Capítulo 9

Produto interno

Neste[1] Capítulo 9 consideramos exclusivamente espaços vetoriais reais

$$(X, +, \cdot, \mathbb{R}), \qquad (E, +, \cdot, \mathbb{R}), \quad n := \dim E < \infty.$$

O uso das letras E, F, G sinaliza dimensão finita enquanto $\dim X \in \mathbb{N}_0 \cup \{\infty\}$.

Produtos internos $\langle \cdot, \cdot \rangle$ só são definidos para espaços vetoriais *reais* ($\mathbb{K} = \mathbb{R}$). Pode-se definir produtos similares, mas não iguais, em espaços vetoriais complexos ($\mathbb{K} = \mathbb{C}$). Eles são chamados de produtos hermitianos, veja Capítulo 13.

9.1 Produto interno, norma, métrica

Definição 9.1.1 (Produto interno)**.** Um **produto interno**[2] num espaço vetorial real X é uma função de duas variáveis

$$\langle \cdot, \cdot \rangle \colon X \times X \to \mathbb{R}, \quad (u, v) \mapsto \langle u, v \rangle$$

a qual satisfaz os três axiomas

(SIM) $\langle u, v \rangle = \langle v, u \rangle$ (simetria)

(BL) $\langle u + \tilde{u}, v \rangle = \langle u, v \rangle + \langle \tilde{u}, v \rangle, \quad \langle \alpha u, v \rangle = \alpha \langle u, v \rangle$ (bi-linearidade)[3]

(POS) $u \neq \mathcal{O} \Rightarrow \langle u, u \rangle > 0$ (positividade)

para todos os vetores $u, v, \tilde{u}, \tilde{v} \in X$ e escalares $\alpha \in \mathbb{R}$. Neste caso X é chamado de **espaço vetorial com produto interno**.

Lema 9.1.2. *Num espaço vetorial X com produto interno vale*

[1] Cap. 9 de MA327 2021-2, autor Joa Weber, atualizado: 2 de maio de 2024

[2] Produtos internos são também chamados de **produtos escalares**.

[3] Note que simetria implica linearidade também na segunda variável, por isso o nome para o axioma dois, abreviando bi-linearidade, é justificado.

(a) $\langle v, \mathcal{O}\rangle = 0 \quad \forall v \in X$;

(ND) $\langle u, v\rangle = \langle \tilde{u}, v\rangle \quad \forall v \in X \quad \Rightarrow \quad u = \tilde{u}$; (não-degenerado)

(ND)′ $\langle u, v\rangle = 0 \quad \forall v \in X \quad \Rightarrow \quad u = \mathcal{O}$. (não-degenerado)′

Demonstração. (a) Dado $v, w \in X$, vale $\langle v, w\rangle + \langle v, \mathcal{O}\rangle \overset{(\mathtt{BL})}{=} \langle v, w + \mathcal{O}\rangle = \langle v, w\rangle$.
(ND) Escrevendo a hipótese na forma $\langle u - \tilde{u}, v\rangle = 0 \ \forall v$, resta aplicar (ND)′.
(ND)′ Suponha por absurdo $u \neq \mathcal{O}$. A hipótese para $v = u$ mostra que $\langle u, u\rangle = 0$ em contradição ao axioma (POS). □

Lema 9.1.3 (Critério para dois operadores são iguais)**.** *Dado operadores* $A, B \in \mathcal{L}(E, F)$ *e bases* $\mathcal{U} = \{\xi_1, \ldots, \xi_n\}$ *e* $\mathcal{V} = \{\eta_1, \ldots, \eta_m\}$, *então*

$$A = B \qquad \Leftrightarrow \qquad \langle A\xi_j, \eta_i\rangle = \langle B\xi_j, \eta_i\rangle \quad \forall i, j.$$

Demonstração. "⇒" Trivial. "⇐" Axioma (BL) na primeira entrada de $\langle \cdot, \cdot\rangle$ e propriedade (ND) em Lema 9.1.2. □

Exemplo 9.1.4 (Produto euclidiano em $\mathbb{R}^n$)**.** A função

$$\langle \cdot, \cdot\rangle_0 \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

$$(x, y) \mapsto x_1 y_1 + \cdots + x_n y_n = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

é chamado de **produto euclidiano** em $\mathbb{R}^n$. O leitor pode verificar os 3 axiomas. Caso não especificamos diferentemente o $\mathbb{R}^n$ sempre será munido do produto euclidiano.

Exemplo 9.1.5 (Produto interno mediante integração)**.** No espaço vetorial $C^0([a, b])$ das funções reais continuas num intervalo $[a, b]$ integração

$$\langle f, g\rangle := \int_a^b f(x) g(x)\, dx \tag{9.1.1}$$

define um produto interno. Deixamos ao leitor verificar os 3 axiomas.

Exemplo 9.1.6 (Integração não dando produto interno)**.** No espaço vetorial $C^0(\mathbb{R})$ das funções reais continuas no $\mathbb{R}$ inteiro, integração, nem sobre $\mathbb{R}$, nem sobre um intervalo $[a, b]$,

$$\langle f, g\rangle_\infty := \int_{-\infty}^{\infty} f(x) g(x)\, dx, \qquad \langle f, g\rangle := \int_a^b f(x) g(x)\, dx$$

define um produto interno em $C^0(\mathbb{R})$, não.

O problema no primeiro caso são valores infinitos, por exemplo $\langle 2, 3\rangle_\infty = \infty$.

O problema no segundo caso é o axioma (POS). Seja $u \colon \mathbb{R} \to \mathbb{R}$ uma função continua a qual anula-se em $[a, b]$, mas não no complemento inteiro. Por exemplo, suponha $u(b + 1) = 1$. Assim $u \neq \mathcal{O}$ não é a função nula, mas a integral $\langle u, u\rangle = \int_a^b u(x)^2\, dx = 0$ não é positivo.

Exemplo 9.1.7 (Polinômios em $\mathbb{R}$ e integração sobre $[a,b]$)**.** Em contraste ao espaço $C^0(\mathbb{R})$, no espaço vetorial dos polinômios $\mathcal{P}(\mathbb{R})$ integração sobre $[a,b]$ produz um produto interno! Deixamos ao leitor mostrar que (9.1.1) realmente satisfaz o axioma (`POS`).

[Dica: Em quantos pontos um polinômio de grau n pode-se anular no máximo?]

Normas – norma induzida

Definição 9.1.8. Uma **norma** num espaço vetorial real X é uma função

$$|\cdot| : X \to [0,\infty), \quad x \mapsto |x|$$

a qual satisfaz os três axiomas

(`HOM`) $|\alpha x| = \alpha\, |x|$ (homogeneidade) $\Rightarrow |\mathcal{O}| = 0$

(Δ) $|x+y| \le |x| + |y|$ (desigualdade triangular)

(`POS`) $x \neq \mathcal{O} \Rightarrow |x| > 0$ (positividade)

para todos os vetores $x, y \in X$ e escalares $\alpha \in \mathbb{R}$. Neste caso X é chamado de **espaço vetorial normado**.

Definição 9.1.9 (Norma induzida)**.** Num espaço vetorial X com produto interno existe para cada um vetor v um número não-negativo

$$|v| = |v|_{\langle\cdot,\cdot\rangle} := \sqrt{\langle v, v\rangle} \ge 0$$

chamado de **norma induzida** de v, ou informalmente o "*comprimento*" do vetor. Um vetor de comprimento $|v| = 1$ é chamado de **vetor unitário**. A notação $\hat{v}$ sinaliza que trata-se de um vetor unitário.

Exercício 9.1.10. (a) A norma induzida é uma norma.
(b) Dado um vetor não-nulo u, então $\hat{u} := \frac{1}{|u|} u$ é um vetor unitário.

Métricas – métrica induzida

Definição 9.1.11. Uma **métrica**[4] num conjunto M é uma função

$$\begin{aligned} d\colon M \times M &\to [0,\infty) \\ (q,p) &\mapsto d(q,p) \end{aligned}$$

a qual satisfaz os três axiomas

(`SIM`) $d(q,p) = d(p,q)$ (simetria)

(Δ) $d(q,r) \le d(q,p) + d(p,r)$ (desigualdade triangular)

(`POS`) $d(q,q) = 0$ mas $q \neq p \Rightarrow d(q,p) > 0$ (positividade)

[4] Métricas são também chamadas de **funções distância** ou simplesmente **distâncias**.

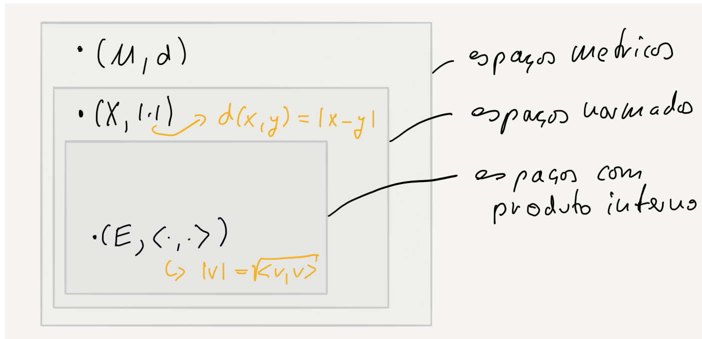

Figura 9.1: Produto interno → norma → função distância

para todos os pontos $q, p \in M$. Neste caso M é chamado de **espaço métrico**.

