Universidade de Brasília
Departamento de Matemática

Álgebra Linear II - Notas de Aula

Profa. Aline Pinto (MAT/UnB)

Conteúdo

Prefácio

Estas notas de aula foram escritas para o curso de Álgebra Linear II, ministrado por mim na L Escola de Verão do MAT/UnB, e foram baseadas no livro [1], de Flávio U. Coelho e Mary L. Lourenço.

Embora a teoria tenha sido guiada pela contida no livro [1], exemplos e explicações adicionais da teoria desenvolvida foram acrescentados por mim. Também, a disposição dos temas incluídos nestas notas aparecem em alguns momentos em ordem distinta da do livro, por julgamento pessoal de melhor enquadramento para a disciplina. Além disso, o livro [2], de K. Hoffman e R. Kunze, esteve sempre em mãos durante a elaboração destas notas para consulta. Por fim, uma abordagem pessoal do conteúdo contruída ao longo de vários semestres em que ministrei Álgebra Linear se somam à motivação para escrever estas notas. Desejo boa leitura e aprendizado.

"E a beleza do lugar, pra se entender
Tem que se achar
Que a vida não é só isso que se vê
É um pouco mais..."

(Paulinho da Viola)

Dezembro de 2021

1 Preliminares

1.1 Corpos

Definição 1.1. *Um conjunto não vazio K é um **corpo** se existem duas operações*

$$\begin{array}{rccc} + : & K \times K & \to & K \\ & (a,b) & \mapsto & a+b \end{array} \qquad e \qquad \begin{array}{rccc} \cdot : & K \times K & \to & K \\ & (a,b) & \mapsto & ab, \end{array}$$

que chamamos de adição e multiplicação, respectivamente, tais que

A1) $a + b = b + a,\ \forall a, b \in K$;

A2) $a + (b + c) = (a + b) + c,\ \forall a, b, c \in K$;

A3) $\exists\ 0 \in K$ *tal que* $a + 0 = 0,\ \forall a \in K$;

A4) $\forall a \in K,\ \exists - a \in K$ *tal que* $a + (-a) = 0$;

M1) $ab = ba,\ \forall a, b \in K$;

M2) $(ab)c = a(bc),\ \forall a, b, c \in K$;

M3) $\exists\ 1 \in K$ *tal que* $1a = a,\ \forall a \in K$;

M4) $\forall\ a \in K$, *com* $a \neq 0$, $\exists\ a^{-1} \in K$ *tal que* $aa^{-1} = 1$;

M5) $(a+b)c = ac + bc$, $\forall a, b, c \in K$.

Exemplo 1.2. *Os conjuntos* $\mathbb{Q}$, $\mathbb{R}$ *e* $\mathbb{C}$, *com a adição e multiplicação usuais, são corpos;*

O conjunto dos inteiros módulo p, com p primo, $\mathbb{Z}_p = \{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{p-1}\}$ *com a adição e multiplicação de classes é um corpo;*

O conjunto dos números inteiros $\mathbb{Z}$ ***não*** *é corpo, pois não existe* $n \in \mathbb{Z}$ *tal que* $2n = 1$ $(\frac{1}{2} \notin \mathbb{Z})$, *o que contraria a propriedade M4.*

Nomenclatura: Os elementos de um corpo K serão chamados de **escalares**.

Note que $\forall x \in \mathbb{R}$, $x \neq 0$, temos $x^{-1} = \frac{1}{x}$.
E para $z = a + bi \in \mathbb{C}$, $z \neq 0$, temos $z^{-1} = \frac{1}{z} = x + iy$, com $x, y \in \mathbb{R}$.

Exercício. Encontre x e y em função de a e b.

Lembre que $zz^{-1} = 1$ e, para $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, com $x_1, x_2, y_1, y_2 \in \mathbb{R}$, temos $z_1 = z_2$ se, e somente se, $x_1 = x_2$ e $y_1 = y_2$.

Solução:

1.2 Sistemas Lineares

Seja K um corpo. Considere o problema de determinar n escalares (i.e. elementos de K) $x_1, x_2, \ldots, x_n$ que satisfaçam as condições

$$S : \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = & b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = & b_m \end{cases}$$

onde b_i, a_{ij}, $1 \leq i \leq m$ e $1 \leq j \leq n$, são elementos de K dados.

Chamamos S de um **sistema linear sobre** K **de** m **equações e** n **incógnitas** $x_1, x_2, \ldots, x_n$. Se $m = n$, dizemos que S é um **sistema linear de ordem** n. Se $b_1 = b_2 = \cdots = b_m = 0$, dizemos que S é um **sistema linear homogêneo**. Os escalares a_{ij}, $1 \leq i \leq m$ e $1 \leq j \leq n$, são chamados de **coeficientes de** S e $b_1, b_2, \ldots, b_m$ de **termos independentes** de S.

Uma n-upla $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ de elementos em K é uma **solução de** S quando ao tomarmos $x_1 = \alpha_1, x_2 = \alpha_2, \ldots, x_n = \alpha_n$, todas as esquações de S são satisfeitas. O **conjunto-solução** de S é o conjunto formado por todas as suas soluções.

Notação: Denotamos a i-ésima equação de S por L_i.

Exemplo 1.3. *Considere o sistema linear sobre* $\mathbb{R}$

$$S : \begin{cases} 2x & -\ y & +z & = 1 \\ x & +2y & & = 6 \end{cases}$$

O Conjunto Solução de S *é*

$$\{(6-2y, y, -11+5y) \mid y \in \mathbb{R}\},$$

pois para cada $y \in \mathbb{R}$*, pela equação* L_2 *obtemos* $x = 6 - 2y$ *e,com isso, pela equação* L_1 *obtemos* $z = -11 + 5y$*. Neste caso,* S *é um sistemas linear que possui infinitas soluções.*

Exemplo 1.4. *O sistema linear sobre* $\mathbb{R}$

$$S : \begin{cases} x_1 & -2x_2 & +\ x_3 & = 1 \\ & x_2 & -\ x_3 & = 2 \\ & & 2x_3 & = 2 \end{cases}$$

tem solução única dada por $(6, 3, 1)$*.*

Exemplo 1.5. *O sistema linear sobre* $\mathbb{C}$

$$S : \begin{cases} 2x & +3i\ y & = 2+i \\ (1+i)x & +\quad y & = 3i \end{cases}$$

tem solução única dada por $(\frac{11+10i}{17}, \frac{-1-4i}{17})$*.*

Exemplo 1.6. *Um sistema linear sobre* $\mathbb{R}$ *ou* $\mathbb{C}$ *do tipo*

$$S : \begin{cases} a_{11}x_1 & +a_{12}x_2 & +\cdots & +a_{1n}x_n & = b_1 \\ \cdots \\ 0.x_1 & +\ 0.x_2 & +\cdots & +\ 0.x_n & = b_i \\ \cdots \end{cases}$$

com $b_i \neq 0$ *não possui solução, pois nenhuma* n*-upla satisfaz a equação* L_i*.*

Exemplo 1.7. (Sistemas Lineares de Ordem 2 sobre $\mathbb{R}$)

Considere o seguinte sistema linear de ordem 2 sobre $\mathbb{R}$

$$S : \begin{cases} a_1x & + & b_1y & = & c_1 \\ a_2x & + & b_2y & = & c_2 \end{cases}$$

onde $(a_i, b_i) \neq (0,0),\ i = 1, 2$*. Cada equação* L_1 *e* L_2 *de* S *determina uma reta no plano-xy e suas soluções são os pontos que pertencem às retas. As posições relativas destas duas retas são:*

- L_1 *e* L_2 *se interceptam em um único ponto: neste caso* S *tem uma* ***única solução****, pois um único par* (x, y) *satizfaz as duas equações;*
- L_1 *e* L_2 *são paralelas: neste caso* S ***não possui solução****, pois não existe* (x, y) *que satisfaça as duas equações simultaneamente.*
- L_1 *e* L_2 *são coincidentes: neste caso* S *tem* ***infinitas soluções****, pois todos os pontos* (x, y) *na reta de* L_1 *são também solução da equação de* L_2*.*

Veremos que o mesmo acontece em qualquer sistema linear sobre K *com* m *equações e* n *incógnitas, ou seja, que todo sistema linear possui solução única, não possui solução ou admite infinitas soluções.*

1.3 Operações Elementares

Exemplo 1.8. Considere o seguinte sistema linear sobre $\mathbb{R}$.

$$S : \left\{ \begin{array}{rcrcrcr} 2x & + & y & + & 2z & = & 0 \\ x & - & y & + & z & = & 1 \\ 3x & - & y & + & z & = & 1 \end{array} \right.$$

Vamos fazer operações com as equações de S da seguinte maneira. Primeiro vamos permutar L_1 e L_2:

$$S \quad \sim^{L_1 \leftrightarrow L_2} \left\{ \begin{array}{rcrcrcr} x & - & y & + & z & = & 1 \\ 2x & + & y & + & 2z & = & 0 \\ 3x & - & y & + & z & = & 1 \end{array} \right.$$

Agora, no sistema resultante, vamos zerar o coeficiente de x nas equações L_2 e L_3 somando tais equações com múltiplos convenientes da equação L_1:

$$S \quad \sim^{L_1 \leftrightarrow L_2} \left\{ \begin{array}{rcrcrcr} x & - & y & + & z & = & 1 \\ 2x & + & y & + & 2z & = & 0 \\ 3x & - & y & + & z & = & 1 \end{array} \right. \sim^{L_2 \to L_2 - 2L_1}_{L_3 \to L_3 - 3L_1} \left\{ \begin{array}{rcrcrcr} x & - & y & + & z & = & 1 \\ & & 3y & & & = & -2 \\ & & 2y & - & 2z & = & -2 \end{array} \right.$$

No sistema resultante, vamos multiplicar a equação L_3 por $\frac{1}{2}$ ($L_3 \to \frac{1}{2}L_3$). Com isso, obtemos o sistema

$$S_1 : \left\{ \begin{array}{rcrcrcr} x & - & y & + & z & = & 1 \\ & & 3y & & & = & -2 \\ & & y & - & z & = & -1 \end{array} \right.$$

que tem solução única dada por $(0, -\frac{2}{3}, \frac{1}{3})$.

No exemplo acima, fazendo operações com as equações de S, chegamos ao sistema S_1, um sistema mais simples e encontramos sua solução. Veremos que a solução encontrada do sistema S_1 é exatamente a solução do sistema original S. Para isso, vamos definir as as operações que podemos fazer com as equações um sistema linear arbitrário de forma que o conjunto solução não seja alterado.

Definição 1.9. *Seja S um sistema linear com m equações e n incógnitas sobre K. Vamos considerar os sistemas obtidos a partir de S quando aplicamos as seguintes operações:*

(I) *permutar duas linhas de S ($L_i \leftrightarrow L_j$);*

(II) *multiplicar uma das equações de S por um escalar $k \neq 0$ ($L_i \to kL_i$);*

(III) *substuir uma das equações de S pela soma dessa equação e uma outra equação multiplicada por um escalar $k \neq 0$ ($L_i \to L_i + kL_j$).*

*Estas operações são chamadas **elementares**. Se um sistema S_1 foi obtido a partir de S aplicando-se de um número finito de operações elementares, dizemos que S_1 é **equivalente** a S, e escrevemos $S \sim S_1$. (Obs.: Esta é uma relação de equivalência.)*

Teorema 1.10. *Sistemas lineares equivalentes possuem o mesmo conjunto-solução.*

Demonstração.

Seja S_1 um sistema linear obtido a partir de S aplicando-se uma operação elementar, onde

$$S : \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = & b_1 \\ \vdots & \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = & b_i \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = & b_m \end{cases}.$$

Seja $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ uma solução de S_1. Vamos mostrar que α é também solução de S.

Se S_1 foi obtido ao aplicarmos a operação elementar (I), então é imediato que α também é solução de S, uma vez que as equações foram apenas permutadas.

Se S_1 foi obtido ao aplicarmos a operação elementar (II), com $L_i \to kL_i$ onde $k \neq 0$, como α é solução de S_1 temos que

$$S : \begin{cases} a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1n}\alpha_n = & b_1 \\ \vdots & \vdots \\ ka_{i1}\alpha_1 + ka_{i2}\alpha_2 + \cdots + ka_{in}\alpha_n = & kb_i \\ \vdots & \vdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \cdots + a_{mn}\alpha_n = & b_m \end{cases}$$

Como as equações $L_1, \ldots, L_{i-1}, L_{i+1}, \ldots, L_m$ de S e S_1 são iguais, α é solução destas equações e, $k \neq 0$, temos

$$\begin{aligned} & ka_{i1}\alpha_1 + ka_{i2}\alpha_2 + \cdots + ka_{in}\alpha_n = kb_i \\ \Rightarrow\ & k(a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n) = kb_i \\ \overset{k\neq 0}{\Rightarrow}\ & a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n = b_i \end{aligned}$$

e assim α também é solução da equação L_i de S. Logo α é solução de S.

Se S_1 foi obtido ao aplicarmos a operação elementar (III), com $L_i \to L_i + kL_j$ onde $k \neq 0$ e $j \neq i$, como α é solução de S_1 e as equações $L_1, \ldots, L_{i-1}, L_{i+1}, \ldots, L_m$ de S e S_1 são iguais, α é solução destas equações. Além disso, α satisfaz a equação L_i de S_1, logo

$$(a_{i1} + ka_{j1})\alpha_1 + (a_{i2} + ka_{j2})\alpha_2 + \cdots + (a_{in} + ka_{jn})\alpha_n = b_i + kb_j.$$

Assim

$$a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n + k(a_{j1}\alpha_1 + a_{j2}\alpha_2 + \cdots + a_{jn}\alpha_n = b_i + kb_j.$$

Como α é solução de L_j, temos $a_{j1}\alpha_1 + a_{j2}\alpha_2 + \cdots + a_{jn}\alpha_n = b_j$. Então

$$a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n + kb_j = b_i + kb_j$$

e assim

$$a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n = b_i.$$

Com isso, α também é solução da equação L_i de S e, portanto, α é solução de S.

Assim, toda solução de S_1 é também solução de S. Agora, como a operação inversa de uma operação elementar é também uma operação elementar, segue que as soluções de S e S_1 são exatamente as mesmas.

■

Exemplo 1.11. *Considere o seguinte sistema linear sobre* $\mathbb{R}$

$$S : \begin{cases} x & -y & +z & +t & = 1 \\ 2x & +y & +2z & -2t & = 0 \\ 3x & -y & +z & -t & = 1 \\ 3x & & +3z & -t & = 1 \end{cases}$$

Temos

$$S \sim^{L_2 \to L_2 - 2L_1} \begin{cases} x & -y & +z & +t & = 1 \\ & 3y & & -4t & = -2 \\ 3x & -y & +z & -t & = 1 \\ 3x & & +3z & -t & = 1 \end{cases} \sim^{L_3 \to L_3 - 3L_1}_{L_4 \to L_4 - 3L_1} \begin{cases} x & -y & +z & +t & = 1 \\ & 3y & & -4t & = -2 \\ & 2y & -2z & -4t & = -2 \\ & 3y & & -4t & = -2 \end{cases}$$

$$\sim^{L_3 \to \frac{1}{2} L_3}_{L_4 \to L_4 - L_2} \begin{cases} x & -y & +z & +t & = 1 \\ & 3y & & -4t & = -2 \\ & y & -z & -2t & = -1 \\ & & & 0 & = 0 \end{cases} \sim^{L_2 \leftrightarrow L_3} \begin{cases} x & -y & +z & +t & = 1 \\ & y & -z & -2t & = -1 \\ & 3y & & -4t & = -2 \\ & & & 0 & = 0 \end{cases}$$

$$\sim^{L_1 \to L_1 + L_2}_{L_3 \to L_3 - 3L_2} \begin{cases} x & & & -t & = 0 \\ & y & -z & -2t & = -1 \\ & & 3z & +2t & = 1 \\ & & & 0 & = 0 \end{cases} \sim^{L_3 \to \frac{1}{3} L_3} \begin{cases} x & & & -t & = 0 \\ & y & -z & -2t & = -1 \\ & & z & +\frac{2}{3}t & = \frac{1}{3} \\ & & & 0 & = 0 \end{cases}$$

$$\sim^{L_2 \to L_2 + L_3} \begin{cases} x & & & -t & = 0 \\ & y & & -\frac{4}{3}t & = -\frac{2}{3} \\ & & z & +\frac{2}{3}t & = \frac{1}{3} \\ & & & 0 & = 0 \end{cases}$$

Com isso, vemos que S *tem infinitas soluções e seu conjunto-solução é dado por*

$$\left\{ \left(t,\ -\frac{2}{3} + \frac{4}{3}t,\ \frac{1}{3} - \frac{2}{3}t,\ t \right) \ \mid t \in \mathbb{R} \right\}.$$

O último sistema acima, que foi obtido a partir de S *e que nos deu o conjunto-solução de* S*, é chamado de forma escalonada de* S*, como veremos a seguir.*

1.4 A Forma Escalonada de um Sistema Linear

Definição 1.12. *Um sistema linear* S *está na* ***forma escalonada*** *quando*

i) o primeiro coeficiente não nulo de cada equação não nula é igual a 1;

ii) se uma equação possui o primeiro coeficiente não nulo na incógnita x_i*, então em todas as outras equações os coeficientes de* x_i *são iguais a zero;*

iii) se as equações $L_1, \ldots, L_r$ *são as equações não nulas de* S *e se o primeiro coeficiente não nulo da equação* L_i *ocorre na incógnita* x_{k_i}*, então* $k_1 < k_2 < \cdots < k_r$*.*

Exemplo 1.13. *Os sitemas lineares*

$$S_1 : \begin{cases} x_1 & & & = 2 \\ & x_2 & -x_3 & = 1 \\ & & x_3 & = -1 \end{cases} \qquad e \qquad S_2 : \begin{cases} & x_2 & & = 1 \\ x_1 & & -x_3 & = -3 \end{cases}$$

não estão *na forma escalonada (Por quê? Quais são as forma escalonadas destes sistemas?).*

Em S_1, a equação L_3 possui o primeiro coeficiente não nulo na incógnita x_3, mas x_3 na equação L_2 não é igual a zero, o que fura o item (ii) da definição. Encontre a forma escalonada:
Em S_2, o item (iii) da definição não está satisfeito. Encontre a forma escalonada:

Os sistemas lineares

$$S_3 : \begin{cases} x_1 & -x_2 & & & = 1 \\ & & x_3 & & = 4 \\ & & & x_4 & = -2 \\ & & & 0 & = 0 \end{cases} \qquad e \qquad S_4 : \begin{cases} x_1 & & & = 2 \\ & x_2 & & = -1 \\ & & x_3 & = 5 \\ & & 0 & = 2 \end{cases}$$

estão *na forma escalonada (Verifique que os itens i), ii) e iii) da definição estão satisfeitos).*

Nota: Todo sistema linear com m equações e n incógnitas sobre K, para quaisquer $m, n > 0$, é equivalente a um único sistema linear na forma escalonada.

Após o processo de escalonamento de um sitema linear S, ao retirarmos as equações do tipo $0 = 0$, chegaremos a um dos seguintes casos:

1°) O sistema escalonado possui uma equação do tipo $0x_1 + 0x_2 + \cdots + 0x_n = b_r$, com $b_r \neq 0$. Neste caso, S **não possui solução**.

Exemplo: Sobre $\mathbb{R}$,

$$S : \begin{cases} x & +y & +z & = 1 \\ x & -y & -z & = 2 \\ 2x & +y & +z & = 3 \end{cases} \sim \begin{cases} x & & & = 2 \\ & y & +z & = -1 \\ & & 0 & = -1 \end{cases}$$

Note que quando aparecem equações do tipo $0 = b$, com $b \neq 0$, no processo de escalonamento do sistema, isso significa que o sistema possui equações incompatíveis entre si.

2°) Não acontece o 1° caso e o número de equações no sistema escalonado é igual ao número de incógnitas. Neste caso o sistema tem **solução única**.

Exemplo: Sobre $\mathbb{R}$,

$$S: \left\{ \begin{array}{rrrl} x & +2y & -4z & =-4 \\ & y & -z & =-2 \\ & & 7z & =7 \end{array} \right. \sim \left\{ \begin{array}{rrrl} x & & & =2 \\ & y & & =-1 \\ & & z & =1 \end{array} \right.$$

3°) Não acontece o 1° caso e o número de equações no sistema escalonado é menor do que o número de incógnitas. Neste caso o sistema tem **infinitas soluções**.

Exemplo: Sobre $\mathbb{R}$,

$$S: \left\{ \begin{array}{rrrl} x & +2y & +z & =-1 \\ & -5y & +z & =2 \\ -2x & -9y & -z & =4 \end{array} \right. \sim \left\{ \begin{array}{rrrl} x & & -\frac{7}{5}z & =-\frac{1}{5} \\ & y & -\frac{1}{5}z & =-\frac{2}{5} \end{array} \right.$$

Tomando z como variável livre, o conjunto solução de S é $\{(\frac{1}{5}+\frac{7}{5}z, -\frac{2}{5}+\frac{1}{5}z, z) \mid z \in \mathbb{R}\}$.

$$\boxed{\text{n}^{\underline{\text{o}}} \text{ de variáveis livres} = \text{n}^{\underline{\text{o}}} \text{ de incógnitas} - \text{n}^{\underline{\text{o}}} \text{ de equações}}$$

Observação: O processo de escalonamento garante que um sistema escalonado que não está no 1° não pode ter mais equações do que incógnitas (Explique!).

Estudando os casos possíveis para a forma escalonada de um sistema, vemos então que **um sistema linear qualquer não tem solução, tem solução única ou possui infinitas soluções.**

Exercício 1.14. *Determine para quais valores de $a \in K$ o sistema linear*

$$\left\{ \begin{array}{rrrl} x & +2y & -3z & =4 \\ 3x & -y & +5z & =2 \\ 4x & +y & +(a^2-14)z & =a+2 \end{array} \right.$$

i) não tem solução;

ii) tem solução única. Determine a solução neste caso;

iii) admite infinitas soluções. Determine o conjunto solução neste caso.

1.5 Sistemas Lineares Homogêneos

Sistemas homogêneos são sistemas lineares onde os termos independentes são todos nulos.

$$S: \left\{ \begin{array}{rl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = & 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = & 0 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = & 0 \end{array} \right.$$

Note que um sistema homogêneo sempre possui solução, a solução nula $x_1 = 0, x_2 = 0, \dots, x_n = 0$. Então S ou possui solução única (que é a solução nula) ou admite infinitas soluções (e portanto tem solução diferente de zero). Se **após o escalonamento** de S

- o número de equações é igual ao número de incógnitas, então S tem solução única (que é a solução nula $x_1 = 0, x_2 = 0, \dots, x_n = 0$);
- o número de equações é menor do que o número de incógnitas, então S possui infinitas soluções (e portanto tem solução não nula).

Exercício 1.15. *Seja S um sistema linear homogêneo com m equações e n incógnitas sobre K, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ e $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ soluções de S e $k \in K$. Defina*

$$\alpha + \beta := (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n) \quad e \quad k\alpha := (k\alpha_1, k\alpha_2, \dots, k\alpha_n).$$

Mostre que $\alpha + \beta$ e $k\alpha$ são também soluções de S (ou seja, soma de soluções e múltiplo escalar de soluções são também soluções). Isso é verdade se S não é homogêneo?

1.6 Sistemas Lineares e Matrizes

Considere um sistema linear

$$S : \left\{ \begin{array}{rl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = & b_2 \\ \vdots \qquad\qquad & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = & b_m \end{array} \right.$$

sobre K com m equações e n incógnitas $x_1, x_2, \dots, x_n$. Podemos reescrever S na forma matricial

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

ou, simplesmente,

$$\boldsymbol{AX{=}B},$$

onde

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m\times n}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{n\times 1} \quad \text{e} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}_{m\times 1}.$$

Chamamos A de **matriz dos coeficientes** de S, X de **matriz das incógnitas** e B de **matriz dos termos independentes** S.

A partir de propriedades da matriz A dos coeficientes de S é possível estudar o conjunto solução de S, como veremos mais adiante. Para tanto, iremos estudar um pouco as matrizes sobre K e suas propriedades. No decorrer desse estudo de matrizes, faremos inferências sobre as consequências para sistemas lineares.

1.7 Matrizes

Notação: Representamos por $\mathcal{M}_{m\times n}(K)$ o conjunto

$$\mathcal{M}_{m\times n}(K) = \{A = (a_{i,j})_{m\times n} \mid a_{ij} \in K, \forall i, j\}$$

de todas as as matrizes $m \times n$ com entradas no corpo K.

Por exemplo, $\mathcal{M}_{2\times 3}(\mathbb{R})$ é o conjunto de todas as matrizes 2×3 com entradas reais. Temos

$$A = \begin{pmatrix} 2 & -1 & 5 \\ 1 & 0 & \frac{1}{2} \end{pmatrix} \in \mathcal{M}_{2\times 3}(\mathbb{R}).$$

1.8 Operações com matrizes

• **Adição:** Dadas $A = (a_{ij})$, $B = (b_{ij}) \in \mathcal{M}_{m\times n}(K)$, a adição de A e B é a matriz em $\mathcal{M}_{m\times n}(K)$ definida por

$$A + B := (a_{ij} + b_{ij}).$$

• **Multiplicação por escalar:** Dada $A = (a_{ij}) \in \mathcal{M}_{m\times n}(K)$ e $r \in K$, a multiplicação de A pelo escalar r é a matriz em $\mathcal{M}_{m\times n}(K)$ definida por

$$rA := (ra_{ij}).$$

Propriedades 1.16. *Para quaisquer $r, s \in K$ e $A, B, C \in \mathcal{M}_{m\times n}(K)$, temos*

A1) $A + B = B + A$;

A2) $A + (B + C) = (A + B) + C$;

A3) $A + 0 = A$, *onde* $0 = (0)_{m\times N}$ *é a matriz nula (i.e. com todas as entradas iguais a zero);*

A4) $A + (-A) = 0$ *(matriz nula), onde se* $A = (a_{ij})$ *então* $-A = (-a_{ij})$;

M1) $(rs)A = r(sA)$;

M2) $(r + s)A = rA + sA$;

M3) $r(A + B) = rA + rB$;

M4) $1A = A$ *(onde* $1 \in K$*).*

Obs.: As propriedades acima nos dizem que $(\mathcal{M}_{m\times n}(K), +, \cdot)$ é um espaço vetorial sobre K, como veremos mais adiante.

Demonstração. Vamos provar algumas das propriedades e as demais ficam como exercício. Sejam $A = (a_{ij}), B = (b_{ij})$ e $C = (c_{ij})$. Então

A1) $A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = (b_{ij}) + (a_{ij}) = B + A$;

A3) $A + 0 = (a_{ij}) + (0) = (a_{ij} + 0) = (a_{ij}) = A$;

M2) $(r+s)A = (r+s).(a_{ij}) = ((r+s)a_{ij}) = (ra_{ij} + sa_{ij}) = (ra_{ij}) + (sa_{ij}) = r(a_{ij}) + s(a_{ij}) = rA + sA$.

■

• **Multiplicação de matrizes:** Sejam $A = (a_{ij})_{m\times n}$ e $B = (b_{jk})_{n\times p}$ matrizes sobre K, onde o número n de colunas de A é igual ao número de linhas de B, e m, p são inteiros positivos quaisquer. Definimos a multiplicação de A por B pela matriz

$$AB := (c_{ik})_{m\times p},$$

onde c_{ik} é obtido ao multiplicarmos as entradas da linha i de A pelas entradas da coluna k de B e somarmos, ou seja,

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}.$$

Por exemplo, veja a entrada c_{11}:

$$AB = \begin{pmatrix} \underline{a_{11}} & \underline{a_{12}} & \cdots & \underline{a_{1n}} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} \underline{b_{11}} & b_{12} & \cdots & b_{1p} \\ \underline{b_{21}} & b_{22} & \cdots & b_{2p} \\ \vdots & & & \vdots \\ \underline{b_{n1}} & b_{n2} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & \cdots \\ \cdots & \ddots \end{pmatrix}_{m\times p}$$

$$= \begin{pmatrix} \sum_{j=1}^{n} a_{1j}b_{j1} & \cdots \\ \cdots & \ddots \end{pmatrix}_{m\times p}$$

Propriedades 1.17. *Para quaisquer matrizes A, B e C sobre K para as quais os produtos estão definidos, temos*

1) $(AB)C = A(BC)$ *(a multiplicação de matrizes é associativa);*

2) $A(B + C) = AB + AC$ *(vale a lei disbributiva à esquerda);*

3) $(A + B)C = AC + AB$ *(vale a lei disbributiva à direita).*

Demonstração. Pela forma como o produto de matrizes é definido, não é óbvio que tais propriedades são válidas. Vamos demonstrar a propriedade 1), que diz que o produto de matrizes é associativo (que é a propriedade mais difícil de provar entre as três enunciadas). A prova das leis distibutivas ficam como **exercício**.

1) Sejam $A = (a_{ij})_{m\times n}$, $B = (b_{jk})_{n\times p}$ e $C = (c_{kl})_{p\times q}$. Então

$$AB = (d_{ik})_{m\times p}, \quad \text{onde} \quad d_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$$

e

$$BC = (e_{jl})_{n\times q}, \quad \text{onde} \quad e_{jl} = \sum_{k=1}^{p} b_{jk}c_{kl}.$$

Assim,

$$(AB)C = (d_{ik})_{m\times p}(c_{kl})_{p\times q} = (f_{il})_{m\times q}, \quad \text{onde} \quad f_{il} = \sum_{k=1}^{p} d_{ik}c_{kl}$$

e

$$A(BC) = (a_{ij})_{m\times n}(e_{jl})_{n\times q} = (g_{il})_{m\times q}, \quad \text{onde} \quad g_{il} = \sum_{j=1}^{n} a_{ij}e_{jl}.$$

Para mostar que $(AB)C = A(BC)$, temos que mostrar que $f_{il} = g_{il}$, para todos i, l. De fato

$$f_{il} = \sum_{k=1}^{p} d_{ik}c_{kl} = \sum_{k=1}^{p}(\sum_{j=1}^{n} a_{ij}b_{jk})c_{kl} = \sum_{k=1}^{p}\sum_{j=1}^{n} a_{ij}b_{jk}c_{kl}$$

$$\sum_{j=1}^{n}\sum_{k=1}^{p} a_{ij}b_{jk}c_{kl} = \sum_{j=1}^{n} a_{ij}(\sum_{k=1}^{p} b_{jk}c_{kl}) = \sum_{j=1}^{n} a_{ij}e_{jl}) = g_{il}.$$

■

Observação 1.18. *AB pode ser diferente de BA (mesmo que os dois produtos estejam definidos).*

Exercícios: Para mostrar que uma afirmação é falsa, basta dar um *contra-exemplo*, ou seja, um exemplo mostrando que a afirmação não é válida.

1) Mostre que $AB = BA$ não é verdadeiro (mesmo quando os produtos estão definidos).

2) **V ou F:** $(A+B)^2 = A^2 + 2AB + B^2$

1.9 Matrizes inversíveis

Definição 1.19. *A matriz*

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}_{n\times n}$$

*é chamada de **matriz identidade** $n \times n$ (de ordem n).*

Em particular,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

são a matrizes identidade de ordens 2, 3 e 4, respectivamente.

I_n tem a seguinte propriedade

$$I_n A = A \quad \text{e} \quad BI_n = B$$

sempre que os produtos estão definidos. Em particular, $I_n A = A = AI_n$, para toda matriz $A \in \mathcal{M}_{n\times n}(K)$.

Notação: *O conjunto $\mathcal{M}_{n\times n}(K)$ das matrizes quadradas $n \times n$ sobre K será denotado simplesmente por $\mathcal{M}_n(K)$.*

Definição 1.20. *Dizemos que uma matriz $A \in \mathcal{M}_n(K)$ é **inversível** se existe uma matriz $B \in \mathcal{M}_n(K)$ tal que*

$$AB = I_n \quad e \quad BA = I_n.$$

Neste caso, a matriz B é chamada de inversa de A.

Proposição 1.21. *Sobre matrizes inversíveis, valem as seguintes propriedades:*

i) *Seja $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz inversível. Então a matriz inversa de A é única.*

   Notação: *Quando A é inversível, denotamos então por A^{-1} sua (única) inversa. Com isso,*

$$AA^{-1} = I_n = A^{-1}A.$$

ii) *Seja $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz inversível. Então A^{-1} também é inversível e*

$$(A^{-1})^{-1} = A.$$

iii) *Sejam $A, B \in \mathcal{M}_n(\mathbb{R})$ matrizes inversíveis. Então AB também é inversível e*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Demonstração. i) Sejam C e D matrizes inversas de A, ou seja, tais que

$$AC = I_n = CA \quad \text{e} \quad AD = I_n = DA.$$

Então

$$C = CI_n = C(AD) = (CA)D = I_n D = D.$$

ii) Como $AA^{-1} = I_n = A^{-1}A$, temos que A^{-1}é invesível e sua inversa é A.

iii) Temos $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$.
Analogamente, $(B^{-1}A^{-1})(AB) = I_n$. Portanto AB é inversível e sua inversa é $B^{-1}A^{-1}$.

∎

Ainda não sabemos dizer se matrizes A são ou não inversíveis e nem determinar a inversa A^{-1} de matrizes inversíveis. Por isso, um exemplo agora seria bastante artificial. Ao longo dessa seção resolveremos isso. Entretanto somente com a definição de uma matriz inversível, já podemos tirar conclusões interessantes sobre sistemas lineares, como a seguir.

Observação 1.22. *Seja $AX = B$ um sistema linear de ordem n sobre K, onde*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}_{n\times n}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{n\times 1} \quad e \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}_{n\times 1}.$$

1) *Se A é inversível, então o sistema $AX = B$ tem solução única, dada por $X = A^{-1}B$. De fato,*

$$AX = B \Rightarrow A^{-1}AX = A^{-1}B \Rightarrow I_nX = A^{-1}B \Rightarrow X = A^{-1}B.$$

2) *Se A é inversível, então, como o sistema $AX = B$ tem solução única, o sistema escalonado equivalente tem número de equações igual ao número de incógnitas.*

$$S : \left\{ \begin{array}{rl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = & b_2 \\ \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = & b_n \end{array} \right. \sim \quad S_1 : \left\{ \begin{array}{llll} x_1 & & & = c_1 \\ & x_2 & & = c_2 \\ & & \ddots & \\ & & x_n & = c_n \end{array} \right.$$

Note que fazer operações elementares em S é o mesmo que fazer tais operações elementares com as linhas da matriz

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right), \quad \textit{chamada de matriz ampliada de } S.$$

Assim a matriz ampliada de S é então equivalente à matriz

$$\left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_n \end{array}\right), \quad \textit{que é a matriz ampliada de } S_1.$$

Segue então que, ao fazermos as operações elementares para o escalonamento do sistema $AX = B$ e observamos o que acontece com a matriz A, obtemos $A \sim I_n$, ou seja, A é equivalente à matriz identidade I_n (onde por matrizes equivalentes entende-se matrizes que podem ser obtidas uma da outra aplicando-se operações elementares com suas linhas).
Conclusão: *Se A é inversível, então $A \sim I_n$.*

Pergunta: Como saber se uma dada matriz é inversível? E se for inversível, como determinar sua inversa?

Exemplo 1.23. $A = \begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix}$ *não é inversível, pois*

$$\begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Leftrightarrow \left\{\begin{array}{rcrl} x & - & z & = 1 \\ y & - & t & = 0 \\ -3x & + & 3z & = 0 \\ -3y & + & 3t & = 1 \end{array}\right. \sim_{\substack{L_3 \to L_3 + 3L_1 \\ L_4 \to L_4 + 3L_2}} \left\{\begin{array}{rcrl} x & - & z & = 1 \\ y & - & t & = 0 \\ & & 0 & = 3 \\ & & 0 & = 4 \end{array}\right.$$

não tem solução.

Como vimos no exemplo acima, nem toda matriz é inversível. E na observação mais acima, vimos que se uma matriz A é inversível, então $A \sim I_n$. Vamos estudar melhor esse fato.

Definição 1.24. *Uma* ***matriz elementar*** *é uma matriz obtida a partir da matriz identidade ao aplicarmos uma (e somente uma) operação elementar.*

Exemplo 1.25. $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$I_3 \sim^{L_2 \to kL_2} E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad I_3 \sim^{L_3 \to L_3 + kL_1} E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{pmatrix}$$

E_1 e E_2 são matrizes elementares.

Exemplo 1.26. *Considere a matriz* $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix}$. *Temos*

$$A \overset{L_2 \leftrightarrow L_3}{\sim} A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 2 & 1 & 0 \end{pmatrix} \overset{L_2 \to \frac{1}{2}L_2}{\sim} A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix} \overset{L_3 \to L_3 - 2L_2}{\sim} A_3 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & -3 & -6 \end{pmatrix}.$$

Seja E_1 a matriz elementar obtida por $L_2 \leftrightarrow L_3$: $E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. *Note que*

$$E_1 A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 2 & 1 & 0 \end{pmatrix} = A_1.$$

Seja E_2 a matriz elementar obtida por $L_2 \to \frac{1}{2}L_2$: $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$. *Note que*

$$E_2 A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix} = A_2.$$

Seja E_3 a matriz elementar obtida por $L_3 \to L_3 - 2L_2$: $E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$. Note que

$$E_3A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & -3 & -6 \end{pmatrix} = A_3.$$

Com isso, obtemos

$$A_3 = E_3E_2E_1A.$$

O exemplo acima nos mostra o seguinte resultado.

Teorema 1.27. *Se A é uma matriz, o resultado de uma operação elementar sobre A é o mesmo resultado da multiplicação da matriz elementar E por A, onde E corresponde à operação elementar aplicada em A.*

Observação 1.28. *Note então que A e B são matrizes equivalentes se, e somente se,*

$$B = E_mE_{m-1}\cdots E_2E_1A,$$

ou seja, a partir de A fazemos uma sequência de operações elementares (que é igual a multiplicar à esquerda pelas matrizes elementares correspondentes) e chegamos em B. Então $A \sim I_n$ implica que

$$I_n = E_mE_{m-1}\cdots E_2E_1A,$$

Assim estamos muito perto de uma matriz ($= E_mE_{m-1}\cdots E_2E_1$) que multiplicada por A é igual a identidade. Mas ainda não temos a multiplicação à direita dando a identidade também.

Corolário 1.29. *Toda matriz elementar E é inversível e sua inversa é a matriz elementar E^{-1} que corresponde à operação elementar inversa da operação efetuada em E.*

Por exemplo, sejam $E = \begin{pmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ a matriz elementar obtida por $L_1 \to kL_1$ e $E^{-1} = \begin{pmatrix} \frac{1}{k} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ a matriz elementar obtida pela operação elementar inversa $L_1 \to \frac{1}{k}L_1$. Como multiplicar à esquerda por uma matriz elementar é aplicar a operação elementar correpondente na matriz, vemos que

$$E^{-1}E = I_3 \quad \text{e} \quad EE^{-1} = I_3.$$

Os resultados e observações acima sobre matrizes elementares nos levam ao seguinte critério para determinar se matrizes são inversíveis e, em caso afirmativo, encontrar sua inversa.

Teorema 1.30. *Uma matriz $A \in \mathcal{M}_n(K)$ é inversível se, e somente se, $A \sim I_n$. Neste caso, a mesma sucessão de operações elementares que transforma A em I_n transforma I_n em A^{-1}.*

Demonstração. ($\Rightarrow$) Já vimos que se A é inversível, então o sistema linear $AX = 0$ tem solução única. Assim o sistema linear escalonado equivalente tem número de equações igual ao número de incógnitas, de onde segue que A é equivalente à matriz identidade.

($\Leftarrow$) Suponha que $A \sim I_n$. Então

$$I_n = E_m E_{m-1} \cdots E_2 E_1 A,$$

onde $E_1,\ E_2, \ldots,\ E_m$ são matrizes elementares. Como cada matriz elementar é inversível, obtemos

$$\begin{aligned}
&E_m^{-1} I_n = E_m^{-1} E_m E_{m-1} \cdots E_2 E_1 A \\
&\Rightarrow E_m^{-1} = E_{m-1} \cdots E_2 E_1 A \\
&\Rightarrow E_{m-1}^{-1} E_m^{-1} I_n = E_{m-2} \cdots E_2 E_1 A \\
&\quad \vdots \\
&\Rightarrow E_1^{-1} E_2^{-1} \cdots E_{m-1}^{-1} E_m^{-1} I_n = A
\end{aligned}$$

Então A é um produto de matrizes inversíveis, logo A é inversível. Além disso, como $I_n = E_m E_{m-1} \cdots E_2 E_1 A$ e A é inversível, obtemos

$$I_n A^{-1} = E_m E_{m-1} \cdots E_2 E_1 A A^{-1}$$

$$A^{-1} = E_m E_{m-1} \cdots E_2 E_1 I_n,$$

de onde vemos que a mesma sucessão de operações elementares que transforma A em I_n transforma I_n em A^{-1}.

■

Exemplo 1.31. 1) *Seja*

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Usando o teorema acima, vamos verificar se A é inversível e em caso afirmativo determinar sua inversa. Temos pelo teorema que, se A for inversível, a mesma sucessão de operações elementares nos dará a inversa de A, então vamos fazer as operações elementares simultaneamente em A e na matriz identidade I_3:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \overset{L_3 \to L_3 - L_1}{\sim} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{array}\right) \overset{L_2 \to -L_3}{\sim} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 1 & 0 \end{array}\right)$$

$$\sim^{\substack{L_1 \to L_1 - 2L_2 \\ L_3 \to L_3 - L_2}} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & -1 & 0 & 2 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 2 & -1 & 1 & 1 \end{array}\right) \overset{L_3 \to \frac{1}{2} L_3}{\sim} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & -1 & 0 & 2 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right)$$

$$\sim^{L_1 \to L_1 - L_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right)$$

Com isso, como $A \sim I_n$, temos que A é inversível e que $A^{-1} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \\ 1 & 0 & -1 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

2) Seja

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 1 & 3 & 4 \end{pmatrix}.$$

Usando o teorema acima, vamos verificar se B é inversível e em caso afirmativo determinar sua inversa.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 3 & 4 & 0 & 0 & 1 \end{array}\right) \overset{L_3 \to L_3 - L_1}{\sim} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \end{array}\right) \overset{L_3 \to L_3 - L_2}{\sim} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{array}\right)$$

Como no processo de escalonamento de B obtemos uma linha toda nula, temos que $B \not\sim I_n$. Logo B não é inversível.

Com a teoria desenvolvida até aqui podemos reunir os resultados conforme o seguinte teorema.

Teorema 1.32. *Seja $A \in \mathcal{M}_n(K)$. As seguintes afirmações são equivalentes:*

(i) A é inversível;

(ii) O sistema linear homogêneo $AX = 0$ tem somente a solução nula;

(iii) O sistema linear $AX = B$ tem solução única, para qualquer $B \in \mathcal{M}_{n\times 1}(K)$;

(iv) $A \sim I_n$, ou seja, A é um produto de matrizes elementares.

Observação. 1) Em um sistema homogêneo $AX = 0$, se A não é inversível, o sistema tem infinitas soluções (portamto tem solução diferente de zero).

2) Em um sistema não homogêneo $AX = B$, se A não é inversível, ou o sistema não tem solução ou o sistema possui infinitas soluções. Para saber qual das possibilidades acontece, temos que escalonar o sistema.

Exemplo 1.33. 1) *Considere o sistema linear*

$$S : \begin{cases} x + 2y + z = 2 \\ \quad\;\; y + 2z = -1 \\ x + y + z = 5 \end{cases}$$

A matriz dos coeficientes de S é

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Vimos no Exemplo 1.31.1) que $A \sim I_3$, logo é inversível, e sua inversa é dada por

$$A^{-1} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \\ 1 & 0 & -1 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Então, como A é inversível, o sistema linear S tem solução única. Na forma matricial $AX = B$, a solução é dada por

$$X = A^{-1}B.$$

Assim

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \\ 1 & 0 & -1 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \\ 1 \end{pmatrix}.$$

Portanto, S tem solução única dada por $(7, -3, 1)$.

