

Módulo 7

ALGA I. Espaços vectoriais com produto interno

Contents

7.1	Espaços Euclidianos reais	92
7.2	Espaços Hermitianos (ou Unitários) complexos . . .	95
7.3	Norma	96
7.4	Ortogonalidade	98
7.5	Bases ortonormadas num espaço vectorial com produto interno	99
7.6	Método de ortogonalização de Gram-Schmidt	100
7.7	Decomposição	ortogonal.
	Teorema da aproximação óptima	103
7.8	Aplicações. Mínimos quadrados	108
7.9	Método dos mínimos quadrados. Aproximação de dados por uma recta	111
7.10	Transformações ortogonais e unitárias. Exemplos . .	112
7.11	Transformações unitárias em $\mathbb{C}^2$. Os grupos $\mathcal{U}(2)$ e $SU(2)$	114
7.12	Exercícios	114

7.1 Espaços Euclidianos reais

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

$$\begin{aligned} \langle | \rangle : \mathcal{V} \times \mathcal{V} &\longrightarrow \mathbb{R} \\ (\mathbf{u}, \mathbf{v}) &\longmapsto \langle \mathbf{u} | \mathbf{v} \rangle \end{aligned} \quad (7.1.1)$$

que satisfaz as três propriedades seguintes:

[PI1]. **é uma forma bilinear:**

$$\begin{aligned} \langle (\mathbf{u} + \mathbf{v}) | \mathbf{w} \rangle &= \langle \mathbf{u} | \mathbf{w} \rangle + \langle \mathbf{v} | \mathbf{w} \rangle \\ \langle \mathbf{u} | (\mathbf{v} + \mathbf{w}) \rangle &= \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{u} | \mathbf{w} \rangle \\ \langle \lambda \mathbf{u} | \mathbf{v} \rangle &= \langle \mathbf{u} | \lambda \mathbf{v} \rangle = \lambda \langle \mathbf{u} | \mathbf{v} \rangle \end{aligned} \quad (7.1.2)$$

[PI2]. **é uma forma simétrica:**

$$\langle \mathbf{u} | \mathbf{v} \rangle = \langle \mathbf{v} | \mathbf{u} \rangle \quad (7.1.3)$$

[PI3]. **é não degenerada:**

$$\langle \mathbf{u} | \mathbf{v} \rangle = 0 \quad \forall \mathbf{v} \in \mathcal{V} \quad \Rightarrow \quad \mathbf{u} = \mathbf{0} \quad (7.1.4)$$

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

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

$$\langle \mathbf{u} | \mathbf{u} \rangle \geq 0, \quad \forall \mathbf{u} \in \mathcal{V} \quad (7.1.5)$$

Um espaço vectorial real, munido de um produto interno Euclideo chama-se um **espaço Euclideo**. 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 $\mathbf{u} | \mathbf{v}$.

► **7.2 Exemplo [Produto interno Euclideo 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 (Euclideo)**, como sendo o escalar $\mathbf{x} \cdot \mathbf{y} \in \mathbb{R}$, dado por:

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &\stackrel{\text{def}}{=} \sum_{i=1}^n x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\ &= \mathbf{x}^t \mathbf{y} \quad \text{em notação matricial} \end{aligned} \quad (7.1.6)$$

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

► **7.3 Exemplo [Produto interno L^2 em $C^0([a, b], \mathbb{R})$]** ... Consideremos o espaço vectorial real constituído pelas funções contínuas reais, definidas no intervalo $[a, b] \subset \mathbb{R}$. Dadas duas funções $f, g \in C^0([a, b], \mathbb{R})$, define-se o respectivo **produto interno L^2** , como sendo o escalar $\langle f | g \rangle \in \mathbb{R}$, dado por:

$$\langle f | g \rangle \stackrel{\text{def}}{=} \int_a^b f(t)g(t) dt \quad (7.1.7)$$

► **7.4 Exemplo [Produto interno de Minkowski em $\mathbb{R}^4$]** ... Dados dois

vectores $\mathbf{x} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$ e $\mathbf{y} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix}$, 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:

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= -x_0 y_0 + x_1 y_1 + x_2 y_2 + x_3 y_3 \\ &= \begin{bmatrix} -x_0 & x_1 & x_2 & x_3 \end{bmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= \mathbf{x}^t \boldsymbol{\eta} \mathbf{y} \end{aligned} \quad (7.1.8)$$

onde η representa a matriz simétrica:

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (7.1.9)$$

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.