Definição 9.1.12 (Métrica induzida)**.** Num espaço vetorial normado a função

$$d(x, y) = d_{|\cdot|}(x, y) := |x - y|$$

é chamado de **métrica induzida** ou **distância** entre dois pontos.

Exercício 9.1.13. A métrica induzida $d(x, y) := |x - y|$ é uma métrica.

Então produtos internos disponibilizam normas e normas disponibilizam distâncias. As inclusões são ilustrados na Figura 9.1.

9.1.1 Produto interno e espaço dual – dualidade

Um produto interno $\langle\cdot,\cdot\rangle$ num espaço vetorial E com $\dim E = n$ finita disponibiliza um isomorfismo canônico[5] entre E e seu espaço dual

$$\begin{aligned} D\colon E &\to E^* = \mathcal{L}(E, \mathbb{R}) \\ v &\mapsto \langle v, \cdot\rangle \end{aligned} \tag{9.1.2}$$

chamado de **dualidade** e onde $\langle v, \cdot\rangle$ é a transformação linear abreviada de

$$v^* := \langle v, \cdot\rangle \colon E \to \mathbb{R}, \quad u \mapsto \langle v, u\rangle.$$

Teorema 9.1.14 (Dualidade)**.** *O operador D em (9.1.2) é um isomorfismo.*

Demonstração. Linearidade de (9.1.2) vale segundo o axioma (`BL`) de bilinearidade. Injetividade vale segundo o axioma (`POS`) na sua incarnação (`ND`)$'$. Sobrejetividade é equivalente a injetividade segundo Corolário 6.5.2 porque as dimensões $\dim E = n = \dim E^*$ são iguais segundo Lema 4.1.20. □

[5] **Canônico** significa sem a necessidade de fazer escolhas das quais o objeto construído eventualmente vai depender. Por exemplo, se $\dim E = \dim F$, então E e F são isomorfos segundo Corolário 6.4.9, mas não tem um isomorfismo canônico geralmente. Mas um isomorfismo com muita escolha geralmente não pode extrair informações intrínsecas.

Exercício 9.1.15 (Produto interno induzido no espaço dual)**.** Mostre que

$$\langle\cdot,\cdot\rangle_* := \langle D^{-1}\cdot, D^{-1}\cdot\rangle\colon E^*\times E^*\to\mathbb{R}$$

é um produto interno no espaço dual de E.

Exercício 9.1.16. Seja E um espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$ e seja $\mathcal{B}=\{\xi_1,\dots,\xi_n\}$ uma base de E. Dado números $\alpha_1,\dots,\alpha_n\in\mathbb{R}$, prove que existe um único vetor $w\in E$ tal que

$$\langle w,\xi_1\rangle=\alpha_1,\ \dots,\ \langle w,\xi_n\rangle=\alpha_n.$$

As afirmações continuam em Lema 9.1.17.

[Dica: Proposição 4.1.12 diz que uma transformação linear $\psi\colon E\to\mathbb{R}$ é determinada por seus valores numa base, dizemos $\psi\xi_i:=\alpha_i$. Defina $w:=D^{-1}\psi$.]

Lema 9.1.17 (Continuando Exercício 9.1.16)**.** *Seja E um espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$ e seja $\mathcal{B}=\{\xi_1,\dots,\xi_n\}$ uma base de E. Prove que existe uma única base $\mathcal{V}=\{\eta_1,\dots,\eta_n\}$ de E tal que*

$$\langle\eta_i,\xi_j\rangle=\delta_{ij},\qquad i,j=1,\dots,n.$$

Defina $a_{ij}:=\langle\xi_i,\xi_j\rangle$ e $b_{ij}:=\langle\eta_i,\eta_j\rangle$, onde $i,j=1,\dots,n$. Prove que as matrizes $\mathbf{a}=(a_{ij})$ e $\mathbf{b}=(b_{ij})$ são inversas uma da outra.

Demonstração. Dado uma base $\mathcal{B}=\{\xi_1,\dots,\xi_n\}$ de E, seja $\mathcal{B}^*=\{\phi_1,\dots\phi_n\}$ a base dual (4.1.6) de E^*. Isomorfismos levam base em base (Teorema 6.4.7), assim $\mathcal{V}:=D^{-1}\mathcal{B}^*$ é uma base de E. Para os elementos $\eta_i:=D^{-1}\phi_i$ de $\mathcal{V}$ vale

$$\langle\eta_i,\xi_j\rangle=\langle D^{-1}\phi_i,\xi_j\rangle\overset{(9.1.2)}{=}\left(D(D^{-1}\phi_i)\right)\xi_j=\phi_i\xi_j\overset{(4.1.6)}{=}\delta_{ij}.$$

Segundo Teorema 5.2.7 $I_E=D^{-1}D$ traduz em $[I_E]_{\mathcal{B},\mathcal{B}}=\left[D^{-1}D\right]_{\mathcal{B},\mathcal{B}}$. Assim

$$\mathbb{1}\overset{(5.2.4)}{=}[I_E]_{\mathcal{B},\mathcal{B}}=\left[D^{-1}D\right]_{\mathcal{B},\mathcal{B}}\overset{(5.2.6)}{=}\left[D^{-1}\right]_{\mathcal{B}^*,\mathcal{B}}[D]_{\mathcal{B},\mathcal{B}^*}=\mathbf{ba}$$

onde resta provar a última identidade. (Como as matrizes são quadradas $\mathbb{1}=\mathbf{ba}$ é equivalente a $\mathbb{1}=\mathbf{ab}$.) Mais detalhado, resta provar que

$$\mathbf{c}:=\left[D^{-1}\right]_{\mathcal{B}^*,\mathcal{B}}=\mathbf{b}:=(\langle\eta_i,\eta_j\rangle)\quad,\quad\mathbf{d}:=[D]_{\mathcal{B},\mathcal{B}^*}=\mathbf{a}:=(\langle\xi_i,\xi_j\rangle).$$

Começamos com a definição de

$$\begin{aligned}
b_{ij}:&=\langle\eta_i,\eta_j\rangle\\
&=\langle D^{-1}\phi_i,\eta_j\rangle\\
&=\left(D(D^{-1}\phi_i)\right)\eta_j\\
&=\phi_i\eta_j\\
&=\phi_i\left(D^{-1}\phi_j\right)\\
&\overset{*}{=}\phi_i\left(\xi_1c_{1j}+\dots\xi_nc_{nj}\right)\\
&=(\phi_i\xi_i)c_{ij}\\
&=c_{ij}
\end{aligned}$$

onde $*$ vale por definição (5.2.1) da matriz $\left[D^{-1}\right]_{\mathcal{B}^*,\mathcal{B}} =: (c_{ij})$. Deixamos ao leitor provar similarmente $a_{ij} = d_{ij}$. □

9.1.2 Produto interno e matrizes simétricas positivas

Consideramos um espaço vetorial real E de dimensão finita n. Escolhendo uma base ordenada $\mathcal{B}$ permite traduzir cada um vetor numa lista de n números reais, o vetor coordenada. Vamos ver como identificar produtos internos com uma classe de matrizes quadradas.

Vai ser útil escrever vetores coordenadas como matrizes coluna. Denotamos de

$$\mathbf{u} := \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathrm{M}(n \times 1)$$

o vetor coordenada $[u]_{\mathcal{B}}$ de um vetor u em respeito à base $\mathcal{B}$, veja (5.1.2).

Produto interno associado a uma base ordenada

Proposição 9.1.18 (Existência de produtos internos)**.** *Um espaço vetorial real E de dimensão finita admite um produto interno.*

Demonstração. Dado uma base ordenada $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ de E, defina

$$\langle u, v\rangle_{\mathcal{B}} := \langle \mathbf{u}, \mathbf{v}\rangle_0 := \mathbf{u}^t \mathbf{v} \tag{9.1.3}$$

para $u, v \in E$ onde $\mathbf{u} = [u]_{\mathcal{B}} \in \mathrm{M}(n \times 1)$ é o vetor coordenada escrito como matriz coluna. Os axiomas do produto euclidiano $\langle\cdot,\cdot\rangle_0$ no $\mathbb{R}^n$ implicam os axiomas correspondentes para $\langle\cdot,\cdot\rangle_{\mathcal{B}}$. □

Exercício 9.1.19. Seja $E = \mathbb{R}^n$. Mostre que o produto interno associado à base canônica $\langle\cdot,\cdot\rangle_{\mathcal{E}} = \langle\cdot,\cdot\rangle_0$ reproduz o produto euclidiano no $\mathbb{R}^n$.

A matriz de um produto interno

Definição 9.1.20 (Matriz de um produto interno)**.** Seja E um espaço vetorial munido de um produto interno $\langle\cdot,\cdot\rangle$. Dado uma base ordenada $\mathcal{B} = \{\xi_1, \ldots, \xi_n\}$ de E, calcule todos os números reais

$$g_{ij} := \langle \xi_i, \xi_j\rangle$$

e coloque numa matriz quadrada denotada, dependente do contexto, de

$$\mathbf{g} = \mathbf{g}_{\mathcal{B}} := (g_{ij})_{i,j=1}^n \in \mathrm{M}(n \times n).$$

Esta matriz é chamada de **matriz do produto interno** em respeito à base $\mathcal{B}$. Verifique que a matriz real quadrada $\mathbf{g}$ é simétrica e também **positiva**, ou seja

$$\sum_{i,j=1}^n g_{ij} u_i u_j > 0$$

para todas as listas não nulas $u \in \mathbb{R}^n$. Denotamos de

$$\mathrm{S}^+(n) \subset \mathrm{M}(n \times n)$$

o subconjunto composto das matrizes reais $n \times n$ simétricas e positivas.