2) *Considere o sistema linear sobre $\mathbb{R}$*

$$S: \begin{cases} x + 2y + 2z = 2 \\ \phantom{x + {}} y + 2z = -1 \\ x + 3y + 4z = 5 \end{cases}$$

A matriz dos coeficientes de S é

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 1 & 3 & 4 \end{pmatrix}.$$

*Vimos no Exemplo 1.31.2) que $A \not\sim I_3$, logo **não é** inversível. Isso nos diz que S não tem solução única. Precisamos ainda saber se S tem então infinitas soluções ou não possui solução. Para isso, vamos escalonar escalonar S (usando a matriz ampliada de S):*

$$\left(\begin{array}{ccc|c} 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & -1 \\ 1 & 3 & 4 & 5 \end{array}\right) \overset{L_3 \to L_3 - L_1}{\sim} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & 3 \end{array}\right) \overset{L_3 \to L_3 - L_2}{\sim} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 4 \end{array}\right)$$

No processo de escalonamento, obtemos a equação $0 = 4$, o que nos diz que S não tem solução.

1.10 Determinante

Seja $A = (a_{ij}) \in \mathcal{M}_n(K)$. O **determinante** de A, denotado por $\det(A)$, é um escalar que associamos a A da seguinte forma:

n=1: $A = (a)_{1\times 1}$, $\det(A) = a$

n=2: $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$

n=3: $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{23}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$

Que podemos reescrever como

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

Com isso, vemos que o determinante de matrizes 3×3 pode ser obtido a partir de determinantes de submatrizes 2×2 de A. O determinante de uma matriz $n \times n$, para $n \geq 2$, é definido a partir de determinantes de submatrizes $(n-1) \times (n-1)$ de A, como veremos a seguir.

Sejam

$$A_{ij} = \text{ submatriz de } A \text{ onde a linha } i \text{ e a coluna } j \text{ foram retiradas}$$

e

$$\Delta_{ij} = (-1)^{i+j} \det(A_{ij}), \qquad \text{que chamamos de } \textbf{cofator} \text{ de } a_{ij}.$$

Definimos então, para qualquer $n \geq 2$,

$$\boxed{\det(A) = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in} = \textstyle\sum_{j=1}^{n} a_{ij}\Delta_{ij}}$$

onde i é uma linha fixa qualquer de A.

Fato: Pode-se escolher qualquer linha (ou qualquer coluna) para calcular $\det(A)$, usando a expressão acima, que o valor obtido é o mesmo. Esta expressão para o cálculo do determinante é chamada de **Desenvolvimento de Laplace**.

Exemplo 1.34. *Seja*

$$A = \begin{pmatrix} 2 & 0 & 0 & -3 \\ 0 & -1 & 0 & 0 \\ 7 & 4 & 3 & 5 \\ -6 & 2 & 2 & 4 \end{pmatrix}$$

Aplicando o Desenvolvimento de Laplace pela linha 1, obtemos

$$\begin{aligned} \det(A) &= 2\ (-1)^{1+1}\begin{vmatrix} -1 & 0 & 0 \\ 4 & 3 & 5 \\ 2 & 2 & 4 \end{vmatrix} + (-3)\ (-1)^{1+4}\begin{vmatrix} 0 & -1 & 0 \\ 7 & 4 & 3 \\ -6 & 2 & 2 \end{vmatrix} \\ &= 2\ (-1)\ (-1)^{1+1}\begin{vmatrix} 3 & 5 \\ 2 & 4 \end{vmatrix} + 3\ (-1)\ (-1)^{1+2}\begin{vmatrix} 7 & 3 \\ -6 & 2 \end{vmatrix} \\ &= -2\ (12-10) + 3\ (14+18) = 92 \end{aligned}$$

Calculando novamente o determinante de A pelo Desenvolvimento de Laplace pela linha 2, temos que

$$\det(A) = -1\ (-1)^{2+2} \begin{vmatrix} 2 & 0 & -3 \\ 7 & 3 & 5 \\ -6 & 2 & 4 \end{vmatrix} = -(24 - 42 - 54 - 20) = 92$$

1.11 Propriedades do Determinante

A seguir veremos uma série de propriedades do determinante de uma matriz e suas justificativas, utilizando a definição pelo Desenvolvimento de Laplace.

1) Se todos os elementos de uma linha (ou coluna) de uma matriz $A \in \mathcal{M}_n(K)$ são nulos, então

$$\det(A) = 0.$$

Prova: Basta aplicar o Desenvolvimento de Laplace pela linha (ou coluna) que é nula.

■

2) Se B é uma matriz obtida a partir de $A \in \mathcal{M}_n(K)$ pela multiplicação de uma única linha de A por $k \in K$, então

$$\det(B) = k \det(A).$$

Prova: Seja $A = (a_{ij})$ e assuma que B foi obtida de A ao multiplicarmos a linha i de A por $k \in K$. Então, calculando o determinante de B pela linha i,

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

obtemos

$$\det(B) = ka_{i1}\Delta_{i1} + ka_{i2}\Delta_{i2} + \cdots + ka_{in}\Delta_{in} = k(a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in}) = k\det(A).$$

■

Exemplo: $\begin{vmatrix} 2 & -1 & 0 \\ -2 & 4 & -6 \\ 0 & 5 & -3 \end{vmatrix} = 2 \begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & -3 \\ 0 & 5 & -3 \end{vmatrix}$

Observação 1.35. *A operação elementar* **(II)** *altera o determinante.*

3) Para qualquer matriz $A \in \mathcal{M}_n(K)$, tem-se

$$\det(kA) = k^n \det(A).$$

Prova: Segue de 2).

$\blacksquare$

4) Se B é obtida de $A \in \mathcal{M}_n(K)$ pela permutação de duas linhas, então

$$\det(B) = -\det(A).$$

Prova: Daremos a ideia da prova em aula.

$\blacksquare$

Observação 1.36. *A operação elementar* **(I)** *altera o sinal do determinante.*

5) Se duas linhas (ou colunas) de $A \in \mathcal{M}_n(K)$ são idênticas, então

$$\det(A) = 0.$$

Prova: Segue de 4).

$\blacksquare$

6) Seja

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{in}+b_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Então o determinante de B é dado por

$$\det(B) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Prova: Calculando o determinante de B pela linha i, obtemos

$$\begin{aligned}
\det(B) &= (a_{i1}+b_{i1})\Delta_{i1} + (a_{i2}+b_{i2})\Delta_{i2} + \cdots + (a_{in}+b_{in})\Delta_{in} \\
&= (a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in}) + (b_{i1}\Delta_{i1} + b_{i2}\Delta_{i2} + \cdots + b_{in}\Delta_{in}) \\
&= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}
\end{aligned}$$

■

Observação 1.37. *A propriedade acima nos diz que somando elementos b_{ij} em uma linha i de A, o determinante da matriz resultante é a soma do determinante de A com o determinante da matriz em que todas as entradas são iguais as de A exceto as da linha i, onde aparecem os elementos b_{ij}. Com isso podemos ver que o determinante da soma $A+B$ de duas matrizes $A, B \in \mathcal{M}_n(K)$ não é em geral a soma de* $\det(A)$ *e* $\det(B)$.

Exercício: Dê um exemplo em que $\det(A+B) \neq \det(A) + \det(B)$.

7) A operação elementar **(III) não altera** o determinante da matriz.

Prova: Seja $A = (a_{ij}) \in \mathcal{M}_n(K)$ e B a matriz obtida a partir de A ao realizarmos a operação elementar $L_i \to L_i + kL_j$, onde $j < i$. Então

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{i1}+ka_{j1} & a_{i2}+ka_{j2} & \cdots & a_{in}+ka_{jn} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Calculando o determinante de B pela linha i, obtemos

$$\begin{aligned}
\det(B) &= (a_{i1}+ka_{j1})\Delta_{i1} + (a_{i2}+ka_{j2})\Delta_{i2} + \cdots + (a_{in}+ka_{jn})\Delta_{in} \\
&= (a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in}) + k(a_{j1}\Delta_{i1} + ka_{j2}\Delta_{i2} + \cdots + a_{jn}\Delta_{in}) \\
&= \det(A) + k \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \overset{\mathbf{5)}}{=} \det(A) + k\cdot 0 = \det(A).
\end{aligned}$$

■

Com relação às operações elementares, vimos nas propriedades acima que

- a operação elementar (I) troca o sinal do determinante;
- a operação elementar (II) multiplica o determinante pelo escalar $k \neq 0$ da operação;
- a operação elementar (III) não altera o determinante.

Com isso obtemos as seguintes informações sobre o determinante das matrizes elementares.

(i) Seja E uma matriz elementar. Então $I_n \sim E$, ao aplicarmos exatamente uma operação elementar na matriz identidade I_n. Como $\det(I_n) = 1$, obtemos então

$$\det(E) = \begin{cases} -\det(I_n) &= -1, & \text{(op. elementar I)} \\ k\det(I_n) &= k, & \text{(op. elementar II)} \\ \det(I_n) &= 1, & \text{(op. elementar III)} \end{cases}$$

Note que então $\det(E) \neq 0$.

(ii) Seja agora $B \in \mathcal{M}_n(K)$ e E uma matriz elementar. Como a multiplicação EB é equivalente à fazer a operação elementar de E em B, obtemos

$$\det(EB) = \begin{cases} -\det(B) &= \det(E)\det(B), & \text{(op. elementar I)} \\ k\det(B) &= \det(E)\det(B), & \text{(op. elementar II)} \\ \det(B) &= \det(E)\det(B), & \text{(op. elementar III)} \end{cases}$$

Assim, se E é uma matriz elementar, então

$$\det(EB) = \det(E)\det(B).$$

Com estas observações, podemos provar o seguinte.

Teorema 1.38. $\boxed{A \in \mathcal{M}_n(K) \text{ é inversível se, e somente se, } \det(A) \neq 0.}$

Demonstração. Se A é inversível, sabemos que então $A \sim I_n$ e assim

$$I_n = E_r \cdots E_2 E_1 A,$$

onde $E_1, E_2, \ldots, E_r$ são matrizes elementares. Calculando o determinante dos dois lados da igualdade acima e aplicando a observação (ii), obtemos

$$1 = \det(I_n) = \det(E_r \cdots E_2 E_1 A) \overset{(ii)}{=} \det(E_r) \cdots \det(E_2) \det(E_1) \det(A).$$

Assim,

$$\det(E_r) \cdots \det(E_2) \det(E_1) \det(A) = 1.$$

Agora, pela observação (i), temos que o determinante de matrizes elementares é sempre diferente de zero. Então, já que o produto acima é igual a 1, segue que $\det(A) \neq 0$. Com isso mostramos que, se A é inversível, então $\det(A) \neq 0$.

Suponha agora que A não é inversível e seja R a forma escalonada de A. Então

$$R = E_s \cdots E_2 E_1 A,$$

onde $E_1, E_2, \ldots, E_s$ são matrizes elementares. Como A não é inversível, temos que sua forma escalonada R possui uma linha toda nula, e assim $\det(R) = 0$. Com isso, obtemos

$$0 = \det(R) = \det(E_s) \cdots \det(E_2) \det(E_1) \det(A).$$

E como $\det(E_i) \neq 0$, para todo i, segue que $\det(A) = 0$. Com isso mostramos que, se A não é inversível, então $\det(A) = 0$. Assim, se $\det(A) \neq 0$, temos que A é inversível (já que matrizes não inversíveis tem determinante igual a zero).

■

Teorema 1.39. *Para quaisquer matrizes $A, B \in \mathcal{M}_n(K)$,*

$$\boxed{\det(AB) = \det(A)\det(B)}.$$

Demonstração. A igualdade independe do fato da matriz A ser inversível ou não, mas vamos fazer a prova separando em dois casos.

Se A é inversível, então $A \sim I_n$ ou, equivalentemente, $I_n \sim A$. Então

$$A = E_m \cdots E_2 E_1 I_n,$$

onde $E_1, E_2, \ldots, E_m$ são matrizes elementares. Com isso e a observação (ii) acima, obtemos

$$\det(A) = \det(E_m) \cdots \det(E_2) \det(E_1) \det(I_n) = \det(E_m) \cdots \det(E_2) \det(E_1).$$

Obtemos também,

$$AB = E_m \cdots E_2 E_1 I_n B = E_m \cdots E_2 E_1 B.$$

Assim,

$$\det(AB) = \det(E_m)\cdots\det(E_2)\det(E_1)\det(B) = \det(A)\det(B)$$

como queríamos demonstrar.

Suponha agora que A não é inversível. Então, pelo teorema anterior, temos $\det(A) = 0$. Vamos mostrar que $\det(AB) = 0$, daí vale a igualdade $\det(AB) = 0 = \det(A)\det(B)$.

Como A não é inversível, a forma escalonada R de A tem uma linha toda nula e $R \sim A$, ou seja,

$$A = E_s \cdots E_2 E_1 R,$$

onde $E_1, E_2, \ldots, E_s$ são matrizes elementares e R tem uma linha toda nula. Com isso e a observação (ii), obtemos

$$AB = E_s \cdots E_2 E_1 RB$$

e

$$\det(AB) = \det(E_s)\cdots\det(E_2)\det(E_1)\det(RB).$$

Agora, como R tem uma linha toda nula, temos que RB também tem uma linha toda nula. Portanto $\det(RB) = 0$. Então, da igualdade acima, obtemos $\det(AB) = 0$, como queríamos.

■

Dos teoremas acima segue a seguinte relação entre o determinante de uma matriz inversível e o determinante de sua inversa.

Corolário 1.40. *Se $A \in \mathcal{M}_n(K)$ é inversível, então*

$$\boxed{\det(A^{-1}) = \tfrac{1}{\det(A)}}.$$

Demonstração. Como A é inversível, temos $\det(A) \neq 0$ e $AA^{-1} = I_n$. Com isso

$$AA^{-1} = I_n \Rightarrow \det(AA^{-1}) = \det(I_n) \Rightarrow \det(A)\det(A^{-1}) = 1 \Rightarrow \det(A^{-1}) = \frac{1}{\det(A)}.$$

■

Com a teoria desenvolvida até aqui, podemos reunir os resultados conforme o seguinte teorema.

Teorema 1.41. *Seja $A \in \mathcal{M}_n(K)$. As seguintes afirmações são equivalentes:*

i) *A é inversível;*

ii) *$A \sim I_n$, ou seja, A é um produto de matrizes elementares;*

iii) $\det(A) \neq 0$*;*

iv) *O sistema linear homogêneo $AX = 0$ tem solução única, que é a solução nula;*

v) *O sistema linear $AX = B$ tem solução única, para qualquer matriz de termos independentes B.*

Exemplo 1.42. *Considere o seguinte sistema linear sobre* $\mathbb{R}$.

$$\begin{cases} x_1 & +4x_2 & +5x_3 & = 2 \\ 4x_1 & +2x_2 & +5x_3 & = 3 \\ -3x_1 & +3x_2 & -\, x_3 & = 1 \end{cases}$$

Sua matriz A dos coeficientes tem determinante igual a

$$\det(A) = \begin{vmatrix} 1 & 4 & 5 \\ 4 & 2 & 5 \\ -3 & 3 & -1 \end{vmatrix} = 29.$$

Como $\det(A) \neq 0$, *o sitema tem solução única. (Encontre!)*

Exemplo 1.43. *Considere o sistema linear*

$$S : \begin{cases} x & +\,y & -az & = 0 \\ ax & +\,y & -\,z & = 2-a \\ x & +ay & -\,z & = 0 \end{cases}$$

Determine para quais valores de $a \in \mathbb{R}$

i) S tem solução única;

ii) S não tem solução;

iii) S tem infinitas soluções;

iv) a matriz dos coeficientes de S é inversível;

v) a matriz dos coeficientes de S não é inversível.

Solução: *A matriz A dos coeficientes tem determinante igual a*

$$\det(A) = \begin{vmatrix} 1 & 1 & -a \\ a & 1 & -1 \\ 1 & a & -1 \end{vmatrix} = -a^3 + 3a - 2 = -(a-1)^2(a+2).$$

Sabemos que S tem solução única, se e somente se, a matriz A dos coeficientes de S é inversível. E A é inversível se, e somente se, $\det(A) \neq 0$. *Então*

i) S tem solução única para todo $a \in \mathbb{R}$, com $a \neq 1$ e $a \neq -2$.

iv) A é inversível para todo $a \in \mathbb{R}$, com $a \neq 1$ e $a \neq -2$.

v) A não é inversível se $a = 1$ ou $a = -2$.

Para $a = 1$, temos

$$S : \left\{ \begin{array}{cccc} x & +y & -z & = 0 \\ x & +y & -z & = 1 \\ x & +y & -z & = 0 \end{array} \right. \sim \left\{ \begin{array}{cccc} x & +y & -z & = 0 \\ & & 0 & = 1 \\ & & 0 & = 0 \end{array} \right.$$

Então S não tem solução para $a = 1$.

Para $a = -2$, temos

$$S : \left\{ \begin{array}{cccc} x & +y & +2z & = 0 \\ -2x & +y & -z & = 4 \\ x & -2y & -z & = 0 \end{array} \right. \sim \left\{ \begin{array}{cccc} x & +y & +2z & = 0 \\ & +3y & +3z & = 4 \\ & -3y & -3z & = 0 \end{array} \right. \sim \left\{ \begin{array}{cccc} x & +y & +2z & = 0 \\ & +3y & +3z & = 4 \\ & & 0 & = 4 \end{array} \right.$$

Então S também não tem solução para $a = -2$. Com isso, temos

ii) S não tem solução para $a = 2$ ou $a = -2$.

iii) Não há valor de $a \in \mathbb{R}$ tal que S tenha infinitas soluções.

1.12 Regra de Cramer

Usando o determinante de matrizes e suas propriedades, apresentamos nesta seção uma forma de encontrar a solução de um sistema linear de ordem n quando este sistema tem solução única, chamada de **Regra de Cramer**. Sabemos que um sistema linear de ordem n tem solução única quando sua matriz A dos coeficientes é inversível, ou equivalentemente, A tem determinante diferente de zero.

Seja S um sistema linear sobre K de ordem n,

$$S : \left\{ \begin{array}{cc} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = & b_2 \\ \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = & b_n \end{array} \right. ,$$

que possui solução única $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in K^n$.

Considere agora a matriz obtida a partir da matriz A dos coeficientes de S ao substituirmos os elementos da coluna 1 de A pelos termos independentes de S:

$$\begin{pmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Como α é solução de S, cada b_i pode ser escrito como $a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n = b_i$. Substituindo tais igualdades na matriz acima e calculando seu determinante, com as propriedades dos determinantes da Seção 1.11 (indicadas na igualdades), obtemos

$$\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1n}\alpha_n & a_{12} & \cdots & a_{1n} \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2n}\alpha_n & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1}\alpha_1 + a_{n2}\alpha_2 + \cdots + a_{nn}\alpha_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$\overset{\mathbf{6)}}{=} \begin{vmatrix} a_{11}\alpha_1 & a_{12} & \cdots & a_{1n} \\ a_{21}\alpha_1 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1}\alpha_1 & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{12}\alpha_2 & a_{12} & \cdots & a_{1n} \\ a_{22}\alpha_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n2}\alpha_2 & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{1n}\alpha_n & a_{12} & \cdots & a_{1n} \\ a_{2n}\alpha_n & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{nn}\alpha_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$\overset{\mathbf{2)}}{=} \alpha_1 \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \alpha_2 \begin{vmatrix} a_{12} & a_{12} & \cdots & a_{1n} \\ a_{22} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \alpha_n \begin{vmatrix} a_{1n} & a_{12} & \cdots & a_{1n} \\ a_{2n} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{nn} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$\overset{\mathbf{5)}}{=} \alpha_1 \cdot \det(A) + \alpha_2 \cdot 0 + \cdots + \alpha_n \cdot 0$$

Assim, obtemos a igualdade

$$\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \alpha_1 \cdot \det(A).$$

Como $\det(A) \neq 0$, segue que

$$\alpha_1 = \frac{\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}}{\det(A)}$$

Como um raciocínio análogo, substituindo a i-ésima coluna de A pelos termos independentes de S, encontramos α_i. Assim

$$\alpha_1 = \frac{\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}}{\det(A)}, \quad \alpha_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} & \cdots & a_{1n} \\ a_{21} & b_2 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & b_n & a_{n3} & \cdots & a_{nn} \end{vmatrix}}{\det(A)},$$

$$\ldots, \alpha_n = \frac{\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & b_n \end{vmatrix}}{\det(A)}$$

Exemplo 1.44. *Considere o seguinte sistema linear sobre $\mathbb{R}$ (Veja Exemplo 1.42).*

$$\begin{cases} x_1 & +4x_2 & +5x_3 & = 2 \\ 4x_1 & +2x_2 & +5x_3 & = 3 \\ -3x_1 & +3x_2 & -\ x_3 & = 1 \end{cases}$$

Sua matriz A dos coeficientes tem determinante igual a

$$\det(A) = \begin{vmatrix} 1 & 4 & 5 \\ 4 & 2 & 5 \\ -3 & 3 & -1 \end{vmatrix} = 29.$$

Então S tem solução única. Pela Regra de Cramer, a solução de S é

$$x_1 = \frac{\begin{vmatrix} 2 & 4 & 5 \\ 3 & 2 & 5 \\ 1 & 3 & -1 \end{vmatrix}}{29} = -\frac{12}{29}, \quad x_2 = \frac{\begin{vmatrix} 1 & 2 & 5 \\ 4 & 3 & 5 \\ -3 & 1 & -1 \end{vmatrix}}{29} = \frac{20}{29}$$

$$e \quad x_3 = \frac{\begin{vmatrix} 1 & 4 & 2 \\ 4 & 2 & 3 \\ -3 & 3 & 1 \end{vmatrix}}{29} = -\frac{23}{29}.$$

1.13 Matriz Adjunta

Nesta seção apresentamos uma maneira diferente de se obter a matriz inversa de uma matriz inversível A, através da matriz adjunta de A. Computacionalmente, encontrar a inversa fazendo operações elementares é bem mais eficiente do que através da matriz adjunta. Mas, além de o desenvolvimento dessa seção ser uma aplicação interessante das propriedades do determinante, a matriz adjunta terá papel importante em demostrações de resultados que desenvolveremos mais adiante.

Seja $A = (a_{ij}) \in \mathcal{M}_n(K)$. Lembre que **cofator** da entrada a_{ij} é

$$\Delta_{ij} = (-1)^{i+j} \det(A_{ij}),$$

onde A_{ij} é a submatriz de A obtida quando retiramos a linha i e a coluna j.

Encontrando Δ_{ij} para todo i e j, podemos formar uma outra matriz

$$\overline{A} = (\Delta_{ij})_{n\times n},$$

chamada de **matriz dos cofatores** de A.

Exemplo 1.45. *Seja* $A = \begin{pmatrix} 2 & 1 & 0 \\ -3 & 1 & 4 \\ 1 & 6 & 5 \end{pmatrix}$. *Então*

$$\Delta_{11} = (-1)^{1+1} \begin{vmatrix} 1 & 4 \\ 6 & 5 \end{vmatrix} = 5 - 24 = -19$$

Calculando todos os Δ_{ij}, vemos que

$$\overline{A} = \begin{pmatrix} -19 & 19 & -19 \\ -5 & 10 & -11 \\ 4 & -8 & 5 \end{pmatrix}.$$

Lembre que, se $B = (b_{ij})_{m\times n}$ é uma matriz, a *transposta* de B é definida por $B^t = (b_{ji})_{n\times m}$ (o que é linha vira coluna e o que é coluna vira linha). Com isso, podemos definir a matriz adjunta.

Definição 1.46. *Dada uma matriz $A \in \mathcal{M}_n(K)$, a* ***matriz adjunta*** *de A é a transposta da matriz dos cofatores de A:*

$$\boxed{\operatorname{Adj}(A) = \overline{A}^t.}$$

Exemplo 1.47. *Voltando no exemplo anterior, a matriz adjunta de A é*

$$\operatorname{Adj}(A) = \begin{pmatrix} -19 & -5 & 4 \\ 19 & 10 & -8 \\ -19 & -11 & 5 \end{pmatrix}.$$

Sobre a matriz adjunta, temos a seguinte propriedade.

Teorema 1.48. *Para $A \in \mathcal{M}_n(K)$, temos*

$$\boxed{A \cdot \operatorname{Adj}(A) = \det(A) \cdot I_n.}$$

Demonstração. Como $\operatorname{Adj}(A) = (\Delta_{ij})^t$, temos

$$A \cdot \operatorname{Adj}(A) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{pmatrix}.$$

Então

$$A \cdot \operatorname{Adj}(A) = (c_{ik})_{n\times n},$$

onde

$$c_{11} = a_{11}\Delta_{11} + a_{12}\Delta_{12} + \cdots a_{1n}\Delta_{1n} = \det(A)$$

$$c_{12} = a_{11}\Delta_{21} + a_{12}\Delta_{22} + \cdots a_{1n}\Delta_{2n} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0$$

$$\vdots$$

$$c_{ik} = a_{i1}\Delta_{k1} + a_{i2}\Delta_{k2} + \cdots a_{in}\Delta_{kn} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{k-1,1} & a_{k-1,2} & \cdots & a_{k-1,n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{cases} \det(A) & , \text{ se } i = k; \\ 0 & , \text{ se } i \neq k. \end{cases}$$

Então

$$A \cdot \mathrm{Adj}(A) = \begin{pmatrix} \det(A) & 0 & \cdots & 0 \\ 0 & \det(A) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \det(A) \end{pmatrix} = \det(A) \cdot I_n.$$

■

Corolário 1.49. *Se $A \in \mathcal{M}_n(K)$ é inversível, então*

$$A^{-1} = \frac{1}{\det(A)} \cdot \mathrm{Adj}(A).$$

Demonstração. Com A é inversível, temos $\det(A) \neq 0$. Então

$$A \cdot \mathrm{Adj}(A) = \det(A) \cdot I_n$$

$$\Rightarrow A^{-1} \cdot A \cdot \mathrm{Adj}(A) = A^{-1}(\det(A) \cdot I_n)$$

$$\Rightarrow \mathrm{Adj}(A) = \det(A) \cdot A^{-1}$$

$$\Rightarrow A^{-1} = \frac{1}{\det(A)} \cdot \mathrm{Adj}(A).$$

■

Exemplo 1.50. *Considere a matriz* $A = \begin{pmatrix} 2 & 1 & 0 \\ -3 & 1 & 4 \\ 1 & 6 & 5 \end{pmatrix}$ *do Exemplo 1.45. Então sua matriz dos cofatores é*

$$\overline{A} = \begin{pmatrix} -19 & 19 & -19 \\ -5 & 10 & -11 \\ 4 & -8 & 5 \end{pmatrix}$$

e sua matriz adjunta é

$$\mathrm{Adj}(A) = \begin{pmatrix} -19 & -5 & 4 \\ 19 & 10 & -8 \\ -19 & -11 & 5 \end{pmatrix}.$$

Além disso, calculando o determinante de A, obtemos $\det(A) = -19$. *Então*

$$A^{-1} = \frac{1}{\det(A)} = \frac{1}{(-19)} \begin{pmatrix} -19 & -5 & 4 \\ 19 & 10 & -8 \\ -19 & -11 & 5 \end{pmatrix} = \begin{pmatrix} 1 & \frac{5}{19} & -\frac{4}{19} \\ -1 & -\frac{10}{19} & \frac{8}{19} \\ 1 & \frac{11}{19} & -\frac{5}{19} \end{pmatrix}.$$

2 Espaços Vetoriais

Definição 2.1. *Seja K um corpo. Um* ***espaço vetorial*** *V* ***sobre*** *K é um conjunto não vazio V, cujos elementos são chamados (abstratamente) de vetores, com duas operações*

(A) *a adição de vetores, que a cada par de vetores $u, v \in V$ associa um vetor $u + v \in V$, de maneira que*

  (A1) $u + v = v + u$, *para todos $u, v \in V$;*

  (A2) $(u + v) + w = u + (v + w)$, *para todos $u, v, w \in V$;*

  (A3) *existe um vetor em V, denotado por 0 e chamado de vetor nulo, tal que $v + 0 = v$, para todo $v \in V$;*

  (A4) *para todo vetor $v \in V$, existe um vetor $-v \in V$ tal que a soma $v + (-v) = 0$, o vetor nulo;*

(M) *a multiplicação por escalar, que a cada escalar $k \in K$ e cada $v \in V$ associa um vetor $kv \in V$, de maneira que*

  (M1) $1.v = v$, *para todo $v \in V$;*

  (M2) $(k_1 k_2)v = k_1(k_2 v)$, *para todos $k_1, k_2 \in K$ e $v \in V$;*

  (M3) $k(u + v) = ku + kv$, *para todos $k \in K$ e $u, v \in V$;*

  (M4) $(k_1 + k_2)v = k_1 v + k_2 v$, *para todos $k_1, k_2 \in K$ e $v \in V$.*

Exemplo 2.2. *Para cada inteiro $n \geq 1$,*

$$K^n := \{(x_1, x_2, \ldots, x_n) \mid x_i \in K,\ i = 1, 2, \ldots, n\},$$

o conjunto de todas a n-uplas com coordenadas em K, com adição e multiplicação por escalar definidas por, para cada $u = (x_1, x_2, \ldots, x_n), v = (y_1, y_2, \ldots, y_n) \in K^n$ e $k \in K$,

$$u + v := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \quad e \quad ku := (kx_1, kx_2, \ldots, kx_n),$$

é um espaço vetorial sobre K, simplesmente chamado de ***o espaço vetorial*** *K^n.*

O vetor nulo é a n-upla $0 = (0, 0, \ldots, 0)$, com todas as n coordenadas nulas e, dado $v = (y_1, y_2, \ldots, y_n) \in K^n$, o vetor $-v \in K^n$ é dado por $-v = (-y_1, -y_2, \ldots, -y_n)$. Fica como exercício verificar que as propriedades (A1)-(A4) e (M1)-(M4) são satisfeitas.

Observações:

i) *Quando $n = 1$, temos K um espaço vetorial sobre ele mesmo, onde a adição de vetores é a adição em K e a multiplicação por escalar é a multiplicação em K.*

ii) *$\mathbb{C}^n$ é um espaço vetorial sobre $\mathbb{C}$,* ***o espaço vetorial*** *$\mathbb{C}^n$;*
$\mathbb{R}^n$ é um espaço vetorial sobre $\mathbb{R}$, ***o espaço vetorial*** *$\mathbb{R}^n$.*

Exemplo 2.3. *A matrizes $\mathcal{M}_{m\times n}(K)$, com a adição de matrizes e multiplicação por escalar definidas na Subseção 1.8 são espaços vetoriais sobre K. Vimos nas Propriedades 1.16 que* (A1)-(A4) *e* (M1)-(M4) *são satisfeitas.*

Exemplo 2.4. *O conjunto dos números complexos*

$$\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}\}$$

com a adição de números complexos: se $u = x_1 + y_1 i$ e $v = x_2 + y_2 i$

$$u + v = (x_1 + x_2) + (y_1 + y_2)i \ ;$$

e a multiplicação por escalar: se $u = x + yi$ e $r \in \mathbb{R}$,

$$ru = rx + ryi,$$

*é um **espaço vetorial sobre** $\mathbb{R}$.*

Exemplo 2.5. *$\mathbb{R}$ é um espaço vetorial sobre $\mathbb{Q}$, onde para todos $x, y \in \mathbb{R}$ e $r \in \mathbb{Q}$, a adição $x + y$ é a adição usual em $\mathbb{R}$ e a multiplicação por escalar rx é a multiplicação usual em $\mathbb{R}$.*

Exemplo 2.6. *Seja X um conjunto não-vazio arbitrário e*

$$\mathcal{F}(X, K) := \{f \mid f : X \to K\}$$

o conjunto de todas as funções de X no corpo K. Defina as seguintes operações em $\mathcal{F}(X, K)$:

(A) *para $f, g \in \mathcal{F}(X, K)$, defina $f + g$ por*

$$(f + g)(x) = f(x) + g(x), \quad \text{para todo } x \in X;$$

(M) *para $f \in \mathcal{F}(X, K)$ e $k \in K$, defina kf por*

$$(kf)(x) = kf(x), \quad \text{para todo } x \in X.$$

*$\mathcal{F}(X, K)$ com tais operaçãos é um espaço vetorial sobre K, chamado de **espaço de funções**. O vetor nulo é a função constante igual a zero e, dada $f \in \mathcal{F}(X, K)$, a função $-f$ é tal que $(-f)(x) := -f(x)$, $\forall x \in X$.*

Exemplo 2.7. *Considere o intervalo fechado $[a, b] \subset \mathbb{R}$.*

$$\mathcal{C}([a, b], \mathbb{R}) = \{f : [a, b] \to \mathbb{R} \mid f \text{ é contínua}\},$$

com as operações definidas no Exemplo 2.6 acima, é um espaço vetorial sobre $\mathbb{R}$. De fato, se f e g são contínuas e $r \in \mathbb{R}$, então $f + g$ e rf também são contínuas; o vetor nulo é a função constante igual a zero, que é contínua; e (A1)-(A4) *e* (M1)-(M4) *são satisfeitas.*

Exemplo 2.8. *Considere o conjunto de todos os polinômios* $f(t) = a_0+a_1t+a_2t^2+\cdots+a_mt^m$, *com coeficientes* $a_0, a_1, \ldots, a_m$ *no corpo* K *e qualquer grau* m, *que denotamos por* $\mathcal{P}(K)$:

$$\mathcal{P}(K) := \{a_0 + a_1t + a_2t^2 + \cdots + a_mt^m \mid a_0, a_1, \ldots, a_m \in K \quad e \quad m \in \mathbb{Z}, m \geq 0\}.$$

Dois tais polinômios $f(t) = a_0 + a_1t + a_2t^2 + \cdots + a_mt^m$ *e* $g(t) = b_0 + b_1t + b_2t^2 + \cdots + b_nt^n$ *são iguais quando* $a_i = b_i, \forall i$. *Quando todos os coeficientes de um polinômio* $f(t)$ *são iguais a zero, o chamamos polinômio nulo. Assim, a igualdade* $a_0 + a_1t + a_2t^2 + \cdots + a_mt^m = 0$ *vale se, e somente se,* $a_0 = a_1 = a_2 = \cdots = a_m = 0$.

Dados $f(t) = a_0 + a_1t + a_2t^2 + \cdots + a_mt^m$, $g(t) = b_0 + b_1t + 2_2t^2 + \cdots + b_mt^m \in \mathcal{P}(K)$ *e* $k \in K$, *defina a adição e multiplicação por escalar em* $\mathcal{P}(K)$ *por*

$$f(t) + g(t) := a_0 + b_0 + (a_1 + b_1)t + (a_2 + b_2)t^2 + \cdots + (a_m + b_m)t^m$$

e

$$kf(t) := ka_0 + ka_1t + ka_2t^2 + \cdots + ka_mt^m.$$

Por exemplo, se $f(t) = 1 - 2t + 7t^2 - 5t^3$ *e* $g(t) = -2 + 5t^2 - t^3 + 2t^4$, *então*

$$f(t) + g(t) = -1 - 2t + 12t^2 - 6t^3 + 2t^4 \quad e \quad 2f(t) = 2 - 4t + 14t^2 - 10t^3.$$

Então $\mathcal{P}(K)$, *com tais operações, é um espaço vetorial,* ***o espaço vetorial dos polinômios sobre*** K. *De fato, para quaisquer* $f(t) = a_0 + a_1t + a_2t^2 + \cdots + a_mt^m$, $g(t) = b_0 + b_1t + 2_2t^2 + \cdots + b_mt^m$, $h(t) = c_0 + c_1t + c_2t^2 + \cdots + c_mt^m \in \mathcal{P}(K)$ *e* $r, s \in K$

(A1) $f(t) + g(t) = a_0 + b_0 + (a_1 + b_1)t + (a_2 + b_2)t^2 + \cdots + (a_m + b_m)t^m = b_0 + a_0 + (b_1 + a_1)t + (b_2 + a_2)t^2 + \cdots + (b_m + a_m)t^m = g(t) + f(t)$, *logo vale* (A1);

(A2) $(f(t) + g(t)) + h(t) = [a_0 + b_0 + (a_1 + b_1)t + (a_2 + b_2)t^2 + \cdots + (a_m + b_m)t^m] + h(t) = [(a_0+b_0)+c_0]+[(a_1+b_1)+c_1]t+[(a_2+b_2)+c_2]t^2+\cdots+[(a_m+b_m)+c_m]t^m = [a_0+(b_0+c_0)]+[a_1+(b_1+c_1)]t+[a_2+(b_2+c_2)]t^2+\cdots+[a_m+(b_m+c_m)]t^m = f(t)+[g(t)+h(t)]$, *logo vale* (A2);

(A3) *o vetor nulo é o polinômio nulo, pois* $f(t) + 0 = f(t)$, *para todo* $f(t) \in \mathcal{P}(K)$, *logo vale* (A3);

(A4) $-f(t) = -a_0 - a_1t - a_2t^2 - \cdots - a_mt^m$, *pois* $-f(t) \in \mathcal{P}(K)$ *e* $f(t) + (-f(t)) = 0$ *é igual ao polinômio nulo, logo vale* (A4);

(M1) $1.f(t) = f(t)$;

(M2) $(rs)f(t) = (rs)a_0 + (rs)a_1t + (rs)a_2t^2 + \cdots + (rs)a_mt^m = r(sa_0 + sa_1t + sa_2t^2 + \cdots + sa_mt^m) = r(sf(t))$, *logo vale* (M2);

(M3) $r(f(t) + g(t)) = r(a_0 + b_0 + (a_1 + b_1)t + (a_2 + b_2)t^2 + \cdots + (a_m + b_m)t^m) = ra_0 + rb_0 + (ra_1 + rb_1)t + (ra_2 + rb_2)t^2 + \cdots + (ra_m + rb_m)t^m = rf(t) + rg(t)$, *logo vale* (M3);

(M4) $(r + s)f(t) = (r + s)a_0 + (r + s)a_1t + (r + s)a_2t^2 + \cdots + (r + s)a_mt^m = ra_0 + sa_0 + ra_1t + sa_1t + ra_2t^2 + sa_2t^2 + \cdots + ra_mt^m + sa_mt^m = ra_0 + ra_1t + ra_2t^2 + \cdots + ra_mt^m + sa_0 + sa_1t + sa_2t^2 + \cdots + sa_mt^m = rf(t) + sf(t)$, *logo vale* (M4).

Exemplo 2.9.
$$\mathbb{C}^n = \{(u_1, u_2, \ldots, u_n) \mid u_i \in \mathbb{C},\ i = 1, 2, \ldots, n\}$$
é um ***espaço vetorial sobre*** $\mathbb{R}$*, onde, se* $u = (u_1, u_2, \ldots, u_n), v = (v_1, v_2, \ldots, v_n) \in \mathbb{C}^n$ *e* $r \in \mathbb{R}$*,*
$$u + v := (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \quad e \quad ru = (ru_1, ru_2, \ldots, ru_n).$$
Este espaço vetorial real é bem diferente do espaço vetorial $\mathbb{C}^n$*, definido no Exemplo 2.2.*

2.1 Subespaços

Nesta seção vamos estudar subconjuntos especiais dos espaços vetoriais, subconjuntos que são também espaços vetoriais por si só.

Definição 2.10. *Seja* V *um espaço vetorial sobre um corpo* K*. Um* ***subespaço de*** V *é um subconjunto* $W \subseteq V$ *que é também um espaço vetorial sobre* K *com as operações de adição de vetores e multiplicação por escalar definidas em* V*.*

Na definição, ao dizer que W é também um espaço vetorial sobre K com as operações de adição de vetores e multiplicação por escalar definidas em V, está subentendido que a adição de vetores pertencentes ao subconjunto W e a multiplicação de um vetor em W por um escalar permanecem no subconjunto W, além de W ser não-vazio. De fato, para garantir que $W \subseteq V$ é um subspaço de V, basta que isso aconteça.

Proposição 2.11. *Seja* V *um espaço vetorial sobre um corpo* K *e* $W \subseteq V$*. Então* W *é subespaço de* V *se, e somente se,* W *é não-vazio e,* $\forall w_1, w_2 \in W$ *e* $k \in K$*,*
$$w_1 + w_2 \in W \quad e \quad kw_1 \in W,$$
ou seja, W *é não vazio e fechado para soma e para multiplicação por escalar de* V*.*

Demonstração. Se W é subespaço de V, então W é um espaço vetorial, logo $W \neq \emptyset$. Além disso, se $w_1, w_2 \in W$ e $k \in K$, temos $w_1 + w_2 \in W$ e $kw_1 \in W$.

Reciprocamente, sendo W fechado para adição de vetores e para a multiplicação por escalar de V, (A1)-(A4) e (M1)-(M4) valem. De fato, (A1), (A2) e (M1)-(M4) valem pois valem para todos os vetores de V, logo valem em particular para vetores de W. Além disso, como W é não-vazio e como ao multiplicarmos um vetor de W por escalares continuamos com vetores em W, tomando $w \in W$ qualquer, obtemos que $0.u = 0 \in W$ e $-1.u = -u \in W$, ou seja, o vetor nulo pertende a W e $-w \in W$, para qualquer vetor $w \in W$. Assim (A3) e (A4) também são satisfeitas. ∎

Notação: Escrevemos $W \leq V$ para indicar que W é subespaço de V.

Exemplo 2.12. *Como espaços vetoriais sobre* $\mathbb{Q}$*, temos*
$$\mathbb{Q} \ \leq \ \mathbb{Q}[\sqrt{2}] \ \leq \ \mathbb{R} \ \leq \ \mathbb{C}.$$

Exemplo 2.13. *Seja V um espaço vetorial sobre K. Então $\{0\}$ e V são subespaços de V, chamados de* ***subespaços triviais****.*

Exemplo 2.14. $\mathcal{C}([a,b],\mathbb{R}) \ \leq \ \mathcal{F}([a,b],\mathbb{R})$

Exemplo 2.15. *Considere o espaço vetorial $\mathcal{P}(K)$ dos polinômios sobre K e, para $m \geq 0$ fixo, seja*

$$\mathcal{P}_m(K) := \{a_0 + a_1t + a_2t^2 + \cdots + a_mt^m \mid a_0, a_1, \ldots, a_m \in K\}$$

o subconjunto dos polinômios de grau no máximo m.

Por exemplo, $\mathcal{P}_3(\mathbb{R}) := \{a_0 + a_1t + a_2t^2 + a_3t^3 \mid a_0, a_1, a_2, a_3 \in \mathbb{R}\}$ é o conjunto dos polinômios de grau ≤ 3.

Para cada m, temos que $\mathcal{P}_m(K)$ é um subespaço do espaço vetorial $\mathcal{P}(K)$. De fato, quando somamos polinômios de grau no máximo m, a soma é um polinômio de grau no máximo m; e ao multiplicarmos um polinômio de grau no máximo m por um escalar qualquer, obtemos um polinômio de grau no máximo n. Assim, $\mathcal{P}_m(K)$ é ***o espaço vetorial dos polinômios de grau*** *$\leq m$* ***sobre*** *K****.***

Exemplo 2.16. *Vamos verificar se*

a) $U = \{(x,y,z,t) \in \mathbb{R}^4 \mid 2x - y + 3z + 2t = 1\}$

b) $W = \{(x,y,z,t) \in \mathbb{R}^4 \mid -x + 3y + 2z - 5t = 0\}$

são subespaços de $\mathbb{R}^4$.

a) Um subespaço de um espaço vetorial contém o vetor nulo. O vetor nulo $0 = (0,0,0,0)$ de $\mathbb{R}^4$ não pertende a U, pois não satisfaz a equação $2x - y + 3z + 2t = 1$. Então U não é subespaço de $\mathbb{R}^4$.

b) W contém o vetor nulo. Vamos verificar se W é fechado para soma de vetores e para a multiplicação por escalar. Temos que um vetor $u = (x,y,z,t) \in W$ se, e somente se, satisfaz $-x + 3y + 2z - 5t = 0$. Sejam $u_1 = (x_1,y_1,z_1,t_1), u_2 = (x_2,y_2,z_2,t_2) \in W$ e $r \in \mathbb{R}$. Então $-x_1 + 3y_1 + 2z_1 - 5t_1 = 0$ e $-x_2 + 3y_2 + 2z_2 - 5t_2 = 0$. Com isso, temos que

$$u_1 + u_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2, t_1 + t_2) \ \text{ e } \ ru_1 = (rx_1, ry_1, rz_1, rt_1)$$

satisfazem

$$\begin{gathered} -(x_1 + x_2) + 3(y_1 + y_2) + 2(z_1 + z_2) - 5(t_1 + t_2) = \\ -x_1 + 3y_1 + 2z_1 - 5t_1 + (-x_2 + 3y_2 + 2z_2 - 5t_2) = \\ 0 + 0 = 0 \end{gathered}$$

e

$$-rx_1 + 3ry_1 + 2rz_1 - 5rt_1 = r(-x_1 + 3y_1 + 2z_1 - 5t_1) = r.0 = 0,$$

logo $u_1 + u_2 \in W$ e $ru_1 \in W$. Portanto W é subespaço de $\mathbb{R}^4$.