► **7.5 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:

$$\begin{aligned} \mathbf{v} &= \sum_i v^i \mathbf{e}_i \\ &= (\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n) \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} \\ &= \mathcal{C}[\mathbf{v}]_{\mathcal{C}} \end{aligned} \quad (7.1.10)$$

onde $[\mathbf{v}]_{\mathcal{C}} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$ 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 $\langle \mathbf{u} | \mathbf{v} \rangle$:

$$\begin{aligned} \langle \mathbf{u} | \mathbf{v} \rangle &= \left\langle \sum_i u^i \mathbf{e}_i \middle| \sum_j v^j \mathbf{e}_j \right\rangle \\ &= \sum_{i,j} u^i v^j \langle \mathbf{e}_i | \mathbf{e}_j \rangle \\ &= \sum_{i,j} g_{ij} u^i v^j \\ &= [\mathbf{u}]_{\mathcal{C}}^t G_{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}} \end{aligned} \quad (7.1.11)$$

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

$$g_{ij} \stackrel{\text{def}}{=} \langle \mathbf{e}_i | \mathbf{e}_j \rangle \quad (7.1.12)$$

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 $G_{\mathcal{E}}$ é definida positiva, isto é:

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

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

$n = 2$

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

$n = 3$

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

7.2 Espaços Hermitianos (ou Unitários) complexos

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

$$\begin{aligned} \langle | \rangle : \mathcal{V} \times \mathcal{V} &\longrightarrow \mathbb{C} \\ (\mathbf{u}, \mathbf{v}) &\longmapsto \langle \mathbf{u} | \mathbf{v} \rangle \end{aligned} \quad (7.2.1)$$

que satisfaz as propriedades seguintes:

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

$$\begin{aligned} \langle (\mathbf{u} + \mathbf{v}) | \mathbf{w} \rangle &= \langle \mathbf{u} | \mathbf{w} \rangle + \langle \mathbf{v} | \mathbf{w} \rangle \\ \langle \mathbf{u} | (\mathbf{v} + \mathbf{w}) \rangle &= \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{u} | \mathbf{w} \rangle \end{aligned} \quad (7.2.2)$$

$$\begin{aligned} \langle \lambda \mathbf{u} | \mathbf{v} \rangle &= \lambda \langle \mathbf{u} | \mathbf{v} \rangle \\ \langle \mathbf{u} | \lambda \mathbf{v} \rangle &= \bar{\lambda} \langle \mathbf{u} | \mathbf{v} \rangle \end{aligned} \quad (7.2.3)$$

[PH2]. **é uma forma Hermitiana:**

$$\langle \mathbf{u} | \mathbf{v} \rangle = \overline{\langle \mathbf{v} | \mathbf{u} \rangle} \quad (7.2.4)$$

[PH3]. **é não degenerada:**

$$\langle \mathbf{u} | \mathbf{v} \rangle = 0 \quad \forall \mathbf{v} \in \mathcal{V} \quad \Rightarrow \quad \mathbf{u} = \mathbf{0} \quad (7.2.5)$$

[PH4]. **é definida positiva:**

$$\langle \mathbf{u} | \mathbf{u} \rangle \geq 0 \quad (7.2.6)$$

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

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

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

► **7.7 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:

$$\begin{aligned} \langle \mathbf{z} | \mathbf{w} \rangle &\stackrel{\text{def}}{=} \sum_{i=1}^n z_i \bar{w}_i = z_1 \bar{w}_1 + z_2 \bar{w}_2 + \cdots + z_n \bar{w}_n \\ &= [z_1 \ z_2 \ \cdots \ z_n] \begin{pmatrix} \bar{w}_1 \\ \bar{w}_2 \\ \vdots \\ \bar{w}_n \end{pmatrix} \\ &= \mathbf{z}^t \bar{\mathbf{w}} \quad \text{em notação matricial} \end{aligned} \quad (7.2.7)$$

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

► **7.8 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:

$$\langle f | g \rangle \stackrel{\text{def}}{=} \int_a^b f(t) \overline{g(t)} dt \quad (7.2.8)$$

7.3 Norma

► **7.9 Definição [Norma]** ... Seja $(\mathcal{V}, \langle | \rangle)$ um espaço com um produto interno (*Euclideano 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:

$$\|\mathbf{v}\| \stackrel{\text{def}}{=} \sqrt{\langle \mathbf{v} | \mathbf{v} \rangle} \quad (7.3.1)$$

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

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

$$\|\mathbf{v}\| \geq 0 \quad \text{e} \quad \|\mathbf{v}\| = 0 \quad \text{sse} \quad \mathbf{v} = \mathbf{0} \quad (7.3.2)$$

[N2]. **é homogénea (positiva):**

$$\|\lambda \mathbf{v}\| = |\lambda| \|\mathbf{v}\| \quad (7.3.3)$$

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

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\| \quad (7.3.4)$$

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

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

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

$$|\langle \mathbf{v} | \mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|, \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V} \quad (7.3.5)$$

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:

$$\begin{aligned} 0 \leq \|\mathbf{u}\|^2 &= \left\langle \left(\mathbf{v} - \frac{\langle \mathbf{v} | \mathbf{w} \rangle}{\|\mathbf{w}\|^2} \mathbf{w} \right) \middle| \left(\mathbf{v} - \frac{\langle \mathbf{v} | \mathbf{w} \rangle}{\|\mathbf{w}\|^2} \mathbf{w} \right) \right\rangle \\ &= \langle \mathbf{v} | \mathbf{v} \rangle - \frac{\langle \mathbf{v} | \mathbf{w} \rangle \langle \mathbf{w} | \mathbf{v} \rangle}{\|\mathbf{w}\|^2} \\ &= \|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v} | \mathbf{w} \rangle|^2}{\|\mathbf{w}\|^2} \end{aligned} \quad (7.3.6)$$

o que demonstra a desigualdade.

