{x}_o) \end{aligned} \quad (6.3.19)$$

Note que  $\det \mathbf{P} = 1$ . Estas fórmulas permitem pois concluir que:

► **6.25 Teorema** ... A matriz  $\mathbf{A}$  dos termos quadráticos, o determinante e o rank da matriz  $\mathbf{B}$  permanecem invariantes quando trasladamos a origem das coordenadas:

$$\mathbf{A}' = \mathbf{A}, \quad \det \mathbf{B}' = \det \mathbf{B}, \quad \text{rank } \mathbf{B}' = \text{rank } \mathbf{B} \quad (6.3.20)$$

► **6.26 Centro** ... Uma cónica (ou uma quádrica) diz-se **central** se  $\det \mathbf{A} \neq 0$ . Neste caso, existe um único ponto  $\mathbf{x}_o$ , chamado o **centro** da quádrica, tal que:

$$\mathbf{b}' = \mathbf{A}\mathbf{x}_o + \mathbf{b} = \mathbf{0}$$

De facto, basta pôr  $\mathbf{x}_o = -\mathbf{A}^{-1}\mathbf{b}$  e, com esta escolha para a origem do novo referencial acima referido, a equação (6.3.15) fica na forma:

$$Q(\mathbf{x}) = \mathbf{x}'^t \mathbf{A}' \mathbf{x}' + C \quad (6.3.21)$$

► **6.27 Centro de uma cónica** ... Um ponto  $O' = (x_o, y_o)$  diz-se um centro da cónica (6.3.1), se:

$$\mathbf{b}' = \mathbf{A}\mathbf{x}_o + \mathbf{b} = \mathbf{0} \quad (6.3.22)$$

isto é:

$$\begin{cases} ax_o + cy_o + d = 0 \\ by_o + cx_o + e = 0 \end{cases} \quad (6.3.23)$$

Um centro é pois uma intersecção das rectas dadas pelas equações:

$$\begin{cases} ax + cy + d = 0 \\ by + cx + e = 0 \end{cases} \quad (6.3.24)$$

e portanto podem ocorrer 3 hipóteses:

- As rectas intersectam-se num único ponto. A cónica tem pois um único centro e diz-se então uma **cónica central**. Isto acontece quando:

$$\delta = \begin{vmatrix} a & c \\ c & b \end{vmatrix} \neq 0 \quad (6.3.25)$$

- As rectas são paralelas e não se intersectam. Neste caso a cónica não tem centro. Isto acontece quando:

$$\delta = \begin{vmatrix} a & c \\ c & b \end{vmatrix} = 0 \quad \text{e} \quad \Delta = \begin{vmatrix} a & c & d \\ c & b & e \\ d & e & f \end{vmatrix} \neq 0 \quad (6.3.26)$$

- As rectas coincidem. Neste caso a cónica tem uma recta de centros. Isto acontece quando:

$$\delta = \begin{vmatrix} a & c \\ c & b \end{vmatrix} = 0 \quad \text{e} \quad \Delta = \begin{vmatrix} a & c & d \\ c & b & e \\ d & e & f \end{vmatrix} = 0 \quad (6.3.27)$$

► **6.28** Quando a cónica é central, devemos escolher a nova origem  $O'$  do referencial  $\mathcal{R}'$ , coincidente com esse centro. Neste caso os termos lineares anulam-se e a equação da cónica, nas novas coordenadas  $x', y'$  é:

$$Q(x', y') = ax'^2 + by'^2 + 2cx'y' + Q(x_o, y_o) = 0 \quad (6.3.28)$$

► **6.29 Efeito da mudança de base ortonormada** ... Escolhamos agora uma nova base ortonormada  $\mathcal{B} = \{\mathbf{u}_i\}$ . Nesta nova base, a matriz de Gram  $\mathbf{A}$ , que representa a parte quadrática  $\mathbf{x}^t \mathbf{A} \mathbf{x}$ , transforma-se, como sabemos, da seguinte forma:

$$\mathcal{C} \rightarrow \mathcal{C}P = \mathcal{B} \quad \Rightarrow \quad \mathbf{A} \rightarrow P^t \mathbf{A} P$$

enquanto que  $\mathbf{b}$  transforma-se como um vector:

$$\mathcal{C} \rightarrow \mathcal{C}P = \mathcal{B} \quad \Rightarrow \quad \mathbf{b} \rightarrow P^t \mathbf{b}$$

(recorde que  $P$  é uma matriz ortogonal:  $P^{-1} = P^t$ ).

Portanto a função quadrática, que nas  $x$ -coordenadas (relativamente à base  $\{O; \mathbf{e}_i\}$ ) foi escrita na forma:

$$Q(\mathbf{x}) = \mathbf{x}^t \mathbf{A} \mathbf{x} + 2\mathbf{x}^t \mathbf{b} + C$$

escreve-se agora, nas  $x'$ -coordenadas, relativamente à base  $\{O; \mathbf{u}_i\}$ , na forma:

$$Q(\mathbf{x}') = \mathbf{x}'^t (P^t \mathbf{A} P) \mathbf{x}' + 2\mathbf{x}'^t (P^t \mathbf{b}) + C \quad (6.3.29)$$

isto é:

$$\begin{aligned} \mathbf{A}' &= P^t \mathbf{A} P \\ \mathbf{b}' &= P^t \mathbf{b} \\ C' &= C \end{aligned} \quad (6.3.30)$$

► **6.30** Mas mais uma vez podemos usar a escrita matricial. Vem então que:

$$\begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} = \begin{bmatrix} P & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}' \\ 1 \end{bmatrix} \quad (6.3.31)$$

Substituindo directamente em (6.3.4) ou (6.3.6) vem que:

$$\begin{aligned} Q(\mathbf{x}) &= \begin{bmatrix} \mathbf{x} & | & 1 \end{bmatrix} \mathbf{B} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{x}' & | & 1 \end{bmatrix} \begin{bmatrix} P^t & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^t & C \end{bmatrix} \begin{bmatrix} P & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}' \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{x}' & | & 1 \end{bmatrix} \begin{bmatrix} P^t \mathbf{A} P & P^t \mathbf{b} \\ \mathbf{b}^t P & C \end{bmatrix} \begin{bmatrix} \mathbf{x}' \\ 1 \end{bmatrix} \end{aligned} \quad (6.3.32)$$

Destas fórmulas deduzimos o seguinte:

► **6.31 Teorema** ... *O determinante e o rank das matrizes  $\mathbf{A}$  e  $\mathbf{B}$  são invariantes sob mudanças de origem e de base ortonormada.*