Exemplo 2.17. *Seja S um sistema linear homogêneo sobre K em n incógnitas. Então o conjunto-solução de S, que é formado por todas as n-uplas que satisfazem as equações de S, é um subconjunto de K^n. Mais ainda, como a soma de soluções e a multiplicação por escalar são também soluções de S (confira Exercício 1.15), temos que o conjunto-solução de S é um subespaço de K^n. Assim,* ***o conjunto-solução de um sistema linear homogêneo é um espaço vetorial****.*

Note que no Exemplo 2.16 acima, $W = \{(x, y, z, t) \in \mathbb{R}^4 \mid -x + 3y + 2z - 5t = 0\}$ é o conunjo-solução da equação homogênea $-x + 3y + 2z - 5t = 0$, logo um subespaço de $\mathbb{R}^4$.

Exemplo 2.18. *O conjunto da matrizes $n \times n$ sobre K simétricas, dado por*

$$W = \{A \in \mathcal{M}_n(K) \mid A^t = A\}$$

e o conjunto das matrizes $n \times n$ sobre K antissimétricas, dado por

$$U = \{A \in \mathcal{M}_n(K) \mid A^t = -A\}$$

são subespaços de $\mathcal{M}_n(K)$. De fato, para quaisquer matrizes $A, B \in \mathcal{M}_n(K)$ e $k \in K$,

$$(A + B)^t = A^t + B^t \quad e \quad (kA)^t = kA^t.$$

Então, se A e B são simétricas e $k \in K$, temos que $(A + B)^t = A^t + B^t = A + B$ e $(kA)^t = kA^t = kA$. Logo $A + B$ e kA também são simétricas.
E se A e B são antissimétricas e $k \in K$, temos que $(A+B)^t = A^t+B^t = -A-B = -(A+B)$ e $(kA)^t = kA^t = k(-A) = -(kA)$. Logo $A + B$ e kA também são antissimétricas.

Exemplo 2.19. *No espaço $\mathcal{M}_n(\mathbb{C})$ de matrizes $n \times n$ complexas, considere o subconjunto das matrizes chamadas hermitianas (ou auto-adjuntas) dado por*

$$U = \{A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C}) \mid a_{ij} = \overline{a}_{ji}\},$$

onde $\overline{a}_{ji}$ é o conjugado complexo de a_{ji}.

Note que, as entradas a_{ii} da diagonal principal de uma matriz hermitiana $A = (a_{ij})$ são reais, pois a condição $a_{ii} = \overline{a}_{ii}$ implica que $a_{ii} \in \mathbb{R}$. Então U não é fechado para a multiplicação por escalares complexos, logo U não é subespaço do espaço vetorial $\mathcal{M}_n(\mathbb{C})$ sobre $\mathbb{C}$.

Mas considerando o espaco vetorial $\mathcal{M}_n(\mathbb{C})$ sobre $\mathbb{R}$, com a adição de matrizes usual e a multiplicação por escalar restrita a escalares reais, temos que o conjunto das matrizes complexas hermitianas é um subespaço de $\mathcal{M}_n(\mathbb{C})$ sobre $\mathbb{R}$.

Exemplo 2.20. *Seja V um espaço vetorial sobre um corpo K e U e W subespaços de V. Então a interseção*

$$U \cap W := \{v \in V \mid v \in U \text{ e } v \in W\}$$

dos subespaços U e W é também um subespaço de V. De fato,

- *como U e W são subespaços, temos que o vetor nulo $0 \in U$ e $0 \in W$, logo $0 \in U \cap W$ e assim $U \cap W$ é não vazio;*

- *sejam $v_1, v_2 \in U \cap W$ e $k \in K$. Então $v_1, v_2 \in U$ e, com U é subespaço, obtemos $v_1 + v_2 \in U$ e $kv_1 \in U$. Analogamente, $v_1 + v_2 \in W$ e $kv_1 \in W$. Logo $v_1 + v_2 \in U \cap W$ e $kv_1 \in U \cap W$.*

Agora observe que a união de subespaços pode não ser um subespaço. Por exemplo, em $\mathbb{R}^2$, considerando duas retas distintas que passam pela origem (retas passando pela origem são subespaços), a união é o desenho em formato de X, o conjunto formado pelas duas retas. Esta união não é um subespaço, ao tomarmos um vetor em cada reta o vetor soma não pertence a nenhuma das retas, logo não está na união. Na verdade vale o seguinte

Exercício: Sejam U e W subespaços de um espaço vetorial V. Então $U \cup W$ é subespaço de V se, e somente se, $U \subseteq W$ ou $W \subseteq U$.

Embora $U \cup W$ não seja subespaço, podemos definir o seguinte subespaço de V que contém a união, que chamamos de subespaço soma e é, de fato, o subespaço gerado pela união (vamos entender melhor isso mais adiante):

$$U + W := \{u + w \mid u \in U \text{ e } w \in W\},$$

ou seja, o conjunto de todos os vetores soma $u + w$, com $u \in U$ e $w \in W$. Note que $U + W$ contém tanto U quanto W, pois $u = u + 0 \in U + W$ e $w = 0 + w \in U + W$, para todo $u \in U$ e para todo $w \in W$. Além disso, o vetor nulo $0 = 0 + 0 \in U + W$.

Agora vamos mostrar que $U + W$ é subespaço de V: sejam $v_1 = u_1 + w_1$ e $v_2 = u_2 + w_2$, com $u_1, u_2 \in U$ e $w_1, w_2 \in W$ e $k \in K$. Então

$$v_1 + v_2 = u_1 + w_1 + u_2 + w_2 = (u_1 + u_2) + (w_1 + w_2) \in U + W$$

e

$$kv_1 = ku_1 + kw_1 \in U + W.$$

Logo $U + W$ é não vazio e é fechado para soma e para a multiplicação por escalar, portanto é subespaço de V.

Exemplo 2.21. *O Exemplo 2.20 acima nos permite uma construção interessante. Considere um espaço vetorial não nulo V e $v_1 \in V$, com $v_1 \neq 0$. Então o conjunto*

$$\{kv_1 \mid k \in K\}$$

de todos os múltiplos escalares de v_1 é um subespaço de V: de fato, a soma de múltiplos escalares de v_1 é também um múltiplo escalar de v_1, pois $kv_1 + lv_1 = (k + l)v_1$, e ao multiplicarmos um múltiplo escalar de v_1 por qualquer escalar continuamos com um múltiplo de v_1, pois $l(kv_1) = (lk)v_1$. Vamos denotar esse subespaço por

$$[v_1] := \{kv_1 \mid k \in K\}$$

e dizemos que $[v_1]$ é o ***subespaço de V gerado por*** *v_1.*

Por exemplo, em $\mathbb{R}^2$, o subespaço gerado por $v_1 = (1, 2)$ é a reta $[(1, 2)] = \{r(1, 2) \mid r \in \mathbb{R}\}$ gerada pelo vetor $(1, 2)$; em $\mathbb{R}^3$, o subespaço gerado por $v_1 = (-1, 3, 2)$ é a reta $[(-1, 3, 2)] = \{r(-1, 3, 2) \mid r \in \mathbb{R}\}$ gerada pelo vetor $(-1, 3, 2)$.

Agora considere um outro vetor $v_2 \in V$. Da mesma forma, temos o subespaço $[v_2] = \{lv_2 \mid l \in K\}$ gerado por v_2. Então a soma dos subespaços $[v_1]$ e $[v_2]$ é também um subespaço de V

$$[v_1] + [v_2] = \{u + w \mid u \in [v_1] \text{ e } w \in [v_2]\} = \{kv_1 + lv_2 \mid k, l \in K\},$$

*e ele é formado por todas as **combinações lineares** $kv_1 + lv_2$ de v_1 e v_2. Mudamos então a notação e escrevemos*

$$[v_1, v_2] := \{kv_1 + lv_2 \mid k, l \in K\}.$$

Por exemplo, em $\mathbb{R}^3$, o subespaço gerado por $v_1 = (-1, 3, 2)$ e $v_2 = (1, 0, 1)$ é formado por todas as combinações lineares destes vetores

$$[(-1, 3, 2), (1, 0, 1)] = \{r(-1, 3, 2) + s(1, 0, 1) \mid r, s \in \mathbb{R}\},$$

que é o plano gerado por eles.

Tomando um outro vetor v_3 e somando o subespaço $[v_3]$ gerado por v_3 com o subespaço $[v_1, v_2]$ gerado por v_1 e v_2, obtemos o subespaço

$$[v_1, v_2, v_3] := \{kv_1 + lv_2 + rv_3 \mid k, l, r \in K\}$$

gerado por v_1, v_2 e v_3, que é formado por todas as combinações lineares dos três vetores. E assim por diante.

*Em geral, dados $v_1, v_2, \ldots, v_n$ vetores de V, dizemos que um vetor $w \in V$ é **combinação linear dos vetores** $v_1, v_2, \ldots, v_n$ quando existem escalares $k_1, k_2, \ldots, k_n \in K$ tais que*

$$w = k_1v_1 + k_2v_2 + \cdots + k_nv_n.$$

E o conjunto formado por todas as combinações lineares de $v_1, v_2, \ldots, v_n$, denotado por

$$[v_1, v_2, \ldots, v_n] := \{k_1v_1 + k_2v_2 + \cdots + k_nv_n \mid k_1, k_2 \ldots, k_n \in K\},$$

*é um subespaço de V, o **subespaço gerado por** $v_1, v_2, \ldots, v_n$. De fato, $[v_1, v_2, \ldots, v_n]$ pode ser interpretado pela soma dos n subespaços $[v_i]$, $i = 1, 2, \ldots, n$, e portanto é subespaço, ou pode-se provar diretamente pela definição. Prove!*

2.2 Combinações Lineares e Conjuntos Geradores

Definição 2.22. *Seja V um espaço vetorial sobre um corpo K.*

i) *Um vetor $v \in V$ é uma **combinação linear dos vetores** $v_1, v_2, \ldots, v_n \in V$ quando existem escalares $k_1, k_2, \ldots, k_n \in K$ tais que*

$$v = k_1v_1 + k_2v_2 + \cdots + k_nv_n = \sum_{i=1}^{n} k_iv_i.$$

ii) *Seja $\mathcal{B}$ um subconjunto de V. Dizemos que $\mathcal{B}$ é um **conjunto de geradores de** V, ou que **gera** V, quando todo vetor de V pode ser escrito como uma combinação linear de um número finito de vetores de $\mathcal{B}$.*

Em particular, se $\mathcal{B} = \{v_1, v_2, \ldots, v_n\} \subset V$ é um conjunto finito, $\mathcal{B}$ gera o subespaço

$$[v_1, v_2, \ldots, v_n] = \{k_1 v_1 + k_2 v_2 + \cdots + k_n v_n \mid k_1, k_2 \ldots, k_n \in K\},$$

construído no Exemplo 2.21.

Observação 2.23. 1) Por convenção, dizemos que o conjunto vazio gera o espaço vetorial nulo $\{0\}$.

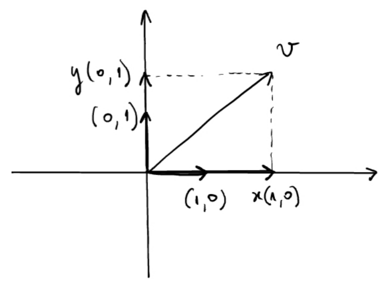
2) Todo espaço vetorial possui um conjunto de geradores, basta tomar $\mathcal{B}$ como sendo o próprio espaço.

3) Se $\mathcal{B}$ gera V, então todo subconjunto contendo $\mathcal{B}$ também gera V.

Exemplo 2.24. $V = \mathbb{R}^2$

$\mathcal{B} = \{(1,0), (0,1)\}$ gera $\mathbb{R}^2$, pois para todo $v = (x,y) \in \mathbb{R}^2$ podemos escrever

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



Exemplo 2.25. $V = \mathbb{R}^3$

$\mathcal{B} = \{(1,0,0), (0,1,0), (0,0,1)\}$ *gera* $\mathbb{R}^3$, *pois para todo* $v = (x,y,z) \in \mathbb{R}^3$ *podemos escrever* $v = x(1,0,0) + y(0,1,0) + z(0,0,1)$.

$\mathcal{B}' = \{(1,0,1), (1,1,0), (1,1,1), (-1,0,0), (-1,-1,0)\}$ *gera* $\mathbb{R}^3$. *De fato,* $\forall v = (a,b,c) \in \mathbb{R}^3$,

$$(a,b,c) = x(1,0,1) + y(1,1,0) + z(1,1,1) + t(-1,0,0) + r(-1,-1,0)$$

$$\Leftrightarrow \quad S : \left\{ \begin{array}{l} x + y + z - t - r = a \\ \quad\;\; y + z \quad\;\; - r = b \\ x \quad\;\; + z \qquad\quad = c \end{array} \right. \quad \textit{tem solução}$$

E

$$S \sim_{L_3 \to L_3 - L_1} \left\{ \begin{array}{l} x + y + z - t - r = a \\ \quad\;\; y + z \quad\;\; - r = b \\ \quad -y \quad\;\; + t + r = c - a \end{array} \right. \sim_{L_3 \to L_3 + L_2} \left\{ \begin{array}{l} x + y + z - t - r = a \\ \quad\;\; y + z \quad\;\; - r = b \\ \qquad\quad z + t \quad\; = c - a + b \end{array} \right.$$

que tem solução (infinitas soluções), para quaisquer $a, b, c \in \mathbb{R}$.

Exemplo 2.26. *$V = \mathcal{P}(K)$, o espaço dos polinômios sobre K*

$\mathcal{B} = \{1, x, x^2, \ldots, x^n, \ldots\}$ gera $\mathcal{P}(K)$, pois dado $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m$, com $a_i \in K$ e $n \in \mathbb{N}$, podemos escrever

$$f(x) = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 + \cdots + a_m \cdot x^m,$$

uma combinação linear de um número finito de elementos de $\mathcal{B}$.

$\tilde{\mathcal{B}} = \{2, 1+x, 1+x^2, \ldots, 1+x^n, \ldots\}$ gera $\mathcal{P}(K)$. De fato, dado $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m$, com $a_i \in K$ e $n \in \mathbb{N}$, temos que

$$\begin{aligned} a_0 + a_1x + a_2x^2 + \cdots + a_mx^m &= b_0 \cdot 2 + b_1(1+x) + b_2(1+x^2) + \cdots + b_m(1+x^m) \\ &= (2b_0 + b_1 + b_2 + \cdots + b_m) + b_1x + b_2x^2 + \cdots + b_mx^m \end{aligned}$$

$$\Leftrightarrow \quad S: \left\{ \begin{array}{cccccl} 2b_0 & +b_1 & +b_2 & +\cdots+ & b_m & = a_0 \\ & b_1 & & & & = a_1 \\ & & b_2 & & & = a_2 \\ & & & \ddots & & \\ & & & & b_m & = a_m \end{array} \right.$$

e o sistema linear S obtido tem solução (única), para quaisquer $a_0, a_1, a_2, \ldots, a_m \in K$.

Exemplo 2.27. **a)** *$V = \mathbb{C}^2$ (espaço vetorial sobre $\mathbb{C}$)*

$\mathcal{B} = \{(1,0), (0,1)\}$ gera $\mathbb{C}^2$. De fato, $\forall (z_1, z_2) \in \mathbb{C}^2$,

$$(z_1, z_2) = z_1(1,0) + z_2(0,1).$$

b) *$V = \mathbb{C}^2$ como espaço vetorial sobre $\mathbb{R}$*

$\mathcal{B} = \{(1,0), (i,0), (0,1), (0,i)\}$ gera $\mathbb{C}^2$ sobre $\mathbb{R}$. De fato, $\forall z_1 = a_1 + b_1i,\ z_2 = a_2 + b_2i$, onde $a_1, a_2, b_1, b_2 \in \mathbb{R}$,

$$(z_1, z_2) = a_1(1,0) + b_1(i,0) + a_2(0,1) + b_2(0,i).$$

Exemplo 2.28. *$V = \mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$*

$\mathbb{Q}[\sqrt{2}]$ é espaço vetorial sobre $\mathbb{Q}$;

$\mathcal{B} = \{1, \sqrt{2}\}$ gera $\mathbb{Q}[\sqrt{2}]$.

Exemplo 2.29. *$V = \mathcal{M}_2(K)$*

$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ gera $\mathcal{M}_2(K)$ sobre K.

Exemplo 2.30. *Seja V um espaço vetorial sobre K e*

$$U := [u_1, u_2, \ldots, u_m] \quad e \quad W := [w_1, w_2, \ldots, w_k],$$

$u_i, w_j \in K$, subespaços de V. Então

$$U + W = [u_1, u_2, \ldots, u_m, w_1, w_2, \ldots, w_k].$$

(Prove!) *Em geral, se $\mathcal{B}$ gera um subespaço U e $\mathcal{B}'$ gera um subespaço W, então $\mathcal{B} \cup \mathcal{B}'$ gera $U + W$.*

Exemplo 2.31. *Vamos determinar conjuntos de geradores para os espaços vetoriais abaixo:*

a) $U_1 = \{(x, y) \in \mathbb{R}^2 \mid x = 2y\}$

Primeiro, temos que U_1 é um espaço vetorial. Para mostrar isso, mostre que U_1 é um subespaço de $\mathbb{R}^2$, logo é um espaço vetorial. Geometricamente, note que U_1 é uma reta passando pela origem.

Agora vamos ao conjunto de geradores de U_1. Temos que

$$U_1 = \{(x, y) \in \mathbb{R}^2 \mid x = 2y\} = \{(2y, y) \mid y \in \mathbb{R}\} = \{y(2, 1) \mid y \in \mathbb{R}\} = [(2, 1)],$$

então U_1 é gerado pelo vetor $(2, 1)$ (é a reta gerada pelo vetor $(2, 1)$).

b) $U_2 = \{(x, y, z) \in \mathbb{R}^3 \mid x + y - 2z = 0\}$

Primeiro, verifique que U_2 é um subespaço de $\mathbb{R}^3$, logo é um espaço vetorial.

Agora vamos ao conjunto de geradores. Temos

$$\begin{aligned} U_2 &= \{(x, y, z) \in \mathbb{R}^3 \mid x + y - 2z = 0\} \\ &= \{(x, y, z) \in \mathbb{R}^3 \mid x = 2z - y\} \\ &= \{(2z - y, y, z) \mid y, z \in \mathbb{R}\} \\ &= \{z(2, 0, 1) + y(-1, 1, 0) \mid y, z \in \mathbb{R}\} \\ &= [(2, 0, 1), (-1, 1, 0)] \end{aligned}$$

então U_2 é gerado pelos vetores $(2, 0, 1)$ e $(-1, 1, 0)$ (é o plano gerado por estes dois vetores).

c) $U_3 = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$

Verifique que U_3 é um subespaço do espaço vetorial das matrizes 2×2 $\mathcal{M}_2(\mathbb{R})$, e assim é um espaço vetorial.

Agora vamos encontrar um conjunto de geradores. Temos

$$\begin{aligned} U_3 &= \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \mid a, b \in \mathbb{R} \right\} \\ &= \left\{ a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mid a, b \in \mathbb{R} \right\} \\ &= \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right]. \end{aligned}$$

d) $W = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + y + t = 0,\ 2x - y + z = 0$ *e* $3x + z + t = 0\}$

W é um espaço vetorial, pois é subespaço de $\mathbb{R}^4$, uma vez que é o conjunto-solução do sistema linear homogêneo

$$S : \left\{ \begin{array}{rrrrl} x & +y & & +t & = 0 \\ 2x & -y & +z & & = 0 \\ 3x & & +z & +t & = 0 \end{array} \right.$$

Agora vamos encontrar um conjunto de geradores para W. Para isso, vamos usar que W é exatamente o conjunto solução do sistema linear homogêneo S. Então vamos encontrar o conjunto solução de S:

$$S : \left\{ \begin{array}{rrrrl} x & +y & & +t & = 0 \\ 2x & -y & +z & & = 0 \\ 3x & & +z & +t & = 0 \end{array} \right. \sim_{\substack{L_2 \to L_2 - 2L_1 \\ L_3 \to L_3 - 3L_1}} \left\{ \begin{array}{rrrrl} x & +y & & +t & = 0 \\ & -3y & +z & -2t & = 0 \\ & -3y & +z & -2t & = 0 \end{array} \right.$$

$$\sim_{L_3 \to L_3 - L_2} \left\{ \begin{array}{rrrrl} x & +y & & +t & = 0 \\ & -3y & +z & -2t & = 0 \end{array} \right.$$

Tomando y e t como variáveis livres, obtemos $z = 3y + 2t$ e $x = -y - t$. Assim,

$$W = \{(-y - t, y, 3y + 2t, t) \mid y, t \in \mathbb{R}\}.$$

Descrevendo W dessa maneira, encontramos um conjunto de geradores:

$$\begin{aligned} W &= \{(-y - t, y, 3y + 2t, t) \mid y, t \in \mathbb{R}\} \\ &= \{y(-1, 1, 3, 0) + t(-1, 0, 2, 1) \mid y, t \in \mathbb{R}\} \\ &= [(-1, 1, 3, 0), (-1, 0, 2, 1)]. \end{aligned}$$

2.3 Dependência e Independência Linear

Vamos voltar no Exemplo 2.25. Ao mostrarmos que

$$\mathcal{B}' = \{(1,0,1),(1,1,0),(1,1,1),(-1,0,0),(-1,-1,0)\}$$

gera $\mathbb{R}^3$, vimos que podemos escrever um vetor $v = (a,b,c) \in \mathbb{R}^3$ como combinação linear dos vetores de $\mathcal{B}'$,

$$(a,b,c) = x(1,0,1) + y(1,1,0) + z(1,1,1) + t(-1,0,0) + r(-1,-1,0),$$

de infinitas maneiras, uma vez que o sistema linear resultante da combinação linear tem infinitas soluções. Isso significa que o conjunto $\mathcal{B}'$ contém vetores que já são gerados pelos outros vetores do próprio conjunto $\mathcal{B}'$, que podem ser retirados de $\mathcal{B}$ sem que se altere o espaço gerado.

Em geral, seja V um espaço vetorial sobre K, $\mathcal{B}$ um conjunto de geradores de V e $v \in V$. E suponha que existam duas combinações lineares distintas para v,

$$v = a_1v_1 + a_2v_2 + \cdots + a_mv_m = b_1v_1 + b_2v_2 + \cdots + b_mv_m,$$

com $a_i, b_i \in K$ e $v_i \in \mathcal{B}$. $i = 1,2,\ldots,m$, ou seja, $a_i \neq b_i$ para algum i. Sem perda de generalidade, suponha que $a_1 \neq b_1$. Então

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = b_1v_1 + b_2v_2 + \cdots + b_mv_m \quad \Rightarrow \quad v_1 = \frac{a_2 - b_2}{a_1 - b_1}v_2 + \cdots + \frac{a_m - b_m}{a_1 - b_1}v_m$$

de onde vemos que v_1 é combinação linear de $v_2,\ldots,v_m$. Com isso, podemos reescrever v como combinação linear de $v_2,\ldots,v_m$e obtemos que $\mathcal{B}\setminus\{v_1\}$ gera V.

O fato de um vetor do conjunto de geradores já ser gerado por outros vetores do conjunto é também caracterizado da seguinte maneira.

Proposição 2.32. *Seja V um espaço vetorial sobre K e $v_1, v_2, \ldots, v_n \in V$. As seguintes afirmações são equivalentes:*

a) *Para algum i, v_i é combinação linear de $v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$, ou seja,*

$$v_i \in [v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n].$$

b) *Existe uma combinação linear não trivial de $v_1, v_2, \ldots, v_n$ dando o vetor nulo, ou seja, existem escalares $k_1, k_2, \ldots, k_n \in K$ **não todos nulos** tais que*

$$k_1v_1 + k_2v_2 + \cdots + k_nv_n = 0 \quad \textit{(vetor nulo)}.$$

Demonstração. Se v_i é combinação linear de $v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$, temos

$$v_i = a_1v_1 + \cdots + a_{i-1}v_{i-1} + a_{i+1}v_{i+1} + \cdots + a_nv_n,$$

com $a_i \in K$. Assim, obtemos que

$$-a_1v_1 - \cdots - a_{i-1}v_{i-1} + 1.v_i - a_{i+1}v_{i+1} - \cdots - a_nv_n = 0$$

é uma combinação linear dos vetores $v_1, v_2, \ldots, v_n$, com escalares não todos nulos, dando o vetor nulo.

Reciprocamente, se existem escalares não todos nulos $k_1, k_2, \ldots, k_n \in K$ tais que $k_1v_1 + k_2v_2 + \cdots + k_nv_n = 0$, como $k_1, k_2, \ldots, k_n$ não são todos nulos, temos algum $k_i \neq 0$. Assim, da combinação linear $k_1v_1 + k_2v_2 + \cdots + k_nv_n = 0$, obtemos

$$v_i = -\frac{k_1}{k_i}v_1 - \cdots - \frac{k_{i-1}}{k_i}v_{i-1} - \frac{k_{i+1}}{k_i}v_{i+1} - \cdots - \frac{k_n}{k_i}v_n.$$

■

Definição 2.33. *Seja V um espaço vetorial sobre K e $\mathcal{B} \subset V$.*

i) *$\mathcal{B}$ é dito **linearmente dependente** (ou **L.D.**) quando existem vetores distintos $v_1, v_2, \ldots, v_n \in \mathcal{B}$ e escalares $k_1, k_2, \ldots, k_n$ não todos nulos tais que $k_1v_1 + k_2v_2 + \cdots + k_nv_n = 0$.*

ii) *$\mathcal{B}$ é dito **linearmente independente** (ou **L.I.**) quando não é L.D., ou seja, $k_1v_1 + k_2v_2 + \cdots + k_nv_n = 0$, com $v_i \in \mathcal{B}$ e $k_i \in K$, só é possível com $k_1 = k_2 = \cdots = k_n = 0$.*

Da equivalência descrita na Proposição 2.32, obtemos que $\mathcal{B}$ é L.D. quando um de seus vetores é combinação linear de outros vetores em $\mathcal{B}$ e um tal vetor pode ser retirado do conjunto de geradores sem alterar o espaço gerado. E $\mathcal{B}$ é L.I. quando isso não acontece, ou seja, nenhum vetor de $\mathcal{B}$ é combinação linear de outros vetores em $\mathcal{B}$ e a retirada de vetores de $\mathcal{B}$ altera o espaço gerado.

Observação 2.34. 1) Por convenção, o conjunto vazio é L.I.

2) Todo conjunto que contém o vetor nulo é L.D.

3) Note que a Definição 2.33 depende do corpo K fixado.

Por exemplo,

$$\{(1,0), (i,0), (0,1), (0,i)\} \subset \mathbb{C}^2$$

é L.D. sobre $\mathbb{C}$, pois $(i,0) = i(1,0)$. Mas é L.I. sobre $\mathbb{R}$, pois

$$a(1,0) + b(i,0) + c(0,1) + d(0,i) = (0,0), \quad \textit{com } a,b,c,d \in \mathbb{R}$$

$$\Leftrightarrow a + bi = 0 \quad e \quad c + di = 0, \quad \textit{com } a,b,c,d \in \mathbb{R}$$

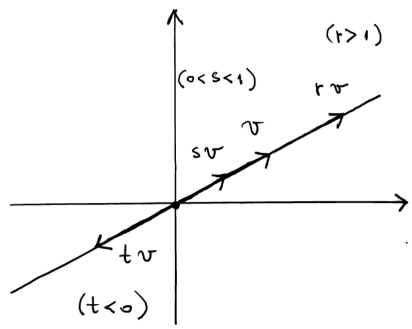
$$\Leftrightarrow \quad a = b = c = d = 0.$$

4) Todo espaço vetorial V não nulo possui um conjunto L.I. Basta tomar um vetor não nulo $v \in V$, assim $\{v\}$ é L.I.

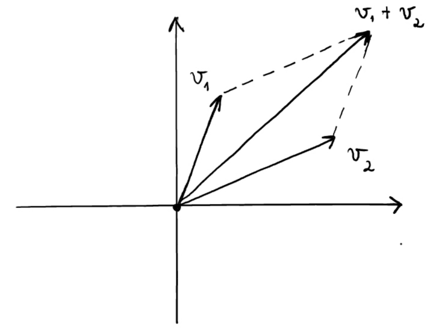
5) Todo subconjunto de um conjunto L.I. é L.I.

6) Todo conjunto que contém um subconjunto L.D é também L.D.

Exemplo 2.35. **(i)** *No espaço vetorial $\mathbb{R}^2$, os vetores $v = (a, b)$ são comumente representados no plano-xy pelo segmento orientado com ponto inicial na origem $0 = (0, 0)$ e ponto final (a, b). Com isso, a adição de vetores e a multiplicação por escalar têm interpretações geométricas:*



(a) Multiplicação por escalar



(b) Adição de vetores: Regra do paralelogramo

- *Se $v = (0, 0)$, então $\{v\}$ é L.D., pois $r(0, 0) = (0, 0)$, $\forall r \in \mathbb{R}, r \neq 0$.*
- *Se $v = (x, y) \neq (0, 0)$, então $\{v\}$ é L.I., pois*

$$rv = 0 \Leftrightarrow r(x, y) = (0, 0) \Leftrightarrow \begin{cases} rx = 0 \\ ry = 0 \end{cases} \overset{(x,y)\neq(0,0)}{\Leftrightarrow} r = 0.$$

  E $[v]$ é a reta pela origem na direção de v.

- *$\{v_1, v_2\} \subset \mathbb{R}^2$ é L.D. se, e somente se, $v_1 = kv_2$ ou $v_2 = kv_1$, o que é equivalente a v_1 e v_2 estarem em uma mesma reta.*

  E $[v_1, v_2] = [v_i]$, $i = 1$ ou $i = 2$, é uma reta pela origem.

- *$\{v_1, v_2\} \subset \mathbb{R}^2$ é L.I. se, e somente se, v_1 e v_2 não estão em uma mesma reta.*

  Ou ainda, se $v_1 = (a_1, b_1)$ e $v_2 = (a_2, b_2)$,

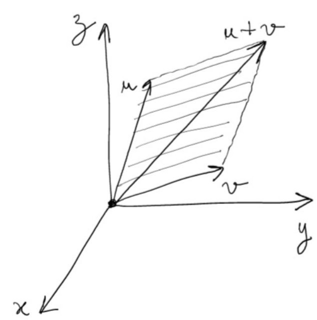
$$\begin{aligned} \{v_1, v_2\} \text{ é L.I.} &\Leftrightarrow rv_1 + sv_2 = 0, \text{ somente se } r = s = 0 \\ &\Leftrightarrow r(a_1, b_1) + s(a_2, b_2) = (0, 0), \text{ somente se } r = s = 0 \\ &\Leftrightarrow \begin{cases} a_1 r + a_2 s = 0 \\ b_1 r + b_2 s = 0 \end{cases} \text{só possui solução nula } r = s = 0 \\ &\Leftrightarrow \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \text{ é inversível.} \end{aligned}$$

  Além disso, quando $\{v_1, v_2\} \subset \mathbb{R}^2$ é L.I., todo vetor $v = (a, b) \in \mathbb{R}^2$ é combinação linear de v_1 e v_2. De fato,

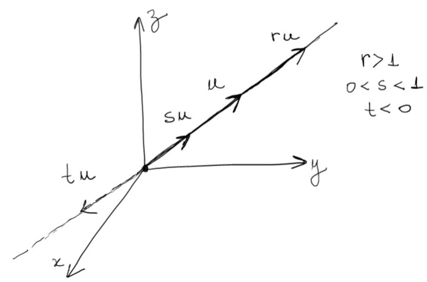
$$rv_1 + sv_2 = v \;\Leftrightarrow\; r(a_1, b_1) + s(a_2, b_2) = (a, b) \;\Leftrightarrow\; \begin{cases} a_1 r + a_2 s = a \\ b_1 r + b_2 s \; = b \end{cases}$$

Como $\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$ *é inversível, o sistema tem solução (única) para quaisquer* $a, b \in \mathbb{R}$.
Então existem $r, s \in \mathbb{R}$ *tais que* $rv_1 + sv_2 = v, \ \forall v \in \mathbb{R}^2$.
Assim $[v_1, v_2] = \mathbb{R}^2$.

(ii) *No espaço vetorial* $\mathbb{R}^3$*, os vetores* $v = (a, b, c)$ *podem ser representados como segmentos orientados no espaço de coordenadas-xyz com ponto inicial na origem* $0 = (0, 0, 0)$ *e ponto final* (a, b, c)*. Com isso, em* $\mathbb{R}^3$ *também temos a regra do paralelogramo para a soma de vetores. Para obtermos o vetor* $u + v$*, tomamos o plano contendo o dois vetores e aplicamos a regra do paralelogramo neste plano. E a multiplicação por escalar mantém o vetor sobre a reta determinada por ele.*



(c) Adição de vetores: Regra do paralelogramo no plano determinado pelos vetores



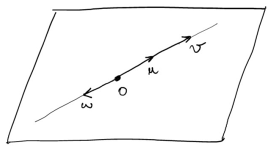
(d) Multiplicação por escalar

- *Se* $v = (0, 0, 0)$*, então* $\{v\}$ *é L.D.*
- *Se* $v \neq (0, 0, 0)$*, então* $\{v\}$ *é L.I., pois* $rv = 0$ *se, e somente se,* $r = 0$.

  E $[v]$ *é a reta pela origem na direção de* v.
- $\{v_1, v_2\} \subset \mathbb{R}^3$ *é L.D. se, e somente se,* $v_1 = kv_2$ *ou* $v_2 = kv_1$*, o que é equivalente a* v_1 *e* v_2 *estarem em uma mesma reta.*

  E $[v_1, v_2] = [v_i]$*,* $i = 1$ *ou* $i = 2$*, é uma reta pela origem.*
- $\{v_1, v_2\} \subset \mathbb{R}^3$ *é L.I. se, e somente se,* v_1 *e* v_2 *não estão em uma mesma reta.*

  Neste caso, $[v_1, v_2] = \{r_1v_1 + r_2v_2 \mid r_1, r_2 \in \mathbb{R}\}$ *é o plano contendo* v_1 *e* v_2 *(unicamente determinado por* v_1 *e* v_2*).*
- $\{v_1, v_2, v_3\} \subset \mathbb{R}^3$ *é L.D. se, e somente se, um dos vetores é combinação linear dos outros, o que é equivalente a* v_1*,*v_2 *e* v_3 *estarem em um mesmo plano.*

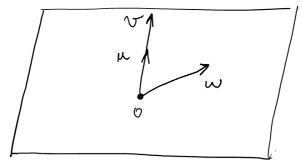


(e) Os três vetores estão em uma mesma reta.
Neste caso, dois vetores são múltiplos de um dos vetores, digamos

$$v = ru \text{ e } w = su.$$

De $v = ru$ obtemos que

$$-r.u + 1.v + 0.w = 0$$

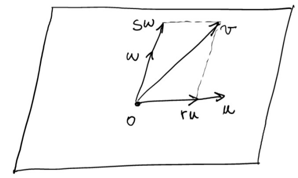


(f) Dois vetores estão sobre uma mesma reta e o outro está fora da reta. Neste caso, um dos vetores é múltiplo de outro, digamos

$$v = ru$$

Com isso, obtemos

$$-r.u + 1.v + 0.w = 0$$



(g) Nenhum dos vetores é múltiplo do outro, mas os três estão no mesmo plano. Neste caso,

$$v = ru + sw$$

e com isso obtemos

$$-r.u + 1.v - s.w = 0$$

- *$\{v_1, v_2, v_3\} \subset \mathbb{R}^3$ é L.I. se, e somente se, v_1, v_2 e v_3 não estão em um mesmo plano.*

  Ou ainda, se $v_1 = (a_1, b_1, c_1)$, $v_2 = (a_2, b_2, c_2)$ e $v_3 = (a_3, b_3, c_3)$,

$$\begin{aligned} \{v_1, v_2, v_3\} \text{ é L.I.} \quad &\Leftrightarrow \quad rv_1 + sv_2 + tv_3 = 0, \text{ somente se } r = s = t = 0 \\ &\Leftrightarrow \quad \begin{cases} a_1 r + a_2 s + a_3 t = 0 \\ b_1 r + b_2 s + b_3 t = 0 \\ c_1 r + c_2 s + c_3 t = 0 \end{cases} \text{ só possui solução nula } r = s = t = 0 \\ &\Leftrightarrow \quad \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \text{ é inversível.} \end{aligned}$$

  Além disso, quando $\{v_1, v_2, v_3\} \subset \mathbb{R}^3$ é L.I., todo vetor $v = (a, b, c) \in \mathbb{R}^3$ é combinação linear de v_1, v_2 e v_3. De fato,

$$rv_1 + sv_2 + tv_3 = v \;\Leftrightarrow\; \begin{cases} a_1 r + a_2 s + a_3 t = a \\ b_1 r + b_2 s + b_3 t = b \\ c_1 r + c_2 s + c_3 t = c \end{cases}$$

  Como $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$ é inversível, o sistema tem solução (única) para quaisquer $a, b, c \in \mathbb{R}$. Então existem $r, s, t \in \mathbb{R}$ tais que $rv_1 + sv_2 + tv_3 = v, \;\; \forall v \in \mathbb{R}^3$.
  Assim $[v_1, v_2, v_3] = \mathbb{R}^3$.

Exemplo 2.36. *Seja $A = (a_{ij}) \in \mathcal{M}_n(K)$. Os* ***vetores-coluna de*** *A são os vetores em K^n dados por $v_j = (a_{1j}, a_{2j}, \ldots, a_{nj})$, $j = 1, 2, \ldots, n$. Temos que*

$$A \text{ é inversível} \;\Leftrightarrow\; \text{os vetores-coluna de } A \text{ são L.I.}$$

De fato, temos que A é inversível se, e somente se,

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

tem solução única $x_1 = x_2 = \cdots = x_n = 0$. Mas a igualdade matricial acima pode ser reescrita como

$$x_1(a_{11}, a_{21}, \ldots, a_{n1}) + x_2(a_{12}, a_{22}, \ldots, a_{n2}) + \cdots + x_n(a_{1n}, a_{2n}, \ldots, a_{nn}) = (0, 0, \ldots, 0).$$

Então A é inversível se, e somente, se a combinação linear acima tem solução única $x_1 = x_2 = \cdots = x_n = 0$, ou seja, se, e somente se, os vetores-coluna de A são L.I.

Exemplo 2.37. *$\mathcal{B} = \{2, 1+x, 1+x^2, \ldots, 1+x^n\} \subset \mathcal{P}_n(K)$ é L.I.*

De fato, temos que

$$\begin{aligned} & a_0 \cdot 2 + a_1(1+x) + a_2(1+x^2) + \cdots + a_n(1+x^n) = 0 \\ \Leftrightarrow \quad & (2a_0 + a_1 + a_2 + \cdots + a_m) + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0 \\ \Leftrightarrow \quad & S : \left\{ \begin{array}{cccccl} 2a_0 & +a_1 & +a_2 & +\cdots+ & a_n & = 0 \\ & a_1 & & & & = 0 \\ & & a_2 & & & = 0 \\ & & & \ddots & & \\ & & & & a_n & = 0 \end{array} \right. \end{aligned}$$

e o sistema linear S obtido tem solução (única) $a_0 = a_1 = a_2 = \cdots = a_n = 0$.

Exemplo 2.38. *$\mathcal{B} = \{\operatorname{sen} x, \cos x\} \subset \mathcal{C}([0, 2\pi], \mathbb{R})$ é L.I. no espaço vetorial sobre $\mathbb{R}$ das funções contínuas $\mathcal{C}([0, 2\pi], \mathbb{R})$.*

De fato, se

$$r \operatorname{sen} x + s \cos x = 0$$

(igualdade de funções, onde aqui o vetor nulo 0 é a função identicamente nula), então

$$r \operatorname{sen} x + s \cos x = 0, \quad \forall x \in [0, 2\pi].$$

Para $x = 0$, obtemos que $s = 0$. E para $x = \pi/2$, obtemos $r = 0$. Portanto a combinação linear $r \operatorname{sen} x + s \cos x = 0$ só é possível com $r = s = 0$ e, assim, $\mathcal{B}$ é L.I.

2.4 Base e Dimensão

Conjuntos geradores de espaços vetoriais que são linearmente independentes recebem um nome especial.

Definição 2.39. *Seja V um espaço vetorial sobre um corpo K e $\mathcal{B} \subset V$, um conjunto de vetores em V. Dizemos que $\mathcal{B}$ é uma **base de** V quando*

i) $\mathcal{B}$ gera V; e

ii) $\mathcal{B}$ é linearmente independente.

Exemplo 2.40. *Seja K um corpo e $V = K^n$. Considere os vetores $e_1, e_2, \ldots, e_n \in K^n$ definidos por*

$$e_1 := (1, 0, 0, \ldots, 0),\ e_2 := (0, 1, 0 \ldots, 0), \ldots,\ e_n := (0, \ldots, 0, 1).$$

Temos que

$$\mathcal{C} = \{e_1, e_2, \ldots, e_n\}$$

*é uma base de K^n (prove!), chamada de **base canônica** de K^n.*

Exemplo 2.41. *Voltando ao Exemplo 2.35, temos que qualquer par de vetores $v_1, v_2 \in \mathbb{R}^2$ linearmente independentes forma uma base $\{v_1, v_2\}$ de $\mathbb{R}^2$. E qualquer conjunto $\{v_1, v_2, v_3\} \subset \mathbb{R}^3$ linearmente independente é uma base de $\mathbb{R}^3$.*

Mais geralmente, segue do Exemplo 2.36 que qualquer conjunto com n vetores $\mathcal{B} = \{v_1, v_2, \ldots, v_n\} \subset K^n$ (K um corpo qualquer) linearmente independente é uma base de K^n. De fato, tomando $v_j = (a_{1j}, a_{2j}, \ldots, a_{nj})$, $j = 1, 2, \ldots, n$, a matriz $A = (a_{ij})$ é inversível. Assim, para qualquer $v = (b_1, b_2, \ldots, b_n) \in K^n$, o sistema linear $AX = B$ tem solução e, portanto, v pode ser escrito como combinação linear de $v_1, v_2, \ldots, v_n$. Logo $\mathcal{B}$ gera K^n.

Exemplo 2.42. $\mathcal{B} = \{1, x, x^2, \ldots, x^n, \ldots\}$ *e* $\mathcal{B}' = \{2, 1+x, 1+x^2, \ldots, 1+x^n, \ldots\}$ *são bases de $\mathcal{P}(K)$, pois são L.I.* (Exemplo 2.37) *e geram $\mathcal{P}(K)$* (Exemplo 2.26). *Note que $\mathcal{B}$ e $\mathcal{B}'$ possuem infinitos elementos. Além disso, observe que $\mathcal{P}(K)$ não pode ser gerado por um número finito de elementos (Por quê? Prove.)*

$\mathcal{B} = \{1, x, x^2, \ldots, x^n\}$ *e* $\mathcal{B}' = \{2, 1+x, 1+x^2, \ldots, 1+x^n\}$ *são bases (finitas) de $\mathcal{P}_n(K)$.*

Exemplo 2.43. $\{(1, 0), (0, 1)\}$ *é uma base de $\mathbb{C}^2$ sobre $\mathbb{C}$;*

$\{(1, 0), (i, 0), (0, 1), (0, i)\}$ *é base de $\mathbb{C}^2$ sobre $\mathbb{R}$.*

Apresentamos a seguir um resultado que é o coração desta seção. É ele que nos permite definir a dimensão de um espaço vetorial finitamente gerado.

Teorema 2.44. *Seja V um espaço vetorial sobre K gerado por um conjunto finito de vetores $\{v_1, v_2, \ldots, v_m\}$. Então todo conjunto linearmente independente de vetores em V é finito e contém no máximo m elementos.*

Demonstração. Vamos mostrar que se $\mathcal{B}$ é um subconjunto de V com mais do que m vetores, então $\mathcal{B}$ é L.D. Para isso, basta mostrar que todo subconjunto que tem $m+1$ elementos é L.D. (já que todo conjunto que contém um subconjunto L.D é também L.D.).