► **7.12** Demonstremos agora a desigualdade triangular (7.3.4):

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v} | \mathbf{u} + \mathbf{v} \rangle \\ &= \langle \mathbf{u} | \mathbf{u} \rangle + \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{v} | \mathbf{u} \rangle + \langle \mathbf{v} | \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + \langle \mathbf{u} | \mathbf{v} \rangle + \overline{\langle \mathbf{u} | \mathbf{v} \rangle} + \|\mathbf{v}\|^2 \\ &= \|\mathbf{u}\|^2 + 2\operatorname{Re} \langle \mathbf{u} | \mathbf{v} \rangle + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u} | \mathbf{v} \rangle| + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2, \text{ por Cauchy-Schwarz (7.3.5)} \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

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

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

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^t \mathbf{x}} = \left(\sum_{i=1}^n (x_i)^2 \right)^{1/2} \quad (7.3.7)$$

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

$$\|\mathbf{z}\| = \sqrt{\mathbf{z}^t \mathbf{z}} = \left(\sum_{i=1}^n |z_i|^2 \right)^{1/2} \quad (7.3.8)$$

(iii). No espaço Unitário $C^o([a, b], \mathbb{C})$, munido do produto interno L^2 , dado por (7.2.8): $\langle f | g \rangle \stackrel{\text{def}}{=} \int_a^b f(t) \overline{g(t)} dt$, a norma de uma função $f \in C^o([a, b], \mathbb{C})$ é dada por:

$$\|f\| = \sqrt{\langle f | f \rangle} = \left(\int_a^b |f(t)|^2 dt \right)^{1/2} \quad (7.3.9)$$

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

$$\left| \int_a^b f(t) \overline{g(t)} dt \right| \leq \left(\int_a^b |f(t)|^2 dt \right)^{1/2} \left(\int_a^b |g(t)|^2 dt \right)^{1/2} \quad (7.3.10)$$

enquanto que a desigualdade triangular tem o aspecto seguinte:

$$\left(\int_a^b |f(t) + g(t)|^2 dt \right)^{1/2} \leq \left(\int_a^b |f(t)|^2 dt \right)^{1/2} + \left(\int_a^b |g(t)|^2 dt \right)^{1/2} \quad (7.3.11)$$

7.4 Ortogonalidade

► **7.14 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:

$$\langle \mathbf{u} | \mathbf{v} \rangle = 0 \quad (7.4.1)$$

► **7.15 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:

$$-1 \leq \frac{\langle \mathbf{u} | \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1 \quad (7.4.2)$$

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:

$$\cos \theta = \frac{\langle \mathbf{u} | \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \in [-1, 1] \quad (7.4.3)$$

Portanto:

$$\langle \mathbf{u} | \mathbf{v} \rangle = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta(\mathbf{u}, \mathbf{v}) \quad (7.4.4)$$

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 $\pi/2$.

► **7.16 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 :

$$S^\perp \stackrel{\text{def}}{=} \{ \mathbf{u} \in \mathcal{V} : \langle \mathbf{u} | \mathbf{s} \rangle = 0, \quad \forall \mathbf{s} \in S \} \quad (7.4.5)$$

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 $\mathbf{u} \in S^\perp$.

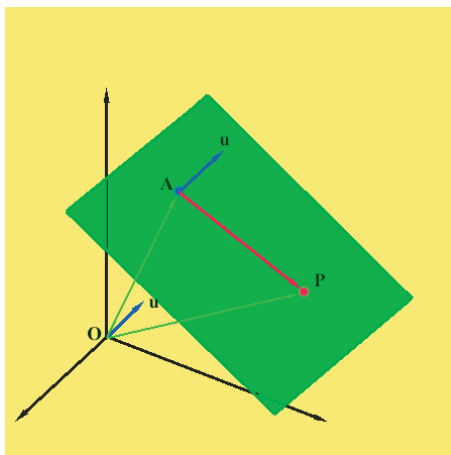
► **7.17 Hiperplanos vectoriais** ... No espaço Euclidiano $\mathbb{E}^n$, dado um vector não nulo $\mathbf{u} \in \mathbb{R}^n - \{\mathbf{0}\}$, o conjunto dos vectores $\mathbf{x} \in \mathbb{E}^n$ que são ortogonais a $\mathbf{u}$:

$$\{ \mathbf{x} \in \mathbb{E}^n : \mathbf{x} \cdot \mathbf{u} = 0 \} \quad (7.4.6)$$