Comentário 9.1.21. Seja $\mathbf{g} = \mathbf{g}_{\mathcal{B}}$ a matriz de um produto interno $\langle\cdot,\cdot\rangle$ em respeito a uma base ordenada $\mathcal{B}$ de E. Usando vetores coordenadas $\mathbf{u} = [u]_{\mathcal{B}} \in \mathrm{M}(n \times 1)$ podemos exprimir o produto interno em E através do produto euclidiano assim

$$\begin{aligned} \langle u, v\rangle = \left\langle \sum_i u_i \xi_i, \sum_j v_j \xi_j \right\rangle &= \sum_{i,j=1}^n u_i g_{ij} v_j \\ &= \underbrace{\mathbf{u}^t}_{1\times n} \underbrace{\mathbf{g}}_{n\times n} \underbrace{\mathbf{v}}_{n\times 1} = \langle \mathbf{u}, \mathbf{g}\mathbf{v}\rangle_0 \end{aligned} \tag{9.1.4}$$

para todos os vetores u, v de E.

Vice versa, dada uma base ordenada $\mathcal{B}$ de E, a fórmula (9.1.4) define um produto interno em E para cada uma matriz simétrica positiva $\mathbf{s}$, em símbolos

$$\Phi_{\mathcal{B}} \colon \mathrm{S}^+(n) \xrightarrow{\text{bij.}} \{\text{produtos internos em } E\}, \quad \mathbf{s} \mapsto \langle [\cdot]_{\mathcal{B}}, \mathbf{s}\, [\cdot]_{\mathcal{B}}\rangle_0.$$

A aplicação $\Phi_{\mathcal{B}}$ é uma bijeção entre conjuntos. Sobrejetivo: escolha um produto interno em E e use para $\mathbf{s}$ a matriz dele. Injetivo: aplique Lema 9.1.3.

Exemplo 9.1.22. Nos polinômios reais de grau menor ou igual um

$$\mathcal{P}_1(\mathbb{R}) := \{p(x) = a_0 + a_1 x \mid a_0, a_1 \in \mathbb{R}\}$$

considere a base ordenada $\mathcal{B} = (\xi_1, \xi_2) = (3, 1 + x)$. Integração

$$\langle p, q\rangle := \int_{-1}^1 p(x) q(x)\, dx$$

da um produto interno em $\mathcal{P}_1(\mathbb{R})$ cuja matriz em respeito a $\mathcal{B}$ tem entradas

$$\begin{aligned} g_{11} &:= \langle \xi_1, \xi_1\rangle = \int_{-1}^1 3 \cdot 3\, dx = 9x|_{-1}^1 = 18, \\ g_{22} &:= \langle \xi_2, \xi_2\rangle = \int_{-1}^1 (1 + x) \cdot (1 + x)\, dx = (x + x^2 + x^3/3)\big|_{-1}^1 = 8/3, \\ g_{12} &:= \langle \xi_1, \xi_2\rangle = \int_{-1}^1 3(1 + x)\, dx = (3x + 3x^2/2)\big|_{-1}^1 = 6, \\ g_{21} &:= \langle \xi_2, \xi_1\rangle \overset{\texttt{(SIM)}}{=} \langle \xi_1, \xi_2\rangle = g_{12} = 6. \end{aligned}$$

Exercício 9.1.23 (Continuamos Exemplo 9.1.22)**.** Determine a distância

$$d(\xi_1, \xi_2) := |\xi_1 - \xi_2| := \sqrt{\langle \xi_1 - \xi_2, \xi_1 - \xi_2\rangle}$$

dos dois membros da base de $E = \mathcal{P}_1$.

Uma solução em *E*. Inserindo na fórmula obtemos para o quadrado

$$d(\xi_1, \xi_2)^2 = \int_{-1}^{1} \underbrace{(3 - (1+x))^2}_{4+4x+x^2} \, dx = (4x + 2x^2 + \tfrac{1}{3}x^3)\Big|_{x=-1}^{1} = \frac{26}{3}.$$

Outra solução em coordenadas. A matriz $[g]_{\mathcal{B}}$ já conhecemos, calculamos

$$[\xi_1 - \xi_2]_{\mathcal{B}} = [\xi_1]_{\mathcal{B}} - [\xi_2]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Com isso, usando (9.1.4) no primeiro passo, obtemos

$$\begin{aligned} d(\xi_1, \xi_2)^2 &= \langle [\xi_1 - \xi_2]_{\mathcal{B}}, [g]_{\mathcal{B}} [\xi_1 - \xi_2]_{\mathcal{B}} \rangle_0 \\ &= \left\langle \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 18 & 6 \\ 6 & 8/3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\rangle_0 \\ &= \left\langle \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 12 \\ 10/3 \end{bmatrix} \right\rangle_0 \\ &= \frac{26}{3}. \end{aligned}$$

9.2 Plano euclidiano: Ângulos e comprimentos

Depois as definições abstratas da prévia Seção 9.1 uns leitores deviam-se perguntar como os matemáticos chegaram a estas fórmulas? Como chegaram à noção de produto interno? Isso tem uma significância no dia a dia?

O curso "Álgebra Linear", bem abstrato e algébrico, é baseado no curso "Geometria Analítica", bem geométrico. Esta ordem reflete o desenvolvimento histórico. Os gregos Pitágoras (aprox. 570-495 antes do Cristo) e Euclides (aprox. 325-270 antes do Cristo), entre outros, estudaram a geometria no plano e o francês René Descartes (1596-1650) introduziu uma ferramenta – o sistema de coordenadas Cartesianas (1637) veja Definição 0.0.4 – para traduzir os estudos geométricos do plano no universo da álgebra.

Vamos lembrar o lado da álgebra primeiro. No espaço vetorial $\mathbb{R}^2$ das listas ordenadas $u = (u_1, u_2)$ de dois números reais temos introduzido uma função

$$\langle \cdot, \cdot \rangle_0 \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \quad (u, v) \mapsto u_1 v_1 + u_2 v_2$$

chamado de **produto interno euclidiano** e uma funcao associada

$$|\cdot|_0 : \mathbb{R}^2 \to [0, \infty), \quad u \mapsto \sqrt{\langle u, u \rangle_0} = \sqrt{u_1^2 + u_2^2}$$

chamado de **norma euclidiana**.

Entramos então a geometria. Fixando no plano Π um sistema de coordenadas Cartesianas OXY as duas funções acima ganham significância geométrica:

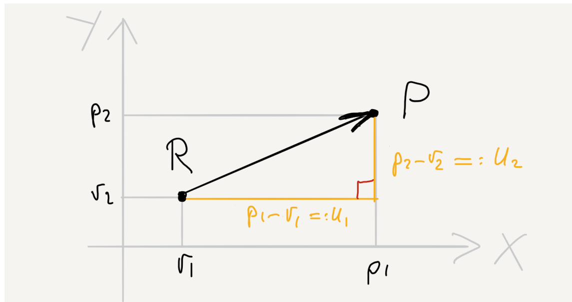

Figura 9.2: Pelo teorema de Pitágoras $\operatorname{dist}(R,P)^2 = u_1^2 + u_2^2 =: |u|_0^2$

Lema 9.2.1 (Comprimento e ângulo)**.** *Seja OXY um sistema de coordenadas Cartesianas no plano Π. Uma flecha $\overrightarrow{RP}$ entre dois pontos, dizemos R com coordenadas (r_1, r_2) e P com coordenadas (p_1, p_2), têm por definição o vetor coordenada $u = (u_1, u_2) := (p_1 - r_1, p_2 - r_2)$.*[6] *Denotamos de v o vetor coordenada da flecha de um ponto S para um ponto Q. Então vale o seguinte.*

(i) *A norma euclidiana do vetor coordenada u da flecha $\overrightarrow{RP}$, ou seja*

$$|u|_0 := \sqrt{u_1^2 + u_2^2} = \operatorname{dist}(R, P) \tag{9.2.1}$$

*é igual ao **comprimento** da flecha $\overrightarrow{RP}$.*

(ii) *O produto interno dos vetores coordenadas u, v de flechas $\overrightarrow{RP}, \overrightarrow{SQ}$, ou seja*

$$\begin{aligned} \langle u, v\rangle_0 := u_1 v_1 + u_2 v_2 &= \operatorname{dist}(R,P) \cdot \operatorname{dist}(S,Q) \cdot \cos\theta \\ &= |u| \cdot |v| \cdot cos\angle(u,v) \end{aligned} \tag{9.2.2}$$

*é igual ao produto dos comprimentos das flechas $\overrightarrow{RP}$ e $\overrightarrow{SQ}$ vezes o coseno do **ângulo** menor*[7] *θ entre as retas contendo as flechas.*

Comentário 9.2.2. Nunca esqueça uma coisa: Num sistema de coordenadas não-Cartesianas, ou seja onde os dois eixos OX e OY não são ortogonais um ao outro, a interpretação geométrica acima é errada.