## 6.4 Redução à forma canónica da equação geral de uma cónica

► **6.32** Consideremos de novo um referencial afim ortonormado  $\mathcal{R} = \{O; \mathbf{e}_1, \mathbf{e}_2\}$  e uma cónica afim em  $\mathbb{E}^2$  de equação:

$$Q(x, y) = ax^2 + by^2 + 2cxy + 2dx + 2ey + f = 0 \quad (6.4.1)$$

Designemos por  $\mathcal{C} = \{\mathbf{e}_1, \mathbf{e}_2\}$ . Sem mudar a origem, escolhamos agora uma nova base ortonormada  $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$ , constituída por vectores próprios associados aos valores próprios  $\lambda, \eta$  da matriz simétrica:

$$\mathbf{A} = \begin{bmatrix} a & c \\ c & b \end{bmatrix} \quad (6.4.2)$$

Nesta nova base, a parte quadrática  $ax^2 + by^2 + 2cxy$  reduz-se à forma diagonal. Mais detalhadamente, se:

$$\overrightarrow{OP} = x\mathbf{e}_1 + y\mathbf{e}_2 = x'\mathbf{u}_1 + y'\mathbf{u}_2$$

então:

$$Q(x', y') = \lambda(x')^2 + \eta(y')^2 + 2d'x' + 2e'y' + f = 0 \quad (6.4.3)$$

► **6.33** Distinguímos agora várias situações possíveis:

**1. Ambos os valores próprios são não nulos:**  $\lambda \neq 0$  e  $\eta \neq 0$ . Neste caso, completamos quadrados em (6.4.3):

$$\begin{aligned} Q(x', y') &= \lambda(x')^2 + \eta(y')^2 + 2d'x' + 2e'y' + f \\ &= \lambda \left( x' + \frac{d}{\lambda} \right)^2 - \frac{d^2}{\lambda} + \eta \left( y' + \frac{e}{\eta} \right)^2 - \frac{e^2}{\eta} + f \\ &= \lambda \left( x' + \frac{d}{\lambda} \right)^2 + \eta \left( y' + \frac{e}{\eta} \right)^2 + \left( f - \frac{d^2}{\lambda} - \frac{e^2}{\eta} \right) \end{aligned} \quad (6.4.4)$$

Transladamos então a origem para a nova origem através das fórmulas:

$$\begin{aligned} x' &= \tilde{x} - \frac{d}{\lambda} \\ y' &= \tilde{y} - \frac{e}{\eta} \end{aligned} \quad (6.4.5)$$

e a nova equação, nas coordenadas  $\tilde{x}, \tilde{y}$  fica na seguinte forma canónica:

$$\boxed{\lambda\tilde{x}^2 + \eta\tilde{y}^2 = C} \quad (6.4.6)$$

**2. Um dos valores próprios é nulo.** Por exemplo,  $\lambda \neq 0$  e  $\eta = 0$ .

Neste caso decompomos o vector  $\mathbf{b} = \begin{bmatrix} d \\ e \end{bmatrix}$  segundo a base ortonormada de vectores  $\mathbf{u}_1, \mathbf{u}_2$  associados aos valores próprios  $\lambda$  e  $\eta$ , respectivamente:

$$\mathbf{b} = \beta\mathbf{u}_1 - \mu\mathbf{u}_2 \quad (6.4.7)$$

A parte linear muda então como segue:

$$\begin{aligned} 2dx + 2ey &= 2\mathbf{b} \cdot \mathbf{x} \\ &= 2(\beta\mathbf{u}_1 - \mu\mathbf{u}_2) \cdot \mathbf{x} \\ &= 2\beta x' - 2\mu y' \end{aligned} \quad (6.4.8)$$

Nas coordenadas  $x', y'$  a equação da cónica fica então na forma:

$$\begin{aligned} Q(x', y') &= \lambda x'^2 + 2\beta x' - 2\mu y' + f \\ &= \lambda \left( x' + \frac{\beta}{\lambda} \right)^2 - \frac{\beta^2}{\lambda} - 2\mu y' + f \\ &= \lambda \left( x' + \frac{\beta}{\lambda} \right)^2 - 2\mu y' + \left( f - \frac{\beta^2}{\lambda} \right) \end{aligned} \quad (6.4.9)$$

- **2(i)**. Se  $\mu = 0$ , a equação fica:

$$\lambda \underbrace{\left(x' + \frac{\beta}{\lambda}\right)^2}_{\tilde{x}^2} + \underbrace{\left(f - \frac{\beta^2}{\lambda}\right)}_C = 0$$

isto é:

$$\boxed{\lambda \tilde{x}^2 = C} \quad (6.4.10)$$

- **2(ii)**. Se  $\mu \neq 0$ , a equação fica:

$$\lambda \underbrace{\left(x' + \frac{\beta}{\lambda}\right)^2}_{\tilde{x}^2} - 2\mu y' + \underbrace{\left(f - \frac{\beta^2}{\lambda}\right)}_C = \lambda \tilde{x}^2 - 2\mu \underbrace{\left(y' - \frac{C}{\mu}\right)}_{\tilde{y}} = 0$$

isto é:

$$\boxed{\lambda \tilde{x}^2 - 2\mu \tilde{y} = 0} \quad (6.4.11)$$

► **6.34 Resumindo** ... temos as 3 formas canónicas seguintes (omitindo os tildes):

$$\begin{array}{ll} \text{(I).} & \lambda x^2 + \eta y^2 = C, \quad \lambda, \eta \neq 0 \\ \text{(II).} & \lambda x^2 - 2\mu y = 0, \quad \lambda, \mu \neq 0 \\ \text{(III).} & \lambda x^2 = C, \quad \lambda \neq 0 \end{array}$$

Conforme os valores de  $\lambda, \eta, \mu$  e  $C$  temos as seguintes possibilidades (no campo real):

$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ $a \geq b > 0$ elipse	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$ $a \geq b > 0$ ponto
$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ $a > 0, b > 0$ hipérbole	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$ $a > 0, b > 0$ duas rectas
$y^2 = 2px$ $p > 0$ parábola	$y^2 - b^2 = 0$ $b > 0$ duas rectas paralelas distintas
	$y^2 = 0$ duas rectas paralelas iguais

► **6.35 Exemplo** ... Reduzir à forma canónica a cónica:

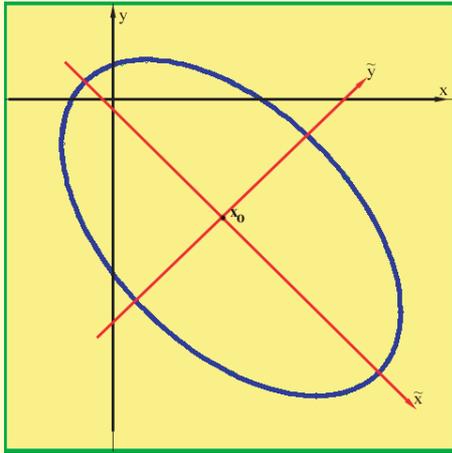
$$q(x, y) = x^2 + xy + y^2 - 3x + 4y - 5 = 0$$

Escrevendo na forma matricial, vem que:

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{x}^t \mathbf{A} \mathbf{x} + 2\mathbf{x}^t \mathbf{b} + c = 0 \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + 2 \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} -3/2 & 2 \end{bmatrix} - 5 \end{aligned} \quad (6.4.12)$$

Como  $\delta = \det \mathbf{A} = \det \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} = 3/4 \neq 0$ , a cónica é central de centro:

$$\mathbf{x}_o = -\mathbf{A}^{-1} \mathbf{b} = - \begin{bmatrix} 4/3 & -2/3 \\ -2/3 & 4/3 \end{bmatrix} \begin{bmatrix} -3/2 \\ 2 \end{bmatrix} = \begin{bmatrix} 10/3 \\ -11/3 \end{bmatrix}$$



Escolhendo o centro para nova origem, e relativamente às coordenadas  $x' = x - 10/3$ ,  $y' = y + 11/3$ , a cónica tem por equação:

$$\begin{aligned} Q(x', y') &= \mathbf{x}'^t \mathbf{A} \mathbf{x}' + Q(\mathbf{x}_0) \\ &= \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} - 52/3 \end{aligned}$$

Como  $\det \begin{bmatrix} 1 - \lambda & 1/2 \\ 1/2 & 1 - \lambda \end{bmatrix} = (1 - \lambda)^2 - 1/4 = 0$ , concluímos que os valores próprios de  $\mathbf{A}$  são  $\lambda = 1/2$ , e  $\lambda = 3/2$ .

A base  $\mathbf{u}_1 = \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$ ,  $\mathbf{u}_2 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$  é uma base de vectores próprios de  $\mathbf{A}$ . Como  $P = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$ , se representarmos as coordenadas relativas à base  $\mathbf{u}_1, \mathbf{u}_2$ , por  $\tilde{x}, \tilde{y}$ , então:

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

isto é:

$$\tilde{x} = \sqrt{2}/2 x' - \sqrt{2}/2 y' = \sqrt{2}/2 (x - 10/3) - \sqrt{2}/2 (y + 11/3)$$

$$\tilde{y} = \sqrt{2}/2 x' + \sqrt{2}/2 y' = \sqrt{2}/2 (x - 10/3) + \sqrt{2}/2 (y + 11/3)$$

e nas coordenadas  $\tilde{x}, \tilde{y}$  a cónica tem por equação canónica:

$$\frac{1}{2} \tilde{x}^2 + \frac{3}{2} \tilde{y}^2 - 52/3 = 0$$

ou ainda:

$$\frac{\tilde{x}^2}{(\sqrt{104}/3)^2} + \frac{\tilde{y}^2}{(\sqrt{104}/9)^2} = 1 \quad (6.4.13)$$

que é uma elipse de centro  $(-10/3, 11/3)$  e semi-eixos  $\sqrt{104}/3$  e  $\sqrt{104}/9$ .

► **6.36 Exemplo** ... Reduzir à forma canónica a cónica:

$$q(x, y) = 4x^2 - 4xy + y^2 - 2x - 14y + 7 = 0$$

A matriz de Gram da parte quadrática  $4x^2 - 4xy + y^2$  é:

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

cujos valores próprios são  $\lambda = 5, \eta = 0$ . Note que esta cónica não é central uma vez que  $\det \mathbf{A} = 0$ . O vector  $\mathbf{u}_1 = \frac{\sqrt{5}}{5}(1, 2)$  é um vector próprio associado ao valor próprio  $\eta = 0$ , enquanto que o vector  $\mathbf{u}_2 = \frac{\sqrt{5}}{5}(2, -1)$  é um vector próprio associado ao valor próprio  $\lambda = 5$ .  $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$  é uma base ortonormada na qual a parte quadrática se reduz à forma diagonal  $5(y')^2$ .

Decompomos agora o vector  $\mathbf{b} = (-2, -14)$  segundo a base  $\mathcal{B}$ :

$$\begin{aligned}\mathbf{b} &= (\mathbf{b} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{b} \cdot \mathbf{u}_2) \mathbf{u}_2 \\ &= \left( (-2, -14) \cdot \frac{\sqrt{5}}{5} (1, 2) \right) \mathbf{u}_1 + \left( (-2, -14) \cdot \frac{\sqrt{5}}{5} (2, -1) \right) \mathbf{u}_2 \\ &= -6\sqrt{5}\mathbf{u}_1 + 2\sqrt{5}\mathbf{u}_2\end{aligned}\tag{6.4.14}$$

A parte linear  $-2x - 14y$  muda então de acordo com:

$$\begin{aligned}-2x - 14y &= (-2, -14) \cdot \mathbf{x}, \quad \text{onde } \mathbf{x} = (x, y) \\ &= (-6\sqrt{5}\mathbf{u}_1 + 2\sqrt{5}\mathbf{u}_2) \cdot \mathbf{x} \\ &= -6\sqrt{5}x' + 2\sqrt{5}y'\end{aligned}\tag{6.4.15}$$

onde pusemos  $\mathbf{x} = (\mathbf{x} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{x} \cdot \mathbf{u}_2)\mathbf{u}_2 = x'\mathbf{u}_1 + y'\mathbf{u}_2$ .