Seja $\mathcal{B} = \{u_1, u_2, \ldots, u_m, u_{m+1}\} \subset V$. Vamos mostrar que $\mathcal{B}$ é L.D. De fato, como $v_1, v_2, \ldots, v_m$ geram V, cada $u_i \in \mathcal{B}$ pode ser escrito como combinação linear dos v_i's:

$$\begin{aligned}
&u_1 = a_{11}v_1 + a_{12}v_2 + \cdots + a_{1m}v_m \\
&u_2 = a_{21}v_1 + a_{22}v_2 + \cdots + a_{2m}v_m \\
&\;\vdots \\
&u_i = a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{im}v_m = \textstyle\sum_{j=1}^{m} a_{ij}v_j, \quad a_{ij} \in K.
\end{aligned}$$

Com isso, uma combinação linear dos u_i's dando o vetor nulo,

$$k_1u_1 + k_2u_2 + \cdots + k_mu_m + k_{m+1}u_{m+1} = \sum_{i=1}^{m+1} k_iu_i = 0, \quad \text{com } k_i \in K$$

pode ser escrita como

$$\begin{aligned}
\sum_{i=1}^{m+1} k_iu_i = 0 \quad &\Leftrightarrow \quad \sum_{i=1}^{m+1} k_i \sum_{j=1}^{m} a_{ij}v_j = 0 \\
\Leftrightarrow \quad \sum_{i=1}^{m+1}\sum_{j=1}^{m} k_ia_{ij}v_j = 0 \quad &\Leftrightarrow \quad \sum_{j=1}^{m}\left(\sum_{i=1}^{m+1} k_ia_{ij}\right)v_j = 0.
\end{aligned}$$

Para mostrar que $\mathcal{B}$ é L.D. temos que mostrar que existem $k_1, k_2, \ldots, k_{m+1} \in K$ não todos nulos satisfazendo a combinação linear $\sum_{i=1}^{m+1} k_iu_i = 0$. Para isso, vamos mostrar que existem $k_1, k_2, \ldots, k_{m+1} \in K$ não todos nulos tais que

$$\sum_{i=1}^{m+1} k_ia_{ij} = k_1a_{1j} + k_2a_{2j} + \cdots + k_{m+1}a_{m+1,j} = 0, \quad \text{para cada } j = 1, 2, \ldots, m.$$

De fato, abrindo as somas acima para $j = 1, 2, \ldots, m$, obtemos

$$\begin{cases}
a_{11}k_1 + a_{21}k_2 + \cdots + a_{m+1,1}k_{m+1} = 0 \\
a_{12}k_1 + a_{22}k_2 + \cdots + a_{m+1,2}k_{m+1} = 0 \\
\;\vdots \\
a_{1m}k_1 + a_{2m}k_2 + \cdots + a_{m+1,m}k_{m+1} = 0
\end{cases}$$

um sistema linear homogêneo nas $m+1$ incógnitas $k_1, k_2, \ldots, k_{m+1}$, com m equações. E um sistema linear homogêneo com mais incógnitas do que equações tem infinitas soluções, logo tem solução não nula. Então existem $k_1, k_2, \ldots, k_{m+1} \in K$ não todos nulos tais que $\sum_{i=1}^{m+1} k_ia_{ij} = k_1a_{1j} + k_2a_{2j} + \cdots + k_{m+1}a_{m+1,j} = 0$, para cada $j = 1, 2, \ldots, m$. Com isso, pelas equivalências acima, estes $k_1, k_2, \ldots, k_{m+1} \in K$ não todos nulos satisfazem a combinação linear

$$k_1u_1 + k_2u_2 + \cdots + k_mu_m + k_{m+1}u_{m+1} = 0,$$

e portanto $\mathcal{B}$ é L.D. ∎

Definição 2.45. *Um espaço vetorial V sobre K é dito de* ***dimensão finita*** *quando V possui uma base finita, ou seja, uma base com um número finito de vetores.*

Corolário 2.46. *Se V é um espaço vetorial sobre K de dimensão finita, então quaisquer duas bases de V são finitas e possuem a mesma quantidade de elementos.*

Demonstração. Como V tem dimensão finita, V possui uma base $\mathcal{B}$ com um número finito n de vetores. Então, pelo Teorema 2.44, qualquer subconjunto de V com mais do que n vetores é L.D. Portanto qualquer outra base $\mathcal{B}'$ de V tem um número finito $m \leq n$ elementos (já que uma base é um subconjunto L.I. de V). Por outro lado, como $\mathcal{B}'$, por ser base, também gera V, qualquer conjunto com mais do que m elementos de V é L.D. Com isso, devemos ter $n \leq m$. Assim $m = n$, ou seja, $\mathcal{B}$ e $\mathcal{B}'$ possuem a mesma quantidade de elementos. ∎

Definição 2.47. *Seja V um espaço vetorial sobre K de dimensão finita. A* ***dimensão de V*** *(sobre K) é definida como o número de elementos de uma base de V.*

Notação: $\dim V$ ou $\dim_K V$ (quando for importante destacar o corpo).

Observação: $\dim_K\{0\} = 0$, para qualquer corpo K (a base é o conjunto vazio).

Exemplo 2.48. $\dim K^n = n$;

Em particular, $\dim \mathbb{R}^n = n$, $\dim \mathbb{C}^n = n$.

$\dim_{\mathbb{R}} \mathbb{C} = 2$;

$\dim_{\mathbb{R}} \mathbb{C}^n = 2n$;

$\dim \mathcal{P}_n(K) = n + 1$;

$\dim \mathcal{M}_{m\times n}(K) = mn$.

Para justificar a dimensão de um dado espaço vetorial, basta exibir uma base do espaço. Fica como exercício a justificativa das dimensões dos espaço acima.

Corolário 2.49. *Seja V um espaço vetorial sobre K de dimensão finita igual a n. Então*

i) todo suconjunto de vetores de V que contém mais do que n vetores é L.D.;

ii) nenhum subconjunto de V com menos do que n vetores gera V
(se gerasse, então todo subconjunto com n vetores seria L.D.);

Ainda não garantimos que todo espaço vetorial tem base. Vamos fazer aqui a prova de que todo espaço vetorial finitamente gerado tem base. Espaços vetoriais que não podem ser gerados por um número finito de elementos também sempre possuem base, mas para estes espaços a demonstração é diferente e comumentemente usa o Lema de Zorn. Deixo aqui o convite para pesquisarem a demonstração do caso geral e o Lema de Zorn.

A proposição a seguir, que usaremos na prova de que todo espaço vetorial finitamente gerados tem base, vale para espaços vetoriais em geral. Mas por ela já conseguimos perceber que o argumento que vamos usar para espaços vetoriais finitamente gerados não funciona para espaços vetoriais não finitamente gerados.

Proposição 2.50. *Seja V um espaço vetorial sobre K e considere $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ um subconjunto L.I. de V. Se $u \notin [v_1, v_2, \ldots, v_m]$, então o conjunto $\{v_1, v_2, \ldots, v_m, u\}$ é L.I..*

Demonstração. Suponha que

$$x_1 v_1 + x_2 v_2 + \cdots + x_m v_m + x_{m+1} u = 0, \quad \text{com } x_i \in K.$$

Queremos mostrar que $x_1 = x_2 = \cdots = x_{m+1} = 0$. Se $x_{m+1} \neq 0$, então

$$u = -\frac{x_1}{x_{m+1}} v_1 - \frac{x_2}{x_{m+1}} v_2 + \cdots + -\frac{x_n}{x_{m+1}} v_m,$$

o que contradiz o fato de $u \notin [v_1, v_2, \ldots, v_m]$. Então $x_{m+1} = 0$ e com isso

$$x_1 v_1 + x_2 v_2 + \cdots + x_m v_m = 0.$$

Como $\mathcal{B}$ é L.I., segue que $x_1 = x_2 = \cdots = x_m = 0$, como queríamos mostrar.

■

Corolário 2.51. *Seja V um espaço vetorial finitamente gerado sobre K. Todo conjunto linearmente independente de V pode ser completado para formar uma base de V.*

Demonstração. Se V é o espaço nulo, V não possui conjuntos linearmente independentes. Então podemos supor que V é não nulo. Como V é finitamente gerado, pelo Teorema 2.44, um subconjunto L.I. de V é finito e tem um número máximo de elementos. Seja $\mathcal{B} = \{v_1, v_2, \ldots, v_m\} \subset V$ linearmente independente. Se $[v_1, v_2, \ldots, v_m] = V$, então $\mathcal{B}$ é base. Caso contrário, existe $u_1 \in V$ com $u_1 \notin [v_1, v_2, \ldots, v_m]$. Pela Proposição 2.50, $\mathcal{B}_1 := \{v_1, v_2, \ldots, v_m, u_1\}$ é L.I. Se $\mathcal{B}_1$ gera V, então $\mathcal{B}_1$ é base. Caso contrário, podemos acrescentar $u_2 \in V$ tal que $\mathcal{B}_2 = \mathcal{B}_1 \cup \{u_2\}$ é L.I. Como V é finitamente gerado, esse processo terminará, pois para algum $l \geq 0$ o conjunto L.I. $\mathcal{B}_l = \{v_1, v_2, \ldots, v_m, u_1, u_2, \ldots, u_l\}$ terá que gerar V pois a inclusão de outro vetor tornará o conjunto L.D., pelo Teorema 2.44.

■

Teorema 2.52. *Todo espaço vetorial finitamente gerado tem base.*

Demonstração. Primeiro, o espaço nulo tem base, o conjunto vazio. Podemos, então, supor que V é não nulo. Tome então $v_1 \in V$, com $v_1 \neq 0$. Então $\{v_1\}$ é um subconjunto L.I. de V e, pelo colorário acima, pode ser completado até formar uma base de V.

■

No Exemplo 2.41, vimos que qualquer subconjunto L.I. com n vetores de K^n é uma base de K^n. Os resultados desenvolvidos acima nos permitem generalizar este fato para qualquer espaço vetorial de dimensão n.

Teorema 2.53. *Seja V um espaço vetorial sobre K. Se $\dim V = n$, então qualquer conjunto com n vetores L.I. em V forma uma base de V.*

Demonstração. Primeiro lembre que como $\dim V = n$, V possui uma base com n vetores. Logo V é gerado por um conjunto com n vetores e assim, pelo Teorema 2.44, todo conjunto com mais do que n vetores em V é L.D..

Seja, agora, $\mathcal{B}$ um conjunto L.I. em V com n vetores. Para mostrar que $\mathcal{B}$ é base falta garantir que $\mathcal{B}$ gera V. Suponha que $\mathcal{B}$ não gera V. Então existe um vetor em V que não é combinação linear dos vetores de $\mathcal{B}$. Pela Proposição 2.50, adicionando este vetor ao conjunto $\mathcal{B}$, obtemos um conjunto com $n+1$ vetores que é L.I.. Mas isso não pode acontecer, pois qualquer conjunto com mais do que n vetores em V é L.D. Logo $\mathcal{B}$ gera V.

∎

Sobre a dimensão de subespaços de espaços vetoriais de dimensão finita temos os seguintes resultados.

Proposição 2.54. *Seja V um espaço vetorial de dimensão finita e U um subespaço de V. Então*

i) $\dim U \leq \dim V$;

ii) $\dim U = \dim V$ *se, e somente se,* $U = V$;

iii) se U é subespaço próprio de V, ou seja, $U \leq V$ com $U \neq V$, então $\dim U < \dim V$;

iv) $\dim U = 0 \Leftrightarrow U = \{0\}$.

Demonstração. Seja $n = \dim V$. Então existem no máximo n vetores L.I. em V e assim, como todo vetor de U é um vetor de V, em U também temos no máximo n vetores *L.I.*. Sendo assim, $\dim U \leq n = \dim V$.

Se $\dim U = \dim V$, então existem n vetores L.I. em U. Como esses n vetores são também vetores de V e $\dim V = n$, pelo Teorema 2.53, eles formam uma base de V, logo geram todo V. Como U é subespaço, combinações lineares de vetores de U continuam em U, logo todo vetor de V pertence a U e portanto $U = V$.

Agora, se $U = \{0\}$, lembrando que por convenção U é então gerado pelo conjunto vazio $\emptyset$, temos que a base de U é o conjunto vazio, que tem zero vetores e assim $\dim U = 0$. Reciprocamente, se $\dim U = 0$, o conjunto vazio é sua base e portanto $U = \{0\}$.

∎

Teorema 2.55. *Seja V um espaço vetorial sobre K e U e W subespaços de V. Então*

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Demonstração. Seja $\mathcal{B} = \{v_1, v_2, \ldots, v_r\}$ uma base de $U \cap W$. Como $\mathcal{B}$ é L.I., pelo Corolário 2.51, podemos completar $\mathcal{B}$ até uma base de U

$$\mathcal{B}_1 = \{v_1, v_2, \ldots, v_r, u_1, u_2, \ldots, u_s\}$$

e até uma base de W

$$\mathcal{B}_2 = \{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_t\}.$$

Vamos mostrar que

$$\mathcal{B}' = \{v_1, v_2, \ldots, v_r, u_1, u_2, \ldots, u_s, w_1, w_2, \ldots, w_t\}$$

é uma base de $U + W$. De fato, dado $v = u + w \in U + W$, com $u \in U$ e $w \in W$, podemos escrever u como combinação linear dos vetores de $\mathcal{B}_1$ e w como combinação dos vetores de $\mathcal{B}_2$. Com isso, obtemos que v é uma combinação linear dos vetores de $\mathcal{B}'$ e, portanto, $\mathcal{B}'$ gera $U + W$. Além disso, $\mathcal{B}'$ é L.I., pois

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r + b_1u_1 + b_2u_2 + \cdots + b_su_s + c_1w_1 + c_2w_2 + \cdots + c_tw_t = 0,$$

$a_i, b_j, c_l \in K$,

$$\Leftrightarrow \sum_{i=1}^{r} a_iv_i + \sum_{j=1}^{s} b_ju_j + \sum_{l=1}^{t} c_lw_l = 0$$

$$\Leftrightarrow \sum_{j=1}^{s} b_ju_j = -\sum_{i=1}^{r} a_iv_i - \sum_{l=1}^{t} c_lw_l.$$

Note a soma do lado esquerdo $\sum_{j=1}^{s} b_ju_j$ pertence a U e a soma a do lado direito pertence a W. Então ambas pertencem a $U \cap W$e portanto podem ser escritas como combinação linear dos vetores de $\mathcal{B}$. Assim, existem escalares $k_1, k_2, \ldots, k_r \in K$ tais que

$$\sum_{j=1}^{s} b_ju_j = -\sum_{i=1}^{r} a_iv_i - \sum_{l=1}^{t} c_lw_l = k_1v_1 + k_2v_2 + \cdots + k_rv_r. \qquad (*)$$

Com isso, obtemos

$$-k_1v_1 - k_2v_2 - \cdots - k_rv_r + b_1u_1 + b_2u_2 + \cdots + b_su_s = 0,$$

uma combinação linear dos vetores de $\mathcal{B}_1$, logo $k_1 = k_2 = \cdots = k_r = 0$ e $b_1 = b_2 = \cdots = b_s = 0$. E voltando a (*) obtemos

$$-\sum_{i=1}^{r} a_iv_i - \sum_{l=1}^{t} c_lw_l = 0,$$

que é uma combinação linear dos vetores de $\mathcal{B}_2$, então $a_1 = a_2 = \cdots = a_r = 0$ e $c_1 = c_2 = \cdots = c_t = 0$. Logo $\mathcal{B}'$ é L.I. e portanto base de $U + W$. Com isso obtemos a igualdade das dimensões do enunciado do teorema.

■

Exemplo 2.56. **a)** *Considere os seguintes subespaços de* $\mathbb{R}^4$

$$U = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + y = 0 \quad e \quad z - t = 0\}$$

e

$$W = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y - z + t = 0\}.$$

Vamos determinar $\dim U$*,* $\dim W$*,* $\dim(U \cap W)$ *e* $\dim(U + W)$*.*

U é dado por

$$\begin{aligned} U &= \{(x,y,z,t) \in \mathbb{R}^4 \mid x+y=0 \quad e \quad z-t=0\} \\ &= \{(x,y,z,t) \in \mathbb{R}^4 \mid y=-x \quad e \quad t=z\} \\ &= \{(x,-x,z,z) \mid x,z \in \mathbb{R}\} \\ &= [(1,-1,0,0),(0,0,1,1)] \end{aligned}$$

Como os vetores geradores acima são L.I., obtemos $\dim U = 2$.

W é dado por

$$\begin{aligned} W &= \{(x,y,z,t) \in \mathbb{R}^4 \mid x-y-z+t=0\} \\ &= \{(x,y,z,t) \in \mathbb{R}^4 \mid t=-x+y+z\} \\ &= \{(x,y,z,-x+y+z) \mid x,y,z \in \mathbb{R}\} \\ &= [(1,0,0,-1),(0,1,0,1),(0,0,1,1)] \end{aligned}$$

Como os vetores geradores acima são L.I., obtemos $\dim W = 3$.

Para encontrar $\dim(U \cap W)$ *e* $\dim(U+W)$*, basta encontrarmos uma delas e a outra é obtida através do Teorema 2.55.*

A dimensão $\dim(U \cap W)$ *pode ser obtida diretamente pelo conjunto-solução do sistema linear homogêneo formado pelas equações de U e W*

$$\begin{cases} x+y \qquad\qquad = 0 \\ \qquad\qquad z-t = 0 \\ x-y-z+t = 0 \end{cases}$$

Como as equações são independentes, temos uma variável livre e portanto $\dim(U \cap W) = 1$. *Com isso, obtemos que*

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W) = 2+3-1 = 4.$$

b) *Considere os seguintes subespaços de* $\mathbb{R}^5$

$$U = [(1,-1,0,0,3),(1,1,-1,0,0),(-1,3,-1,0,-6)]$$

e

$$W = [(-1,-5,3,0,6),(1,-1,0,1,2)].$$

Agora, os espaços U e W são definidos a partir de seus geradores, ao invés de equações. Então para encontrar as dimensões $\dim U$ *e* $\dim W$ *precisamos verificar se os geradores dados são L.I. Sendo L.I, formam uma base. Sendo L.D., precisamos retirar do conjunto de geradores os vetores que são combinações dos outros. E para encontrarmos* $\dim(U \cap W)$ *e* $\dim(U+W)$*? Sabemos que $U+W$ vai ser gerado pela união dos geradores de U com os geradores de W. Mas de novo precisamos extrair uma base do conjunto de geradores. A seguir apresentamos uma ferramenta bastante útil para esta situação.*

Observação 2.57. *Sejam $v_1, v_2, \ldots, v_m \in \mathbb{K}^n$, $U = [v_1, v_2, \ldots, v_m]$ e considere a matriz $A \in \mathcal{M}_{m\times n}(K)$ cujas m linhas são formadas pelas coordenadas dos vetores $v_1, v_2, \ldots, v_m$ (os vetores-linha de A são $v_1, v_2, \ldots, v_m$). E seja R* **a matriz escalonada equivalente a** *A,*

$$A \sim R.$$

Então os vetores-linha $w_1, w_2, \ldots, w_m$ de R são combinações lineares dos vetores-linha de A, logo são combinações lineares de $v_1, v_2, \ldots, v_m$. Com isso, obtemos $[w_1, w_2, \ldots, w_m] \leq [v_1, v_2, \ldots, v_m]$, ou seja, o espaço gerado pelos vetores-linha de R é subespaço do espaço gerado pelos vetores-linha de A. Reciprocamente, pertindo de R podemos fazer as operações elementares inversas até obtermos A. Logo, obtemos também que $[v_1, v_2, \ldots, v_m] \leq [w_1, w_2, \ldots, w_m]$. Assim

$$[v_1, v_2, \ldots, v_m] = [w_1, w_2, \ldots, w_m],$$

os vetores-linha de A e os vetores-linha de R geram o mesmo subespaço de K^n.

Agora, algumas linhas de R pode mser nulas e portanto alguns dos vetores $w_1, w_2, \ldots, w_m$ podem ser nulos. Sejam $\{w_1, w_2, \ldots, w_k\}$, com $k \leq m$, todos os vetores não nulos. Das propriedades do escalonamento, os vetores-linha não nulos de R têm uma posição igual a 1 e todos os outros têm valor 0 nesta posição. Disso segue que $\{w_1, w_2, \ldots, w_k\}$ é L.I. e será portanto base de $U = [v_1, v_2, \ldots, v_m]$.

Exemplo 2.58. *Voltando no item b) do exemplo anterior, temos os subespaços*

$$U = [(1, -1, 0, 0, 3), (1, 1, -1, 0, 0), (-1, 3, -1, 0, -6)]$$

e

$$W = [(-1, -5, 3, 0, 6), (1, -1, 0, 1, 2)]$$

de $\mathbb{R}^5$ e queremos determinar $\dim U$, $\dim W$, $\dim(U \cap W)$ e $\dim(U + W)$.

Em U, temos

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 3 \\ 1 & 1 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 & 0 & 3 \\ 0 & 2 & -1 & 0 & -3 \\ 0 & 2 & -1 & 0 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 & 0 & 3 \\ 0 & 2 & -1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Então $U = [(1, -1, 0, 0, 3), (0, 2, -1, 0, -3)]$ e $\dim U = 2$.

Para W, temos $\dim W = 2$, pois os 2 vetores geradores não são múltiplos um do outro.

Para $U + W$, temos que $U + W$ é gerado pela união dos geradores de U e de W. Então $U + W = [(1, -1, 0, 0, 3), (0, 2, -1, 0, -3), (-1, -5, 3, 0, 6), (1, -1, 0, 1, 2)]$. E temos que

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 3 \\ 0 & 2 & -1 & 0 & -3 \\ -1 & -5 & 3 & 0 & 6 \\ 1 & -1 & 0 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 & 0 & 3 \\ 0 & 2 & -1 & 0 & -3 \\ 0 & -6 & 3 & 0 & 9 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 & 0 & 3 \\ 0 & 2 & -1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

Então $U + W = [(1, -1, 0, 0, 3), (0, 2, -1, 0, -3), (0, 0, 0, 1, -1)]$ e tais vetores são L.I., logo $\dim(U + W) = 3$. Com isso, obtemos também que

$$\dim(U \cap W) = \dim U + \dim W - \dim(U + W) = 2 + 2 - 3 = 1.$$

2.5 Soma Direta

Definição 2.59. *Seja V um espaço vetorial e U e W subespaços de V.*

i) *Dizemos que a soma $U+W$ é direta quando $U \cap W = \{0\}$.*
Neste caso, escrevemos $U \oplus W$.

ii) *Quando $V = U \oplus W$, dizemos que V é a **soma direta** de U e V.*

Exemplo 2.60. *Em $V = \mathbb{R}^2$, considere*

$$U_1 = \{(x,y) \in \mathbb{R}^2 \mid y = 0\} = [(1,0)] \quad e \quad U_2 = \{(x,y) \in \mathbb{R}^2 \mid x = 0\} = [(0,1)].$$

Temos $U_1 \cap U_2 = \{(0,0)\}$ e $U_1 + U_2 = [(1,0),(0,1)] = \mathbb{R}^2$. Então

$$\mathbb{R}^2 = U_1 \oplus U_2.$$

Note que considerando

$$W = \{(x,y) \mid x = y\} = [(1,1)]$$

Temos $W \cap U_i = \{(0,0)\}$ e $W + U_i = \mathbb{R}^2$, $i = 1,2$. Então também

$$\mathbb{R}^2 = U_1 \oplus W \quad e \quad \mathbb{R}^2 = U_2 \oplus W.$$

Exemplo 2.61. *Seja $V = \mathcal{M}_2(K)$, onde K é um subcorpo de $\mathbb{C}$. Considere*

$$U = \{A \in \mathcal{M}_2(K) \mid A = A^t\}, \quad \textit{o subespaço das matrizes simétricas,}$$

e

$$W = \{A \in \mathcal{M}_2(K) \mid A = -A^t\}, \quad \textit{o subespaço das matrizes antissimétricas.}$$

Temos que

$$A \in U \cap W \Leftrightarrow A = A^t \ \ e \ \ A = -A^t \Leftrightarrow A = -A \Leftrightarrow A = 0.$$

Então $U \cap W = \{0\}$. Além disso,

$$A = A^t \Leftrightarrow A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \ a,b,c \in K.$$

Então $\dim U = 3$. E

$$A = -A^t \Leftrightarrow A = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}, \ b \in K.$$

Então $\dim W = 1$. Com isso, obtemos

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W) = 3 + 1 - 0 = 4,$$

Logo $U + W = \mathcal{M}_2(K)$. Mais ainda, $\mathcal{M}_2(K) = U \oplus W$.

Sobre somas diretas, vamos destacar os seguintes resultados que serão bastante utilizados na teoria de Operadores Lineares.

Proposição 2.62. *Seja V um espaço vetorial sobre K e W_1 e W_2 subespaços de V. Então $V = W_1 \oplus W_2$ se, e somente se, cada $v \in V$ pode ser escrito de maneira única $v = w_1 + w_2$, com $w_1 \in W_1$ e $w_2 \in W_2$.*

Demonstração. Se $V = W_1 \oplus W_2$, então cada $v \in V$ pode ser escrito como $v = w_1 + w_2$, com $w_1 \in W_1$ e $w_2 \in W_2$. Suponha que $v = w_1' + w_2'$, com $w_1' \in W_1$ e $w_2' \in W_2$. Então

$$w_1 + w_2 = w_1' + w_2' \quad \Rightarrow \quad w_1 - w_1' = w_2' - w_2 \ \in W_1 \cap W_2.$$

Como $W_1 \cap W_2 = \{0\}$, segue que $w_1 - w_1' = 0$ e $w_2' - w_2 = 0$, ou seja, $w_1 = w_1'$ e $w_2 = w_2'$.

Reciprocamente, suponha que para cada $v \in V$ existem únicos $w_1 \in W_1$ e $w_2 \in W_2$ tais que $v = w_1 + w_2$. Então $V = W_1 + W_2$. Além disso, se $w \in W_1 \cap W_2$, então w pode ser escrito como

$$w = w + 0 \in W_1 + W_2 \quad \text{e} \quad w = 0 + w \in W_1 + W_2.$$

Da hipótese de unicidade da expressão, segue que $w = 0$. Portanto $W_1 \cap W_2 = \{0\}$ e, assim, $V = W_1 \oplus W_2$.

■

Proposição 2.63. *Seja V um espaço vetorial de dimensão finita e W_1 um subespaço de V. Então existe $W_2 \leq V$ tal que $V = W_1 \oplus W_2$, que é chamado de **complemento de** W_1 **em** V.*

Demonstração. Seja $\mathcal{B} = \{u_1, u_2, \ldots, u_k\}$ uma base de W_1. Se $W_1 = V$, então $W_2 = \{0\}$. Caso contrário, podemos completar $\mathcal{B}$ até uma base de V, $\mathcal{B}' = \{u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_l\}$. Tome $W_2 = [v_1, v_2, \ldots, v_l]$. Então $V = W_1 + W_2$, pois na base $\mathcal{B}'$ cada $v \in V$ pode ser escrito como $v = \sum_{i=1}^{k} a_i u_i + \sum_{j=1}^{l} b_j v_j$, com $a_i, b_j \in K$, e temos $\sum_{i=1}^{k} a_i u_i \in W_1$ e $\sum_{j=1}^{l} b_j v_j \in W_2$. Além disso, $W_1 \cap W_2 = \{0\}$, pois

$$v \in W_1 \cap W_2 \ \Rightarrow \ v = \sum_{i=1}^{k} a_i u_i \ \text{ e } \ v = \sum_{j=1}^{l} b_j v_j$$

$$\Rightarrow \ \sum_{i=1}^{k} a_i u_i - \sum_{j=1}^{l} b_j v_j = 0.$$

Como $\mathcal{B}'$ é L.I., seque que $a_i = b_j = 0, \forall i, j$. Assim $v = 0$.

■

Observação 2.64. *O complemento de um subespaço vetorial nem sempre é único, veja o Exemplo 2.60.*

A soma direta de dois subespaços pode ser generalizada para a soma direta de vários subespaços da seguinte maneira. Seja V um espaço vetorial sobre K e $W_1, \ W_2, \ldots, \ W_t$ subespaços de V. Defina

$$W_1 + W_2 + \cdots + W_t = \{w_1 + w_2 + \cdots + w_t \mid w_i \in W_i, \ i = 1, 2, \ldots, t\}.$$

Quando

$$W_i \cap (W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_t) = \{0\}, \quad \text{para todo} \ \ i = 1, 2, \ldots, t,$$

então a soma $W_1 + W_2 + \cdots + W_t$ é chamada **soma direta** de $W_1,\ W_2,\ \ldots,\ W_t$ e denotada por

$$W_1 \oplus W_2 \oplus \cdots \oplus W_t.$$

Se $V = W_1 \oplus W_2 \oplus \cdots \oplus W_t$, dizemos que V **é a soma direta dos subespaços** $W_1,\ W_2, \ldots,\ W_t$.

2.6 Coordenadas de um vetor e Matriz de Mudança de Base

Considere o espaço vetorial $\mathbb{R}^n$ e sua base canônica $\mathcal{C} = \{e_1, e_2, \ldots, e_n\}$. Então, para cada $w = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, temos

$$w = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

Assim, as coordenadas do vetor $w \in \mathbb{R}^n$ coincidem com os escalares que aparecem na expressão de w como combinação linear dos vetores da base $\mathcal{C}$. Motivados por isso, dada uma base de um espaço vetorial, iremos chamar os escalares da expressão de um vetor nessa base de coordenadas desse vetor na base dada. Antes da definição precisa, vamos fazer algumas considerações.

Seja V um espaço vetorial sobre um corpo K e $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ uma base de V. Então cada vetor $w \in V$ pode ser escrito como combinação linear dos vetores da base $\mathcal{B}$, ou seja, existem escalares $k_1, k_2, \ldots, k_n \in K$ tais que

$$w = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n.$$

Além disso, esta expressão para w é única, ou seja, os escalares $k_1, k_2, \ldots, k_n \in K$ são únicos tais que $w = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n$. De fato, se

$$w = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n \quad \text{e} \quad w = l_1 v_1 + l_2 v_2 + \cdots + l_n v_n,$$

então

$$\begin{aligned} & k_1 v_1 + k_2 v_2 + \cdots + k_n v_n = l_1 v_1 + l_2 v_2 + \cdots + l_n v_n \\ \Rightarrow\ & (k_1 - l_1) v_1 + (k_2 - l_2) v_2 + \cdots + (k_n - l_n) v_n = 0. \end{aligned}$$

Como os vetores são L.I., segue que

$$k_1 - l_1 = 0,\ k_2 - l_2 = 0,\ \ldots,\ k_n - l_n = 0,$$

ou seja,

$$k_1 = l_1,\ k_2 = l_2,\ \ldots,\ k_n = l_n,$$

o que prova a afirmação. Isso nos diz que fixada uma base do espaço vetorial, os escalares de um vetor nesta base são unicamente determinados. Com isso, temos a seguinte definição.

Definição 2.65. *Seja $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ uma base de um espaço vetorial V sobre K e $w \in V$ tal que $w = k_1v_1 + k_2v_2 + \cdots + k_nv_n$, com $k_1, k_2, \ldots, k_n \in K$. Os escalares $k_1, k_2, \ldots, k_n$ são chamados de* ***coordenadas de w na base $\mathcal{B}$****, que denotamos por*

$$[w]_{\mathcal{B}} = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}.$$

Note que na definição acima estamos fixando a ordem dos elementos da base $\mathcal{B}$, que chamamos de **base ordenada**. Por simplicidade, iremos nos referir às bases ordenadas como bases, deixando subentendida a ordenação.

Exemplo 2.66. *Seja $V = \mathbb{R}^2$ e $w = (2, 1)$. As coordenadas de w na base canônica $\mathcal{C} = \{(1, 0), (0, 1)\}$ são*

$$[(2,1)]_{\mathcal{C}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \textit{pois} \quad w = (2,1) = 2(1,0) + 1(0,1).$$

Agora considere a base $\mathcal{B} = \{(2, -1), (3, 4)\}$. As coordenadas de w na base $\mathcal{B}$ são:

$$[(2,1)]_{\mathcal{B}} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \textit{onde} \quad w = (2,1) = a(2,-1) + b(3,4).$$

Resolvendo o sistema,

$$\begin{cases} 2a + 3b = 2 \\ -a + 4b = 1 \end{cases}$$

obtemos $a = \frac{5}{11}$ e $b = \frac{4}{11}$. Assim

$$[(2,1)]_{\mathcal{B}} = \begin{pmatrix} \frac{5}{11} \\ \frac{4}{11} \end{pmatrix}.$$

No exemplo acima, como as bases $\mathcal{C}$ e $\mathcal{B}$ de $\mathbb{R}^2$ são distintas, obtemos coordenadas diferentes para um mesmo vetor w. Mas existe uma relação entre tais coordenadas, como veremos a seguir.

Seja V um espaço vetorial sobre K e sejam

$$\mathcal{B} = \{v_1, v_2, \ldots, v_n\} \quad \text{e} \quad \mathcal{B}' = \{u_1, u_2, \ldots, u_n\}$$

bases de V. Então, dado $w \in V$, existem escalares $k_1, k_2, \ldots, k_n \in K$ e $l_1, l_2, \ldots, l_n \in K$ (únicos) tais que

$$w = k_1v_1 + k_2v_2 + \cdots + k_nv_n \quad \text{e} \quad w = l_1u_1 + l_2u_2 + \cdots + l_nu_n.$$

Assim

$$[w]_{\mathcal{B}} = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} \quad \text{e} \quad [w]_{\mathcal{B}'} = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}.$$

Podemos relacionar as coordenadas de w na base $\mathcal{B}$ com suas coordenadas na base $\mathcal{B}'$ da seguinte maneira:

Como $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ é uma base de V e cada vetor u_i da base $\mathcal{B}'$ é um vetor de V, podemos escrevê-los como combinação linear dos vetores da base $\mathcal{B}$. Assim,

$$\begin{aligned} u_1 &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n \\ u_2 &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n \\ &\vdots \\ u_n &= a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{nn}v_n \end{aligned} \tag{1}$$

para certos $a_{ij} \in K, i, j = 1, 2, \ldots, n$. Como

$$w = l_1u_1 + l_2u_2 + \cdots + l_nu_n,$$

substituindo as expressões acima, obtemos

$$\begin{gathered} w = l_1(a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n) + l_2(a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n) + \\ + \cdots + l_n(a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{nn}v_n) \end{gathered}$$

$$\begin{gathered} \Rightarrow w = (l_1a_{11} + l_2a_{12} + \cdots + l_na_{1n})v_1 + (l_1a_{21} + l_2a_{22} + \cdots + l_na_{2n})v_2 + \\ + \cdots + (l_1a_{n1} + l_2a_{n2} + \cdots + l_na_{nn})v_n. \end{gathered} \tag{2}$$

Note que na expressão (2) acima temos w escrito como combinação linear dos vetores v_1, $v_2, \ldots,$ v_n da base $\mathcal{B}$. Como a expressão de um vetor numa base é única e temos também $w = k_1v_1 + k_2v_2 + \cdots + k_nv_n$, segue que

$$\begin{cases} k_1 = l_1a_{11} + l_2a_{12} + \cdots + l_na_{1n} \\ k_2 = l_1a_{21} + l_2a_{22} + \cdots + l_na_{2n} \\ \vdots \\ k_n = l_1a_{n1} + l_2a_{n2} + \cdots + l_na_{nn} \end{cases}$$

que na forma matricial é escrita como

$$\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}.$$

Assim,

$$[w]_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot [w]_{\mathcal{B}'},$$

de onde vemos que as coordenadas $[w]_{\mathcal{B}}$ e $[w]_{\mathcal{B}'}$ de w nas bases $\mathcal{B}$ e $\mathcal{B}'$, respectivamente, estão relacionadas.

Note que: a matriz acima que relaciona $[w]_{\mathcal{B}}$ e $[w]_{\mathcal{B}'}$ tem suas entradas dadas pelos escalares das expressões de (1). Mais precisamente, na coluna 1 da matriz estão as coordenadas de u_1 na base $\mathcal{B}$, na coluna 2 estão as coordenadas de u_2 na base $\mathcal{B}$, ... , na coluna n estão as coordenadas de u_n na base $\mathcal{B}$. Denotamos esta matriz por $[\ I\ ]_{\mathcal{B}}^{\mathcal{B}'}$. Com isso,

$$[\ I\ ]_{\mathcal{B}}^{\mathcal{B}'} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} | & | & & | \\ [u_1]_{\mathcal{B}} & [u_2]_{\mathcal{B}} & \cdots & [u_n]_{\mathcal{B}} \\ | & | & & | \end{pmatrix}$$

e chamamos a matriz $[\ I\ ]_{\mathcal{B}}^{\mathcal{B}'}$ de **matriz de mudança de base**, da base $\mathcal{B}'$ para a base $\mathcal{B}$. Assim

$$\boxed{[w]_{\mathcal{B}} = [\ I\ ]_{\mathcal{B}}^{\mathcal{B}'} \cdot [w]_{\mathcal{B}'}}$$

onde a matriz $[\ I\ ]_{\mathcal{B}}^{\mathcal{B}'}$ é obtida os escrevermos cada vetor da base $\mathcal{B}'$ na base $\mathcal{B}$.

Perceba que olhando a relação em destaque acima da direita para a esquerda, ela diz que pegamos as coordenadas de w na base $\mathcal{B}'$, aplicamos a matriz que muda da base $\mathcal{B}'$ para a base $\mathcal{B}$ e obtemos as coordenadas de w na base $\mathcal{B}$.

Exemplo 2.67. *Voltando no Exemplo 2.66, foram consideradas as bases $\mathcal{C} = \{(1,0),(0,1)\}$ e $\mathcal{B} = \{(2,-1),(3,4)\}$ de $\mathbb{R}^2$ e, para $w = (2,1)$, obtemos*

$$[(2,1)]_{\mathcal{C}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad e \quad [(2,1)]_{\mathcal{B}} = \begin{pmatrix} \frac{5}{11} \\ \frac{4}{11} \end{pmatrix}.$$

Agora, matriz de mudança de base de $\mathcal{C}$ para $\mathcal{B}$ é dada por

$$[\ I\ ]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} | & | \\ [(1,0)]_{\mathcal{B}} & [(0,1)]_{\mathcal{B}} \\ | & | \end{pmatrix}.$$

Vamos então encontrar $[(1,0)]_{\mathcal{B}}$ e $[(0,1)]_{\mathcal{B}}$. Temos

$$(1,0) = a(2,-1) + b(3,4) \iff \begin{cases} 2a + 3b = 1 \\ -a + 4b = 0 \end{cases} \iff a = \frac{4}{11} \quad e \quad b = \frac{1}{11},$$

$$\Rightarrow (1,0) = \frac{4}{11}(2,-1) + \frac{1}{11}(3,4) \Rightarrow [(1,0)]_{\mathcal{B}} = \begin{pmatrix} \frac{4}{11} \\ \frac{1}{11} \end{pmatrix};$$

e

$$(0,1) = c(2,-1) + d(3,4) \iff \begin{cases} 2c + 3d = 0 \\ -c + 4d = 1 \end{cases} \iff c = -\frac{3}{11} \quad e \quad d = \frac{2}{11}$$

$$\Rightarrow (0,1) = -\frac{3}{11}(2,-1) + \frac{2}{11}(3,4) \Rightarrow [(0,1)]_{\mathcal{B}} = \begin{pmatrix} -\frac{3}{11} \\ \frac{2}{11} \end{pmatrix}.$$

Assim,

$$[\ I\ ]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} | & | \\ [(1,0)]_{\mathcal{B}} & [(0,1)]_{\mathcal{B}} \\ | & | \end{pmatrix} = \begin{pmatrix} \frac{4}{11} & -\frac{3}{11} \\ \frac{1}{11} & \frac{2}{11} \end{pmatrix}.$$

Com isso, para $w = (2,1)$, temos relação

$$[(2,1)]_{\mathcal{B}} = [\ I\ ]_{\mathcal{B}}^{\mathcal{C}} \cdot [(2,1)]_{\mathcal{C}},$$

ou seja,

$$\begin{pmatrix} \frac{5}{11} \\ \frac{4}{11} \end{pmatrix} = \begin{pmatrix} \frac{4}{11} & -\frac{3}{11} \\ \frac{1}{11} & \frac{2}{11} \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

O exemplo acima foi apenas para ilustrar como se calcula a matriz de mudança de base e a equação de mudança de base. A seguir, alguns exemplos mais interessantes.

Exemplo 2.68. *Considere o problema de, dado um vetor $w \in \mathbb{R}^2$, encontrar o vetor w' obtido aplicarmos uma rotação de ângulo θ no sentido anti-horário no vetor w.*

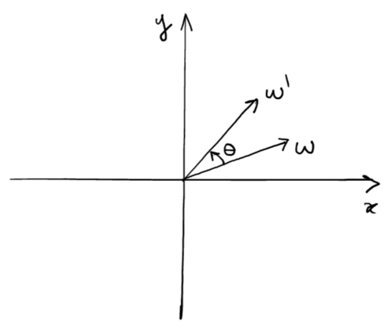
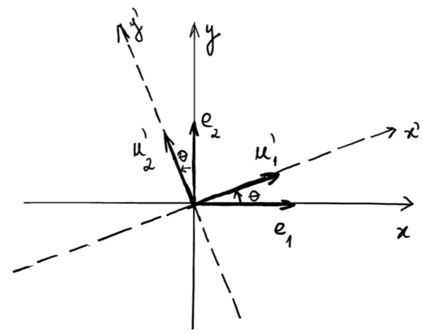


Figura 1: Rotação anti-horária de ângulo θ em $\mathbb{R}^2$

O problema pode ser resolvido usando a matriz de mudança de base da seguinte maneira.

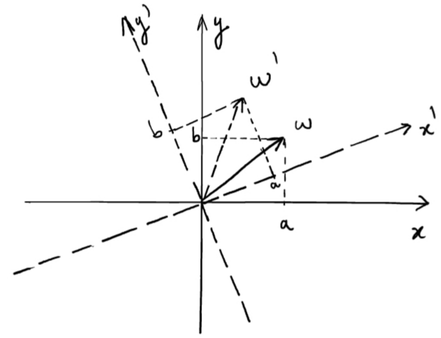
Considere o sistema de coordenadas xy. Aplicando uma rotação anti-horária de ângulo θ no sistema de coordenadas xy, obtemos outro sistema de coordenadas $x'y'$. Com isso, os vetores da base canônica $\mathcal{C} = \{e_1, e_2\}$ são levados em vetores ortogonais e unitários u'_1 e u'_2, respectivamente, sobre os eixos $x'y'$ formando também uma base $\mathcal{B}' = \{u'_1, u'_2\}$ de $\mathbb{R}^2$.



Agora, escrevendo o vetor $w = (a,b)$ e seu vetor obtido pela rotação $w' = (a', b')$, queremos saber quem são as coordenadas a' e b' de w', ou seja, queremos saber quais são as coordenadas de w' na base canônica, $[w']_{\mathcal{C}}$.

