Notas de aula de Álgebra Linear

Programa de Verão 2012

B.6 - Álgebra Linear (turma 2)

Gustavo de Lima Prado - glprado@ime.usp.br

v1.1

Espaço vetorial

Definição 0.1. *Seja X um conjunto não vazio. Uma* **operação binária sobre** X *é uma aplicação*

$$* : X \times X \to X$$

que, a cada par de elementos $(x, y) \in X \times X$, associa um elemento $x * y \in X$.

Exemplo 0.2. *Seja $\mathbb{N} := \{0, 1, 2, 3, \ldots\}$ o conjunto dos números naturais. Então a adição entre números naturais*

$$\begin{array}{rccc} + : & \mathbb{N} \times \mathbb{N} & \to & \mathbb{N} \\ & (m, n) & \mapsto & m + n \end{array}$$

é uma operação binária sobre $\mathbb{N}$.

Observação 0.3. *Denotaremos $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$.*

Definição 0.4. *Um conjunto não vazio $\mathbb{K}$ é um* **corpo** *se pudermos definir duas operações binárias sobre $\mathbb{K}$, $+$ e $\cdot$, satisfazendo:*

(a1) $\alpha + \beta = \beta + \alpha$, $\alpha, \beta \in \mathbb{K}$

(a2) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$, $\alpha, \beta, \gamma \in \mathbb{K}$

(a3) existe $0 \in \mathbb{K}$ *tal que* $\alpha + 0 = \alpha$, $\alpha \in \mathbb{K}$

(a4) para todo $\alpha \in \mathbb{K}$, *existe* $-\alpha \in \mathbb{K}$ *tal que* $\alpha + (-\alpha) = 0$

(m1) $\alpha.\beta = \beta.\alpha$, $\alpha, \beta \in \mathbb{K}$

(m2) $\alpha.(\beta.\gamma) = (\alpha.\beta).\gamma$, $\alpha, \beta, \gamma \in \mathbb{K}$

(m3) existe $1 \in \mathbb{K}$ *tal que* $\alpha.1 = \alpha$, $\alpha \in \mathbb{K}$

(m4) para todo $\alpha \in \mathbb{K} \setminus \{0\}$, *existe* $\alpha^{-1} \in \mathbb{K}$ *tal que* $\alpha.\alpha^{-1} = 1$

(d) $(\alpha+\beta).\gamma = \alpha.\gamma + \beta.\gamma$, $\alpha, \beta, \gamma \in \mathbb{K}$

Exemplo 0.5. *Seja* $\mathbb{R}$ *o conjunto dos números reais. Então* $\mathbb{R}$ *com a adição entre números reais,* $+$, *e a multiplicação entre números reais,* $\cdot$, *é um corpo. Ainda, o conjunto dos números complexos,* $\mathbb{C} := \{a + bi : a, b \in \mathbb{R}\}$, *com a adição e a multiplicação entre números complexos também é um corpo.*
Entretanto $\mathbb{N}$ *com a adição entre números naturais,* $+$, *e a multiplicação entre números naturais,* $\cdot$, *não é um corpo. De fato, já não vale (a4) (na verdade, não valem (a4) e (m4)).*
Mostre que o conjunto dos números inteiros $\mathbb{Z}$ *com a adição e a multiplicação entre números inteiros não é um corpo.*

Observação 0.6. *De agora em diante, procuraremos denotar a multiplicação entre dois elementos* $\alpha, \beta \in \mathbb{K}$, $\mathbb{K}$ *um corpo, simplesmente por* $\alpha\beta := \alpha.\beta$. *Ainda, observamos que, neste texto, os elementos de um corpo são chamados de escalares.*

Definição 0.7. *Um conjunto não vazio* V *é um* **espaço vetorial sobre** *(um corpo)* $\mathbb{K}$ *se pudermos definir uma operação binária,* $\boldsymbol{+}$, *e uma multiplicação por escalar,* $\cdot$, *sobre* V*:*

$$\begin{array}{rccc} \boldsymbol{+}: & V \times V & \to & V \\ & (u,v) & \mapsto & u\boldsymbol{+}v \\ \cdot: & \mathbb{K} \times V & \to & V \\ & (\alpha, v) & \mapsto & \alpha.v \end{array}$$

satisfazendo:

(A1) $u\boldsymbol{+}v = v\boldsymbol{+}u$, $u, v \in V$

(A2) $u\boldsymbol{+}(v\boldsymbol{+}w) = (u\boldsymbol{+}v)\boldsymbol{+}w$, $u, v, w \in V$

(A3) existe $\vec{0} \in V$ *tal que* $u\boldsymbol{+}\vec{0} = u$, $u \in V$

(A4) para todo $u \in V$*, existe* $-u \in V$ *tal que* $u\mathbf{+}(-u) = \overrightarrow{0}$

(M1) $\alpha.(\beta.u) = (\alpha\beta).u$*,* $\alpha, \beta \in \mathbb{K}$*,* $u \in V$

(M2) $1.u = u$*,* $u \in V$*, onde* 1 *é o elemento dado por (m3)*

(D1) $\alpha.(u\mathbf{+}v) = \alpha.u\mathbf{+}\alpha.v$*,* $\alpha \in \mathbb{K}$*,* $u, v \in V$

(D2) $(\alpha + \beta).u = \alpha.u\mathbf{+}\beta.u$*,* $\alpha, \beta \in \mathbb{K}$*,* $u \in V$

Observamos que os elementos de um espaço vetorial são chamados de vetores e, em particular, $\overrightarrow{0}$ *é chamado de o vetor nulo de* V*. Ainda, um espaço vetorial sobre* $\mathbb{K}$ *também é chamado de um* $\mathbb{K}$*-espaço vetorial.*

Vejamos agora alguns exemplos de espaços vetoriais.

Exemplo 0.8. *Todo corpo* $\mathbb{K}$ *é* $\mathbb{K}$*-espaço vetorial.*
Com efeito, seja $\mathbb{K}$ *um corpo. Então temos definidas duas operações binárias,* $+$ *e* $\cdot$*, sobre* $\mathbb{K}$*, satisfazendo as propriedades mencionadas na definição de corpo. Daí, pelas propriedades de (a1) até (a4), verificamos que* $+$ *satisfaz as propriedades desde (A1) até (A4). Agora, por (m2), notamos que* $\cdot$ *satisfaz (M1) e, por (m1) e (m3), vemos que* $1.u = u.1 = u$*, para todo* $u \in \mathbb{K}$*, donde* $\cdot$ *satisfaz* $(M2)$*. Verifique agora que* $+$ *e* $\cdot$ *satisfazem (D1) e (D2).*

Exemplo 0.9. *Sejam* $\mathbb{K}$ *um corpo e* $n \in \mathbb{N}^* = \{1, 2, 3, \ldots\}$*. Então* $\mathbb{K}^n := \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{K} \times \cdots \times \mathbb{K} : \alpha_i \in \mathbb{K}, i = 1, \ldots, n\}$ *é um espaço vetorial sobre* $\mathbb{K}$ *com a adição entre vetores e a multiplicação por escalar feitas componente a componente, isto é:*

$$\begin{array}{rccc}
\mathbf{+}: & \mathbb{K}^n \times \mathbb{K}^n & \to & \mathbb{K}^n \\
& ((a_1, \ldots, a_n), (b_1, \ldots, b_n)) & \to & (a_1, \ldots, a_n)\mathbf{+}(b_1, \ldots, b_n) := \\
& & & (a_1 + b_1, \ldots, a_n + b_n) \\
\cdot: & \mathbb{K} \times \mathbb{K}^n & \to & \mathbb{K}^n \\
& (\alpha, (b_1, \ldots, b_n)) & \to & \alpha.(b_1, \ldots, b_n) := \\
& & & (\alpha b_1, \ldots, \alpha b_n)
\end{array}$$

Verifique este fato, ou seja, mostre que $\boldsymbol{+}$ *e* $\cdot$ *satisfazem as propriedades mencionadas na definição de espaço vetorial.*

Exemplo 0.10. $\mathbb{C}$ *é um* $\mathbb{R}$*-espaço vetorial com* $\boldsymbol{+}$ *e* $\cdot$ *sendo tais que:*

$$(a+bi)\boldsymbol{+}(c+di) := (a+c)+(b+d)i,$$

$$\alpha.(c+di) := (\alpha c)+(\alpha d)i,$$

$a, b, c, d, \alpha \in \mathbb{R}$. *Verifique isto.*

Exemplo 0.11. *Seja* $\mathbb{K}$ *um corpo. Para todo* $m \in \mathbb{N}^*$*, temos que o conjunto* $\mathcal{P}_m(\mathbb{K}) := \{p(x) = a_n x^n + \cdots + a_1 x + a_0 : a_0, \ldots, a_n \in \mathbb{K}, 0 \leqslant n \leqslant m\}$ *dos polinômios de grau menor ou igual a* m *(mais o nulo) com coeficientes em* $\mathbb{K}$ *é um espaço vetorial sobre* $\mathbb{K}$ *com as operações usuais de adição entre polinômios e multiplicação por escalar.*
Em particular, $(a_3x^3+a_2x^2+a_1x+a_0)\boldsymbol{+}(b_1x+b_0) = (a_3+0)x^3+(a_2+0)x^2+(a_1+b_1)x+(a_0+b_0)$ *e* $\alpha.(b_1x+b_0) = (\alpha b_1)x+(\alpha b_0)$.
Ainda, com estas mesmas operações, o conjunto $\mathcal{P}(\mathbb{K}) := \{p(x) = a_n x^n + \cdots + a_1 x + a_0 : a_0, \ldots, a_n \in \mathbb{K}, n \geqslant 0\}$ *dos polinômios com coeficientes em* $\mathbb{K}$ *também é um espaço vetorial sobre* $\mathbb{K}$.

Exemplo 0.12. *Sejam* X *um conjunto não vazio e* $\mathbb{K}$ *um corpo. Então o conjunto das funções de* X *em* $\mathbb{K}$*,* $\mathcal{F}(X, \mathbb{K})$*, é um* $\mathbb{K}$*-espaço vetorial com* $\boldsymbol{+}$ *e* $\cdot$ *sendo tais que:*

$$(f\boldsymbol{+}g)(x) := f(x)+g(x),$$

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

$\alpha \in \mathbb{R}$, $f, g \in \mathcal{F}(X, \mathbb{K})$. *Verifique este fato, isto é, veja que* $\boldsymbol{+}$ *e* $\cdot$ *satisfazem as propriedades mencionadas na definição de espaço vetorial.*

Exemplo 0.13. *Seja* $\mathbb{K}$ *um corpo. O conjunto das matrizes* $m \times n$ *com coeficientes em* $\mathbb{K}$*,* $\mathbb{M}_{m\times n}(\mathbb{K})$ *é um espaço vetorial sobre* $\mathbb{K}$ *com a adição* *entre matrizes e a multiplicação por escalar feitas entrada a entrada. Em particular,*

$$\left(\begin{smallmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{smallmatrix}\right) \boldsymbol{+} \left(\begin{smallmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{smallmatrix}\right) = \left(\begin{smallmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{smallmatrix}\right) \in \mathbb{M}_{2\times 3}(\mathbb{K}),$$

$$\alpha.\left(\begin{smallmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{smallmatrix}\right) = \left(\begin{smallmatrix} \alpha b_{11} & \alpha b_{12} & \alpha b_{13} \\ \alpha b_{21} & \alpha b_{22} & \alpha b_{23} \end{smallmatrix}\right) \in \mathbb{M}_{2\times 3}(\mathbb{K}).$$

A fim de termos mais com o que trabalhar, vejamos agora algumas propriedades de espaços vetoriais.

Proposição 0.14. *Seja V um espaço vetorial sobre $\mathbb{K}$. Então o vetor nulo de V é único.*

Com efeito, suponhamos que existe $\tilde{0} \in V$ tal que $u+\tilde{0} = u$, para todo $u \in V$. Então, em particular, tomando $u = \vec{0}$, temos que $\vec{0} = \vec{0} + \tilde{0} \overset{(A1)}{=} \tilde{0} + \vec{0} \overset{(A3)}{=} \tilde{0}$. ■

Proposição 0.15. *Sejam V um espaço vetorial sobre $\mathbb{K}$ e $u \in V$. Então o vetor oposto de u, $-u \in V$, é único.*

De fato, suponhamos que existe $\tilde{u} \in V$ tal que $u + \tilde{u} = \vec{0}$. Então $-u \overset{(A3)}{=} (-u) + \vec{0} = (-u) + (u + \tilde{u}) \overset{(A2)}{=} ((-u) + u) + \tilde{u} \overset{(A1,A4)}{=} \vec{0} + \tilde{u} \overset{(A1,A3)}{=} \tilde{u}$. ■

Proposição 0.16. *Sejam V um espaço vetorial sobre $\mathbb{K}$, $\alpha \in \mathbb{K}$ e $u \in V$. Então $0.u = \vec{0}$ e $\alpha.\vec{0} = \vec{0}$. Ainda, se $\alpha.u = \vec{0}$, então $\alpha = 0$ ou $u = \vec{0}$.*

Mostremos que $0.u = \vec{0}$. De fato, por $(A4)$, temos que existe $-0.u \in V$ tal que $0.u + (-0.u) = \vec{0}$. Daí, $\vec{0} \overset{(A4)}{=} 0.u + (-0.u) \overset{(a3)}{=} (0+0).u + (-0.u) \overset{(D2)}{=} (0.u + 0.u) + (-0.u) \overset{(A2)}{=} 0.u + (0.u + (-0.u)) \overset{(A4)}{=} 0.u + \vec{0} \overset{(A3)}{=} 0.u$.
Mostremos agora que $\alpha.\vec{0} = \vec{0}$. Com efeito, $\alpha.\vec{0} \overset{(A3)}{=} \alpha.(\vec{0} + \vec{0}) \overset{(D1)}{=} \alpha.\vec{0} + \alpha.\vec{0}$, donde $\vec{0} \overset{(A4)}{=} \alpha.\vec{0} + (-\alpha.\vec{0}) = (\alpha.\vec{0} + \alpha.\vec{0}) + (-\alpha.\vec{0}) \overset{(A2)}{=} \alpha.\vec{0} + (\alpha.\vec{0} + (-\alpha.\vec{0})) \overset{(A4)}{=} \alpha.\vec{0} + \vec{0} \overset{(A3)}{=} \alpha.\vec{0}$.
Finalmente, seja $\alpha.u = \vec{0}$. Suponhamos que $\alpha \neq 0$. Então, por $(m4)$, existe α^{-1} tal que $\alpha\alpha^{-1} = 1$. Daí, $\vec{0} = \alpha^{-1}.\vec{0} = \alpha^{-1}.(\alpha.u) \overset{(M1)}{=} (\alpha^{-1}\alpha).u \overset{(m1,m4)}{=} 1.u \overset{(M2)}{=} u$. ■

Proposição 0.17. *Sejam V um espaço vetorial sobre $\mathbb{K}$, $\alpha \in \mathbb{K}$ e $u, v, w \in V$. Então*

1. $(-\alpha).u = \alpha.(-u) = -(\alpha.u)$
2. $-(-u) = u$
3. $u + v = u + w \Rightarrow v = w$

Exercício. ∎

Observação 0.18. *Denotaremos $u - v := u + (-v)$. Ainda, muitas vezes, denotaremos $\alpha u := \alpha.u$, isto é, também omitiremos o ponto da multiplicação por escalar.*

Base

Definição 0.19. *Seja V um espaço vetorial sobre $\mathbb{K}$.*

1. *um vetor $v \in V$ é uma* **combinação linear** *(finita) de vetores $v_1, \dots,$ $v_n \in V$ se existem escalares $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tais que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$.*
2. *Seja $\mathcal{G} \subset V$ um subconjunto qualquer de V. Dizemos que $\mathcal{G}$* **gera** *V (ou que $\mathcal{G}$ é um* **conjunto gerador de** *V) se todo vetor de V for uma combinação linear de elementos de $\mathcal{G}$.*

Observação 0.20. *Destaquemos que uma combinação linear sempre é feita com uma quantidade finita de vetores.*
Ainda, dizemos que uma combinação linear $\alpha_1 v_1 + \cdots + \alpha_n v_n$ é trivial se $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0 \in \mathbb{K}$.

Observação 0.21. *Se V é um $\mathbb{K}$-espaço vetorial e $\{v_1, \ldots, v_n\} \subset V$, então o conjunto de todas as combinações lineares de $v_1, \ldots, v_n$ é novamente um $\mathbb{K}$-espaço vetorial, denotado por $[v_1, \ldots, v_n]$. Note que $[v_1, \ldots, v_n] := \{\alpha_1 v_1 + \cdots + \alpha_n v_n : \alpha_1, \ldots, \alpha_n \in \mathbb{K}\}$.*
Para o caso $n = 2$, se $\gamma \in \mathbb{K}$ e $(\alpha_1 v_1 + \alpha_2 v_2), (\beta_1 v_1 + \beta_2 v_2) \in [v_1, v_2]$ são elementos quaisquer, então, usando as propriedades de V ser $\mathbb{K}$-espaço vetorial, obtemos que $(\alpha_1 v_1 + \alpha_2 v_2) + (\beta_1 v_1 + \beta_2 v_2) = (\alpha_1 + \beta_1) v_1 + (\alpha_2 + \beta_2) v_2 \in [v_1, v_2]$ e $\gamma(\beta_1 v_1 + \beta_2 v_2) = (\gamma\beta_1) v_1 + (\gamma\beta_2) v_2 \in [v_1, v_2]$. Convença-se de que, com esta adição e com esta multiplicação por escalar, $[v_1, v_2]$ é um $\mathbb{K}$-espaço vetorial.

Vejamos alguns exemplos.

Exemplo 0.22. *Sejam $\mathbb{K}$ um corpo e $n \in \mathbb{N}^*$. Então o conjunto das n-uplas $\{(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\} \subset \mathbb{K}^n$ é um conjunto gerador de $\mathbb{K}^n$ (sobre $\mathbb{K}$). Com efeito, seja $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{K}^n$ é um elemento qualquer. Como $(\alpha_1, \alpha_2, \ldots, \alpha_n) = \alpha_1.(1, 0, \ldots, 0) + \cdots + \alpha_n.(0, \ldots, 0, 1)$, segue que um elemento qualquer de $\mathbb{K}^n$ é uma combinação linear de elementos do conjunto acima.*

Exemplo 0.23. *Consideremos $\mathbb{R}^2$ como sendo espaço vetorial (sobre $\mathbb{R}$). Então $\mathcal{C} := \{(1, 0), (0, 1), (1, 1)\} \subset \mathbb{R}^2$ gera $\mathbb{R}^2$. De fato, seja $(r, s) \in \mathbb{R}^2$. Como existem $\alpha, \beta, \gamma \in \mathbb{R}$, $\gamma = \frac{s}{2}$, $\beta = \frac{s}{2}$ e $\alpha = r - \frac{s}{2}$, tais que $(r, s) = \alpha(1, 0) + \beta(0, 1) + \gamma(1, 1)$, segue que $\mathcal{C}$ gera $\mathbb{R}^2$. Assim, por este exemplo e pelo anterior, vemos que um espaço vetorial (sobre algum corpo) pode admitir diferentes conjuntos geradores.*

Exemplo 0.24. $\{1, i\} \subset \mathbb{C}$ *é um conjunto gerador de $\mathbb{C}$ (sobre $\mathbb{R}$).*

Exemplo 0.25. $\{1, x, x^2, \ldots, x^m\} \subset \mathcal{P}_m(\mathbb{K})$ *é um conjunto gerador de $\mathcal{P}_m(\mathbb{K})$ (sobre $\mathbb{K}$) e $\{1, x, x^2, \ldots\} \subset \mathcal{P}(\mathbb{K})$ é um conjunto gerador de $\mathcal{P}(\mathbb{K})$ (sobre $\mathbb{K}$).*

Exemplo 0.26. $\{(1, 0), (0, 1)\} \subset \mathbb{C}^2$ *não é um conjunto gerador de $\mathbb{C}^2$ sobre $\mathbb{R}$, pois existe $(i, 0) \in \mathbb{C}^2$ tal que $(i, 0) \neq \alpha(1, 0) + \beta(0, 1)$, para todo*

$\alpha, \beta \in \mathbb{R}$. *Aqui vemos que a estrutura do espaço vetorial (o conjunto não vazio, o corpo e as operações satisfazendo as propriedades) é essencial na hora de dizermos se um conjunto é ou não gerador dele. Verificamos que* $\{(1,0),(0,1)\}$ *não é gerador. Precisamos então de mais "gente". Por exemplo,* $\{(1,0),(0,1),(i,0),(0,i)\}$ *já é um conjunto gerador de* $\mathbb{C}^2$ *sobre* $\mathbb{R}$. *Verifique.*

Proposição 0.27. *Seja* V *um* $\mathbb{K}$*-espaço vetorial. Se* $\mathcal{B} \subset \mathcal{C} \subset V$ *e* $\mathcal{B}$ *é um conjunto gerador de* V*, então* $\mathcal{C}$ *também o é.*

De fato, seja $u \in V$. Como $\mathcal{B}$ gera V, então existem $\alpha_i \in \mathbb{K}$, $u_i \in \mathcal{B}$, com $i = 1, \ldots, m$, tais que $u = \alpha_1 u_1 + \cdots + \alpha_m u_m$. Mas $\mathcal{B} \subset \mathcal{C}$. Logo, $u_i \in \mathcal{C}$, para todo $i = 1, \ldots, m$. Portanto, $\mathcal{C}$ gera V. ■

Observação 0.28. *Inicialmente, consideremos o espaço vetorial unitário* $\{\vec{0}\}$ *(defina operações,* $+$ *e* $\cdot$*, neste conjunto e convença-se de que ele é um espaço vetorial (sobre qualquer corpo)). Feito isto, temos que conjunto* $\{\vec{0}\} \subset \{\vec{0}\}$ *gera o espaço vetorial* $\{\vec{0}\}$*, pois sempre é possível escrevermos a seguinte combinação linear:* $\vec{0} = 0.\ \vec{0}$.
De um modo mais geral, se V *é um espaço vetorial sobre* $\mathbb{K}$*, então o conjunto* $V \subset V$ *sempre é um conjunto gerador do espaço vetorial* V*. De fato, seja* $v \in V$ *um vetor qualquer. Desde que, por* $(M2)$*,* $v = 1.v$*, segue que* v *é uma combinação linear de elementos de* V *(só precisamos usar o próprio vetor). Portanto,* V *sempre gera* V*. Será que existe* $\mathcal{G} \subset V$ *tal que* $\mathcal{G} \neq V$ *e* $\mathcal{G}$ *gera* V*?*

Definição 0.29. *Sejam* V *um* $\mathbb{K}$*-espaço vetorial e* $\mathcal{B} \subset V$ *um subconjunto qualquer. Dizemos que* $\mathcal{B}$ *é* **linearmente independente** *(LI) se toda combinação linear de vetores de* $\mathcal{B}$ *resultando no vetor nulo for tal que os escalares já sejam todos nulos, isto é, se* $\alpha_1 v_1 + \cdots + \alpha_n v_n = \vec{0}$*, com* $\alpha_i \in \mathbb{K}$*,* $v_i \in \mathcal{B}$*, implicar que* $\alpha_1 = \cdots = \alpha_n = 0$*. E dizemos que* $\mathcal{B}$ *é* **linearmente dependente** *(LD) se não for LI.*

Observação 0.30. *Explicitamente, dizemos que $\mathcal{B} \subset V$, V um $\mathbb{K}$-espaço vetorial, é LD quando o vetor nulo for uma combinação linear não trivial de elementos de $\mathcal{B}$.*

Observação 0.31. *Todo conjunto contendo o vetor nulo é LD. Por quê? Sejam V um espaço vetorial (sobre algum corpo) e $v \in V$ um vetor não nulo. Então $\{v\}$ é LI. Por quê?*

Exemplo 0.32. *Notemos que se consideramos $\mathbb{C}$ como $\mathbb{C}$-espaço vetorial obtemos que $\{1, i\} \subset \mathbb{C}$ é LD e se consideramos $\mathbb{C}$ como $\mathbb{R}$-espaço vetorial segue que $\{1, i\} \subset \mathbb{C}$ é LI. Aqui verificamos que a estrutura do espaço vetorial (o conjunto não vazio, o corpo e as operações satisfazendo as propriedades) é crucial no momento de dizermos se um conjunto é ou não LI.*

Exemplo 0.33. $\{\mathrm{sen} x, \cos x\}$ *subconjunto do $\mathbb{R}$-espaço vetorial das funções contínuas de $[0, 2\pi]$ em $\mathbb{R}$, $\mathcal{C}([0, 2\pi], \mathbb{R})$, é LI. De fato, sejam $\alpha, \beta \in \mathbb{R}$ tais que $\alpha \mathrm{sen} x + \beta \cos x = \vec{0}$. Mas esta igualdade deve valer para todo o domínio das funções. Logo, tomando ora $x = 0$, ora $x = \pi/2$, obtemos que $\beta = 0 = \alpha$.*

Proposição 0.34. *Seja V um $\mathbb{K}$-espaço vetorial. Se $\mathcal{B} \subset \mathcal{C} \subset V$ e $\mathcal{C}$ é um conjunto LI, então $\mathcal{B}$ também o é.*

Com efeito, sejam $\alpha_i \in \mathbb{K}$, $u_i \in \mathcal{B}$, com $i = 1, \ldots, m$, tais que $\alpha_1 u_1 + \cdots + \alpha_m u_m = \vec{0}$. Mas $\mathcal{B} \subset \mathcal{C}$. Logo, $u_i \in \mathcal{C}$, para todo $i = 1, \ldots, m$. Como $\mathcal{C}$ é LI, segue que $\alpha_1 = \cdots = \alpha_m = 0 \in \mathbb{K}$. Portanto, $\mathcal{B}$ é LI. ■

Definição 0.35. *Seja V um espaço vetorial (sobre algum corpo). Dizemos que um subconjunto $\mathcal{B} \subset V$ é uma* **base** *de V se $\mathcal{B}$ for um conjunto gerador de V e LI.*

Retomemos alguns exemplos.

Exemplo 0.36. *Sejam $\mathbb{K}$ um corpo e $n \in \mathbb{N}^*$. Então o conjunto das n-uplas $\mathcal{B}_C := \{(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\} \subset \mathbb{K}^n$ é uma base de $\mathbb{K}^n$ (sobre $\mathbb{K}$), chamada de base canônica de $\mathbb{K}^n$.*

Exemplo 0.37. $\{(1,2),(3,1)\} \subset \mathbb{R}^2$ *é uma base de* $\mathbb{R}^2$ *(sobre* $\mathbb{R}$*). De fato, mostremos inicialmente que* $\mathcal{A} := \{(1,2),(3,1)\}$ *gera* $\mathbb{R}^2$*. Seja* $(r,s) \in \mathbb{R}^2$ *um elemento qualquer. Escrevendo* $(r,s) = x(1,2) + y(3,1)$*, com* $x, y \in \mathbb{R}$*, e resolvendo o sistema (em* x *e* y*), obtemos que* $x = \frac{3s-r}{5}$ *e* $y = \frac{2r-s}{5}$ *são escalares tais que* (r,s) *é combinação linear de vetores de* $\mathcal{A}$*. Vejamos agora que* $\mathcal{A}$ *é LI. Escrevendo* $(0,0) = \alpha(1,2) + \beta(3,1)$*, com* $\alpha, \beta \in \mathbb{R}$*, segue que* $\alpha + 3\beta = 0$ *e* $2\alpha + \beta = 0$*, donde* $\alpha = -3\beta$ *e* $-5\beta = 0$*. Logo, devemos ter* $\alpha = \beta = 0$*. Portanto,* $\mathcal{A}$ *é base. Notemos que, pelo exemplo anterior,* $\mathcal{B}_C = \{(1,0),(0,1)\}$ *também é base de* $\mathbb{R}^2$*. Logo, um espaço vetorial pode ter mais de uma base.*

Exemplo 0.38. $\{1, i\} \subset \mathbb{C}$ *é uma base de* $\mathbb{C}$ *(considerado como* $\mathbb{R}$*-espaço vetorial).*

Exemplo 0.39. $\{1, x, x^2, \ldots, x^m\}$ *é uma base de* $\mathcal{P}_m(\mathbb{K})$ *e* $\{1, x, x^2, \ldots\}$ *é uma base de* $\mathcal{P}(\mathbb{K})$*, chamadas de bases canônicas respectivamente de* $\mathcal{P}_m(\mathbb{K})$ *e* $\mathcal{P}(\mathbb{K})$*. Verifique que* $\{1, x, x^2+1\} \subset \mathcal{P}_2(\mathbb{K})$ *também é uma base de* $\mathcal{P}_2(\mathbb{K})$*.*

Espaço vetorial finitamente gerado

Definição 0.40. *Dizemos que um espaço vetorial (sobre algum corpo) é* **finitamente gerado** *se admitir um conjunto gerador finito.*

Proposição 0.41. *Seja* V *um* $\mathbb{K}$*-espaço vetorial finitamente gerado por um conjunto* $\{v_1, \ldots, v_m\} \subset V$*. Então todo subconjunto LI de* V *tem no máximo* m *elementos.*

Seja $\mathcal{A} := \{u_1, \ldots, u_n\} \subset V$, com $n > m$. Mostremos que $\mathcal{A}$ é LD. Seja $1 \leqslant j \leqslant n$. Então $u_j = \alpha_{1j}v_1 + \cdots \alpha_{mj}v_m$ (*), para algum $\alpha_{1j}, \ldots, \alpha_{mj} \in \mathbb{K}$, pois $\{v_1, \ldots, v_m\}$ é gerador. Consideremos uma combinação linear qualquer de vetores de $\mathcal{A}$, $\lambda_1 u_1 + \cdots + \lambda_n u_n$, com $\lambda_i \in \mathbb{K}$, para todo $i =$

$1,\ldots,n$. Então $\lambda_1 u_1 + \cdots + \lambda_n u_n = \sum_{j=1}^{n} \lambda_j u_j \overset{(*)}{=} \sum_{j=1}^{n} \lambda_j \big(\sum_{i=1}^{m} \alpha_{ij} v_i\big) \overset{(D1)}{=} \sum_{j=1}^{n} \sum_{i=1}^{m} \lambda_j (\alpha_{ij} v_i) \overset{(M1)}{=} \sum_{j=1}^{n} \sum_{i=1}^{m} (\lambda_j \alpha_{ij}) v_i = \sum_{i=1}^{m} \big(\sum_{j=1}^{n} (\lambda_j \alpha_{ij})\big) v_i$ (**). Agora, simplesmente, consideremos o seguinte sistema de equações:

$$\begin{cases} \sum_{j=1}^{n} (\lambda_j \alpha_{1j}) & = & 0 \\ \sum_{j=1}^{n} (\lambda_j \alpha_{2j}) & = & 0 \\ & \vdots & \\ \sum_{j=1}^{n} (\lambda_j \alpha_{mj}) & = & 0 \end{cases}$$

Como este sistema homogêneo tem mais incógnitas $\lambda_1, \ldots, \lambda_n$ do que equações (pois $n > m$), segue que ele admite uma solução não trivial, ou seja, existem $\gamma_1, \ldots, \gamma_n \in \mathbb{K}$ não todos nulos tais que $\sum_{j=1}^{n} (\gamma_j \alpha_{ij}) = 0$, para todo $i = 1, \ldots, m$. Logo, tomando $\lambda_j = \gamma_j$ em (**), obtemos $\gamma_1 u_1 + \cdots + \gamma_n u_n = \vec{0}$, com $\gamma_1, \ldots, \gamma_n \in \mathbb{K}$ não todos nulos. Logo, $\mathcal{A}$ é LD. ∎

Exercício 0.42. *Usando as propriedades de espaço vetorial, verifique que* $\sum_{j=1}^{3} \big(\sum_{i=1}^{2} (\lambda_j \alpha_{ij}) v_i\big) = \sum_{i=1}^{2} \big(\sum_{j=1}^{3} (\lambda_j \alpha_{ij})\big) v_i$.