Demonstração de Lema 9.2.1. (i) Como os eixos OX e OY são ortogonais o teorema de Pitágoras aplica: o quadrado do comprimento da hipotenusa RP é a soma dos quadrados do comprimento $|u_i| = |p_i - r_i|$ dos catetos; veja Figura 9.2. (ii) Flechas equipolentes (mesmo comprimento, direção, sentido de percurso - veja Exemplo 0.0.1) tem o mesmo vetor coordenada. Por isso suponhamos que as flechas tem ponto inicial na origem $O = R = S$. Assim o vetor coordenada de $\overrightarrow{OP}$ é $u = (u_1, u_2) = (p_1, p_2)$, análogo para $\overrightarrow{OQ}$. Consideramos dois casos.

[6] Exercício: Flechas equipolentes, veja Exemplo 0.0.1, têm o mesmo vetor coordenada.
[7] de fato, a simetria do coseno garante que não importa se escolhe o ângulo menor ou maior

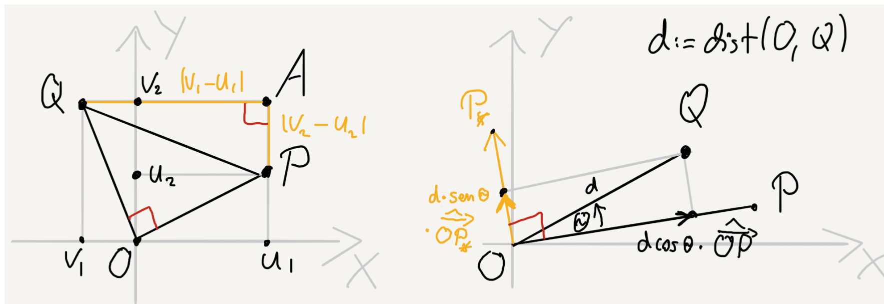

Figura 9.3: Caso perpendicular Caso geral

Caso perpendicular $\overrightarrow{OP} \perp \overrightarrow{OQ}$. De um lado, Pitágoras para o triângulo retângulo OPQ diz que

$$\operatorname{dist}(P,Q)^2 = \operatorname{dist}(O,P)^2 + \operatorname{dist}(O,Q)^2.$$

De outro lado, Pitágoras para o triângulo APQ, veja Figura 9.3, diz que

$$\begin{aligned}
\operatorname{dist}(P,Q)^2 &= (v_1-u_1)^2 + (v_2-u_2)^2 \\
&= u_1^2 + u_2^2 + v_1^2 + v_2^2 - 2(u_1v_1 + u_2v_2) \\
&= \operatorname{dist}(O,P)^2 + \operatorname{dist}(O,Q)^2 - 2\langle u,v\rangle_0.
\end{aligned}$$

Pela comparação $\langle u,v\rangle_0 = 0$, de outro lado $\cos\pi/2 = 0$, assim (9.2.2) vale.

Caso geral. Utilizamos o caso perpendicular junto com o axioma (`BL`) de bilinearidade do produto interno. Denotamos de θ o angulo menor entre as duas flechas $\overrightarrow{OP}$ e $\overrightarrow{OQ}$. Dada a flecha $\overrightarrow{OP}$ com vetor coordenada u, e $\overrightarrow{OQ}$ com v, fixamos um ponto $P_* \neq O$ tal que a flecha $\overrightarrow{OP_*}$ (vetor coordenada u_*) é ortogonal a $\overrightarrow{OP}$. Como ilustrado na Figura 9.3 consideramos a combinação linear

$$\overrightarrow{OQ} = \operatorname{dist}(O,Q)\cdot\cos\theta\cdot\frac{\overrightarrow{OP}}{\operatorname{dist}(O,P)} + \operatorname{dist}(O,Q)\cdot\sin\theta\cdot\frac{\overrightarrow{OP_*}}{\operatorname{dist}(O,P_*)}$$

a qual em coordenadas toma a forma ($\hat{u} := \frac{1}{|u|_0}u$ é vetor unitário)

$$v = |v|_0\cdot\cos\theta\cdot\hat{u} + |v|_0\cdot\sin\theta\cdot\widehat{u_*}.$$

Tomando o produto interno com u e usando o axioma (`BL`) obtemos

$$\langle u,v\rangle_0 = |v|_0\cdot\cos\theta\cdot\underbrace{\langle u,\hat{u}\rangle_0}_{|u|_0} + |v|_0\cdot\sin\theta\cdot\underbrace{\langle u,\widehat{u_*}\rangle_0}_{\overset{(9.2.2)}{=}0} = |u|_0\,|v|_0\cdot\cos\theta.$$

Agora use (9.2.1) para $|u|_0$ e para $|v|_0$. Note que segundo o caso perpendicular já provado a fórmula (9.2.2) aplica e diz que $\langle u,\widehat{u_*}\rangle_0 = \dots\cdot\underbrace{\cos\pi/2}_{=0} = 0.$ □

9.3 Ortogonalidade

Definição 9.3.1. Seja E um espaço vetorial com produto interno.

(i) Chama-se dois **vetores** u e v **ortogonais**, ou **perpendiculares**, símbolo $u \perp v$, se tem produto nulo $\langle u, v\rangle = 0$. (Note $\mathcal{O} \perp v$ para todos vetores.)

Chama-se um **vetor** u **ortogonal a um subconjunto** X, símbolo $u \perp X$, se $u \perp x$ para todos os elementos x de X.

(ii) Chama-se $X \subset E$ um **subconjunto ortogonal** se os vetores de X são dois-a-dois ortogonais.

(iii) Chama-se $X \subset E$ um **subconjunto ortonormal (ON)** se X é composto de vetores unitários (norma 1) dois-a-dois ortogonais.

(iv) Uma base $\mathcal{Z} = \{\varepsilon_1, \dots, \varepsilon_n\}$ é chamado de **base ortonormal (ON)** se

$$\langle \varepsilon_i, \varepsilon_j\rangle = \delta_{ij} := \begin{cases} 1 & , i = j \\ 0 & , i \neq j \end{cases} \tag{9.3.1}$$

onde δ_{ij} é o símbolo de Kronecker.

Teorema 9.3.2 (Conjuntos ortogonais, sem $\mathcal{O}$, são LI)**.**

$$X \subset E \setminus \{\mathcal{O}\} \textit{ conjunto ortogonal} \quad \Rightarrow \quad X \textit{ LI.}$$

Demonstração. Suponha que

$$\alpha_1 x_1 + \dots + \alpha_k x_k = \mathcal{O}$$

para uma escolha finita de elementos $x_i \in X$ e escalares $\alpha_i \in \mathbb{R}$. Tome o produto interno de ambos lados com qualquer um dos elementos, dizemos x_j, segue

$$\alpha_1 \underbrace{\langle x_1, x_j\rangle}_{=0} + \underbrace{\dots}_{=0} + \alpha_j \underbrace{\langle x_j, x_j\rangle}_{=1} + \underbrace{\dots}_{=0} + \alpha_k \underbrace{\langle x_k, x_j\rangle}_{=0} = \langle \mathcal{O}, x_j\rangle = 0.$$

Mas $\alpha_j |x_j|^2 = 0$ implica $\alpha_j = 0$ porque $\langle x_j, x_j\rangle > 0$ pela hipótese $x_j \neq \mathcal{O}$. □

Exercício 9.3.3. Dado uma base ON $\mathcal{Z} = \{\varepsilon_1, \dots, \varepsilon_n\}$ de E, mostre que

$$v = \sum_{i=1}^{n} v_i \varepsilon_i \qquad \Leftrightarrow \qquad v_i = \langle \varepsilon_i, v\rangle, \quad i = 1, \dots, n \tag{9.3.2}$$

para cada um vetor $v \in E$.

Exercício 9.3.4. Seja $\mathcal{B}$ uma base ordenada de um espaço vetorial real E de dimensão finita n. Seja $\langle u, v\rangle_{\mathcal{B}} = \langle [u]_{\mathcal{B}}, [v]_{\mathcal{B}}\rangle_0$ o produto interno associado, veja (9.1.3). Mostre que $\mathcal{B}$ é uma base ON de $\langle u, v\rangle_{\mathcal{B}}$.

Exemplo 9.3.5 (Conjuntos e bases ortogonais)**.**

a) A base canônica $\mathcal{E}^n$ é ortonormal em respeito a $\langle\cdot,\cdot\rangle_0$.
b) O conjunto $\{(0,0),(-1,1)\}$ é ortogonal em $\mathbb{R}^2$.
c) O conjunto $\{(1,1),(-1,1)\}$ é uma base ortogonal de $\mathbb{R}^2$.

Teorema 9.3.6 (Teorema de Pitágoras para espaços com produto interno)**.** *Para dois vetores $u,v\in E$ são equivalente*

$$u\perp v \qquad\Leftrightarrow\qquad |u+v|^2=|u|^2+|v|^2\,. \tag{9.3.3}$$

Demonstração. Usamos os axiomas bi-linearidade e simetria para obter

$$\begin{aligned}|u+v|^2 &= \langle u+v,u+v\rangle\\ &\overset{(\mathtt{BL})}{=} \langle u,u\rangle+\langle u,v\rangle+\langle v,u\rangle+\langle v,v\rangle\\ &\overset{(\mathtt{SIM})}{=} |u|^2+|v|^2+2\langle u,v\rangle\end{aligned}$$

e agora lembramos que por definição escrevemos $u\perp v$ no caso $\langle u,v\rangle=0$. □

Exemplo 9.3.7. Para vetores não-nulos $u,v\in\mathbb{R}^2$ ter produto euclidiano nulo

$$0=\langle u,v\rangle_0=|u|\cdot|v|\cdot\cos\angle(u,v)$$

é equivalente que o ângulo entre eles é rectângulo, ou seja $\angle(u,v)\in\{\frac{\pi}{2},\frac{3\pi}{2}\}$.