**Resumindo** - relativamente às coordenadas  $(x', y')$  relativas à base ortonormada  $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$ ,  $q$  escreve-se na forma:

$$q(x', y') = 5(y')^2 - 6\sqrt{5}x' + 2\sqrt{5}y' + 7 = 0$$

ou ainda:

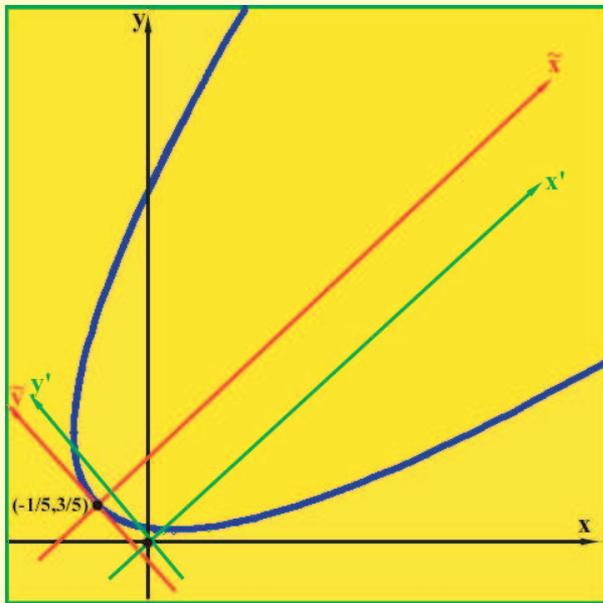
$$q(x', y') = (y')^2 - 6\frac{\sqrt{5}}{5}x' + 2\frac{\sqrt{5}}{5}y' + \frac{7}{5} = 0\tag{6.4.16}$$

Completando quadrados vem então que:

$$\begin{aligned}q(x', y') &= \left( y' + \frac{\sqrt{5}}{5} \right)^2 - \frac{1}{5} - 6\frac{\sqrt{5}}{5}x' + \frac{7}{5} \\ &= \underbrace{\left( y' + \frac{\sqrt{5}}{5} \right)^2}_{\tilde{y}^2} - \underbrace{6\frac{\sqrt{5}}{5} \left( x' - \frac{\sqrt{5}}{5} \right)}_{\tilde{x}} \\ &= \tilde{y}^2 - \frac{6\sqrt{5}}{5}\tilde{x} = 0\end{aligned}\tag{6.4.17}$$

que é da forma  $y^2 = 2px$ , e é portanto uma parábola de parâmetro  $p = \frac{3\sqrt{5}}{5}$ , com vértice no ponto de coordenadas  $(\tilde{x}_o, \tilde{y}_o) = (0, 0)$ , isto é, no ponto  $(x'_o, y'_o) = \left( \frac{\sqrt{5}}{5}, -\frac{\sqrt{5}}{5} \right)$ , ou ainda no ponto:

$$(x_o, y_o) = \left( -\frac{1}{5}, \frac{3}{5} \right)$$



► **6.37 Exemplo** ... Considere a cônica afim Euclideana  $\mathcal{C}$  em  $\mathbb{E}^2$ , definida por:

$$7x^2 + 7y^2 + 2xy + 4x - 20y - 4 = 0$$

- a.) Verifique se  $\mathcal{C}$  é central e, em caso afirmativo, calcule o seu centro.  
 b.) Reduza  $\mathcal{C}$  à forma canónica e identifique a cónica  $\mathcal{C}$ .  
 c.) Calcule as coordenadas do(s) foco(s) de  $\mathcal{C}$  relativamente ao referencial original  $\{O; x, y\}$

**Resolução ...**

a.) Escrevendo na forma matricial, vem que:

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{x}^t \mathbf{A} \mathbf{x} + 2\mathbf{x}^t \mathbf{b} + c = 0 \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 7 & 1 \\ 1 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + 2 \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 \\ -10 \end{bmatrix} - 4 = 0 \end{aligned} \quad (6.4.18)$$

Como  $\delta = \det \mathbf{A} = \det \begin{bmatrix} 7 & 1 \\ 1 & 7 \end{bmatrix} = 48 \neq 0$ , a cónica é central de centro:

$$\mathbf{x}_o = -\mathbf{A}^{-1} \mathbf{b} = -\frac{1}{48} \begin{bmatrix} 7 & -1 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -10 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 3/2 \end{bmatrix}$$

b.) Valores próprios da matriz  $\mathbf{A} = \begin{bmatrix} 7 & 1 \\ 1 & 7 \end{bmatrix}$ :  $\lambda = 6, 8$ .

Base ortonormada de vectores próprios:  $\mathcal{B} = \left\{ \mathbf{u}_1 = \frac{(1, -1)}{\sqrt{2}}, \mathbf{u}_2 = \frac{(1, 1)}{\sqrt{2}} \right\}$

A matriz de passagem da base canónica  $\mathcal{C}$  para a base  $\mathcal{B}$  é matriz ortogonal  $P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ .

O vector  $\mathbf{x}$  muda de acordo com:

$$\mathcal{C} \rightarrow \mathcal{B} = \mathcal{C}P \Rightarrow \mathbf{x}_{\mathcal{B}} = \tilde{\mathbf{x}} = \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = P^t \mathbf{x}_{\mathcal{C}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

e analogamente para o vector  $\mathbf{b}$ :

$$\mathcal{C} \rightarrow \mathcal{B} = \mathcal{C}P \Rightarrow \mathbf{b}_{\mathcal{B}} = \tilde{\mathbf{b}} = P^t \mathbf{b}_{\mathcal{C}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -10 \end{pmatrix} = \begin{pmatrix} 12/\sqrt{2} \\ -8/\sqrt{2} \end{pmatrix}$$

Na nova base  $\mathcal{B}$  a equação da cónica é:

$$\begin{aligned} q(\tilde{\mathbf{x}}) &= \tilde{\mathbf{x}}^t \mathbf{diag}(6, 8) \tilde{\mathbf{x}} + 2\tilde{\mathbf{x}}^t \tilde{\mathbf{b}} + c = 0 \\ &= \begin{bmatrix} \tilde{x} & \tilde{y} \end{bmatrix} \begin{bmatrix} 6 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} + 2 \begin{bmatrix} \tilde{x} & \tilde{y} \end{bmatrix} \begin{pmatrix} 12/\sqrt{2} \\ -8/\sqrt{2} \end{pmatrix} - 4 \\ &= 6\tilde{x}^2 + 8\tilde{y}^2 + 12\sqrt{2}\tilde{x} - 8\sqrt{2}\tilde{y} - 4 = 0 \end{aligned} \quad (6.4.19)$$

Completando quadrados vem que:

$$6(\tilde{x} + \sqrt{2})^2 + 8(\tilde{y} - \sqrt{2}/2)^2 = 20$$

ou:

$$\frac{X^2}{\left(\sqrt{\frac{20}{6}}\right)^2} + \frac{Y^2}{\left(\sqrt{\frac{20}{8}}\right)^2} = 1$$

onde pusemos  $X = \tilde{x} + \sqrt{2}$ ,  $Y = \tilde{y} - \sqrt{2}/2$ . A cônica é pois uma elipse com semieixos iguais a  $a = \sqrt{\frac{20}{6}}$  e  $b = \sqrt{\frac{20}{8}}$ . Como:

$$\begin{cases} \tilde{x} = \frac{x-y}{\sqrt{2}} \\ \tilde{y} = \frac{x+y}{\sqrt{2}} \end{cases}, \quad \text{e} \quad \begin{cases} X = \tilde{x} + \sqrt{2} \\ Y = \tilde{y} - \sqrt{2}/2 \end{cases}$$

vem que:

$$\begin{cases} X = \frac{x-y}{\sqrt{2}} + \sqrt{2} \\ Y = \frac{x+y}{\sqrt{2}} - \sqrt{2}/2 \end{cases}$$