Note que, como estamos rotacionando todo o sistema, as coordenadas de w na base $\mathcal{C}$ são levadas nas coordenadas de w' na base $\mathcal{B}'$, ou seja, $w' = au'_1 + bu'_2$. Assim

$$[w']_{\mathcal{B}'} = \begin{pmatrix} a \\ b \end{pmatrix}.$$



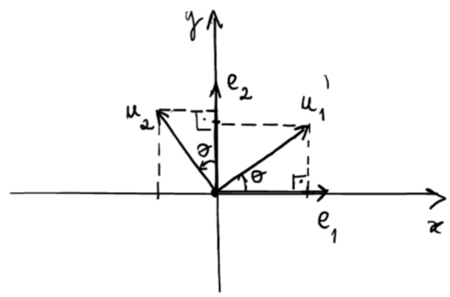
Pela teoria que vimos, as coordenadas $[w']_{\mathcal{C}}$ e $[w']_{\mathcal{B}'}$ estão relacionadas por

$$[w']_{\mathcal{C}} = [\ I\ ]_{\mathcal{C}}^{\mathcal{B}'}[w']_{\mathcal{B}'},$$

onde

$$[\ I\ ]_{\mathcal{C}}^{\mathcal{B}'} = \begin{pmatrix} | & | \\ [u'_1]_{\mathcal{C}} & [u'_2]_{\mathcal{C}} \\ | & | \end{pmatrix}.$$

Vamos então encontrar as coordenadas $[u'_1]_{\mathcal{C}}$ e $[u'_2]_{\mathcal{C}}$. Temos



$$u'_1 = (\cos\theta, \operatorname{sen}\theta)$$

$$u'_2 = (-\operatorname{sen}\theta, \cos\theta)$$

Então

$$[u'_1]_{\mathcal{C}} = \begin{pmatrix} \cos\theta \\ \operatorname{sen}\theta \end{pmatrix} \quad e \quad [u'_2]_{\mathcal{C}} = \begin{pmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{pmatrix},$$

logo

$$[\ I\ ]_{\mathcal{C}}^{\mathcal{B}'} = \begin{pmatrix} | & | \\ [u'_1]_{\mathcal{C}} & [u'_2]_{\mathcal{C}} \\ | & | \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

e portanto a relação $[w']_{\mathcal{C}} = [\ I\ ]_{\mathcal{C}}^{\mathcal{B}'}[w']_{\mathcal{B}'}$ é

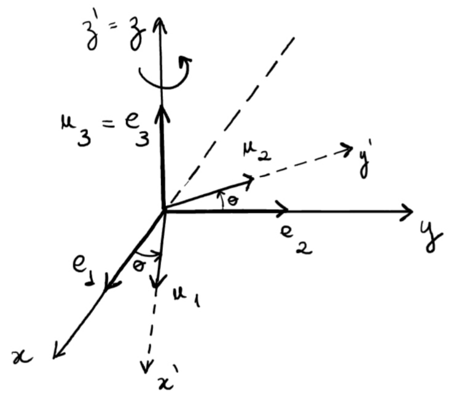
$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix}.$$

Assim, dado um vetor $w = (a, b) \in \mathbb{R}^2$, o vetor w' obtido aplicarmos uma rotação de ângulo θ no sentido anti-horário no vetor w é

$$w' = (a \cdot \cos\theta - b \cdot \operatorname{sen}\theta, a \cdot \operatorname{sen}\theta + b \cdot \cos\theta).$$

Exemplo 2.69 (Rotações em torno dos eixos em $\mathbb{R}^3$)**.** *De maneira análoga à que foi feita no Exemplo 2.68, dado $w \in \mathbb{R}^3$, vamos encontrar a rotação anti-horária de ângulo θ em torno do eixo-z do vetor w.*

Seja $\mathcal{C} = \{e_1, e_2, e_3\}$ a base canônica de $\mathbb{R}^3$. Aplicando a rotação nos vetores da base canônica, obtemos vetores u_1, u_2, u_3, respectivamente, que formam uma nova base $\mathcal{B}' = \{u_1', u_2', u_3'\}$ de $\mathbb{R}^3$. Tais vetores são dados por:



$$[u_1']_{\mathcal{C}} = \begin{pmatrix} \cos\theta \\ \text{sen}\theta \\ 0 \end{pmatrix}$$

$$[u_2']_{\mathcal{C}} = \begin{pmatrix} -\text{sen}\theta \\ \cos\theta \\ 0 \end{pmatrix}$$

$$[u_3']_{\mathcal{C}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Assim a matriz de mudança de base da base $\mathcal{B}'$ para a base $\mathcal{C}$ é dada por

$$[\ I\ ]_{\mathcal{C}}^{\mathcal{B}'} = \begin{pmatrix} | & | & | \\ [u_1']_{\mathcal{C}} & [u_2']_{\mathcal{C}} & [u_3']_{\mathcal{C}} \\ | & | & | \end{pmatrix} = \begin{pmatrix} \cos\theta & -\text{sen}\theta & 0 \\ \text{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Assim, o vetor $w' = (a', b', c')$ obtido ao aplicarmos a rotação no vetor $w = (a, b, c)$, é dado por

$$\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = [w']_{\mathcal{C}} = [\ I\ ]_{\mathcal{C}}^{\mathcal{B}'} [w']_{\mathcal{B}'} = \begin{pmatrix} \cos\theta & -\text{sen}\theta & 0 \\ \text{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

*Deixamos como **exercício** encontrar as matrizes de rotação anti-horária de ângulo θ em torno do eixo-x e em torno do eixo-y.*

Vamos agora fazer algumas observações teóricas importantes sobre as matrizes de mudança de base.

Observação 2.70. *Sejam $\mathcal{B}$ e $\mathcal{B}'$ bases de um espaço vetorial V de dimensão $n > 0$.*

i) Se $\mathcal{B} = \mathcal{B}'$, então a matriz de mudança de base é a matriz identidade. De fato, sejam $\mathcal{B} = \mathcal{B}' = \{v_1, v_2, \ldots, v_n\}$. Por definição, para determinar a matriz de mudança de base $[\ I\ ]_{\mathcal{B}}^{\mathcal{B}'}$, escrevemos cada vetor da base $\mathcal{B}'$ como combinação dos vetores da base $\mathcal{B}$. Fazendo isso, obtemos

$$\begin{aligned} v_1 &= 1.v_1 + 0.v_2 + \cdots + 0.v_n \\ v_2 &= 0.v_1 + 1.v_2 + \cdots + 0.v_n \\ &\vdots \\ v_n &= 0.v_1 + 0.v_2 + \cdots + 1.v_n \end{aligned}$$

Assim

$$[\ I\ ]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} | & | & & | \\ [v]_{\mathcal{B}} & [v_2]_{\mathcal{B}} & \cdots & [v_n]_{\mathcal{B}} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I_n.$$

Logo

$$\boxed{[\ I\ ]^{\mathcal{B}}_{\mathcal{B}} = I_n}$$

ii) Considere mais uma base $\mathcal{B}''$ de V, ou seja, considere três bases $\mathcal{B}$, $\mathcal{B}'$ e $\mathcal{B}''$ de V.

$$\begin{tikzcd} \mathcal{B} & \mathcal{B}' \arrow[l, "{[I]^{\mathcal{B}'}_{\mathcal{B}}}"'] & \mathcal{B}'' \arrow[l, "{[I]^{\mathcal{B}''}_{\mathcal{B}'}}"'] \arrow[ll, bend left, "{[I]^{\mathcal{B}''}_{\mathcal{B}}}"] \end{tikzcd}$$

Então, dado $w \in V$, temos as coordenadas de w nestas três bases:

$$[w]_{\mathcal{B}}, \quad [w]_{\mathcal{B}'} \quad e \quad [w]_{\mathcal{B}''}.$$

Para mudar da base $\mathcal{B}''$ para a base $\mathcal{B}'$, temos a relação

$$[w]_{\mathcal{B}'} = [\ I\ ]^{\mathcal{B}''}_{\mathcal{B}'}[w]_{\mathcal{B}''} \tag{3}$$

e para mudar da base $\mathcal{B}'$ para a base $\mathcal{B}$, temos a relação

$$[w]_{\mathcal{B}} = [\ I\ ]^{\mathcal{B}'}_{\mathcal{B}}[w]_{\mathcal{B}'}. \tag{4}$$

Substituindo (3) em (4), obtemos

$$[w]_{\mathcal{B}} = [\ I\ ]^{\mathcal{B}'}_{\mathcal{B}}[w]_{\mathcal{B}'} = [\ I\ ]^{\mathcal{B}'}_{\mathcal{B}}[\ I\ ]^{\mathcal{B}''}_{\mathcal{B}'}[w]_{\mathcal{B}''}. \tag{5}$$

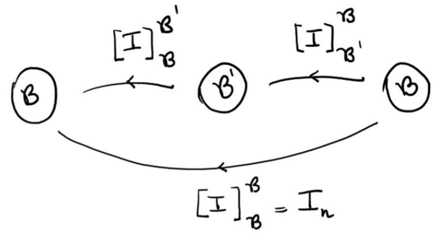
Note que a relação acima (leia da direita para a esquerda) está mudando as coordenadas de w na base $\mathcal{B}''$ para as coordenadas de w na base $\mathcal{B}$. Sabemos que a matriz $[\ I\ ]^{\mathcal{B}''}_{\mathcal{B}}$ é a matriz que faz essa mudança de base, ou seja,

$$[w]_{\mathcal{B}} = [\ I\ ]^{\mathcal{B}''}_{\mathcal{B}}[w]_{\mathcal{B}''}. \tag{6}$$

Assim, das relações (5) e (6), obtemos

$$\boxed{[\ I\ ]^{\mathcal{B}''}_{\mathcal{B}} = [\ I\ ]^{\mathcal{B}'}_{\mathcal{B}} \cdot [\ I\ ]^{\mathcal{B}''}_{\mathcal{B}'}}$$

iii) Das observações i) e ii), como um caso particular, segue uma outra propriedade importante. Considerando as bases $\mathcal{B}$ e $\mathcal{B}'$ de V, por ii) e i) obtemos



$$[\ I\ ]_{\mathcal{B}}^{\mathcal{B}'} \cdot [\ I\ ]_{\mathcal{B}'}^{\mathcal{B}} = [\ I\ ]_{\mathcal{B}}^{\mathcal{B}} = I_n$$

Analogamente,

$$[\ I\ ]_{\mathcal{B}'}^{\mathcal{B}} \cdot [\ I\ ]_{\mathcal{B}}^{\mathcal{B}'} = [\ I\ ]_{\mathcal{B}'}^{\mathcal{B}'} = I_n$$

*Então vemos que as matrizes $[\ I\ ]_{\mathcal{B}'}^{\mathcal{B}}$ e $[\ I\ ]_{\mathcal{B}}^{\mathcal{B}'}$ multiplicadas à direita e a esquerda dá a matriz identidade. Isso nos diz que as matrizes são inversíveis e que são uma a inversa da outra. Assim, **toda matriz de mudança de base é inversível** e*

$$\boxed{([\ I\ ]_{\mathcal{B}'}^{\mathcal{B}})^{-1} = [\ I\ ]_{\mathcal{B}}^{\mathcal{B}'}}.$$

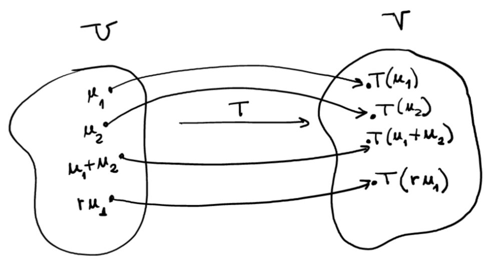
3 Transformações Lineares

Nesta seção vamos considerar funções entre espaços vetoriais, mas não quaisquer funções e sim funções que sejam compatíveis com a estrutura algébrica dos espaços.

Para entender a definição a seguir, considere uma função

$$T : U \to V$$

entre dois espaços vetoriais U e V (sobre $\mathbb{R}$).



Então, dados $u_1, u_2 \in U$, temos suas imagens $T(u_1)$ e $T(u_2)$ em V já que, como T é uma função, todo elemento de U tem uma imagem em V. Agora, como U é um espaço vetorial, temos o vetor soma $u_1 + u_2 \in U$ e assim esse vetor também tem sua imagem $T(u_1 + u_2)$ em V. Por outro lado, os vetores $T(u_1)$ e $T(u_2)$ pertencem a V, que também é um espaço vetorial, logo temos o vetor soma $T(u_1) + T(u_2) \in V$. Para uma função qualquer, temos os dois elementos $T(u_1 + u_2)$ e $T(u_1) + T(u_2)$ em V. Da mesma maneira em relação à multiplicação por escalar, ou seja, para $r \in K$, temos os vetores $T(ru_1)$, que é a imagem do vetor ru_1 pela função T, e temos $rT(u_1)$, que é a multiplicação do vetor $T(u_1)$ pelo escalar r. A compatibilidade com a estrutura que mencionamos pede que aconteça a igualdade entre os vetores $T(u_1 + u_2)$ e $T(u_1) + T(u_2)$ e a igualdade entre os vetores $T(ru_1)$ e $rT(u_1)$.

Definição 3.1. *Sejam U e V espaços vetoriais sobre um corpo K. Uma **transformação linear de** U **em** V é uma aplicação $T : U \to V$ tal que, para quaisquer $u_1, u_2 \in U$ e $k \in K$,*

i) $T(u_1 + u_2) = T(u_1) + T(u_2)$;

ii) $T(ku_1) = kT(u_1)$;

ou, equivamentemente, (juntando os dois itens em um)

$$T(ku_1 + u_2) = kT(u_1) + T(u_2).$$

Notação: A imagem de u por T pode ser representada por

$$T(u) \text{ ou, simplesmente, por } Tu \text{ (quando não há ambiguidade).}$$

Exemplo 3.2. *Sejam U e V espaços vetoriais sobre K.*

A função nula $T : U \to V$, onde $Tu = 0$, $\forall u \in U$, é transformação linear:

$$T(ku_1 + u_2) = 0 \quad e \quad kTu_1 + Tu_2 = k0 + 0 = 0, \quad \forall u_1, u_2 \in U \quad e \quad k \in K;$$

A função identidade $I : U \to U$, onde $I(u) = u$, $\forall u \in U$, é transformação linear:

$$I(ku_1 + u_2) = ku_1 + u_2 = kI(u_1) + I(u_2), \quad \forall u_1, u_2 \in U \quad e \quad k \in K.$$

Exemplo 3.3. *Seja $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$, dada pela derivada, ou seja, dado $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$, a função D é definida por*

$$D(p(t)) = p'(t) := a_1 + 2a_2 t + \cdots + n a_n t^{n-1}.$$

Temos que D é uma transformação linear. De fato, sejam $p(t), q(t) \in \mathcal{P}(\mathbb{R})$ e $r \in \mathbb{R}$. Então

$$D(p(t) + q(t)) = [p(t) + q(t)]' = p'(t) + q'(t) = D(p(t)) + D(q(t))$$

e

$$D(rp(t)) = [rp(t)]' = rp'(t) = rD(p(t)).$$

Exemplo 3.4. *A aplicação*

$$\begin{array}{rccl} T: & \mathcal{C}([a,b],\mathbb{R}) & \to & \mathbb{R} \\ & f & \mapsto & \int_a^b f(x)dx \end{array}$$

é uma transformação linear. De fato, para quaisquer $f, g \in \mathcal{C}([a,b],\mathbb{R})$ e $k \in K$,

$$T(kf+g) = \int_a^b (kf+g)(x)dx = \int_a^b [kf(x)+g(x)]dx = k\int_a^b f(x)dx + \int_a^b g(x)dx = kT(f)+T(g).$$

Exemplo 3.5. *Seja $A \in \mathcal{M}_{m\times n}(K)$, K corpo, e considere*

$$\begin{array}{rccl} T: & \mathcal{M}_{n\times 1}(K) & \to & \mathcal{M}_{m\times 1}(K) \\ & X & \mapsto & T(X) = AX \end{array}$$

Temos que T é uma transformação linear. De fato, $\forall X_1, X_2 \in \mathcal{M}_{n\times 1}(K)$ e $k \in K$,

$$T(kX_1 + X_2) = A(kX_1 + X_2) = kAX_1 + AX_2 = kT(X_1) + T(X_2).$$

Exemplo 3.6. *Vamos verificar se as seguintes aplicações entre espaços vetoriais são ou não transformações lineares.*

a) $T : \mathbb{R}^3 \to \mathbb{R}^2$, *onde* $T(x, y, z) = (x + y, 2x - z)$.

Vamos verificar se T é transformação linear. Sejam $u_1 = (x_1, y_1, z_1), u_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$ e $r \in \mathbb{R}$. Temos que

$$\begin{aligned} T(ru_1 + u_2) &= T(rx_1 + x_2, ry_1 + y_2, rz_1 + z_2) \\ &= ((rx_1 + x_2) + (ry_1 + y_2), 2(rx_1 + x_2) - (rz_1 + z_2)) \\ &= (r(x_1 + y_1) + x_2 + y_2, r(2x_1 - z_1) + 2x_2 - z_2) \\ &= r(x_1 + y_1, 2x_1 - z_1) + (x_2 + y_2, 2x_2 - z_2) \\ &= rT(x_1, y_1, z_1) + T(x_2, y_2, z_2) \\ &= rT(u_1) + T(u_2) \end{aligned}$$

Então T é uma transformação linear.

b) $T : \mathbb{R}^2 \to \mathbb{R}^2$, *dada por* $T(x, y) = (x + y, 2)$.

Vamos verificar se T *é transformação linear. Sejam* $u_1 = (a, b), u_2 = (c, d) \in \mathbb{R}^2$. *Então*

$$T(u_1 + u_2) = T(a + c, b + d) = (a + c + b + d, 2)$$

e

$$T(u_1)+T(u_2) = T(a, b)+T(c, d) = (a+b, 2)+(c+d, 2) = (a+b+c+d, 4) \neq T(u_1+u_2).$$

Como T *não é compatível com a adição de vetores, temos que* T ***não é*** *uma transformação linear.*

Observação: Note que, dada uma transformação linear $T : U \to V$ e $u \in U$ qualquer, temos que $T(0) = T(0.u) = 0.T(u) = 0$, ou seja, T leva o vetor nulo de U no vetor nulo de V. No exemplo, temos $T(0, 0) = (0, 2)$. Essa observação já justificaria que T não é transformação linear.

c) $T : \mathbb{R}^2 \to \mathbb{R}^2$, *dada por* $T(x, y) = (x^2 + y^2, x)$.

Vamos verificar se T *é transformação linear. Sejam* $u_1 = (a, b), u_2 = (c, d) \in \mathbb{R}^2$ *e* $r \in \mathbb{R}$. *Então*

$$T(ru_1) = T(ra, rb) = ((ra)^2 + (rb)^2, ra) = (r^2a^2 + r^2b^2, ra) = r(ra^2 + rb^2, a)$$

e

$$rT(u_1) = rT(a, b) = r(a^2 + b^2, a).$$

Tomando, por exemplo, $r = 3$, *vemos que* $T(ru_1) \neq rT(u_1)$. *Então* T ***não é*** *uma transformação linear.*

Observação: Neste exemplo tem-se $T(0, 0) = (0, 0)$, mas ainda assim T não é transformação linear.

Exemplo 3.7 (Transformações do plano no plano)**.** *A seguir algumas transformações lineares geométricas de* $\mathbb{R}^2$ *em* $\mathbb{R}^2$.

Expansão ou contração uniformes*: dado* $k \in \mathbb{R}$, *com* $k > 0$, *defina*

$$\begin{array}{rccc} T: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x, y) & \mapsto & k(x, y) \end{array}$$

Temos que T *é uma transformação linear, pois, dados* $v_1 = (x_1, y_1), v_2 = (x_2, y_2) \in \mathbb{R}^2$ *e* $r \in \mathbb{R}$, *temos*

$$T(v_1+v_2) = T(x_1+x_2, y_1+y_2) = k(x_1+x_2, y_1+y_2) = k(x_1, y_1)+k(x_2, y_2) = T(v_1)+T(v_2)$$

$$e \ \ T(rv_1) = T(rx_1, ry_1) = k(rx_1, ry_1) = rk(x_1, y_1) = rT(v_1).$$

*Quando $k > 1$, dizemos que T é uma **expansão uniforme**.*

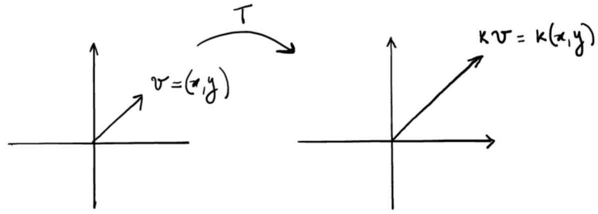
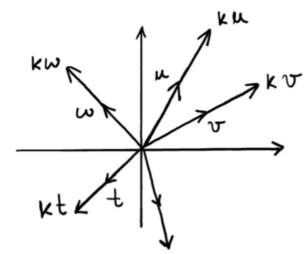


Figura 2: $(k > 1)$ Expansão Uniforme

Observe que, neste caso, T pega cada vetor de $\mathbb{R}^2$ e expande seu comprimento multiplicando-o por k. Como a transformação faz essa expansão em todos os vetores de $\mathbb{R}^2$, o efeito da transformação é uma expansão radial de todo o $\mathbb{R}^2$ mantendo a origem fixa.



*Quando $0 < k < 1$, dizemos que T é uma **contração uniforme**.*

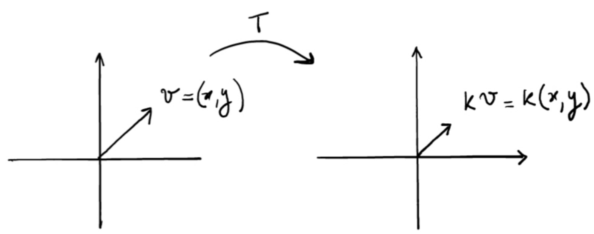
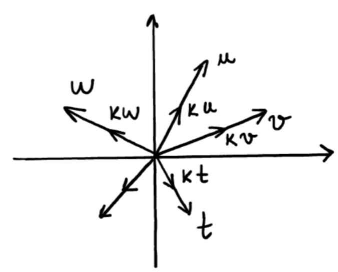
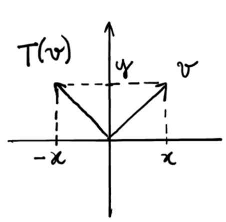


Figura 3: $(0 < k < 1)$ Contração Uniforme

Observe que, neste caso, T pega cada vetor de $\mathbb{R}^2$ e contrai seu comprimento multiplicando-o por k. Como a transformação faz essa contração em todos os vetores de $\mathbb{R}^2$, o efeito da transformação é uma contração radial de todo o $\mathbb{R}^2$ mantendo a origem fixa.



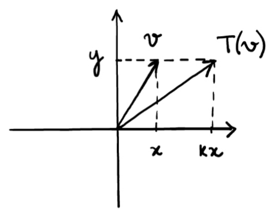
***Reflexão em relação ao eixo-**y:*



$$\begin{array}{rccc} T: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x,y) & \mapsto & (-x,y) \end{array}$$

(Verifique que T é transformação linear.)

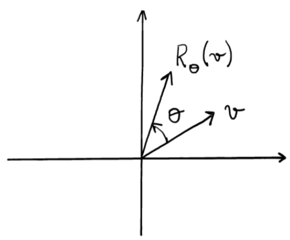
Expansão de um fator $k > 1$ ***na direção*** x:



$$\begin{array}{rccc} T: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x,y) & \mapsto & (kx,y) \end{array}$$

(Verifique que T é transformação linear.)

Rotação de ângulo θ ***no sentido anti-horário***:



$$R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$$

dada por

$R_\theta(x,y) = (x\cos\theta - y\,\text{sen}\theta, x\,\text{sen}\theta + y\cos\theta)$

(Verifique que R_θ é transformação linear.)

Observação 3.8. *Se $T : U \to V$ é uma transformação linear, então*

1) $T(0) = 0$;

2) $T(-u) = -T(u)$, $\forall u \in U$;

3) *se* $u = \sum_{i=1}^{m} k_i u_i$, *com* $k_i \in K$ *e* $u_i \in U$, *então* $T(u) = \sum_{i=1}^{m} k_i T(u_i)$.

O item 3) da observação acima, nos diz que se u é combinação linear de vetores u_i, então $T(u)$ fica completamente determinado pelas imagens $T(u_i)$ dos vetores u_i. Mais além, temos o seguinte.

Teorema 3.9. *Sejam U e V espaço vetoriais sobre K e $\{u_1, u_2, \ldots, u_n\}$ uma base de U. Então, para cada subconjunto $\{v_1, v_2, \ldots, v_n\} \subset V$, existe uma única transformação linear $T : U \to V$ tal que $T(u_i) = v_i$, $i = 1, 2, \ldots, n$.*

Demonstração. Como $\{u_1, u_2, \ldots, u_n\}$ é base de U, todo $u \in U$ pode ser escrito, de maneira única, como $u = k_1u_1 + k_2u_2 + \cdots + k_nu_n$, com $k_i \in K$. Com isso, defina

$$T : U \to V \quad \text{por} \quad T(u) = k_1v_1 + k_2v_2 + \cdots + k_nv_n.$$

Temos que:

(i) Como os k_i's são únicos, T está bem definida.

(ii) Como cada u_i é escrito como $u_i = 0u_1 + 0u_2 + \cdots + 1u_i + \cdots + 0u_n$, temos que

$$T(u_i) = 0v_1 + 0v_2 + \cdots + 1v_i + \cdots + 0v_n = v_i.$$

(iii) T é transformação linear. De fato, dados $u = k_1u_1 + k_2u_2 + \cdots + k_nu_n$ e $w = l_1u_1 + l_2u_2 + \cdots + l_nu_n$ em U e $r \in K$,

$$ru + w = (rk_1 + l_1)u_1 + (rk_2 + l_2)u_2 + \cdots + (rk_n + l_n)u_n$$

e

$$\begin{aligned} T(ru + w) &= (rk_1 + l_1)v_1 + (rk_2 + l_2)v_2 + \cdots + (rk_n + l_n)v_n \\ &= r(k_1v_1 + k_2v_2 + \cdots + k_nv_n) + l_1v_1 + l_2v_2 + \cdots + l_nv_n \\ &= rT(u) + T(w). \end{aligned}$$

(iv) T é única tal que $T(u_i) = v_i$. De fato, suponha que $S : U \to V$ é uma transformação linear tal que $S(u_i) = v_i$. Então, $\forall u \in U$, $u = k_1u_1 + k_2u_2 + \cdots + k_nu_n$, tem-se

$$\begin{aligned} S(u) &= S(k_1u_1 + k_2u_2 + \cdots + k_nu_n) \\ &= k_1S(u_1) + k_2S(u_2) + \cdots + k_nS(u_n) \\ &= k_1v_1 + k_2v_2 + \cdots + k_nv_n \\ &= T(u). \end{aligned}$$

Logo $S = T$. ∎

Exemplo 3.10. *Sabendo que $T : \mathbb{R}^2 \to \mathbb{R}^2$ é uma transformação linear e que*

$$T(1, 2) = (3, -1) \quad e \quad T(0, 1) = (-2, 1),$$

vamos encontrar $T(x, y)$, para $(x, y) \in \mathbb{R}^2$ qualquer.

Primeiro observamos que os vetores $(1, 2)$ e $(0, 1)$ são L.I. e assim formam uma base de R^2. Então cada $(x, y) \in \mathbb{R}^2$ pode ser escrito como combinação linear deles:

$$a(1, 2) + b(0, 1) = (x, y) \Leftrightarrow \begin{cases} a = x \\ 2a + b = y \quad \Rightarrow b = y - 2a = y - 2x \end{cases}$$

Então

$$x(1, 2) + (y - 2x)(0, 1) = (x, y).$$

Agora, como T é transformação linear, obtemos

$$\begin{aligned} T(x, y) &= T(x(1, 2) + (y - 2x)(0, 1)) \\ &= x\ T(1, 2) + (y - 2x)\ T(0, 1) \\ &= x\ (3, -1) + (y - 2x)\ (-2, 1) \\ &= (3x - 2y + 4x, -x + y - 2x) \\ &= (7x - 2y, -3x + y) \end{aligned}$$

Portanto $T(x, y) = (7x - 2y, -3x + y)$ é a transformação linear dada.

3.1 Núcleo e Imagem

Sejam U e V espaços vetoriais sobre K e $T : U \to V$ uma transformação linear. Então, via T, estamos levando o espaço U para um subconjunto do espaço V, que é a **imagem de** T, ou seja, o subconjunto de V dado por

$$\mathrm{Im}(T) := \{T(u) \mid u \in U\} = \{v \in V \mid v = T(u), \ \text{ para algum } \ u \in U\}.$$

Mas para uma tal transformação linear geral $T : U \to V$, alguma perda pode acontecer. Por exemplo, a transfomação linear do Exemplo 3.6 a), leva um espaço vetorial de dimensão 3 em um espaço de dimensão 2, logo algo se perdeu, digamos assim. Para compreender essa "perda", vamos considerar o subconjunto de U, que chamamos de **núcleo de** T, dado por

$$\mathrm{Ker}(T) := \{u \in U \mid T(u) = 0 \ \text{ (vetor nulo de } V)\}.$$

Exemplo 3.11. *Voltando no Exemplo 3.6 a), considere $T : \mathbb{R}^3 \to \mathbb{R}^2$, onde $T(x, y, z) = (x + y, 2x - z)$.*

O núcleo de T é dado por

$$\begin{aligned} \mathrm{Ker}(T) &= \{(x, y, z) \in \mathbb{R}^3 \mid T(x, y, z) = (0, 0)\} \\ &= \{(x, y, z) \in \mathbb{R}^3 \mid (x + y, 2x - z) = (0, 0)\} \\ &= \{(x, y, z) \in \mathbb{R}^3 \mid x + y = 0 \ \ e \ \ 2x - z = 0\} \end{aligned}$$

Então Ker$T$ *é o conjunto-solução do sistema linear homogêneo*

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

Assim $\mathrm{Ker}(T) = \{(x, -x, 2x) \in \mathbb{R}^3 \mid x \in \mathbb{R}\} = [(1, -1, 2)]$.

E a imagem de T é

$$\begin{aligned} \mathrm{Im}(T) &= \{T(x, y, z) \mid (x, y, z) \in \mathbb{R}^3\} \\ &= \{(x + y, 2x - z) \mid x, y, z \in \mathbb{R}\} \\ &= \{x(1, 2) + y(1, 0) + z(0, -1) \mid x, y, z \in \mathbb{R}\} \\ &= [(1, 2), (1, 0), (0, -1)]. \end{aligned}$$

Assim, $\mathrm{Im}T = \mathbb{R}^2$.

Exemplo 3.12. *Sejam U e V espaços vetoriais sobre K. O núcleo da transformação nula $T : U \to V$, onde $Tu = 0$, $\forall u \in U$, é*

$$\mathrm{Ker}T = \{u \in U \mid Tu = 0\} = U$$

e sua imagem é

$$\mathrm{Im}T = \{T(u) \mid u \in U\} = \{0\}.$$

O núcleo da transformação indentidade $I : U \to U$, $I(u) = u$, $\forall u \in U$, é dado por

$$\mathrm{Ker}I = \{u \in U \mid I(u) = 0\} = \{u \in U \mid u = 0\} = \{0\}$$

e sua imagem é

$$\mathrm{Im}I = \{I(u) \mid u \in U\} = \{u \mid u \in U\} = U.$$

Exemplo 3.13. *Seja $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$, dada pela derivada, como no Exemplo 3.3.*
O núcleo de D é

$$\begin{aligned}\mathrm{Ker}D &= \{p(t) \in \mathcal{P}(\mathbb{R}) \mid D(p(t)) = 0\}\\ &= \{p(t) \in \mathcal{P}(\mathbb{R}) \mid p'(t) = 0\}\\ &= \{a_0 \mid a_0 \in \mathbb{R}\}\end{aligned}$$

e sua imagem é

$$\begin{aligned}\mathrm{Im}D &= \{p(t) \in \mathcal{P}(\mathbb{R}) \mid p(t) = D(q(t)), \quad \text{para algum } q(t) \in \mathcal{P}(\mathbb{R})\}\\ &= \mathcal{P}(\mathbb{R}).\end{aligned}$$

Exemplo 3.14. *Seja $A \in \mathcal{M}_{m\times n}(K)$, K corpo, e considere*

$$\begin{array}{rccc} T: & \mathcal{M}_{n\times 1}(K) & \to & \mathcal{M}_{m\times 1}(K)\\ & X & \mapsto & T(X) = AX\end{array}$$

como no Exemplo 3.5. O núcleo de T é

$$\mathrm{Ker}T = \{X \mid AX = 0\}$$

e a imagem de T é

$$\mathrm{Im}T = \{AX \mid X \in \mathcal{M}_{n\times 1}(K)\}.$$

Para além de subconjuntos, o núcleo e a imagem são subespaços, como veremos a seguir.

Proposição 3.15. *Seja $T : U \to V$ uma transformação linear. Então*

i) $\mathrm{Ker}(T)$ *é um subespaço de U;*

ii) $\mathrm{Im}(T)$ *é um subespaço de V.*

Demonstração. i) Por definição, $\mathrm{Ker}(T) := \{u \in U \mid T(u) = 0\}$. Então um vetor $u \in U$ pertence ao núcleo de T quando é levado no vetor nulo por T.

Primeiro, temos que, como T é transformação linear, T leva o vetor nulo de U no vetor nulo de V, ou seja, $T(0) = 0$. Logo $0 \in \mathrm{Ker}(T)$.

Agora vamos mostar que $\mathrm{Ker}(T)$ é fechado para soma e para a multiplicação por escalar. Sejam $u_1, u_2 \in \mathrm{Ker}(T)$ e $k \in K$. Como $u_1, u_2 \in \mathrm{Ker}(T)$, temos que $T(u_1) = 0$ e $T(u_2) = 0$. Com isso,

$$T(ku_1 + u_2) = kT(u_1) + T(u_2) = k.0 + 0 = 0,$$

logo $T(ku_1 + u_2) = 0$ e com isso $ku_1 + u_2 \in \mathrm{Ker}(T)$. Portanto $\mathrm{Ker}(T)$ é um subespaço de U.

ii) Por definição, $\mathrm{Im}(T) = \{T(u) \mid u \in U\}$. Como temos $0 = T(0)$, temos que $0 \in \mathrm{Im}(T)$. Sejam $T(u_1), T(u_2) \in \mathrm{Im}(T)$ e $k \in K$. Então

$$kT(u_1) + T(u_2) = T(ku_1 + u_2) \in \mathrm{Im}(T),$$

o que mostra que $\mathrm{Im}(T)$ é um subespaço de V.

■

Sejam U e V espaços vetoriais sobre K. Lembremos que uma função qualquer $T : U \to V$ é **sobrejetora** quando $\text{Im}(T) = V$. E T é **injetora** quando, dados $u_1, u_2 \in U$, com $u_1 \neq u_2$, então $T(u_1) \neq T(u_2)$; ou, equivalentemente, dados $u_1, u_2 \in U$, se $T(u_1) = T(u_2)$ então $u_1 = u_2$.

Por definição, a imagem de uma transformação linear $T : U \to V$ já descreve sua sobrejetividade. Agora, em relação à injetividade, notemos o seguinte. Toda transformação linear $T : U \to V$ leva vetor nulo em vetor nulo. Com isso, se o núcleo da transformação contiver algum vetor não nulo, então teremos vetores distintos em U sendo levados no vetor nulo de V e assim a transformação não será injetora. De fato, veremos a seguir que o núcleo de uma transformação linear caracteriza sua injetividade.

Proposição 3.16. *Seja $T : U \to V$ uma transformação linear.*

i) *Se V tem dimensão finita, então T é sobrejetora se, e somente se,* $\dim \text{Im}T = \dim V$.

ii) *T é injetora se, e somente se,* $\text{Ker}(T) = \{0\}$.

Demonstração. (i) Por definição, T é sobrejetora quando $\text{Im}T = V$, e temos que $\text{Im}T = V$ se, e somente se, $\dim \text{Im}T = \dim V$.

(ii) Primeiro suponha que T é injetora e vamos mostrar que isso implica que $\text{Ker}(T) = \{0\}$. Para isso, seja $u \in \text{Ker}(T)$ um vetor qualquer no núcleo de T. Como $u \in \text{Ker}(T)$, temos $T(u) = 0$. Mas também $T(0) = 0$, assim $T(u) = 0 = T(0)$. Como T é injetora e $T(u) = T(0)$, segue que $u = 0$. Mostramos então que se um vetor pertence ao núcleo do T, ele é nulo, assim $\text{Ker}(T) = \{0\}$.

Reciprocamente, suponha que $\text{Ker}(T) = \{0\}$ e vamos mostrar que isso implica que T é injetora. Para isso, sejam $u_1, u_2 \in U$ tais que $T(u_1) = T(u_2)$. Para que T seja injetora, devemos mostrar que $u_1 = u_2$. De fato,

$$T(u_1) = T(u_2) \Rightarrow T(u_1) - T(u_2) = 0 \Rightarrow T(u_1 - u_2) = 0.$$

Então, como $T(u_1 - u_2) = 0$, obtemos $u_1 - u_2 \in \text{Ker}(T)$. Mas, por hipótese, $\text{Ker}(T) = \{0\}$. Logo $u_1 - u_2 = 0$ e portanto $u_1 = u_2$. Assim, T é injetora.

■

Voltando aos exemplos anteriores:

Exemplo 3.17. *$T : \mathbb{R}^3 \to \mathbb{R}^2$, onde $T(x, y, z) = (x + y, 2x - z)$ não é injetora, pois* $\text{Ker}(T) = [(1, -1, 2)] \neq \{(0, 0, 0)\}$. *E T é sobrejetora, pois* $\text{Im}T = \mathbb{R}^2$.

Exemplo 3.18. *A transformação indentidade $I : U \to U$, $I(u) = u$, $\forall u \in U$, é injetora e sobrejetora, pois* $\text{Ker}I = \{0\}$ *e* $\text{Im}I = U$.

Exemplo 3.19. *$D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$, dada pela derivada, não é injetora, pois* $\text{Ker}D = \{a_0 \mid a_0 \in \mathbb{R}\} \neq \{0\}$, *e é sobrejetora, pois* $\text{Im}D = \mathcal{P}(\mathbb{R})$. *De fato, $\forall p(t) \in \mathcal{P}(\mathbb{R})$, pelo Teorema Fundamental do Cálculo, tome q(t)=P(t) tal que $P'(t) = p(t)$; assim $p(t) = D(q(t)) \in \text{Im}D$.*

Exemplo 3.20. *Seja $A \in \mathcal{M}_n(K)$, K corpo, e considere $T : \mathcal{M}_{n\times 1}(K) \to \mathcal{M}_{n\times 1}(K)$, dada por $T(X) = AX$. Temos que*

$$\mathrm{Ker}T = \{X \mid AX = 0\} = \{0\} \quad \Leftrightarrow \quad A \ \textit{é inversível.}$$

Então A é injetora se, e somente se, A é inversível; e, neste caso, T é sobrejetora, pois, para toda matriz coluna B, o sistema linear $AX = B$ tem solução, ou seja, existe X tal que $AX = B$, de onde obtemos que $B = T(X) \in \mathrm{Im}T$.

Seja $T : U \to V$ uma transformação linear e $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ um conjunto de geradores de U. Então, $\forall u \in U$, existem escalares $k_1, k_2, \ldots, k_n \in K$ tais que $u = \sum_{i=1}^n k_i u_i$. Logo

$$Tu = \sum_{i=1}^{n} k_i T u_i,$$

de onde vemos que $\{Tu_1, Tu_2, \ldots, Tu_n\}$ gera $\mathrm{Im}T$. Quando $\mathcal{B}$ é base de U, obtemos em particular que $\{Tu_1, Tu_2, \ldots, Tu_n\}$ gera $\mathrm{Im}T$. Em geral, não podemos afirmar que $T(\mathcal{B})$ é base de $\mathrm{Im}T$, isso depende da injetividade de T.

Teorema 3.21 (Teorema do Núcleo e da Imagem). *Seja $T : U \to V$ uma transformação linear, onde U e V são espaços vetoriais sobre K e U é de dimensão finita. Então*

$$\dim U = \dim \mathrm{Ker}(T) + \dim \mathrm{Im}(T).$$

Demonstração. Seja $\{u_1, u_2, \ldots, u_r\}$ uma base de $\mathrm{Ker}(T)$. O núcleo de T está contido em U, então os vetores $u_1, u_2, \ldots, u_r$ pertencem a U e, por serem base de $\mathrm{Ker}(T)$, são L.I. Assim $\{u_1, u_2, \ldots, u_r\}$ é um conjunto L.I. contido em U. Podemos então completar esse conjunto até obtermos uma base de U

$$\mathcal{B} = \{u_1, u_2, \ldots, u_r, w_1, w_2, \ldots, w_s\}.$$

Com isso,

$$\dim U = r + s, \quad \text{onde} \quad r = \dim \mathrm{Ker}(T).$$

Agora vamos olhar para a imagem de T, $\mathrm{Im}(T) = \{T(u) \mid u \in U\}$. Seja $u \in U$ um vetor qualquer de U. Então, com $\mathcal{B}$ é base de U, podemos escrever u como combinação linear

$$u = a_1u_1 + a_2u_2 + \cdots + a_ru_r + b_1w_1 + b_2w_2 + \cdots + b_sw_s$$

com $a_1, a_2, \ldots, a_r, b_1, b_2, \ldots, b_s \in K$. Aplicando T, obtemos

$$\begin{aligned} T(u) &= T(a_1u_1 + a_2u_2 + \cdots + a_ru_r + b_1w_1 + b_2w_2 + \cdots + b_sw_s) \\ &= a_1T(u_1) + a_2T(u_2) + \cdots + a_rT(u_r) + b_1T(w_1) + b_2T(w_2) + \cdots + b_sT(w_s) \end{aligned}$$

Como $u_1, u_2, \ldots, u_r \in \mathrm{Ker}(T)$, temos que $T(u_1) = 0$, $T(u_2) = 0$, …, $T(u_r) = 0$. Com isso, obtemos

$$T(u) = b_1T(w_1) + b_2T(w_2) + \cdots + b_sT(w_s).$$

Segue então que todo vetor $T(u)$ pode ser escrito como combinação linear de $T(w_1)$, $T(w_2)$, ..., $T(w_s)$, ou seja, o conjunto

$$\{T(w_1), T(w_2), \ldots, T(w_s)\}$$

gera $\mathrm{Im}(T)$. Vamos mostrar que este conjunto é L.I.:

$$b_1T(w_1) + b_2T(w_2) + \cdots + b_sT(w_s) = 0$$

$$\Leftrightarrow T(b_1w_1 + b_2w_2 + \cdots + b_sw_s) = 0$$

$$\Leftrightarrow b_1w_1 + b_2w_2 + \cdots + b_sw_s \ \in \mathrm{Ker}(T)$$

Como $\{u_1, u_2, \ldots, u_r\}$ é uma base de $\mathrm{Ker}(T)$, o vetor acima pode ser escrito como combinação linear dos vetores $u_1, u_2, \ldots, u_r$ ou seja, existem escalares $a_1, a_2, \ldots, a_r \in K$ tais que

$$b_1w_1 + b_2w_2 + \cdots + b_sw_s = a_1u_1 + a_2u_2 + \cdots + a_ru_r.$$

Com isso, obtemos

$$-a_1u_1 - a_2u_2 - \cdots - a_ru_r + b_1w_1 + b_2w_2 + \cdots + b_sw_s = 0.$$

Mas $\mathcal{B}$ é L.I. então tal combinação linear só é possível com escalares todos nulos. Assim $a_1 = a_2 = \cdots = a_r = 0$ e $b_1 = b_2 = \cdots = b_s = 0$. Como então $b_1 = b_2 = \cdots = b_s = 0$, obtemos que $\{T(w_1), T(w_2), \ldots, T(w_s)\}$ é L.I.

Portanto $\{T(w_1), T(w_2), \ldots, T(w_s)\}$ gera $\mathrm{Im}(T)$ e é L.I., assim é uma base de $\mathrm{Im}(T)$ e $\dim \mathrm{Im}(T) = s$. Com isso, obtemos a relação desejada

$$\dim U = r + s = \dim \mathrm{Ker}(T) + \dim \mathrm{Im}(T).$$

$\blacksquare$

3.2 Isomorfismos

Definição 3.22. *Sejam U e V espaços vetoriais sobre um corpo K e $T : U \to V$ uma transformação linear. Se T é uma bijeção, ou seja, injetora e sobrejetora, dizemos que T é um isomorfismo de U em V. E dizemos que U e V são espaços vetoriais isomorfos quando existe um isomorfismo de U em V.*

Notação: $U \cong V$

Observação 3.23. *Note que se U tem dimensão finita e $U \cong V$, então existe um isomorfimo $T : U \to V$. Com isso,* $\mathrm{Im}T = V$ *e* $\dim U = \dim \mathrm{Im}T = \dim V$. *Assim, V tem dimensão finita e* $\dim V = \dim U$.

Teorema 3.24. *Sejam U e V espaços vetoriais sobre K. Se* $\dim U = \dim V < \infty$ *e $T : U \to V$ é uma transformação linear, então são equivalentes:*

i) *T é isomorfismo;*

ii) *T é injetora;*

iii) *T é sobrejetora.*

A prova do teorema segue do Teorema do Núcleo e da Imagem.