Corolário 0.43. *Seja $V \neq \{\vec{0}\}$ um $\mathbb{K}$-espaço vetorial finitamente gerado. Então duas bases quaisquer de V têm o mesmo número de elementos.*

De fato, sejam $\mathcal{C}, \mathcal{D}$ duas bases de V. Desde que V é finitamente gerado e $\mathcal{C}$ e $\mathcal{D}$ são LI, então, pela proposição anterior, $\mathcal{C}$ e $\mathcal{D}$ são conjuntos finitos (digamos que com $c > 0$ e $d > 0$ elementos respectivamente). Mas $\mathcal{C}$ é um conjunto LI num espaço vetorial finitamente gerado por $\mathcal{D}$, donde $c \leqslant d$. Analogamente, $\mathcal{D}$ é um conjunto LI num espaço vetorial finitamente gerado $\mathcal{C}$, donde $d \leqslant c$. Portanto, $c = d$. ∎

Definição 0.44. *Seja V um $\mathbb{K}$-espaço vetorial. Se V for finitamente gerado, chamamos o número de elementos de uma base de V simplesmente de* **dimensão de** *V (sobre $\mathbb{K}$). Caso contrário, isto é, se V não for finitamente gerado, dizemos que a dimensão de V (sobre $\mathbb{K}$) é infinita. Notação:* $\dim_{\mathbb{K}} V$.

Observação 0.45. *Usando a convenção de que o conjunto vazio* $\emptyset$ *é uma base para o espaço vetorial unitário* $U := \{\vec{0}\}$, *temos que* $\dim_{\mathbb{K}} U = 0$. *De fato, os únicos subconjuntos de* U *são* $\emptyset$ *e* $\{\vec{0}\}$. *Como* $\{\vec{0}\} \subset U$ *gera* U, *mas é LD, segue que não pode ser base de* U. *Daí, a única base de* U *é* $\emptyset$ *e faz sentido escrevermos* $\dim_{\mathbb{K}}\{\vec{0}\} = 0$ *(que é o número de elementos de uma base qualquer de* $\{\vec{0}\}$*). Na verdade,* $\dim_{\mathbb{K}} V = 0$ *se, e somente se,* $V = \{\vec{0}\}$.

Exemplo 0.46.
$\dim_{\mathbb{K}} \mathbb{K} = 1$,
$\dim_{\mathbb{C}} \mathbb{C} = 1$,
$\dim_{\mathbb{R}} \mathbb{C} = 2$,
$\dim_{\mathbb{K}} \mathbb{K}^n = n$,
$\dim_{\mathbb{K}} \mathcal{P}_m(\mathbb{K}) = m + 1$, *pois* $\{1, x, \ldots, x^m\} \subset \mathcal{P}_m(\mathbb{K})$ *é base,*
$\dim_{\mathbb{K}} \mathcal{P}(\mathbb{K}) = \infty$, *pois* $\{1, x, \ldots\} \subset \mathcal{P}(\mathbb{K})$ *é base,*
$\dim_{\mathbb{K}} \mathbb{M}_{m\times n}(\mathbb{K}) = m.n$. *Por exemplo, para o caso* $m = n = 2$, $\{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)\} \subset \mathbb{M}_{2\times 2}(\mathbb{K})$ *é base, chamada de base canônica de* $\mathbb{M}_{2\times 2}(\mathbb{K})$, *donde* $\dim_{\mathbb{K}} \mathbb{M}_{2\times 2}(\mathbb{K}) = 4 = 2.2$. *A partir deste exemplo, que subconjunto de* $\mathbb{M}_{3\times 2}(\mathbb{K})$ *seria a base canônica de* $\mathbb{M}_{3\times 2}(\mathbb{K})$*?*

Corolário 0.47. *Seja* V *um* $\mathbb{K}$*-espaço vetorial, com* $\dim_{\mathbb{K}} V = n > 0$. *Então*

1. *todo conjunto de vetores de* V *com mais do que* n *vetores é LD.*

2. *todo conjunto de vetores de* V *com menos do que* n *vetores não gera* V.

Exercício (use a proposição anterior). ∎

Proposição 0.48. *Seja* V *um* $\mathbb{K}$*-espaço vetorial. Seja* $\mathcal{C} \subset V$, $\mathcal{C}$ *LI. Se existe* $v \in V$ *que não é combinação linear de elementos de* $\mathcal{C}$, *então* $\mathcal{C} \cup \{v\}$ *é LI.*

De fato, seja $v \in V$ qualquer. Seja $c := \alpha_1 v_1 + \cdots + \alpha_n v_n + \beta v$ uma combinação linear qualquer de vetores de $\mathcal{C} \cup \{v\}$. Suponhamos que c é o vetor nulo. Se $\beta \neq 0$, então $v = \frac{\alpha_1}{-\beta} v_1 + \cdots + \frac{\alpha_n}{-\beta} v_n$. Logo, se v não é combinação linear de elementos de $\mathcal{C}$, então β deve ser 0. Daí, $\alpha_1 v_1 + \cdots + \alpha_n v_n = \vec{0}$ e portanto, como $\mathcal{C}$ é LI e $v_1, \ldots, v_n \in \mathcal{C}$, temos que $\alpha_1 = \cdots = \alpha_n = 0$. Assim $\mathcal{C} \cup \{v\}$ é LI. ■

Teorema 0.49. *Todo espaço vetorial finitamente gerado possui uma base.*

De fato, seja V um $\mathbb{K}$-espaço vetorial finitamente gerado. Suponhamos que V é gerado por um conjunto com $n > 0$ elementos. Então existe $v_1 \in V$, $v_1 \neq \vec{0}$. Temos que $\{v_1\}$ é LI. Se $\{v_1\}$ gera V, então $\{v_1\}$ é base de V e acabou. Caso contrário, então existe $v_2 \in V$ tal que v_2 não é combinação linear de v_1 e pela preposição anterior $\{v_1, v_2\}$ é LI. Se $\{v_1, v_2\}$ gera V, então é base e acabou. Senão, como, por uma proposição anterior, não é possível termos um subconjunto LI de V com $n+1$ elementos, então, pelo menos no n-ésimo passo, obteremos um conjunto $\{v_1, v_2, \ldots, v_n\}$ (LI pela preposição anterior aplicada várias vezes) que deverá gerar V e portanto será base de V (notemos que este processo pode parar antes de n). Finalmente, observemos que se V é gerado por um conjunto com 0 elementos, então $V = \{\vec{0}\}$ e $\emptyset$ é uma base de $\{\vec{0}\}$. ■

Usando a mesma idéia da demonstração acima, verifiquemos o seguinte resultado:

Proposição 0.50. *Seja V um $\mathbb{K}$-espaço vetorial finitamente gerado. Seja $\mathcal{C} \subset V$, $\mathcal{C}$ LI. Então existe uma base de V que contém $\mathcal{C}$.*

Com efeito, suponhamos que V é gerado por um conjunto com $m \geqslant 0$ vetores. Se $\mathcal{C}$ gera V, então $\mathcal{C}$ é base de V e acabou. Caso contrário, então existe $u_1 \in V$ tal que u_1 não é combinação linear de elementos de $\mathcal{C}$ e portanto $\mathcal{C} \cup \{u_1\}$ é LI. Se $\mathcal{C} \cup \{u_1\}$ gera V, então $\mathcal{C} \cup \{u_1\}$ é base de V e acabou. Senão, desde que não é possível termos um subconjunto LI de V com $m+1$ vetores, então, pelo menos no m-ésimo passo, obteremos um conjunto $\mathcal{C}\cup\{u_1,\ldots,u_m\}$ (LI pela preposição anterior aplicada várias vezes) que deverá gerar V e portanto será base de V (notemos que este processo pode parar antes de m). ■

Exemplo 0.51. *Consideremos* $(1,3,1)\in\mathbb{R}^3$. *Como exibir uma base de* $\mathbb{R}^3$ *que contém* $\mathcal{C}:=\{(1,3,1)\}$*? Consideremos o conjunto de todas as combinações lineares de* $(1,3,1)$*, isto é,* $[(1,3,1)]=\{(\alpha,3\alpha,\alpha)|\alpha\in\mathbb{R}\}$*. Tomando, por exemplo,* $(1,3,0)\in\mathbb{R}^3$*, temos que* $(1,3,0)\neq(\alpha,3\alpha,\alpha)$*, para todo* $\alpha\in\mathbb{R}$*, donde* $\mathcal{D}:=\mathcal{C}\cup\{(1,3,0)\}$ *é LI. Consideremos agora o conjunto de todas as combinações lineares de* $(1,3,1),(1,3,0)$*, isto é,* $[(1,3,1),(1,3,0)]=\{(\alpha+\beta,3\alpha+3\beta,\alpha)|\alpha,\beta\in\mathbb{R}\}$*. Tomando, por exemplo,* $(1,0,0)\in\mathbb{R}^3$*, segue que* $(1,0,0)\neq(\alpha+\beta,3\alpha+3\beta,\alpha)$*, para todo* $\alpha,\beta\in\mathbb{R}$*, donde* $\mathcal{E}:=\mathcal{C}\cup\{(1,3,0),(1,0,0)\}$ *é LI. Como existe uma base de* $\mathbb{R}^3$ *que contém* $\mathcal{E}$ *e como toda base de* $\mathbb{R}^3$ *tem exatamente 3 elementos (*$\dim_{\mathbb{R}}\mathbb{R}^3=3$*), então* $\mathcal{E}$ *é uma base de* $\mathbb{R}^3$.

Proposição 0.52. *Seja* V *um* $\mathbb{K}$*-espaço vetorial, com* $\dim_{\mathbb{K}}V=n>0$*. Seja* $\mathcal{C}\subset V$*. Então as condições abaixo são equivalentes:*

1. $\mathcal{C}$ *é base de* V*;*

2. *cada vetor de* V *se escreve de maneira única como combinação linear de elementos de* $\mathcal{C}$.

De fato, suponhamos inicialmente que $\mathcal{C}$ é base de V. Como $\dim_{\mathbb{K}}V=n$, podemos escrever $\mathcal{C}:=\{v_1,\ldots,v_n\}$. Seja $v\in V$ qualquer. Desde que $\mathcal{C}$ gera V, existem $\alpha_1,\ldots,\alpha_n\in\mathbb{K}$ tais que $v=\alpha_1v_1+\cdots+\alpha_nv_n$. Suponhamos então que $v=\beta_1v_1+\cdots+\beta_nv_n$, para algum $\beta_1,\ldots,\beta_n\in\mathbb{K}$. Daí, $\alpha_1v_1+\cdots+\alpha_nv_n=\beta_1v_1+\cdots+\beta_nv_n$ e portanto $(\alpha_1-\beta_1)v_1+\cdots+(\alpha_n-\beta_n)v_n=\vec{0}$. Como $\mathcal{C}$ é LI, segue que $(\alpha_1-\beta_1)=\cdots=(\alpha_n-\beta_n)=0$ e portanto v se escreve de

forma única como combinação linear de elementos de $\mathcal{C}$.
Reciprocamente, suponhamos que todo vetor de V se escreve de modo único como combinação linear de elementos de $\mathcal{C}$. Daí, já obtemos que $\mathcal{C}$ gera V. Mostremos que $\mathcal{C}$ é LI. Com efeito, consideremos uma combinação linear qualquer, $c := \gamma_1 u_1 + \cdots + \gamma_m u_m$, de vetores de $\mathcal{C}$. Suponhamos que c é o vetor nulo, isto é, $\gamma_1 u_1 + \cdots + \gamma_m u_m = \overrightarrow{0}$. Mas $\overrightarrow{0} = 0u_1 + \cdots + 0u_m$. Desde que $\overrightarrow{0} \in V$ se escreve de maneira única como combinação linear de elementos de $\mathcal{C}$, então $\gamma_1 = \cdots = \gamma_m = 0$. Portanto, $\mathcal{C}$ é base de V. ■

Exemplo 0.53. *Consideremos* $\mathcal{C} := \{(1,0,0),(0,1,0),(1,1,0)\} \subset \mathbb{R}^3$. *Notemos que* $(1,1,0) = 0.(1,0,0) + 0.(0,1,0) + 1.(1,1,0)$ *e também* $(1,1,0) = 1.(1,0,0) + 1.(0,1,0) + 0.(1,1,0)$. *Daí,* $(1,1,0)$ não *se escreve de maneira única como combinação linear de elementos de* $\mathcal{C}$ *e portanto* $\mathcal{C}$ *não é base de* $\mathbb{R}^3$.

Exemplo 0.54. *Consideremos* $\mathcal{C} := \{2,\ x,\ 3x+5\} \subset \mathcal{P}_1(\mathbb{R})$. *Observemos que* $x + 1 = \frac{1}{2}.\mathbf{2} + 1.\boldsymbol{x} + 0.(\mathbf{3x+5})$ *e também* $x + 1 = (-2).\mathbf{2} + (-2).\boldsymbol{x} + 1.(\mathbf{3x+5})$. *Assim,* $x+1$ não *se escreve de forma única como combinação linear de elementos de* $\mathcal{C}$ *e portanto* $\mathcal{C}$ *não é base de* $\mathcal{P}_1(\mathbb{R})$.

Coordenadas

Usando a proposição anterior, vejamos uma forma de identificar um $\mathbb{K}$-espaço vetorial finitamente gerado qualquer com $\mathbb{K}^n$, para algum $n \in \mathbb{N}^*$.

Observação 0.55. *Seja* V *um* $\mathbb{K}$*-espaço vetorial, com* $\dim_{\mathbb{K}} V = n > 0$. *Notemos que, ao considerarmos uma base* $\mathcal{B} := \{v_1, \ldots, v_n\}$*, então, pela proposição anterior, dado* $v \in V$*, existem únicos escalares* $\alpha_1, \ldots, \alpha_n$ *tais que* $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. *Se considerarmos ainda que* $\mathcal{B}$ *é uma base ordenada, isto é, uma base onde a ordem em que os vetores são dados importa, podemos identificar* $v \in V$ *com* $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}$. *Ao considerarmos bases ordenadas,* *temos, por exemplo, que* $\mathcal{B} = \{v_1, v_2, v_3, \ldots, v_n\} \neq \{v_2, v_1, v_3, \ldots, v_n\} =: \mathcal{B}'$ *(ambos,* $\mathcal{B}$ *e* $\mathcal{B}$*', são bases, mas vistos como bases ordenadas são diferentes).*

Definição 0.56. *Seja* V *um* $\mathbb{K}$*-espaço vetorial finitamente gerado. Seja* $\mathcal{B} := \{v_1, \ldots, v_n\}$ *uma base ordenada de* V*. Seja* $v \in V$*. Dizemos que* $(\alpha_1, \ldots, \alpha_n)_{\mathcal{B}} \in \mathbb{K}^n$ *é a n-upla das coordenadas de* v *com relação à base* $\mathcal{B}$ *se* $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$*. Notação:* $[v]_{\mathcal{B}} := (\alpha_1, \ldots, \alpha_n)_{\mathcal{B}}$*.*

Observação 0.57. *Quando não houver risco de confusão, escreveremos simplesmente* $[v]_{\mathcal{B}} := (\alpha_1, \ldots, \alpha_n)$*.*

Exemplo 0.58. *Já vimos que* $\mathcal{C} := \{(1,2),(3,1)\}, \mathcal{B}_C = \{(1,0),(0,1)\} \subset \mathbb{R}^2$ *são bases de* $\mathbb{R}^2$*. Considerando-as ordenadas, temos que* $\widetilde{\mathcal{C}} := \{(3,1),(1,2)\}$*,* $\widetilde{\mathcal{B}_C} := \{(0,1),(1,0)\} \subset \mathbb{R}^2$ *são outras duas bases de* $\mathbb{R}^2$*. Consideremos* $(2,-1) \in \mathbb{R}^2$*. Temos que* $[(2,-1)]_{\mathcal{C}} = (-1,1)_{\mathcal{C}}$*, pois* $(2,-1) = (-1).(1,2) + 1.(3,1)$*. Ainda,* $[(2,-1)]_{\mathcal{B}_C} = (2,-1)_{\mathcal{B}_C}$*, pois* $(2,-1) = 2.(1,0) + (-1).(0,1)$*. Notemos ainda que* $[(2,-1)]_{\widetilde{\mathcal{C}}} = (1,-1)_{\widetilde{\mathcal{C}}}$ *e* $[(2,-1)]_{\widetilde{\mathcal{B}_C}} = (-1,2)_{\widetilde{\mathcal{B}_C}}$*. Logo, as coordenadas de um dado vetor variam conforme mudamos a base do espaço vetorial.*

Exemplo 0.59. *Seja* $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{K}^n$ *qualquer. Consideremos* $\mathcal{B}_C = \{(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)\} \subset \mathbb{K}^n$ *a base canônica de* $\mathbb{K}^n$*. Desde que* $(\alpha_1, \alpha_2, \ldots, \alpha_n) = \alpha_1(1,0,\ldots,0) + \alpha_2(0,1,0,\ldots,0) + \cdots + \alpha_n(0,\ldots,0,1)$*, então as coordenadas do vetor* $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ *com relação à* $\mathcal{B}_C$ *são* $[(\alpha_1, \alpha_2, \ldots, \alpha_n)]_{\mathcal{B}_C} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$*.*

Exemplo 0.60. *Seja* $p(x) := a_2x^2 + a_1x + a_0 \in \mathcal{P}_2(\mathbb{K})$ *qualquer. Consideremos* $\mathcal{B}_C = \{1, x, x^2\} \subset \mathcal{P}_2(\mathbb{K})$ *a base canônica de* $\mathcal{P}_2(\mathbb{K})$*. Como* $p(x) = a_2.\boldsymbol{x^2} + a_1.\boldsymbol{x} + a_0.\mathbf{1}$*, então as coordenadas do vetor* $p(x)$ *com relação à* $\mathcal{B}_C$ *são* $[p(x)]_{\mathcal{B}_C} = (a_0, a_1, a_2)$*.*

Subespaço vetorial

Definição 0.61. *Seja V um espaço vetorial sobre $\mathbb{K}$. Dizemos que um subconjunto $W \subset V$ é um* **subespaço vetorial de** V *se a restrição das operações de V a W o tornarem um espaço vetorial sobre $\mathbb{K}$.*

Observação 0.62. *Ou seja, a restrição das operações $\boldsymbol{+}$ e $\cdot$ de V a W devem ser tais que*

$$\begin{array}{rlcl} \boldsymbol{+}: & W \times W \ni (u,v) & \mapsto & u \boldsymbol{+} v \in W \\ \cdot: & \mathbb{K} \times W \ni (\alpha, v) & \mapsto & \alpha.v \in W \end{array}$$

e as propriedades de espaço vetorial sejam satisfeitas (com W no lugar de V).

Exemplo 0.63. *Se V é um $\mathbb{K}$-espaço vetorial, então $\{\vec{0}\} \subset V$ e $V \subset V$ são subespaços vetoriais de V, chamados de subespaços triviais de V.*

Proposição 0.64. *Seja V um $\mathbb{K}$-espaço vetorial. Seja $W \subset V$ um subconjunto qualquer. Então W é um subespaço vetorial de V se, e somente se, as condições abaixo valem:*

1. *se $\vec{0}$ é o vetor nulo de V, então $\vec{0} \in W$,*
2. *se $u, v \in W$, então $u \boldsymbol{+} v \in W$,*
3. *se $\alpha \in \mathbb{K}$ e $u \in W$, então $\alpha.u \in W$.*

Exercício. ■

Observação 0.65. *Se V é um $\mathbb{K}$-espaço vetorial e $\mathcal{C} \subset V$ é um subconjunto qualquer de V, então podemos considerar o conjunto de todas as combinações lineares de elementos de $\mathcal{C}$, denotado por $[\mathcal{C}]$. Então $[\mathcal{C}]$ é um subespaço vetorial de V, chamado de subespaço de V gerado por $\mathcal{C}$. Se $\mathcal{C}$ é LI, então $\mathcal{C}$ é uma base de $[\mathcal{C}]$. Notemos que se $\mathcal{C} := \{v_1, \ldots, v_n\}$ (um conjunto finito de vetores), então temos que $[\mathcal{C}] = [v_1, \ldots, v_n]$.*

Proposição 0.66. *Sejam V um $\mathbb{K}$-espaço vetorial e $W_1, W_2 \subset V$ subespaços vetoriais de V. Então $W_1 \cap W_2$ e $W_1 + W_2 := \{w_1 + w_2 : w_1 \in W_1, w_2 \in W_2\}$ também são subespaços vetoriais de V.*

Exercício (use a caracterização de subespaço vetorial dada pela proposição acima). ■

Exemplo 0.67. *Consideremos $\mathbb{C}$ como $\mathbb{R}$-espaço vetorial. Consideremos $\{1, i\} \subset \mathbb{C}$. Então $[1] \cap [i]$ e $[1] + [i]$ são subespaços vetoriais de $\mathbb{C}$. Com efeito, $[1] \cap [i] = \{0\}$ e $[1] + [i] = \mathbb{C}$. Verifique isto. Notemos que $[1] \cup [i]$ não é subespaço vetorial de $\mathbb{C}$, pois, por exemplo, $1 + i \notin [1]$ e $1 + i \notin [i]$, donde $1 + i \notin [1] \cup [i]$. Na verdade, se W_1, W_2 são subespaços de V, não segue, em geral, que $W_1 \cup W_2$ é um subespaço de V.*

Exemplo 0.68. *Consideremos $\mathbb{R}^3$ como $\mathbb{R}$-espaço vetorial. Consideremos ainda $W_1 := [(1, 0, 0), (0, 1, 0)] \subset \mathbb{R}^3$ e $W_2 := [(1, 0, 0), (0, 0, 1)] \subset \mathbb{R}^3$ subespaços vetoriais de $\mathbb{R}^3$. Então $W_1 \cap W_2$ e $W_1 + W_2$ também o são.*

Proposição 0.69. *Sejam V um $\mathbb{K}$-espaço vetorial e $W \subset V$ um subespaço vetorial de V, com $W \neq V$ e $\dim_{\mathbb{K}} W$ finita. Então $\dim_{\mathbb{K}} W < \dim_{\mathbb{K}} V$.*

Se $W = \{\vec{0}\}$, então, como $\{\vec{0}\} \neq V$, segue que $\dim_{\mathbb{K}}\{\vec{0}\} = 0 < \dim_{\mathbb{K}} V$.
Se $W \neq \{\vec{0}\}$, existe $\mathcal{B} := \{v_1, \ldots, v_n\} \subset W$ base de W, onde $n = \dim_{\mathbb{K}} W$. Daí, se $\dim_{\mathbb{K}} V = \infty$, é claro que $\dim_{\mathbb{K}} W = n < \infty$. Suponhamos agora que $\dim_{\mathbb{K}} V < \infty$ (ou seja, que V é finitamente gerado). Como $W \neq V$, então $\mathcal{B}$ não gera V, donde existe $v \in V$ tal que v não é combinação linear de $v_1, \ldots, v_n$. Daí, $\mathcal{B} \cup \{v\} \subset V$ é LI e portanto existe uma base de V que contém $\mathcal{B} \cup \{v\}$. Logo, $\dim_{\mathbb{K}} V \geqslant n + 1 > n = \dim_{\mathbb{K}} W$. ■

Proposição 0.70. *Sejam V um $\mathbb{K}$-espaço vetorial e $W_1, W_2 \subset V$ subespaços vetoriais de V, ambos de dimensão finita. Então $\dim_{\mathbb{K}}(W_1 + W_2) = \dim_{\mathbb{K}} W_1 + \dim_{\mathbb{K}} W_2 - \dim_{\mathbb{K}} W_1 \cap W_2$.*

Primeiramente, como W_1 é finitamente gerado, então $W_1 \cap W_2 \subset W_1$ também o é. Se $W_1 \cap W_2 \neq \{\vec{0}\}$, seja $\mathcal{B} := \{v_1, \ldots, v_n\} \subset W_1 \cap W_2$ uma base de $W_1 \cap W_2$. Como $\mathcal{B}$ é um conjunto LI contido em W_1, então existe $\mathcal{B}_1 \subset W_1$ uma base de W_1 que contém $\mathcal{B}$, digamos $\mathcal{B}_1 := \{v_1, \ldots, v_n, w_1, \ldots, w_p\}$. Analogamente, existe $\mathcal{B}_2 \subset W_2$ uma base de W_2 que contém $\mathcal{B}$, digamos $\mathcal{B}_2 := \{v_1, \ldots, v_n, u_1, \ldots, u_m\}$.
Consideremos agora $\mathcal{C} := \{v_1, \ldots, v_n, w_1, \ldots, w_p, u_1, \ldots, u_m\} \subset V$. Na verdade, $\mathcal{C} \subset W_1 + W_2$. Mostremos que $\mathcal{C}$ é uma base de $W_1 + W_2$. Inicialmente, verifiquemos que é $\mathcal{C}$ é LI. Com efeito, seja $c := \sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^p \beta_i w_i + \sum_{i=1}^m \gamma_i u_i$ uma combinação linear qualquer de elementos de $\mathcal{C}$. Suponhamos que c é o vetor nulo, ou seja, $\sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^p \beta_i w_i + \sum_{i=1}^m \gamma_i u_i = \vec{0}$ (*). Logo,

$$W_1 \ni \sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^p \beta_i w_i = -\sum_{i=1}^m \gamma_i u_i \in W_2,$$

isto é, $-\sum_{i=1}^m \gamma_i u_i \in W_1 \cap W_2$. Logo, como $\mathcal{B}$ gera $W_1 \cap W_2$, existem $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tais que $-\sum_{i=1}^m \gamma_i u_i = \sum_{i=1}^n \lambda_i v_i$. Daí,

$$\sum_{i=1}^n \lambda_i v_i + \sum_{i=1}^m \gamma_i u_i = \vec{0}\ .$$

Como $v_1, \ldots, v_n, u_1, \ldots, u_m \in \mathcal{B}_2$ que é LI, segue que $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$ e $\gamma_1 = \gamma_2 = \cdots = \gamma_m = 0$. Substituindo estes valores em $(*)$, segue que

$$\sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^p \beta_i w_i = \vec{0}\ .$$

Desde que $v_1, \ldots, v_n, w_1, \ldots, w_p \in \mathcal{B}_1$ que é LI, então $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ e $\beta_1 = \beta_2 = \cdots = \beta_p = 0$. Logo, os escalares são todos nulos em $(*)$ e portanto $\mathcal{C}$ é LI. Mostremos agora que $\mathcal{C}$ gera $W_1 + W_2$. Seja $d := \left(\sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^p \beta_i w_i\right) + \left(\sum_{i=1}^n \lambda_i v_i + \sum_{i=1}^m \gamma_i u_i\right) \in W_1 + W_2$ um elemento qualquer de $W_1 + W_2$. Como $d = \sum_{i=1}^n (\alpha_i + \lambda_i) v_i + \sum_{i=1}^p \beta_i w_i + \sum_{i=1}^m \gamma_i u_i$, ou seja, d se escreve como combinação linear de elementos de $\mathcal{C}$, obtemos que $\mathcal{C}$ gera

W_1+W_2. Portanto, $\mathcal{C}$ é base de W_1+W_2 e $\dim_{\mathbb{K}}(W_1+W_2) = n+p+m$. Logo, $\dim_{\mathbb{K}}(W_1+W_2) = (n+p)+(n+m)-n = \dim_{\mathbb{K}} W_1 + \dim_{\mathbb{K}} W_2 - \dim_{\mathbb{K}} W_1 \cap W_2$. Se $W_1 \cap W_2 = \{\overrightarrow{0}\}$, consideremos $\widetilde{\mathcal{B}_1} := \{w_1, \ldots, w_p\} \subset W_1$ uma base de W_1 e $\widetilde{\mathcal{B}_2} := \{u_1, \ldots, u_m\} \subset W_2$ uma base de W_2 e mostremos que $\widetilde{\mathcal{C}} := \widetilde{\mathcal{B}_1} \cup \widetilde{\mathcal{B}_2}$ é uma base de $W_1 + W_2$. É fácil ver que $\widetilde{\mathcal{C}}$ gera $W_1 + W_2$. Verifiquemos então que $\widetilde{\mathcal{C}}$ é LI. De fato, suponhamos $\sum_{i=1}^{p} \beta_i w_i + \sum_{i=1}^{m} \gamma_i u_i = \overrightarrow{0}$. Daí, $W_1 \ni \sum_{i=1}^{p} \beta_i w_i = -\sum_{i=1}^{m} \gamma_i u_i \in W_2$ e portanto $-\sum_{i=1}^{m} \gamma_i u_i \in W_1 \cap W_2 = \{\overrightarrow{0}\}$. Logo, como $\widetilde{\mathcal{B}_1}$ é LI, os escalares γ_i são todos nulos. Mas $\sum_{i=1}^{p} \beta_i w_i = -\sum_{i=1}^{m} \gamma_i u_i = \overrightarrow{0}$. Então, como $\widetilde{\mathcal{B}_2}$ é LI, os escalares β_i são todos nulos. Logo, $\widetilde{\mathcal{C}}$ é LI e portanto é base de $W_1 + W_2$ e $\dim_{\mathbb{K}}(W_1 + W_2) = p + m - 0 = \dim_{\mathbb{K}} W_1 + \dim_{\mathbb{K}} W_2 - \dim_{\mathbb{K}} W_1 \cap W_2$. ■

Definição 0.71. *Seja V um espaço vetorial (sobre algum corpo) e sejam W_1, W_2 subespaços vetoriais de V. Dizemos que a* **soma** $W_1 + W_2$ *é* **direta** *se $W_1 \cap W_2 = \{\overrightarrow{0}\}$. Notação: $W_1 \oplus W_2$.*
Ainda, dizemos que V é a **soma direta de** W_1 *e* W_2 *se $V = W_1 \oplus W_2$.*

Exemplo 0.72. *Temos que $\mathbb{C} = [1] \oplus [i]$, pois $[1] + [i] = \mathbb{C}$ e $[1] \cap [i] = \{\overrightarrow{0}\}$.*

Exemplo 0.73. *Sejam $W_1 := [(1,0,0),(0,1,0)]$ e $W_2 := [(1,0,0),(0,0,1)]$ dois subespaços vetoriais de $\mathbb{R}^3$. Como $W_1 \cap W_2 = \{(\alpha,0,0) : \alpha \in \mathbb{R}\} \neq \{\overrightarrow{0}\}$, então a soma $W_1 + W_2 = \mathbb{R}^3$ não é direta.*

O próximo resultado nos dá uma caracterização de quando um espaço vetorial é uma soma direta de dois subespaços vetoriais seus.

Proposição 0.74. *Seja V um $\mathbb{K}$-espaço vetorial e sejam W_1, W_2 subespaços vetoriais de V. Então $V = W_1 \oplus W_2$ se, e só se, cada vetor $v \in V$ se escreve de forma única como $w_1 + w_2$, com $w_1 \in W_1$ e $w_2 \in W_2$.*

Com efeito, suponhamos inicialmente que $V = W_1 \oplus W_2$. Daí, se $v \in V$, então existem $w_1 \in W_1, w_2 \in W_2$ tais que $v = w_1 + w_2$. Suponhamos então que $v = u_1 + u_2$, para algum $u_1 \in W_1, u_2 \in W_2$. Daí, $w_1 + w_2 = u_1 + u_2$ e portanto $W_1 \ni -u_1 + w_1 = u_2 - w_2 \in W_2$. Logo, como $W_1 \cap W_2 = \{\vec{0}\}$, segue que $-u_1 + w_1 = u_2 - w_2 = \vec{0}$, isto é, a decomposição de v é única. Reciprocamente, suponhamos que todo vetor $v \in V$ se escreve de forma única como $w_1 + w_2$, com $w_1 \in W_1$ e $w_2 \in W_2$. Em particular, segue que $V = W_1 + W_2$. Seja, agora, $w \in W_1 \cap W_2$ qualquer. Então

$$W_1 + W_2 \ni w + \vec{0} = w = \vec{0} + w \in W_1 + W_2.$$

Daí, desde que a decomposição de w é única, então $w = \vec{0}$ e portanto $W_1 \cap W_2 = \{\vec{0}\}$. ■

Exemplo 0.75. *Sejam* $W_1 := [(1,0,0),(0,1,0)]$ *e* $W_2 := [(1,0,0),(0,0,1)]$ *dois subespaços vetoriais de* $\mathbb{R}^3$*. Notemos, por exemplo, que podemos escrever* $(3,0,0)$ *ora como* $(1,0,0)+(2,0,0) \in W_1+W_2$*, ora como* $(3,0,0)+(0,0,0) \in W_1 + W_2$*, isto é, a decomposição de* $(3,0,0)$ *não é única.*

Exemplo 0.76. *Sejam* $W_1 := \left[\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right]$ *e* $W_2 := \left[\left(\begin{smallmatrix}1&0\\0&2\end{smallmatrix}\right), \left(\begin{smallmatrix}0&0\\0&4\end{smallmatrix}\right)\right]$ *dois subespaços vetoriais de* $\mathbb{M}_{2\times 2}(\mathbb{R})$*. Notemos, por exemplo, que podemos escrever* $\left(\begin{smallmatrix}4&0\\0&4\end{smallmatrix}\right)$ *ora como* $(2\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)) + (2\left(\begin{smallmatrix}1&0\\0&2\end{smallmatrix}\right) + (-\frac{1}{2})\left(\begin{smallmatrix}0&0\\0&4\end{smallmatrix}\right)) \in W_1 + W_2$*, ora como* $(0\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)) + (4\left(\begin{smallmatrix}1&0\\0&2\end{smallmatrix}\right) + (-1)\left(\begin{smallmatrix}0&0\\0&4\end{smallmatrix}\right)) \in W_1+W_2$*, isto é, a decomposição de* $\left(\begin{smallmatrix}4&0\\0&4\end{smallmatrix}\right)$ *não é única. Logo, a soma* $W_1 + W_2$ *não é direta.*

O próximo resultado nos diz quando é possível "completar"um subespaço vetorial a fim de obter o espaço vetorial todo.