A nova origem do referencial  $\{\tilde{O}; X, Y\}$  está situada no ponto cujas coordenadas  $x, y$  obtêm-se através de:

$$\begin{cases} \frac{x-y}{\sqrt{2}} + \sqrt{2} = 0 \\ \frac{x+y}{\sqrt{2}} - \sqrt{2}/2 = 0 \end{cases}$$

Resolvendo vem:

$$x = -1/2, \quad y = 3/2$$

que são exactamente as coordenadas  $x, y$  do centro da cônica.

Os focos da elipse estão situados nos pontos de coordenadas  $X, Y$  iguais, respectivamente,  $(\pm\sqrt{5/6}, 0)$ , uma vez que a distância semi-focal é dada por  $c = \sqrt{a^2 - b^2} = \sqrt{5/6}$ . As correspondentes coordenadas  $x, y$  obtêm-se resolvendo, em ordem a  $x$  e  $y$ , o sistema:

$$\begin{cases} \frac{x-y}{\sqrt{2}} + \sqrt{2} = \pm\sqrt{5/6} \\ \frac{x+y}{\sqrt{2}} - \sqrt{2}/2 = 0 \end{cases}$$

## Capítulo 7

# Quaterniões e Rotações

► **7.1 Quaterniões** ... Um **quaternião** é uma matriz  $2 \times 2$  da forma:

$$q = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}, \quad a, b, c, d \in \mathbb{R} \quad (7.0.1)$$

Podemos ainda escrevê-lo na forma:

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \quad (7.0.2)$$

usando as matrizes:

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathbf{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad (7.0.3)$$

Um quaternião da forma:

$$\mathbf{p} = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$$

diz-se **puro**.

► **7.2 Matrizes de Pauli** ... Note que as matrizes  $\mathbf{i}, \mathbf{j}$  e  $\mathbf{k}$  relacionam-se com as chamadas **matrizes de Pauli**,  $\sigma_1, \sigma_2, \sigma_3$ , definidas por:

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (7.0.4)$$

através das fórmulas:

$$\mathbf{i} = i\sigma_1, \quad \mathbf{j} = i\sigma_2, \quad \mathbf{k} = i\sigma_3 \quad (7.0.5)$$

► **7.3 Corpo  $\mathbb{H}$  dos quaterniões** ... É fácil mostrar que estas matrizes satisfazem as relações seguintes:

$$\begin{aligned} \mathbf{ij} &= \mathbf{k}, & \mathbf{jk} &= \mathbf{i}, & \mathbf{ki} &= \mathbf{j} \\ \mathbf{ij} &= -\mathbf{ji}, & \mathbf{jk} &= -\mathbf{kj}, & \mathbf{ik} &= -\mathbf{ki} \\ \mathbf{i}^2 &= \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1} \end{aligned} \quad (7.0.6)$$

com as quais é extremamente simples multiplicar dois quaterniões escritos na forma (7.0.2).

Com esta multiplicação o conjunto dos quaterniões fica munido de estrutura de corpo **não comutativo**, notado por  $\mathbb{H}$ , designado por **corpo dos quaterniões de Hamilton**.

É claro que, como espaço vectorial,  $\mathbb{H}$  é isomorfo a  $\mathbb{R}^4$ .

► **7.4 Conjugação. Norma** ... O **conjugado** de um quaterniã  $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$  define-se por:

$$\bar{q} = a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k} \quad (7.0.7)$$

A **norma de um quaterniã**  $q \in \mathbb{H}$ , escrito na forma (7.0.1) ou (7.0.2), nota-se por  $|q|$  e coincide, por definição, com a norma de  $q$  visto como um vector de  $\mathbb{R}^4$  com a sua estrutura Euclideana usual:

$$|q|^2 = a^2 + b^2 + c^2 + d^2 = \det q = \det \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix} \quad (7.0.8)$$

A distância entre dois quaterniões  $q_1, q_2 \in \mathbb{H}$  define-se como habitualmente à custa da norma, através de:

$$d(q_1, q_2) = |q_1 - q_2| \quad (7.0.9)$$

► **7.5 Propriedades** ... As propriedades seguintes são fáceis de verificar (fazer como exercício e notar a analogia com propriedades análogas familiares para números complexos):

(1).  $\overline{pq} = \bar{q}\bar{p}$

(2).  $q\bar{q} = \bar{q}q = |q|^2$

(3).  $q^{-1} = \frac{\bar{q}}{|q|^2}$ , se  $q \neq 0$

(4).  $|pq| = (\det(pq))^{1/2} = (\det p)^{1/2}(\det q)^{1/2} = |p||q|$

(5).  $|q^{-1}| = |q|^{-1}$

(6). Um quaterniã  $p$  é puro sse  $\bar{p} = -p$ . Designaremos por  $\mathcal{P}$  o subespaço dos quaterniões puros:

$$\mathcal{P} = \{\mathbf{p} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} : x, y, z \in \mathbb{R}\} \cong \mathbb{R}^3$$

que é pois isomorfo a  $\mathbb{R}^3$ . Um quaterniã puro  $\mathbf{p}$  será sempre identificado com o vector correspondente de  $\mathbb{R}^3$ .

(7). O produto de dois quaterniões puros não é, em geral, puro. De facto, é válida a seguinte fórmula:

$$\mathbf{p}\mathbf{q} = -(\mathbf{p} \cdot \mathbf{q})\mathbf{1} + \mathbf{p} \wedge \mathbf{q} \quad (7.0.10)$$

onde  $\mathbf{p} \cdot \mathbf{q}$  e  $\mathbf{p} \wedge \mathbf{q}$  representam, respectivamente, o produto interno e o produto vectorial usuais em  $\mathcal{P} \cong \mathbb{R}^3$ .