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**:

$$\sum_i u_i x_i = u_1 x_1 + u_2 x_2 + \cdots + u_n x_n = 0 \quad (7.4.7)$$

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



No espaço Euclídeo $\mathbb{E}^n$, com a estrutura afim canónica, dado um ponto A e um vector 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 $\mathbf{u}$:

$$\{P \in \mathbb{E}^n : \overrightarrow{AP} \cdot \mathbf{u} = 0\} \quad (7.4.8)$$

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**:

$$\sum_i u_i x_i = u_1 x_1 + u_2 x_2 + \cdots + u_n x_n = c \quad (7.4.9)$$

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

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

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \quad (7.4.10)$$

Dem.:

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v} | \mathbf{u} + \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{v} | \mathbf{u} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \end{aligned} \quad (7.4.11)$$

7.5 Bases ortonormadas num espaço vectorial com produto interno

► **7.20 Definição [Base ortonormada]** ... Seja $(\mathcal{V}, \langle | \rangle)$ um espaço vectorial de dimensão n com um produto interno (Euclídeo 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:

$$\langle \mathbf{e}_i | \mathbf{e}_j \rangle = \delta_{ij} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \quad (7.5.1)$$

► **7.21 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 $\mathbf{v} \in \mathcal{V}$:*

$$\mathbf{v} = \sum_{i=1}^n \langle \mathbf{v} | \mathbf{e}_i \rangle \mathbf{e}_i \quad (7.5.2)$$

e:

$$\|\mathbf{v}\|^2 = \sum_{i=1}^n |\langle \mathbf{v} | \mathbf{e}_i \rangle|^2 \quad (7.5.3)$$

Dem.: Cálculo directo.

7.6 Método de ortogonalização de Gram-Schmidt

► **7.22 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 $\mathcal{V}$:

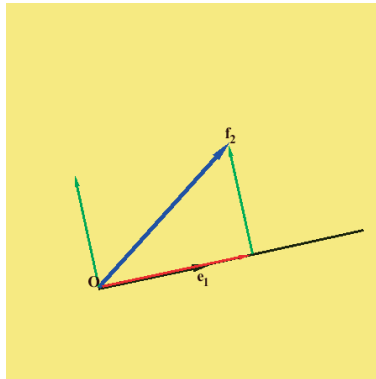
$$\langle \mathbf{e}_i | \mathbf{e}_j \rangle = 0, \quad i \neq j$$

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

[1.] Em primeiro lugar põmos:

$$\mathbf{e}_1 = \mathbf{f}_1 \quad (7.6.1)$$

[2.]



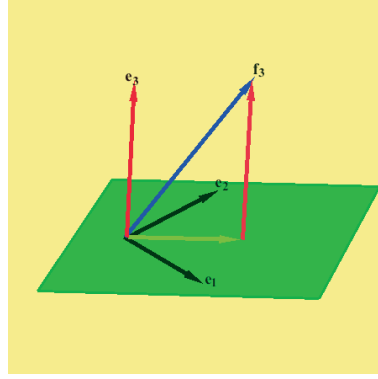
Em segundo lugar, começamos por calcular a chamada **projectção ortogonal de $\mathbf{f}_2$ sobre a recta gerada por $\mathbf{f}_1 = \mathbf{e}_1$** . Esta projectção ortogonal, por estar na recta gerada por $\mathbf{f}_1 = \mathbf{e}_1$, vai ser um vector do tipo $\lambda \mathbf{e}_1$, onde $\lambda \in \mathbb{K}$ 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:

$$\mathbf{e}_2 = \mathbf{f}_2 - \frac{\langle \mathbf{f}_2 | \mathbf{e}_1 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 \quad (7.6.2)$$

[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:

$$\mathbf{e}_3 = \mathbf{f}_3 - \frac{\langle \mathbf{f}_3 | \mathbf{e}_1 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 - \frac{\langle \mathbf{f}_3 | \mathbf{e}_2 \rangle}{\|\mathbf{e}_2\|^2} \mathbf{e}_2 \quad (7.6.3)$$

[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:

$$\mathbf{e}_{k+1} = \mathbf{f}_{k+1} - \sum_{i=1}^k \frac{\langle \mathbf{f}_{k+1} | \mathbf{e}_i \rangle}{\|\mathbf{e}_i\|^2} \mathbf{e}_i \quad (7.6.4)$$

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.

► **7.23 Polinômios 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 pômos:

$$\psi_0(t) = 1$$

Depois pômos:

$$\begin{aligned} \psi_1 &= t - \frac{\langle t|1 \rangle}{\|1\|^2} 1 \\ &= t - \frac{\int_{-1}^1 t dt}{\|\int_{-1}^1 1^2 dt\|^2} 1 \\ &= t \end{aligned} \quad (7.6.5)$$