Proposição 0.77. *Seja* V *um* $\mathbb{K}$*-espaço vetorial finitamente gerado e seja* W_1 *um subespaço vetorial de* V*. Então existe* $W_2 \subset V$ *subespaço vetorial de* V *tal que* $V = W_1 \oplus W_2$*.*

De fato, se $V = \{\vec{0}\}$, então $W_1 = \{\vec{0}\}$ e basta tomar $W_2 = \{\vec{0}\}$. Seja então $V \neq \{\vec{0}\}$. Notemos que se $W_1 = \{\vec{0}\}$ (ou se $W_1 = V$), basta tomar $W_2 = V$ (ou $W_2 = \{\vec{0}\}$). Como V é finitamente gerado, então W_1 também o é. Logo, existe $\mathcal{B}_1 \subset W_1$ uma base de W_1, digamos $\mathcal{B}_1 := \{v_1, \ldots, v_n\}$. Como $\mathcal{B}_1$ é um conjunto LI contido em V, então existe $\mathcal{C} \subset V$ uma base de V que contém $\mathcal{B}_1$, digamos $\mathcal{C} := \{v_1, \ldots, v_n, u_1, \ldots, u_m\}$. Escrevamos $\mathcal{B}_2 := \{u_1, \ldots, u_m\}$ e consideremos $W_2 := [\mathcal{B}_2]$. Desde que cada vetor $v \in V$ se escreve de forma única como combinação linear de elementos de $\mathcal{C}$, $v = \sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^m \beta_i u_i$, pois $\mathcal{C}$ é base, e em particular $\sum_{i=1}^n \alpha_i v_i \in W_1$ e $\sum_{i=1}^m \beta_i u_i \in W_2$, então a decomposição de v em $W_1 + W_2$ é única e, pela proposição anterior, $V = W_1 \oplus W_2$. ■

Exemplo 0.78. *Consideremos novamente* $W_1 := \left[\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\right] \subset \mathbb{M}_{2\times 2}(\mathbb{R})$. *Seja* $\mathcal{C} := \left\{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\right\} \subset \mathbb{M}_{2\times 2}(\mathbb{R})$. *Desde que* $\left\{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\right\} \cup \mathcal{C}$ *é uma base de* $\mathbb{M}_{2\times 2}(\mathbb{R})$ *(verifique isto), temos que* $\mathbb{M}_{2\times 2}(\mathbb{R}) = W_1 \oplus [\mathcal{C}]$.

Transformação linear

Definição 0.79. *Sejam* U, V *espaços vetoriais sobre* $\mathbb{K}$. *Dizemos que uma aplicação* $T : U \to V$, $U \ni u \mapsto T(u) \in V$, *é uma* **transformação linear** *se* $T(u + w) = T(u) + T(w)$ *e* $T(\alpha u) = \alpha T(u)$, *para todo* $\alpha \in \mathbb{K}$, $u, w \in U$.

Exemplo 0.80. *Consideremos* $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T((x, y, z)) = (z, 0)$, *para todo* $x, y, z \in \mathbb{R}$. *Então* T *é linear. De fato,* $T((x_1, y_1, z_1) + (x_2, y_2, z_2)) = T((x_1 + x_2, y_1 + y_2, z_1 + z_2)) = (z_1 + z_2, 0) = (z_1, 0) + (z_2, 0) = T((x_1, y_1, z_1)) + T((x_2, y_2, z_2))$ *e* $T(\alpha(x_1, y_1, z_1)) = T((\alpha x_1, \alpha y_1, \alpha z_1)) = (\alpha z_1, 0) = \alpha(z_1, 0) = \alpha T((x_1, y_1, z_1))$, *para todo* $\alpha \in \mathbb{R}$, $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$. *Portanto,* T *é linear.*

Exemplo 0.81. *Consideremos* $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T((x, y, z)) = (z, 1)$, *para todo* $x, y, z \in \mathbb{R}$. *Então* T *não é linear. Com efeito,* $T((\sqrt{2}, \sqrt{3}, 2) + (0, 1, 3)) = T((\sqrt{2}, \sqrt{3} + 1, 5)) = (5, 1) \neq (5, 2) = (2, 1) + (3, 1) = T((\sqrt{2}, \sqrt{3}, 2)) + T((0, 1, 3))$. *Logo,* T *não é linear.*

Observação 0.82. *Muitas vezes escreveremos apenas $T(x_1, x_2, \dots, x_n)$ para denotar $T((x_1, x_2, \dots, x_n))$, isto é, omitiremos o excesso de parênteses.*

Observação 0.83. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$ e $T : U \to V$ uma* aplicação *tal que $T(\alpha u + w) = \alpha T(u) + T(w)$, para todo $\alpha \in \mathbb{K}$, $u, w \in U$. Em particular, tomando $\alpha = -1$ e $w = u$, temos que $T((-1)u + u) = (-1)T(u) + T(u) = ((-1) + 1)T(u) = 0T(u) = \vec{0}$. Logo, $T(\vec{0}) = \vec{0}$.*

Proposição 0.84. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$ e $T : U \to V$ uma aplicação. Então T é linear se, e só se, $T(\alpha u + w) = \alpha T(u) + T(w)$, para todo $\alpha \in \mathbb{K}$, $u, w \in U$.*

De fato, suponhamos inicialmente que T é linear. Daí, usando as duas propriedades de T ser linear, obtemos que $T(\alpha u + w) = T(\alpha u) + T(w) = \alpha T(u) + T(w)$, para todo $\alpha \in \mathbb{K}$, $u, w \in U$.
Reciprocamente, suponhamos agora que $T(\alpha u + w) = \alpha T(u) + T(w)$, para todo $\alpha \in \mathbb{K}$, $u, w \in U$. Mostremos que T é linear. Sejam $\alpha \in \mathbb{K}$, $u, w \in U$ quaisquer. Como $T(u + w) = T(1.u + w) = 1.T(u) + T(w) = T(u) + T(w)$, obtemos a primeira propriedade para T ser linear. Desde que $T(\alpha u) = T(\alpha u + \vec{0}) = \alpha T(u) + T(\vec{0}) = \alpha T(u) + \vec{0} = \alpha T(u)$, obtemos a segunda propriedade para T ser linear. ■

Proposição 0.85. *Seja $T : U \to V$ uma transformação linear. Então, para todo $\alpha_1, \dots, \alpha_n \in \mathbb{K}$, $u, u_1, \dots, u_n \in U$, T satisfaz:*

1. $T(\vec{0}) = \vec{0}$
2. $T(-u) = -T(u)$
3. $T(\sum_{i=1}^{n} \alpha_i u_i) = \sum_{i=1}^{n} \alpha_i T(u_i)$

Exercício (use a caracterização de transformação linear dada pela proposição acima). ■

Exemplo 0.86. *Consideremos* $T : \mathcal{P}_1(\mathbb{R}) \to \mathbb{R}^2$, $T(a_1x+a_0) = (a_0, a_1)$, *para todo* $a_1x + a_0 \in \mathcal{P}_1(\mathbb{R})$. *Então* T *é linear. Verifique.*

Exemplo 0.87. *Consideremos* $T : \mathbb{M}_{2\times 2}(\mathbb{R}) \to \mathbb{R}^3$, $T\left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right) = (a_{11}, a_{12}, 0)$, *para todo* $\left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right) \in \mathbb{M}_{2\times 2}(\mathbb{R})$. *Então* T *é linear. Verifique.*

Exemplo 0.88. *Consideremos* $T : \mathbb{R} \to \mathcal{P}_2(\mathbb{R})$, $T(a) = ax^2 + a$, *para todo* $a \in \mathbb{R}$. *Então* T *é linear. De fato,* $T(\alpha a + b) = (\alpha a + b)x^2 + (\alpha a + b)$ *e* $\alpha T(a) + T(b) = \alpha(ax^2 + a) + (bx^2 + b) = (\alpha a + b)x^2 + (\alpha a + b)$. *Portanto,* $T(\alpha a + b) = \alpha T(a) + T(b)$, *para todo* $\alpha, a, b \in \mathbb{R}$ *e* T *é linear.*

Exemplo 0.89. *Consideremos* $T : \mathbb{R}^2 \to \mathbb{M}_{2\times 3}(\mathbb{R})$, $T(a, b) = \left(\begin{smallmatrix} a & b & a \\ 0 & a & 0 \end{smallmatrix}\right)$, *para todo* $(a, b) \in \mathbb{R}^2$. *Então* T *é linear. De fato,*

$$T(\alpha(a, b) + (a', b')) = T(\alpha a + a', \alpha b + b') = \begin{pmatrix} (\alpha a+a') & (\alpha b+b') & (\alpha a+a') \\ 0 & (\alpha a+a') & 0 \end{pmatrix} \text{ e}$$

$$\alpha T(a, b) + T(a', b') = \alpha \left(\begin{smallmatrix} a & b & a \\ 0 & a & 0 \end{smallmatrix}\right) + \left(\begin{smallmatrix} a' & b' & a' \\ 0 & a' & 0 \end{smallmatrix}\right) = \begin{pmatrix} (\alpha a+a') & (\alpha b+b') & (\alpha a+a') \\ 0 & (\alpha a+a') & 0 \end{pmatrix}.$$

Logo, $T(\alpha(a, b) + (a', b')) = \alpha T(a, b) + T(a', b')$, *para todo* $\alpha, a, b, a', b' \in \mathbb{R}$ *e* T *é linear.*

Exemplo 0.90. *Seja* $T : U \to V$ *uma aplicação tal que* $T(u) = \vec{0}$, *para todo* $u \in U$ *(com* U, V $\mathbb{K}$*-espaços vetoriais). Então* T *é linear pois, para todo* $\alpha \in \mathbb{K}, u, w \in U$, $T(\alpha u + w) = \vec{0} = \vec{0} + \vec{0} = \alpha\vec{0} + \vec{0} = \alpha T(u) + T(w)$. *Esta* T *é chamada de transformação identicamente nula.*

Exemplo 0.91. *Seja* $T : U \to U$ *uma aplicação tal que* $T(u) = u$, *para todo* $u \in U$ *(com* U *um* $\mathbb{K}$*-espaço vetorial). Então* T *é linear pois, para todo* $\alpha \in \mathbb{K}, u, w \in U$, $T(\alpha u + w) = \alpha u + w = \alpha T(u) + T(w)$. *Esta* T *é chamada de identidade (de* U*) e é denotada por* Id_U *(ou simplesmente por* Id*).*

Exemplo 0.92. *Seja* $\beta \in \mathbb{K}$ *qualquer. Seja* $T : U \to U$ *uma aplicação tal que* $T(u) = \beta u$, *para todo* $u \in U$ *(com* U *um* $\mathbb{K}$*-espaço vetorial). Então* T *é linear pois, para todo* $\alpha \in \mathbb{K}, u, w \in U$, $T(\alpha u + w) = \beta(\alpha u + w) = \alpha(\beta u) + \beta w =$

$\alpha T(u) + T(w)$. Esta T é chamada de transformação multiplicação por β e é denotada por T_β.

Exemplo 0.93. *Seja $T : \mathbb{K} \to \mathbb{K}$ uma* transformação linear qualquer. *Então existe $\beta \in \mathbb{K}$ tal que $T = T_\beta$. De fato, se $\alpha \in \mathbb{K}$, temos que $T(\alpha) = T(1.\alpha) = T(\alpha.1) = \alpha.T(1) = T(1).\alpha$. Logo, tomando $\beta := T(1) \in \mathbb{K}$, temos que $T(\alpha) = \beta\alpha$, para todo $\alpha \in \mathbb{K}$, ou seja, $T = T_\beta$, com $\beta = T(1) \in \mathbb{K}$.*

Exemplo 0.94. *Seja $T : \mathbb{K}^n \to \mathbb{K}$ uma* transformação linear qualquer. *Então existem $\beta_1, \ldots, \beta_n \in \mathbb{K}$ tais que $T(\alpha_1, \ldots, \alpha_n) = \sum_{i=1}^n \beta_i\alpha_i$. Com efeito, se $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$, então, como $(\alpha_1, \ldots, \alpha_n) = \alpha_1(1, 0, \ldots, 0) + \cdots + \alpha_n(0, \ldots, 0, 1)$, segue que $T(\alpha_1, \ldots, \alpha_n) = T(\alpha_1(1, 0, \ldots, 0) + \cdots + \alpha_n(0, \ldots, 0, 1)) = \alpha_1 T(1, 0, \ldots, 0) + \cdots + \alpha_n T(0, \ldots, 0, 1)$. Daí, tomando $\beta_1 := T(1, 0, \ldots, 0), \ldots, \beta_n := T(0, \ldots, 0, 1) \in \mathbb{K}$, obtemos que $T(\alpha_1, \ldots, \alpha_n) = \sum_{i=1}^n \beta_i\alpha_i$, para todo $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$.*

Teorema 0.95. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$. Seja $\mathcal{B} := \{u_1, \ldots, u_n\}$ uma base de U. Se $\mathcal{C} := \{v_1, \ldots, v_n\}$ é um subconjunto qualquer de V, então existe uma única transformação linear $T : U \to V$ tal que $T(u_i) = v_i$, para todo $i = 1, \ldots, n$.*

De fato, seja $u \in U$. Então, como $\mathcal{B}$ é base de U, existem $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que $u = \sum_{i=1}^n \alpha_i u_i$. Definamos $T(u) = \sum_{i=1}^n \alpha_i v_i$, para todo $u = \sum_{i=1}^n \alpha_i u_i \in U$. Primeiramente, esta é uma boa definição para T, pois se $\sum_{i=1}^n \alpha_i u_i = \sum_{i=1}^n \beta_i u_i$, temos do fato de que cada vetor de U se escreve de maneira única como combinação linear de elementos de $\mathcal{B}$, que $\alpha_i = \beta_i$, $i = 1, \ldots, n$ e portanto que $T(\sum_{i=1}^n \alpha_i u_i) = \sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \beta_i v_i = T(\sum_{i=1}^n \beta_i u_i)$. Ainda, desta definição, temos que $T(u_i) = T(1.u_i) = 1.v_i = v_i$, para todo $i = 1, \ldots, n$. Agora, verifiquemos que T é linear. Com efeito, $T(\beta(\sum_{i=1}^n \alpha_i u_i) + \sum_{i=1}^n \gamma_i u_i) = T(\sum_{i=1}^n (\beta\alpha_i + \gamma_i)u_i) = \sum_{i=1}^n (\beta\alpha_i + \gamma_i)v_i = \beta T(\sum_{i=1}^n \alpha_i u_i) + T(\sum_{i=1}^n \gamma_i u_i)$, para todo $\beta, \alpha_1, \ldots, \alpha_n, \gamma_1, \ldots, \gamma_n \in \mathbb{K}$. Finalmente, vejamos que T é a única transformação linear tal que $T(u_i) = v_i$.

Para isto, suponhamos que existe $S : U \to V$ uma transformação linear tal que $S(u_i) = v_i$, para todo $i = 1, \ldots, n$. Então $S(u) = S(\sum_{i=1}^{n} \alpha_i u_i) \overset{S \text{ linear}}{=} \sum_{i=1}^{n} \alpha_i S(u_i) = \sum_{i=1}^{n} \alpha_i v_i \overset{\text{def. } T}{=} T(u)$, para todo $u \in U$, ou seja $S = T$. $\blacksquare$

Definição 0.96. *Seja $T : U \to V$ uma transformação linear.*

1. *Chamamos o subconjunto de U, $\{u \in U | T(u) = \vec{0}\}$, dos vetores que são "anulados pela T"simplesmente de* **núcleo de** T *(ou kernel de T) e denotamos-no por* $\mathrm{Nuc}T$ *(ou* $\mathrm{Ker}T$*).*

2. *Chamamos o subconjunto de V, $\{v \in V | v = T(u)$, para algum $u \in U\}$, dos vetores que são "T de alguém"simplesmente de* **imagem de** T *e denotamos-no por* $\mathrm{Im}T$*.*

Observação 0.97. *Com as mesmas notações,* $\mathrm{Im}T = \{v \in V | v = T(u)$, *para algum* $u \in U\} = \{T(u) | u \in U\} =: T(U)$*.*

Relembremos algumas definições.

Observação 0.98. *Uma aplicação $f : X \to Y$ é* **injetora** *se $f(a) = f(b)$ implicar que $a = b$, e f é* **sobrejetora** *se para cada $y \in Y$ existe $x \in X$ tal que $y = f(x)$. Ainda, f é* **bijetora** *se for injetora e sobrejetora, e neste caso temos que a* **inversa de** f *existe e a denotamos por $f^{-1} : Y \to X$.*
Se $f : X \to Y$ e $g : Y \to Z$ são aplicações, denotamos por $g \circ f : X \to Z$ a aplicação tal que $(g \circ f)(x) = g(f(x))$, com $x \in X$. Se $g \circ f$ é injetora (respectivamente, sobrejetora), então f é injetora (respectivamente, g é sobrejetora). De fato, se $f(a) = f(b)$, então $g(f(a)) = g(f(b))$ e portanto, se $g \circ f$ é injetora, segue que $a = b$, isto é, f é injetora. Se $z \in Z$, então, se $g \circ f$ é sobrejetora, existe $x \in X$ tal que $(g \circ f)(x) = z$. Mas, daí, $f(x) \in Y$ é tal que $g(f(x)) = z$, ou seja, g é sobrejetora.
Se $f : X \to Y$ é bijetora, então $f \circ f^{-1} = Id_Y$ e $f^{-1} \circ f = Id_X$.

Proposição 0.99. *Sejam U, V, W espaços vetoriais sobre $\mathbb{K}$ e sejam $T : U \to V$ e $S : V \to W$ transformações lineares. Então $S \circ T : U \to W$ é uma transformação linear.*

De fato, sejam $u_1, u_2 \in U$ e $\alpha \in \mathbb{K}$ quaisquer. Como $(S \circ T)(\alpha u_1 + u_2) = S(T(\alpha u_1 + u_2)) \overset{T\,\text{linear}}{=} S(\alpha T(u_1) + T(u_2)) \overset{S\,\text{linear}}{=} \alpha S(T(u_1)) + S(T(u_2)) = \alpha(S \circ T)(u_1) + (S \circ T)(u_2)$, então $S \circ T$ é linear. ■

Vejamos agora caracterizações de T ser injetora e de T ser sobrejetora.

Proposição 0.100. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$ e $T : U \to V$ uma transformação linear. Então:*

1. $\mathrm{Ker}T$ *é um subespaço de U e* $\mathrm{Im}T$ *é um subespaço de V*
2. *T é sobrejetora se, e somente se,* $\mathrm{Im}T = V$
3. *T é injetora se, e somente se,* $\mathrm{Ker}T = \{\vec{0}\}$

Mostremos (1). Com efeito, sejam $\alpha \in \mathbb{K}$, $u, w \in \mathrm{Ker}T$. Em particular, temos que $T(u) = \vec{0}$ e $T(w) = \vec{0}$. Desde que $T(\alpha u + w) \overset{T\,\text{linear}}{=} \alpha T(u) + T(w) = \alpha.\vec{0} + \vec{0} = \vec{0}$, então $\alpha u + w \in \mathrm{Ker}T$, e verificamos que tanto a soma de vetores como a multiplicação por escalar pertencem a $\mathrm{Ker}T$. Daí, como $T(\vec{0}) = \vec{0}$, segue que $\vec{0} \in \mathrm{Ker}T$, e $\mathrm{Ker}T$ é subespaço. Sejam, agora, $\beta \in \mathbb{K}$, $v_1, v_2 \in \mathrm{Im}T$. Em particular, existem $u_1, u_2 \in U$ tais que $v_1 = T(u_1)$ e $v_2 = T(u_2)$. Uma vez que $\beta v_1 + v_2 = \beta T(u_1) + T(u_2) \overset{T\,\text{linear}}{=} T(\beta u_1 + u_2)$, então $\beta v_1 + v_2 \in \mathrm{Im}T$, e vemos que tanto a soma de vetores como a multiplicação por escalar pertencem a $\mathrm{Im}T$. Daí, como $\vec{0} = T(\vec{0})$, segue que $\vec{0} \in \mathrm{Im}T$, e $\mathrm{Im}T$ é subespaço.

Mostremos, agora, (2). Seja T sobrejetora. Já temos que $\mathrm{Im}T \subset V$, donde basta vermos que $V \subset \mathrm{Im}T$. De fato, seja $v \in V$. Como T é sobrejetora, existe $u \in U$ tal que $T(u) = v$ e portanto $v \in \mathrm{Im}T$. Suponhamos agora que $\mathrm{Im}T = V$. Daí, dado $v \in V$, então $v \in \mathrm{Im}T$ e existe $u \in U$ tal que $v = T(u)$. Logo, T é sobrejetora.

Finalmente, mostremos (3). Seja T injetora. Se $u \in \mathrm{Ker}T$, então $T(u) = \vec{0}$. Mas $T(\vec{0}) = \vec{0}$. Daí, como T é injetora e $T(u) = T(\vec{0})$, segue que $u = \vec{0}$ e $\mathrm{Ker}T = \{\vec{0}\}$. Suponhamos agora que $\mathrm{Ker}T = \{\vec{0}\}$. Se $u_1, u_2 \in U$ são tais que $T(u_1) = T(u_2)$, então, como T é linear, $T(u_1 - u_2) = \vec{0}$ e portanto $u_1 - u_2 \in \mathrm{Ker}T = \{\vec{0}\}$. Logo, $u_1 - u_2 = \vec{0}$, isto é, $u_1 = u_2$ e T é injetora. ■

Observação 0.101. *Notemos que o item 2 da proposição acima vale independentemente de T ser linear.*
Ainda, pela proposição acima, T é bijetora se, e somente se, $\mathrm{Im}T = V$ *e* $\mathrm{Ker}T = \{\vec{0}\}$.

Definição 0.102. *Seja $T : U \to V$ uma transformação linear. Chamamos a dimensão da imagem de T, $\dim_{\mathbb{K}} \mathrm{Im}T$, de* **posto de** *$T$ e chamamos a dimensão do núcleo de T, $\dim_{\mathbb{K}} \mathrm{Ker}T$, de* **nulidade de** *$T$.*

Exemplo 0.103. *Seja $T : \mathbb{R}^3 \to \mathbb{M}_{2\times 2}(\mathbb{R})$ uma aplicação tal que $T(a,b,c) = \left(\begin{smallmatrix} c & 0 \\ 0 & a-b \end{smallmatrix}\right)$, para todo $(a,b,c) \in \mathbb{R}^3$. Então T é linear, a nulidade de T é 1 e o posto de T é 2. Ainda, T não é injetora, nem sobrejetora. De fato,*

$T(\alpha(a,b,c)+(a',b',c')) = T(\alpha a+a', \alpha b+b', \alpha c+c') = \begin{pmatrix} (\alpha c+c') & 0 \\ 0 & (\alpha a+a')-(\alpha b+b') \end{pmatrix}$ *e*

$$\alpha T(a,b,c) + T(a',b',c') = \alpha \left(\begin{smallmatrix} c & 0 \\ 0 & a-b \end{smallmatrix}\right) + \left(\begin{smallmatrix} c' & 0 \\ 0 & a'-b' \end{smallmatrix}\right) = \begin{pmatrix} (\alpha c+c') & 0 \\ 0 & (\alpha a+a')-(\alpha b+b') \end{pmatrix}.$$

Logo, $T(\alpha(a,b,c) + (a',b',c')) = \alpha T(a,b,c) + T(a',b',c')$, para todo $\alpha, a, b, c, a', b', c' \in \mathbb{R}$ e portanto T é linear.
Agora, se $T(a,b,c) = \vec{0}$, então $\left(\begin{smallmatrix} c & 0 \\ 0 & a-b \end{smallmatrix}\right) = \vec{0}$ e portanto $c = 0$ e $a = b$. Logo, $\mathrm{Ker}T = \{(a,b,c) \in \mathbb{R}^3 | T(a,b,c) = \vec{0}\} = \{(a,a,0) | a \in \mathbb{R}\} = [(1,1,0)]$. *Como $\{(1,1,0)\}$ é LI, segue que é uma base de* $\mathrm{Ker}T$. *Logo, a nulidade de T é 1.*
Agora, notemos que $T(a,b,c) = \left(\begin{smallmatrix} c & 0 \\ 0 & a-b \end{smallmatrix}\right) = \left(\begin{smallmatrix} c & 0 \\ 0 & 0 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 0 & 0 \\ 0 & a-b \end{smallmatrix}\right) = c\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) + (a-b)\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.
Daí, $\mathrm{Im}T = \{T(a,b,c) | (a,b,c) \in \mathbb{R}^3\} = \{c\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) + (a-b)\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right) | a,b,c \in \mathbb{R}\} = [\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)]$. *Desde que $\{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)\}$ é LI (convença-se disto), então é uma base de* $\mathrm{Im}T$. *Assim o posto de T é 2. Notemos, ainda, que T não é injetora, pois* $\mathrm{Ker}T \neq \{\vec{0}\}$, *nem sobrejetora, pois* $\mathrm{Im}T \neq \mathbb{M}_{2\times 2}(\mathbb{K})$.

Proposição 0.104. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$, com U sendo finitamente gerado. Seja $T : U \to V$ uma transformação linear. Se $\{u_1, \ldots, u_n\}$ é uma base de U, então $\{T(u_1), \ldots, T(u_n)\}$ gera* $\mathrm{Im}T$.

De fato, seja $\mathcal{B} := \{u_1, \ldots, u_n\}$ uma base de U. Queremos mostrar que $\mathcal{G} := \{T(u_1), \ldots, T(u_n)\}$ é um conjunto gerador de $\mathrm{Im}T$. Consideremos então $v \in \mathrm{Im}T$ qualquer. Daí, existe $u \in U$ tal que $v = T(u)$. Como $u \in U$ e $\mathcal{B}$ é base de U, então existem $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que $u = \alpha_1 u_1 + \cdots + \alpha_n u_n$. Logo, $T(u) = T(\alpha_1 u_1 + \cdots + \alpha_n u_n) \overset{T \text{ linear}}{=} \alpha_1 T(u_1) + \cdots + \alpha_n T(u_n)$. Daí, desde que v é uma combinação linear de vetores de $\mathcal{G}$, então $\mathcal{G}$ gera $\mathrm{Im}T$. ■

Observação 0.105. *Notemos que, na demonstração acima, só usamos o fato de que $\mathcal{B}$ é um conjunto gerador de U, donde o resultado é mais geral do que o enunciado.*

Vejamos agora o "teorema do núcleo e da imagem".

Teorema 0.106. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$, com U sendo finitamente gerado. Seja $T : U \to V$ uma transformação linear. Então* $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} \mathrm{Ker}T + \dim_{\mathbb{K}} \mathrm{Im}T$.

Com efeito, como $\mathrm{Ker}T$ é um subespaço vetorial de U e U é finitamente gerado, segue que $\mathrm{Ker}T$ também o é. Daí, existe $\mathcal{B} := \{u_1, \ldots, u_m\} \subset \mathrm{Ker}T$ uma base de $\mathrm{Ker}T$. Como $\mathcal{B}$ é um conjunto LI contido em U, então existe uma base, $\mathcal{C}$, de U contendo $\mathcal{B}$, digamos $\mathcal{C} := \{u_1, \ldots, u_m, w_1, \ldots, w_p\}$. Daí, se mostrarmos que $\mathcal{D} := \{T(w_1), \ldots, T(w_p)\} \subset \mathrm{Im}T$ é uma base de $\mathrm{Im}T$, então $\dim_{\mathbb{K}} U = m + p = \dim_{\mathbb{K}} \mathrm{Ker}T + \dim_{\mathbb{K}} \mathrm{Im}T$. Façamos isto. Pelo resultado anterior, como $\mathcal{C}$ é uma base de U, então $\{T(u_1), \ldots, T(u_m), T(w_1), \ldots, T(w_p)\}$ gera $\mathrm{Im}T$. Desde que $T(u_1) = T(u_2) = \cdots = T(u_m) = \vec{0}$, pois $\mathcal{B} \subset \mathrm{Ker}T$, então $\mathcal{D}$ já gera $\mathrm{Im}T$. Mostremos que é LI. Seja $c := \alpha_1 T(w_1) + \alpha_2 T(w_2) + \cdots + \alpha_p T(w_p)$ uma combinação linear qualquer de vetores de $\mathcal{D}$. Suponhamos que c é o vetor nulo. Daí, $\vec{0} = \alpha_1 T(w_1) + \alpha_2 T(w_2) + \cdots + \alpha_p T(w_p) \overset{T \text{ linear}}{=}$

$T(\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_p w_p)$, isto é, $\sum_{i=1}^{p} \alpha_i w_i \in \mathrm{Ker}T$. Como $\mathcal{B}$ é uma base de $\mathrm{Ker}T$, segue que existem escalares $\beta_1, \ldots, \beta_m$ tais que $\sum_{i=1}^{p} \alpha_i w_i = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_m u_m$. Já que $\sum_{i=1}^{p} \alpha_i w_i - \sum_{i=1}^{m} \beta_i u_i = \vec{0}$ é uma combinação linear de vetores de $\mathcal{C}$ resultando no vetor nulo, então $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$, $\beta_1 = \beta_2 = \cdots = \beta_m = 0$ e concluímos que $\mathcal{D}$ é LI e portanto que é uma base de $\mathrm{Im}T$. ■

Isomorfismo de espaços vetoriais

Definição 0.107. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$.*

1. *Seja $T : U \to V$ uma transformação linear. Dizemos que T é um isomorfismo (de espaços vetoriais) se T for bijetora.*

2. *Dizemos que U e V são isomorfos, se existe um isomorfismo (de espaços vetoriais) entre eles. Notação: $U \cong V$.*

Exemplo 0.108. *Seja U um espaço vetorial sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U > 0$. Então a identidade de U, $Id : U \to U$ dada por $Id(u) = u$, para todo $u \in U$, é um isomorfismo, mas a transformação identicamente nula, $T : U \to U$ dada por $T(u) = \vec{0}$, para todo $u \in U$, não o é. Aqui vemos que $U \cong U$, mas que nem toda transformação linear de U em U é um isomorfismo.*

Observação 0.109. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$, com U sendo finitamente gerado. Se $U \cong V$, então $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} V$. De fato, como $U \cong V$, segue que existe um isomorfismo $T : U \to V$. Então, se $\mathcal{B} := \{u_1, \ldots, u_m\}$ é uma base de U, temos que $\mathcal{C} := \{T(u_1), \ldots, T(u_m)\}$ gera $\mathrm{Im}T = V$, pois T é sobrejetora. Daí, se mostrarmos que $\mathcal{C}$ é também LI, então $\dim_{\mathbb{K}} U = m = \dim_{\mathbb{K}} V$. Façamos isto. Seja $c := \alpha_1 T(u_1) + \alpha_2 T(u_2) + \cdots + \alpha_m T(u_m)$ uma combinação linear qualquer de elementos de $\mathcal{C}$. Suponhamos que c é o vetor nulo. Daí, $\vec{0} = \alpha_1 T(u_1) + \alpha_2 T(u_2) + \cdots + \alpha_m T(u_m) \overset{T \text{ linear}}{=} T(\alpha_1 u_1 + \alpha_2 u_2 +$*

$\cdots + \alpha_m u_m)$ *e portanto* $\sum_{i=1}^{m} \alpha_i u_i \in \mathrm{Ker}T = \{\vec{0}\}$*, pois* T *é injetora. Assim, como* $\sum_{i=1}^{m} \alpha_i u_i = \vec{0}$ *e* $\mathcal{B}$ *é LI, então* $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$ *e portanto* $\mathcal{C}$ *é LI.*

Observação 0.110. *Assim, se dois espaços são isomorfos e conhecemos a dimensão de um deles, então já sabemos que a dimensão do outro é a mesma coisa.*

Veremos mais adiante a recíproca deste resultado, isto é, que se dois espaços têm a mesma dimensão, então eles são isomorfos.