9.3.1 Projeção ortogonal sobre uma reta

Definição 9.3.8 (Projeção ortogonal sobre uma reta $\mathbb{R}\hat u$)**.** Seja E um espaço vetorial com produto interno. Dado um vetor $u\in E$ não-nulo, seja $\hat u=\frac{1}{|u|}u$ o vetor unitário correspondente. Definimos a transformação linear

$$\begin{aligned}\mathrm{pr}_u\colon E&\to E\\ v&\mapsto \frac{\langle u,v\rangle}{\langle u,u\rangle}u=\underbrace{\langle\hat u,v\rangle\hat u}_{=\mathrm{pr}_{\hat u}v}\in\mathbb{R}\hat u\end{aligned}$$

chamada de **projeção ortogonal sobre a reta $\mathbb{R}\hat{\mathbf{u}}$**. Note que $\mathrm{pr}_u=\mathrm{pr}_{\hat u}$.

Vamos justificar o nome de pr_u. Linearidade segue do axioma (BL) e

$$\mathrm{pr}_{\hat u}(\mathrm{pr}_{\hat u}v)=\langle\hat u,\mathrm{pr}_{\hat u}v\rangle\,\hat u=\langle\hat u,\langle\hat u,v\rangle\hat u\rangle\,\hat u=\langle\hat u,v\rangle\langle\hat u,\hat u\rangle\hat u=\mathrm{pr}_{\hat u}v$$

mostra que $\mathrm{pr}_{\hat u}=(\mathrm{pr}_{\hat u})^2$ é uma projeção. Assim

$$\mathrm{Im}(\mathrm{pr}_{\hat u})\overset{(7.0.1)}{=}\mathrm{Fix}(\mathrm{pr}_{\hat u})=\mathbb{R}\hat u \tag{9.3.4}$$

onde identidade dois segue imediatamente da definição de $\mathrm{pr}_{\hat u}$, similarmente

$$\mathrm{N}(\mathrm{pr}_{\hat u})=\{v\in E\mid \langle v,\alpha\hat u\rangle=0,\ \forall\alpha\hat u\in\mathbb{R}\hat u\}=:(\mathbb{R}\hat u)^\perp\,.$$

Para justificar o termo *ortogonal* no nome de pr_u escolha $v \in E$ e note que

$$\langle \mathrm{pr}_{\hat{u}} v - v, \alpha\hat{u}\rangle = \langle\langle \hat{u}, v\rangle \hat{u}, \alpha\hat{u}\rangle - \langle v, \alpha\hat{u}\rangle = \alpha \langle \hat{u}, v\rangle \langle \hat{u}, \hat{u}\rangle - \alpha \langle v, \hat{u}\rangle = 0.$$

Assim o vetor w conectando os pontos v e $\mathrm{pr}_{\hat{u}} v$ é ortogonal à reta $\mathbb{R}\hat{u}$, ou seja

$$w := (v - \mathrm{pr}_{\hat{u}} v) \perp \mathbb{R}\hat{u}, \quad v \in E.$$

Na notação (7.1.1) temos que

$$\mathrm{pr}_{\hat{u}} = P_{\mathbb{R}\hat{u}.(\mathbb{R}\hat{u})^\perp}.$$

Lema 9.3.9 (Projeções ortogonais não alongam)**.** *Para $u, v \in E$ não-nulos vale*

$$|\mathrm{pr}_u v| \le |v|$$

onde "$=$" vale exatamente ao longo dos pontos fixos (9.3.4), ou seja $\mathrm{pr}_u v = v$.

Demonstração. Como $w := (v - \mathrm{pr}_{\hat{u}} v) \perp \mathbb{R}\hat{u}$ obtemos do Pitágoras (9.3.3) que

$$|v|^2 = |v - \mathrm{pr}_{\hat{u}} v + \mathrm{pr}_{\hat{u}} v|^2 = |w + \mathrm{pr}_{\hat{u}} v|^2 = |w|^2 + |\mathrm{pr}_{\hat{u}} v|^2 \ge |\mathrm{pr}_{\hat{u}} v|^2$$

onde "$=$" vale exatamente se $|w|^2 = 0$, equivalentemente $\mathcal{O} = w = v - \mathrm{pr}_{\hat{u}} v$. □

9.4 Desigualdades

Proposição 9.4.1 (Desigualdade de Schwarz)**.** *Para $u, v \in E$ não-nulos*[8] *vale*

$$|\langle u, v\rangle| \le |u| \cdot |v| \tag{9.4.1}$$

onde "$=$" é equivalente a cada um de u, v é múltiplo (não-nulo) do outro.

Demonstração. Caso um de u ou v é nulo as afirmações valem. Suponha no seguinte $u \ne \mathcal{O}$ e $v \ne \mathcal{O}$: "$\le$" Usando Lema 9.3.9 no último passo obtemos

$$\frac{|\langle u, v\rangle|}{|u|} = |\langle \hat{u}, v\rangle \, \hat{u}| = |\mathrm{pr}_{\hat{u}} v| \le |v|.$$

"$=$" Segundo Lema 9.3.9 temos igualdade se e somente se $v = \mathrm{pr}_{\hat{u}} v$. Mas $\mathrm{pr}_{\hat{u}} v \in \mathrm{Im}(\mathrm{pr}_{\hat{u}}) = \mathbb{R}\hat{u}$. Daí $v \in \mathbb{R}\hat{u} = \mathbb{R}u$ é da forma $v = \alpha u$ com $\alpha \ne 0$. □

Proposição 9.4.2 (Desigualdade triangular)**.** *Para $u, v \in E$ não-nulos*[9] *vale*

$$|u + v| \le |u| + |v| \tag{9.4.2}$$

onde "$=$" é equivalente a cada um de u, v é múltiplo positivo do outro.

Demonstração. Utilizamos a desigualdade de Schwarz (9.4.1) para obter

$$|u + v|^2 = |u|^2 + |v|^2 + 2\langle u, v\rangle \le |u|^2 + |v|^2 + 2|u| \cdot |v| = (|u| + |v|)^2$$

onde igualdade "$=$" é equivalente a $\langle u, v\rangle = |u| \cdot |v| > 0$. Assim, por Schwarz, cada um de u, v é múltiplo do outro, dizemos $v = \alpha u$ para um real $\alpha \ne 0$. Mas neste caso de $0 < \langle u, v\rangle = \alpha |u|^2$ e $|u|^2 > 0$ segue $\alpha > 0$. □

[8] Para u, ou v, nulo (9.4.1) também vale: $0 = 0$. Mas só o nulo é múltiplo do outro.
[9] Para u, ou v, nulo (9.4.2) também vale: $0 = 0$. Mas só o nulo é múltiplo do outro.

9.5 Ortonormalização segundo Gram-Schmidt

Hipótese. Seja $\mathcal{X} = \{\xi_1, \dots, \xi_n\}$ uma base ordenada de um espaço vetorial E com produto interno. Denotamos de

$$F_1 := \langle \xi_1 \rangle \subset \cdots \subset \boxed{F_k := \langle \xi_1, \dots, \xi_k \rangle} \subset \cdots \subset F_n := \langle \xi_1, \dots, \xi_n \rangle = E$$

os subespaços gerados pelos primeiros $1, 2, \dots, n$ membros da base $\mathcal{X}$.

Passo A. Vamos construir iterativamente bases ortogonais:

(1) base ortogonal $\{\eta_1\}$ de F_1: escolha $\eta_1 := \xi_1$ e já pronto;

(k) dado $k \geq 1$ suponha que $\{\eta_1, \dots, \eta_k\}$ é base ortogonal de F_k, defina

$$\eta_{k+1} := \xi_{k+1} - \sum_{i=1}^{k} \mathrm{pr}_{\eta_i} \xi_{k+1} = \xi_{k+1} - \sum_{i=1}^{k} \frac{\langle \eta_i, \xi_{k+1} \rangle}{\langle \eta_i, \eta_i \rangle} \eta_i; \tag{9.5.1}$$

($k+1$) então $\{\eta_1, \dots, \eta_k, \eta_{k+1}\}$ é uma base ortogonal de F_{k+1};

o processo usa o último membro ξ_n de $\mathcal{X}$ quando $k = n - 1 \Rightarrow k + 1 = n$

(n) ao fim obtemos uma base ortogonal

$$\mathcal{Y} = \{\eta_1, \dots, \eta_k, \eta_{k+1}, \dots, \eta_n\}$$

de $F_n = E$.