(8). Se  $\mathbf{p}, \mathbf{q} \in \mathcal{P}$  são dois quaterniões puros, então o seu anti-comutador  $\{\mathbf{p}, \mathbf{q}\}$ , é dado por:

$$\{\mathbf{p}, \mathbf{q}\} \stackrel{\text{def}}{=} \mathbf{p}\mathbf{q} + \mathbf{q}\mathbf{p} = -2(\mathbf{p} \cdot \mathbf{q}) \quad (7.0.11)$$

(9). Se  $\mathbf{p}, \mathbf{q} \in \mathcal{P}$  são dois quaterniões puros, então o seu comutador  $[\mathbf{p}, \mathbf{q}]$ , é dado por:

$$[\mathbf{p}, \mathbf{q}] \stackrel{\text{def}}{=} \mathbf{p}\mathbf{q} - \mathbf{q}\mathbf{p} = 2(\mathbf{p} \wedge \mathbf{q}) \quad (7.0.12)$$

Em particular o comutador de dois quaterniões puros é um quaterniã puro.

(10). Para cada  $q \in \mathbb{H}$ , a aplicação:

$$L_q : \mathbb{H} \rightarrow \mathbb{H}, \quad r \mapsto L_q(r) = qr$$

multiplica todas as distâncias por  $|q|$ . De facto:

$$\begin{aligned} d(L_q(r_1), L_q(r_2)) &= |qr_1 - qr_2| \\ &= |q(r_1 - r_2)| \\ &= |q||r_1 - r_2| \\ &= |q|d(r_1, r_2) \end{aligned} \tag{7.0.13}$$

Em particular, se  $|q| = 1$ , a aplicação  $L_q$  é uma isometria de  $\mathbb{H} \cong \mathbb{R}^4$ .

► **7.6 Rotações no espaço dos quaterniões puros** ... Como já vimos, o subespaço  $\mathcal{P}$  de  $\mathbb{H}$  constituído pelos quaterniões puros, é isomorfo a  $\mathbb{R}^3$ :

$$\mathcal{P} = \{\mathbf{p} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} : x, y, z \in \mathbb{R}\} \cong \mathbb{R}^3$$

O quaternião puro  $\mathbf{p} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$  será pois identificado ao vector  $\mathbf{p} = (x, y, z) \in \mathbb{R}^3$ . Em particular, os quaterniões puros  $\mathbf{i}, \mathbf{j}$  e  $\mathbf{k}$  são identificados aos vectores da base canónica de  $\mathbb{R}^3$ , designados pelas mesmas letras.

Se  $q \in \mathbb{H}$  é um quaternião arbitrário, tem-se que:

$$q\mathbf{p}q^{-1} \in \mathcal{P} \quad \text{sempre que } \mathbf{p} \in \mathcal{P} \tag{7.0.14}$$

De facto:

$$\overline{q\mathbf{p}q^{-1}} = \frac{1}{|q|^2} \overline{q\mathbf{p}\bar{q}} = \frac{1}{|q|^2} q \bar{\mathbf{p}} \bar{q} = -q\mathbf{p}q^{-1}$$

(recorde que um quaternião  $\mathbf{p}$  é puro sse  $\bar{\mathbf{p}} = -\mathbf{p}$ ).

► **7.7 Teorema** ... Se  $q \in \mathbb{H}$  a aplicação:

$$\begin{aligned} \mathcal{R}_q : \mathcal{P} \cong \mathbb{R}^3 &\longrightarrow \mathcal{P} \cong \mathbb{R}^3 \\ \mathbf{p} &\longmapsto q\mathbf{p}q^{-1} \end{aligned} \tag{7.0.15}$$

é uma isometria de  $\mathcal{P} \cong \mathbb{R}^3$ .

**Dem.:** Como já vimos em (7.0.14),  $\mathcal{R}_q(\mathcal{P}) \subseteq \mathcal{P}$ . É fácil ver que  $\mathcal{R}_q$  é linear. Por último, tem-se que:

$$\|\mathcal{R}_q(\mathbf{p})\| = |q\mathbf{p}q^{-1}| = |q||\mathbf{p}||q|^{-1} = \|\mathbf{p}\|, \quad \forall \mathbf{p} \in \mathcal{P} \cong \mathbb{R}^3$$

e portanto  $\mathcal{R}_q$  é uma isometria de  $\mathbb{R}^3$ .

► **7.8 Representação quaterniónica de uma simetria  $S_\pi$**  ... Consideremos agora uma simetria  $S_\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , relativamente ao plano  $\pi = \mathbf{n}^\perp$ , onde  $\mathbf{n}$  é um vector não nulo em  $\mathbb{R}^3 \cong \mathcal{P}$ . Como sabemos:

$$S_\pi(\mathbf{p}) = \mathbf{p} - 2 \frac{\mathbf{p} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n}, \quad \mathbf{p} \in \mathbb{R}^3 \cong \mathcal{P} \tag{7.0.16}$$

Mas, em  $\mathbb{H}$ , esta fórmula escreve-se na forma:

$$\begin{aligned}
 S_\pi(\mathbf{p}) &= \mathbf{p} - 2\frac{\mathbf{p} \cdot \mathbf{n}}{|\mathbf{n}|^2} \mathbf{n} \\
 &= \mathbf{p} + (\mathbf{p}\mathbf{n} + \mathbf{n}\mathbf{p})\frac{\mathbf{n}}{|\mathbf{n}|^2} \\
 &= \mathbf{p} - (\mathbf{p}\mathbf{n} + \mathbf{n}\mathbf{p})\mathbf{n}^{-1} \\
 &= -\mathbf{n}\mathbf{p}\mathbf{n}^{-1}
 \end{aligned} \tag{7.0.17}$$

onde usamos os factos seguintes:  $\mathbf{p}\mathbf{n} + \mathbf{n}\mathbf{p} = -2(\mathbf{p} \cdot \mathbf{n})$ ,  $\mathbf{p}, \mathbf{n} \in \mathcal{P}$  (ver (7.0.11)),  $\mathbf{n}^{-1} = \bar{\mathbf{n}}/|\mathbf{n}|^2$  e  $\bar{\mathbf{n}} = -\mathbf{n}$  para um quaterniãõ puro.

**Concluindo:** a simetria  $S_\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  relativamente ao plano  $\pi = \mathbf{n}^\perp$ , pode ser escrita na forma:

$$S_\pi(\mathbf{p}) = -\mathbf{n}\mathbf{p}\mathbf{n}^{-1} \tag{7.0.18}$$

### ► 7.9 Teorema ...

(1). Qualquer rotação de  $\mathbb{R}^3 \cong \mathcal{P}$  pode ser representada na forma:

$$\mathcal{R}_q : \mathbf{p} \mapsto \mathcal{R}_q(\mathbf{p}) = q\mathbf{p}q^{-1}, \quad \mathbf{p} \in \mathcal{P} \cong \mathbb{R}^3 \tag{7.0.19}$$

onde  $q \in \mathbb{H}$  é um quaterniãõ não nulo.