Proposição 0.111. *Seja* $T : U \to V$ *um isomorfismo. Então a inversa de* T*,* $T^{-1} : V \to U$*, é uma transformação linear.*

De fato, lembremos que $T \circ T^{-1} = Id_V$ e $T^{-1} \circ T = Id_U$. Em particular, ambas as compostas são transformações lineares. Sejam $v_1, v_2 \in V$, $\alpha \in \mathbb{K}$ quaisquer e mostremos que $T^{-1}(\alpha v_1 + v_2) = \alpha T^{-1}(v_1) + T^{-1}(v_2)$. Com efeito, como $V = \mathrm{Im}T$, então existem $u_1, u_2 \in U$ tais que $v_1 = T(u_1)$ e $v_2 = T(u_2)$. Então $T^{-1}(\alpha v_1 + v_2) = T^{-1}(\alpha T(u_1) + T(u_2)) \overset{T \text{ linear}}{=} T^{-1}(T(\alpha u_1 + u_2)) = (T^{-1} \circ T)(\alpha u_1 + u_2) \overset{T^{-1} \circ T \text{ linear}}{=} \alpha (T^{-1} \circ T)(u_1) + (T^{-1} \circ T)(u_2) = \alpha T^{-1}(T(u_1)) + T^{-1}(T(u_2)) = \alpha T^{-1}(v_1) + T^{-1}(v_2)$ e portanto T^{-1} é linear. ■

Proposição 0.112. *Sejam* U, V *espaços vetoriais sobre* $\mathbb{K}$*, com* $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} V = n$*. Seja* $T : U \to V$ *uma transformação linear. Então as condições abaixo são equivalentes:*

1. *T é um isomorfismo*
2. *T é sobrejetora*
3. *T é injetora*

Façamos $(1 \Leftrightarrow 2)$. De fato, se T é um isomorfismo, é claro que T é sobrejetora. Agora, se T é sobrejetora, então $\mathrm{Im}T = V$. Daí, pelo teorema do núcleo e da

imagem, $n = \dim_{\mathbb{K}} U = \dim_{\mathbb{K}} \mathrm{Ker}T + \dim_{\mathbb{K}} \mathrm{Im}T = \dim_{\mathbb{K}} \mathrm{Ker}T + \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \mathrm{Ker}T + n$, donde $\dim_{\mathbb{K}} \mathrm{Ker}T = 0$ e portanto $\mathrm{Ker}T = \{\vec{0}\}$, isto é, T é injetora. Logo, como T é linear e bijetora, segue que é um isomorfismo.
Façamos agora $(1 \Leftrightarrow 3)$. Com efeito, se T é um isomorfismo, é direto que T é injetora. Agora, se T é injetora, então $\mathrm{Ker}T = \{\vec{0}\}$. Assim, pelo teorema do núcleo e da imagem, $n = \dim_{\mathbb{K}} U = \dim_{\mathbb{K}} \mathrm{Ker}T + \dim_{\mathbb{K}} \mathrm{Im}T = 0 + \dim_{\mathbb{K}} \mathrm{Im}T$, donde $\dim_{\mathbb{K}} \mathrm{Im}T = n$ e portanto $\mathrm{Im}T = V$, ou seja, T é sobrejetora. Logo, desde que T é linear e bijetora, então é um isomorfismo. ■

Exemplo 0.113. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ uma aplicação tal que $T(x, y) = (y, x, y - x)$, para todo $(x, y) \in \mathbb{R}^2$. Então T é linear, injetora e não é sobrejetora. Com efeito,*

$$T(\alpha(x, y) + (\tilde{x}, \tilde{y})) = T(\alpha x + \tilde{x}, \alpha y + \tilde{y}) = (\alpha y + \tilde{y}, \alpha x + \tilde{x}, (\alpha y + \tilde{y}) - (\alpha x + \tilde{x}))\ e$$
$$\alpha T(x, y) + T(\tilde{x}, \tilde{y}) = \alpha(y, x, y - x) + (\tilde{y}, \tilde{x}, \tilde{y} - \tilde{x}) = (\alpha y + \tilde{y}, \alpha x + \tilde{x}, (\alpha y + \tilde{y}) - (\alpha x + \tilde{x})).$$

Então $T(\alpha(x, y) + (\tilde{x}, \tilde{y})) = \alpha T(x, y) + T(\tilde{x}, \tilde{y})$, para todo $\alpha, x, y, \tilde{x}, \tilde{y} \in \mathbb{R}$ e T é linear.
A fim de verificarmos se T é, ou não, injetora, analisemos o núcleo de T. Se $T(x, y) = (0, 0, 0)$, então $(y, x, y - x) = (0, 0, 0)$ e portanto $y = 0 = x$, ou seja, se $(x, y) \in \mathrm{Ker}T$, então $(x, y) = (0, 0)$. Logo, $\mathrm{Ker}T = \{(0, 0)\}$ e T é injetora. Um segundo jeito seria tomarmos $(x, y), (\tilde{x}, \tilde{y}) \in \mathbb{R}^2$ tais que $T(x, y) = T(\tilde{x}, \tilde{y})$, fazermos as contas e concluirmos que $x = \tilde{x}$ e $y = \tilde{y}$.
Com o intuito de vermos se T é, ou não, sobrejetora, analisemos a imagem de T. Como $(y, x, y - x) = y(1, 0, 1) + x(0, 1, -1)$, então $\mathrm{Im}T = [(1, 0, 1), (0, 1, -1)]$. Daí, $\dim_{\mathbb{R}} \mathrm{Im}T < 3 = \dim_{\mathbb{R}} \mathbb{R}^3$ e $\mathrm{Im}T$ não pode ser todo o $\mathbb{R}^3$, ou seja, T não é sobrejetora. Uma segunda maneira seria exibirmos uma tripla $(x_0, y_0, z_0) \neq (y, x, y - x)$, para todo $x, y \in \mathbb{R}$. Por exemplo, tomando $z_0 := 0$ e $x_0 := 1 \neq 2 =: y_0$, temos que não existem $x, y \in \mathbb{R}$ tais que $(1, 2, 0) = (y, x, y - x)$.

Exemplo 0.114. *Seja* $T : \mathcal{P}_3(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ *uma aplicação tal que* $T(p(x)) = p'(x)$, *onde se* $p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$, *então* $p'(x) = 3a_3x^2 + 2a_2x + a_1$. *Então* T *é linear (verifique), sobrejetora e não é injetora.*
De fato, para mostrarmos que T *é sobrejetora, basta tomarmos* $q(x) \in \mathcal{P}_2(\mathbb{R})$ *qualquer e exibirmos* $p(x) \in \mathcal{P}_3(\mathbb{R})$ *tal que* $q(x) = p'(x)$. *Dado* $q(x) = b_2x^2 + b_1x + b_0$, *queremos escrever* $q(x) = 3a_3x^2 + 2a_2x + a_1$, *para algum* $a_3, a_2, a_1 \in \mathbb{R}$, *donde basta escolhermos* $a_3 := \frac{b_2}{3}$, $a_2 := \frac{b_1}{2}$ *e* $a_1 := b_0$. *Daí, temos que* $p(x) = \frac{b_2}{3}x^3 + \frac{b_1}{2}x^2 + b_0x \in \mathcal{P}_3(\mathbb{R})$ *é tal que* $T(p(x)) = q(x)$ *e* T *é sobrejetora.*
Um outro jeito seria vermos que $\mathrm{Im}T$ *é gerada por um conjunto com* 3 *vetores que, por sua vez, também gera* $\mathcal{P}_2(\mathbb{R})$, *e concluirmos que* $\mathrm{Im}T = \mathcal{P}_2(\mathbb{R})$.
Para verificarmos que T *não é injetora, basta vermos que o núcleo de* T *é diferente de* $\{\vec{0}\}$. *Assim, se* $T(a_3x^3 + a_2x^2 + a_1x + a_0) = \vec{0}$, *então* $3a_3x^2 + 2a_2x + a_1 = \vec{0}$ *e portanto* $a_3 = a_2 = a_1 = 0$. *Daí, concluímos que* $T(a_0) = \vec{0}$, *para todo* $a_0 \in \mathbb{R}$, *isto é, todos os polinômios constantes de* $\mathcal{P}_3(\mathbb{R})$ *estão no núcleo de* T. *Logo,* T *não é injetora (notemos que mostramos também que* $\mathrm{Ker}T = [1]$*).*

Os exemplos acima ilustram um fato mais geral:

Observação 0.115. *Sejam* U, V *espaços vetoriais sobre* $\mathbb{K}$, *com* U *sendo finitamente gerado. Seja* $T : U \to V$ *uma transformação linear.*

1. *Se* $\dim_{\mathbb{K}} U > \dim_{\mathbb{K}} V$, *então* T *não é injetora*

2. *Se* $\dim_{\mathbb{K}} U < \dim_{\mathbb{K}} V$, *então* T *não é sobrejetora*

De fato, temos que $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} \mathrm{Ker}T + \dim_{\mathbb{K}} \mathrm{Im}T$. *Se* $\dim_{\mathbb{K}} U > \dim_{\mathbb{K}} V \geqslant \dim_{\mathbb{K}} \mathrm{Im}T$, *então* $\dim_{\mathbb{K}} \mathrm{Ker}T > 0$ *(isto é, não pode ser* 0*), donde* $\mathrm{Ker}T \neq \{\vec{0}\}$ *e* T *não é injetora. Por outro lado, se* $\dim_{\mathbb{K}} V > \dim_{\mathbb{K}} U$, *então* $\dim_{\mathbb{K}} V > \dim_{\mathbb{K}} \mathrm{Im}T$ *(ou seja,* $\dim_{\mathbb{K}} \mathrm{Im}T$ *não pode ser* $\dim_{\mathbb{K}} V$*), donde* $\mathrm{Im}T \neq V$ *e* T *não é sobrejetora.*

Exemplo 0.116. *Seja* $T : \mathbb{R}^2 \to \mathbb{R}^3$ *uma aplicação dada por* $T(x, y) = (2x - 2y, 0, y - x)$*. Então* T *é linear (verifique), mas não é injetora, nem sobrejetora. De fato, pela observação acima, já sabemos que* T *não é sobrejetora. Para vermos que* T *não é injetora, basta notarmos, por exemplo, que* $T(2, 3) = T(1, 2)$*. Um outro jeito seria verificarmos que* $\mathrm{Ker}T = [(1, 1)] \neq \{\vec{0}\}$*.*

Exemplo 0.117. *Seja* $T : \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_1(\mathbb{R})$ *uma aplicação dada por* $T(a_2x^2 + a_1x + a_0) = (a_2 - a_1 + a_0)x$*. Então* T *é linear (verifique), mas não é injetora, nem sobrejetora. Com efeito, pela observação acima, já sabemos que* T *não é injetora. Para vermos que* T *não é sobrejetora, basta notarmos, por exemplo, que* $1 \in \mathcal{P}_1(\mathbb{R})$ *não está na imagem de* T *(na verdade, nenhum polinômio que tem termo independente diferente de zero está na imagem de* T*). Um outro jeito seria verificarmos que* $\mathrm{Im}T = [x] \neq [1, x] = \mathcal{P}_1(\mathbb{R})$*.*

Teorema 0.118. *Dois espaços vetoriais finitamente gerados de mesma dimensão são isomorfos.*

De fato, sejam U, V espaços vetoriais sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} V = m$. Sejam $\mathcal{B} := \{u_1, \ldots, u_m\}$ e $\mathcal{C} := \{v_1, \ldots, v_m\}$ bases de U e V, respectivamente. Por um lado, como $\mathcal{B}$ é base de U, sabemos, por uma proposição anterior, que existe uma única transformação linear $T : U \to V$ tal que $T(u_i) = v_i$, com $i = 1, \ldots, m$. Por outro lado, desde que $\mathcal{C}$ é base de V, sabemos, pela mesma proposição, que existe uma única transformação linear $S : V \to U$ tal que $S(v_i) = u_i$, com $i = 1, \ldots, m$. Daí, obtemos que $(S \circ T)(u_i) = S(T(u_i)) = S(v_i) = u_i$ (*) e $(T \circ S)(v_i) = v_i$, para todo $i = 1, \ldots, m$. Seja $u \in U$ qualquer. Então $u = \sum_{i=1}^{m} \alpha_i u_i$, para algum $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$. Logo, como $S \circ T$ é linear, segue que $(S \circ T)(u) = (S \circ T)(\sum_{i=1}^{m} \alpha_i u_i) = \sum_{i=1}^{m} \alpha_i (S \circ T)(u_i) \overset{(*)}{=} \sum_{i=1}^{m} \alpha_i u_i = u$, para todo $u \in U$ e portanto $S \circ T = Id_U$ (transformação identidade de U). Analogamente, $T \circ S = Id_V$. Logo, T é bijetora e portanto isomorfismo. ∎

Exemplo 0.119. *Sejam* $S, T : \mathbb{R}^2 \to \mathcal{P}_1(\mathbb{R})$ *transformações lineares dadas por* $S(a,b) = ax + b$ *e* $T(a,b) = b$. *Então* S *é um isomorfismo (como os espaços têm a mesma dimensão, basta ver que* S *é injetora, por exemplo) e* T *não é um isomorfismo (pois* T *não é injetora). Verifique.*

Observação 0.120. *Aqui vemos que* $\mathbb{R}^2 \cong \mathcal{P}_1(\mathbb{R})$ *(*), mas que nem toda transformação linear de um no outro é um isomorfismo. Podemos verificar (*) ao exibirmos um isomorfismo* S, *ou ao constatarmos que ambos têm a mesma dimensão e usarmos o resultado anterior.*

Corolário 0.121. *Se* V *é um espaço vetorial sobre* $\mathbb{K}$, *com* $\dim_{\mathbb{K}} V = n$, *então* $V \cong \mathbb{K}^n$.

Basta notarmos que $\dim_{\mathbb{K}} \mathbb{K}^n = n$ e aplicarmos o teorema anterior. ■

Matriz de transformação linear

Sejam U, V espaços vetoriais sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U = n$, $\dim_{\mathbb{K}} V = m$. Sejam $\mathcal{B} := \{u_1, \ldots, u_n\}$ e $\mathcal{C} := \{v_1, \ldots, v_m\}$ bases de U e V, respectivamente. Seja $T : U \to V$ uma transformação linear. Desde que $T(u_1) \in V$ e $\mathcal{C}$ é uma base de V, existem (únicos) escalares $a_{11}, \ldots, a_{m1}$ tais que $T(u_1) = a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_m$. Analogamente, para todo j, existem (únicos) escalares $a_{1j}, \ldots, a_{mj}$ tais que $T(u_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{mj}v_m$.
Seja $u \in U$. Desde que $\mathcal{B}$ é uma base de U, existem (únicos) escalares $\alpha_1, \ldots, \alpha_n$ tais que $[u]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)_{\mathcal{B}}$. Como $T(u) \in V$ e $\mathcal{C}$ é uma base de V, existem (únicos) escalares $\beta_1, \ldots, \beta_m$ tais que $[T(u)]_{\mathcal{C}} = (\beta_1, \ldots, \beta_m)_{\mathcal{C}}$. Mas será que não conseguimos explicitar $\beta_1, \ldots, \beta_m$ em função das informações já fixadas? Vejamos: por um lado, $T(u) = \sum_{i=1}^m \beta_i v_i$. Por outro lado, $T(u) = T(\sum_{j=1}^n \alpha_j u_j) \overset{T\text{ linear}}{=} \sum_{j=1}^n \alpha_j T(u_j) = \sum_{j=1}^n \alpha_j \left(\sum_{i=1}^m a_{ij} v_i\right) = \sum_{i=1}^m \left(\sum_{j=1}^n \alpha_j a_{ij}\right) v_i$. Desde que a combinação linear é única, pois $\mathcal{C}$ é base,

então $\beta_1 = \sum_{j=1}^n \alpha_j a_{1j}$, $\ldots$, $\beta_m = \sum_{j=1}^n \alpha_j a_{mj}$. Em notação matricial, obtemos que:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}$$

Definição 0.122. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U = n$, $\dim_{\mathbb{K}} V = m$. Sejam $\mathcal{B}$ e $\mathcal{C}$ bases de U e V, respectivamente, digamos $\mathcal{B} := \{u_1, \ldots, u_n\}$. Seja $T : U \to V$ uma transformação linear. Definimos a* **matriz de T com relação às bases $\mathcal{B}$ e $\mathcal{C}$** *como sendo a matriz $(a_{ij}) \in \mathbb{M}_{m\times n}(\mathbb{K})$, onde $(a_{1j}, a_{2j}, \ldots, a_{mj}) = [T(u_j)]_{\mathcal{C}}$, para todo $j = 1, \ldots, n$. Notação: $[T]_{\mathcal{B},\mathcal{C}} := (a_{ij})$.*

Observação 0.123. *Com as mesmas notações, usando a notação matricial acima, notemos que $[T]_{\mathcal{B},\mathcal{C}}.[u]_{\mathcal{B}} = [T(u)]_{\mathcal{C}}$, onde vemos as coordenadas, por exemplo, de u com relação à base $\mathcal{B}$ como sendo um vetor coluna $n \times 1$, isto é, se $[u]_{\mathcal{B}} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, então consideramos*

$$[u]_{\mathcal{B}} := \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Relembremos a definição de multiplicação entre matrizes.

Observação 0.124. *Sabemos multiplicar duas matrizes do seguinte modo:*

$$\begin{array}{rccc} \cdot : & \mathbb{M}_{m\times n}(\mathbb{K}) \times \mathbb{M}_{n\times p}(\mathbb{K}) & \to & \mathbb{M}_{m\times p}(\mathbb{K}) \\ & ((a_{ij})_{i,j}, (b_{jk})_{j,k}) & \mapsto & (a_{ij})_{i,j}.(b_{jk})_{j,k} := \left(\sum_{j=1}^n a_{ij} b_{jk}\right)_{i,k}, \end{array}$$

isto é,

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} . \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix} :=$$

$$\begin{pmatrix} \sum_{j=1}^{n} a_{1j}b_{j1} & \sum_{j=1}^{n} a_{1j}b_{j2} & \cdots & \sum_{j=1}^{n} a_{1j}b_{jp} \\ \sum_{j=1}^{n} a_{2j}b_{j1} & \sum_{j=1}^{n} a_{2j}b_{j2} & \cdots & \sum_{j=1}^{n} a_{2j}b_{jp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} a_{mj}b_{j1} & \sum_{j=1}^{n} a_{mj}b_{j2} & \cdots & \sum_{j=1}^{n} a_{mj}b_{jp} \end{pmatrix}.$$

Por exemplo,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} . \begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \end{pmatrix}.$$

Exemplo 0.125. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ uma transformação linear dada por $T(x,y) = (y, x, y-x)$. Sejam $\mathcal{B}_C^2$ e $\mathcal{B}_C^3$ as bases canônicas de $\mathbb{R}^2$ e $\mathbb{R}^3$, respectivamente. Sejam $\mathcal{B} := \{(1,1),(1,2)\}$ e $\mathcal{C} := \{(1,1,0),(1,1,1),(1,0,0)\}$ outras bases de $\mathbb{R}^2$ e $\mathbb{R}^3$, respectivamente. Então*

$$\begin{aligned} T(1,1) = (1,1,0) &= \mathbf{1}.(1,1,0) + \mathbf{0}.(1,1,1) + \mathbf{0}.(1,0,0) \Rightarrow [T(1,1)]_{\mathcal{C}} = (1,0,0) \\ &= \mathbf{1}.(1,0,0) + \mathbf{1}.(0,1,0) + \mathbf{0}.(0,0,1) \Rightarrow [T(1,1)]_{\mathcal{B}_C^3} = (1,1,0), \end{aligned}$$

$$\begin{aligned} T(1,2) = (2,1,1) &= \mathbf{0}.(1,1,0) + \mathbf{1}.(1,1,1) + \mathbf{1}.(1,0,0) \Rightarrow [T(1,2)]_{\mathcal{C}} = (0,1,1) \\ &= \mathbf{2}.(1,0,0) + \mathbf{1}.(0,1,0) + \mathbf{1}.(0,0,1) \Rightarrow [T(1,2)]_{\mathcal{B}_C^3} = (2,1,1). \end{aligned}$$

Logo,

$$[T]_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \text{ e } [T]_{\mathcal{B},\mathcal{B}_C^3} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Ainda, como $[T(1,0)]_{\mathcal{C}} = (2,-1,-1)$, $[T(0,1)]_{\mathcal{C}} = (-1,1,1)$, $[T(1,0)]_{\mathcal{B}_C^3} = (0,1,-1)$ e $[T(0,1)]_{\mathcal{B}_C^3} = (1,0,1)$, segue que

$$[T]_{\mathcal{B}_C^2,\mathcal{C}} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \\ -1 & 1 \end{pmatrix} \text{ e } [T]_{\mathcal{B}_C^2,\mathcal{B}_C^3} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Exemplo 0.126. *Seja $S : \mathcal{P}_3(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ uma transformação linear dada por $S(p(x)) = p'(x)$, onde se $p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$, então $p'(x) = 3a_3x^2 + 2a_2x + a_1$. Sejam $\mathcal{B}_C^3 = \{1, x, x^2, x^3\}$ e $\mathcal{B}_C^2 = \{1, x, x^2\}$ as bases canônicas de $\mathcal{P}_3(\mathbb{R})$ e $\mathcal{P}_2(\mathbb{R})$, respectivamente. Como*

$$\begin{cases} S(1) & = & 0 & = & 0.1 + 0.x + 0.x^2 + 0.x^3 \\ S(x) & = & 1 & = & 1.1 + 0.x + 0.x^2 + 0.x^3 \\ S(x^2) & = & 2x & = & 0.1 + 2.x + 0.x^2 + 0.x^3 \\ S(x^3) & = & 3x^2 & = & 0.1 + 0.x + 3.x^2 + 0.x^3 \end{cases}$$

então

$$[S]_{\mathcal{B}_C^3,\mathcal{B}_C^2} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

é a matriz de S com relação às bases $\mathcal{B}_C^3$ e $\mathcal{B}_C^2$.

Observação 0.127. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$, e $\mathcal{B}, \mathcal{C}$ bases de U, V, respectivamente, digamos $\mathcal{B} := \{u_1, \ldots, u_n\}$ e $\mathcal{C} := \{v_1, \ldots, v_m\}$. Por um lado, se $T : U \to V$ é uma transformação linear, então existe uma única matriz $[T]_{\mathcal{B},\mathcal{C}} \in \mathbb{M}_{m\times n}(\mathbb{K})$, em vista da unicidade das coordenadas de $T(u_j)$ com relação à base $\mathcal{C}$, para todo $j = 1, \ldots, n$.*

Por outro lado, dada uma matriz $M = (a_{ij}) \in \mathbb{M}_{m\times n}(\mathbb{K})$, existe uma única transformação linear $S : U \to V$ tal que $[S]_{\mathcal{B},\mathcal{C}} = M$. De fato, consideremos $\mathcal{D} := \{\sum_{i=1}^m a_{i1}v_i, \sum_{i=1}^m a_{i2}v_i, \ldots, \sum_{i=1}^m a_{in}v_i\}$ um subconjunto de vetores de V. Notemos que $\mathcal{D}$ tem o mesmo número de vetores da base $\mathcal{B}$ de U. Daí, temos um resultado que diz que existe uma única transformação linear $S : U \to V$ tal que $S(u_j) = \sum_{i=1}^m a_{ij}v_i$, para todo $j = 1, \ldots, n$. Assim, $[S(u_j)]_{\mathcal{C}} = (a_{1j}, a_{2j}, \ldots, a_{mj})$ e portanto $[S]_{\mathcal{B},\mathcal{C}} = (a_{ij}) = M$.

Proposição 0.128. *Sejam U, V, W espaços vetoriais sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U = n$, $\dim_{\mathbb{K}} V = m$ e $\dim_{\mathbb{K}} W = p$. Sejam $\mathcal{B}, \mathcal{C}, \mathcal{D}$ bases de U, V, W, respectivamente. Sejam $T : U \to V$ e $S : V \to W$ transformações lineares. Então $[S \circ T]_{\mathcal{B},\mathcal{D}} = [S]_{\mathcal{C},\mathcal{D}}.[T]_{\mathcal{B},\mathcal{C}}$.*

Inicialmente, notemos que $[S]_{\mathcal{C},\mathcal{D}} \in \mathbb{M}_{p\times m}(\mathbb{K})$ e $[T]_{\mathcal{B},\mathcal{C}} \in \mathbb{M}_{m\times n}(\mathbb{K})$, donde $[S]_{\mathcal{C},\mathcal{D}}.[T]_{\mathcal{B},\mathcal{C}} \in \mathbb{M}_{p\times n}(\mathbb{K})$ e faz sentido nos perguntarmos se esta é igual a $[S \circ T]_{\mathcal{B},\mathcal{D}} \in \mathbb{M}_{p\times n}(\mathbb{K})$. Seja $\mathcal{B} := \{u_1, \ldots, u_n\}$. Então $([S]_{\mathcal{C},\mathcal{D}}.[T]_{\mathcal{B},\mathcal{C}}).[u_j]_{\mathcal{B}} = [S]_{\mathcal{C},\mathcal{D}}.([T]_{\mathcal{B},\mathcal{C}}.[u_j]_{\mathcal{B}}) = [S]_{\mathcal{C},\mathcal{D}}.[T(u_j)]_{\mathcal{C}} = [S(T(u_j))]_{\mathcal{D}} = [(S \circ T)(u_j)]_{\mathcal{D}}$, isto é, $[S]_{\mathcal{C},\mathcal{D}}.[T]_{\mathcal{B},\mathcal{C}}$ é a matriz $(a_{ij}) \in \mathbb{M}_{p\times n}(\mathbb{K})$, onde $(a_{1j}, a_{2j}, \ldots, a_{pj}) = [(S \circ T)(u_j)]_{\mathcal{D}}$, para todo $j = 1\ldots, n$. Daí, pela definição de $[S \circ T]_{\mathcal{B},\mathcal{D}}$, então $[S]_{\mathcal{C},\mathcal{D}}.[T]_{\mathcal{B},\mathcal{C}} = [S \circ T]_{\mathcal{B},\mathcal{D}}$. ■

Observação 0.129. *Notemos que uma transformação linear é* **invertível** *(isto é, tal que existe a inversa) se, e somente se, é um isomorfismo.*

Corolário 0.130. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} V = n > 0$. Sejam $\mathcal{B}, \mathcal{C}$ bases de U, V, respectivamente. Seja $T : U \to V$ uma transformação linear. Então T é invertível se, e somente se, $[T]_{\mathcal{B},\mathcal{C}}$ é invertível (e neste caso, $([T]_{\mathcal{B},\mathcal{C}})^{-1} = [T^{-1}]_{\mathcal{C},\mathcal{B}}$).*

Se T é invertível, então existe $T^{-1} : V \to U$, que é uma transformação linear, tal que $T \circ T^{-1} = Id_V$ e $T^{-1} \circ T = Id_U$. Notemos que $[T]_{\mathcal{B},\mathcal{C}}, [T^{-1}]_{\mathcal{C},\mathcal{B}} \in \mathbb{M}_{n\times n}(\mathbb{K})$. Daí, pelo resultado anterior, $[T]_{\mathcal{B},\mathcal{C}}.[T^{-1}]_{\mathcal{C},\mathcal{B}} = [T \circ T^{-1}]_{\mathcal{C},\mathcal{C}} = [Id_V]_{\mathcal{C},\mathcal{C}} = I_n$ e $[T^{-1}]_{\mathcal{C},\mathcal{B}}.[T]_{\mathcal{B},\mathcal{C}} = [T^{-1} \circ T]_{\mathcal{B},\mathcal{B}} = [Id_U]_{\mathcal{B},\mathcal{B}} = I_n$, onde $I_n \in \mathbb{M}_{n\times n}(\mathbb{K})$ denota a matriz identidade $n \times n$. Logo, $[T]_{\mathcal{B},\mathcal{C}}$ é invertível e $([T]_{\mathcal{B},\mathcal{C}})^{-1} = [T^{-1}]_{\mathcal{C},\mathcal{B}}$.
Agora, se $[T]_{\mathcal{B},\mathcal{C}}$ é invertível, então existe $([T]_{\mathcal{B},\mathcal{C}})^{-1} \in \mathbb{M}_{n\times n}(\mathbb{K})$ tal que $[T]_{\mathcal{B},\mathcal{C}}.([T]_{\mathcal{B},\mathcal{C}})^{-1} = I_n = ([T]_{\mathcal{B},\mathcal{C}})^{-1}.[T]_{\mathcal{B},\mathcal{C}}$. Mas, dada a matriz $([T]_{\mathcal{B},\mathcal{C}})^{-1}$, sabemos que existe uma única transformação linear $S : V \to U$ tal que $[S]_{\mathcal{C},\mathcal{B}} = ([T]_{\mathcal{B},\mathcal{C}})^{-1}$. Daí, pelo resultado anterior, temos que $[T \circ S]_{\mathcal{C},\mathcal{C}} = I_n = [S \circ T]_{\mathcal{B},\mathcal{B}}$.

Seja $u \in U$ qualquer. Como $[(S \circ T)(u)]_{\mathcal{B}} = [S \circ T]_{\mathcal{B},\mathcal{B}}.[u]_{\mathcal{B}} = I_n.[u]_{\mathcal{B}} = [u]_{\mathcal{B}}$, isto é, os vetores $(S \circ T)(u)$ e u têm as mesmas coordenadas com relação à base $\mathcal{B}$, então $(S \circ T)(u) = u$, para todo $u \in U$. Logo, $S \circ T = Id_U$. Analogamente, obtemos que $T \circ S = Id_V$. Portanto, T é invertível, $T^{-1} = S$ e $[T^{-1}]_{\mathcal{C},\mathcal{B}} = ([T]_{\mathcal{B},\mathcal{C}})^{-1}$. ■

Observação 0.131. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$. Sejam $S, T : U \to V$ transformações lineares e $\lambda \in \mathbb{K}$. Definindo $S + T : U \to V$ por $(S+T)(u) = S(u) + T(u)$ e $\lambda S : U \to V$ por $(\lambda S)(u) = \lambda.S(u)$, temos que ambas são transformações lineares (verifique). Consideremos o conjunto, $\mathcal{L}(U, V)$, de todas as transformações lineares de U em V, e as operações sobre $\mathcal{L}(U, V)$:*

$$\begin{array}{rccc} \boldsymbol{+}: & \mathcal{L}(U,V) \times \mathcal{L}(U,V) & \to & \mathcal{L}(U,V) \\ & (S,T) & \to & S+T \\ \cdot: & \mathbb{K} \times \mathcal{L}(U,V) & \to & \mathcal{L}(U,V) \\ & (\lambda, S) & \to & \lambda S \end{array}$$

Então, com estas operações, $\mathcal{L}(U, V)$ é um espaço vetorial sobre $\mathbb{K}$ (verifique as propriedades). Notemos apenas que o vetor nulo de $\mathcal{L}(U, V)$ é a transformação identicamente nula.