Demonstração $(k) \Rightarrow (k+1)$. Suponha (k) e defina η_{k+1}, então

a) $\eta_{k+1} \perp \eta_1, \dots, \eta_k$; $\quad \langle \eta_{k+1}, \eta_j \rangle \overset{(k)}{=} \langle \xi_{k+1}, \eta_j \rangle - \langle \xi_{k+1}, \eta_j \rangle = 0$

b) $\eta_{k+1} \notin F_k \overset{(k)}{=} \langle \eta_1, \dots, \eta_k \rangle \ni \mathcal{O}$; $\quad$ suponha por absurdo $\eta_{k+1} \in \langle \eta_1, \dots, \eta_k \rangle$ $\Rightarrow \xi_{k+1} \in \langle \eta_1, \dots, \eta_k \rangle = F_k := \langle \xi_1, \dots, \xi_k \rangle$ contradição

c) $\eta_{k+1} \in F_{k+1}$. $\quad \eta_{k+1} \in \langle \eta_1 \dots, \eta_k, \xi_{k+1} \rangle \overset{(k)}{=} \langle \xi_1 \dots, \xi_k, \xi_{k+1} \rangle =: F_{k+1}$

Segundo hipótese (k) o conjunto $\{\eta_1, \dots, \eta_k\}$ é LI. Além disso η_{k+1} é não-nulo segundo b) e ortogonal a $\eta_1, \dots, \eta_k$ segundo a). Sendo assim o conjunto ortogonal $\{\eta_1, \dots, \eta_k, \eta_{k+1}\}$ é LI segundo Teorema 9.3.2. Note que o subespaço

$$\langle \eta_1, \dots, \eta_{k+1} \rangle \subset \langle \xi_1, \dots, \xi_{k+1} \rangle$$

é contido num subespaço da mesma dimensão $k + 1$. Então os dois são iguais segundo Teorema 3.2.1 (d). Assim $\{\eta_1, \dots, \eta_{k+1}\}$ é base ortogonal de F_{k+1}.

Passo B. A base $\mathcal{Z} := \{\hat{\eta}_1, \dots, \hat{\eta}_n\}$ de E é ortonormal onde $\hat{\eta}_i := |\eta_i|^{-1} \eta_i$.

Comentário 9.5.1. No caso que ξ_{k+1} já é ortogonal a $\eta_1, \dots, \eta_k$ a definição de η_{k+1} mostra que $\eta_{k+1} = \xi_{k+1}$. O processo de Gram-Schmidt não muda ξ_{k+1}.

Exercício 9.5.2 (Listas arbitrárias)**.** Seja $(\xi_1,\dots,\xi_\ell)$ uma lista arbitrária de ℓ vetores $\xi_i \in E$, dobros e o vetor nulo, tudo permitido. Pode-se aplicar o processo de Gram-Schmidt com a seguinte modificação pequena da hipótese

(k) dado $k \geq 1$ suponha a lista $(\eta_1,\dots,\eta_k)$ gera F_k e seus membros são dois-a-dois ortogonais, defina

$$\eta_{k+1} := \xi_{k+1} - \sum_{\substack{i=1\\ \eta_i\neq\mathcal{O}}}^{k} \mathrm{pr}_{\eta_i}\xi_{k+1} = \xi_{k+1} - \sum_{\substack{i=1\\ \eta_i\neq\mathcal{O}}}^{k} \langle\hat\eta_i,\xi_{k+1}\rangle\,\hat\eta_i.$$

Obtém-se também uma lista $(\eta_1,\dots,\eta_\ell)$ cujos membros são dois-a-dois ortogonais, só agora é possível que uns são nulos. Com efeito, mostre que

$$\xi_{k+1} \in \langle\xi_1,\dots,\xi_k\rangle \overset{(k)}{=} \langle\eta_1,\dots,\eta_k\rangle \quad\Rightarrow\quad \eta_{k+1} = \mathcal{O}$$

[Dica: Note que $\langle\hat\eta_i,\xi_{k+1}\rangle$ é a i-ésima coordenada do vetor ξ_{k+1} na base ON composto daqueles $\hat\eta_i$ onde $\eta_i\neq\mathcal{O}$ é não-nulo. Exercício 9.3.3.]

Exemplo 9.5.3. Determine uma base ON do subespaço $F\subset\mathbb{R}^3$ gerado por

$$\xi_1 = \begin{bmatrix}1\\-1\\1\end{bmatrix},\quad \xi_2 = \begin{bmatrix}1\\1\\1\end{bmatrix},\quad \xi_3 = \begin{bmatrix}1\\0\\1\end{bmatrix},\qquad \langle\cdot,\cdot\rangle := \langle\cdot,\cdot\rangle_0.$$

Solução com Gram-Schmidt (GS). Definição (9.5.1) dos η_{k+1} diz que

$$\eta_1 := \xi_1 = \begin{bmatrix}1\\-1\\1\end{bmatrix},\quad |\eta_1|^2 := \langle\eta_1,\eta_1\rangle = 1^2+(-1)^2+1^2 = 3,$$

$$\eta_2 := \xi_2 - \frac{\langle\eta_1,\xi_2\rangle}{|\eta_1|^2}\,\eta_1 = \begin{bmatrix}1\\1\\1\end{bmatrix} - \frac{1}{3}\underbrace{\left\langle\begin{bmatrix}1\\-1\\1\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}\right\rangle}_{=1}\begin{bmatrix}1\\-1\\1\end{bmatrix} = \tfrac{2}{3}\begin{bmatrix}1\\2\\1\end{bmatrix},\ |\eta_2|^2 = \frac{8}{3},$$

$$\begin{aligned}
\eta_3 &:= \xi_3 \quad - \frac{\langle\eta_1,\xi_3\rangle}{|\eta_1|^2}\,\eta_1 \qquad\qquad - \frac{\langle\eta_2,\xi_3\rangle}{|\eta_2|^2}\,\eta_2\\
&= \begin{bmatrix}1\\0\\1\end{bmatrix} - \frac{1}{3}\underbrace{\left\langle\begin{bmatrix}1\\-1\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix}\right\rangle}_{=2}\begin{bmatrix}1\\-1\\1\end{bmatrix} - \frac{3}{8}\underbrace{\left\langle\tfrac{2}{3}\begin{bmatrix}1\\2\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix}\right\rangle}_{=\frac{2}{3}\cdot 2}\tfrac{2}{3}\begin{bmatrix}1\\2\\1\end{bmatrix}\\
&= \begin{bmatrix}1\\0\\1\end{bmatrix} - \tfrac{2}{3}\begin{bmatrix}1\\-1\\1\end{bmatrix} - \tfrac{1}{3}\begin{bmatrix}1\\2\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.
\end{aligned}$$

Segundo GS e como $\eta_3 = \mathcal{O}$ sabemos que $F := \langle\xi_1,\xi_2,\xi_3\rangle = \langle\eta_1,\eta_2,\eta_3\rangle = \langle\eta_1,\eta_2\rangle$. GS diz que o conjunto $\{\eta_1,\eta_2\}$ é ortogonal, então LI segundo Teorema 9.3.2 usando que $\eta_1,\eta_2\neq\mathcal{O}$. O conjunto ortogonal $\{\eta_1,\eta_2\}$ é uma base ortogonal de F porque é LI e gera F. Uma base ON é composto dos vetores

$$\varepsilon_1 := \tfrac{1}{|\eta_1|}\,\eta_1 = \tfrac{1}{\sqrt{3}}\begin{bmatrix}1\\-1\\1\end{bmatrix}, \quad \varepsilon_2 := \tfrac{1}{|\eta_2|}\,\eta_2 = \tfrac{\sqrt{3}}{\sqrt{8}}\tfrac{2}{3}\begin{bmatrix}1\\2\\1\end{bmatrix} = \tfrac{1}{\sqrt{6}}\begin{bmatrix}1\\2\\1\end{bmatrix}. \tag{9.5.2}$$

9.5.1 Existência e extensão de bases ortogonais

Lembramos que um espaço vetorial de dimensão finita admite um produto interno – cada uma base ordenada $\mathcal{B}$ induz um, notação $\langle\cdot,\cdot\rangle_{\mathcal{B}}$, veja (9.1.3).

Teorema 9.5.4 (Existência)**.** *Um espaço vetorial E com produto interno e de dimensão finita n admite uma base ON $\mathcal{Z} = \{\varepsilon_1, \dots, \varepsilon_n\}$.*

Demonstração. Pegue uma base ordenada $\mathcal{X} = \{\xi_1, \dots, \xi_n\}$ de E e aplique o processo de ortonormalização de Gram-Schmidt. □

Proposição 9.5.5 (Extensão)**.** *Seja E um espaço vetorial com produto interno. Toda base ON $\mathcal{X}$ de um subespaço F estende-se a uma base ON $\mathcal{Z}$ de E.*

Demonstração. Segundo Teorema 3.2.1 (b) a base $\mathcal{X} = \{\varepsilon_1, \dots \varepsilon_k\}$ de F é contida numa base ordenada $\mathcal{Y}$ de E, dizemos $\mathcal{Y} = \{\varepsilon_1, \dots \varepsilon_k, \xi_{k+1}, \dots, \xi_n\}$. Aplique Gram-Schmidt para obter $\mathcal{Z} = \{\varepsilon_1, \dots \varepsilon_k, \varepsilon_{k+1}, \dots, \varepsilon_n\}$. □

9.5.2 Projeção ortogonal sobre um subespaço

O processo de Gram-Schmidt prova a existência de bases ONs: pegue qualquer base e aplique o processo; veja Proposição 9.5.5. É importante usar uma base ON nesta definição:

Definição 9.5.6 (Projeção ortogonal sobre um subespaço F)**.** Escolha uma base ordenada ON $\mathcal{Z} = \{\varepsilon_1, \dots, \varepsilon_k\}$ de F e defina a transformação linear

$$\begin{aligned} \mathrm{pr}_F \colon E &\to E \\ v &\mapsto \sum_{i=1}^{k} \mathrm{pr}_{\varepsilon_i} v = \sum_{i=1}^{k} \langle \varepsilon_i, v\rangle \varepsilon_i. \end{aligned} \tag{9.5.3}$$

Teorema 9.5.7 (Propriedades da projeção ortogonal)**.**

1. *pr_F é linear e bem definido (independente da base ON $\mathcal{Z}$ de F).*
2. *$(\mathrm{pr}_F)^2 = \mathrm{pr}_F$.*
3. *$\mathrm{Im}(\mathrm{pr}_F) = \mathrm{Fix}(\mathrm{pr}_F)$ e $E = \mathrm{Im}(\mathrm{pr}_F) \oplus \mathrm{N}(\mathrm{pr}_F)$.*
4. *$\mathrm{Im}(\mathrm{pr}_F) = F$.*
5. *$\mathrm{pr}_F|_F = I_F \in \mathcal{L}(F)$.*
6. *$w := (v - \mathrm{pr}_F v) \perp f \quad \forall f \in F$.*

7. $\forall v \in E$ *vale*[10] $\operatorname{dist}(v, F) := \inf_{f\in F} \underbrace{\operatorname{dist}(v, f)}_{:=|v-f|} = |v - \operatorname{pr}_F v|$.