Observação 3.25. *Note que a hipótese de dimensão finita é necessária. Por exemplo, a transformação linear D no Exemplo 3.19 é sobrejetora e não é injetora.*

Teorema 3.26. *Dois espaços vetoriais sobre um corpo K de mesma dimensão finita são isomorfos.*

Demonstração. Suponha que $\dim U = \dim V = n$ e sejam $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ e $\mathcal{B}' = \{v_1, v_2, \ldots, v_n\}$ bases de U e V, respectivamente. Então, pelo Teorema 3.9, existe uma única transformação linear $T : U \to V$ tal que $T(u_i) = v_i$, $i = 1, 2, \ldots, n$. Temos que T é sobrejetora, pois $\dim \mathrm{Im}T = n = \dim V$. Logo, pelo Teorema 3.24, T é isomorfismo. ∎

Corolário 3.27. *Todo espaço vetorial de dimensão finita n sobre K é isomorfo a K^n.*

3.3 Matrizes de uma Transformação Linear

Exemplo 3.28. *Considere a transformação linear* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *dada por*

$$T(x, y, z) = (4x - 2y + z, 2x + z, 2x - 2y + 3z).$$

Representando o vetor (x, y, z) *na base canônica* $\mathcal{C}$*, temos que* $T(x, y, z)$ *pode ser representado matricialmente por*

$$\begin{pmatrix} 4 & -2 & 1 \\ 2 & 0 & 1 \\ 2 & -2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Agora, observe que $[(x, y, z)]_{\mathcal{C}} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ *são as coordenadas do vetor* (x, y, z) *na base canônica. Além disso, as colunas da matriz*

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

são as coordenadas de $T(1, 0, 0)$*,* $T(0, 1, 0)$ *e* $T(0, 0, 1)$*: de fato,*

$$T(1, 0, 0) = (4, 2, 2), \quad T(0, 1, 0) = (-2, 0, -2) \quad e \quad T(0, 0, 1) = (1, 1, 3),$$

assim

$$[T(1, 0, 0)]_{\mathcal{C}} = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}, \quad [T(0, 1, 0)]_{\mathcal{C}} = \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix} \quad e \quad [T(0, 0, 1)]_{\mathcal{C}} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

Por isso, denotamos a matriz A *por*

$$[T]_{\mathcal{C}},$$

pois ela é obtida ao aplicarmos T *nos vetores de* $\mathcal{C}$ *e escrevermos suas imagens na base* $\mathcal{C}$*. Essa matriz é então chamada de* ***matriz de*** T ***na base*** $\mathcal{C}$*.*

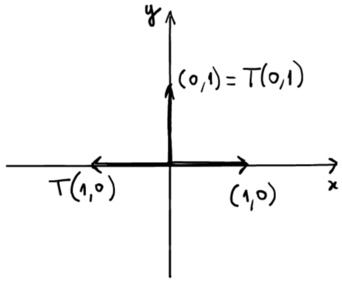
Exemplo 3.29. *Vamos encontrar as matrizes de algumas das transformações lineares geométricas de* $\mathbb{R}^2$ *em* $\mathbb{R}^2$ *do Exemplo 3.7 na base canônica* $\mathcal{C}$*.*

1. *Expansão ou contração uniformes:* $T : \mathbb{R}^2 \to \mathbb{R}^2$*, dada por* $T(x, y) = k(x, y)$*, onde* $k \in \mathbb{R}$*, com* $k > 0$*. A matriz de* T *em relação à base canônica é:*

$$T(1, 0) = k(1, 0) = k(1, 0) + 0(0, 1)$$
$$T(0, 1) = k(0, 1) = 0(1, 0) + k(0, 1)$$

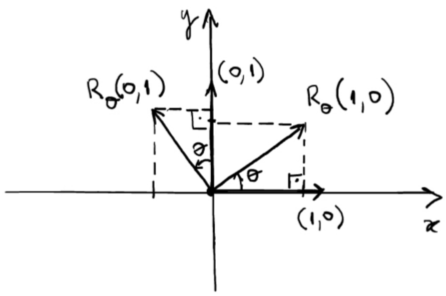
$$\Rightarrow [T]_{\mathcal{C}} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$$

2. Reflexão em relação ao eixo-y:



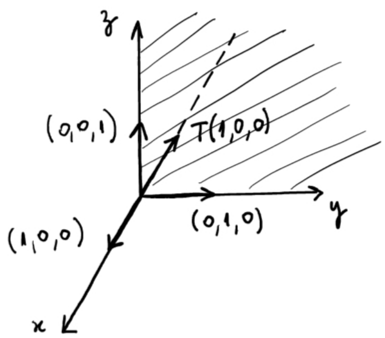
$$T : \mathbb{R}^2 \to \mathbb{R}^2$$
$$T(1,0) = (-1,0) \quad \text{e} \quad T(0,1) = (0,1)$$
$$\Rightarrow [T]_{\mathcal{C}} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

3. Rotação de ângulo θ no sentido anti-horário:



$$R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$$
$$R_\theta(1,0) = (\cos\theta,\, \operatorname{sen}\theta)$$
$$R_\theta(0,1) = (-\operatorname{sen}\theta, \cos\theta)$$
$$\Rightarrow [T]_{\mathcal{C}} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

Exemplo 3.30. *Considere a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada pela reflexão em relação ao plano $x = 0$. Temos que*



$$T(1,0,0) = (-1,0,0)$$
$$T(0,1,0) = (0,1,0)$$
$$T(0,0,1) = (0,0,1)$$

Logo

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

A representação matricial de uma transfomação linear pode ser feita em relação a bases quaisquer dos espaços envolvidos na transformação, como veremos a seguir.

Sejam V e W espaços vetoriais de dimensão finita e $T : V \to W$ uma transformação linear. Considere

$$\mathcal{B} = \{v_1, v_2, \ldots, v_n\} \ \text{ uma base de } \ V$$

e

$$\mathcal{B}' = \{w_1, w_2, \ldots, w_m\} \ \text{ uma base de } \ W.$$

Vamos associar à T uma matriz $m \times n$ da seguinte maneira. Como $T(v_1)$, $T(v_2)$,..., $T(v_n)$ pertencem a W e $\mathcal{B}'$ é base de W, podemos escrever cada $T(v_j)$ como combinação dos vetores da base $\mathcal{B}'$:

$$\begin{aligned} T(v_1) &= a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m \\ T(v_2) &= a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m \\ &\vdots \\ T(v_n) &= a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m \end{aligned} \tag{7}$$

com $a_{ij} \in K$, $\forall i, j$. Com isso, definimos a **matriz de T em relação às bases $\mathcal{B}$ e $\mathcal{B}'$** por

$$[T]_{\mathcal{B}'}^{\mathcal{B}} := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} | & | & & | \\ [T(v_1)]_{\mathcal{B}'} & [T(v_2)]_{\mathcal{B}'} & \cdots & [T(v_n)]_{\mathcal{B}'} \\ | & | & & | \end{pmatrix}$$

Então, por definição, para encontrarmos a matriz $[T]_{\mathcal{B}'}^{\mathcal{B}}$, aplicamos T em cada um dos vetores v_j de $\mathcal{B}$, escrevemos cada $T(v_j)$ na base $\mathcal{B}'$ e colocamos os escalares desta combinação linear na coluna j de $[T]_{\mathcal{B}'}^{\mathcal{B}}$.

Exemplo 3.31. *Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$, dada por*

$$T(x, y, z) = (2x + y - z, 3x - 2y + 4z)$$

e considere as bases

$$\mathcal{B} = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\} \quad e \quad \mathcal{B}' = \{(1, 0), (1, 1)\}$$

de $\mathbb{R}^3$ e $\mathbb{R}^2$, respectivamente. Vamos encontrar a matriz $[T]_{\mathcal{B}'}^{\mathcal{B}}$, ou seja, a matriz de T em relação às bases $\mathcal{B}$ e $\mathcal{B}'$. Temos

$$\begin{aligned} T(1, 1, 1) &= (2, 5) = -3(1, 0) + 5(1, 1) \\ T(1, 1, 0) &= (3, 1) = 2(1, 0) + 1(1, 1) \\ T(1, 0, 0) &= (2, 3) = -1(1, 0) + 3(1, 1) \end{aligned}$$

Assim

$$[T]_{\mathcal{B}'}^{\mathcal{B}} = \begin{pmatrix} -3 & 2 & -1 \\ 5 & 1 & 3 \end{pmatrix}$$

Exemplo 3.32. *Considere a transformação identidade*

$$I : V \to V, \quad \textit{dada por } I(v) = v,$$

e bases $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ e $\mathcal{B}' = \{w_1, w_2, \ldots, w_n\}$ de V. A matriz de I em relação às bases $\mathcal{B}$ e $\mathcal{B}'$ é dada por

$$[I]_{\mathcal{B}'}^{\mathcal{B}} = \begin{pmatrix} | & | & & | \\ [I(v_1)]_{\mathcal{B}'} & [I(v_2)]_{\mathcal{B}'} & \cdots & [I(v_n)]_{\mathcal{B}'} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ [v_1]_{\mathcal{B}'} & [v_2]_{\mathcal{B}'} & \cdots & [v_n]_{\mathcal{B}'} \\ | & | & & | \end{pmatrix},$$

ou seja, $[I]_{\mathcal{B}'}^{\mathcal{B}}$ é a matriz de mudança de base da base $\mathcal{B}$ para a $\mathcal{B}'$.

Voltando ao caso geral, para cada $v \in V$, escrevendo v na base $\mathcal{B}$ de V

$$v = x_1v_1 + x_2v_2 + \cdots + x_nv_n, \quad \text{com} \quad x_1, x_2, \ldots, x_n \in K,$$

temos

$$[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Aplicando T em v, obtemos

$$\begin{aligned} T(v) &= x_1T(v_1) + x_2T(v_2) + \cdots + x_nT(v_n) \\ &= \textstyle\sum_{j=1}^{n} x_jT(v_j) \\ &= \textstyle\sum_{j=1}^{n} x_j \left(\sum_{i=1}^{n} a_{ij}w_i\right) && \text{[pelas relações de (7)]} \\ &= \textstyle\sum_{j=1}^{n} \sum_{i=1}^{m} x_ja_{ij}w_i \\ &= \textstyle\sum_{i=1}^{m} \left(\sum_{j=1}^{n} x_ja_{ij}\right) w_i \end{aligned}$$

Note que, na última igualdade acima, temos $T(v)$ escrito como combinação linear dos vetores w_i da base $\mathcal{B}'$. Assim as coordenadas de $T(v)$na base $\mathcal{B}'$ são

$$[T(v)]_{\mathcal{B}'} = \begin{pmatrix} \sum_{j=1}^{n} x_ja_{1j} \\ \sum_{j=1}^{n} x_ja_{2j} \\ \vdots \\ \sum_{j=1}^{n} x_ja_{mj} \end{pmatrix} = \begin{pmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = [T]_{\mathcal{B}'}^{\mathcal{B}} \cdot [v]_{\mathcal{B}}.$$

Obtemos então a relação

$$\boxed{[T(v)]_{\mathcal{B}'} = [T]_{\mathcal{B}'}^{\mathcal{B}} \cdot [v]_{\mathcal{B}}} \tag{8}$$

Forma matricial de T

Exemplo 3.33. *Voltando no Exemplo 3.31, para*

$$T : \mathbb{R}^3 \to \mathbb{R}^2, \quad \textit{dada por} \quad T(x, y, z) = (2x + y - z, 3x - 2y + 4z),$$

e bases

$$\mathcal{B} = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\} \quad e \quad \mathcal{B}' = \{(1, 0), (1, 1)\},$$

vimos que a matriz de T nas bases $\mathcal{B}$ e $\mathcal{B}'$ é

$$[T]^{\mathcal{B}}_{\mathcal{B}'} = \begin{pmatrix} -3 & 2 & -1 \\ 5 & 1 & 3 \end{pmatrix}.$$

Considere agora o vetor $v = (2, -3, -1) \in \mathbb{R}^3$. Aplicando T, obtemos

$$T(2, -3, -1) = (4 - 3 - 1, 6 + 6 - 4) = (2, 8).$$

Por outro lado, escrevendo v na base $\mathcal{B}$, obtemos

$$(2, -3, -1) = -1(1, 1, 1) + (-2)(1, 1, 0) + 5(1, 0, 0), \quad \textit{assim} \ \ [(2, -3, -1)]_{\mathcal{B}} = \begin{pmatrix} -1 \\ -2 \\ 5 \end{pmatrix}.$$

Então

$$\begin{aligned} [T(2, -3, -1)]_{\mathcal{B}'} &= [T]^{\mathcal{B}}_{\mathcal{B}'} \cdot [(2, -3, -1)]_{\mathcal{B}} \\ & \begin{pmatrix} -3 & 2 & -1 \\ 5 & 1 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ -2 \\ 5 \end{pmatrix} = \begin{pmatrix} -6 \\ 8 \end{pmatrix} \end{aligned}$$

e com isso

$$T(2, -3, -1) = -6(1, 0) + 8(1, 1) = (2, 8).$$

Isso ilustra o uso da relação (8).

Exemplo 3.34. *Sejam*

$V = \mathcal{P}_3(\mathbb{R}) = \{a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_0, a_1, a_2, a_3 \in \mathbb{R}\}$*;*

$D : \mathcal{P}_3(\mathbb{R}) \to \mathcal{P}_3(\mathbb{R})$ *a derivação;*

$\mathcal{B} = \{1, x, x^2, x^3\}$.

A matriz de D na base $\mathcal{B}$ é dada por

$$[D]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} | & | & | & | \\ [D(1)]_{\mathcal{B}} & [D(x)]_{\mathcal{B}} & [D(x^2)]_{\mathcal{B}} & [D(x^3)]_{\mathcal{B}} \\ | & | & | & | \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Com isso, tomando $p(x) = 2 + x - 3x^2 + x^3$, temos que

$$\begin{aligned} [D(2 + x - 3x^2 + x^3)]_{\mathcal{B}} &= [D]^{\mathcal{B}}_{\mathcal{B}} \cdot [2 + x - 3x^2 + x^3]_{\mathcal{B}} \\ &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -6 \\ 3 \\ 0 \end{pmatrix}, \end{aligned}$$

o que nos diz que

$$D(2 + x - 3x^2 + x^3) = 1 - 6x + 3x^2 + 0x^3.$$

Observação 3.35. *Sejam U e V espaços vetoriais sobre K de dimensões finitas* $\dim U = n > 0$ *e* $\dim V = m > 0$ *e $\mathcal{B}$ e $\mathcal{B}'$ bases de U e V, respectivamente.*

(i) Dada uma transformação linear $T : U \to V$, fixadas as bases $\mathcal{B}$ e $\mathcal{B}'$, a matriz $[T]^{\mathcal{B}}_{\mathcal{B}'}$ é única.

(ii) A recíproca é também verdadeira, ou seja, dada uma matriz $A \in \mathcal{M}_{m\times n}(K)$, fixadas as bases $\mathcal{B}$ e $\mathcal{B}'$, existe uma única transformação linear $T : U \to V$ tal que $[T]^{\mathcal{B}}_{\mathcal{B}'} = A$.

De fato, vimos no Teorema 3.9 que para cada base $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ de U e qualquer subconjunto $\{v_1, v_2, \ldots, v_n\} \subset V$, existe uma única transformação linear $T : U \to V$ tal que $T(u_i) = v_i$. Basta então definir $[T(u_i)]_{\mathcal{B}'}$ como a i-ésima coluna de A.

A seguir veremos uma série de consequências importantes da álgebra de matrizes ao se representar transformações lineares por suas matrizes.

Teorema 3.36. *Sejam $T : U \to V$ e $S : V \to W$ transformações lineares e $m = \dim_K U$, $n = \dim_K V$ e $l = \dim_K W$. Fixadas bases $\mathcal{B}$, $\mathcal{B}'$ e $\mathcal{B}''$ de U, V e W, respectivamente, temos*

$$\boxed{[S \circ T]^{\mathcal{B}}_{\mathcal{B}''} = [S]^{\mathcal{B}'}_{\mathcal{B}''} \cdot [T]^{\mathcal{B}}_{\mathcal{B}'}}$$

Matriz da transformação composta

ou seja, a matriz da transformação composta $S \circ T$ é o produto das matrizes de S e de T.

Demonstração. Sejam

$$\mathcal{B} = \{u_1, u_2, \ldots, u_n\}, \quad \mathcal{B}' = \{v_1, v_2, \ldots, v_m\} \quad \text{e} \quad \mathcal{B}'' = \{w_1, w_2, \ldots, w_l\}.$$

Para encontrarmos as matrizes

$$[T]^{\mathcal{B}}_{\mathcal{B}'} \qquad \text{e} \qquad [S]^{\mathcal{B}'}_{\mathcal{B}''}$$

por definição, escrevemos

$$\begin{array}{ll}
T(u_1) = a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_m & \qquad S(v_1) = b_{11}w_1 + b_{21}w_2 + \cdots + b_{l1}w_l \\
T(u_2) = a_{12}v_1 + a_{22}v_2 + \cdots + a_{m2}v_m & \qquad S(v_2) = b_{12}w_1 + b_{22}w_2 + \cdots + b_{l2}w_l \\
\quad\vdots & \qquad\quad \vdots \\
T(u_n) = a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{mn}v_m & \qquad S(v_m) = b_{1m}w_1 + b_{2m}w_2 + \cdots + b_{lm}w_l
\end{array}$$

que podemos escrever como

$$T(u_j) = \sum_{i=1}^{m} a_{ij}v_i \qquad \text{e} \qquad S(v_i) = \sum_{k=1}^{l} b_{ki}w_k.$$

Com isso,

$$[T]^{\mathcal{B}}_{\mathcal{B}'} = (a_{ij}) = A \qquad \text{e} \qquad [S]^{\mathcal{B}'}_{\mathcal{B}''} = (b_{ki}) = B$$

Agora, para determinar a matriz $[S \circ T]^{\mathcal{B}}_{\mathcal{B}''}$, vamos encontrar $(S \circ T)(u_j)$ para cada $u_j \in \mathcal{B}$:

$$\begin{aligned}(S \circ T)(u_j) &= S(T(u_j)) = S\left(\sum_{i=1}^{m} a_{ij} v_i\right) \\ &= \sum_{i=1}^{m} a_{ij} S(v_i) = \sum_{i=1}^{m} a_{ij} (\sum_{k=1}^{l} b_{ki} w_k) \\ &= \sum_{k=1}^{l} \left(\sum_{i=1}^{m} b_{ki} a_{ij}\right) w_k.\end{aligned}$$

Logo

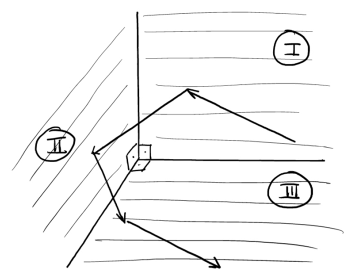
$$[S \circ T]^{\mathcal{B}}_{\mathcal{B}''} = (c_{kj}), \qquad \text{onde} \quad c_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij}.$$

Portanto

$$[S \circ T]^{\mathcal{B}}_{\mathcal{B}''} = (b_{ki}) \cdot (a_{ij}) = BA = [S]^{\mathcal{B}'}_{\mathcal{B}''} \cdot [T]^{\mathcal{B}}_{\mathcal{B}'}.$$

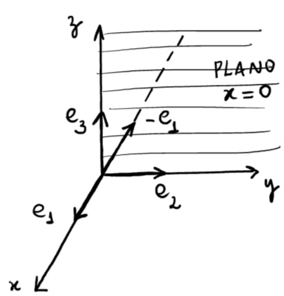
■

Exemplo 3.37. *Considere um conjunto de três espelhos (perfeitos) dispostos de forma que sejam dois-a-dois ortogonais. Vamos mostrar que qualquer feixo de luz que incide sobre o conjunto sairá paralelo à direção de incidência.*



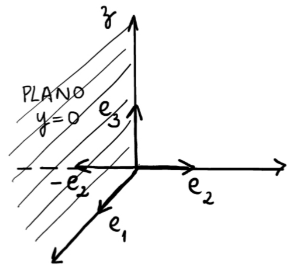
A reflexão final sobre o conjunto é a composição de três reflexões:

Espelho I*: Matriz da reflexão em relação ao plano* $x = 0$



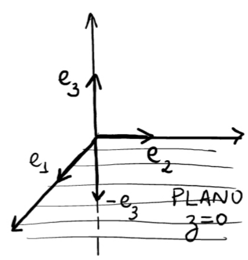
$$M_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Espelho II*: Matriz da reflexão em relação ao plano* $y = 0$



$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Espelho III*: Matriz da reflexão em relação ao plano $z = 0$*



$$M_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Note que as matrizes M_1, M_2 e M_3 comutam. Isso nos diz que a reflexão final sobre o conjunto de três espelhos independe da posição do feixe de luz incidente. E a matriz que representa a reflexão final sobre o conjunto é

$$M = M_3 \cdot M_2 \cdot M_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Assim, se o feixe de luz incidente corresponde ao vetor (a, b, c), então o feixe de luz refletido terá coordenadas $(-a, -b, -c)$, sendo assim paralelo ao feixe de luz incidente.

Corolário 3.38. *Sejam U e V espaços vetoriais de dimensão n sobre K e $\mathcal{B}$ e $\mathcal{B}'$ bases de U e V, respectivamente. Uma transformação linear $T : U \to V$ é um isomorfismo se, se somente se, a matriz $[T]^{\mathcal{B}}_{\mathcal{B}'}$ é inversível. Neste caso,*

$$\boxed{[T^{-1}]^{\mathcal{B}'}_{\mathcal{B}} = ([T]^{\mathcal{B}}_{\mathcal{B}'})^{-1}}$$

Matriz da transformação inversa

Demonstração. Se T é um isomorfismo, então temos a transformação linear inversa $T^{-1} : V \to U$ de forma que

$$T^{-1} \circ T = I_U \quad \text{e} \quad T \circ T^{-1} = I_V$$

é são as identidades I_U em U e I_V em V. Assim, pelo Teorema 3.36 acima e Observação 2.70 sobre matrizes de mudança de base, obtemos

$$[T^{-1} \circ T]^{\mathcal{B}}_{\mathcal{B}} = [I]^{\mathcal{B}}_{\mathcal{B}} \quad \Rightarrow \quad [T^{-1}]^{\mathcal{B}'}_{\mathcal{B}} \cdot [T]^{\mathcal{B}}_{\mathcal{B}'} = I_n$$

e

$$[T \circ T^{-1}]_{\mathcal{B}'}^{\mathcal{B}'} = [I]_{\mathcal{B}'}^{\mathcal{B}'} \quad \Rightarrow \quad [T]_{\mathcal{B}'}^{\mathcal{B}} \cdot [T^{-1}]_{\mathcal{B}}^{\mathcal{B}'} = I_n$$

Segue então que $[T]_{\mathcal{B}'}^{\mathcal{B}}$ é inversível e sua matriz inversa é a matriz da transfomação linear inversa T^{-1}, nas bases $\mathcal{B}'$ e $\mathcal{B}$, ou seja, $[T^{-1}]_{\mathcal{B}}^{\mathcal{B}'} = ([T]_{\mathcal{B}'}^{\mathcal{B}})^{-1}$.

Reciprocamente, se $[T]_{\mathcal{B}'}^{\mathcal{B}}$ é inversível, então $[T]_{\mathcal{B}'}^{\mathcal{B}} \cong I_n$. Portanto $T(u_1), T(u_2), \ldots, T(u_n)$ são L.I., onde $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$. Consequentemente T é isomorfismo.

■

Observação 3.39. *i) Note que o Corolário 3.38 independe das bases $\mathcal{B}$ e $\mathcal{B}'$ fixadas, ou seja, T é um isomorfismo se, e somente se, qualquer matriz de T é inversível. Além disso, se uma matriz de T em bases $\mathcal{B}$ e $\mathcal{B}'$ é inversível, qualquer outra matriz de T também é inversível.*

ii) Considere o caso particular em que $T : U \to U$ é uma transformação linear de U nele mesmo, onde $\dim_K U = n > 0$, e $\mathcal{B}$ e $\mathcal{B}'$ são duas bases de U. E considere as matrizes

$$[T]_{\mathcal{B}}^{\mathcal{B}} \quad e \quad [T]_{\mathcal{B}'}^{\mathcal{B}'}$$

de T na base $\mathcal{B}$ e de T na base $\mathcal{B}'$, respectivamente. Agora, considere a composição $I \circ T \circ I$:

$$\boxed{\begin{array}{ccccccc} \overset{\mathcal{B}}{V} & \xrightarrow{I} & \overset{\mathcal{B}'}{V} & \xrightarrow{T} & \overset{\mathcal{B}'}{V} & \xrightarrow{I} & \overset{\mathcal{B}}{V} \\ v & \longmapsto & v & \longmapsto & T(v) & \longmapsto & T(v) \end{array}}$$

Figura: Diagrama da composta $I \circ T \circ I$

Temos que

$$I \circ T \circ I = T.$$

Então, considerando as bases conforme o diagrama, obtemos

$$[T]_{\mathcal{B}}^{\mathcal{B}} = [I \circ T \circ I]_{\mathcal{B}}^{\mathcal{B}} = [I]_{\mathcal{B}}^{\mathcal{B}'} \cdot [T]_{\mathcal{B}'}^{\mathcal{B}'} \cdot [I]_{\mathcal{B}'}^{\mathcal{B}}. \tag{9}$$

Lembrando que $[I]_{\mathcal{B}}^{\mathcal{B}'}$ é a matriz de mudança de base de $\mathcal{B}'$ para $\mathcal{B}$ e que $[I]_{\mathcal{B}'}^{\mathcal{B}} = ([I]_{\mathcal{B}}^{\mathcal{B}'})^{-1}$, denotando

$$P = [I]_{\mathcal{B}'}^{\mathcal{B}}$$

obtemos de (9) a relação

$$\boxed{[T]_{\mathcal{B}}^{\mathcal{B}} = P^{-1} \cdot [T]_{\mathcal{B}'}^{\mathcal{B}'} \cdot P} \tag{10}$$

Semelhança entre as matrizes de T

*Dizemos então que as matrizes $[T]_{\mathcal{B}}^{\mathcal{B}}$ e $[T]_{\mathcal{B}'}^{\mathcal{B}'}$ são **semelhantes**, pois estão relacionadas através de uma matriz inversível P, como na equação acima.*

Exemplo 3.40. *Considere a transformação linear* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *dada por*

$$T(x, y, z) = (x + y - z, -2x + 4y - z, -4x + 4y + z)$$

e considere a base canônica $\mathcal{C}$ *de* $\mathbb{R}^3$ *e base* $\mathcal{B} = \{(1, 1, 1), (-1, 0, 2), (1, 1, 0)\}$. *A matriz* $[T]_{\mathcal{C}}^{\mathcal{C}}$ *de* T *em relação à base canônica é*

$$[T]_{\mathcal{C}}^{\mathcal{C}} = \begin{pmatrix} 1 & 1 & -1 \\ -2 & 4 & -1 \\ -4 & 4 & 1 \end{pmatrix}.$$

E temos que

$$\begin{aligned} T(1, 1, 1) &= (1, 1, 1) = 1(1, 1, 1) + 0(-1, 0, 2) + 0(1, 1, 0) \\ T(-1, 0, 2) &= (-3, 0, 6) = 0(1, 1, 1) + 3(-1, 0, 2) + 0(1, 1, 0) \\ T(1, 1, 0) &= (2, 2, 0) = 0(1, 1, 1) + 0(-1, 0, 2) + 2(1, 1, 0) \end{aligned}$$

logo a matriz de T *na base* $\mathcal{B}$ *é*

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

As matrizes $[T]_{\mathcal{C}}^{\mathcal{C}}$ *e* $[T]_{\mathcal{B}}^{\mathcal{B}}$ *estão relacionadas por*

$$[T]_{\mathcal{B}}^{\mathcal{B}} = P^{-1} \cdot [T]_{\mathcal{C}}^{\mathcal{C}} \cdot P,$$

onde $P = [I]_{\mathcal{C}}^{\mathcal{B}}$ *é a matriz de mudança de base de* $\mathcal{B}$ *para* $\mathcal{C}$. *Temos que*

$$[I]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix} \quad e \quad [I]_{\mathcal{B}}^{\mathcal{C}} = \left([I]_{\mathcal{C}}^{\mathcal{B}}\right)^{-1} = \begin{pmatrix} 2 & -2 & 1 \\ -1 & 1 & 0 \\ -2 & 3 & -1 \end{pmatrix}.$$

Note que

$$\det[T]_{\mathcal{C}}^{\mathcal{C}} = \det(P \cdot [T]_{\mathcal{B}}^{\mathcal{B}} \cdot P^{-1}) = \det[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{vmatrix} = 6 \neq 0,$$

portanto T *é isomorfismo. Vamos encontrar* T^{-1}. *Temos*

$$[T^{-1}]_{\mathcal{C}}^{\mathcal{C}} = ([T]_{\mathcal{C}}^{\mathcal{C}})^{-1} = \begin{pmatrix} \frac{4}{3} & -\frac{5}{6} & \frac{1}{2} \\ 1 & -\frac{1}{2} & \frac{1}{2} \\ \frac{4}{3} & -\frac{4}{3} & 1 \end{pmatrix},$$

Logo

$$T^{-1}(x, y, z) = \left(\frac{4x}{3} - \frac{5y}{6} + \frac{z}{2}\ ,\ x - \frac{y}{2} + \frac{z}{2}\ ,\ \frac{4x}{3} - \frac{4y}{3} + z\right).$$

Por fim, vamos observar que a relação de semelhança $[T]_{\mathcal{B}}^{\mathcal{B}} = P^{-1} \cdot [T]_{\mathcal{C}}^{\mathcal{C}} \cdot P$, nos dá uma ferramenta interessante. Suponha que queremos calcular a décima potência A^{10} da matriz A dada por

$$A = [T]_{\mathcal{C}}^{\mathcal{C}} = \begin{pmatrix} 1 & 1 & -1 \\ -2 & 4 & -1 \\ -4 & 4 & 1 \end{pmatrix}.$$

Como A é semelhante à matriz diagonal

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

onde $A = PDP^{-1}$, temos que

$$A^{10} = (PDP^{-1})^{10} = \underbrace{(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}_{\textit{10 vezes}} = PD^{10}P^{-1},$$

pois as matrizes P^{-1} e P se cancelam. Assim

$$A^{10} = PD^{10}P^{-1}.$$

Além disso, como D é uma matriz diagonal, podemos ver que

$$D^{10} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{10} = \begin{pmatrix} 1^{10} & 0 & 0 \\ 0 & 3^{10} & 0 \\ 0 & 0 & 2^{10} \end{pmatrix}$$

Assim,

$$A^{10} = PD^{10}P^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1^{10} & 0 & 0 \\ 0 & 3^{10} & 0 \\ 0 & 0 & 2^{10} \end{pmatrix} \begin{pmatrix} 2 & -2 & 1 \\ -1 & 1 & 0 \\ -2 & 3 & -1 \end{pmatrix}$$

A base $\mathcal{B}$ não foi dada por acaso. Veremos na seção seguinte que podemos encontrar, quando existe, uma tal base de forma que a matriz da transformação linear seja uma matriz diagonal nesta base.

4 Formas Canônicas

4.1 Operadores Diagonalizáveis

Definição 4.1. *Seja V um espaço vetorial sobre K. Uma transformação linear*

$$T : V \to V$$

de um espaço vetorial V nele mesmo é também chamada de um ***operador linear*** *de V. Se V tem dimensão finita e $\mathcal{B}$ é uma base de V, a matriz de T na base $\mathcal{B}$ é, por simplicidade, denotada por*

$$[T]_{\mathcal{B}}.$$

Seja $T : V \to V$ um operador linear de V e suponha que exista uma base $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ tal que a matriz $[T]_{\mathcal{B}}$ seja uma matriz diagonal:

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}, \qquad \lambda_i \in K,\ i = 1, 2, \ldots, n.$$

Sabemos que, por definição, a matriz $[T]_{\mathcal{B}}$ tem suas colunas j formadas ao tomarmos cada vetor v_j de $\mathcal{B}$, aplicarmos T e escrevermos suas coordenadas na base $\mathcal{B}$. Com isso, vemos que

$$\begin{aligned} T(v_1) &= \lambda_1 v_1 + 0\ v_2 + \cdots + 0\ v_n = \lambda_1 v_1 \\ T(v_2) &= 0\ v_1 + \lambda_2 v_2 + \cdots + 0\ v_n = \lambda_2 v_2 \\ &\vdots \\ T(v_n) &= 0\ v_1 + 0\ v_2 + \cdots + \lambda_n v_n = \lambda_n v_n \end{aligned}$$

Assim, os vetores v_j da base $\mathcal{B}$ satisfazem

$$T(v_j) = \lambda_j v_j$$

ou seja, a imagem de cada vetor da base $\mathcal{B}$ é um múltiplo escalar dele mesmo.

Objetivo: Determinar se existe uma tal base $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ seja diagonal e, em caso afirmativo, encontrar $\mathcal{B}$.

Como observado acima, uma tal base é formada por vetores $v \in V$ que satisfazem

$$T(v) = \lambda v, \quad \text{para algum escalar } \lambda \in K.$$

Definição 4.2. *Seja $T : V \to V$ um operador linear.*

i) *Um **autovalor de** T é um escalar $\lambda \in K$ tal que existe um vetor não nulo $v \in V$ com $Tv = \lambda v$.*

ii) *Se λ é um autovalor de T, então todo vetor não nulo $v \in V$ tal que $Tv = \lambda v$ é chamado de **autovetor de** T **associado a** λ. Denotamos por $Aut_T(\lambda)$ o subespaço de V gerado por todos os autovetores associados a λ.*

iii) *Suponha que $\dim_K V = n < \infty$. Dizemos que T **é diagonalizável** se existe uma base $\mathcal{B}$ de V tal que a matriz $[T]_{\mathcal{B}}$ é diagonal, o que é equivalente a dizer que existe uma base de V formada por autovetores de T.*

Exemplo 4.3. *Seja $T : V \to V$ um operador linear não injetor. Então existe $v \in V$, $v \neq 0$, tal que $Tv = 0 = 0\ v$. Logo $\lambda = 0$ é um autovalor de T.*

Exemplo 4.4. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a reflexão em relação à reta $y = x$. Então T é a transformação linear dada por $T(x, y) = (y, x)$.*

Queremos encontrar $(x, y) \in \mathbb{R}^2$, não nulo, tal que

$$T(x, y) = \lambda(x, y).$$

Temos que

$$T(x,y) = \lambda(x,y) \Leftrightarrow (y,x) = \lambda(x,y) \Leftrightarrow \left\{ \begin{array}{l} y = \lambda x \\ x = \lambda y \end{array} \right. \Leftrightarrow \left\{ \begin{array}{r} \lambda x - \ \ y = 0 \\ -x + \lambda y = 0 \end{array} \right.$$

Para ter solução (x, y) não nula, o determinante da matriz do sistema linear obtido deve ser igual a zero:

$$\left| \begin{array}{cc} \lambda & -1 \\ -1 & \lambda \end{array} \right| = \lambda^2 - 1 = 0 \Leftrightarrow \lambda = 1 \ \ ou \ \ \lambda = -1.$$

Então para $\lambda = 1$ e para $\lambda = -1$, existe $(x, y) \neq (0, 0)$ tal que $T(x, y) = \lambda(x, y)$, ou seja, $\lambda = 1$ e $\lambda = -1$ são autovalores de T. Vamos encontrar os autovetores associados:

$\underline{\lambda = 1}$*: Do sistema linear acima, para $\lambda = 1$, obtemos $y = x$. Então os vetores (x, x), com $x \neq 0$, são autovetores de T associados ao autovalor $\lambda = 1$.*

$\underline{\lambda = -1}$*: Do sistema linear acima, para $\lambda = -1$, obtemos $y = -x$. Então os vetores $(x, -x)$, com $x \neq 0$, são autovetores de T associados ao autovalor $\lambda = -1$.*

Tomando então o autovetor $(1, 1)$ associado ao autovalor $\lambda = 1$ e o autovetor (1,-1) associado ao autovalor $\lambda = -1$, considere

$$\mathcal{B} = \{(1, 1), (1, -1)\}.$$

Temos que $\mathcal{B}$ é uma base de $\mathbb{R}^2$ formada por autovetores de T. Com isso, temos

$$\begin{aligned} T(1,\ 1) \ &= \ 1\ (1,1) = 1\ (1,1) + 0\ (1,-1) \\ T(1,-1) &= -1\ (1,-1) = 0\ (1,1) - 1\ (1,-1) \end{aligned}$$

e portanto

$$[T]_{\mathcal{B}} = \left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right)$$

é uma matriz diagonal. Logo T é diagonalizável.

Exemplo 4.5. *Nem todo operador linear possui autovetores. Considere a rotação anti-horária R_θ em $\mathbb{R}^2$ de ângulo θ, onde $0 < \theta < \pi$. A matriz de R_θ na base canônica $\mathbb{C}$ é*

$$[R_\theta]_{\mathcal{C}} = \left(\begin{array}{cc} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{array} \right)$$

e assim $R_\theta(x, y) = (x\cos\theta - y\operatorname{sen}\theta, x\operatorname{sen}\theta + y\cos\theta)$.

Queremos encontrar $(x, y) \in \mathbb{R}^2$, *não nulo, tal que*

$$R_\theta(x, y) = \lambda(x, y).$$

Temos que

$$R_\theta(x, y) = \lambda(x, y) \;\Leftrightarrow\; (x\cos\theta - y \operatorname{sen}\theta, x\operatorname{sen}\theta + y\cos\theta) = \lambda(x, y)$$

$$\Leftrightarrow \left\{ \begin{array}{l} \cos\theta\; x - \operatorname{sen}\theta\; y = \lambda x \\ \operatorname{sen}\theta\; x + \cos\theta\; y = \lambda y \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} (\lambda - \cos\theta)\; x + \operatorname{sen}\theta \quad y = 0 \\ -\operatorname{sen}\theta \quad x + (\lambda - \cos\theta)\; y = 0 \end{array} \right.$$

Para ter solução (x, y) *não nula, o determinante da matriz do sistema linear obtido deve ser igual a zero:*

$$\left| \begin{array}{cc} \lambda - \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \lambda - \cos\theta \end{array} \right| = (\lambda - \cos\theta)^2 + \operatorname{sen}^2\theta = \lambda^2 - 2\cos\theta\;\lambda + 1 \neq 0, \quad \textit{para todo} \;\; \lambda \in \mathbb{R},$$

pois $\Delta = 4\cos^2\theta - 4 = 4(\cos^2\theta - 1) < 0$, *para* $0 < \theta < \pi$.

Como o determinante da matriz do sistema é diferente de zero, o sistema só tem a solução nula, logo não existe $(x, y) \neq (0, 0)$ *tal que* $R_\theta(x, y) = \lambda(x, y)$. *Com isso vemos que* R_θ *não possui autovetores.*

[*Geometricamente, podemos ver que não há vetores não nulos de* $\mathbb{R}^2$ *que ao serem rotacionados sejam um múltiplo escalar de si mesmo.*]

Voltemos ao caso geral. Seja $T : V \to V$ um operador linear, onde V é uma espaço vetorial sobre K de dimensão n, e $\mathcal{C}$ uma base qualquer de V. Note que

$$T(v) = \lambda v \;\Leftrightarrow\; \lambda v - T(v) = 0 \;\Leftrightarrow\; (\lambda I - T)(v) = 0 \;\Leftrightarrow\; v \in \operatorname{Ker}(\lambda I - T),$$

onde $I : V \to V$ é a transfomação identidade. Assim,

$$\begin{aligned} \lambda \text{ é autovalor de } T \;\;&\Leftrightarrow\;\; \text{existe } v \neq 0 \text{ tal que } T(v) = \lambda v \\ &\Leftrightarrow\;\; \text{existe } v \neq 0 \text{ tal que } v \in \operatorname{Ker}(\lambda I - T) \\ &\Leftrightarrow\;\; \operatorname{Ker}(\lambda I - T) \neq \{0\} \\ &\Leftrightarrow\;\; \text{a matriz } [\lambda I - T]_\mathcal{C} \text{ não é inversível} \\ &\Leftrightarrow\;\; \det[\lambda I - T]_\mathcal{C} = 0 \end{aligned}$$

Agora, temos que

$$[\lambda I - T]_\mathcal{C} = \lambda[I]_\mathcal{C} - [T]_\mathcal{C} = \lambda I_n - [T]_\mathcal{C},$$

onde I_n é a matriz identidade. Logo, obtemos que

$$\lambda \;\text{ é autovalor de } T \;\Leftrightarrow\; \det(\lambda I_n - [T]_\mathcal{C}) = 0.$$

Definição 4.6. *Seja* V *um espaço vetorial de dimensão finita* n, $T : V \to V$ *um operador linear e* $\mathcal{C}$ *uma base de* V. *Chamamos o polinômio*

$$p_T(x) = \det(xI_n - [T]_\mathcal{C}),$$

de ***polinômio característico de*** T.

Pelo que vimos acima, os autovalores de T são (exatamente) as raízes do polinômio característico $p_T(x)$ de T.

Exemplo 4.7. *Voltando no Exemplo 4.4, temos* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *dada por* $T(x, y) = (y, x)$. *Seja* $\mathcal{C}$ *a base canônica de* $\mathbb{R}^2$. *Então*

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

O polinômio característico de T *é*

$$p_T(x) = \det(xI_n - [T]_{\mathcal{C}}) = \det\left[x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right] = \begin{vmatrix} x & -1 \\ -1 & x \end{vmatrix} = x^2 - 1.$$

E as raízes de $p_T(x)$ *são os autovalores de* T, *logo* $\lambda_1 = 1$ *e* $\lambda_2 = -1$ *são os autovalores de* T.