Observação 0.132. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U = n$ e $\dim_{\mathbb{K}} V = m$. Sejam $\mathcal{B}, \mathcal{C}$ bases de U, V, respectivamente, digamos $\mathcal{C} := \{v_1, \ldots, v_m\}$. Sejam $S, T \in \mathcal{L}(U, V)$ e $\lambda \in \mathbb{K}$. Então $[S + T]_{\mathcal{B},\mathcal{C}} = [S]_{\mathcal{B},\mathcal{C}} + [T]_{\mathcal{B},\mathcal{C}}$ e $[\lambda S]_{\mathcal{B},\mathcal{C}} = \lambda.[S]_{\mathcal{B},\mathcal{C}}$. De fato, seja $u \in U$. Se $[S(u)]_{\mathcal{C}} = (\alpha_1, \ldots, \alpha_m)$ e $[T(u)]_{\mathcal{C}} = (\beta_1, \ldots, \beta_m)$, então $(S + T)(u) = S(u) + T(u) = \sum_{i=1}^m \alpha_i v_i + \sum_{i=1}^m \beta_i v_i = \sum_{i=1}^m (\alpha_i + \beta_i) v_i$, ou seja, $[(S + T)(u)]_{\mathcal{C}} = (\alpha_1 + \beta_1, \ldots, \alpha_m + \beta_m) = [S(u)]_{\mathcal{C}} + [T(u)]_{\mathcal{C}}$. Da mesma forma, obtemos que $[(\lambda S)(u)]_{\mathcal{C}} = (\lambda\beta_1, \ldots, \lambda\beta_m) = \lambda[S(u)]_{\mathcal{C}}$. Logo, $[S + T]_{\mathcal{B},\mathcal{C}}.[u]_{\mathcal{B}} = ([S]_{\mathcal{B},\mathcal{C}} + [T]_{\mathcal{B},\mathcal{C}}).[u]_{\mathcal{B}}$ e $[\lambda S]_{\mathcal{B},\mathcal{C}}.[u]_{\mathcal{B}} = (\lambda.[S]_{\mathcal{B},\mathcal{C}}).[u]_{\mathcal{B}}$, para todo $u \in U$ e portanto $[S + T]_{\mathcal{B},\mathcal{C}} = [S]_{\mathcal{B},\mathcal{C}} + [T]_{\mathcal{B},\mathcal{C}}$ e $[\lambda S]_{\mathcal{B},\mathcal{C}} = \lambda.[S]_{\mathcal{B},\mathcal{C}}$.*

Proposição 0.133. *Sejam U, V espaços vetoriais sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U = n$ e $\dim_{\mathbb{K}} V = m$. Então $\dim_{\mathbb{K}} \mathcal{L}(U, V) = m.n$.*

De fato, sejam $\mathcal{B}$ e $\mathcal{C}$ bases de U e V, respectivamente. Sabemos que $\dim_{\mathbb{K}} \mathbb{M}_{m\times n}(\mathbb{K}) = m.n$ e que, para toda $T \in \mathcal{L}(U, V)$, existe uma única $[T]_{\mathcal{B},\mathcal{C}} \in \mathbb{M}_{m\times n}(\mathbb{K})$. Assim, se mostrarmos que $\mathcal{L}(U, V) \cong \mathbb{M}_{m\times n}(\mathbb{K})$, então $\dim_{\mathbb{K}} \mathcal{L}(U, V) = \dim_{\mathbb{K}} \mathbb{M}_{m\times n}(\mathbb{K}) = m.n$. Façamos isto. Definamos $\varphi : \mathcal{L}(U, V) \to \mathbb{M}_{m\times n}(\mathbb{K})$ por $\varphi(T) = [T]_{\mathcal{B},\mathcal{C}}$ e verifiquemos que φ é um isomorfismo. Com efeito, φ é bem definida, pois dada T, existe uma única $[T]_{\mathcal{B},\mathcal{C}}$. Ainda, φ é linear, pois $\varphi(\lambda S + T) = [\lambda S + T]_{\mathcal{B},\mathcal{C}} = \lambda.[S]_{\mathcal{B},\mathcal{C}} + [T]_{\mathcal{B},\mathcal{C}} = \lambda\varphi(S) + \varphi(T)$, com $S, T \in \mathcal{L}(U, V)$, $\lambda \in \mathbb{K}$. Agora, φ é sobrejetora, pois, dada uma matriz $M \in \mathbb{M}_{m\times n}(\mathbb{K})$, existe uma única transformação linear $S : U \to V$ tal que $[S]_{\mathcal{B},\mathcal{C}} = M$, isto é, $\varphi(S) = M$. Finalmente, φ é injetora, pois, se $\varphi(T) = \vec{0}_{\mathbb{M}_{m\times n}(\mathbb{K})}$, então $[T]_{\mathcal{B},\mathcal{C}} = \vec{0}_{\mathbb{M}_{m\times n}(\mathbb{K})}$ e portanto $[T(u)]_{\mathcal{C}} = \vec{0}_{\mathbb{K}^m}$, para todo $u \in U$. Daí, $T(u) = \vec{0}_V$, para todo $u \in U$ e portanto T é a transformação identicamente nula. Logo, $\mathrm{Ker}\varphi = \{\vec{0}_{\mathcal{L}(U,V)}\}$. ■

Observação 0.134. *Notemos que, na demonstração acima, desde que, dada $M \in \mathbb{M}_{m\times n}(\mathbb{K})$, existe uma única $S \in \mathcal{L}(U, V)$ tal que $\varphi(S) = M$, então já poderíamos ter concluído direto que φ era bijetora, pois $\varphi(S) = \varphi(T)$ implica, da unicidade de S, que $S = T$.*

Definição 0.135. *Seja U um espaço vetorial sobre $\mathbb{K}$. Chamamos uma transformação linear $T : U \to U$ de um* **operador linear**.

Corolário 0.136. *Seja U um espaço vetorial sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U = n$. Então $\dim_{\mathbb{K}} \mathcal{L}(U, U) = n^2$.*

Basta aplicarmos a proposição anterior. ■

Observação 0.137. *Seja U um espaço vetorial sobre $\mathbb{K}$. Observemos que, sobre $\mathcal{L}(U, U)$, além das operações $+$ e $\cdot$, que o tornam um espaço vetorial,* *podemos considerar a operação* $\circ : \mathcal{L}(U,U) \times \mathcal{L}(U,U) \to \mathcal{L}(U,U)$ *dada por* $(S,T) \mapsto S \circ T$. *Daí, denotamos*

$$T^k := \underbrace{T \circ T \circ \cdots \circ T}_{k} \text{ e } T^0 := Id.$$

Ainda, dado $p(x) \in \mathcal{P}(\mathbb{K})$, $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, *denotamos o operador linear* $a_n T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0 Id \in \mathcal{L}(U,U)$ *simplesmente por* $p(T)$. *Daí, temos que* $p(T)(u) = a_n.T^n(u) + a_{n-1}.T^{n-1}(u) + \cdots + a_1.T(u) + a_0.u$. *Notemos que as três operações,* $+, \cdot, \circ$, *aparecem em* $p(T)$.

Observação 0.138. *Seja* U *um espaço vetorial sobre* $\mathbb{K}$, *com* $\dim_{\mathbb{K}} U = n > 0$. *Sejam* $\mathcal{B}, \mathcal{C}$ *bases de* U. *Consideremos o operador linear identidade (de* U*),* $Id : U \to U$. *Já temos que* $[Id]_{\mathcal{B},\mathcal{C}}.[u]_{\mathcal{B}} = [Id(u)]_{\mathcal{C}} = [u]_{\mathcal{C}}$, *ou seja,* $[Id]_{\mathcal{B},\mathcal{C}}$ *é tal que a partir das coordenadas de* u *com relação à base* $\mathcal{B}$, *obtemos as coordenadas de* u *com relação à base* $\mathcal{C}$, *para todo* $u \in U$. *Notemos ainda que* $[Id]_{\mathcal{B},\mathcal{B}} = I_n = [Id]_{\mathcal{C},\mathcal{C}}$, *isto é, ambas são iguais à matriz identidade* $n \times n$.

Definição 0.139. *Chamamos* $[Id]_{\mathcal{B},\mathcal{C}}$ *de* **matriz de mudança (de bases) com relação às bases $\mathcal{B}$ e $\mathcal{C}$**.

Exemplo 0.140. *Sejam* $\mathcal{B} := \{(0,1),(1,1)\}$ *e* $\mathcal{C} := \{(1,1),(1,2)\}$ *bases de* $\mathbb{R}^2$. *Se* $(2,3)_{\mathcal{C}}$ *são as coordenadas de um vetor* (x,y) *com relação à base* $\mathcal{C}$, *quais são suas coordenadas com relação à base* $\mathcal{B}$*? Calculemos* $[Id]_{\mathcal{C},\mathcal{B}}$. *Bem,* $[Id(1,1)]_{\mathcal{B}} = [(1,1)]_{\mathcal{B}} = (0,1)$, *pois* $(1,1) = \mathbf{0}.(0,1) + \mathbf{1}.(1,1)$ *e* $[Id(1,2)]_{\mathcal{B}} = [(1,2)]_{\mathcal{B}} = (1,1)$, *pois* $(1,2) = \mathbf{1}.(0,1) + \mathbf{1}.(1,1)$, *donde* $[Id]_{\mathcal{C},\mathcal{B}} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$. *Daí,* $[(x,y)]_{\mathcal{B}} = [Id]_{\mathcal{C},\mathcal{B}}.[(x,y)]_{\mathcal{C}} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 3 \\ 5 \end{smallmatrix}\right)$. *Ou de outra forma, se* $(x,y) = 2(1,1) + 3(1,2) = (5,8) = \alpha(0,1) + \beta(1,1)$, *então* $\alpha = 3$ *e* $\beta = 5$, *donde* $[(x,y)]_{\mathcal{B}} = (3,5)$.

Observação 0.141. *Em geral, a matriz de mudança com relação às bases* $\mathcal{C}$ *e* $\mathcal{B}$, $[Id]_{\mathcal{C},\mathcal{B}}$, *é diferente de* $[Id]_{\mathcal{B},\mathcal{C}}$. *Mas, como* Id *é invertível e* $Id^{-1} = Id$ *(*), segue, por um resultado anterior, que* $[Id]_{\mathcal{C},\mathcal{B}} \overset{(*)}{=} [Id^{-1}]_{\mathcal{C},\mathcal{B}} = ([Id]_{\mathcal{B},\mathcal{C}})^{-1}$, *ou seja, uma é a inversa da outra.*

Vejamos a definição de semelhança entre matrizes.

Observação 0.142. *Sejam $M, N \in \mathbb{M}_{n\times n}(\mathbb{K})$. Lembremos que M e N são* **semelhantes** *se existe existe uma matriz invertível $P \in \mathbb{M}_{n\times n}(\mathbb{K})$ tal que $M = P^{-1}NP$. Notação: $M \sim N$.*

Proposição 0.143. *Seja U um espaço vetorial sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} U = n > 0$. Sejam $\mathcal{B}, \mathcal{C}$ bases de U. Seja $T \in \mathcal{L}(U, U)$. Então $[T]_{\mathcal{B},\mathcal{B}} \sim [T]_{\mathcal{C},\mathcal{C}}$.*

De fato, consideremos a matriz invertível $[Id]_{\mathcal{B},\mathcal{C}} \in \mathbb{M}_{n\times n}(\mathbb{K})$. Como podemos ver T como sendo a composta $Id \circ T \circ Id$, então

$$\begin{aligned} [T]_{\mathcal{B},\mathcal{B}} &= [Id \circ T \circ Id]_{\mathcal{B},\mathcal{B}} \\ &= [Id]_{\mathcal{C},\mathcal{B}}.[T]_{\mathcal{C},\mathcal{C}}.[Id]_{\mathcal{B},\mathcal{C}} \\ &= ([Id]_{\mathcal{B},\mathcal{C}})^{-1}.[T]_{\mathcal{C},\mathcal{C}}.[Id]_{\mathcal{B},\mathcal{C}} \end{aligned}$$

e portanto $[T]_{\mathcal{B},\mathcal{B}} \sim [T]_{\mathcal{C},\mathcal{C}}$. ■

Observação 0.144. *Chamamos a matriz de T com relação às bases $\mathcal{B}$ e $\mathcal{B}$, $[T]_{\mathcal{B},\mathcal{B}}$, simplesmente de matriz de T com relação à base $\mathcal{B}$ e a denotamos por $[T]_{\mathcal{B}}$. Assim, o resultado anterior nos diz que $[T]_{\mathcal{B}}$ e $[T]_{\mathcal{C}}$ são semelhantes.*

Exemplo 0.145. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ um operador linear dado por $T(x, y) = (x, x + y)$. Sejam $\mathcal{B} := \{(0, 1), (1, 1)\}$ e $\mathcal{C} := \{(1, 1), (1, 2)\}$ bases de $\mathbb{R}^2$. Então*

$$\begin{aligned} T(0,1) = (0,1) &= \mathbf{1}.(0,1) + \mathbf{0}.(1,1) \Rightarrow [T(0,1)]_{\mathcal{B}} = (1,0) \\ &= \mathbf{-1}.(1,1) + \mathbf{1}.(1,2) \Rightarrow [T(0,1)]_{\mathcal{C}} = (-1,1), \end{aligned}$$

$$\begin{aligned} T(1,1) = (1,2) &= \mathbf{1}.(0,1) + \mathbf{1}.(1,1) \Rightarrow [T(1,1)]_{\mathcal{B}} = (1,1) \\ &= \mathbf{0}.(1,1) + \mathbf{1}.(1,2) \Rightarrow [T(1,1)]_{\mathcal{C}} = (0,1). \end{aligned}$$

Logo, $[T]_{\mathcal{B}} = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ e $[T]_{\mathcal{B},\mathcal{C}} = \left(\begin{smallmatrix} -1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$. Ainda, como $[T(1,1)]_{\mathcal{C}} = (0,1)$ e $[T(1,2)]_{\mathcal{C}} = (-1,2)$, então $[T]_{\mathcal{C}} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 2 \end{smallmatrix}\right)$ (do mesmo modo, obtenha $[T]_{\mathcal{C},\mathcal{B}}$). Agora, desde que $[Id(0,1)]_{\mathcal{C}} = [(0,1)]_{\mathcal{C}} = (-1,1)$ e $[Id(1,1)]_{\mathcal{C}} = [(1,1)]_{\mathcal{C}} = (1,0)$, então $[Id]_{\mathcal{B},\mathcal{C}} = \left(\begin{smallmatrix} -1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Notemos que, pelo resultado anterior, $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \sim \left(\begin{smallmatrix} 0 & -1 \\ 1 & 2 \end{smallmatrix}\right)$, ou melhor, $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 & 1 \\ 1 & 0 \end{smallmatrix}\right)^{-1} \left(\begin{smallmatrix} 0 & -1 \\ 1 & 2 \end{smallmatrix}\right) \left(\begin{smallmatrix} -1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

Observação 0.146. *Seja U um espaço vetorial sobre $\mathbb{K}$ de dimensão finita. Sejam $\mathcal{B}, \mathcal{C}, \mathcal{D}$ bases de U. Como podemos ver Id como sendo a composta $Id \circ Id$, então $[Id]_{\mathcal{B},\mathcal{C}} = [Id \circ Id]_{\mathcal{B},\mathcal{C}} = [Id]_{\mathcal{D},\mathcal{C}}.[Id]_{\mathcal{B},\mathcal{D}} = ([Id]_{\mathcal{C},\mathcal{D}})^{-1}.[Id]_{\mathcal{B},\mathcal{D}}$, donde se conhecemos estas duas últimas matrizes de mudança, obtemos $[Id]_{\mathcal{B},\mathcal{C}}$.*

Exemplo 0.147. *Sejam $\mathcal{B} := \{(0,1),(1,1)\}$ e $\mathcal{C} := \{(1,1),(1,2)\}$ bases de $\mathbb{R}^2$. Como $[Id]_{\mathcal{B},\mathcal{B}_C} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ e $[Id]_{\mathcal{C},\mathcal{B}_C} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$, então*

$$[Id]_{\mathcal{B},\mathcal{C}} = ([Id]_{\mathcal{C},\mathcal{B}_C})^{-1}[Id]_{\mathcal{B},\mathcal{B}_C} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)^{-1}\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & -1 \\ -1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 & 1 \\ 1 & 0 \end{smallmatrix}\right).$$

Operador diagonalizável

Sejam V um espaço vetorial sobre $\mathbb{K}$ e $T : V \to V$ um operador linear. Se $\mathcal{B} := \{v_1, \ldots, v_n\}$ é uma base de V tal que

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

para algum $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, então, pela construção de $[T]_{\mathcal{B}}$, temos que $T(v_1) = \lambda_1 v_1 + 0v_2 + \cdots + 0v_n = \lambda_1 v_1$, $T(v_2) = \lambda_2 v_2, \ldots, T(v_n) = \lambda_n v_n$.

Observação 0.148. *Seja $M := (a_{ij}) \in \mathbb{M}_{n\times n}(\mathbb{K})$. Dizemos que M é* **diagonal***, se $a_{ij} = 0$, para todo $i \neq j$.*

Exemplo 0.149. *Temos que $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right) \in \mathbb{M}_{2\times 2}(\mathbb{R})$ são diagonais. Temos que $\left(\begin{smallmatrix} 0 & 0 \\ 3 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 3 & 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 3 & 2 \end{smallmatrix}\right) \in \mathbb{M}_{2\times 2}(\mathbb{R})$ não são diagonais.*

Definição 0.150. *Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T : V \to V$ um operador linear.*

1. *Se, para algum $v \in V$, $v \neq \overrightarrow{0}$, existe $\lambda \in \mathbb{K}$ tal que $T(v) = \lambda v$, dizemos que λ é um* **autovalor de T**

2. *Seja $\lambda \in \mathbb{K}$ um autovalor de T. Todo vetor $v \in V$, $v \neq \overrightarrow{0}$, é chamado de* **autovetor de T associado a λ**. *Denotamos por $\mathrm{Aut}_T(\lambda)$ o subespaço de V gerado pelos autovetores de T associados a λ e chamamos-no de* **autoespaço de T associado a λ**

3. *Dizemos que T é* **diagonalizável** *se existe uma base $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ é diagonal, ou seja, se V admite uma base formada por autovetores de T*

Exemplo 0.151. *Considere $\mathbb{R}^2$ como espaço vetorial sobre $\mathbb{R}$. Seja $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ dado por $T(x,y) = (-y, 5x+6y)$. Então T é diagonalizável. De fato, suponhamos que $T(x,y) = \lambda(x,y)$, isto é, $(-y, 5x+6y) = \lambda(x,y)$, para algum $\lambda \in \mathbb{R}$. Então*

$$\begin{cases} -y = \lambda x \\ 5x + 6y = \lambda y \end{cases}$$

e daí $(\lambda^2 - 6\lambda + 5)x = 0$. Se $x = 0$, então $y = 0$, e não obtemos informações acerca de autovalores (estamos dizendo que $T(0,0) = (0,0)$, o que é verdade). Se $x \neq 0$, então $\lambda^2 - 6\lambda + 5 = 0$ e portanto ou $\lambda = 1$, ou $\lambda = 5$. Se $\lambda = 1$, temos que $y = -x$ e portanto os autovetores de T associados a 1 são da forma $(x, -x)$, com $x \in \mathbb{R}$, e $\mathrm{Aut}_T(1) = [(1,-1)]$. Agora, se $\lambda = 5$, segue que $y = -5x$ e daí os autovetores de T associados a 5 são da forma $(x, -5x)$, com $x \in \mathbb{R}$, e $\mathrm{Aut}_T(5) = [(1,-5)]$. Como $\mathcal{B} := \{(1,-1), (1,-5)\}$ é uma base de $\mathbb{R}^2$ formada por autovetores de T (ou equivalentemente, como $[T]_{\mathcal{B}} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 5 \end{smallmatrix}\right)$ diagonal), então T é diagonalizável.

Exemplo 0.152. *Considere $\mathbb{R}^2$ como espaço vetorial sobre $\mathbb{R}$. Seja $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ dado por $T(x,y) = (-y, x)$. Então T não é diagonalizável. Com efeito, suponhamos que $T(x,y) = \alpha(x,y)$, isto é, $(-y, x) = \alpha(x,y)$, para* *algum $\alpha \in \mathbb{R}$. Então*

$$\begin{cases} -y = \alpha x \\ \ \ x = \alpha y \end{cases}$$

e daí $(1-\alpha^2)x = 0$. Se $x = 0$, então $y = 0$, e não obtemos informações sobre de autovalores (estamos dizendo que $T(0,0) = (0,0)$, o que é verdade). Se $x \neq 0$, então $1 - \alpha^2 = 0$, o que não ocorre, pois $\alpha \in \mathbb{R}$. Como T não admite autovalores, segue que também não admite autovetores. Logo, V não admite uma base formada por autovetores e daí T não é diagonalizável.

Observação 0.153. *Se $T \in \mathcal{L}(V,V)$ é não injetor, então $0 \in \mathbb{K}$ é um autovalor de T. De fato, como $\mathrm{Ker}T \neq \{\vec{0}\}$, então existe $v \in V$, $v \neq \vec{0}$, tal que $T(v) = \vec{0}$. Daí, $T(v) = 0v$, com $v \neq \vec{0}$, e portanto 0 é autovalor de T.*

Exemplo 0.154. *Considere $\mathbb{R}^2$ como espaço vetorial sobre $\mathbb{R}$. Seja $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ dado por $T(x,y) = (x,0)$. Então 1 e 0 são autovalores de T. Com efeito, se $(x,0) = \beta(x,y)$, para algum $\beta \in \mathbb{R}$, então $x = \beta x$ e $0 = \beta y$. Se $x \neq 0$, então $\beta = 1$ e $y = 0$, donde 1 é autovalor e os autovetores associados são da forma $(x,0)$, com $x \in \mathbb{R}$. Se $y \neq 0$, então $\beta = 0$ e $x = 0$, donde 0 é autovalor e os autovetores associados são da forma $(0,y)$, com $y \in \mathbb{R}$. Assim, $\mathrm{Aut}_T(1) = [(1,0)]$ e $\mathrm{Aut}_T(0) = [(0,1)]$.*

Relembremos a definição de determinante.

Observação 0.155. *Seja $A := (a_{ij}) \in \mathbb{M}_{n\times n}(\mathbb{K})$. Se $n = 1$, definimos $\det A = a_{11}$. Se $n = 2$, definimos $\det A = a_{11}\det A_{11} - a_{12}\det A_{12}$, onde A_{11} (respectivamente, A_{12}) é a matriz 1×1 dada quando desconsideramos a primeira linha e a primeira (respectivamente, segunda) coluna da matriz A. Isto é, se $n = 2$, $\boldsymbol{\det A = a_{11}a_{22} - a_{12}a_{21}}$. Assim, se A_{ij} é a matriz $(n-1)\times(n-1)$ dada quando desconsideramos a i-ésima linha e a j-ésima coluna de A, definimos, a partir de uma linha i fixada, o* **determinante de** *A como sendo $\det A = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}.\det A_{ij}$ (ou da mesma forma, a partir*

de uma linha j fixada, $\det A = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}.\det A_{ij}$*). Observemos que, em geral, o melhor para fazermos este cálculo é escolhermos ou uma linha, ou uma coluna, com o maior número de zeros possível. Observemos ainda que* $\det A \in \mathbb{K}$*, para todo* $A \in \mathbb{M}_{n\times n}(\mathbb{K})$*.*
Notemos que uma matriz $M \in \mathbb{M}_{n\times n}(\mathbb{K})$ é invertível se, e somente se, $\det M \neq 0$*.*

Exemplo 0.156. *Seja $A := (a_{ij}) \in \mathbb{M}_{3\times 3}(\mathbb{K})$. Então, a partir da primeira linha por exemplo,* $\det A = a_{11}\det A_{11} + a_{12}\det A_{12} + a_{13}\det A_{13}$*, onde* $A_{11} = \left(\begin{smallmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{smallmatrix}\right), A_{12} = \left(\begin{smallmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{smallmatrix}\right), A_{13} = \left(\begin{smallmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{smallmatrix}\right)$*.*
Seja

$$B = \begin{pmatrix} 1 & 0 & 7 \\ 2 & 1 & 6 \\ 3 & 0 & 5 \end{pmatrix} \in \mathbb{M}_{3\times 3}(\mathbb{R}).$$

Como entre todas as linhas e entre todas as colunas, a que tem mais zeros é a segunda coluna, então calculemos o determinante de B a partir desta. Daí, $\det B = 0\det A_{12} + 1\det A_{22} + 0\det A_{32} = \det A_{22} = \det\left(\begin{smallmatrix} 1 & 7 \\ 3 & 5 \end{smallmatrix}\right) = 1.5 - 7.3 = -16$.

Relembremos agora a definição de matriz adjunta.

Observação 0.157. *Seja $A := (a_{ij}) \in \mathbb{M}_{n\times n}(\mathbb{K})$. Com as mesmas notações acima, definimos a* **matriz adjunta de** *A como sendo $Adj(A) := (b_{ji}) \in \mathbb{M}_{n\times n}(\mathbb{K})$, onde* $b_{ij} = (-1)^{i+j}\det A_{ij}$*. Notemos que $Adj(A).A = A.Adj(A) =$* $(\det A)I$*, donde, se* $\det A \neq 0$*, então* $A^{-1} = \frac{Adj(A)}{\det A}$*.*

Exemplo 0.158. *Seja $A := (a_{ij}) \in \mathbb{M}_{2\times 2}(\mathbb{K})$. Se* $\det A \neq 0$*, então* $A^{-1} = \frac{1}{\det A}\left(\begin{smallmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{smallmatrix}\right)$*. Seja*

$$B = \begin{pmatrix} 1 & 0 & 7 \\ 2 & 1 & 6 \\ 3 & 0 & 5 \end{pmatrix} \in \mathbb{M}_{3\times 3}(\mathbb{R}).$$

Então

$$Adj(B) = \begin{pmatrix} 5 & 0 & -7 \\ 8 & -16 & 8 \\ -3 & 0 & 1 \end{pmatrix} \in \mathbb{M}_{3\times 3}(\mathbb{R})$$

e $B^{-1} = \frac{Adj(B)}{-16}$.

Observação 0.159. *Sejam* V *um espaço vetorial sobre* $\mathbb{K}$*, com* $\dim_{\mathbb{K}} V = n > 0$*, e* $T \in \mathcal{L}(V,V)$*. Se* $\mathcal{C}$ *é uma base de* V*, podemos considerar* $[T]_{\mathcal{C}} \in \mathbb{M}_{n\times n}(\mathbb{K})$*. Daí, para toda base* $\mathcal{B}$ *de* V*, existe* P *invertível tal que* $[T]_{\mathcal{B}} = P^{-1}[T]_{\mathcal{C}}P$ *e, por propriedades de determinante,*

$$\det([T]_{\mathcal{B}}) = \det(P^{-1})\det([T]_{\mathcal{C}})\det P = (\det P)^{-1}\det([T]_{\mathcal{C}})\det P = \det([T]_{\mathcal{C}}).$$

Notemos que se $\lambda \in \mathbb{K}$*, então* $\lambda Id - T$ *é um operador linear de* V *em* V *e* $\det([\lambda Id - T]_{\mathcal{B}}) = \det([\lambda Id - T]_{\mathcal{C}})$*. Ainda, se* $[T]_{\mathcal{C}} := (a_{ij})$*, notemos que*

$$\det([xId - T]_{\mathcal{C}}) = \det \begin{pmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1(n-1)} & -a_{1n} \\ -a_{21} & x - a_{22} & & -a_{2(n-1)} & -a_{2n} \\ \vdots & & \ddots & \vdots & \vdots \\ -a_{(n-1)1} & -a_{(n-1)2} & \cdots & x - a_{(n-1)(n-1)} & -a_{(n-1)n} \\ -a_{n1} & -a_{n2} & \cdots & -a_{n(n-1)} & x - a_{nn} \end{pmatrix},$$

que é um polinômio em x *de grau* n *com coeficientes em* $\mathbb{K}$ *e é tal que* $\det([xId - T]_{\mathcal{B}}) = \det([xId - T]_{\mathcal{C}})$*, isto é, independe da base escolhida.*

Definição 0.160. *Sejam* V *um espaço vetorial sobre* $\mathbb{K}$*, com* $\dim_{\mathbb{K}} V = n > 0$*, e* $T \in \mathcal{L}(V,V)$*. Seja* $\mathcal{C}$ *uma base de* V*. Chamamos o polinômio* $\det([xId - T]_{\mathcal{C}})$ *(que é de grau* n *com coeficientes em* $\mathbb{K}$*) de* **polinômio característico de** T *e denotamos-no por* $p_T(x)$*.*

Proposição 0.161. *Sejam* V *um espaço vetorial sobre* $\mathbb{K}$ *de dimensão finita e* $T \in \mathcal{L}(V,V)$*. Então as condições abaixo são equivalentes:*

1. *$\lambda \in \mathbb{K}$ é um autovalor de T*

2. $\operatorname{Ker}(\lambda Id - T) \neq \{\vec{0}\}$

3. $\det([\lambda Id - T]_{\mathcal{C}}) = 0$*, para alguma base $\mathcal{C}$ de V*

4. *$\lambda \in \mathbb{K}$ é uma raiz de $p_T(x)$*

De fato, seja $\mathcal{C}$ uma base de V.
Mostremos ($1 \Leftrightarrow 2$). Se λ é um autovalor de T, então existe $v \neq \vec{0}$ tal que $T(v) = \lambda v = \lambda Id(v)$. Daí, $(\lambda Id - T)(v) = \vec{0}$ e portanto $v \in \operatorname{Ker}(\lambda Id - T)$. Logo, como $v \neq \vec{0}$, $\operatorname{Ker}(\lambda Id - T) \neq \{\vec{0}\}$. Agora, se $\operatorname{Ker}(\lambda Id - T) \neq \{\vec{0}\}$, então existe $v \neq \vec{0}$ tal que $(\lambda Id - T)(v) = \vec{0}$. Daí, $T(v) = \lambda Id(v) = \lambda v$ e portanto, como $v \neq \vec{0}$, λ é um autovalor de T.
Mostremos agora ($2 \Leftrightarrow 3$). Se $\operatorname{Ker}(\lambda Id - T) \neq \{\vec{0}\}$, então $\lambda Id - T$ não é injetor e portanto $\lambda Id - T$ não é invertível. Daí, $[\lambda Id - T]_{\mathcal{C}}$ não é invertível e portanto $\det([\lambda Id - T]_{\mathcal{C}}) = 0$. Se agora $\det([\lambda Id - T]_{\mathcal{C}}) = 0$, então $[\lambda Id - T]_{\mathcal{C}}$ não é invertível e portanto $\lambda Id - T$ não é invertível. Em particular, como as dimensões do domínio e do contra-domínio de $\lambda Id - T$ são as mesmas, $\lambda Id - T$ não é sobrejetor, nem injetor. Deste último, segue que, $\operatorname{Ker}(\lambda Id - T) \neq \{\vec{0}\}$.
Mostremos finalmente ($3 \Leftrightarrow 4$). Se $\det([\lambda Id - T]_{\mathcal{C}}) = 0$, então $p_T(\lambda) = \det([\lambda Id - T]_{\mathcal{C}}) = 0$ e portanto λ é uma raiz de $p_T(x)$. Se agora $p_T(\lambda) = 0$, então é claro que $\det([\lambda Id - T]_{\mathcal{C}}) = 0$. ■

Observação 0.162. *Ainda, com as mesmas notações acima, se $\lambda \in \mathbb{K}$ é um autovalor de T, segue, da demonstração, que* $\operatorname{Aut}_T(\lambda) = \operatorname{Ker}(\lambda Id - T)$.