Em seguida:

$$\begin{aligned} \psi_2 &= t^2 - \frac{\langle t^2|1 \rangle}{\|1\|^2} 1 - \frac{\langle t^2|t \rangle}{\|t\|^2} t \\ &= t^2 - \frac{\int_{-1}^1 t^2 dt}{\|\int_{-1}^1 1^2 dt\|^2} 1 - \frac{\int_{-1}^1 t^3 dt}{\|\int_{-1}^1 t^2 dt\|^2} t \\ &= t^2 - \frac{1}{3} \end{aligned} \quad (7.6.6)$$

e procedendo da mesma forma:

$$\begin{aligned} \psi_3 &= t^3 - \frac{3}{5}t \\ \psi_4 &= t^4 - \frac{6}{7}t^2 + \frac{3}{35} \\ &\vdots \end{aligned} \quad (7.6.7)$$

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

$$\begin{aligned} \varphi_0 &= \sqrt{\frac{1}{2}} \\ \varphi_1 &= \sqrt{\frac{3}{2}}t \\ \varphi_2 &= \frac{1}{2}\sqrt{\frac{5}{2}}(3t^2 - 1) \\ \varphi_3 &= \frac{1}{2}\sqrt{\frac{7}{2}}(5t^3 - 3t) \\ &\vdots \end{aligned} \quad (7.6.8)$$

7.7 Decomposição ortogonal. Teorema da aproximação óptima

► **7.24 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:

$$\mathcal{V} = \mathcal{S} \oplus \mathcal{S}^\perp \quad (7.7.1)$$

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

$$\mathbf{v} = \mathbf{s} + (\mathbf{v} - \mathbf{s}), \quad \text{onde } \mathbf{s} \in \mathcal{S} \text{ e } \mathbf{v} - \mathbf{s} \in \mathcal{S}^\perp \quad (7.7.2)$$

Além disso:

$$\|\mathbf{v}\|^2 = \|\mathbf{s}\|^2 + \|\mathbf{v} - \mathbf{s}\|^2 \quad (7.7.3)$$

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:

$$\mathbf{s} \stackrel{\text{def}}{=} \sum_{i=1}^m \langle \mathbf{v} | \mathbf{e}_i \rangle \mathbf{e}_i \quad (7.7.4)$$

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 (7.7.2).

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 (7.7.3) deduz-se do Teorema de Pitágoras (ver o teorema 7.19).

► **7.25 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:

$$\mathbf{P}_\mathcal{S} : \mathcal{V} \longrightarrow \mathcal{V} \quad (7.7.5)$$

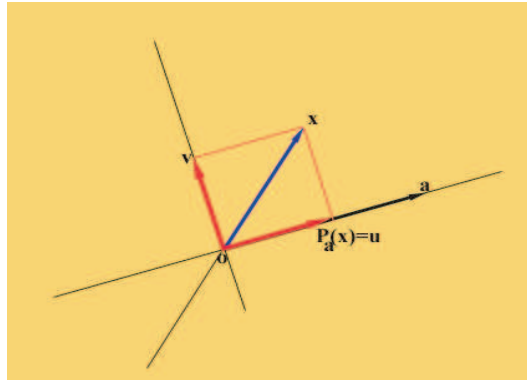
chamada a **projectão ortogonal sobre $\mathcal{S}$** da seguinte forma. Por definição de soma directa, todo o vector $\mathbf{v} \in \mathcal{V}$ admite uma decomposição única da forma: $\mathbf{v} = \mathbf{s} + (\mathbf{v} - \mathbf{s})$, onde $\mathbf{s} \in \mathcal{S}$ e $\mathbf{v} - \mathbf{s}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

- $\text{im } \mathbf{P}_\mathcal{S} = \mathcal{S}$
- $\ker \mathbf{P}_\mathcal{S} = \mathcal{S}^\perp$
- $\mathbf{P}_\mathcal{S}^2 = \mathbf{P}_\mathcal{S}$

- $\|\mathbf{P}_S(\mathbf{v})\| \leq \|\mathbf{v}\|$, $\forall \mathbf{v} \in \mathcal{V}$
- Se $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ é uma base ortonormada para $\mathcal{S}$, então:

$$\mathbf{P}_S(\mathbf{v}) = \sum_{i=1}^m \langle \mathbf{v} | \mathbf{e}_i \rangle \mathbf{e}_i \quad (7.7.6)$$

► 7.26 Exemplo [Projectão 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 **projectão ortogonal** de $\mathbf{x}$ sobre a recta gerada por $\mathbf{a}$, e é dado por:

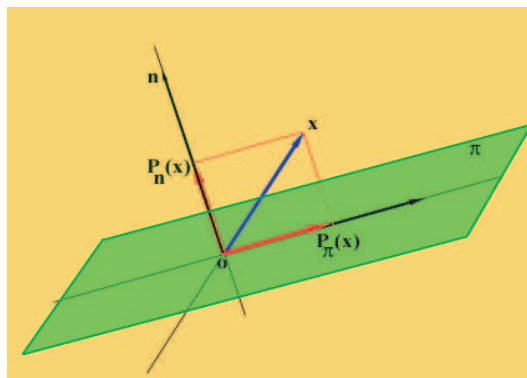
$$\mathbf{P}_a(\mathbf{x}) = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} \quad (7.7.7)$$

A aplicação $\mathbf{P}_a : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida por (7.7.7), 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 $\mathbf{a}$.