Exemplo 4.8. *Voltando no Exemplo 4.5, temos* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *dada pela rotação anti-horária de ângulo* θ, *onde* $0 < \theta < \pi$. *Seja* $\mathcal{C}$ *a base canônica de* $\mathbb{R}^2$. *Então*

$$[T]_{\mathcal{C}} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

O polinômio característico de T *é*

$$p_T(x) = \det(xI_n - [T]_{\mathcal{C}}) = \det\left[x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\right]$$

$$= \begin{vmatrix} x - \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & x - \cos\theta \end{vmatrix} = (x - \cos\theta)^2 + \operatorname{sen}^2\theta = x^2 - 2\cos\theta\ x + 1$$

As raízes de $p_T(x)$ *são os autovalores de* T. *Como*

$$p_T(\lambda) = \lambda^2 - 2\cos\theta\ \lambda + 1 \neq 0, \quad \textit{para todo} \quad \lambda \in \mathbb{R},$$

T *não possui autovalores.*

Observação 4.9. **i)** *A definição de* $p_T(x)$ *não depende da base de* $\mathcal{C}$ *de* V *escolhida. De fato, se* $\mathcal{C}$ *e* $\mathcal{B}$ *são bases de* V, *sabemos que as matrizes* $[T]_{\mathcal{C}}$ *e* $[T]_{\mathcal{B}}$ *são semelhantes, ou seja, estão relacionadas por*

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P,$$

onde $P = [I]_{\mathcal{B}}^{\mathcal{C}}$ *é a matriz de mudança de base. Então, obtemos que*

$$\begin{aligned} \det(xI_n - [T]_{\mathcal{C}}) &= \det\left(xI_n - P^{-1}[T]_{\mathcal{B}}P\right) \\ &= \det\left(\ xP^{-1}I_nP - P^{-1}[T]_{\mathcal{B}}P\ \right) \\ &= \det\left(\ P^{-1}\ (xI_n - [T]_{\mathcal{B}})\ P\ \right) \\ &= \det(P^{-1})\ \det(xI_n - [T]_{\mathcal{B}})\ \det(P) \\ &= \frac{1}{\det(P)}\ \det(xI_n - [T]_{\mathcal{B}})\ \det(P) \\ &= \det(xI_n - [T]_{\mathcal{B}}). \end{aligned}$$

Assim $\det(xI_n - [T]_{\mathcal{C}}) = \det(xI_n - [T]_{\mathcal{B}})$ *para quaisquer duas bases* $\mathcal{C}$ *e* $\mathcal{B}$ *de* V.

ii) *Lembre que $Aut_T(\lambda)$ é o subespaço de V gerado pelos autovetores associados ao autovalor λ. Então*

$$v \in Aut_T(\lambda) \;\Leftrightarrow\; v = a_1v_1 + a_2v_2 + \cdots a_rv_r, \quad \textit{onde } a_i \in K \;\; e \;\; Tv_i = \lambda v_i, \quad i = 1, 2\ldots, r.$$

Com isso,

$$\begin{aligned} T(v) &= a_1Tv_1 + a_2Tv_2 + \cdots a_rTv_r \\ &= a_1\lambda v_1 + a_2\lambda v_2 + \cdots a_r\lambda v_r \\ &= \lambda(a_1v_1 + a_2v_2 + \cdots a_rv_r) \\ &= \lambda v \end{aligned}$$

Então

$$v \in Aut_T(\lambda) \;\Leftrightarrow\; Tv = \lambda v \;\Leftrightarrow\; v \in \operatorname{Ker}(\lambda I - T).$$

Portanto, temos que

$$\boxed{\; Aut_\lambda(T) = \; \operatorname{Ker}(\lambda I - T) \;} \tag{11}$$

Auto-espaço associado a λ

Exemplo 4.10. *Considere o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ dado por*

$$T(x, y) = (-3x + 4y, -x + 2y).$$

Seja $\mathcal{C}$ a base canônica de $\mathbb{R}^2$. O polinômio característico de T é $p_T(x) = \det(xI_2 - [T]_\mathcal{C})$. Temos que

$$[T]_\mathcal{C} = \begin{pmatrix} -3 & 4 \\ -1 & 2 \end{pmatrix}$$

e

$$xI_2 - [T]_\mathcal{C} = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} - \begin{pmatrix} -3 & 4 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} x+3 & -4 \\ 1 & x-2 \end{pmatrix}.$$

Então o polinômio característico de T é

$$\begin{aligned} p_T(x) = \det(xI_2 - [T]_\mathcal{C}) &= \begin{vmatrix} x+3 & -4 \\ 1 & x-2 \end{vmatrix} \\ &= (x+3)(x-2) + 4 \\ &= x^2 + x - 2 \\ &= (x+2)(x-1) \end{aligned}$$

Os autovalores de T são as raízes do polinômio característico $p_T(x)$ de T, então os autovalores de T são

$$\lambda_1 = -2 \quad e \quad \lambda_2 = 1.$$

Os autoespaços associados aos autovalores são:

$\underline{\lambda_1 = -2}$: *Temos que $Aut_T(\lambda_1) = \operatorname{Ker}(\lambda_1 I - T)$.*

Vamos calcular o núcleo usando a matriz de $\lambda_1 I - T$:

$$\lambda_1 I_2 - [T]_{\mathcal{C}} = \begin{pmatrix} \lambda_1 + 3 & -4 \\ 1 & \lambda_1 - 2 \end{pmatrix} = \begin{pmatrix} 1 & -4 \\ 1 & -4 \end{pmatrix}$$

Então $Aut_T(\lambda_1) = \mathrm{Ker}(\lambda_1 I - T)$ *é formado pelos vetores* (x, y) *tais que*

$$\begin{pmatrix} 1 & -4 \\ 1 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Temos que

$$\begin{pmatrix} 1 & -4 \\ 1 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x - 4y = 0.$$

Tomando y *como variável livre, temos* $x = 4y$*. Assim*

$$Aut_T(\lambda_1) = \mathrm{Ker}(\lambda_1 I - T) = \{(4y, y) \mid y \in \mathbb{R}\} = [(4, 1)].$$

$\underline{\lambda_2 = 1}$*: Temos que* $Aut_{\lambda_2}(T) = \mathrm{Ker}(\lambda_2 I - T)$*.*

Vamos calcular o núcleo usando a matriz de $\lambda_2 I - T$*:*

$$\lambda_2 I_2 - [T]_{\mathcal{C}} = \begin{pmatrix} \lambda_2 + 3 & -4 \\ 1 & \lambda_2 - 2 \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ 1 & -1 \end{pmatrix}$$

Então $Aut_T(\lambda_2) = \mathrm{Ker}(\lambda_2 I - T)$ *é formado pelos vetores* (x, y) *tais que*

$$\begin{pmatrix} 4 & -4 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Temos que

$$\begin{pmatrix} 4 & -4 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{r} 4x - 4y = 0 \\ x - \ y = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{r} x - \ y = 0 \\ 0 = 0 \end{array} \right.$$

Tomando y *como variável livre, temos* $x = y$*. Assim*

$$Aut_T(\lambda_2) = \mathrm{Ker}(\lambda_2 I - T) = \{(y, y) \mid y \in \mathbb{R}\} = [(1, 1)].$$

Vamos, agora vamos tomar a base $\mathcal{B} = \{(4, 1), (1, 1)\}$ *formada pelos autovetores geradores dos autoespaços de* T*. A matriz de* T *nesta base é então a matriz diagonal*

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

cujos elementos da diagonal são os autovalores associados aos autovetores da base $\mathcal{B}$*.*

Então T *é diagonalizável e a matriz* $[T]_{\mathcal{C}}$ *é semelhante a uma matriz diagonal:*

$$[T]_{\mathcal{B}} = P^{-1}[T]_{\mathcal{C}} P, \quad \textit{onde} \quad P = [I]_{\mathcal{C}}^{\mathcal{B}}.$$

Temos que

$$P = [I]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix},$$

então P é a matriz formada pela base $\mathcal{B}$ de autovetores de T, ou seja, as colunas de P são os autovetores da base $\mathcal{B}$. E

$$P^{-1} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{4}{3} \end{pmatrix},$$

Assim, a relação $[T]_{\mathcal{B}} = P^{-1}[T]_{\mathcal{C}}P$ é então

$$\begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{4}{3} \end{pmatrix} \begin{pmatrix} -3 & 4 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}.$$

Exemplo 4.11. *Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x, y, z) = (3x - 4z, 3y + 5z, -z)$. A matriz de T na base canônica é*

$$A = \begin{pmatrix} 3 & 0 & -4 \\ 0 & 3 & 5 \\ 0 & 0 & -1 \end{pmatrix}.$$

O polinômio característico de T é

$$p_T(x) = \det(xI_3 - A) = \begin{vmatrix} x-3 & 0 & 4 \\ 0 & x-3 & -5 \\ 0 & 0 & x+1 \end{vmatrix} = (x-3)^2(x+1)$$

Então os autovalores distintos de T são $\lambda_1 = 3$ e $\lambda_2 = -1$. Os auto-espaços associados são:

$\underline{\lambda_1 = 3}$*: $Aut_T(\lambda_1) = \mathrm{Ker}(\lambda_1 I - T)$. A matriz de $\lambda_1 I - T$ na base canônica é*

$$\lambda_1 I_3 - A = \begin{pmatrix} \lambda_1 - 3 & 0 & 4 \\ 0 & \lambda_1 - 3 & -5 \\ 0 & 0 & \lambda_1 + 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & -5 \\ 0 & 0 & 4 \end{pmatrix}$$

Então $Aut_T(\lambda_1) = \mathrm{Ker}(\lambda_1 I - T)$ é formado pelos vetores (x, y, z) tais que

$$\begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & -5 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{r} 4z = 0 \\ -5z = 0 \\ 4z = 0 \end{array} \right. \Leftrightarrow z = 0.$$

Assim

$$Aut_T(\lambda_1) = \mathrm{Ker}(\lambda_1 I - T) = \{(x, y, 0) \mid x, y \in \mathbb{R}\} = [(1, 0, 0), (0, 1, 0)].$$

$\underline{\lambda_2 = -1}$*: $Aut_T(\lambda_2) = \mathrm{Ker}(\lambda_2 I - T)$. A matriz de $\lambda_2 I - T$ na base canônica é*

$$\lambda_2 I_3 - A = \begin{pmatrix} \lambda_2 - 3 & 0 & 4 \\ 0 & \lambda_2 - 3 & -5 \\ 0 & 0 & \lambda_2 + 1 \end{pmatrix} = \begin{pmatrix} -4 & 0 & 4 \\ 0 & -4 & -5 \\ 0 & 0 & 0 \end{pmatrix}$$

Então $Aut_T(\lambda_2) = \mathrm{Ker}(\lambda_2 I - T)$ é formado pelos vetores (x, y, z) tais que

$$\begin{pmatrix} -4 & 0 & 4 \\ 0 & -4 & -5 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{r} -4x \quad + 4z = 0 \\ -4y - 5z = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{r} x \quad - z = 0 \\ 4y + 5z = 0 \end{array} \right.$$

Tomando z como variável livre, obtemos $x = z$ e $y = -5z/4$. Assim

$$Aut_T(\lambda_2) = \operatorname{Ker}(\lambda_2 I - T) = \{(z, -5z/4, z) \mid z \in \mathbb{R}\} = [(1, -5/4, 1)].$$

Tomando a base $\mathcal{B} = \{(1,0,0),(0,1,0),(1,-5/4,1)\}$ formada pelos autovetores geradores dos auto-espaços de T, a matriz de T nesta base é a matriz diagonal

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

No exemplo acima temos um operador linear de $\mathbb{R}^3$ que possui apenas dois autovalores distintos, mas que ainda assim é diagonalizável. Em geral, se $T : V \to V$ é um operador linear e se conseguimos uma base $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ de V formada por autovetores de T, então a matriz de T na base $\mathcal{B}$ é

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

onde os λ_i' não são necessariamente distintos. O número de vezes que um autovetor λ aparece na diagonal é igual a $\dim Aut_T(\lambda)$, como veremos a seguir.

Teorema 4.12. *Seja $T : V \to V$ um operador linear, onde V é um espaço vetorial sobre K de dimensão finita e sejam λ_1, $\lambda_2, \ldots, \lambda_r$, $r \geq 1$, os (distintos) autovalores de T. Para cada i, seja $\mathcal{B}_i \subset Aut_T(\lambda_i)$ um subconjunto L.I.. Então*

$$\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r \quad \textit{é L.I..}$$

Para provar o Teorema acima, vamos primeiro provar o seguinte resultado auxiliar.

Lema 4.13. *Seja $T : V \to V$ um operador linear, onde V é um espaço vetorial sobre K de dimensão finita e sejam λ_1, $\lambda_2, \ldots, \lambda_r$, $r \geq 1$, os (distintos) autovalores de T. Se $v_i \in Aut_T(\lambda_i)$, $i = 1, 2, \ldots, r$, são tais que*

$$v_1 + v_2 + \cdots + v_r = 0,$$

então

$$v_1 = v_2 = \cdots = v_r = 0.$$

Demonstração. Vamos provar o Lema por indução sobre r. Para $r = 1$, temos $v_1 = 0$. Seja $r > 1$ e suponha que a afrimação seja verdadeira para m vetores, $\forall m \leq r - 1$. Se

$$v_1 + v_2 + \cdots + v_r = 0, \tag{12}$$

com $v_i \in Aut_T(\lambda_i)$, então $Tv_1 + Tv_2 + \cdots + Tv_r = T0 = 0$, logo

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r = 0 \tag{13}$$

Multiplicando (12) por λ_1 e subtraindo (13), obtemos

$$(\lambda_1 - \lambda_2)v_2 + \cdots + (\lambda_1 - \lambda_r)v_r = 0.$$

Como $v_i \in Aut_T(\lambda_i)$, qualquer múltiplo escalar de v_i também pertende a $Aut_T(\lambda_i)$. Assim $w_i = (\lambda_1 - \lambda_i)v_i \in Aut_T(\lambda_i)$, $i = 2, \ldots, r$, e

$$w_2 + \cdots + w_r = 0.$$

Por hipótese de indução, seque que $w_2 = \cdots = w_r = 0$, de onde obtemos $v_2 = \cdots = v_r = 0$ e, consequentemente, $v_1 = 0$. Logo a afirmação é verdadeira para r.

■

Agora vamos provar o Teorema.

Demonstração [**Teorema 4.12**] Para cada i, seja $\mathcal{B}_i = \{v_{i1}, v_{i2}, \ldots, v_{in_i}\} \subset Aut_T(\lambda_i)$, com $\mathcal{B}_i$ linearmente independente. Para mostrar que $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$ é L.I., suponha que

$$\sum_{j=1}^{n_1} a_{1j}v_{1j} + \sum_{j=1}^{n_2} a_{2j}v_{2j} + \cdots + \sum_{j=1}^{n_r} a_{rj}v_{rj} = 0$$

seja uma combinação linear dos vetores da união, com escalares em K, dando o vetor nulo. Na combinação linear, temos que

$$\sum_{j=1}^{n_1} a_{1j}v_{1j} \in Aut_T(\lambda_1), \quad \sum_{j=1}^{n_2} a_{2j}v_{2j} \in Aut_T(\lambda_2), \ \ldots, \ \sum_{j=1}^{n_r} a_{rj}v_{rj} \in Aut_T(\lambda_r).$$

Assim, do Lema 4.13, segue que

$$\sum_{j=1}^{n_i} a_{ij}v_{ij} = 0, \quad \forall i = 1, 2, \ldots, r.$$

E como cada $\mathcal{B}_i$ é L.I., segue que $a_{ij} = 0$, $\forall i, j$. Logo $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$ é L.I.

■

Corolário 4.14. *Seja $T : V \to V$ um operador linear, onde V é um espaço vetorial de dimensão finita sobre K. Se $\lambda_1, \lambda_2, \ldots, \lambda_r$ são todos os distintos autovalores de T, então T é diagonalizável se, e somente se,*

$$\dim_K V = \sum_{i=1}^{r} \dim_K Aut_T(\lambda_i).$$

Exemplo 4.15. *Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x, y, z) = (3x - 4y - 4z, 3y + 5z, -z)$. A matriz de T na base canônica é*

$$A = \begin{pmatrix} 3 & -4 & -4 \\ 0 & 3 & 5 \\ 0 & 0 & -1 \end{pmatrix}.$$

O polinômio característico de T é

$$p_T(x) = \det(xI_3 - A) = \begin{vmatrix} x-3 & 4 & 4 \\ 0 & x-3 & -5 \\ 0 & 0 & x+1 \end{vmatrix} = (x-3)^2(x+1)$$

Então os autovalores distintos de T são $\lambda_1 = 3$ e $\lambda_2 = -1$. Os autoespaços associados são:

$\underline{\lambda_1 = 3}$: *$Aut_T(\lambda_1) = \mathrm{Ker}(\lambda_1 I - T)$. A matriz de $\lambda_1 I - T$ na base canônica é*

$$\lambda_1 I_3 - A = \begin{pmatrix} 0 & 4 & 4 \\ 0 & 0 & -5 \\ 0 & 0 & 4 \end{pmatrix}$$

Então $Aut_T(\lambda_1) = \mathrm{Ker}(\lambda_1 I - T)$ é formado pelos vetores (x, y, z) tais que

$$\begin{pmatrix} 0 & 4 & 4 \\ 0 & 0 & -5 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 4y + 4z = 0 \\ \quad -5z = 0 \\ \quad\ \ 4z = 0 \end{cases} \Leftrightarrow y = z = 0.$$

Assim

$$Aut_T(\lambda_1) = \mathrm{Ker}(\lambda_1 I - T) = \{(x, 0, 0) \mid x \in \mathbb{R}\} = [(1, 0, 0)]$$

$\underline{\lambda_2 = -1}$: *$Aut_T(\lambda_2) = \mathrm{Ker}(\lambda_2 I - T)$. A matriz de $\lambda_2 I - T$ na base canônica é*

$$\lambda_2 I_3 - A = \begin{pmatrix} -4 & 4 & 4 \\ 0 & -4 & -5 \\ 0 & 0 & 0 \end{pmatrix}$$

Então $Aut_T(\lambda_2) = \mathrm{Ker}(\lambda_2 - T)$ é formado pelos vetores (x, y, z) tais que

$$\begin{pmatrix} -4 & 4 & 4 \\ 0 & -4 & -5 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} -4x + 4y + 4z = 0 \\ \quad -4y - 5z = 0 \end{cases} \Leftrightarrow \begin{cases} x - \ y - \ z = 0 \\ \quad 4y + 5z = 0 \end{cases}$$

Tomando z como variável livre, obtemos $y = -5z/4$ e $x = -z/4$. Assim

$$Aut_T(\lambda_2) = \mathrm{Ker}(\lambda_2 I - T) = \{(-z/4, -5z/4, z) \mid z \in \mathbb{R}\} = [(-1/4, -5/4, 1)]$$

*Temos que $Aut_T(\lambda_1) = [(1, 0, 0)]$ e $Aut_T(\lambda_2) = [(-1/4, -5/4, 1)]$ e estes são todos os auto-espaços de T. Como $\dim Aut_T(\lambda_1) + \dim Aut_T(\lambda_2) = 2 < 3$, T **não é** diagonalizável.*

O próximo passo será relacionar $\dim Aut_T(\lambda)$ com a multiplicidade de λ como raiz do polinômio característico $p_T(\lambda)$.

Definição 4.16. *Seja λ um autovalor de um operador linear $T : V \to V$, onde V é um espaço vetorial de dimensão finita sobre K, e suponha que*

$$p_T(x) = (x - \lambda)^m q(x),$$

com $q(\lambda) \neq 0$, seja o polinômio característico de T.

i) *A* ***multiplicidade algébrica*** *de λ é o inteiro m, que denotamos por $ma(\lambda)$. Assim, $ma(\lambda)$ é o maior inteiro M tal que $(x-\lambda)^m$ divide $p_T(x)$.*

ii) *A* ***multiplicidade geométrica*** *de λ é a dimensão de seu auto-espaço associado $Aut_T(\lambda)$, que denotamos por $mg(\lambda)$. Assim*

$$mg(\lambda) = \dim_K Aut_T(\lambda).$$

Proposição 4.17. *Seja λ um autovalor de $T : V \to V$, onde $\dim_K V = n$. Então*

$$1 \leq mg(\lambda) \leq ma(\lambda).$$

Demonstração. Primeiro, como λ é autovalor, temos que $Aut_T(\lambda) \neq \{0\}$, logo $mg(\lambda) = \dim Aut_T(\lambda) \geq 1$. Agora, sejam

$$W = Aut_T(\lambda) \quad \text{e } s = mg(\lambda) = \dim Aut_T(\lambda).$$

Tome $\mathcal{B}' = \{w_1, w_2, \ldots, w_s\}$ uma base de W e $\mathcal{B} = \{w_1, w_2, \ldots, w_s, w_{s+1}, \ldots, w_n\}$ uma base de V contendo $\mathcal{B}'$. Temos $T(w_i) = \lambda w_i$, para todo $i = 1, 2, \ldots, s$. Então a matriz de T na base $\mathcal{B}$ é da forma

$$[T]_{\mathcal{B}} = \left(\begin{array}{cccc|c} \lambda & 0 & \cdots & 0 & \\ 0 & \lambda & \cdots & 0 & A_{s\times(n-s)} \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & \lambda & \\ \hline 0 & 0 & \cdots & 0 & \\ \vdots & \vdots & \ddots & \vdots & B_{(n-s)\times(n-s)} \\ 0 & 0 & \cdots & 0 & \end{array}\right)_{n\times n}.$$

Assim

$$p_T(x) = \det(xI_n - [T]_{\mathcal{B}}) = (x-\lambda)^s \det(xI_{(n-s)} - B),$$

logo $(x-\lambda)^s$ divide $p_T(x)$. Como $ma(\lambda)$ é o maior inteiro m tal que $(x-\lambda)^m$ divide $p_T(x)$, segue que

$$s \leq ma(\lambda).$$

■

Observação 4.18. *Seja $T : V \to V$ um operador linear, onde $\dim_K V = n$, e suponha que o polinômio característico de T é da forma*

$$p_T(x) = (x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_r)^{n_r},$$

onde $\lambda_1, \lambda_2, \ldots, \lambda_r$ são distintos. Como

$$\text{grau } p_T(x) = \dim V,$$

segue que

$$\dim V = n_1 + n_2 + \cdots + n_r.$$

Como, pela Proposição 4.17,

$$n_i = ma(\lambda_i) \geq mg(\lambda_i) = \dim_K Aut_T(\lambda_i), \quad i = 1, 2, \ldots, r,$$

para que aconteça

$$\dim V = \dim Aut_T(\lambda_1) + \dim Aut_T(\lambda_2) + \cdots + \dim Aut_T(\lambda_r),$$

devemos ter $n_i = \dim Aut_T(\lambda_i)$*, para todo* $i = 1, 2, \ldots, r$*. Portanto, segue que*

$$T \text{ é diagonalizável} \quad \Leftrightarrow \quad ma(\lambda_i) = mg(\lambda_i).$$

Sobre a diagonalização de operadores lineares, temos então o seguinte resultado.

Teorema 4.19. *Seja* $T : V \to V$ *um operador linear, onde* $\dim_K V < \infty$*, e sejam* $\lambda_1, \lambda_2, \ldots, \lambda_r$ *os autovalores distintos de* T*. Então as seguintes afirmações são equivalentes:*

a) T *é diagonalizável;*

b) $p_T(x) = (x - \lambda_1)^{n_1}(x - \lambda_2)^{n_2} \cdots (x - \lambda_r)^{n_r}$*, com* $n_i \geq 1$ *e* $mg(\lambda_i) = ma(\lambda_i)$*, para todo* $i = 1, 2, \ldots, r$*;*

c) $\dim_K V = \sum_{i=1}^{r} \dim_K Aut_T(\lambda_i)$.

4.2 Espaço $L(U,V)$ e o Polinômio Minimal

Sejam U e V espaços vetoriais sobre um corpo K e considere

$$L(U,V) = \{T : U \to V \mid T \text{ é transformação linear}\},$$

o conjunto de todas as transformações lineares de U em V. Note que $L(U,V)$ é um espaço vetorial, com a soma e a multiplicação por escalar de funções: para $T, S \in L(U,V)$ e $k \in K$, $T + S$ e kT são dadas por

$$(T+S)(u) = T(u) + S(u) \quad \text{e} \quad (kT)(u) = kT(u), \quad \forall u \in U.$$

Note que $kT + S$ é uma trasnformação linear:

$$\begin{aligned}
(kT+S)(ru+w) &= kT(ru+w) + S(ru+w) \\
& kuTu + kTw + rSu + Sw \\
& r(kTu + Su) + kTu + Sw \\
& r(kT+S)(u) + (kT+S)(w), \quad \forall u, w \in U, \ r \in K.
\end{aligned}$$

Com isso, vemos que $L(U,V)$ é um subespaço do espaço das funções $\mathcal{F}(U,V)$, logo é um espaço vetorial. Além disso, temos o seguinte.

Teorema 4.20. *Se* $\dim_K U = n$ *e* $\dim_K V = m$, *então* $\dim_K L(U,V) = nm$.

Demonstração. Seja $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ uma base de U e $\mathcal{B}' = \{v_1, v_2, \ldots, v_m\}$ uma base de V. Então, para toda $T \in L(U,V)$, existe uma única matriz $m \times n$ associada a T, a matriz $[T]^{\mathcal{B}}_{\mathcal{B}'}$. Defina então

$$\begin{array}{rccc} \varphi: & L(U,V) & \to & \mathcal{M}_{m\times n}(K) \\ & T & \mapsto & [T]^{\mathcal{B}}_{\mathcal{B}'} \end{array}$$

Vamos mostrar que φ é isomorfismo:

(i) φ é uma transformação linear, pois

$$\varphi(kT+S) = [kT+S]^{\mathcal{B}}_{\mathcal{B}'} = k[T]^{\mathcal{B}}_{\mathcal{B}'} + [S]^{\mathcal{B}}_{\mathcal{B}'} = k\varphi(T) + \varphi(S),$$

para quaisquer $T, S \in L(U,V)$ e $k \in K$.

(ii) φ é injetora, pois

$$\varphi(T) = 0 \in \mathcal{M}_{m\times n}(K) \;\Rightarrow\; [T]^{\mathcal{B}}_{\mathcal{B}'} = (0)_{m\times n} \;\Rightarrow\; T = 0.$$

(iii) φ é sobrejetora. De fato, pela Observação 3.35(ii), dada uma matriz $A \in \mathcal{M}_{m\times n}(K)$, existe $T \in L(U,V)$ tal que $[T]\mathcal{B}_{\mathcal{B}'} = A$.

Logo φ é isomorfismo. Assim $\dim L(U,V) = \dim \mathcal{M}_{m\times n}(K) = mn$.

■

Consideremos, agora, o **caso especial** $L(V,V)$, ou seja, o espaço vetorial dos operadores lineares de V. Note que se $T \in L(V,V)$, então

$$T^n = \underbrace{T \circ T \circ \cdots \circ T}_{n} \in L(V,V),$$

já que a composição de transformações lineares é também transformação linear. Com isso, se $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_m t^m \in \mathcal{P}(K)$, podemos definir o operador linear

$$p(T) : V \to V$$

dado por

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m,$$

onde, para cada $v \in V$,

$$p(T)(v) = a_0 v + a_1 T(v) + a_2 T^2(v) + \cdots + a_m T^m(v).$$

Exemplo 4.21. *Considere* $T : \mathbb{R}^2 \to \mathbb{R}^2$, *onde* $T(x,y) = (x-y, 3x)$, *e seja* $p(t) = t^3 - 2t + 4$. *Então*

$$p(T) = T^3 - 2T + 4$$

e

$$p(T)(x,y) = T^3(x,y) - 2T(x,y) + 4(x,y).$$

Temos

$$T^3(x,y) = T^2(T(x,y)) = T^2(x-y,3x) = T(x-y-3x,3(x-y)) =$$
$$= T(-2x-y,3x-3y) = (-2x-y-3x+3y,-6x-3y) = (-5x+2y,-6x-3y),$$

assim, $p(T)(x,y)$ *é dado por*

$$p(T)(x,y) = (-5x+2y,-6x-3y) - 2(x-y,3x) + 4(x,y) = (-3x+4y,-8x+y)$$

Agora, voltando ao caso especial $L(V,V)$, pelo Teorema 4.20, se $\dim_K V = n$, então $\dim L(V,V) = n^2$. Logo existe $m \geq 1$ tal que $T^0,\ T,\ T^2,\dots,T^{m-1}$ é L.I. em $L(V,V)$, mas $T^0,\ T,\ T^2,\dots,T^{m-1},T^m$ é L.D.. Assim existem $a_0,a_1,a_2,\dots,a_{m-1} \in K$ tais que

$$T^m = a_0 I + a_1 T + a_2 T^2 + \cdots + a_{m-1}T^{m-1},$$

ou seja,

$$T^m(v) = a_0 v + a_1 T(v) + a_2 T^2(v) + \cdots + a_{m-1}T^{m-1}(v), \quad \forall v \in V.$$

Considere, então, o polinômio

$$m_T(x) = x^m - \sum_{i=0}^{m-1} a_i x^i,$$

onde $a_0,a_1,a_2,\dots,a_{m-1}$ são os escalares dados acima. Então

$$m_T(T)(v) = 0, \quad \forall v \in V,$$

e portanto

$$m_T(T) = 0 \ \text{ é o operador nulo.}$$

Além disso, por construção, m é o menor inteiro positivo tal que $T^0,\ T,\ T^2,\dots,T^{m-1},T^m$ é L.D., o que implica no seguinte.

Afirmação: *Se* $p(x) \in \mathcal{P}(K)$ *é tal que* $p(T) = 0$*, então* $p(x)$ *é um múltiplo de* $m_T(x)$*.*

De fato, dividindo $p(x)$ por $m_T(x)$, temos que

$$p(x) = q(x)m_T(x) + r(x),$$

onde $q(x), r(x) \in \mathcal{P}(K)$ e $r(x) = 0$ ou $\text{grau}(r(x)) < \text{grau}(m_T(x))$. Suponha que $r(x) \neq 0$. Então $r(x) = b_0 + b_1 x + \cdots + b_s x^s$, com $b_s \neq 0$ e $s = \text{grau}(r(x)) < \text{grau}(m_T(x)) = m$. Como por hipótese $p(T) = 0$, obtemos

$$0 = p(T) = q(T) \circ m_T(T) + r(T) = r(T)$$

pois $m_T(T) = 0$ é nulo. Assim

$$r(T) = b_0 I + b_1 T + \cdots + b_s T^s = 0.$$

Mas, como $b_s \neq 0$, segue que $\{I, T, T^2, \dots, T^s\}$ é L.D., com $s < m$, o que contradiz a definição de m. Portanto $r(x) = 0$ e $p(x)$ é um múltiplo de $m_T(x)$, o que prova a afirmação.

Definição 4.22. *O **polinômio minimal** de um operador linear $T \in L(V,V)$ é o polinômio mônico $m_T(x)$ de menor grau tal que $m_T(T)(v) = 0$, $\forall v \in V$.*

Observação 4.23. *Vimos acima que se $p(x)$ é um polinômio tal que $p(T)(v) = 0$, $\forall v \in V$, então além de $\text{grau}(p(x)) > \text{grau}(m_T(x))$, temos algo mais forte, temos que $m_T(x)$ divide $p(x)$.*

Teorema 4.24 (Cayley-Hamilton). *Seja V um espaço vetorial de dimensão finita n sobre K, $T \in L(V,V)$ e $p_T(x)$ o polinômio característico de T. Então*

$$p_T(T) = 0.$$

Demonstração. Seja $\mathcal{B}$ uma base de V e $A = [T]_{\mathcal{B}}$. Considere $A' = xI_n - A$. Com isso,

$$p_T(x) = \det A'.$$

Agora, seja

$$B = \text{Adj}(A') = (b_{ij})$$

, a matriz adjunta de A'. Então as entradas b_{ij} de B são os cofatores da matriz $A' = xI_n - A$, portanto representam polinômios de graus no máximo $n-1$. Vamos escrever tais polinômios b_{ij} como

$$b_{ij} = b_{ij}^{(0)} + b_{ij}^{(1)}x + \cdots + b_{ij}^{(n-1)}x^{n-1},$$

onde $b_{ij}^{(k)} \in K$, $k = 1, 2, \ldots, n-1$. Assim

$$B = \begin{pmatrix} b_{11}^{(0)} + b_{11}^{(1)}x + \cdots + b_{11}^{(n-1)}x^{n-1} & \cdots & b_{1n}^{(0)} + b_{1n}^{(1)}x + \cdots + b_{1n}^{(n-1)}x^{n-1} \\ \vdots & & \vdots \\ b_{n1}^{(0)} + b_{n1}^{(1)}x + \cdots + b_{n1}^{(n-1)}x^{n-1} & \cdots & b_{nn}^{(0)} + b_{nn}^{(1)}x + \cdots + b_{nn}^{(n-1)}x^{n-1} \end{pmatrix},$$

que podemos escrever como

$$B = B^{(0)} + B^{(1)}x + \cdots + B^{(n-1)}x^{n-1}, \quad \text{onde} \quad B^{(k)} = \begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} & \cdots & b_{1n}^{(k)} \\ \vdots & \vdots & & \vdots \\ b_{n1}^{(k)} & b_{n2}^{(k)} & \cdots & b_{nn}^{(k)} \end{pmatrix}.$$

Lembrando que $B = \text{Adj}(A')$, temos que

$$B.A' = \text{Adj}(A').A' = \det A' \cdot I_n = p_T(x) \cdot I_n.$$

Escrevendo $p_T(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$, a igualdade $B \cdot A' = p_T(x) \cdot I_n$ acima, é reescrita como

$$(B^{(0)} + B^{(1)}x + \cdots + B^{(n-1)}x^{n-1})(xI_n - A) = (a_0 + a_1x + a_2x^2 + \cdots + a_nx^n)I_n.$$

Assim

$$(B^{(0)} + B^{(1)}x + \cdots + B^{(n-1)}x^{n-1})(xI_n - A) = a_0I_n + a_1I_nx + a_2x^2 + \cdots + a_nI_nx^n.$$

Comparando os coeficientes dos polinômios na igualdade acima, vemos que

$$\begin{array}{lcll}
a_0I_n & = & -B^{(0)}A & \qquad (\times I_n) \\
a_1I_n & = & B^{(0)} - B^{(1)}A & \qquad (\times A) \\
a_2I_n & = & B^{(1)} - B^{(2)}A & \qquad (\times A^2) \\
 & \vdots & & \\
a_{n-1}I_n & = & B^{(n-2)} - B^{(n-1)}A & \qquad (\times A^{n-1}) \\
a_nI_n & = & B^{(n-1)} & \qquad (\times A^n)
\end{array}$$

Então

$$a_0I_n + a_1A + a_2A^2 + \cdots + a_{n-1}A^{n-1} + a_nA^n =$$

$$= -B^{(0)}A + (B^{(0)} - B^{(1)}A)A + (B^{(1)} - B^{(2)}A)A^2 + \cdots + (B^{(n-2)} - B^{(n-1)}A)A^{n-1} + B^{(n-1)}A^n =$$

$$= 0.$$

Assim, como $A = [T]_{\mathcal{B}}$, temos que

$$p_T(T) = a_0I + a_1T + a_2T^2 + \cdots + a_{n-1}T^{n-1} + a_nT^n = 0.$$

■

Corolário 4.25. *Seja $T \in L(V,V)$, onde $\dim_K V < \infty$. Então o polinômio minimal $m_T(x)$ de T divide o polinômio característico $p_T(x)$ de T.*

Demonstração. Como, pelo Teorema de Cayley-Hamilton, $p_T(x)$ é um polinômio tal que $p_T(T) = 0$, pela Observação 4.23, temos que o polinômio minimal $m_T(x)$ de T divide o polinômio característico $p_T(x)$ de T.

■

Exemplo 4.26. *Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$, tal que a matriz na base canônica é*

$$A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}.$$

Vamos encontrar o polinômio minimal $m_T(x)$ de T. O polinômio característico de T é

$$\begin{aligned}
p_T(x) &= \det(xI_3 - A) = \begin{vmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{vmatrix} \\
&= (x-5)(x-4)(x+4) + 36 + 36 + 18(x-4) + 12(x-5) - 6(x+4) \\
&= x^3 - 5x^2 + 8x - 4 = (x-1)(x-2)^2
\end{aligned}$$

Sabemos que $m_T(x)$ divide $p_T(x)$, então temos as seguintes possibilidades para o minimal:

$$(x-1),\ (x-2),\ (x-1)(x-2),\ (x-2)^2,\ (x-1)(x-2)^2.$$

Vamos testar:

$$A-1I=\begin{pmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{pmatrix} \neq 0$$

$$A-2I=\begin{pmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{pmatrix} \neq 0$$

$$(A-1I)(A-2I)=\begin{pmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{pmatrix}\begin{pmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{pmatrix} = 0$$

Portanto, obtemos que $m_T(x)=(x-1)(x-2)$.

Como $m_T(x)$ divide $p_T(x)$, temos que toda raiz de $m_T(x)$ é também raiz de $p_T(x)$. Veremos a seguir que vale a recíproca, ou seja, que toda raíz de $p_T(x)$ é também raiz de $m_T(x)$.

Proposição 4.27. *Seja V um espaço vetorial de dimensão finita $n \geq 1$ e $T \in L(V,V)$. Então os polinômios característico e minimal de T possuem as mesmas raízes, a menos de multiplicidade.*

Demonstração. Já observamos que toda raiz de $m_T(x)$ é também raiz de $p_T(x)$. Reciprocamente, seja $\lambda \in K$ uma raiz do polinômio característico $p_T(x)$. Então λ é um autovalor de T e assim existe $v \in V$, com $v \neq 0$, tal que $Tv=\lambda v$. Note que $T^i(v)=\lambda^i v$, $\forall i \geq 1$. Logo, se

$$m_T(x)=\sum_{i=0}^{m} a_i x^i,$$

temos que

$$0=m_T(T)(v)=\sum_{i=0}^{m} a_i T^i(v)=\left(\sum_{i=0}^{m} a_i \lambda^i\right) v.$$

Como $v \neq 0$, segue que

$$\sum_{i=0}^{m} a_i \lambda^i = 0.$$

Portanto λ é uma raiz de $m_T(x)$.

■

Observação 4.28. *No exemplo anterior, os possíveis candidatos par $m_T(x)$ eram então, na verdade,*

$$(x-1)(x-2) \quad e \quad (x-1)(x-2)^2.$$

Além disso, o minimal é o de menor grau que anula T. Como $(x-1)(x-2)^2$ é o polinômio característico de T, sabemos de antemão que ele anula T. Então bastava saber se T era anulado, antes, por $(x-1)(x-2)$.

4.3 Espaços Vetoriais T-cíclicos

Seja V um espaço vetorial de dimensão finita sobre K e $T \in L(V,V)$. Vimos na seção anterior que

i) $m_T(x)$ divide $p_T(x)$

ii) as raízes de $m_T(x)$ e $p_T(x)$ são as mesmas a menos de multiplicidade.

Agora, nesta seção, vamos estudar quando $m_T(x) = p_T(x)$ acontece.

Considere $T \in L(V,V)$, onde $\dim_K V = n > 0$, e seja $v \in V$, com $v \neq 0$. Como $\dim V < \infty$, existe $l \geq 0$ tal que $\{v, Tv, \ldots, T^l v\}$ é L.I., mas $\{v, Tv, \ldots, T^{l+1}v\}$ é L.D.. Então

$$T^{l+1}v = a_0 v + a_1 Tv + \cdots + a_l T^l v, \quad \text{para certos} \;\; a_i \in K.$$

Assim

$$(T^{l+1} - \sum_{i=0}^{l} a_i T^i)(v) = 0.$$

Denotando $m_{T,v}(x) = x^{l+1} - \sum_{i=0}^{l} a_i x^i$, temos $m_{T,v}(T)(v) = 0$. E temos que

- $m_{T,v}(x)$ é o polinômio de menor grau tal que $m_{T,v}(T)(v) = 0$;
- $m_{T,v}(x)$ divide $m_T(x)$. De fato, sabemos que $m_T(T)(w) = 0$, $\forall w \in V$, então em particular $m_T(T)(v) = 0$. Com isso, como $m_{T,v}(x)$ é o polinômio de menor grau tal que $m_{T,v}(T)(v) = 0$, concluimos que $m_{T,v}(x)$ divide $m_T(x)$ fazendo um argumento análogo ao da Observação 4.23.

Seja $\mathcal{B} = \{v, Tv, \ldots, T^l v\}$ e $C_T(v) = [v, Tv, \ldots, T^l v]$ o subespaço determinado por $\mathcal{B}$. Note que $\dim_K C_T(v) = 1$ se, e somente se, v é um autovetor de T.

Definição 4.29. *$v \in V$ é um **vetor T-cíclico** de $V = C_T(v)$. E se V possui um vetor T-cíclico, dizemos que V é um **espaço T-cíclico**.*

Exemplo 4.30. *Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$, dado por $T(x,y,z) = (0,x,y)$. Temos que $T(1,0,0) = (0,1,0)$, $T^2(1,0,0) = (0,0,1)$ e $T^3(1,0,0) = (0,0,0)$. Logo $C_T(1,0,0) = \mathbb{R}^3$ e tomando $\mathcal{B} = \{(1,0,0), T(1,0,0), T^2(1,0,0)\}$, temos que*

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

Em geral, seja $V = C_T(v)$ um espaço T-cíclico de dimensão n. Seja $m_{t,v}(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ e considere $\mathcal{B} = \{v, Tv, \ldots, T^{n-1}v\}$. Então a matriz de T na base m$\mathcal{B}$ é

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 & \cdots & -a_0 \\ 1 & 0 & 0 & \cdots & -a_1 \\ 0 & 1 & 0 & \cdots & -a_2 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -a_{n-1} \end{pmatrix}.$$

Além disso, como $\dim V = n$, temos que grau$(p_T(x)) = n$, $m_{T,v}(x)$ divide $m_T(x)$ e $m_T(x)$ divide $p_T(x)$. Então

$$n = \text{grau}(m_{T,v}(x)) \leq \text{grau}(m_T(x)) \leq \text{grau}(p_T(x)) = n,$$

logo grau$(m_T(x)) = n$. Como $m_T(x)$ divide $p_T(x)$, segue que $m_T(x) = p_T(x)$. Portanto

$$\text{se } V \text{ espaço } T\text{-cíclico de dimensão } n, \text{ então } m_T(x) = p_T(x).$$

A recíproca também é verdadeira. A demonstração fica como exercício de leitura (Livro Principal [1], Teorema 5.4.6, pg 156).

4.4 Subespaços T-invariantes

Definição 4.31. *Seja $T : V \to V$ um operador linear, onde V é um espaço vetorial sobre um corpo K. Um subespaço W de V é dito um **subespaço T-invariante** de V quando $T(W) \subseteq W$, ou seja, quando $Tw \in W$, $\forall w \in W$.*

Exemplo 4.32. *Seja $\lambda \in K$ um autovalor de $T \in L(V, V)$ e considere*

$$W = Aut_T(\lambda) = \{v \in V \mid Tv = \lambda v\}.$$

Temos que, $\forall v \in W$, $Tv = \lambda v \in W$. Logo cada auto-espaço W de V é subespaço T-invariante.

Exemplo 4.33. *Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$, dado por $T(x, y, z) = (0, x, y)$.*

Tome $W = [(1, 0, 0), (0, 1, 0)]$. Então $T(W) = [T(1, 0, 0), T(0, 1, 0)] = [(0, 1, 0), (0, 0, 1)] \not\subseteq W$, logo W não é T-invariante.

Agora tome $W' = [(0, 1, 0), (0, 0, 1)]$. Então $T(W') = [(0, 0, 1)] \subseteq W'$, logo W' é T-invariante.

Observação 4.34. *Seja $T : V \to V$ um operador linear. Listamos a seguir uma série de observações importantes a respeito de subespaços T-invariantes de V.*

i) $\text{Ker}T$ *e* $\text{Im}T$ *são subespaços T-invariantes.*

   De fato, $\forall v \in \text{Ker}T$, $Tv = 0 \in \text{Ker}T$. E, $\forall v \in \text{Im}T$, $Tv \in \text{Im}T$.

ii) *Se λ é um autovalor de T, então $Aut_T(\lambda)$ é T-invariante.*

iii) *Se W é subespaço T invariante, como $T(W) \subseteq W$, temos que $T|_W : W \to W$ é um operador linear de W.*

iv) *Seja $n = \dim_K V \geq 1$ e W um subespaço T-invariante de V, com $\dim_K W = m$ e $1 \leq m < n$. Sejam $\mathcal{B}'$ uma base de W e $\mathcal{B}$ uma base de V contendo $\mathcal{B}'$. Já observamos que $T' = T|_W : W \to W$ é um operador linear de W. Observe também que como $T'(w) = T(w) \in W$, $\forall w \in W$, a matriz de T na base $\mathcal{B}$ é do tipo*

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T']_{\mathcal{B}'} & A_{m\times(n-m)} \\ & \\ 0 & B_{(n-m)\times(n-m)} \end{pmatrix}.$$

v) *Suponha que V pode ser escrito como uma soma direta*

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$$

de subespaços T-invariantes W_i, $i = 1, 2, \ldots, r$. Então, se $\mathcal{B}_i$ é uma base de W_i, $i = 1, 2, \ldots, r$, temos que

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$$

é base de V. Também temos que a restição $T_i = T|_{W_i} : W_i \to W_i$ é um operador linear dde W_i, $i = 1, 2, \ldots, r$ e, com isso, a matriz de T na base $\mathcal{B}$ é da forma

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T_1]_{\mathcal{B}_1} & 0 & \cdots & 0 \\ 0 & [T_2]_{\mathcal{B}_2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & [T_r]_{\mathcal{B}_r} \end{pmatrix},$$

onde $[T_i]_{\mathcal{B}_i}$ é a matriz de $T_i = T|_{W_i}$ na base $\mathcal{B}_i$ e os zeros são matrizes nulas com as ordens correspondentes. Assim $[T]_{\mathcal{B}}$ é uma matriz "diagonal"em blocos.

Nota também que podemos escrever T como

$$T = T_1 \oplus T_2 \oplus \cdots \oplus T_r,$$

onde se $v = v_1 + v_2 + \cdots + v_r$ é a (única) expressão de $v \in V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$, então

$$Tv = (T_1 \oplus T_2 \oplus \cdots \oplus T_r)(v_1 + v_2 + \cdots + v_r) := T_2 v_1 + T_2 v_2 + \cdots + T_r v_r.$$

*Dizemos, neste caso, que T é a **soma direta dos operadores** T_1, $T_2, \ldots,$ T_r.*

4.5 Operadores Nilpotentes

Definição 4.35. *$T \in L(V, V)$ é dito **nilpotente** quando existe $m > 0$ tal que $T^m = 0$. O **índice de nilpotência** de T é o menor tal m.*

Exemplo 4.36. *O operador derivada $D : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_n(\mathbb{R})$ é nilpotente de índice $n + 1$.*

Exemplo 4.37. $A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} \in \mathcal{M}_3(K)$ *é nilpotente, pois* $A^3 = 0$.

Observação 4.38. i) *Se* $T \in L(V, V)$ *é nilpotente, então* $\mathrm{Ker}T \neq \{0\}$.