Exemplo 0.163. *Considere $\mathbb{R}^2$ como espaço vetorial sobre $\mathbb{R}$. Seja $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ dado por $T(x, y) = (-y, 5x+6y)$. Já vimos que T é diagonalizável. Temos que $[T]_{\mathcal{B}_C} = \left(\begin{smallmatrix} 0 & -1 \\ 5 & 6 \end{smallmatrix}\right)$. Daí, $p_T(x) = \det\left(\begin{smallmatrix} x & 1 \\ -5 & x-6 \end{smallmatrix}\right) = x(x-6) + 5 = x^2 - 6x + 5 = (x-1)(x-5)$ e portanto 1 e 5 são autovalores de T.*

Exemplo 0.164. *Considere* $\mathbb{R}^2$ *como espaço vetorial sobre* $\mathbb{R}$. *Seja* $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ *dado por* $T(x, y) = (-y, x)$. *Já vimos que* T *não é diagonalizável. Temos que* $[T]_{\mathcal{B}_C} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. *Daí,* $p_T(x) = \det\left(\begin{smallmatrix} x & 1 \\ -1 & x \end{smallmatrix}\right) = x^2 + 1$ *que não possui raízes reais.*

Exemplo 0.165. *Considerando* $\mathbb{C}^2$ *como* $\mathbb{C}$ *espaço vetorial e* $S \in \mathcal{L}(\mathbb{C}^2, \mathbb{C}^2)$ *dado por* $S(x, y) = (-y, x)$, *então* $p_S(x) = x^2 + 1 = (x - i)(x + i)$ *e portanto* i *e* $-i$ *são autovalores de* S. *Ainda, existe uma base* $\mathcal{B}$ *de* $\mathbb{C}^2$ *(determine* $\mathcal{B}$*) tal que* $[S]_{\mathcal{B}} = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right) \in \mathbb{M}_{2\times 2}(\mathbb{C})$.

Exemplo 0.166. *Considere* $\mathbb{R}^3$ *como espaço vetorial sobre* $\mathbb{R}$. *Seja* $T \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^3)$ *tal que*

$$[T]_{\mathcal{C}} = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix},$$

para alguma base $\mathcal{C}$ *de* $\mathbb{R}^3$. *Determine os autovalores e os autovetores de* T. *Bem,*

$$[xId - T]_{\mathcal{C}} = \begin{pmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{pmatrix} \ e$$

$$\begin{aligned} p_T(x) &= \det([xId - T]_{\mathcal{C}}) \\ &= (x-3)(x-2)x + 2 + 4 + 2(x-2) + 2(x-3) - 2x \\ &= (x-3)(x-2)x + 2 + 4 + 2x - 4 + 2x - 6 - 2x \\ &= (x-3)(x-2)x + 2(x-2) \\ &= ((x-3)x + 2)(x-2) \\ &= (x^2 - 3x + 2)(x-2) \end{aligned}$$

e daí 1 *e* 2 *são autovalores de* T *(pois são as raízes de* $p_T(x)$*). Consideremos as matrizes*

$$[1Id - T]_{\mathcal{C}} = \begin{pmatrix} 1-3 & -1 & 1 \\ -2 & 1-2 & 1 \\ -2 & -2 & 1 \end{pmatrix} \text{ e } [2Id - T]_{\mathcal{C}} = \begin{pmatrix} 2-3 & -1 & 1 \\ -2 & 2-2 & 1 \\ -2 & -2 & 2 \end{pmatrix}.$$

Temos que $\mathrm{Aut}_T(1) = \mathrm{Ker}(1Id - T)$*, donde os autovetores de* T *associados a* 1 *são os vetores* $v \in \mathbb{R}^3 \setminus \{\vec{0}\}$ *tais que* $(1Id - T)v = \vec{0}$*, isto é,*

$$\begin{pmatrix} -2 & -1 & 1 \\ -2 & -1 & 1 \\ -2 & -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Notemos que, na verdade, (x, y, z) *são as coordenadas de* v *na base* $\mathcal{C}$ *(entretanto trataremos os dois termos como sendo a mesma coisa quando não houver mais de uma base envolvida). Assim, os autovetores de* T *associados a* 1 *são os vetores da forma* $(x, 0, 2x)$*, isto é,* $\mathrm{Aut}_T(1) = [(1, 0, 2)]$*. Da mesma forma,* $\mathrm{Aut}_T(2) = \mathrm{Ker}(2Id - T)$ *e resolvendo*

$$\begin{pmatrix} -1 & -1 & 1 \\ -2 & 0 & 1 \\ -2 & -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

obtemos que os autovetores de T *associados a* 2 *são os vetores da forma* $(x, x, 2x)$*, ou seja,* $\mathrm{Aut}_T(2) = [(1, 1, 2)]$*. Como* $\{(1, 0, 2), (1, 1, 2)\}$ *não é uma base de* $\mathbb{R}^2$ *(pois não gera), segue que* T *não é diagonalizável.*

Exemplo 0.167. *Considere* $\mathbb{R}^3$ *como espaço vetorial sobre* $\mathbb{R}$*. Seja* $S \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^3)$ *tal que*

$$[S]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 3 \end{pmatrix},$$

para alguma base $\mathcal{B}$ *de* $\mathbb{R}^3$*. Determine os autovalores e os autovetores de* S*. Bem,*

$$[xId - S]_{\mathcal{B}} = \begin{pmatrix} x-1 & 0 & 0 \\ 1 & x-2 & -1 \\ 2 & 0 & x-3 \end{pmatrix} \text{ e}$$

$$\begin{aligned} p_S(x) &= \det([xId - S]_{\mathcal{B}}) \\ &= (x-1)(x-2)(x-3) \end{aligned}$$

e daí 1, 2 *e* 3 *são autovalores de* S *(pois são as raízes de* $p_S(x)$*). Então* $\mathrm{Aut}_S(1) = [(1,0,1)]$, $\mathrm{Aut}_S(2) = [(0,1,0)]$ *e* $\mathrm{Aut}_S(3) = [(0,1,1)]$*. Como* $\{(1,0,1),(0,1,0),(0,1,1)\}$ *é uma base de* $\mathbb{R}^3$*, então* S *é diagonalizável.*

O próximo resultado nos dá uma forma de obter uma base formada por autovetores de um certo operador linear.

Proposição 0.168. *Sejam* V *um espaço vetorial sobre* $\mathbb{K}$ *de dimensão finita e* $T \in \mathcal{L}(V,V)$*. Sejam* $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ *autovalores de* T *distintos entre si.*

1. *Se* $v_1 + v_2 + \cdots + v_k = \vec{0}$*, com* $v_i \in \mathrm{Aut}_T(\lambda_i)$*,* $i = 1, 2, \ldots, k$*, então* $v_1 = v_2 = \cdots = v_k = \vec{0}$*.*

2. *Se* $\mathcal{B}_i$ *é um subconjunto LI de* $\mathrm{Aut}_T(\lambda_i)$*, então* $\bigcup_{i=1}^{k} \mathcal{B}_i$ *é um subconjunto LI de* V*.*

De fato, mostremos (1). Façamos por indução em k. Se $k = 1$, acabou. Se $k > 1$, suponhamos que o resultado vale para $k-1$ (HI) e mostremos que vale para k. Sejam $v_i \in \mathrm{Aut}_T(\lambda_i)$, $i = 1, 2, \ldots, k$, tais que $v_1 + v_2 + \cdots + v_k = \vec{0}$. Daí, $T(v_1 + v_2 + \cdots + v_k) = \vec{0}$ e portanto, como T é linear, $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = \vec{0}$ (*). Ainda, $\lambda_k(v_1 + v_2 + \cdots + v_k) = \vec{0}$, isto é, $\lambda_k v_1 + \lambda_k v_2 + \cdots + \lambda_k v_k = \vec{0}$ (**). Fazendo (*) menos (**), obtemos que $(\lambda_1 - \lambda_k)v_1 + (\lambda_2 - \lambda_k)v_2 + \cdots + (\lambda_{k-1} - \lambda_k)v_{k-1} = \vec{0}$. Notemos que $(\lambda_j - \lambda_k)v_j \in \mathrm{Aut}_T(\lambda_j)$. Então, por (HI), segue que $(\lambda_j - \lambda_k)v_j = 0$, para todo $j = 1, 2, \ldots, k-1$. Como $\lambda_j - \lambda_k \neq 0$, para todo $j \neq k$ (pois os autovalores são distintos entre si), então $v_1 = v_2 = \cdots = v_{k-1} = \vec{0}$. Daí, a partir

da igualdade inicial, obtemos também que $v_k = \vec{0}$ o que encerra a primeira demonstração.
Mostremos agora (2). Seja $\mathcal{B}_i := \{v_{1i}, \ldots, v_{n_i i}\}$ um subconjunto LI de $\mathrm{Aut}_T(\lambda_i)$. Seja $c := \sum_{j=1}^{n_1} \alpha_{j1} v_{j1} + \sum_{j=1}^{n_2} \alpha_{j2} v_{j2} + \cdots + \sum_{j=1}^{n_k} \alpha_{jk} v_{jk}$ uma combinação linear qualquer de elementos de $\bigcup_{i=1}^{k} \mathcal{B}_i$. Suponhamos que c é o vetor nulo. Daí, como $\sum_{j=1}^{n_i} \alpha_{ji} v_{ji} \in \mathrm{Aut}_T(\lambda_i)$ (pois este é subespaço), segue pela primeira parte que $\sum_{j=1}^{n_i} \alpha_{ji} v_{ji} = \vec{0}$ e portanto, como $\mathcal{B}_i$ é LI, obtemos que $\alpha_{1i} = \alpha_{2i} = \cdots = \alpha_{n_i i} = 0$, para todo $i = 1, 2, \ldots, k$. Logo, $\bigcup_{i=1}^{k} \mathcal{B}_i$ é LI.
■

Observação 0.169. *Com as mesmas notações, este resultado nos diz, em particular, que se* $v_1 \in \mathrm{Aut}_T(\lambda_1) \setminus \{\vec{0}\}$ *e* $v_2 \in \mathrm{Aut}_T(\lambda_2) \setminus \{\vec{0}\}$, *com* $\lambda_1 \neq \lambda_2$, *então* $\{v_1, v_2\}$ *é LI.*

Corolário 0.170. *Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathcal{L}(V, V)$. Sejam $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ todos os autovalores de T (distintos entre si). Então T é diagonalizável se, e somente se,* $\dim_{\mathbb{K}} V = \sum_{i=1}^{k} \dim_{\mathbb{K}} \mathrm{Aut}_T(\lambda_i)$.

Seja $\mathcal{B}_i$ uma base de $\mathrm{Aut}_T(\lambda_i)$, com n_i vetores. Então, pelo resultado anterior, $\bigcup_{i=1}^{k} \mathcal{B}_i$ é um subconjunto LI de V, com $n_1 + n_2 + \cdots + n_k$ vetores. Daí, $W := [\bigcup_{i=1}^{k} \mathcal{B}_i]$ é um subespaço vetorial de V, com $\dim_{\mathbb{K}} W = \sum_{i=1}^{k} \dim_{\mathbb{K}} \mathrm{Aut}_T(\lambda_i)$. Notemos que W é o subespaço gerado por todos os autovetores de T, uma vez que $\lambda_1, \lambda_2, \ldots, \lambda_k$ são todos os autovalores de T. Ainda, temos que $\dim_{\mathbb{K}} W \leqslant \dim_{\mathbb{K}} V$, pois W é subespaço vetorial. Por um lado, se $\dim_{\mathbb{K}} W = \dim_{\mathbb{K}} V$, então $W = V$ e portanto $\bigcup_{i=1}^{k} \mathcal{B}_i$ é uma base de V formada por autovetores de T, donde T é diagonalizável. Por outro lado, se T é diagonalizável, então V admite uma base, $\mathcal{C}$, formada por autovetores de T. Desde que os elementos de $\mathcal{C}$ são autovetores de T, então $\mathcal{C} \subset W$, donde $\dim_{\mathbb{K}} V \leqslant \dim_{\mathbb{K}} W$ e portanto $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$. ■

Definição 0.171. *Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita, $T \in \mathcal{L}(V, V)$ e $\lambda \in \mathbb{K}$ um autovalor de T. Se $p_T(x) = (x - \lambda)^m q(x)$, onde $q(\lambda) \neq 0$, dizemos que $m \in \mathbb{N}$ é a* **multiplicidade algébrica de** *λ e a denotamos por $ma(\lambda)$. Ainda, dizemos que $\dim_{\mathbb{K}} \operatorname{Aut}_T(\lambda) \in \mathbb{N}$ é a* **multiplicidade geométrica de** *λ e a denotamos por $mg(\lambda)$.*

Observação 0.172. *Com as mesmas notações, observemos que $ma(\lambda)$ é o maior inteiro j tal que $(x - \lambda)^j$ divide $p_T(x)$.*

Proposição 0.173. *Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita, $T \in \mathcal{L}(V, V)$ e $\lambda \in \mathbb{K}$ um autovalor de T. Então $mg(\lambda) \leqslant ma(\lambda)$.*

Seja $W := \operatorname{Aut}_T(\lambda)$ e seja $\mathcal{B} := \{u_1, u_2, \ldots, u_m\}$ uma base de W (em particular, notemos que $mg(\lambda) = m$). Como $\mathcal{B}$ é LI, existe uma base de V, $\mathcal{C} := \{u_1, u_2, \ldots, u_m, v_1, v_2, \ldots, v_n\}$, contendo $\mathcal{B}$. Notemos que $T(u_i) = \lambda u_i$, donde $T(u_i) = 0u_1 + \cdots + 0u_{i-1} + \lambda u_i + 0u_{i+1} + \cdots + 0u_m + 0v_1 + 0v_2 + \cdots + 0v_n$, para todo $i = 1, \ldots, m$. Daí,

$$[T]_{\mathcal{C}} = \begin{pmatrix} \lambda I_m & A \\ \vec{0} & B \end{pmatrix},$$

para alguma matriz $A \in \mathbb{M}_{m \times n}(\mathbb{K})$, $B \in \mathbb{M}_{n \times n}(\mathbb{K})$, onde I_m é a matriz identidade $m \times m$ e $\vec{0}$ é a matriz nula $n \times m$. Por uma propriedade de determinante, $\det(xI_{m+n} - [T]_{\mathcal{C}}) = \det(xI_m - \lambda I_m).\det(xI_n - B) = (x - \lambda)^m.q(x)$, com $q(x) := \det(xI_n - B)$. Daí, $p_T(x) = (x - \lambda)^m.q(x)$, donde já sabemos que $ma(\lambda)$ é maior ou igual a m (pode ser que λ também seja raiz de $q(x)$). Logo, $mg(\lambda) = m \leqslant ma(\lambda)$. ■

Observação 0.174. *Sejam V um espaço vetorial sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} V = n > 0$, e $T \in \mathcal{L}(V, V)$. Sejam $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ autovalores de T (distintos entre si). Observemos que, se $ma(\lambda_i) = n_i$, $i = 1, \ldots, k$, então $p_T(x) = (x - \lambda_1)^{n_1}.(x - \lambda_2)^{n_2}. \cdots . (x - \lambda_k)^{n_k}.q(x)$, para algum $q(x) \in \mathcal{P}(\mathbb{K})$, pois $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ são raízes de $p_T(x)$. Ainda, desde que o grau de $p_T(x)$*

é $n = \dim_{\mathbb{K}} V$, temos que $n \geqslant \sum_{i=1}^{k} n_i$. Notemos ainda que $q(x)$ é tal que $q(\lambda_i) \neq 0$, pois $n_i = ma(\lambda_i)$, para todo $i = 1, \ldots, k$.

Teorema 0.175. *Sejam V um espaço vetorial sobre $\mathbb{K}$, com $\dim_{\mathbb{K}} V = n > 0$, e $T \in \mathcal{L}(V, V)$. Sejam $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ todos os autovalores de T (distintos entre si). Então as condições abaixo são equivalentes:*

1. *T é diagonalizável*

2. $\dim_{\mathbb{K}} V = \sum_{i=1}^{k} \dim_{\mathbb{K}} \mathrm{Aut}_T(\lambda_i)$

3. *$p_T(x) = (x - \lambda_1)^{n_1}.(x - \lambda_2)^{n_2}. \cdots .(x - \lambda_k)^{n_k}$, onde $n_i = ma(\lambda_i) = mg(\lambda_i)$, $i = 1, \ldots, k$*

De fato, já mostramos ($1 \Leftrightarrow 2$). Mostremos então ($2 \Leftrightarrow 3$). Se $p_T(x) = (x - \lambda_1)^{n_1}.(x - \lambda_2)^{n_2}. \cdots .(x - \lambda_k)^{n_k}$, onde $n_i = ma(\lambda_i) = mg(\lambda_i)$, então o grau de $p_T(x)$ é igual a $\sum_{i=1}^{k} n_i$. Mas, por definição, o grau de $p_T(x)$ é igual a $\dim_{\mathbb{K}} V$. Logo, $\dim_{\mathbb{K}} V = \sum_{i=1}^{k} n_i = \sum_{i=1}^{k} mg(\lambda_i) = \sum_{i=1}^{k} \dim_{\mathbb{K}} \mathrm{Aut}_T(\lambda_i)$. Suponhamos agora que $\dim_{\mathbb{K}} V = \sum_{i=1}^{k} \dim_{\mathbb{K}} \mathrm{Aut}_T(\lambda_i)$. A princípio, se $ma(\lambda_i) = n_i$, $i = 1, \ldots, k$, sabemos que $p_T(x) = (x - \lambda_1)^{n_1}.(x - \lambda_2)^{n_2}. \cdots . (x - \lambda_k)^{n_k}.q(x)$, para algum $q(x) \in \mathcal{P}(\mathbb{K})$. Então, pelo resultado anterior, $\dim_{\mathbb{K}} V = \sum_{i=1}^{k} \dim_{\mathbb{K}} \mathrm{Aut}_T(\lambda_i) = \sum_{i=1}^{k} mg(\lambda_i) \leqslant \sum_{i=1}^{k} ma(\lambda_i) = \sum_{i=1}^{k} n_i \leqslant \dim_{\mathbb{K}} V$, pois o grau de $p_T(x)$ é igual a $\dim_{\mathbb{K}} V$. Logo, $q(x) = 1$, pois $\dim_{\mathbb{K}} V = \sum_{i=1}^{k} mg(\lambda_i) = \sum_{i=1}^{k} ma(\lambda_i)$. Agora, sabemos que $mg(\lambda_i) \leqslant ma(\lambda_i)$, com $i = 1, \ldots, n$. Suponhamos, por absurdo, que $mg(\lambda_j) < ma(\lambda_j)$, para algum $j = 1, \ldots, k$. Daí, $\sum_{i=1}^{k} mg(\lambda_i) < \sum_{i=1}^{k} ma(\lambda_i)$, o que não ocorre. Logo, $mg(\lambda_i) = ma(\lambda_i)$, para todo $i = 1, \ldots, k$ e portanto $p_T(x) = (x-\lambda_1)^{n_1}.(x-\lambda_2)^{n_2}. \cdots .(x-\lambda_k)^{n_k}$, onde $n_i = ma(\lambda_i) = mg(\lambda_i)$, $i = 1, \ldots, n$.
■

Observação 0.176. *Sejam V um espaço vetorial sobre $\mathbb{K}$ e $T \in \mathcal{L}(V, V)$. Notemos que se $\lambda \in \mathbb{K}$ é um autovalor de T, segue que $\mathrm{Aut}_T(\lambda) \neq \{\vec{0}\}$,* *donde* $\dim_{\mathbb{K}} \operatorname{Aut}_T(\lambda) \geqslant 1$ *(ou seja, não é* 0*). Logo, se* $ma(\lambda) = 1$*, então* $1 = ma(\lambda) \geqslant mg(\lambda) = \dim_{\mathbb{K}} \operatorname{Aut}_T(\lambda) \geqslant 1$*, isto é,* $1 = ma(\lambda) = mg(\lambda)$*.*

Exemplo 0.177. *Seja* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *um operador linear tal que* $p_T(x) = (x-1)(x-2)(x-3)$*. Como* $ma(1) = 1$*, segue, da observação acima, que* $1 = ma(1) = mg(1)$*. Analogamente,* $1 = ma(2) = mg(2)$ *e* $1 = ma(3) = mg(3)$*. Logo, pelo teorema acima,* T *é diagonalizável. Daí, existe uma base* $\mathcal{B}$ *de* $\mathbb{R}^3$ *tal que*

$$[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Exemplo 0.178. *Seja* $S : \mathbb{R}^3 \to \mathbb{R}^3$ *um operador linear dado por*

$$[S]_{\mathcal{C}} = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix},$$

para alguma base $\mathcal{C}$ *de* $\mathbb{R}^3$*. Já sabemos que* $p_S(x) = (x-1)(x-2)^2$*. Como* $\operatorname{Aut}_S(2) = [(1,1,2)]$*, então* $\{(1,1,2)\}$ *é uma base de* $\operatorname{Aut}_S(2)$ *e portanto* $mg(2) = \dim_{\mathbb{K}} \operatorname{Aut}_S(2) = 1 \neq 2 = ma(2)$*. Logo,* S *não é diagonalizável. Notemos apenas que, como* $ma(1) = 1$*, segue, da observação acima, que* $1 = ma(1) = mg(1)$*.*

Exemplo 0.179. *Seja* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *um operador linear dado por*

$$[T]_{\mathcal{C}} = \begin{pmatrix} 1 & 2 & -1 \\ -2 & -3 & 1 \\ 2 & 2 & -2 \end{pmatrix},$$

para alguma base $\mathcal{C}$ *de* $\mathbb{R}^3$*.* T *é diagonalizável? Bem,*

$$[xId - T]_{\mathcal{C}} = \begin{pmatrix} x-1 & -2 & 1 \\ 2 & x+3 & -1 \\ -2 & -2 & x+2 \end{pmatrix} \text{ e}$$

$$\begin{aligned} p_T(x) &= \det([xId - T]_{\mathcal{C}}) \\ &= (x-1)(x+3)(x+2) - 4 - 4 + 2(x+3) - 2(x-1) + 4(x+2) \\ &= (x-1)(x+3)(x+2) - 8 + 2x + 6 - 2x + 2 + 4(x+2) \\ &= (x-1)(x+3)(x+2) + 4(x+2) \\ &= ((x-1)(x+3) + 4)(x+2) \\ &= (x+1)^2(x+2) \end{aligned}$$

e daí 1 *e* 2 *são autovalores de* T *(pois são as raízes de* $p_T(x)$*). Da observação acima, como* $ma(-2) = 1$*, já temos que* $1 = ma(-2) = mg(-2)$*. Calculemos então* $\mathrm{Aut}_T(-1)$*. Consideremos a matriz*

$$[-1Id - T]_{\mathcal{C}} = \begin{pmatrix} -1-1 & -2 & 1 \\ 2 & -1+3 & -1 \\ -2 & -2 & -1+2 \end{pmatrix}.$$

Temos que $\mathrm{Aut}_T(-1) = \mathrm{Ker}(-1Id - T)$*, donde os autovetores de* T *associados a* (-1) *são os vetores tais que*

$$\begin{pmatrix} -2 & -2 & 1 \\ 2 & 2 & -1 \\ -2 & -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

isto é, são os vetores da forma $(x, y, 2x+2y)$*. Ou seja, temos que* $\mathrm{Aut}_T(-1) = \{(x, y, 2x+2y) | x, y \in \mathbb{R}\} = \{x(1,0,2) + y(0,1,2) | x, y \in \mathbb{R}\} = [(1,0,2), (0,1,2)]$*. Como* $\{(1,0,2), (0,1,2)\}$ *é LI, pois um não é múltiplo do outro, então é uma base de* $\mathrm{Aut}_T(-1)$*. Em particular,* $\dim_{\mathbb{K}} \mathrm{Aut}_T(-1) = 2$*. Portanto, como* $p_T(x) = (x+1)^2(x+2)$*, com* $2 = ma(-1) = mg(-1)$ *e* $1 = ma(-2) = mg(-2)$*, segue que* T *é diagonalizável.*

Polinômio minimal

Observação 0.180. *Sejam V um espaço vetorial sobre $\mathbb{K}$, $T \in \mathcal{L}(V,V)$ e $p(x) \in \mathcal{P}(\mathbb{K})$. Então podemos considerar o operador linear $p(T) \in \mathcal{L}(V,V)$. Lembremos que se $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, então $p(T) = a_n T^n + +a_{n-1}T^{n-1} + \cdots + a_1 T + a_0 Id$, onde*

$$T^k = \underbrace{T \circ T \circ \cdots \circ T}_{k}.$$

Agora, dizemos que **T é raiz de $p(x)$**, *se $p(T) = \vec{0}_{\mathcal{L}(V,V)}$, isto é, se $p(T)$ for o operador identicamente nulo. Lembremos que se $\dim_{\mathbb{K}} V = n$, então $\dim_{\mathbb{K}} \mathcal{L}(V,V) = n^2$. Daí, se $m \geqslant n^2 = \dim_{\mathbb{K}} \mathcal{L}(V,V)$, temos que o conjunto, $\{Id, T, T^2, \ldots, T^m\}$, de vetores de $\mathcal{L}(V,V)$ é LD, donde existem escalares $a_0, a_1, a_2, \ldots, a_m$ não todos nulos tais que $a_m T^m + a_{m-1}T^{m-1} + \cdots + a_1 T + a_0 Id = \vec{0}_{\mathcal{L}(V,V)}$. Em particular, T é raiz de $q(x)$, com $q(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$.*

Exemplo 0.181. *Seja $T \in \mathcal{L}(\mathbb{R},\mathbb{R})$. Como $\dim_{\mathbb{R}} \mathcal{L}(\mathbb{R},\mathbb{R}) = 1^2 = 1$, então o conjunto $\{Id, T\}$ é LD. De fato, sabemos que $T = T_\beta$, para algum $\beta \in \mathbb{R}$, isto é, T é dada por $T(x) = \beta x$, para algum $\beta \in \mathbb{R}$. Se $\beta = 0$, então $T = \vec{0}_{\mathcal{L}(\mathbb{R},\mathbb{R})}$ e o conjunto é LD. Suponhamos então $\beta \neq 0$. Seja $c := aId + bT$ uma combinação linear qualquer de Id, T. Suponhamos que c é o vetor nulo. Assim, $aId + bT = \vec{0}_{\mathcal{L}(\mathbb{R},\mathbb{R})}$, donde em particular $(aId + bT)(1) = 0$ (isto é, $a + b\beta = 0$). Portanto, $(-\beta)Id + T = \vec{0}_{\mathcal{L}(\mathbb{R},\mathbb{R})}$ é uma combinação linear não trivial de Id, T resultando no vetor nulo. Logo, o conjunto é LD.*

Observação 0.182. *Mas pode ser que $\{Id, T, T^2, \ldots, T^k\}$ já seja LD, para algum $k < n^2$. Consideremos $k > 0$ tal que $\{Id, T, T^2, \ldots, T^{k-1}\}$ é LI, mas $\{Id, T, T^2, \ldots, T^k\}$ é LD. Mas, então, T^k deve ser combinação linear de $Id, T, T^2, \ldots, T^{k-1}$, ou seja, existem escalares $a_0, a_1, a_2, \ldots, a_{k-1}$ tais que $T^k = a_{k-1}T^{k-1} + \cdots + a_1 T + a_0 Id$. Daí, T é raiz de $m_T(x)$, onde $m_T(x) =$*

$x^k - a_{k-1}x^{k-1} - \cdots - a_1x - a_0$. Notemos $m_T(x)$ é um **polinômio mônico**, *pois $a_k := 1$.*

Exemplo 0.183. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ um operador linear dado por $T(x,y) = (x+y, x-y)$. Temos que $T^2(x,y) = T(T(x,y)) = T(x+y, x-y) = (x+y+(x-y), x+y-(x-y)) = (2x, 2y) = 2Id(x,y)$, donde $T^2 = 2Id$. Já sabemos que $\{Id\}$ é LI. Verifiquemos que $\{Id, T\}$ é LI. De fato, se $\alpha Id + \beta T = \vec{0}_{\mathcal{L}(\mathbb{R}^2,\mathbb{R}^2)}$, então $\alpha(x,y) + \beta(x+y, x-y) = (0,0)$, para todo $(x,y) \in \mathbb{R}^2$, donde $(\alpha+\beta)x + \beta y = 0$ e $\beta x + (\alpha - \beta)y = 0$. Se $x = 0$ e $y = 1$, obtemos que $\beta = 0 = \alpha$, donde $\{Id, T\}$ é LI. Agora, desde que $T^2 = 2Id$, obtemos que $\{Id, T, T^2\}$ é LD. Então, pela construção de $m_T(x)$, como $T^2 - 2Id = \vec{0}_{\mathcal{L}(\mathbb{R}^2,\mathbb{R}^2)}$, segue que $m_T(x) = x^2 - 2$.*

Proposição 0.184. *Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathcal{L}(V,V)$. Se $p(x) \in \mathcal{P}(\mathbb{K})$ é tal que $p(T) = \vec{0}_{\mathcal{L}(V,V)}$, então $p(x)$ é um múltiplo de $m_T(x)$.*

De fato, seja $p(x) \in \mathcal{P}(\mathbb{K})$ tal que $p(T) = \vec{0}_{\mathcal{L}(V,V)}$. Consideremos a divisão de $p(x)$ por $m_T(x)$. Daí, existem $q(x), r(x) \in \mathcal{P}(\mathbb{K})$ tais que $p(x) = m_T(x)q(x) + r(x)$, com ou $r(x) = \vec{0}_{\mathcal{P}(\mathbb{K})}$, ou o grau de $r(x)$ menor do que o grau de $m_T(x)$. Daí, $\vec{0}_{\mathcal{L}(V,V)} = P(T) = m_T(T) \circ q(T) + r(T) = \vec{0}_{\mathcal{L}(V,V)} \circ q(T) + r(T) = \vec{0}_{\mathcal{L}(V,V)} + r(T) = r(T)$. Como $m_T(x)$ é o polinômio mônico de menor grau tal que T é raiz, e $r(T) = \vec{0}_{\mathcal{L}(V,V)}$, segue que $r(x) = \vec{0}_{\mathcal{P}(\mathbb{K})}$, isto é, $r(x)$ é o polinômio identicamente nulo. Logo, $p(x) = m_T(x)q(x)$ e portanto $p(x)$ é um múltiplo de $m_T(x)$. ■

Definição 0.185. *Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathcal{L}(V,V)$. Chamamos o polinômio $m_T(x) \in \mathcal{P}(\mathbb{K})$ (que é mônico, de menor grau tal que $m_T(T) = \vec{0}_{\mathcal{L}(V,V)}$) de* **polinômio minimal de T**.

Exemplo 0.186. *Seja $S : \mathbb{R}^3 \to \mathbb{R}^3$ um operador linear dado por*

$$[S]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

para alguma base $\mathcal{B}$ de $\mathbb{R}^3$. S é diagonalizável? Bem,

$$[xId - S]_{\mathcal{B}} = \begin{pmatrix} x & 0 & 0 \\ -1 & x & 0 \\ 0 & -1 & x \end{pmatrix}$$

e portanto $p_S(x) = \det([xId - S]_{\mathcal{B}}) = x^3$, donde 0 é o único autovalor de S (pois é a única raiz de $p_T(x)$, a menos de multiplicade). É fácil ver que $\mathrm{Aut}_S(0) = [(0,0,1)]$, donde S não é diagonalizável (pois $mg(0) = \dim_{\mathbb{K}} \mathrm{Aut}_T(0) = 1 \neq 3 = ma(0)$.
Observemos agora que $[S]_{\mathcal{B}} \neq \overrightarrow{0}$, $([S]_{\mathcal{B}})^2 \neq \overrightarrow{0}$, mas $([S]_{\mathcal{B}})^3 = \overrightarrow{0}_{\mathbb{M}_{3\times 3}(\mathbb{R})}$, donde $S \neq \overrightarrow{0}$, $S^2 \neq \overrightarrow{0}$, mas $S^3 = \overrightarrow{0}_{\mathcal{L}(\mathbb{R}^3,\mathbb{R}^3)}$. Então S é raiz de x^3 e portanto x^3 é um múltiplo de $m_S(x)$. Como os únicos divisores de x^3 são x, x^2 e x^3, e S não anula os dois primeiros segue que $m_S(x) = x^3$.