► 7.27 Exemplo [Projectão 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:

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

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

De acordo com (7.7.7), temos que:

$$\mathbf{P}_{\mathbf{n}^\perp} \stackrel{\text{def}}{=} \mathbf{x} - \mathbf{P}_{\mathbf{n}}(\mathbf{x}) = \mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n} \quad (7.7.8)$$

A aplicação $\mathbf{P}_{\mathbf{n}^\perp} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida por (7.7.8), 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:

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

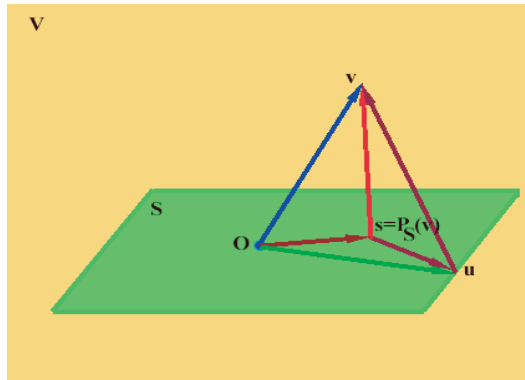
► **7.28 Teorema [da aproximação óptima]** ... Consideremos um espaço vectorial com um produto interno $(\mathcal{V}, \langle \cdot | \cdot \rangle)$ (Euclideano se $\mathcal{V}$ é real ou Hermitiano se $\mathcal{V}$ é complexo), e seja $\mathcal{S}$ um subespaço de dimensão finita. Dado um vector $\mathbf{v}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

$$\|\mathbf{v} - \mathbf{P}_{\mathcal{S}}(\mathbf{v})\| \leq \|\mathbf{v} - \mathbf{u}\|, \quad \forall \mathbf{u} \in \mathcal{S} \quad (7.7.9)$$

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 (7.7.2), 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 $\mathbf{s} = \mathbf{u}$.

► **7.29 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:

$$\mathcal{S}_n = \text{span}_{\mathbb{R}} \left\{ \varphi_0(t) = \frac{1}{\sqrt{2}}, \varphi_{2k-1}(t) = \frac{\cos kt}{\sqrt{\pi}}, \varphi_{2k}(t) = \frac{\sin kt}{\sqrt{\pi}} : k = 1, \dots, n \right\} \quad (7.7.10)$$

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 isto²).

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 (7.7.6), temos que:

$$\mathcal{F}_n(f) = \sum_{k=0}^{2n} \langle f|\varphi_k \rangle \varphi_k \quad (7.7.11)$$

onde:

$$\langle f|\varphi_k \rangle = \int_0^{2\pi} f(t)\varphi_k(t) dt \quad (7.7.12)$$

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

$$\mathcal{F}_n(f) = \frac{1}{2}a_0 + \sum_{k=1}^n (a_k \cos kt + b_k \sin kt) \quad (7.7.13)$$

onde os coeficientes de Fourier são dados por:

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt dt \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt dt \end{aligned} \quad (7.7.14)$$

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 (7.7.13), 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.

► **7.30 Exemplo** Seja $\mathcal{V} = C^o([-1, 1]; \mathbb{R})$ o espaço das funções reais contínuas definidas em $[-1, 1]$, munido do produto L^2 :

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

²Usar as relações trigonométricas seguintes:

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

e $\mathcal{S}_n$ o subespaço de dimensão $n + 1$ gerado pelos polinómios de Legendre normalizados, introduzidos no exemplo 7.23:

$$\mathcal{S}_n = \text{span}_{\mathbb{R}} \{\varphi_0, \varphi_1, \dots, \varphi_n\} \quad (7.7.15)$$

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 (7.7.6), temos que:

$$\mathbf{P}_n(f) = \sum_{k=0}^n \langle f | \varphi_k \rangle \varphi_k, \quad \text{onde} \quad \langle f | \varphi_k \rangle = \int_{-1}^1 f(t) \varphi_k(t) dt \quad (7.7.16)$$

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.

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

► **7.31 Exemplo ...** Considere o espaço vectorial $\mathbb{R}_3[t]$ das funções polinomiais $p(t)$, de grau ≤ 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:

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

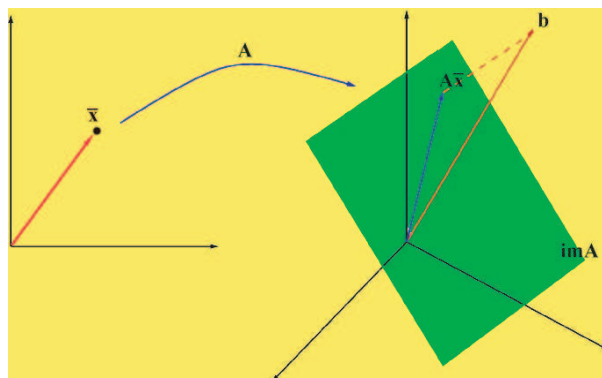
e portanto $\text{im } \mathbf{T} = \{P(t) = A + Bt + Ct^2 + Dt^3 \in \mathbb{R}_3[t] : 2A + C = 0\}$.