Demonstração. Teorema C.6.1. □

9.6 Complemento ortogonal

Subconjuntos não-vazios

Definição 9.6.1. O **complemento ortogonal** de um subconjunto não-vazio $S \subset E$ é definido assim

$$S^\perp := \{v \in E \mid \langle v, x\rangle = 0 \ \ \forall x \in S\}.$$

Exercício 9.6.2. Seja $S \subset E$ um subconjunto não-vazio. Mostre que

(i) o complemento ortogonal $S^\perp$ é um subespaço de E;

(ii) ou $S^\perp \cap S = \emptyset$, ou $S^\perp \cap S = \{\mathcal{O}\}$;

(iii) $T \subset S \Rightarrow S^\perp \subset T^\perp$;

(iv) $S^\perp = \langle S\rangle^\perp$.

Subespaços

Exercício 9.6.3. Seja $\mathcal{B}$ uma base de um subespaço $F \subset E$, mostre que

$$F^\perp = \{v \in E \mid \langle v, \xi\rangle = 0 \ \ \forall \xi \in \mathcal{B}\}.$$

Proposição 9.6.4 (Relação entre F e $F^\perp$ e a projeção ortogonal pr_F de (9.5.3))**.** *Para subespaços F de E vale o seguinte*

(i) $F^\perp = \mathrm{N}(\operatorname{pr}_F)$; $\qquad \operatorname{Im}(\operatorname{pr}_F) = F$

(ii) $E = F \oplus F^\perp$ *e* $\dim E = \dim F + \dim F^\perp$*;*

(iii) $\operatorname{pr}_F = P_{F,F^\perp}$ *veja (7.1.1);*

(iv) $(F^\perp)^\perp = F$*.*

Demonstração. (i) "$\subset$" Para $v \in F^\perp$ vale

$$0 = \langle\underbrace{\operatorname{pr}_F v}_{\in F}, v\rangle = \langle \operatorname{pr}_F v, \underbrace{v - \operatorname{pr}_F v}_{w\perp F}\rangle + \langle \operatorname{pr}_F v, \operatorname{pr}_F v\rangle = 0 + \langle \operatorname{pr}_F v, \operatorname{pr}_F v\rangle$$

e consequentemente (`POS`) diz que $\operatorname{pr}_F v = \mathcal{O}$.
"$\supset$" Para $v \in \mathrm{N}(\operatorname{pr}_F)$ vale $v = v - \operatorname{pr}_F v \perp f \ \forall f \in F$ segundo Teorema 9.5.7 6.
(ii) Item 3 de Teorema 9.5.7 junto com item 4 e no teorema presente item (i).
(iii) Item (ii) dize que $(F, F^\perp)$ são subespaços complementares. Teorema 7.1.5.
(iv) É suficiente mostrar inclusão $F \subset (F^\perp)^\perp$ e igualdade de dimensão, veja

[10] como $\dim F < \infty$ um ínfimo é um mínimo

Teorema 3.2.1 (d). Seja $f \in F$, então $\langle f, \tilde f\rangle = 0\, \forall \tilde f \in F^\perp$, mas isso significa que $f \in (F^\perp)^\perp := \{v \in E\, id\langle v, \tilde f\rangle = 0\, \forall \tilde f \in F^\perp\}$.
O Corolário 3.2.3 disponibiliza a fórmula de dimensão para a soma direta. Use item (ii) acima primeiro para $F^\perp$ no lugar de F e segundo para F para obter

$$\dim(F^\perp)^\perp = \dim E - \dim F^\perp = \dim F.$$

$\square$

Exemplo 9.6.5. No Exemplo 9.5.3 temos calculado a base ON $\mathcal{Z} = \{\varepsilon_1, \varepsilon_2\}$, veja (9.5.2), do subespaço $F := \langle \xi_1, \xi_2, \xi_3\rangle$ de $\mathbb{R}^3$. Determine uma base ON do complemento ortogonal

$$F^\perp \overset{\text{def.}}{=} \{v \in \mathbb{R}^3 \mid \langle v, f\rangle = 0 \ \ \forall f \in F\} = \{v \in \mathbb{R}^3 \mid v \perp \varepsilon_1,\, v \perp \varepsilon_2\}.$$

Vale a ultima igualdade porque a condição $\langle v, f\rangle = 0$ é linear em f, então é suficiente checar para os elementos f de uma base só.

Uma solução.

$$v \perp \varepsilon_1: \quad 0 = \left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tfrac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\rangle = \tfrac{1}{\sqrt{3}}\,(x - y + z).$$

$$v \perp \varepsilon_2: \quad 0 = \left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tfrac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\rangle = \tfrac{1}{\sqrt{6}}\,(x + 2y + z).$$

Multiplique a primeira identidade por $\sqrt{3}$ e a segunda por $\sqrt{6}$ e forma a diferença das identidades resultantes para obter

$$0 - 0 = 0 - 3y - 0 \qquad \Rightarrow \qquad y = 0.$$

Com isso obtemos da primeira identidade que

$$z = -x,\ x \in \mathbb{R} \text{ livre}, \qquad F^\perp = \mathbb{R} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Uma base ON de $F^\perp$ é composto do vetor $\frac{1}{\sqrt{2}}(1, 0, -1)$.

Exemplo 9.6.6. Ache uma base ON para o complemento ortogonal do subespaço F de $\mathbb{R}^3$ gerado por os dois vetores

$$\xi_1 = (3, -1, 1), \qquad \xi_2 = (-1, 2, 3).$$

9.7 Exercícios e umas soluções

Exercícios.

1. Prove que $\langle\cdot,\cdot\rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ dado por
$$((x_1, x_2), (y_1, y_2)) \mapsto 2x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2,$$
define um produto interno em $\mathbb{R}^2$.

2. **Lei do paralelogramo.** Seja $|\cdot| := \sqrt{\langle\cdot,\cdot\rangle}$ a norma induzida de um produto interno num espaço vetorial E. Prove o lei do paralelogramo
$$|u+v|^2 + |u-v|^2 = 2\,|u|^2 + 2\,|v|^2 \tag{9.7.1}$$
Interprete (9.7.1) geometricamente mediante um desenho.
No ano 1935 J. v. Neumann e P. Jordan descobriram que, vice versa, uma norma $|\cdot|$ já é induzida de um produto interno quando ela satisfaz o lei do paralelogramo. Neste caso $\langle u, v\rangle := \frac{1}{4}\,|u+v|^2 - \frac{1}{4}\,|u-v|^2$ é um produto interno e induz $|\cdot|$.

3. Considere os vetores $u = (2, -1, 2)$, $v = (1, 2, 1)$ e $w = (-2, 3, 3)$. Determine o vetor de $\mathbb{R}^3$ que é a projeção ortogonal de w sobre o plano gerado por u e v.

4. Considere a base $\mathcal{U} = \{\xi_1, \xi_2, \xi_3\}$ de $\mathbb{R}^3$ onde
$$\xi_1 = (1, 1, 1), \qquad \xi_2 = (1, -1, 1), \qquad \xi_3 = (1, -1, -1).$$
Aplique o método de Gram-Schmidt para obter uma base ortonormal $\mathcal{Z} = \{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$. Determine a matriz $\mathbf{p}$ de passagem da base $\mathcal{U}$ para a base $\mathcal{Z}$.

5. Determine as bases obtidas de $\mathcal{U} = \{\xi_1, \xi_2, \xi_3\}$ pelo processo de Gram-Schmidt nos casos seguintes:
   (a) $\xi_1 = (3, 0, 0)$, $\xi_2 = (-1, 3, 0)$, $\xi_3 = (2, -5, 1)$;
   (b) $\xi_1 = (-1, 1, 0)$, $\xi_2 = (5, 0, 0)$, $\xi_3 = (2, -2, 3)$.

6. Sejam $F_1, F_2 \subset E$ subespaços. Prove que
   (a) $(F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp$ (b) $F_1^\perp + F_2^\perp = (F_1 \cap F_2)^\perp$.