Exemplo 0.187. *Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ um operador linear dado por $T(x,y,z) = (z,z,z)$. Mostre que T é diagonalizável. De fato,*

$$[T]_{\mathcal{B}_C} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

e portanto $p_T(x) = \det([xId - T]_{\mathcal{B}_C}) = x^2(x-1)$, donde 0 e 1 são autovalores de T. Já sabemos que $ma(1) = mg(1) = 1$. Como $\mathrm{Aut}_T(0) = [(1,0,0),(0,1,0)]$ (verifique), então $\{(1,0,0),(0,1,0)\}$ é uma base de $\mathrm{Aut}_T(0)$, donde $mg(0) = \dim_{\mathbb{K}} \mathrm{Aut}_T(0) = 2 = ma(0)$. Logo, T é diagonalizável.
Agora, notemos que

$$[T]_{\mathcal{B}_C}.([T]_{\mathcal{B}_C} - I_3) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

donde, obtemos que $T(T-I) = \vec{0}_{\mathcal{L}(\mathbb{R}^3,\mathbb{R}^3)}$, *isto é,* T *é raiz de* $x(x-1)$ *e portanto* $x(x-1)$ *é um múltiplo de* $m_T(x)$. *Desde que os únicos divisores de* $x(x-1)$ *são* x, $x-1$ *e* $x(x-1)$, *e* T *não anula os dois primeiros segue que* $m_T(x) = x(x-1)$.

Observação 0.188. *Notemos que, nos exemplos anteriores, verificamos que as raízes do polinômio minimal do operador são os autovalores do operador.*

Vejamos agora, sem demonstração, o teorema de Cayley-Hamilton.

Teorema 0.189. *Sejam* V *um espaço vetorial sobre* $\mathbb{K}$ *de dimensão finita e* $T \in \mathcal{L}(V,V)$. *Então* T *é raiz de* $p_T(x)$. ■

Observação 0.190. *Assim, pelos resultados anteriores, obtemos que* $p_T(x)$ *é um múltiplo de* $m_T(x)$.

Proposição 0.191. *Sejam* V *um espaço vetorial sobre* $\mathbb{K}$ *de dimensão finita e* $T \in \mathcal{L}(V,V)$. *Então* $p_T(x)$ *e* $m_T(x)$ *têm as mesmas raízes (a menos de multiplicidade).*

De fato, suponhamos que $\lambda \in \mathbb{K}$ é uma raiz de $p_T(x)$. Então λ é um autovalor de T e portanto existe $v_0 \neq \vec{0}_V$ tal que $T(v_0) = \lambda v_0$. Daí, $T^2(v_0) = \lambda^2 v_0, T^3(v_0) = \lambda^3 v_0, \ldots, T^j(v_0) = \lambda^j v_0$, para todo $j > 0$. Seja $m_T(x) = x^k + a_{k-1}x^{k-1} \cdots + a_1 x + a_0$, para algum $a_0, \ldots, a_k \in \mathbb{K}$. Então $\vec{0}_{\mathcal{L}(V,V)} = m_T(T) = T^k + a_{k-1}T^{k-1} \cdots + a_1 T + a_0 Id$, donde

$$\begin{aligned}
\vec{0}_V = m_T(T)(v_0) &= T^k(v_0) + a_{k-1}T^{k-1}(v_0) + \cdots + a_1 T(v_0) + a_0 Id(v_0) \\
&= \lambda^k v_0 + a_{k-1}\lambda^{k-1} v_0 + \cdots + a_1 \lambda v_0 + a_0 v_0 \\
&= (\lambda^k + a_{k-1}\lambda^{k-1} + \cdots + a_1 \lambda + a_0) v_0
\end{aligned}$$

e portanto, como $v_0 \neq \vec{0}_V$, então $\lambda^k + a_{k-1}\lambda^{k-1} + \cdots + a_1\lambda + a_0 = 0$, isto é, λ é uma raiz de $m_T(x)$.
Suponhamos agora que $\lambda \in \mathbb{K}$ é uma raiz de $m_T(x)$. Então $m_T(x) = (x -$

$\lambda)q(x)$, para algum $q(x) \in \mathcal{P}(\mathbb{K})$. Daí, notemos que $\vec{0}_{\mathcal{L}(V,V)} = m_T(T) = (T - \lambda Id) \circ q(T)$. Agora, como o grau de $q(x)$ é menor do que o grau de $m_T(x)$, então $q(T) \neq \vec{0}_{\mathcal{L}(V,V)}$. Daí, existe $u_0 \in V$ tal que $q(T)(u_0) \neq \vec{0}_V$. Denotemos $v_0 := q(T)(u_0)$. Logo, $\vec{0}_V = m_T(T)(u_0) = ((T - \lambda Id) \circ q(T))(u_0) = (T - \lambda Id)(q(T)(u_0)) = T(v_0) - \lambda v_0$, isto é, $T(v_0) = \lambda v_0$, com $v_0 \neq \vec{0}_V$, donde λ é um autovalor de T. Daí, λ é uma raiz de $p_T(x)$. ■

Proposição 0.192. *Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathcal{L}(V,V)$. Sejam $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ todos os autovalores de T (distintos entre si). Se T é diagonalizável, então $m_T(x) = (x - \lambda_k).(x - \lambda_{k-1}). \cdots .(x - \lambda_1)$.*

De fato, seja $q(x) := (x - \lambda_k).(x - \lambda_{k-1}). \cdots .(x - \lambda_1)$. Desde que $q(x)$ é o polinômio mônico de menor grau que tem as mesmas raízes do que $p_T(x)$, então se mostrarmos que $q(T) = \vec{0}_{\mathcal{L}(V,V)}$, obteremos que $m_T(x) = q(x)$. Façamos isto. Seja $\mathcal{B}$ uma base de V formada por autovetores de T. Se $v \in \mathcal{B}$, então existe um autovalor λ_i tal que $T(v) = \lambda_i v$. Daí, $(T - \lambda_i Id)(v) = \vec{0}_V$. Logo, se $q(T) = (T - \lambda_k Id) \circ (T - \lambda_{k-1} Id) \circ \cdots \circ (T - \lambda_1 Id)$, então $q(T)(v) = \vec{0}_V$, para todo $v \in \mathcal{B}$. Portanto, $q(T)(v) = \vec{0}_V$, para todo $v \in V$, donde $q(T) = \vec{0}_{\mathcal{L}(V,V)}$.

■

Observação 0.193. *Na demonstração acima, suponha que $\dim_{\mathbb{K}} V = 3$ e $\mathcal{B} := \{v_1, \tilde{v_1}, v_2\}$, onde $v_1, \tilde{v_1} \in \mathrm{Aut}_T(\lambda_1)$ e $v_2 \in \mathrm{Aut}_T(\lambda_2)$. Então um vetor v de V se escreve de forma única, digamos $v = \alpha_1 v_1 + \tilde{\alpha_1}\tilde{v_1} + \alpha_2 v_2$. Assim,*

$$\begin{array}{rcl}
\big((T - \lambda_2 Id) \circ (T - \lambda_1 Id)\big)(v) & = & \\
(T - \lambda_2 Id)\big(T(v) - \lambda_1 v\big) & = & \\
(T - \lambda_2 Id)\big(T(\alpha_1 v_1 + \tilde{\alpha_1}\tilde{v_1} + \alpha_2 v_2) - \lambda_1(\alpha_1 v_1 + \tilde{\alpha_1}\tilde{v_1} + \alpha_2 v_2)\big) & \overset{\text{linear}}{=} & \\
(T - \lambda_2 Id)\big(\alpha_1 T(v_1) + \tilde{\alpha_1} T(\tilde{v_1}) + \alpha_2 T(v_2) - \lambda_1(\alpha_1 v_1 + \tilde{\alpha_1}\tilde{v_1} + \alpha_2 v_2)\big) & \overset{\text{autovet.}}{=} & \\
(T - \lambda_2 Id)\big(\alpha_1 \lambda_1 v_1 + \tilde{\alpha_1}\lambda_1\tilde{v_1} + \alpha_2 \lambda_2 v_2 - \lambda_1(\alpha_1 v_1 + \tilde{\alpha_1}\tilde{v_1} + \alpha_2 v_2)\big) & = & \\
(T - \lambda_2 Id)\big(\alpha_2 \lambda_2 v_2 - \lambda_1 \alpha_2 v_2\big) & \overset{\text{linear}}{=} & \\
(\alpha_2 \lambda_2 - \lambda_1 \alpha_2)(T - \lambda_2 Id)(v_2) & \overset{\text{autovet.}}{=} & \\
(\alpha_2 \lambda_2 - \lambda_1 \alpha_2)\,\vec{0} & = & \vec{0}
\end{array}$$

ou seja, neste caso obtemos que $(T - \lambda_2 Id) \circ (T - \lambda_1 Id) = \vec{0}_{\mathcal{L}(V,V)}$.

Agora, sem demonstração, vejamos o resultado abaixo.

Proposição 0.194. *Sejam* V *um espaço vetorial sobre* $\mathbb{K}$ *de dimensão finita e* $T \in \mathcal{L}(V, V)$. *Sejam* $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ *todos os autovalores de* T *(distintos entre si). Se* $m_T(x) = (x - \lambda_k).(x - \lambda_{k-1}). \cdots .(x - \lambda_1)$, *então* T *é diagonalizável.* ■

Observação 0.195. *Com os dois resultados anteriores, temos mais uma caracterização de* T *ser diagonalizável, a saber:* T *é diagonalizável se, e somente se,* $m_T(x)$ *só admite* **raízes simples** *(cada autovalor de* T *deve ser raiz de* $m_T(x)$ *uma única vez).*

Exemplo 0.196. *Seja* $T : \mathbb{R}^{11} \to \mathbb{R}^{11}$ *um operador linear dado por*

$$T(x_1, x_2, \ldots, x_{11}) = (0, x_1, 0, \ldots, 0).$$

Então T *não é diagonalizável. Com efeito,*

$$T^2(x_1, x_2, \ldots, x_{11}) = T(T(x_1, x_2, \ldots, x_{11})) = T(0, x_1, 0 \ldots, 0) = (0, 0, \ldots, 0),$$

donde segue que $T^2 = \vec{0}_{\mathcal{L}(\mathbb{R}^{11}, \mathbb{R}^{11})}$, *ou seja,* T *é raiz de* x^2 *e portanto* x^2 *é um múltiplo de* $m_T(x)$. *Já que os únicos divisores de* x^2 *são* x *e* x^2, *e* T *não anula o primeiro, então* $m_T(x) = x^2$ *e portanto* T *não é diagonalizável (pois* 0 *não é uma raiz simples de* $m_T(x)$*).*

Exemplo 0.197. *Seja* $S : \mathbb{R}^{11} \to \mathbb{R}^{11}$ *um operador linear dado por*

$$S(x_1, x_2, \ldots, x_{11}) = (x_1, 0, \ldots, 0).$$

Então S *é diagonalizável. Com efeito,*

$$S^2(x_1, x_2, \ldots, x_{11}) = S(S(x_1, x_2, \ldots, x_{11})) = S(x_1, 0 \ldots, 0) = (x_1, 0, \ldots, 0),$$

donde segue que $S^2 = S$, ou seja, S é raiz de $x^2 - x$ e portanto $x^2 - x$ é um múltiplo de $m_S(x)$. Como os únicos divisores de $x^2 - x$ são x, $x - 1$ e $x(x-1)$, e S não anula os dois primeiros, então $m_S(x) = x(x-1)$ e portanto S é diagonalizável (pois 0 e 1 são raízes simples de $m_S(x)$).

Exemplo 0.198. *Seja $S : \mathbb{R}^3 \to \mathbb{R}^3$ um operador linear tal que $m_S(x) = x^2-24x+143$. S tem autovalores? S é diagonalizável? Quais são os possíveis $p_S(x)$? Sim, S tem autovalores, pois, como $m_S(x) = x^2 - 24x + 143 = (x - 11)(x - 13)$, então 11 e 13 são todos os autovalores de S. Ainda, S é diagonalizável, pois $m_S(x)$ só tem raízes simples (cada autovalor de S é raiz de $m_S(x)$ uma única vez). Finalmente, como $p_S(x)$ é um múltiplo de $m_S(x)$ que tem as mesmas raízes deste, então $p_S(x) = (x - 11)^i(x - 13)^j$, para algum $i, j \geqslant 1$. Desde que o grau de $p_S(x)$ é 3, pois $\dim_{\mathbb{R}} \mathbb{R}^3 = 3$, então ou $p_S(x) = (x - 11)(x - 13)^2$, ou $p_S(x) = (x - 11)^2(x - 13)$.*

Espaço vetorial com produto interno

De agora em diante, consideraremos apenas espaços vetoriais sobre $\mathbb{R}$ (corpo dos números reais).

Definição 0.199. *Seja V um $\mathbb{R}$-espaço vetorial. Um* **produto interno em** *V é uma função*

$$\begin{aligned} \langle\, , \rangle : V \times V &\to \mathbb{R} \\ (u, v) &\mapsto \langle u, v\rangle \in \mathbb{R} \end{aligned}$$

satisfazendo:

(P1) $\langle \lambda u + v, w\rangle = \lambda\langle u, w\rangle + \langle v, w\rangle$, se $u, v, w \in V$, $\lambda \in \mathbb{R}$

(P2) $\langle v, u\rangle = \langle u, v\rangle$, se $u, v \in V$

(P1) $\langle u, u\rangle > 0$, se $u \neq \vec{0}$, $u \in V$

Assim, um $\mathbb{R}$-espaço vetorial V munido com um produto interno é chamado de um **espaço vetorial** *(real)* **com produto interno**.

Proposição 0.200. *Seja V um espaço vetorial com produto interno $\langle\ ,\ \rangle$. Então:*

1. $\langle v, \vec{0}\rangle = \langle \vec{0}, v\rangle = 0$, *se* $v \in V$
2. $\langle u, \lambda v + w\rangle = \lambda\langle u, v\rangle + \langle u, w\rangle$, *se* $u, v, w \in V$, $\lambda \in \mathbb{R}$
3. $\langle \sum_{i=1}^{m} \alpha_i u_i, \sum_{j=1}^{n} \beta_j v_j\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\beta_j\langle u_i, v_j\rangle$, *se* $\alpha_i, \beta_j \in \mathbb{R}$, $u_i, v_j \in V$, *com* $i = 1, \ldots, m$, $j = 1, \ldots, n$
4. $\langle u, u\rangle = 0$ *se, e somente se,* $u = \vec{0}$

Mostremos (1). Como $\vec{0} = 1.\vec{0} + \vec{0}$, então $\langle \vec{0}, v\rangle = \langle 1.\vec{0} + \vec{0}, v\rangle \overset{(P1)}{=} 1\langle \vec{0}, v\rangle + \langle \vec{0}, v\rangle$ e portanto, somando $-\langle \vec{0}, v\rangle$ dos dois lados igualdade, obtemos que $\langle \vec{0}, v\rangle = 0$. Por $(P2)$, segue então que $\langle v, \vec{0}\rangle = \langle \vec{0}, v\rangle = 0$, para todo $v \in V$.
Mostremos (2). Para todo $u, v, w \in V$, $\lambda \in \mathbb{R}$, $\langle u, \lambda v + w\rangle \overset{(P2)}{=} \langle \lambda v + w, u\rangle \overset{(P1)}{=} \lambda\langle v, u\rangle + \langle w, u\rangle \overset{(P2)}{=} \lambda\langle u, v\rangle + \langle u, w\rangle$.
Mostremos (3). Sejam $u_i, v_j, v \in V$, $\alpha_i, \beta_j \in \mathbb{R}$, com $i = 1, \ldots, m$, $j = 1, \ldots, n$. Mostremos primeiro que $\langle \sum_{i=1}^{m} \alpha_i u_i, v\rangle = \sum_{i=1}^{m} \alpha_i\langle u_i, v\rangle$. Façamos por indução em m. Se $m = 1$, então o resultado segue por $(P1)$. Suponhamos que o resultado vale para $m - 1$ (HI). Logo, $\langle \sum_{i=1}^{m} \alpha_i u_i, v\rangle = \langle \sum_{i=1}^{m-1} \alpha_i u_i + \alpha_m u_m, v\rangle \overset{(P1)}{=} \langle \sum_{i=1}^{m-1} \alpha_i u_i, v\rangle + \langle \alpha_m u_m, v\rangle \overset{(HI)}{=} \sum_{i=1}^{m-1} \alpha_i\langle u_i, v\rangle + \alpha_m\langle u_m, v\rangle$. Agora, por $(P2)$, obtemos que $\langle v, \sum_{j=1}^{n} \beta_j v_j\rangle = \sum_{i=1}^{n} \beta_j\langle v, v_j\rangle$. Usando estas duas igualdades, obtemos o resultado.
Mostremos (4). Se $u \neq \vec{0}$, então, por $(P3)$, $\langle u, u\rangle > 0$ e portanto $\langle u, u\rangle \neq 0$. Se $u = \vec{0}$, então, pela primeira parte, $\langle \vec{0}, \vec{0}\rangle = 0$. Logo, $\langle u, u\rangle = 0$ se, e só se, $u = \vec{0}$. ■

Exemplo 0.201. *Consideremos $\mathbb{R}^n$ como sendo $\mathbb{R}$-espaço vetorial. Definamos $\langle\ ,\ \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ por $\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle := x_1.y_1 + \cdots + x_n.y_n \in \mathbb{R}$.*

Mostremos que $\langle\, ,\, \rangle$ *é um produto interno em* $\mathbb{R}^n$*. De fato,* $\langle \lambda(x_1, \dots, x_n) + (y_1, \dots, y_n), (z_1, \dots, z_n)\rangle = \langle(\lambda x_1 + y_1, \dots, \lambda x_n + y_n), (z_1, \dots, z_n)\rangle = (\lambda x_1 + y_1).z_1 + \cdots + (\lambda x_n + y_n).z_n = \lambda(x_1.z_1 + \cdots + x_n.z_n) + y_1.z_1 + \cdots + y_n.z_n = \lambda\langle(x_1, \dots, x_n), (z_1, \dots, z_n)\rangle + \langle(y_1, \dots, y_n), (z_1, \dots, z_n)\rangle$*. Ainda, é fácil ver que* $(P2)$ *vale. Finalmente, como* $\langle(x_1, \dots, x_n), (x_1, \dots, x_n)\rangle = x_1^2 + \cdots + x_n^2$ *que é maior do que* 0*, se* $(x_1, \dots, x_n) \neq (0, \dots, 0)$*, segue que* $\langle\, ,\, \rangle$ *é um produto interno em* $\mathbb{R}^n$*, chamado de produto interno canônico de* $\mathbb{R}^n$*.*

Exemplo 0.202. *Consideremos* $\mathbb{R}^n$ *como sendo* $\mathbb{R}$*-espaço vetorial. Sejam* $\alpha_1, \alpha_2, \dots, \alpha_n$ *escalares positivos. Definamos* $\langle\, ,\, \rangle_{\alpha_1 \dots \alpha_n} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *por* $\langle(x_1, \dots, x_n), (y_1, \dots, y_n)\rangle_{\alpha_1 \dots \alpha_n} := \alpha_1.x_1.y_1 + \alpha_2.x_2.y_2 + \cdots + \alpha_n.x_n.y_n \in \mathbb{R}$*. Mostremos que* $\langle\, ,\, \rangle_{\alpha_1 \dots \alpha_n}$ *é um produto interno em* $\mathbb{R}^n$*. É fácil ver que* $(P1)$ *e* $(P2)$ *valem. Agora, desde que* $\langle(x_1, \dots, x_n), (x_1, \dots, x_n)\rangle_{\alpha_1 \dots \alpha_n} = \alpha_1.x_1^2 + \alpha_2.x_2^2 + \cdots + \alpha_n.x_n^2$ *que é maior do que* 0*, se* $(x_1, \dots, x_n) \neq (0, \dots, 0)$*, segue que* $\langle\, ,\, \rangle_{\alpha_1 \dots \alpha_n}$ *é um produto interno em* $\mathbb{R}^n$*. Assim, por este exemplo, vemos que um* $\mathbb{R}$*-espaço vetorial pode admitir diferentes produtos internos.*

Exemplo 0.203. *Consideremos o* $\mathbb{R}$*-espaço vetorial das matrizes* $m \times n$ *com coeficientes em* $\mathbb{R}$*,* $\mathbb{M}_{m\times n}(\mathbb{R})$*. Definamos* $\langle\, ,\, \rangle : \mathbb{M}_{m\times n}(\mathbb{R}) \times \mathbb{M}_{m\times n}(\mathbb{R}) \to \mathbb{R}$ *por*

$$\langle(a_{ij})_{i,j}, (b_{ij})_{i,j}\rangle := \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}.b_{ij} \in \mathbb{R}.$$

Então $\langle\, ,\, \rangle$ *é um produto interno em* $\mathbb{M}_{m\times n}(\mathbb{R})$ *(verifique), chamado de produto interno canônico de* $\mathbb{M}_{m\times n}(\mathbb{R})$*.*

Exemplo 0.204. *Consideremos o* $\mathbb{R}$*-espaço vetorial das funções contínuas de* $[a, b]$ *em* $\mathbb{R}$*,* $\mathcal{C}([a, b], \mathbb{R})$*. Definamos* $\langle\, ,\, \rangle : \mathcal{C}([a, b], \mathbb{R}) \times \mathcal{C}([a, b], \mathbb{R}) \to \mathbb{R}$ *por*

$$\langle f, g\rangle := \int_a^b f(x).g(x)dx \in \mathbb{R}.$$

Então $\langle\, ,\, \rangle$ *é um produto interno em* $\mathcal{C}([a, b], \mathbb{R})$*. De fato, basta usarmos que* $\int_a^b (\lambda f(x) + g(x)).h(x)dx = \lambda \int_a^b f(x).h(x)dx + \int_a^b g(x).h(x)dx$*,* $\int_a^b g(x).f(x)dx = \int_a^b f(x).g(x)dx$*, e* $\int_a^b (f(x))^2 dx > 0$*, se* f *não for identicamente nula, isto é,* *se* $f \neq \vec{0}_{\mathcal{C}([a,b],\mathbb{R})}$. *Com estas propriedades, verifique que* $\langle\, , \rangle$ *é um produto interno em* $\mathcal{C}([a,b],\mathbb{R})$

O próximo resultado nos dá uma maneira de obter um produto interno a partir de um outro produto interno.

Proposição 0.205. *Sejam* V, W $\mathbb{R}$*-espaços vetoriais,* V *sendo com produto interno* $\langle\, , \rangle_V$*. Seja* $T : W \to V$ *uma transformação linear injetora. Então a função* $\langle\, , \rangle_T : W \times W \to \mathbb{R}$ *dada por* $\langle w_1, w_2\rangle_T := \langle T(w_1), T(w_2)\rangle_V$ *é um produto interno em* W*.*

Com efeito, sejam $\lambda \in \mathbb{R}$, $w_1, w_2, w_3 \in W$. Então $\langle \lambda w_1 + w_2, w_3\rangle_T = \langle T(\lambda w_1 + w_2), T(w_3)\rangle_V \overset{T\,\text{linear}}{=} \langle \lambda T(w_1) + T(w_2), T(w_3)\rangle_V \overset{\langle\, ,\rangle_V\ \text{p.i.}}{=} \lambda\langle T(w_1), T(w_3)\rangle_V + \langle T(w_2), T(w_3)\rangle_V = \lambda\langle w_1, w_3\rangle_T + \langle w_2, w_3\rangle_T$.
Ainda, $\langle w_1, w_2\rangle_T = \langle T(w_1), T(w_2)\rangle_V \overset{\langle\, ,\rangle_V\ \text{p.i.}}{=} \langle T(w_2), T(w_1)\rangle_V = \langle w_2, w_1\rangle_T$.
Finalmente, $\langle w_1, w_1\rangle_T = \langle T(w_1), T(w_1)\rangle_V \overset{\langle\, ,\rangle_V\ \text{p.i.}}{>} 0$, se $T(w_1) \neq \vec{0}$. Como T é injetora, segue que $T(w_1) \neq \vec{0}$ se, e só se, $w_1 \neq \vec{0}$. Logo, $\langle w_1, w_1\rangle_T > 0$, se $w_1 \neq \vec{0}$. Portanto, $\langle\, , \rangle_T$ é um produto interno em W. ■

Corolário 0.206. *Todo* $\mathbb{R}$*-espaço vetorial de dimensão finita admite um produto interno.*

De fato, sejam V um $\mathbb{R}$-espaço vetorial e $\mathcal{B} := \{v_1, \ldots, v_n\}$ uma base de V. Consideremos $\mathbb{R}^n$ como sendo $\mathbb{R}$-espaço vetorial e $\mathcal{B}_C$ a base canônica de $\mathbb{R}^n$. Desde que $\mathcal{B}$ e $\mathcal{B}_C$ são bases que têm o mesmo número de elementos, existe uma única transformação linear $T : V \to \mathbb{R}^n$ tal que $T(v_1) = (1, 0, \ldots, 0), T(v_2) = (0, 1, 0, \ldots, 0), \ldots, T(v_n) = (0, \ldots, 0, 1)$ e existe uma única $S : \mathbb{R}^n \to V$ que faz o contrário. Assim, T é um isomorfismo (pois verificamos que $S = T^{-1}$) dado por $T(\sum_{i=1}^n \alpha_i v_i) = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$. Logo, $\langle \sum_{i=1}^n \alpha_i v_i, \sum_{i=1}^n \beta_i v_i\rangle_T := \langle(\alpha_1, \alpha_2, \ldots, \alpha_n), (\beta_1, \beta_2, \ldots, \beta_n)\rangle_{\mathbb{R}^n}$ é um produto interno em V, onde $\langle\, , \rangle_{\mathbb{R}^n}$ é o produto interno canônico de $\mathbb{R}^n$. Em outras palavras, se $v, w \in V$, então $\langle v, w\rangle_T := \langle [v]_{\mathcal{B}}, [w]_{\mathcal{B}}\rangle_{\mathbb{R}^n}$. ■

Definição 0.207. *Seja V um espaço vetorial com produto interno $\langle\, ,\, \rangle$. Para cada $v \in V$, chamamos o número $\sqrt{\langle v, v\rangle} \in \mathbb{R}$ de* **norma de** *v e o denotamos por $||v||$.*

Exemplo 0.208. *Consideremos $\mathbb{R}^n$ com o produto interno (canônico)*

$$\langle (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)\rangle = x_1.y_1 + x_2.y_2 + \cdots + x_n.y_n.$$

Então

$$||(x_1, x_2, \ldots, x_n)|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Com efeito, $||(x_1,\, x_2,\, \ldots,\, x_n)|| = \sqrt{\langle (x_1,\, x_2,\, \ldots,\, x_n), (x_1,\, x_2,\, \ldots,\, x_n)\rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.

Exemplo 0.209. *Consideremos $\mathcal{C}([a, b], \mathbb{R})$ com o produto interno*

$$\langle f, g\rangle = \int_a^b f(x).g(x)dx \in \mathbb{R}.$$

Então

$$||f|| = \sqrt{\int_a^b f(x)^2 dx}.$$

De fato, $||f|| = \sqrt{\langle f, f\rangle} = \sqrt{\int_a^b f(x)^2 dx}$.

Proposição 0.210. *Seja V um espaço vetorial com produto interno $\langle\, ,\, \rangle$. Então:*

1. *$||v|| \geqslant 0$, para todo $v \in V$*

2. *$||v|| = 0$ se, e somente se, $v = \vec{0}$*

3. *$||\alpha.v|| = |\alpha|.||v||$, para todo $\alpha \in \mathbb{R}$, $v \in V$*

Com efeito, por um lado, se $v = \vec{0}$, então $\langle v, v\rangle = \langle \vec{0}, \vec{0}\rangle = 0$, donde $||v|| = \sqrt{\langle v, v\rangle} = \sqrt{0} = 0$. Por outro lado, se $v \neq \vec{0}$, então, por $(P3)$, $\langle v, v\rangle > 0$ e portanto $||v|| = \sqrt{\langle v, v\rangle} > 0$. Logo, em qualquer caso, $||v|| \geqslant 0$. Agora, notemos que, com esta argumentação, também já mostramos que $v = \vec{0}$ se, e somente se, $||v|| = 0$. Mostremos, finalmente, (3). De fato, $||\alpha.v|| = \sqrt{\langle \alpha.v, \alpha.v\rangle} = \sqrt{\alpha.\alpha.\langle v, v\rangle} = \sqrt{\alpha^2}.\sqrt{\langle v, v\rangle} = |\alpha|.||v||$, para todo $\alpha \in \mathbb{R}$, $v \in V$. ■

Proposição 0.211. *Seja V um espaço vetorial com produto interno $\langle\, ,\, \rangle$. Sejam $u, v \in V$. Então:*

1. $\langle u, v\rangle = \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2$
2. $|\langle u, v\rangle| \leqslant ||u||.||v||$, *e a igualdade vale se, e somente se, $\{u, v\}$ é LD*
3. $||u+v|| \leqslant ||u|| + ||v||$

Com efeito, mostremos (1). Temos que

$$\begin{aligned}
||u+v||^2 = \langle u+v, u+v\rangle &\overset{(P1)}{=} \langle u, u+v\rangle + \langle v, u+v\rangle \\
&= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
&\overset{(P2)}{=} \langle u, u\rangle + \langle u, v\rangle + \langle u, v\rangle + \langle v, v\rangle \\
&= ||u||^2 + 2\langle u, v\rangle + ||v||^2, \\
||u-v||^2 = \langle u-v, u-v\rangle &\overset{(P1)}{=} \langle u, u-v\rangle - \langle v, u-v\rangle \\
&= \langle u, u\rangle - \langle u, v\rangle - \langle v, u\rangle + \langle v, v\rangle \\
&\overset{(P2)}{=} ||u||^2 - 2\langle u, v\rangle + ||v||^2.
\end{aligned}$$

Logo, $||u+v||^2 - ||u-v||^2 = 4\langle u, v\rangle$, donde o resultado segue.
Mostremos agora (2). Se $\alpha, \beta \in \mathbb{R}$, então

$$\begin{aligned}
0 \leqslant \langle \alpha u - \beta v, \alpha u - \beta v\rangle &\overset{(P1)}{=} \alpha\langle u, \alpha u - \beta v\rangle - \beta\langle v, \alpha u - \beta v\rangle \\
&= \alpha^2\langle u, u\rangle - \alpha\beta\langle u, v\rangle - \beta\alpha\langle v, u\rangle + \beta^2\langle v, v\rangle \\
&\overset{(P2)}{=} \alpha^2||u||^2 - 2\alpha\beta\langle u, v\rangle + \beta^2||v||^2.
\end{aligned}$$

Tomando $\alpha := ||v||^2$ e $\beta := \langle u, v\rangle$, obtemos que

$$0 \leqslant (||v||^2)^2.||u||^2 - 2||v||^2.\langle u, v\rangle.\langle u, v\rangle + \langle u, v\rangle^2.||v||^2,$$

isto é, $0 \leqslant ||v||^2.\big(||u||^2.||v||^2 - \langle u, v\rangle^2\big)$, donde, como $||v|| \geqslant 0$, segue que $||u||^2.||v||^2 - \langle u, v\rangle^2 \geqslant 0$. Daí, $|\langle u, v\rangle| = \sqrt{\langle u, v\rangle^2} \leqslant \sqrt{||u||^2.||v||^2} = ||u||.||v||$.
Agora, por um lado, se $\{u, v\}$ é LD, então $v = \lambda u$, para algum $\lambda \in \mathbb{R}$. Daí, $|\langle u, v\rangle| = |\langle u, \lambda u\rangle| = |\lambda\langle u, u\rangle| = |\lambda|.\langle u, u\rangle = |\lambda|.||u||^2 = ||u||.||v||$, pois $||v|| = |\lambda|.||u||$. Por outro lado, se $|\langle u, v\rangle| = ||u||.||v||$, então $0 = \langle \alpha u - \beta v, \alpha u - \beta v\rangle$, onde $\alpha := ||v||^2$ e $\beta := \langle u, v\rangle$. Logo, $\alpha u + \beta v = \vec{0}$. Se $v = \vec{0}$, acabou, pois $\{u, \vec{0}\}$ é LD. Se $v \neq \vec{0}$, então $\alpha = ||v||^2 \neq 0$ e portanto $u = \frac{\beta}{\alpha}v$, ou seja, u é um múltiplo de v e daí $\{u, v\}$ é LD.
Finalmente, mostremos (3). De fato, por (2), temos que $||u + v||^2 = ||u||^2 + 2\langle u, v\rangle + ||v||^2 \leqslant ||u||^2 + 2|\langle u, v\rangle| + ||v||^2 \leqslant ||u||^2 + 2||u||.||v|| + ||v||^2 = \big(||u|| + ||v||\big)^2$. Logo, $||u + v|| \leqslant ||u|| + ||v||$. ∎

Observação 0.212. *A igualdade do item (1) acima é a chamada de* **identidade de polarização** *(para $\mathbb{R}$). A desigualdade do item* (2) *acima é chamada de* **desigualdade de Schwarz**. *E a desigualdade do item* (3) *acima é chamada de* **desigualdade triangular**.