7.8 Aplicações. Mínimos quadrados

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

$$\mathbf{Ax} = \mathbf{b} \tag{7.8.1}$$

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 (7.8.1) é, por definição, um vector $\hat{\mathbf{x}}$, que satisfaz:

$$\|\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}\| \text{ é mínimo} \quad (7.8.2)$$

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 $\mathbf{A}\hat{\mathbf{x}}$ está mais próxima de $\mathbf{b}$.

► **7.33** 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:

$$\hat{\mathbf{x}} = \mathbf{A}^{-1} \mathbf{P}_{\text{im } A}(\mathbf{b}) \quad (7.8.3)$$

Isto é, para calcular a “solução” dos mínimos quadrados do sistema (7.8.1) procede-se da seguinte forma:

1. Calcula-se a projecção ortogonal $\hat{\mathbf{y}} = \mathbf{P}_{\text{im } A}(\mathbf{b}) \in \text{im } A$, 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:

$$\mathbf{A}\hat{\mathbf{x}} = \hat{\mathbf{y}} = \mathbf{P}_{\text{im } A}(\mathbf{b}) \quad (7.8.4)$$

► **7.34 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 $\mathbf{P}_{\text{im } \mathbf{A}}(\mathbf{b})$:

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

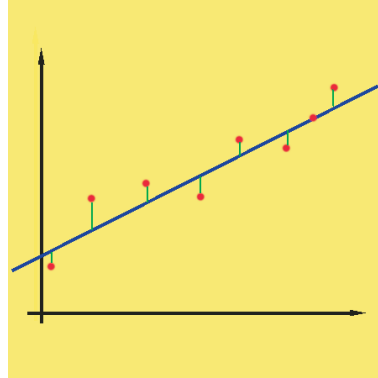
► **7.35 Exemplo ...** Calcular a “solução” dos mínimos quadrados do sistema:

$$\begin{cases} x + 2y = 1 \\ 3x - y + z = 0 \\ -x + 2y + z = -1 \\ x - y - 2z = 2 \\ 2x + y - z = 2 \end{cases} \quad (7.8.5)$$

e o erro correspondente.

7.9 Método dos mínimos quadrados. Aproximação de dados por uma recta

► 7.36 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:

t_1	t_2	t_3	$\dots$	t_n
y_1	y_2	y_3	$\dots$	y_n

(7.9.1)

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:

$$y = \alpha t + \beta \quad (7.9.2)$$

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:

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

Podemos reunir estas equações numa única em forma matricial:

$$\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{Ax} \quad (7.9.3)$$

onde:

$$\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$\boldsymbol{\varepsilon}$ é o chamado **vector de erro** e $\mathbf{y}$ o **vector de dados**. Os coeficientes α e β , 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 $\varepsilon_i = 0$, e poderíamos resolver o sistema $\mathbf{Ax} = \mathbf{y}$. 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:

$$\|\varepsilon\| = (\varepsilon_1^2 + \cdots + \varepsilon_n^2)^{1/2}$$

seja mínimo.

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

$$\|\varepsilon\| = \|\mathbf{Ax} - \mathbf{y}\|$$

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

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

t_i	0	1	3	6
y_i	2	3	7	12

(7.9.4)

Solução: $y = 12/7(1 + t)$.

7.10 Transformações ortogonais e unitárias. Exemplos

► **7.38 Definição ...** [Transformações ortogonais] ... Seja $(\mathcal{V}, \langle | \rangle)$ um espaço Euclideano de dimensão n , isto é, um espaço vectorial real com um produto interno Euclideano. 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.:

$$\langle \mathbf{A}(\mathbf{v}) | \mathbf{A}(\mathbf{w}) \rangle = \langle \mathbf{v} | \mathbf{w} \rangle \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V} \quad (7.10.1)$$

Se A é a matriz de uma tal transformação ortogonal, **relativamente a uma base ortonormada** de $\mathcal{V}$, então (7.10.1) escreve-se na seguinte forma matricial:

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

ou ainda:

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

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

$$A^t A = \mathbf{I} \quad (7.10.2)$$

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

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

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

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

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 $\mathcal{SO}(n)$:

$$\begin{aligned}\mathcal{O}(n) &= \{A \in \mathcal{M}_n(\mathbb{R}) : A^t A = \mathbf{I}\} \\ \mathcal{SO}(n) &= \{A \in \mathcal{M}_n(\mathbb{R}) : A^t A = \mathbf{I}, \text{ e } \det A = 1\}\end{aligned}\quad (7.10.3)$$

► **7.39 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.:

$$\langle \mathbf{A}(\mathbf{v}) | \mathbf{A}(\mathbf{w}) \rangle = \langle \mathbf{v} | \mathbf{w} \rangle \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V} \quad (7.10.4)$$