(2).

$$\mathcal{R}_q = \mathcal{R}_{q'} \quad \text{se e só se} \quad q = \lambda q', \quad \lambda \in \mathbb{R} - \{0\} \tag{7.0.20}$$

(3).

$$\boxed{\mathcal{R}_q \circ \mathcal{R}_{q'} = \mathcal{R}_{qq'}} \tag{7.0.21}$$

**Dem.:** Começemos com (7.0.21):

$$\mathcal{R}_q \circ \mathcal{R}_{q'}(\mathbf{p}) = q(q'\mathbf{p}q'^{-1})q^{-1} = (qq')\mathbf{p}(qq')^{-1} = \mathcal{R}_{qq'}(\mathbf{p}), \quad \forall \mathbf{p} \in \mathcal{P}$$

Quanto a (7.0.20):

$$\mathcal{R}_q(\mathbf{p}) = \mathcal{R}_{q'}(\mathbf{p}), \quad \forall \mathbf{p} \in \mathcal{P} \quad \Leftrightarrow \quad q\mathbf{p}q^{-1} = q'\mathbf{p}q'^{-1} \quad \Leftrightarrow \quad (q'^{-1}q)\mathbf{p} = \mathbf{p}(q'^{-1}q)$$

o que significa que  $q'^{-1}q$  comuta com todo o quaterniãõ puro. Como  $q'^{-1}q$  também comuta com  $\mathbb{R}\mathbf{1}$ , tem-se que  $q'^{-1}q = \lambda$ , para algum escalar  $\lambda \neq 0$ . Portanto  $q = \lambda q'$ .

Finalmente, para demonstrar a parte (1.), basta atender ao facto de que uma rotação de  $\mathbb{R}^3$  é um produto de duas simetrias relativamente a planos de  $\mathbb{R}^3$ . De facto, por (7.0.18), tem-se que:

$$\begin{aligned}
 S_\pi \circ S_{\pi'} &= S_\pi(-\mathbf{n}\mathbf{p}\mathbf{n}'^{-1}) \\
 &= -\mathbf{n}(-\mathbf{n}\mathbf{p}\mathbf{n}'^{-1})\mathbf{n}^{-1} \\
 &= \mathbf{n}\mathbf{n}'\mathbf{p}(\mathbf{n}\mathbf{n}')^{-1} \\
 &= \mathcal{R}_q(\mathbf{p}), \quad \text{onde } q = \mathbf{n}\mathbf{n}'
 \end{aligned} \tag{7.0.22}$$

► **7.10 Exercício ...** Calcular, usando quaterniões, o simétrico do vector  $\mathbf{p} = (-1, 0, 2)$  relativamente ao plano  $\pi : x - y - 3z = 0$  de  $\mathbb{R}^3$

► **7.11 Teorema** ... Seja  $q = q_o + \mathbf{q}$  um quaterniãõ não nulo, onde  $q_o \in \mathbb{R}$  e  $\mathbf{q} \in \mathcal{P}$ . Então:

(1). o eixo da rotação  $\mathcal{R}_q$ , definida por (7.0.19), é a recta gerada por  $\mathbf{q}$ .

(2). o ângulo de rotação é  $\pi/2$  se  $q_o = 0$  e, quando  $q_o \neq 0$ , é o ângulo  $\theta > 0$  tal que:

$$\tan \frac{\theta}{2} = \frac{|\mathbf{q}|}{q_o} \quad (7.0.23)$$

**Dem.:** Para demonstrar (1.), basta verificar que  $\mathbf{q}$  fica invariante sob  $\mathcal{R}_q$ . De facto:

$$\mathcal{R}_q(\mathbf{q}) = q\mathbf{q}q^{-1} = (q_o\mathbf{1} + \mathbf{q})\mathbf{q}(q_o\mathbf{1} + \mathbf{q})^{-1} = \dots = \mathbf{q}$$

(verificar como exercício).

Veamos agora a parte (2.). Dados dois quaterniões puros  $\mathbf{p}, \mathbf{p}' \in \mathcal{P}$ , com a mesma norma, existe sempre um quaterniãõ  $u$  tal que  $\mathbf{p}' = u\mathbf{p}u^{-1}$  (porquê?). Aplicando esta observação aos quaterniões puros  $\mathbf{q}$  e  $\rho\mathbf{i}$ , onde escolhemos  $\rho > 0$  de tal forma a que  $|\mathbf{q}| = |\rho\mathbf{i}|$ , calculamos um quaterniãõ  $u$  tal que  $u\mathbf{q}u^{-1} = \rho\mathbf{i}$ . Como  $\mathcal{R}_{uqu^{-1}} = \mathcal{R}_u\mathcal{R}_q\mathcal{R}_u^{-1}$ , podemos limitarmo-nos ao caso em que  $\mathbf{q} = \rho\mathbf{i}$ , isto é, a uma rotação em torno do eixo gerado por  $\mathbf{i}$  (o eixo dos  $x$ 's).

Suponhamos então que  $q = q_o + \rho\mathbf{i}$ . Como  $q^{-1} = \frac{q_o - \rho\mathbf{i}}{q_o^2 + \rho^2}$ , vem que:

$$\mathcal{R}_q(\mathbf{j}) = \frac{1}{q_o^2 + \rho^2} (q_o + \rho\mathbf{i})\mathbf{j}(q_o - \rho\mathbf{i}) = \frac{q_o^2 - \rho^2}{q_o^2 + \rho^2}\mathbf{j} + \frac{2q_o\rho}{q_o^2 + \rho^2}\mathbf{k}$$

donde se deduz que o ângulo de rotação  $\theta$  satisfaz:

$$\cos \theta = \frac{q_o^2 - \rho^2}{q_o^2 + \rho^2}, \quad \sin \theta = \frac{2q_o\rho}{q_o^2 + \rho^2}$$

Para obter (7.0.23) basta atender à identidade trigonométrica:

$$\tan \frac{\theta}{2} = \frac{\sin \theta}{1 + \cos \theta} = \frac{\frac{2q_o\rho}{q_o^2 + \rho^2}}{1 + \frac{q_o^2 - \rho^2}{q_o^2 + \rho^2}} = \dots = \frac{|\mathbf{q}|}{q_o}$$

► **7.12 Nota** ... Podemos representar qualquer quaterniãõ  $q \in \mathbb{H}$  na **forma polar** seguinte:

$$q = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} \mathbf{n} \quad (7.0.24)$$

onde  $\mathbf{n}$  é um quaterniãõ puro de norma 1:  $|\mathbf{n}| = 1$ .