Observação 0.213. *Seja V um espaço vetorial com produto interno $\langle\, ,\, \rangle$. Definindo $d : V \times V \to \mathbb{R}$ por $d(u, v) = ||u - v||$, para todo $u, v \in V$, temos que d é uma* **métrica** *em V, pois:*

1. $d(u, v) \geqslant 0$, *se* $u, v \in V$
2. $d(u, v) = 0$ *se, e somente se,* $u = v$
3. $d(u, v) = d(v, u)$, *se* $u, v \in V$
4. $d(u, v) \leqslant d(u, w) + d(w, v)$, *se* $u, v, w \in V$

Exemplo 0.214. *Consideremos $\mathbb{R}^3$ com o produto interno*

$$\langle (x_1, x_2, x_3), (y_1, y_2, y_3) \rangle_* = 17x_1.y_1 + 13x_2.y_2 + 11x_3.y_3$$

(verifique que de fato $\langle\ ,\ \rangle_*$ é um produto interno). Então $||(x_1, x_2, x_3)||_* = \sqrt{17x_1^2 + 13x_2^2 + 11x_3^2}$ *e portanto* $||(1, 0, 0)||_* = \sqrt{17}$, $||(0, 1, 0)||_* = \sqrt{13}$ *e* $||(0, 0, 1)||_* = \sqrt{11}$. *Daí,* $||(\frac{1}{\sqrt{17}}, 0, 0)||_* = 1$, $||(0, \frac{1}{\sqrt{13}}, 0)||_* = 1$ *e* $||(0, 0, \frac{1}{\sqrt{11}})||_* = 1$. *Consideremos agora* $\mathbb{R}^3$ *com o produto interno canônico. Então* $||(1, 0, 0)|| = 1$, $||(0, 1, 0)|| = 1$ *e* $||(0, 0, 1)|| = 1$.

Exemplo 0.215. *Seja V um espaço vetorial com produto interno $\langle\ ,\ \rangle$. Sejam $u, v \in V$ tais que:*

a) $||u|| = 3$, $||v|| = 4$ *e* $||u + v|| = 7$. *Calculemos* $\langle u, v\rangle$ *e* $||u - v||$. *Como* $7^2 = ||u+v||^2 = \langle u+v, u+v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle = 3^2 + 2\langle u, v\rangle + 4^2$, *então* $\langle u, v\rangle = 12$. *Daí, como* $||u - v||^2 = \langle u - v, u - v\rangle = \langle u, u\rangle + \langle u, -v\rangle + \langle -v, u\rangle + \langle -v, -v\rangle = 3^2 - 2.12 + 4^2 = 1$, *então* $||u - v|| = 1$.

b) $||u|| = 3$, $||v|| = 4$ *e* $||u - v|| = 5$. *Calculemos* $\langle u, v\rangle$ *e* $||u + v||$. *Desde que* $5^2 = ||u - v||^2 = 3^2 - 2\langle u, v\rangle + 4^2$, *então* $\langle u, v\rangle = 0$. *Daí, como* $||u + v||^2 = 3^2 + 2.0 + 4^2 = 25$, *então* $||u + v|| = 5$.

c) $||u|| = 0$ *e* $||v|| = 1$. *Calculemos* $\langle u, v\rangle$ *e* $||u - v||$ *e* $||u + v||$. *Já que* $||u|| = 0$, *então* $u = \vec{0}$, *donde* $\langle u, v\rangle = \langle \vec{0}, v\rangle = 0$, $||u - v|| = || - v|| = | - 1| ||v|| = 1$ *e* $||u + v|| = ||v|| = 1$.

d) $||u - v|| = 2$ *e* $||u + v|| = 4$. *Calculemos* $\langle u, v\rangle$. *Temos que* $\langle u, v\rangle = \frac{1}{4}||u + v||^2 + \frac{1}{4}||u - v||^2 = \frac{1}{4}4^2 - \frac{1}{4}2^2 = 3$.

Observação 0.216. *Seja V um espaço vetorial com produto interno $\langle\ ,\ \rangle$. Se, para todo $v \in V$, $\langle u, v\rangle = 0$, então $u = \vec{0}$. Com efeito, tomando em particular $v = u$, obtemos que $\langle u, u\rangle = 0$ e portanto $u = \vec{0}$.*

Ortogonalidade

Definição 0.217. *Seja V um espaço vetorial com produto interno $\langle\ ,\ \rangle$.*

1. *Sejam $u, v \in V$. Dizemos que u e v são* **ortogonais** *se $\langle u, v\rangle = 0$.*

2. *Um subconjunto $\mathcal{S}$ de V é chamado de* **ortogonal** *se $\langle u, v\rangle = 0$, para todo $u \neq v$, $u, v \in \mathcal{S}$*

3. *Um subconjunto $\mathcal{S}$ de V é chamado de* **ortonormal** *se for ortogonal e $||u|| = 1$, para todo $u \in \mathcal{S}$*

Observação 0.218. *Notemos que, se V é um espaço vetorial com produto interno $\langle\, ,\, \rangle$, então o vetor nulo de V é ortogonal a todos os vetores de V, pois $\langle \vec{0}, v\rangle = 0$, para todo $v \in V$. Além disto, mostramos que se $\langle u, v\rangle = 0$, para todo $v \in V$, então $u = \vec{0}$, donde $\vec{0}$ é o único vetor de V que é ortogonal a todos os vetores de V.*

Exemplo 0.219. *A base canônica de $\mathbb{R}^n$ (este com o produto interno canônico) é ortonormal.*
A base canônica de $\mathbb{M}_{m\times n}(\mathbb{R})$ (este com o produto interno canônico) é ortonormal. De fato, para $m = n = 2$, lembremos que $\langle \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{smallmatrix}\right), \left(\begin{smallmatrix} b_{11} & b_{12} \\ b_{21} & b_{22}\end{smallmatrix}\right)\rangle = a_{11}.b_{11} + a_{12}.b_{12} + a_{21}.b_{21} + a_{22}.b_{22}$ e $\mathcal{B}_C = \{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1\end{smallmatrix}\right)\}$. Como $\langle \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0\end{smallmatrix}\right)\rangle = 1.1 + 0.0 + 0.0 + 0.0 = 1$, então $||\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0\end{smallmatrix}\right)|| = \sqrt{1} = 1$. Analogamente, $||\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0\end{smallmatrix}\right)|| = 1$, $||\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0\end{smallmatrix}\right)|| = 1$ e $||\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1\end{smallmatrix}\right)|| = 1$. Logo, os quatro vetores de $\mathcal{B}_C$ são unitários. Agora, desde que, $\langle \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0\end{smallmatrix}\right)\rangle = 1.0 + 0.1 + 0.0 + 0.0 = 0$, então $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0\end{smallmatrix}\right)$ e $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0\end{smallmatrix}\right)$ são ortogonais. Analogamente, obtemos que $\langle \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0\end{smallmatrix}\right)\rangle = 0, \langle \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1\end{smallmatrix}\right)\rangle = 0, \langle \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0\end{smallmatrix}\right)\rangle = 0, \langle \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1\end{smallmatrix}\right)\rangle = 0$ e $\langle \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1\end{smallmatrix}\right)\rangle = 0$, donde $\mathcal{B}_C$ é ortonormal.

Proposição 0.220. *Seja V um espaço vetorial com produto interno $\langle\, ,\, \rangle$. Seja $\mathcal{S}$ um subconjunto ortogonal de V formado por vetores não nulos. Então:*

1. *Se $v \in [v_1, v_2, \ldots, v_n]$, com $v_1, v_2, \ldots, v_n \in \mathcal{S}$, segue que $v = \frac{\langle v, v_1\rangle}{||v_1||^2}v_1 + \frac{\langle v, v_2\rangle}{||v_2||^2}v_2 + \cdots + \frac{\langle v, v_n\rangle}{||v_n||^2}v_n$*

2. *S é LI*

De fato, se $v \in [v_1, v_2, \ldots, v_n]$, então existem escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tais que $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. Mas $\langle v_i, v_j \rangle = 0$, se $i \neq j$, pois $v_1, v_2, \ldots, v_n \in \mathcal{S}$ e $\mathcal{S}$ é ortogonal. Daí, $\langle v, v_1 \rangle = \langle \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n, v_1 \rangle \overset{(P1)}{=} \alpha_1 \langle v_1, v_1 \rangle + \alpha_2 \langle v_2, v_1 \rangle + \cdots + \alpha_n \langle v_n, v_1 \rangle = \alpha_1 \langle v_1, v_1 \rangle$. Como $v_1 \neq \vec{0}$, então $\langle v_1, v_1 \rangle \neq 0$, donde $\alpha_1 = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} = \frac{\langle v, v_1 \rangle}{||v_1||^2}$. Da mesma forma, $\alpha_2 = \frac{\langle v, v_2 \rangle}{||v_2||^2}, \ldots, \alpha_n = \frac{\langle v, v_n \rangle}{||v_n||^2}$. Logo, $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \frac{\langle v, v_1 \rangle}{||v_1||^2} v_1 + \frac{\langle v, v_2 \rangle}{||v_2||^2} v_2 + \cdots + \frac{\langle v, v_n \rangle}{||v_n||^2} v_n$.
Mostremos (2). Seja $c := \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_m u_m$ uma combinação linear qualquer de vetores de $\mathcal{S}$ (com $u_1, u_2, \ldots, u_m \in \mathcal{S}$ e $\beta_1, \beta_2, \ldots, \beta_m \in \mathbb{R}$). Suponhamos que c é o vetor nulo. Daí, $\vec{0} = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_m u_m$. Pela prova de (1), como $\vec{0} \in [u_1, u_2, \ldots, u_m]$, obtemos que $\beta_1 = \frac{\langle \vec{0}, u_1 \rangle}{||u_1||^2}, \beta_2 = \frac{\langle \vec{0}, u_2 \rangle}{||u_2||^2}, \ldots, \beta_m = \frac{\langle \vec{0}, u_m \rangle}{||u_m||^2}$, donde $\beta_1 = \beta_2 = \cdots = \beta_m = 0$ e portanto $\mathcal{S}$ é LI. ■

Corolário 0.221. *Seja V um espaço vetorial com produto interno $\langle\, , \rangle$. Seja $\mathcal{B} := \{v_1, v_2, \ldots, v_n\}$ uma base ortonormal de V. Então, para cada $v \in V$, $v = \langle v, v_1 \rangle v_1 + \langle v, v_2 \rangle v_2 + \cdots + \langle v, v_n \rangle v_n$.*

Com efeito, desde que $\mathcal{B}$ é uma base de V, então $[v_1, v_2, \ldots, v_n] = V$. Agora, como $\mathcal{B}$ é, em particular, um subconjunto ortogonal de V formado por vetores não nulos, então, pelo resultado anterior, $v = \frac{\langle v, v_1 \rangle}{||v_1||^2} v_1 + \frac{\langle v, v_2 \rangle}{||v_2||^2} v_2 + \cdots + \frac{\langle v, v_n \rangle}{||v_n||^2} v_n$, para todo $v \in [v_1, v_2, \ldots, v_n] = V$. Finalmente, desde que $||v_1|| = ||v_2|| = \cdots = ||v_n|| = 1$, pois $\mathcal{B}$ é ortonormal, então o resultado segue. ■

O próximo resultado nos fornece um modo de obtermos um conjunto ortogonal a partir de um conjunto LI.

Teorema 0.222. *Seja V um espaço vetorial com produto interno $\langle\, , \rangle$. Se $\mathcal{S} := \{v_1, v_2, \ldots, v_n\}$ é um subconjunto LI de V, então existe $\widetilde{\mathcal{S}} := \{w_1, w_2, \ldots, w_n\}$ um subconjunto ortogonal de V tal que $[\widetilde{\mathcal{S}}] = [\mathcal{S}]$.*

Façamos por indução em n.
Se $n = 2$, tomemos $w_1 := v_1$ e $w_2 := v_2 - \frac{\langle v_2, w_1 \rangle}{||w_1||^2} w_1$. Notemos que $w_2 \neq \vec{0}$, pois v_2 não é um múltiplo de $w_1 = v_1$, já que $\{v_1, v_2\}$ é LI. Como $\langle w_1, w_2 \rangle =$

$\langle v_1, v_2\rangle - \frac{\langle v_2,v_1\rangle}{||v_1||^2}\langle v_1, v_1\rangle = \langle v_1, v_2\rangle - \langle v_2, v_1\rangle = 0$, então $\{w_1, w_2\}$ é ortogonal. Ainda, como $w_1, w_2 \in [v_1, v_2]$, então $[w_1, w_2] \subset [v_1, v_2]$. Daí, desde que $\dim_{\mathbb{R}}[w_1, w_2] = 2 = \dim_{\mathbb{R}}[v_1, v_2]$ (pois $\{w_1, w_2\}$ é base para $[w_1, w_2]$ e $\{v_1, v_2\}$ é base para $[v_1, v_2]$), segue que $[w_1, w_2] = [v_1, v_2]$.

Suponhamos por hipótese de indução (HI) que o resultado vale para $n - 1$. Seja $\mathcal{S} := \{v_1, v_2, \ldots, v_n\}$ um subconjunto LI de V. Como $\{v_1, v_2, \ldots, v_{n-1}\}$ é LI, então, por (HI), existe $\{w_1, w_2, \ldots, w_{n-1}\}$ um subconjunto ortogonal de V tal que $[v_1, v_2, \ldots, v_{n-1}] = [w_1, w_2, \ldots, w_{n-1}]$ (*). Tomemos $w_n := v_n - \sum_{i=1}^{n-1} \frac{\langle v_n,w_i\rangle}{||w_i||^2} w_i$. Seja $\widetilde{\mathcal{S}} := \{w_1, w_2, \ldots, w_n\}$. Notemos que $w_n \neq \vec{0}$, pois v_n não é combinação linear de $v_1, \ldots, v_{n-1}$, já que $\mathcal{S}$ é LI, e portanto também não é combinação linear de $w_1, \ldots, w_{n-1}$, por (*). Agora, para todo $j = 1, 2, \ldots, n-1$,

$$\begin{aligned}
\langle w_n, w_j\rangle &= \langle v_n - \textstyle\sum_{i=1}^{n-1} \frac{\langle v_n,w_i\rangle}{||w_i||^2} w_i, w_j\rangle \\
&= \langle v_n, w_j\rangle - \textstyle\sum_{i=1}^{n-1} \frac{\langle v_n,w_i\rangle}{||w_i||^2}\langle w_i, w_j\rangle \\
&= \langle v_n, w_j\rangle - \frac{\langle v_n,w_j\rangle}{||w_j||^2}\langle w_j, w_j\rangle \\
&= 0,
\end{aligned}$$

pois $\langle w_i, w_j\rangle = 0$, se $i \neq j$, uma vez que $\{w_1, w_2, \ldots, w_{n-1}\}$ é ortogonal. Daí, $\widetilde{S}$ é ortogonal. Ainda, como, pela construção de w_n e por (*), $w_1, w_2, \ldots, w_n \in [v_1, v_2, \ldots, v_n]$, segue que $[\widetilde{\mathcal{S}}] \subset [\mathcal{S}]$. Daí, desde que $\dim_{\mathbb{R}}[\widetilde{\mathcal{S}}] = n = \dim_{\mathbb{R}}[\mathcal{S}]$ (pois $\widetilde{\mathcal{S}}$ é base para $[\widetilde{\mathcal{S}}]$ e $\mathcal{S}$ é base para $[\mathcal{S}]$), segue que $[\widetilde{\mathcal{S}}] = [\mathcal{S}]$. ■

Observação 0.223. *A construção feita na prova acima é chamada de* **processo de ortogonalização de Gram-Schmidt**.

Observação 0.224. *Se V é um espaço vetorial com produto interno $\langle\, ,\, \rangle$ e $v \in V \setminus \{\vec{0}\}$, então o vetor $\frac{v}{||v||}$ é um múltiplo por escalar de v tal que $||\frac{v}{||v||}|| = 1$, pois $||\frac{v}{||v||}|| = |\frac{1}{||v||}|.||v|| = \frac{1}{||v||}.||v|| = 1$.*

Corolário 0.225. *Todo espaço vetorial (de dimensão finita maior do que 0) com produto interno tem uma base ortonormal.*

De fato, seja V um espaço vetorial com produto interno, com $\dim_{\mathbb{R}} V = n > 0$. Seja $\mathcal{B} := \{v_1, v_2, \ldots, v_n\}$ uma base de V. Pelo resultado anterior, existe $\widetilde{\mathcal{B}} := \{w_1, w_2, \ldots, w_n\}$ um subconjunto ortogonal de V tal que $[\widetilde{\mathcal{B}}] = [\mathcal{B}]$. Tomemos $\mathcal{C} := \{\frac{w_1}{||w_1||}, \frac{w_2}{||w_2||}, \ldots, \frac{w_n}{||w_n||}\}$. Como $\mathcal{C}$ é LI, pois é um subconjunto ortonormal de V, e gera V, pois $[\mathcal{C}] = [\widetilde{\mathcal{B}}] = [\mathcal{B}] = V$, segue que $\mathcal{C}$ é uma base ortonormal de V.

Exemplo 0.226. *Consideremos $\mathbb{R}^2$ com o produto interno canônico $\langle\, , \rangle$. Temos que $\{(1,1),(1,2)\}$ é uma base de $\mathbb{R}^2$. Vamos construir uma base ortonormal de $\mathbb{R}^2$. Sejam $v_1 := (1,1)$ e $v_2 := (1,2)$. Tomemos $w_1 := v_1$. Temos que $||w_1||^2 = \langle w_1, w_1\rangle = 1.1 + 1.1 = 2$ e $\langle v_2, w_1\rangle = 1.1 + 2.1 = 3$. Daí, tomemos $w_2 := v_2 - \frac{\langle v_2, w_1\rangle}{||w_1||^2} w_1 = (1,2) - \frac{3}{2}(1,1) = (1 - \frac{3}{2}, 2 - \frac{3}{2}) = (-\frac{1}{2}, \frac{1}{2})$. Logo, $\{w_1, w_2\}$ é ortogonal (pelo processo acima). Daí, $\{\frac{w_1}{||w_1||}, \frac{w_2}{||w_2||}\}$ é uma base ortonormal de $\mathbb{R}^2$ (pois é um conjunto LI com dois vetores). Notemos que $\frac{w_1}{||w_1||} = \frac{(1,1)}{\sqrt{2}} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ e $\frac{w_2}{||w_2||} = \frac{(-\frac{1}{2}, \frac{1}{2})}{\frac{\sqrt{2}}{2}} = (\frac{-\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$.*

Definição 0.227. *Seja V um espaço vetorial com produto interno $\langle\, , \rangle$. Seja $\mathcal{S}$ um subconjunto de V. Chamamos de* **ortogonal a $\mathcal{S}$** *ao conjunto $\mathcal{S}^\perp := \{v \in V | \langle v, u\rangle = 0, \text{ para todo } u \in \mathcal{S}\}$.*

Proposição 0.228. *Seja V um espaço vetorial com produto interno $\langle\, , \rangle$. Se $\mathcal{S}$ é um subconjunto de V, então $\mathcal{S}^\perp$ é um subespaço de V.*

De fato, $\vec{0} \in \mathcal{S}^\perp$, pois $\langle \vec{0}, u\rangle = 0$, para todo $u \in V$. Sejam $\lambda \in \mathbb{R}$, $v, w \in \mathcal{S}^\perp$. Como $\langle \lambda v + w, u\rangle = \lambda\langle v, u\rangle + \langle w, u\rangle = \lambda.0 + 0 = 0$, para todo $u \in V$, então $\lambda v + w \in \mathcal{S}^\perp$. Assim, $\mathcal{S}^\perp$ é um subespaço de V. ■

Observação 0.229. *Notemos que $\mathcal{S}^\perp$ é sempre um subespaço de V, mesmo que $\mathcal{S}$ não o seja.*
Ainda, se $\mathcal{S} = \{\vec{0}\}$, então, para todo $v \in V$, $\langle v, \vec{0}\rangle = 0$, donde $V \subset \mathcal{S}^\perp$ e portanto $\mathcal{S}^\perp = V$.
Se $\mathcal{S}$ contém uma base ortogonal $\{u_1, u_2, \ldots, u_n\}$ de V, então $\mathcal{S}^\perp = \{\vec{0}\}$. De

fato, se $v \in \mathcal{S}^\perp$, existem escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tais que $v = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n$. Daí, $0 = \langle v, u\rangle = \alpha_1\langle u_1, u\rangle + \alpha_2\langle u_2, u\rangle + \cdots + \alpha_n\langle u_n, u\rangle$, para todo $u \in \mathcal{S}$. Em particular, $0 = \langle v, u_1\rangle = \alpha_1\langle u_1, u_1\rangle$, donde, como $\langle u_1, u_1\rangle > 0$, segue que $\alpha_1 = 0$. Analogamente, obtemos que $\alpha_2 = \cdots = \alpha_n = 0$. Portanto, neste caso, se $v \in \mathcal{S}^\perp$, então $v = \vec{0}$, isto é, $\mathcal{S}^\perp = \{\vec{0}\}$.

Exemplo 0.230. *Consideremos $\mathbb{R}^3$ com o produto interno canônico $\langle\, ,\, \rangle$. Seja $\mathcal{S} := \{(1,1,1),(1,1,2)\}$. Calculemos $\mathcal{S}^\perp$ e mostremos que $\mathbb{R}^3 = [\mathcal{S}] \oplus \mathcal{S}^\perp$. Se $(x,y,z) \in \mathcal{S}^\perp$, então $\langle (x,y,z),(1,1,1)\rangle = 0$ e $\langle (x,y,z),(1,1,2)\rangle = 0$, isto é,*

$$\begin{cases} 1x + 1y + 1z = 0 \\ 1x + 1y + 2z = 0 \end{cases}$$

Logo, $z = 0$ e $y = -x$, donde $\mathcal{S}^\perp = \{(x,-x,0) | x \in \mathbb{R}\} = [(1,-1,0)]$. Agora, $[\mathcal{S}] = [(1,1,1),(1,1,2)] = \{\alpha(1,1,1) + \beta(1,1,2) | \alpha, \beta \in \mathbb{R}\} = \{(\alpha+\beta, \alpha+\beta, \alpha+2\beta) | \alpha, \beta \in \mathbb{R}\}$. Se $(\alpha+\beta, \alpha+\beta, \alpha+2\beta) = (x,-x,0)$, então $\alpha+\beta = x = -(\alpha+\beta)$ e $\alpha + 2\beta = 0$. Daí, $\beta = 0$ e $\alpha = 0$, donde $(\alpha+\beta, \alpha+\beta, \alpha+2\beta) = (0,0,0)$ e portanto $[\mathcal{S}] \cap \mathcal{S}^\perp = \{(0,0,0)\}$. Ainda, como $\{(1,1,1),(1,1,2),(1,-1,0)\}$ gera $\mathbb{R}^3$ (verifique), segue que $\mathbb{R}^3 = [\mathcal{S}] + \mathcal{S}^\perp$. Portanto, $\mathbb{R}^3 = [\mathcal{S}] \oplus \mathcal{S}^\perp$.

O próximo resultado nos dá uma caracterização dos vetores do ortogonal a um subespaço.

Proposição 0.231. *Seja V um espaço vetorial com produto interno $\langle\, ,\, \rangle$. Sejam W um subespaço de V e $\mathcal{B} := \{w_1, w_2, \ldots, w_k\}$ um conjunto gerador de W. Então $v \in W^\perp$ se, e somente se, $\langle v, w_i\rangle = 0$, para todo $i = 1, \ldots, k$.*

De fato, se $w \in W$, então existem escalares $\alpha_1, \alpha_2, \ldots, \alpha_k$ tais que $w = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_k w_k$. Daí, se $v \in V$, então $\langle v, w\rangle = \langle v, \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_k w_k\rangle = \alpha_1\langle v, w_1\rangle + \alpha_2\langle v, w_2\rangle + \cdots + \alpha_k\langle v, w_k\rangle$. Por um lado, se $\langle v, w_i\rangle = 0$, para todo $i = 1, \ldots, k$, então $\langle v, w\rangle = 0$, para todo $w \in W$, e portanto $v \in W^{\perp}$. Por outro lado, se $v \in W^{\perp}$, então, para todo $w \in W$, $\langle v, w\rangle = 0$ e daí $\langle v, w_i\rangle = 0$, para todo $i = 1, \ldots, k$, pois $w_1, w_2, \ldots, w_k \in W$. ■

Observação 0.232. *Seja V um espaço vetorial com produto interno $\langle\, ,\, \rangle$. Se $\mathcal{S}$ é um subconjunto de V, então $[\mathcal{S}]^{\perp} = \mathcal{S}^{\perp}$. De fato, por um lado, se $w \in [\mathcal{S}]^{\perp}$, então, em particular, $\langle w, u\rangle = 0$, para todo $u \in \mathcal{S}$, donde $w \in \mathcal{S}^{\perp}$ e portanto $[\mathcal{S}]^{\perp} \subset \mathcal{S}^{\perp}$. Por outro lado, se $w \in \mathcal{S}^{\perp}$, então $\langle w, u\rangle = 0$, para todo $u \in \mathcal{S}$. Daí, como $\mathcal{S}$ é um conjunto gerador de $[\mathcal{S}]$, então, pelo resultado anterior, $w \in [\mathcal{S}]^{\perp}$ e portanto $\mathcal{S}^{\perp} \subset [\mathcal{S}]^{\perp}$. Logo, $[\mathcal{S}]^{\perp} = \mathcal{S}^{\perp}$.*

Proposição 0.233. *Seja V um espaço vetorial com produto interno $\langle\, ,\, \rangle$ de dimensão finita. Se W é um subespaço de V, então $V = W \oplus W^{\perp}$.*

Com efeito, se $V = \{\vec{0}\}$, então $\{\vec{0}\} = \{\vec{0}\} \oplus \{\vec{0}\}$ e o resultado segue. Seja então $V \neq \{\vec{0}\}$. Seja W um subespaço de V. Se $W = \{\vec{0}\}$, então $W^{\perp} = V$ e portanto $V = \{0\} \oplus V$. Seja então $W \neq \{0\}$. Como W é um subespaço de dimensão finita maior do que 0, segue que W possui uma base ortonormal $\mathcal{B} := \{w_1, w_2, \ldots, w_k\}$. Como $\mathcal{B}$ é LI, existe uma base de V que contém $\mathcal{B}$, digamos $\{w_1, w_2, \ldots, w_k, v_1, v_2 \ldots, v_n\}$. Aplicando o processo de ortogonalização de Gram-Schmidt, obtemos $\mathcal{C} := \{u_1, u_2, \ldots, u_{k+n}\}$, com

$$\begin{aligned} u_1 &:= w_1, \\ u_2 &:= w_2 - \tfrac{\langle w_2, u_1\rangle}{||u_1||^2} u_1 = w_2 \\ &\vdots \\ u_k &:= w_k - \textstyle\sum_{i=1}^{k-1} \tfrac{\langle w_k, u_i\rangle}{||u_i||^2} u_i = w_k, \end{aligned}$$

pois $\langle w_j, w_i\rangle = 0$, para todo $j \neq i$. Verifiquemos que $W^{\perp} = [u_{k+1}, \ldots, u_{k+n}]$. De fato, como $\langle u_{k+1}, w_i\rangle = 0$, para todo $i = 1, \ldots, k$, pois $\mathcal{C}$ é ortogonal, então pela proposição anterior segue que $u_{k+1} \in W^{\perp}$. Da mesma forma, $u_{k+2}, \ldots, u_{k+n} \in W^{\perp}$, donde, como $W^{\perp}$ é um subespaço, segue que $[u_{k+1}, \ldots, u_{k+n}] \subset W^{\perp}$. Mas se $w \in W^{\perp}$, então sabemos que $w = \sum_{i=1}^{k+n} \frac{\langle w, u_i\rangle}{||u_i||^2} u_i$, pois em particular $w \in [u_1, \ldots, u_{k+n}] = [\mathcal{C}]$ e $\mathcal{C}$ é um conjunto ortogonal formado por vetores não nulos. Daí, desde que, $\langle w, w_i \rangle = 0$, pois $w \in W^\perp$ e $w_i \in W$, para todo $i = 1, \ldots, k$, então $w = \sum_{i=k+1}^{k+n} \frac{\langle w, u_i \rangle}{||u_i||^2} u_i$ e portanto $w \in [u_{k+1}, \ldots, u_{k+n}]$. Logo, $W^\perp = [u_{k+1}, \ldots, u_{k+n}]$. Finalmente, notemos que $V = W + W^\perp$ pois $\mathcal{C}$ gera V, e $W \cap W^\perp = \{\vec{0}\}$ pois $\mathcal{C}$ é LI. Portanto, $V = W \oplus W^\perp$. ■

Observação 0.234. *Seja V um espaço vetorial com produto interno $\langle \, , \, \rangle$ de dimensão finita. Se $\mathcal{S}$ é um subconjunto de V, então, pelos resultados anteriores, $V = [\mathcal{S}] \oplus \mathcal{S}^\perp$, pois $V = [\mathcal{S}] \oplus [\mathcal{S}]^\perp$ e $[\mathcal{S}]^\perp = \mathcal{S}^\perp$.*

Exemplo 0.235. *Consideremos $\mathbb{R}^2$ com o produto interno canônico $\langle \, , \, \rangle$. Seja $\mathcal{S} := \{(11, 13)\}$. Calculemos $\mathcal{S}^\perp$. Se $(x, y) \in \mathcal{S}^\perp$, então $\langle (x, y), (11, 13) \rangle = 0$, isto é, $11x + 13y = 0$. Logo, $y = -\frac{11}{13}x$, donde $\mathcal{S}^\perp = \{(x, -\frac{11}{13}x) | x \in \mathbb{R}\} = [(1, -\frac{11}{13})]$. Pela observação anterior, $\mathbb{R}^2 = [\mathcal{S}] \oplus \mathcal{S}^\perp = [(11, 13)] \oplus [(1, -\frac{11}{13})]$.*

Corolário 0.236. *Seja V um espaço vetorial com produto interno $\langle \, , \, \rangle$ de dimensão finita. Se W é um subespaço de V, então $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W + \dim_{\mathbb{R}} W^\perp$.*

Com efeito, seja W um subespaço de V. Então, pelo resultado anterior, $V = W \oplus W^\perp$. Daí, $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W + \dim_{\mathbb{R}} W^\perp - \dim_{\mathbb{R}}(W \cap W^\perp) = \dim_{\mathbb{R}} W + \dim_{\mathbb{R}} W^\perp - 0 = \dim_{\mathbb{R}} W + \dim_{\mathbb{R}} W^\perp$. ■

Programa de Verão 2012

04/01/2012 a 17/02/2012

B.6 - Álgebra Linear (turma 2)

2ª a 6ª das 19h às 21h - sala B-16

Gustavo de Lima Prado - glprado@ime.usp.br - www.ime.usp.br/ glprado

Programa: Vetores no $\mathbb{R}^n$. Espaços vetoriais e subespaços. Transformações lineares e matrizes. Semelhança e Diagonalização. Determinantes. Produto interno e ortogonalidade.

Pré-requisitos: 1 a 2 anos de graduação em Ciências Exatas.

Público: Alunos de graduação em Ciências Exatas.

Carga Horária: 120h

Bibliografia:

F. U. Coelho e M. L. Lourenço, Um curso de Álgebra Linear, EDUSP, 2010;

C. A. Callioli, H. H. Domingues e R. C. F. Costa, Álgebra L. e Aplicações, Editora Atual, 1998;

K. Hoffman e R. Kunze, Álgebra Linear, LTC, 1979;

A. Howard e R. C. Busby, Álgebra L. Contemporânea, Editora Bookman, 2006.