Se A é a matriz de uma tal transformação unitária, **relativamente a uma base ortonormada** de $\mathcal{V}$, então (7.11.1) 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:

$$A^t \overline{A} = \mathbf{I} \quad (7.10.5)$$

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

$$A^\dagger = \overline{A}^t \quad (7.10.6)$$

Portanto A é unitária sse:

$$AA^\dagger = \mathbf{I} \quad (7.10.7)$$

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 $\mathcal{SU}(n)$:

$$\begin{aligned}\mathcal{U}(n) &= \{A \in \mathcal{M}_n(\mathbb{C}) : A^\dagger A = \mathbf{I}\} \\ \mathcal{SU}(n) &= \{A \in \mathcal{M}_n(\mathbb{C}) : A^\dagger A = \mathbf{I}, \text{ e } \det A = 1\}\end{aligned}\quad (7.10.8)$$

7.11 Transformações unitárias em $\mathbb{C}^2$. Os grupos $\mathcal{U}(2)$ e $\mathcal{SU}(2)$

► **7.40** 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.:

$$\langle \mathbf{A}(\mathbf{z}) | \mathbf{A}(\mathbf{w}) \rangle = \langle \mathbf{z} | \mathbf{w} \rangle \quad \forall \mathbf{z}, \mathbf{w} \in \mathbb{C}^2 \quad (7.11.1)$$

Se A é a matriz de uma tal transformação unitária, relativamente à base canónica de $\mathbb{C}^2$, então (7.11.1) 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.:

$$A^t \overline{A} = \mathbf{I} \quad (7.11.2)$$

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:

$$AA^\dagger = \mathbf{I} \quad (7.11.3)$$

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

► **7.41** 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 $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ 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:

$$A^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} = A^\dagger = \begin{pmatrix} \overline{\alpha} & \overline{\gamma} \\ \overline{\beta} & \overline{\delta} \end{pmatrix}$$

isto é: $\delta = \overline{\alpha}$ e $\gamma = -\overline{\beta}$. Portanto $\mathcal{SU}(2)$ é o grupo das matrizes que são da forma:

$$A = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} \quad e \quad \det A = |\alpha|^2 + |\beta|^2 = 1 \quad (7.11.4)$$

7.12 Exercícios

► **Exercício 7.1** ... Verifique quais das seguintes funções são produtos internos Euclidianos em $\mathbb{R}^2$ ou $\mathbb{R}^3$:

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

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

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

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

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

► **Exercício 7.2** ... Calcule em cada caso $\langle \mathbf{u}, \mathbf{v} \rangle$ usando o produto interno Euclidiano usual e o produto interno definido em 7.1-a). Depois, calcule $\|\mathbf{u}\|$ e $\|\mathbf{v}\|$ recorrendo também a cada um desses dois produtos internos.

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

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

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

► **Exercício 7.3** ... Calcule em cada caso $\langle \mathbf{u}, \mathbf{v} \rangle$ usando o produto interno euclidiano usual e o produto interno definido em 7.1-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 7.4** ... 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 7.5** ... 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 7.6** ... 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 7.7** ... Em cada caso, determine o cos do ângulo θ entre os vetores $\mathbf{u}$ 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 7.1-d).

b) $\mathbf{u} = 2t - 1$, $\mathbf{v} = t^2$ em $\mathbb{R}(t)$, usando o produto interno Euclidiano definido no exercício 7.14.

► **Exercício 7.8** ... 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 7.9** ... 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 7.10** ... 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 7.11** ... 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 7.12** ... 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 7.13** ... 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 7.14** ... 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 7.15** ... Seja E um espaço vectorial no qual está definido um produto escalar. Mostre que :

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

► **Exercício 7.16** ... Em cada um dos casos, determine uma base ortonormada do subespaço de $\mathbb{R}^3$ gerado pelos seguintes vectores:

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

$$\text{b) } \mathbf{x}_1 = (1, 1, 1), \quad \mathbf{x}_2 = (-1, 1, -1), \quad \mathbf{x}_3 = (1, 0, 1).$$

► **Exercício 7.17** ... Em cada um dos casos, determine uma base ortonormada do subespaço de $\mathbb{R}^4$ gerado pelos seguintes vectores:

$$\text{a) } \mathbf{x}_1 = (1, 1, 0, 0), \quad \mathbf{x}_2 = (0, 1, 1, 0), \quad \mathbf{x}_3 = (0, 0, 1, 1), \quad \mathbf{x}_4 = (1, 0, 0, 1).$$

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

► **Exercício 7.18** ... 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 7.19** ... 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 7.20** ... 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 7.21** ...

► **Exercício 7.22** ... Considere o espaço vectorial real $\mathbb{R}_2(t)$ no qual está definido o produto interno $\langle f, g \rangle = \int_0^1 f(t)g(t) dt$.

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

► **Exercício 7.23** ... Seja V o espaço linear das matrizes 2×2 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 7.24** ... 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 7.25** ... 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 7.26** ... 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 7.27** ... Determine as projecções ortogonais seguintes:

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