Note que  $\mathbf{n}$  satisfaz  $\mathbf{n}^2 = -1$  (porquê?). O quaterniãõ  $q = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} \mathbf{n}$  representa a rotação  $\mathcal{R}_{(\mathbf{n};\theta)}$  de eixo gerado por  $\mathbf{n}$  e ângulo  $\theta$  (no sentido directo).

► **7.13 Exemplos** ... Por exemplo, temos que:

$$q = \cos \frac{\theta}{2} \mathbf{1} + \sin \frac{\theta}{2} \mathbf{i} \longrightarrow \mathcal{R}_q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \quad (7.0.25)$$

$$q = \cos \frac{\theta}{2} \mathbf{1} + \sin \frac{\theta}{2} \mathbf{j} \longrightarrow \mathcal{R}_q = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \quad (7.0.26)$$

$$q = \cos \frac{\theta}{2} \mathbf{1} + \sin \frac{\theta}{2} \mathbf{k} \longrightarrow \mathcal{R}_q = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (7.0.27)$$

► **7.14 Exemplo** ... Considere as duas rotações seguintes:

- rotação  $\mathcal{R}_1$  de ângulo  $\pi/3$  em torno do eixo (orientado) gerado por  $\mathbf{u} = (-1, 1, 0)$ , no sentido directo.

- rotação  $\mathcal{R}_2$  de ângulo  $\pi/2$  em torno do eixo (orientado) gerado por  $\mathbf{v} = (1, 0, -1)$ , no sentido directo.

Calcular  $\mathcal{R}_1 \circ \mathcal{R}_2$  e  $\mathcal{R}_2 \circ \mathcal{R}_1$ .

**Res...** A rotação  $\mathcal{R}_1$  pode ser representada pelo quaternião  $q = q_o + \mathbf{q}$  com parte pura  $\mathbf{q} = \mathbf{u} = (-1, 1, 0) = -\mathbf{i} + \mathbf{j}$  e parte real:

$$q_o = |\mathbf{q}| \tan \frac{\theta}{2} = |-\mathbf{i} + \mathbf{j}| \tan \frac{\pi}{6} = \sqrt{2} \frac{\sqrt{3}}{3} = \frac{\sqrt{6}}{3}$$

Portanto:

$$\mathcal{R}_1 = \mathcal{R}_q, \quad \text{com} \quad q = \frac{\sqrt{6}}{3} - \mathbf{i} + \mathbf{j}$$

Analogamente a segunda rotação  $\mathcal{R}_2$  pode ser representada pelo quaternião  $q' = q'_o + \mathbf{q}'$  com parte pura  $\mathbf{q}' = \mathbf{v} = (1, 0, -1) = \mathbf{i} - \mathbf{k}$  e parte real:

$$q'_o = |\mathbf{q}'| \tan \frac{\theta'}{2} = |\mathbf{i} - \mathbf{k}| \tan \frac{\pi}{4} = \sqrt{2} \frac{\sqrt{2}}{2} = 1$$

Portanto:

$$\mathcal{R}_2 = \mathcal{R}_{q'}, \quad \text{com} \quad q' = 1 + \mathbf{i} - \mathbf{k}$$

Calculemos agora os produtos  $qq'$  e  $q'q$ :

$$\begin{aligned} qq' &= \left( \frac{\sqrt{6}}{3} - \mathbf{i} + \mathbf{j} \right) (1 + \mathbf{i} - \mathbf{k}) \\ &= \left( \frac{\sqrt{6}}{3} + 1 \right) + \left( \frac{\sqrt{6}}{3} - 2 \right) \mathbf{i} - \left( \frac{\sqrt{6}}{3} + 1 \right) \mathbf{k} \\ q'q &= (1 + \mathbf{i} - \mathbf{k}) \left( \frac{\sqrt{6}}{3} - \mathbf{i} + \mathbf{j} \right) \\ &= \dots \end{aligned} \tag{7.0.28}$$

Como:

$$\mathcal{R}_1 \circ \mathcal{R}_2 = \mathcal{R}_q \circ \mathcal{R}_{q'} = \mathcal{R}_{qq'}$$

vemos que a rotação  $\mathcal{R}_1 \circ \mathcal{R}_2$  é representada pelo quaternião:

$$qq' = \left( \frac{\sqrt{6}}{3} + 1 \right) + \left( \frac{\sqrt{6}}{3} - 2 \right) \mathbf{i} - \left( \frac{\sqrt{6}}{3} + 1 \right) \mathbf{k}$$

Logo trata-se de uma rotação em torno do eixo gerado por:

$$\left( \frac{\sqrt{6}}{3} - 2, 0, -\frac{\sqrt{6}}{3} - 1 \right)$$

e de ângulo  $\theta$  que satisfaz:

$$\tan \frac{\theta}{2} = \frac{\left( \left( \frac{\sqrt{6}}{3} + 1 \right)^2 + \left( \frac{\sqrt{6}}{3} - 2 \right)^2 + \left( \frac{\sqrt{6}}{3} + 1 \right)^2 \right)^{1/2}}{\frac{\sqrt{6}}{3} + 1}$$

► **7.15** O teorema 7.9 diz-nos que podemos sempre escrever uma rotação de  $\mathbb{R}^3$ , na forma  $R_s : \mathbf{p} \mapsto s\mathbf{p}s^{-1}$ , onde  $s$  é um quaternião de norma 1, multiplicando por um escalar se necessário.

Representemos por:

$$\mathcal{S} = \{s \in \mathbb{H} : |s| = 1\} \cong \mathbf{S}^3 \subset \mathbb{R}^4 \quad (7.0.29)$$

o conjunto dos quaterniões de norma 1. Como o produto de dois quaterniões de norma 1 é ainda um quaternião de norma 1, vemos que  $\mathcal{S}$  é um grupo (não comutativo).

Os teoremas anteriores mostram que temos um homorfismo sobrejectivo deste grupo sobre o grupo  $SO(3)$  das rotações de  $\mathbb{R}^3$ :

$$\mathcal{S} \rightarrow SO(3)$$

cujo núcleo é o subgrupo de ordem 2 em  $\mathcal{S}$ :  $\mathbf{Z}_2 = \{\pm \mathbf{1}\}$ . Isto significa que a cada rotação  $\phi \in SO(3)$  correspondem dois quaterniões opostos  $\pm s \in \mathcal{S}$ , de norma 1, tais que:

$$R_{\pm s} = \phi$$

FIM