7. Prove que o produto vetorial $\cdot \times \cdot : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ definido no Exercício 5.4.5, satisfaz:
   (a) $u \times v = -v \times u$;
   (b) $u \times (v + \tilde{v}) = u \times v + u \times \tilde{v}$;
   (c) $u \times (\alpha v) = \alpha(u \times v)$, para todo $\alpha \in \mathbb{R}$;
   (d) $u \times v \neq 0 \iff \{u, v\}$ é um conjunto LI;
   (e) $u \times v$ é ortogonal a u e ortogonal a v;
   (f) $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$.

C.6 Projeção ortogonal

Seja E um espaço vetorial com produto interno e F um subespaço.

Teorema C.6.1 (Teorema 9.5.7 – Projeção ortogonal (9.5.3))**.**

1. pr_F *é linear e bem definido (independente da base ON* $\mathcal{Y}$*).*
2. $(\mathrm{pr}_F)^2 = \mathrm{pr}_F$.
3. $\mathrm{Im}(\mathrm{pr}_F) = \mathrm{Fix}(\mathrm{pr}_F)$ *e* $E = \mathrm{Im}(\mathrm{pr}_F) \oplus \mathrm{N}(\mathrm{pr}_F)$.
4. $\mathrm{Im}(\mathrm{pr}_F) = F$.
5. $\mathrm{pr}_F|_F = I_F \in \mathcal{L}(F)$.
6. $w := (v - \mathrm{pr}_F v) \perp f \quad \forall f \in F$.
7. $\forall v \in E$ *vale*[2] $\mathrm{dist}(v, F) := \inf_{f\in F} \underbrace{\mathrm{dist}(v, f)}_{:=|v-f|} = |v - \mathrm{pr}_F v|$.

Demonstração. Sejam $k \le n$ as dimensões de $F \subset E$. 1. O axioma (BL) causa linearidade. Sejam $\mathcal{Y} = \{\varepsilon_1, \dots, \varepsilon_k\}$ e $\widetilde{\mathcal{Y}} = \{\tilde\varepsilon_1, \dots, \tilde\varepsilon_k\}$ bases ONs de F. Escrevemos os vetores $\tilde\varepsilon_j \in F$ na base $\mathcal{Y}$ de F com coeficientes α_{ij}, em símbolos

$$\tilde\varepsilon_j = \sum_{i=1}^k \varepsilon_i \alpha_{ij}, \qquad \text{note que } \langle \varepsilon_i, \tilde\varepsilon_j\rangle = \alpha_{ij},$$

onde $j = 1, \dots, k$. Conforme Proposição 9.5.5 podemos estender a base ON $\mathcal{Y}$ de F tal que obtemos uma base ON $\mathcal{Z} = \{\varepsilon_1, \dots, \varepsilon_k, \varepsilon_{k+1}, \dots, \varepsilon_n\}$ de E. Usando a mesma extensão obtemos a base ON $\widetilde{\mathcal{Z}} := \{\tilde\varepsilon_1, \dots, \tilde\varepsilon_k, \varepsilon_{k+1}, \dots, \varepsilon_n\}$ de E. Escrevendo $v \in E$ na base $\widetilde{\mathcal{Z}}$ de E na forma

$$v = \sum_{j=1}^k \tilde\varepsilon_j \tilde v_j + \sum_{J=k+1}^n \varepsilon_J v_J \tag{C.6.1}$$

e abreviamos $\sum_i = \sum_{i=1}^k$ para chegamos no nosso destino assim

$$\begin{aligned}
\sum_{i=1}^k \langle \varepsilon_i, v\rangle \varepsilon_i &= \sum_i \left\langle \varepsilon_i \sum_j \tilde\varepsilon_j \tilde v_j + \sum_{J=k+1}^n \varepsilon_J v_J \right\rangle \varepsilon_i \\
&\overset{\text{(BL)}}{=} \sum_i \left(\sum_j \underbrace{\langle \varepsilon_i, \tilde\varepsilon_j\rangle}_{\alpha_{ij}} \tilde v_j + \sum_{J=k+1}^n \underbrace{\langle \varepsilon_i, \varepsilon_J\rangle}_{0} v_J \right) \varepsilon_i \\
&= \sum_{i,j} \alpha_{ij} \tilde v_j \varepsilon_i \\
&= \sum_j \tilde v_j \tilde\varepsilon_j \\
&\overset{(9.3.2)}{=} \sum_{j=1}^k \langle \tilde\varepsilon_j, v\rangle \tilde\varepsilon_j .
\end{aligned}$$

[2] como $\dim F < \infty$ um ínfimo é um mínimo

2. Dado $v \in E$, use a definição (9.5.3) duas vezes para obter

$$\begin{aligned}
\mathrm{pr}_F(\mathrm{pr}_F v) &= \sum_{j=1}^k \left\langle \varepsilon_j, \sum_{i=1}^k \langle \varepsilon_i, v\rangle \varepsilon_i \right\rangle \varepsilon_j \\
&= \sum_{i,j=1}^k \langle \varepsilon_i, v\rangle \underbrace{\langle \varepsilon_j, \varepsilon_i\rangle}_{\delta_{ij}} \varepsilon_j \\
&= \sum_{ij=1}^k \langle \varepsilon_i, v\rangle \varepsilon_i \overset{\text{def.}}{=} \mathrm{pr}_F v.
\end{aligned}$$

3. Lema 7.1.2.
4. "$\subset$"Óbvio. "$\supset$"Escreve $f \in F$ como CL na base ON $\mathcal{Y}$, ou seja

$$\begin{aligned}
f = f_1\varepsilon_1 + \dots + f_k\varepsilon_k &\overset{\text{ON}}{=} f_1 \sum_i \langle \varepsilon_i, \varepsilon_1\rangle \varepsilon_i + \dots + f_k \sum_i \langle \varepsilon_i, \varepsilon_k\rangle \varepsilon_i \\
&= f_1\, \mathrm{pr}_F \varepsilon_1 + \dots + f_k\, \mathrm{pr}_F \varepsilon_k \\
&\overset{\text{lin.}}{=} \mathrm{pr}_F(f_1\varepsilon_1 + \dots + f_k\varepsilon_k) \\
&= \mathrm{pr}_F f \in \mathrm{Im}(\mathrm{pr}_F).
\end{aligned}$$

5. São exatamente os pontos fixos $\mathrm{Fix}(\mathrm{pr}_F) \overset{3.}{=} \mathrm{Im}(\mathrm{pr}_F) \overset{4.}{=} F$ nos quais uma aplicação age como a identidade.
6. Escrevendo $v \in E$ na forma (C.6.1) obtemos

$$v - \mathrm{pr}_F v = \sum_j \varepsilon_j v_j + \sum_J \varepsilon_J v_J - \sum_j \underbrace{\langle \varepsilon_j, v\rangle}_{v_j} \varepsilon_j = \sum_J \varepsilon_J v_J.$$

Escrevendo $f \in F$ na forma $f = \sum_{i=1}^k \varepsilon_i f_i$ obtemos

$$\langle v - \mathrm{pr}_F v, f\rangle = \sum_{J=k+1}^n \sum_{i=1}^k v_J \underbrace{\langle \varepsilon_J, \varepsilon_i\rangle}_{0} f_i = 0.$$

7. Dado $v \in E$ e $f \in F$, definindo $w := v - \mathrm{pr}_F v$ e $\tilde f := \mathrm{pr}_F v - f \in F$ obtemos que $v - f = w + \tilde f$. Como $w \perp \tilde f$ segundo item 6 o Pitágoras generalizado diz

$$|v - f|^2 \overset{(9.3.3)}{=} |v - \mathrm{pr}_F v|^2 + |\mathrm{pr}_F v - f|^2 \geq |v - \mathrm{pr}_F v|^2 .$$

Assim $\inf_{f \in F} |v - f| \geq |v - \mathrm{pr}_F v|$. Mas a desigualdade oposta vale também porque $\mathrm{pr}_F v$ é elemento de F. □

Exercício C.6.2 (Exercício 9.6.2)**.** Seja $X \subset E$ um subconjunto não-vazio.

(i) O complemento ortogonal $X^\perp$ é um subespaço de E.

(ii) Ou $X^\perp$ é disjunto a X, ou $X^\perp \cap X = \{\mathcal{O}\}$.

(iii) $Y \subset X \Rightarrow X^\perp \subset Y^\perp$.

(iv) $X^\perp = \langle X \rangle^\perp$.

Solução. (i) A condição para um $v \in E$ de ser elemento de $X^\perp$ é linear. Consequentemente $X^\perp$ é fechado sob adição e multiplicação linear.
(ii) Caso $X \cap X^\perp = \emptyset$: Este caso aparece, por exemplo $X = \{x\}$ onde $x \neq \mathcal{O}$.
Caso $X \cap X^\perp \neq \emptyset$: Seja $y \in X \cap X^\perp$. Como $y \in X^\perp$ vale que $\langle x, y \rangle = 0 \; \forall x \in X$.
Escolha $x = y \in X$ para obter $0 = \langle y, y \rangle$, assim $y = \mathcal{O}$ segundo axioma (`(POS)`).
(iii) Dado $v \in X^\perp$, então $\langle v, x \rangle = 0 \; \forall x \in X$. Obviamente esta condição é satisfeita para os elementos y de um subconjunto Y de X.
(iv) "$\subset$"Seja $v \in X^\perp$, então $\langle v, x \rangle = 0 \; \forall x \in X$. Mas esta condição é linear em x e por isso fica valida para combinações lineares em X. "$\supset$" Item (iii) para a inclusão $X \subset \langle X \rangle$. □