De fato, seja m *o índice de nilpotência de* T. *Então* $T^m = 0$ *e* $T^{m-1} \neq 0$. *Assim, existe* $v \in V$ *tal que* $T^{m-1}(v) \neq 0$. *Mas* $0 = T^m(v) = T(T^{m-1}(v))$, *pois* $T^m = 0$. *Logo* $T^{m-1}(v) \in \mathrm{Ker}T$ *e, portanto,* $\mathrm{Ker}T \neq \{0\}$.

ii) *Nenhum operador linear inversível é nilpotente.*

De fato, todo operador linear T *inversível tem núcleo nulo. Como* $\mathrm{Ker}T = \{0\}$, *pelo item i),* T *não é nilpotente.*

Teorema 4.39. *Seja* $T \in L(V, V)$, *onde* $\dim_K V < \infty$. *Então* T *é uma soma direta de um operador linear nilpotente e um operador linear inversível. Além disso, tal decomposição é única.*

Lembre que a definição de decomposição de um operador $T \in L(V, V)$ como soma direta é, na verdade, uma decomposição do espaço V como soma direta de subespaços T-invariantes. Assim, o teorema diz que $V = W_1 \oplus W_2$, onde W_1 e W_2 são subespaços T-invariantes e $T|_{W_1}$ é nilpotente e $T|_{W_2}$ é inversível.

Demonstração. (*Existência da decomposição*) Note que

$$v \in \mathrm{Ker}T \Rightarrow Tv = 0 \Rightarrow T^2 v = T(T(v)) = T(0) = 0 \Rightarrow v \in \mathrm{Ker}T^2.$$

Assim $\mathrm{Ker}T \leq \mathrm{Ker}T^2$. Em geral, temos a seguinte cadeia de subespaços de T:

$$\mathrm{Ker}T \leq \mathrm{Ker}T^2 \leq \mathrm{Ker}T^3 \leq \cdots \leq \mathrm{Ker}T^l \leq \mathrm{Ker}T^{l+1} \leq \cdots \leq V.$$

Como $\dim V < \infty$, existe $m > 0$ tal que $\mathrm{Ker}T^m = \mathrm{Ker}T^{m+i}$, $\forall i \geq 0$. Vamos tomar o menor m com essa propriedade, ou seja, tal que

$$\mathrm{Ker}T^{m-1} < \mathrm{Ker}T^m \text{ (inclusão própria)}$$

$$\text{e } \mathrm{Ker}T^m = \mathrm{Ker}T^{m+i}, \quad \forall i \geq 0.$$

Com isso, defina $W_1 := \mathrm{Ker}T^m$ e $W_2 := \mathrm{Im}T^m$. Vamos mostrar que W_1 e W_2 assim definidos induzem a decomposição desejada:

i) $W_1 \cap W_2 = \{0\}$

De fato,

$$v \in W_1 \cap W_2 \Leftrightarrow \left\{ \begin{array}{l} v \in W_1 = \mathrm{Ker}T^m \\ \quad e \\ v \in W_2 = \mathrm{Im}T^m \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} T^m(v) = 0 \\ \quad e \\ v = T^m(w), \quad \text{para algum } w \in V \end{array} \right.$$

Então
$$0 = T^m(v) = T^m(T^m(w)) = T^{2m}(w) \Rightarrow w \in \text{Ker}T^{2m}.$$
Pela escolha de m, temos $\text{Ker}T^{2m} = \text{Ker}T^m$, logo
$$\Rightarrow w \in \text{Ker}T^m \Rightarrow T^m(w) = 0 \Rightarrow v = 0.$$

ii) $V = W_1 + W_2$

De fato, temos que
$$\begin{aligned}\dim(W_1 + W_2) &= \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)\\ &= \dim W_1 + \dim W_2 + 0\\ &= \dim \text{Ker}T^m + \dim \text{Im}T^m\\ &= \dim V,\end{aligned}$$
pois $T^m : V \to V$ é um operador linear de V. Logo $W_1 + W_2 = V$.

De *i)* e *ii)* obtemos que $V = W_1 \oplus W_2$.

iii) W_1 e W_2 são T-invariantes.

De fato,

$w_1 \in W_1 = \text{Ker}T^m \;\Rightarrow\; T^m(T(w_1)) = T(T^m(w_1)) = T(0) = 0 \;\Rightarrow\; T(w_1) \in \text{Ker}T^m = W_1;$

e

$w_2 \in W_2 = \text{Im}T^m \;\Rightarrow\; w_2 = T^m(v),\ v \in V \;\Rightarrow\; T(w_2) = T(T^m(v)) = T^m(T(v)) \in \text{Im}T^m = W_2.$

Como W_1 e W_2 são T-invariantes, podemos considerar as restrições $T|_{W_1}$ e $T|_{W_2}$. Vamos, agora, mostar que $T|_{W_1}$ é nilpotente e que $T|_{W_2}$ é inversível, o que conclui a prova da existência da decomposição. De fato,

$(T|_{W_1})^m(w_1) = T^m(w_1) = 0$, $\forall w_1 \in W_1 = \text{Ker}T^m$, então $(T|_{W_1})^m \equiv 0$ e portanto $T|_{W_1}$ é nilpotente;

$$\begin{aligned} w_2 \in \text{Ker}T|_{W_2} \;&\Rightarrow\; T(w_2) = 0 \;\text{ e }\; w_2 = T^m(v), \;\text{ para algum } v \in V\\ &\Rightarrow\; 0 = T(w_2) = T(T^m(v)) = T^{m+1}(v)\\ &\Rightarrow\; v \in \text{Ker}T^{m+1} = \text{Ker}T^m, \;\text{ pela escolha de } m \qquad .\\ &\Rightarrow\; T^m(v) = 0\\ &\Rightarrow\; w_2 = 0.\end{aligned}$$
Logo $\text{Ker}T|_{W_2} = \{0\}$ e assim $T|_{W_2}$ é inversível.

(*Unicidade da decomposição*) Considere a decomposição $V = W_1 \oplus W_2$ construída acima e suponha que $V = U_1 \oplus U_2$, onde U_1 e U_2 são subespaços T-invariantes, $T_1 := T|_{U_1}$ é nilpotente de índice m' e $T_2 := T|_{U_2}$ é inversível. Vamos mostrar que
$$U_1 = W_1 = \text{Ker}T^m \quad \text{e} \quad U_2 = W_2 = \text{Im}T^m.$$

Tome $k = \max\{m, m'\}$. Então

$W_1 \subseteq U_1$: se $w_1 \in W_1$, como $W_1 \leq V = U_1 \oplus U_2$, podemos escrever $w_1 = u_1 + u_2$ com $u_i \in U_i$, $i = 1, 2$. Agora, como $W_1 = \mathrm{Ker} T^m \leq \mathrm{Ker} T^k$, temos $0 = T^k(w_1) = T^k(u_1) + T^k(u_2) = T^k(u_2)$, pois $(T|_{U_1})^{m'} = 0$ e $k \geq m'$. Então temos $T^k(u_2) = 0$. Mas $T|_{U_2}$ é inversível, logo $u_2 = 0$ e assim $w_1 = u_1 \in U_1$.

Analogamente se mostra que $W_1 \subseteq U_1$. Portanto $U_1 = W_1$.

$W_2 \subseteq U_2$: se $w_2 \in W_2$, então $w_2 = T^m(v)$ para algum $v \in V$. Como $v \in V$, podemos escrever $v = u_1 + u_2$ com $u_i \in U_i$, $i = 1, 2$. Mas, como $U_1 = \mathrm{Ker} T^m$, segue que $w_2 = T^m(v) = T^m(u_1) + T^m(u_2) = T^m(u_2)$. Agora, como U_2 é T-invariante, obtemos que $w_2 = T^m(u_2) \in U_2$.

Reciprocamente, se $u_2 \in U_2$, então u_2 pode ser escrito como $u_2 = w_1 + w_2$, com $w_i \in W_i$. Assim, como $W_2 \subseteq U_2$ e $W_1 = U_1$, segue que $w_1 = u_2 - w_2 \in U_1 \cap U_2 = \{0\}$. Logo $u_2 = w_2 \in W_2$. Portanto $U_2 \subseteq W_2$ e, assim, $U_2 = W_2$.

■

O decomposição dada pelo Teorema 4.39 é válida para qualquer operador linear de um espaço vetorial de dimensão finita. Em relação à parte nilpotente, temos o seguinte.

Proposição 4.40. *Seja $T \in L(V, V)$ nilpotente de índice $m \geq 1$, onde $\dim_K V < \infty$. Se $v \in V$ é tal que $T^{m-1}(v) \neq 0$, então*

(a) *$\{v, Tv, T^2v, \ldots, T^{m-1}v\}$ é L.I.;*

(b) *considerando $U = [v, Tv, T^2v, \ldots, T^{m-1}v]$, existe um subespaço T-invariante W de V tal que $V = U \oplus W$.*

Demonstração. (a) Suponha que $a_0v + a_1Tv + a_2T^2v + \cdots + a_{m-1}T^{m-1} = 0$, com $a_i \in K$, $i = 1, 2, \ldots, m-1$. Queremos mostrar que $a_0 = a_1 = a_2 = \cdots = a_{m-1} = 0$. Suponha que não e tome l o menor índice i tal que $a_i \neq 0$. Então $a_0 = a_1 = \cdots = a_{l-1} = 0$ e $a_l \neq 0$, assim

$$T^l v = \sum_{i=l+1}^{m-1} -\frac{a_i}{a_l}\, T^i v.$$

Então

$$T^{m-1}v = (T^{m-1-l} \circ T^l)(v) = T^{m-(l+1)} \left(\sum_{i=l+1}^{m-1} -\frac{a_i}{a_l}\, T^i v \right) = \sum_{i=l+1}^{m-1} -\frac{a_i}{a_l}\, T^{m+i-(l+1)}(v)$$

Como $m+i-(l+1) \geq m$, $\forall i = l+1, \ldots, m-1$, temos que $T^{m+i-(l+1)}(v)$, $\forall i = l+1, \ldots, m-1$. Assim $T^{m-1}v = 0$, o que contradiz a hipótese. Portanto $a_0 = a_1 = a_2 = \cdots = a_{m-1} = 0$ e $\{v, Tv, T^2v, \ldots, T^{m-1}v\}$ é L.I..

(b) Vamos provar a afirmação por indução sobre o índice de nilpotência $m \geq 1$ de T.

Se $m = 1$, então $T = T^1 = 0$. Assim $U = [v]$ e qualquer subespaço de V é T-invariante. Então $\{v\}$ pode ser completado até formar uma base $\{v, w_1, w_2, \ldots, w_r\}$ de V, encontrando $W = [w_1, w_2, \ldots, w_r]$ tal que $V = U \oplus W$.

Suponha $m > 1$ e que o resultado seja verdadeiro para todos os operadores nilpotentes de índice menor do que m. Temos que $\text{Im}T$ é T-invariante, logo a restrição $T|_{\text{Im}T}$ é nilpotente de índice $m-1$, pois $T^{m-1}(T\tilde{v}) = 0$, $\forall T\tilde{v} \in \text{Im}T$, e $T^{m-2}(Tv) \neq 0$.

Além disso, $U \cap \text{Im}T = [Tv, T^2v, \ldots, T^{m-1}v] = [Tv, T(Tv), \ldots, T^{m-2}(Tv)]$, com $T^{m-2}(Tv) \neq 0$. Logo, considerando $U' = U \cap \text{Im}T$, pela hipótese de indução, existe um subespaço T-invariante W' tal que

$$\text{Im}T = U' \oplus W'.$$

Para construirmos W do enunciado, considere

$$W'' = \{w \in V \mid Tw \in W'\}.$$

Afirmação 1: $V = U + W''$

De fato, dado $\tilde{v} \in V$, temos $T\tilde{v} \in \text{Im}T = U' \oplus W'$, logo podemos escrever $\tilde{v} = u' + w'$, com $u' \in U'$ e $w' \in W'$. Como $u' \in U' = [Tv, T^2v, \ldots, T^{m-1}v]$, existem escalares $a_1, a_2, \ldots, a_{m-1} \in K$ tais que

$$u' = a_1Tv + a_2T^2v + \cdots + a_{m-1}T^{m-1}v = T(\underbrace{a_1v + a_2Tv + \cdots + a_{m-1}T^{m-2}v}_{\in U}).$$

Denotando $u'' = a_1v + a_2Tv + \cdots + a_{m-1}T^{m-2}v$, temos $u' = T(u'')$ e, com isso, $T\tilde{v} = T(u'') + w'$.

$$\Rightarrow\ w' = T(\tilde{v} - u'') \in W' \ \Rightarrow\ v - u'' \in W''.$$

Logo

$$\tilde{v} = u'' + v - u'' \in U + W''.$$

Afirmação 2: $U \cap W' = \{0\}$

De fato, seja $u \in U \cap W'$. Como U e W' são T-invariantes, temos que $Tu \in U \cap \text{Im}T = U'$ e $Tu \in W'$. Como $U' \cap W' = \{0\}$, obtemos $Tu = 0$. Assim,

$$\begin{aligned}
u \in U \ &\Rightarrow\ u = a_0v + a_1Tv + \cdots + a_{m-1}T^{m-1}v, \quad \text{com } a_i \in K \\
&\Rightarrow\ 0 = Tu = a_0Tv + a_1T^2v + \cdots + a_{m-2}T^{m-1}v + a_{m-1}T^mv \\
&\Rightarrow\ a_0Tv + a_1T^2v + \cdots + a_{m-2}T^{m-1}v = 0 \quad \text{(pois } T^m = 0) \\
&\overset{\text{(a)}}{\Rightarrow}\ a_0 = a_1 = \cdots = a_{m-2} = 0 \\
&\Rightarrow\ u = a_{m-1}T^{m-1}v \in U'
\end{aligned}$$

Como $U' \cap W' = \{0\}$, segue que $u = 0$, o que prova a Afirmação 2.

Da Afirmação 2 segue que $(U \cap W'') \cap W' = \{0\}$. E, como $U \cap W'' \leq W''$ e $W' \leq W''$, segue que existe um subespaço $\overline{W}$ tal que

$$W'' = \overline{W} \oplus W' \oplus (U \cap W''). \qquad (*)$$

Afirmação 3: $W = \overline{W} \oplus W'$ é o subespaço desejado.

De fato,

(i) $U \cap W = \{0\}$: $W \subseteq W''$ e $W \cap (U \cap W'') = \{0\}$, logo $U \cap W = \{0\}$.

(ii) $V = U + W$: Pela Afirmação 1,

$$\dim V = \dim U + \dim W' - \dim(W \cap W').$$

Da igualdade (*), obtemos

$$\dim W'' = \dim W + \dim(U \cap W'').$$

Assim,

$$\dim V = \dim U + \dim W.$$

Como $U, W \leq V$, segue que $U + W = V$.

De (i) e (ii), obtemos que $V = U \oplus W$. Resta mostrar que W é T invariante. Como $W \subseteq W''$, pela definição de W'', temos que $T(W) \subseteq W'$. E como $W' \subseteq W$, obtemos $T(W) \subseteq W$, como queríamos mostrar.

■

Observação 4.41. *Seja $T \in L(V, V)$, onde $\dim_K V = n \geq 1$.*

1) *Se T é nilpotente de índice m, então existe $v \in V$ tal que $T^{m-1}v \neq 0$ e segue da Proposição 4.40(a) que*

$$m \leq n = \dim V,$$

já que $\mathcal{B} = \{v, Tv, T^2v, \ldots, T^{m-1}v\}$ é um conjunto L.I. em V.

Agora, se $m = n = \dim V$, então $\mathcal{B} = \{v, Tv, t^2v, \ldots, T^{m-1}v\}$ é uma base de V. A matriz de T nesta base é

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

2) *Suponha agora que $p_T(x) = (x - \lambda)^n$. Neste caso, pelo Teorema de Cayley-Hamilton, o operador linear $(T - \lambda I)$ é nilpotente. Se o índice de nilpotência de $(T - \lambda I)$ for n, então usando 1), existe uma base $\mathcal{B}$ de V tal que*

$$[T - \lambda I]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

Então

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda & 0 & \cdots & 0 & 0 \\ 1 & \lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda \end{pmatrix}_{n\times n}$$

O item (2) da observação acima é um caso bastante particular. Veremos a seguir o caso geral para operadores lineares nilpotentes. Para isso, vamos considerar a seguinte definição.

Definição 4.42. *Um* ***bloco de Jordan*** $r \times r$ ***em*** λ *é a matriz* $J_r(\lambda)$ *em* $\mathcal{M}_r(K)$ *que tem* λ *nas entradas da diagonal principal, 1 nas entradas da diagonal abaixo e 0 nas demais entradas, ou seja,*

$$J_r(\lambda) := \begin{pmatrix} \lambda & 0 & \cdots & 0 & 0 \\ 1 & \lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda \end{pmatrix}_{r\times r} .$$

Teorema 4.43. *Seja* $T \in L(V,V)$ *um operador linear nilpotente de índice* $m \geq 1$*, onde* $\dim_K V = n < \infty$*. Então existem inteiros positivos* $t,\ m_1,\ m_2,\ldots,\ m_t$ *e vetores* $v_1, v_2, \ldots, v_t \in V$ *tais que*

(a) $m = m_1 \geq m_2 \geq \cdots \geq m_t$ *;*

(b) *o conjunto* $\mathcal{B} = \{v_1, Tv_1, \ldots, T^{m_1-1}v_1,\ v_2, Tv_2, \ldots, T^{m_2-1}v_2, \ldots,\ v_t, Tv_t, \ldots, T^{m_t-1}v_t\}$ *é uma base de* V *(com isso,* $m_1 + m_2 + \cdots + m_t = n$*);*

(c) $T^{m_i}(v_i) = 0$*, para todo* $i = 1, 2, \ldots, t$ *;*

(d) *Se* S *for um operador linear de um espaço vetorial* W *de dimensão finita, então os inteiros* $t,\ m_1,\ m_2,\ldots,\ m_t$ *associados a* S *e a* T *são iguais se, e somente se, existe um isomorfismo* $\Phi : V \to W$ *com* $\Phi T \Phi^{-1} = S$ *(ou seja, as matrizes associadas a* T *e a* S *são semelhantes).*

Demonstração. Como $T^{m-1} \neq 0$, existe $v_1 \in V$ tal que $T^{m-1}(v_1) \neq 0$. Pela Proposição 4.40, $\mathcal{B}_1 = \{v_1, Tv_1, \ldots, T^{m_1-1}v_1\}$ é L.I. e $V = W_1 \oplus W_2'$, onde W_1 é o subespaço gerado por $\mathcal{B}_1$ e W_2' é T-invariante.

Escreva $m_1 = m$. Note que como T é nilpotente, a restrição de T a W_2' é também nilpotente de índice, digamos, m_2 onde $m_2 \leq m = m_1$. Assim, pela Proposição 4.40, existe $v_2 \in V$ tal que $\mathcal{B}_2 = \{v_2, Tv_2, \ldots, T^{m_2-1}v_2\}$ é L.I. e $W_2' = W_2 \oplus W_3'$, onde $W_2 = [v_2, Tv_2, \ldots, T^{m_2-1}v_2]$ e W_3' é T-invariante. Repetindo essa argumento, como $\dim_K V < \infty$, chegamos aos inteiros $t,\ m_1,\ m_2,\ldots,\ m_t$ do enunciado, satisfazendo (a), (b) e (c). O item (d) fica como exercício. ∎

Observação 4.44. *Se $T \in L(V,V)$ é nilpotente e t, m_1, $m_2, \ldots$, m_t são como no enunciado do teorema acima, então a matriz de T na base $\mathcal{B}$ é da forma*

$$[T]_{\mathcal{B}} = \begin{pmatrix} J_{m_1}(0) & 0 & \cdots & 0 \\ 0 & J_{m_2}(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_t}(0) \end{pmatrix},$$

onde $J_{m_i}(0)$ é um bloco de Jordan $m_i \times m_i$ com 0's na diagonal principal e 1's na diagonal abaixo da principal, ou seja,

$$J_{m_i}(0) := \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}_{m_i \times m_i}.$$

Exemplo 4.45. *Considere $T \in L(\mathbb{C}^3, \mathbb{C}^3)$ tal que $p_T(x) = (x-2)^3$.*
Pelo Teorema de Cayley-Hamilton, temos que

$$p_T(T) = (T - 2I)^3 = 0.$$

Então $T - 2I$ é nilpotente de índice $m \leq 3$. As possibilidades para o polinômio minimal de T são:

$$x - 2, \quad (x-2)^2, \quad (x-2)^3.$$

Vamos então analisar cada uma das possibilidades:

- *Se $m_T(x) = x - 2$, então*

$$\Rightarrow \ T - 2I = 0 \ \Rightarrow \ T - 2I \text{ nilpotente de índice 1.}$$

  Pelo Teorema 4.43, $m_1 = 1 \geq m_2$ e, portanto $m_2 = 1$. Como $m_1 + m_2 = 2 < 3 = \dim \mathbb{C}^3$, existe m_3 com $1 = m_1 \geq m_2 \geq m_3$, logo $m_3 = 1$ e, como $m_1 + m_2 + m_3 = 3$, temos $t = 3$. Além disso, existem $v_1, v_2, v_3 \in \mathbb{C}^3$ tais que $\mathcal{B} = \{v_1, v_2, v_3\}$ é L.I. e

$$[T - 2I]_{\mathcal{B}} = \begin{pmatrix} J_{m_1}(0) & 0 & 0 \\ 0 & J_{m_2}(0) & 0 \\ 0 & 0 & J_{m_3}(0) \end{pmatrix}.$$

  Como $m_1 = m_2 = m_3 = 1$, temos que $J_{m_i}(0) = (0)_{1\times 1}$. Assim

$$[T - 2I]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

e portanto

$$\Rightarrow \ [T]_{\mathcal{B}} - 2I_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ \Rightarrow \ [T]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

[Que é o caso em que T é diagonalizável.]

- *Se $m_T(x) = (x-2)^2$, então*

$$T - 2I \quad \textit{tem índice de nilpotência } m_1 = 2.$$

Pelo Teorema 4.43, existe m_2 tal que $2 = m_1 \geq m_2 \geq 1$. Como $m_1 + m_2 \leq 3 = \dim \mathbb{C}^3$, obtemos $m_2 = 1$ e $t = 2$. Além disso, existem $v_1, v_2 \in V$ tais que

$$\mathcal{B} = \{v_1, (T-2I)(v_1), v_2\} \quad \textit{é base de } \ \mathbb{C}^3$$

e

$$[T-2I]_{\mathcal{B}} = \begin{pmatrix} J_{m_1}(0) & 0 \\ 0 & J_{m_2}(0) \end{pmatrix},$$

onde

$$J_{m_1}(0) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_{2\times 2} \quad e \quad J_{m_2}(0) = (0)_{1\times 1}.$$

Portanto

$$\Rightarrow \quad [T-2I]_{\mathcal{B}} = \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 1 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)_{3\times 3} \quad \Rightarrow \quad [T]_{\mathcal{B}} = \left(\begin{array}{cc|c} 2 & 0 & 0 \\ 1 & 2 & 0 \\ \hline 0 & 0 & 2 \end{array}\right)_{3\times 3}.$$

- *Se $m_T(x) = (x-2)^3$, então*

$$T - 2I \quad \textit{tem índice de nilpotência } m_1 = 3 = \dim \mathbb{C}^3.$$

Então $t = 1$ e existe $v_1 \in V$ tal que $\mathcal{B} = \{v_1, (T-2I)(v_1), (T-2i)^2(v_1)\}$ é base de $\mathbb{C}^3$ e

$$\Rightarrow \quad [T-2I]_{\mathcal{B}} = (J_{m_1}(0)) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}_{3\times 3} \quad \Rightarrow \quad [T]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}_{3\times 3}$$

4.6 Formas de Jordan

Com os Teoremas 4.39 e 4.43 vistos na seção anterior, estamos prontos para provar o Teorema da Forma de Jordan para operadores lineares de espaços vetoriais de dimensão finita.

Teorema 4.46. *Seja $T \in L(V,V)$, onde $\dim_K V = n$, tal que o polinômio característico pode ser completamente fatorado como*

$$p_T(x) = (x-\lambda_1)^{s_1}(x-\lambda_2)^{s_2}\cdots(x-\lambda_r)^{s_r},$$

onde $s_i \geq 1$ e $\lambda_i \neq \lambda_j$ se $i \neq j$, $\forall i,j = 1,2,\ldots,r$. Então

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_r,$$

onde, para cada $i = 1,2,\ldots,r$, temos

a) $\dim_K U_i = s_i$ *;*

b) *U_i é T-invariante ;*

c) *a restrição de $T - \lambda_i I$ a U_i é nilpotente.*

Demonstração. Para cada $i = 1,2,\ldots,r$, considere $T_i = T - \lambda_i I$. Então $T_i : V \to V$ é um operador linear de V e portanto, pelo Teorema 4.39, T_i é uma soma direta de um operador linear nilpotente e um operador linear inversível, ou seja, $V = U_i \oplus W_i$, onde U_i e W_i são T invariantes e as restrições $T_i|_{U_i}$ e $T_i|_{W_i}$ são nilpotente e inversível, respectivamente.

Note que se W é um subespaço qualquer de V que é T_i-invariante, então W é também T-invariante. De fato, se $T_i(W) \subseteq W$, então $T_i(w) \in W$, $\forall w \in W$. Como $T_i(w) = (T - \lambda_i I)(w) = T(w) - \lambda_i w$ e $\lambda_i w \in W$, $\forall w \in W$, segue que $T(w) = \lambda_i w + T_i(w) \in W$, $\forall w \in W$. Logo $T(W) \subseteq W$.

Então, como U_i e W_i são T_i-invariantes, para cada $i = 1,2,\ldots,r$, temos que U_i e W_i são também T-invariantes. Fixado i, sejam

$$T' := T|_{U_i} \quad \text{e} \quad T'' := T|_{W_i}.$$

Como $V = U_i \oplus W_i$, se $\mathcal{C}_i$ é base de U_i e $\mathcal{D}_i$ é base de W_i, então $\mathcal{B}_i := \mathcal{C}_i \cup \mathcal{D}_i$ é base de V e

$$[T]_{\mathcal{B}_i} = \begin{pmatrix} [T']_{\mathcal{C}_i} & 0 \\ 0 & [T'']_{\mathcal{D}_i} \end{pmatrix},$$

de onde concluímos que

$$p_T(x) = p_{T'}(x) \cdot p_{T''}(x).$$

Afirmação: $p_{T'}(x) = (x - \lambda_i I)^{s_i}$.

De fato, como $T_i|_{U_i}$ é nilpotente, temos $T_i^l = 0$ em U_i, para algum l. Agora,

$$T_i^l = 0 \text{ em } U_i \;\Rightarrow\; (T-\lambda_i I)^l = 0 \text{ em } U_i, \text{ ou seja, } (T' - \lambda_i I)^l = 0 \text{ em } U_i.$$

Então o polinômio $f(x) = (x-\lambda_i)^l$ anula $T' = T|_{U_i}$. Com isso, temos que o polinômio minimal $m_{T'}(x)$ de T' divide $f(x)$. Portanto $m_{T'}(x) = (x-\lambda_i)^\alpha$. Além disso, como $m_{T'}(x)$ e $p_{T'}(x)$ possuem as mesmas raízes, obtemos $p_{T'}(x) = (x-\lambda_i)^t$.

Resta mostrar que $t = s_i$. Sabemos que $T_i|_{W_i}$ é inversível, logo $\mathrm{Ker} T_i|_{W_i} = \{0\}$, ou seja, $T - \lambda_i I$ tem núcleo trivial em W_i. Assim, λ_i não é autovalor de $T'' = T|_{W_i}$. Logo $(x - \lambda_i)$

não divide $p_{T''}(x)$. Mas $(x-\lambda_i)^{s_i}$ divide $p_T(x) = p_{T'}(x)\cdot p_{T''}(x)$, portanto $(x-\lambda_i)^{s_i}$ divide $p_{T'}(x)$. Logo $t = s_i$ e $p_{T'}(x) = (x-\lambda_i)^{s_i}$ como afirmamos.

Como $p_{T'}(x) = (x-\lambda_i)^{s_i}$ e $T' = T|_{U_i}$, segue que $\dim_K U_i = s_i$ e $U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_r) = \{0\}$. para cada $i = 1, 2, \ldots, r$. E como $s_1 + s_2 + \cdots + s_r = \dim V$, obtemos $V = U_1 \oplus U_2 \oplus \cdots \oplus U_r$.

■

Do teorema acima obtemos o seguinte: como $V = U_1 \oplus U_2 \oplus \cdots \oplus U_r$ e $T_i = (T - \lambda_i I)|_{U_i}$ é nilpotente, existe uma base $\mathcal{B}_i$ de U_i tal que

$$[T|_{U_i}]_{\mathcal{B}_i} = \begin{pmatrix} J_{m_1}(\lambda_i) & 0 & \cdots & 0 \\ 0 & J_{m_2}(\lambda_i) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_{t_{\lambda_i}}}(\lambda_i) \end{pmatrix},$$

onde, para cada $j = 1, 2, \ldots, t_{\lambda_i}$,

$$J_{m_j}(\lambda_i) = \begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 \\ 1 & \lambda_i & 0 & \cdots & 0 \\ 0 & 1 & \lambda_i & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda_i \end{pmatrix}_{m_j \times m_j}$$

é um bloco de Jordan, $m_1 \geq m_2 \geq \cdots \geq m_{t_{\lambda_i}}$ e $m_1 + m_2 + \cdots + m_{t_{\lambda_i}} = \dim U_i$.

Então $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ é base de V e

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T|_{U_1}]_{\mathcal{B}_1} & 0 & \cdots & 0 \\ 0 & [T|_{U_2}]_{\mathcal{B}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [T|_{U_r}]_{\mathcal{B}_r} \end{pmatrix},$$

que chamamos de **Forma de Jordan associada a** T.

Pelo Teorema 4.43, os números t_{λ_i}e m_j, estão bem determinados a partir de T. Além disso, dois operadores lineares $S \in L(V,V)$ e $T \in L(V',V')$ têm a mesma forma de Jordan se, e somente se, existe um isomorfismo $\Phi : V \to V'$ tal que $\Phi^{-1}T\Phi = S$.

Exemplo 4.47. *Seja* $T : \mathbb{C}^4 \to \mathbb{C}^4$ *dado por*

$$T(z_1, z_2, z_3, z_4) = (8z_1 - z_2, 4z_1 + 12z_2, 9z_3 + 2z_4, 2z_3 + 6z_4).$$

A matriz de T *na base canônica* $\mathcal{C}$ *de* $\mathbb{C}^4$ *é*

$$A = [T]_{\mathcal{B}} = \begin{pmatrix} 8 & -1 & 0 & 0 \\ 4 & 12 & 0 & 0 \\ 0 & 0 & 9 & 2 \\ 0 & 0 & 2 & 6 \end{pmatrix}.$$

Calculando o polinômio característico de T, obtem-se

$$p_T(x) = \det(xI_4 - A) = (x-10)^3(x-5).$$

Com isso temos que $V = U_1 \oplus U_2$, com $\dim_{\mathbb{C}} U_1 = 3$ e $\dim_{\mathbb{C}} U_2 = 1$.

Os possíveis polinômios minimais de T são $(x-10)(x-5)$, $(x-10)^2(x-5)$ e $(x-10)^3(x-5)$. Calculando, obtemos

$$m_T(x) = (x-10)^2(x-5).$$

Então o primeiro bloco de Jordan associado a $\lambda_1 = 10$ é 2×2. E como $\dim U_1 = 3$, temos um segundo bloco 1×1. E para $\lambda_2 = 5$, que é autovalor de multiplicidade algébrica 1, temos um único bloco de Jordan 1×1.

Assim a Forma de Jordan associada a T é

$$J = \left(\begin{array}{ccc|c} 10 & 0 & 0 & 0 \\ 1 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ \hline 0 & 0 & 0 & 5 \end{array}\right)$$

Exemplo 4.48. *Seja $T \in L(\mathbb{R}^2, \mathbb{R}^2)$ tal que $p_T(x) = (x+2)^4$ e $m_T(x) = (x+2)^2$.*

T possui um único autovalor $\lambda_1 = -2$. Como sua multiplicidade em $m_T(x)$ é 2, temos que o primeiro bloco de Jordan associado a λ_1 é 2×2. Então as possíveis formas de Jordan para T são

$$J_1 = \left(\begin{array}{cc|cc} -2 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ \hline 0 & 0 & -2 & 0 \\ 0 & 0 & 1 & -2 \end{array}\right) \quad ou \quad J_2 = \left(\begin{array}{cc|cc} -2 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ \hline 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{array}\right)$$

$$(m_1 = 2 \ \ e \ \ m_2 = 2) \qquad\qquad (m_1 = 2 \ \ e \ \ m_2 = m_3 = 1)$$

Para a forma J_1, a base dada pelo Teorema 4.43 é do tipo

$$\mathcal{B} = \{v_1, (T+2I)(v_1), v_2, (T+2I)(v_2)\}$$

e $J_1 = [T]_{\mathcal{B}}$. Observe que a segunda e quarta colunas de J_1 estão na forma diagonal, digamos assim. De fato o segundo vetor $(T+2I)(v_1)$ e o quarto vetor $(T+2I)(v_2)$ de $\mathcal{B}$ são autovetores de T, já que ao aplicarmos $(T+2I)$ nestes vetores, obtemos $(T+2I)^2(v_j) = 0$, $j = 1,2$. Então, na forma J_1, temos 2 autovetores L.I. associados ao autovalor -2, ou seja, $\dim Aut_T(-2) = 2$.

Para a forma J_2, a base dada pelo Teorema 4.43 é do tipo

$$\mathcal{B} = \{v_1, (T+2I)(v_1), v_2, v_3)\}$$

e $J_2 = [T]_{\mathcal{B}}$. Neste caso, a segunda, terceira e quarta colunas de J_2 estão na forma diagonal. De fato os vetores $(T+2I)(v_1)$, v_2 e v_3 de $\mathcal{B}$ são autovetores de T, já que ao aplicarmos $(T+2I)$ nestes vetores, obtemos $(T+2I)^2(v_1) = 0$, $(T+2I)(v_2) = 0$ e $(T+2I)(v_3) = 0$. Então, na forma J_2, temos 3 autovetores L.I. associados ao autovalor -2, ou seja, $\dim Aut_T(-2) = 3$.

Exemplo 4.49. *Seja $T \in L(\mathbb{C}^2, \mathbb{C}^2)$ qualquer.*

Então o polinômio característico de T é

$$p_T(x) = (x-a)(x-b), \quad \textit{onde } a, b \in \mathbb{C}.$$

i) *Se $a \neq b$, então $m_t(x) = (x-a)(x-b)$ e $\mathbb{C}^2$ possui dois autovetores associados a autovalores distintos. Assim T é diagonalizável e existe uma base de $\mathbb{C}^2$ (formada por autovetores de T) tal que*

$$[T] = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

ii) *Se $a = b$, então $p_T(x) = (x-a)^2$ e temos duas possibilidades para $m_T(x)$: $m_T(x) = (x-a)$ ou $m_T(x) = (x-a)^2$.*

Se $m_T(x) = x - a$, então $m_T(T) = T - aI$ anula $\mathbb{C}^2$ e portanto todo vetor não nulo é um autovetor de T. Assim

$$[T] = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

em alguma base de $\mathbb{C}^2$.

Agora, se $m_T(x) = (x-a)^2$, então a foram de Jordan de T possui um bloco de Jordan $J_2(a) = \begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}$ e assim a forma de Jordan de T é

$$[T] = \begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}.$$

Com isso, podemos concluir que toda matriz $A \in \mathcal{M}_2(\mathbb{C})$ é semelhante a uma das matrizes

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \textit{ou} \quad \begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}, \qquad a, b \in \mathbb{C}.$$

A seguir, reunimos as informações obtidas ao longo da seção sobre a Forma de Jordan de um operador linear.

Resumo da Forma de Jordan de um operador linear:

Seja V um espaço vetorial de dimensão finita n sobre um corpo K e $T \in L(V, V)$ tal que

$$p_T(x) = (x - \lambda_1)^{\alpha_1}(x - \lambda_2)^{\alpha_2} \cdots (x - \lambda_r)^{\alpha_r},$$

onde $\lambda_i \neq \lambda_j$, $\forall i \neq j$, e $\alpha_i \geq 1$, $i, j = 1, 2, \ldots, r$.

- Então λ_i é autovalor de T com multiplicidade algébrica $ma(\lambda_i) = \alpha_i$.
- A multiplicidade geométrica de λ_i é $mg(\lambda_i) = \dim Aut_T(\lambda_i)$, a dimensão do auto-espaço associado a λ_i.

- $V = U_1 \oplus U_2 \oplus \cdots \oplus U_r$, onde U_i é T-invariante, $\dim U_i = \alpha_i$ e $(T - \lambda_i I)|_{U_i}$ é nilpotente.

  Com isso, se $\mathcal{B}_i$ é uma base de U_i, então $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$ é base de V e

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T|_{U_1}]_{\mathcal{B}_1} & 0 & \cdots & 0 \\ 0 & [T|_{U_2}]_{\mathcal{B}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [T|_{U_r}]_{\mathcal{B}_r} \end{pmatrix}_{n \times n},$$

  onde cada $[T|_{U_i}]_{\mathcal{B}_i}$ é uma matriz $\alpha_i \times \alpha_i$.

- Como cada $(T - \lambda_i I)|_{\mathcal{B}_i}$ é nilpotente, segue que existe uma base $\mathcal{B}_i$ de U_i tal que

$$[T|_{U_i}]_{\mathcal{B}_i} = \begin{pmatrix} J_{m_{i_1}(\lambda_i)} & 0 & \cdots & 0 \\ 0 & J_{m_{i_2}(\lambda_i)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_{it_i}(\lambda_i)} \end{pmatrix}_{\alpha_i \times \alpha_i},$$

  onde cada $J_{m_{i_j}(\lambda_i)}$ é um bloco de Jordan, $m_{i_1} \geq m_{i_2} \geq \cdots \geq m_{it_i} \geq 1$ e $m_{i_1} + m_{i_2} + \cdots + m_{it_i} = \alpha_i$.

- A base $\mathcal{B}_i$ que dá a Forma de Jordan para cada $T|_{U_i}$ é da forma

$$\mathcal{B}_i = \Big\{ v_{i_1}, (T - \lambda_i I)(v_{i_1}), \ldots, (T - \lambda_i I)^{m_{i_1}-1}(v_{i_1}), \quad v_{i_2}, (T - \lambda_i I)(v_{i_2}), \ldots,$$
$$(T - \lambda_i I)^{m_{i_2}-1}(v_{i_2}), \ \ldots \ , v_{it_i}, (T - \lambda_i I)(v_{it_i}), \ldots, (T - \lambda_i I)^{m_{it_i}-1}(v_{it_i}) \Big\}.$$

  Os vetores "finais" $(T - \lambda_i I)^{m_{i_1}-1}(v_{i_1})$, $(T - \lambda_i I)^{m_{i_2}-1}(v_{i_2})$, $\ldots$, $(T - \lambda_i I)^{m_{it_i}-1}(v_{it_i})$ são autovetores L.I. de T associados a λ_i.

- O número de blocos de Jordan associados a cada λ_i é igual sua multiplicidade geométrica $mg(\lambda_i)$, ou seja, igual a $\dim Aut_T(\lambda_i)$.

- Se o polinômio monimal de T é

$$m_T(x) = (x - \lambda_1)^{\beta_1}(x - \lambda_2)^{\beta_2} \cdots (x - \lambda_r)^{\beta_r},$$

  então o primeiro bloco de Jordan associado a cada λ_i tem ordem $\beta_i \times \beta_i$, ou seja, $m_{i_1} = \beta_i$.

- A Forma de Jordan de T é única (fixada a ordenação dos λ_i's). E dois operadores T e S possuem a mesma forma de Jordan se, e somente se, suas matrizes são semelhantes, ou seja, existe uma matriz inversível P tal que $[S] = P^{-1}[T]P$.

Exemplo 4.50. *Vamos encontrar a forma de Jordan da matriz*

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \in \mathcal{M}_6(\mathbb{C}).$$

O polinômio característico de A é

$$p_A(x) = \det(xI_6 - A) = (x-2)^4(x-1)(x+1).$$

Então $\lambda_1 = 2, \lambda_2 = 1$ e $\lambda_3 = -1$ são os autovalores de A. O caso sensível é $\lambda_1 = 2$. Vamos calcular o polinômio minimal de A para saber o tamanho do primeiro bloco de Jordan associado a λ_1. Temos que

$$(A-2I)(A-I)(A+I) =$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 & 0 \\ 1 & 1 & 1 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 3 & 0 & 0 & 0 & 0 \\ 5 & 7 & 0 & 3 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 \end{pmatrix} \neq 0$$

$$(A-2I)^2(A-I)(A+I) =$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 3 & 0 & 0 & 0 & 0 \\ 5 & 7 & 0 & 3 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 \\ 7 & 3 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \neq 0$$

$$(A-2I)^3(A-I)(A+I) =$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 \\ 7 & 3 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \neq 0$$

Então $m_T(x) = p_T(x) = (x-2)^4(x-1)(x+1)$. Com isso, temos um único bloco de Jordan 4×4 associado a $\lambda_1 = 2$, 1 bloco de Jordan 1×1 associado a $\lambda_2 = 1$ e 1 bloco de Jordan 1×1 associado a $\lambda_3 = -1$. Assim, a Forma de Jordan de A é

$$J = \left(\begin{array}{cccc|cc} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right)$$

Note que a informação de que a forma de Jordan de A tem apenas um bloco de Jordan associado a $\lambda_1 = 2$ pode também ser vista a partir do auto-espaço $Aut_A(\lambda_1) = \mathrm{Ker}(\lambda_1 I - A)$: Temos

$$\lambda_1 I - A = 2I - A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 3 \end{pmatrix}$$

e

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 3 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} z_1 = 0 \\ z_2 = 0 \\ z_3 = 0 \\ z_4 = 0 \\ -z_5 + 3z_6 = 0 \end{cases}$$

Então

$$Aut_A(\lambda_1) = \{(0,0,0,0,3z_6,z_6) \mid z_6 \in \mathbb{C}\} = [(0,0,0,0,3,1)].$$

Como $\dim Aut_A(\lambda_1) = 1$, temos um bloco de Jordan associado a λ_1.

Exemplo 4.51. *Considere*

$$B = \begin{pmatrix} 1 & 0 & -2 & 0 & 4 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathcal{M}_5(\mathbb{C}).$$

O polinômio característico de B é

$$p_B(x) = \det(xI_5 - B) = (x-1)^5.$$

Então $\lambda_1 = 1$ é o único os autovalor de B.

Vamos encontrar $Aut_B(\lambda_1) = \mathrm{Ker}(\lambda_1 I - B)$*: Temos*

$$\lambda_1 I - B = I - B = \begin{pmatrix} 0 & 0 & 2 & 0 & -4 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

e

$$\begin{pmatrix} 0 & 0 & 2 & 0 & -4 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 2z_3 - 4z_5 = 0 \\ z_4 - z_5 = 0 \\ z_4 = 0 \end{cases} \Leftrightarrow \begin{cases} z_3 = 0 \\ z_4 = 0 \\ z_5 = 0 \end{cases}$$

Então

$$Aut_B(\lambda_1) = \{(z_1, z_2, 0, 0, 0) \mid z_1, z_2 \in \mathbb{C}\} = [(1,0,0,0,0), (0,1,0,0,0)].$$

Como $\dim Aut_B(\lambda_1) = 2$*, temos 2 blocos de Jordan associados a* λ_1*. Como* λ_1 *é o único autovalor, temos* $\dim U_1 = \dim V = 5$*, a soma dos tamanhos dos blocos de Jordan associados a* λ_1 *é igual a 5. Como são dois blocos, temos duas possibilidades: 1 bloco* 3×3 *e 1 bloco* 2×2*; ou 1 bloco* 4×4 *e 1 bloco* 1×1*. O tamanho do primeiro bloco é dado pela multiplicidade de* λ_1 *no polinômio minimal. Temos*

$$B - 1I = \begin{pmatrix} 0 & 0 & -2 & 0 & 4 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (B - 1I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad e \quad (B - 1I)^3 = 0$$

Então o primeiro bloco de Jordan é 3×3 *e a Forma de Jordan de* B *é*

$$J = \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right).$$

5 Espaços com Produto Interno

Nota: Ao longo de todo este capítulo, vamos fixar o corpo K com sendo $\mathbb{R}$ ou $\mathbb{C}$.

5.1 Produto Interno

Definição 5.1. *Seja V um espaço vetorial sobre K. Um* ***produto interno sobre*** *V é uma função $\langle\ ,\ \rangle : V \times V \to K$ que satisfaz:*

(P1) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$;

(P2) $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$;

(P3) $\langle u, v\rangle = \overline{\langle v, u\rangle}$, *onde a barra significa o conjugado;*

(P4) $\langle u, u\rangle > 0$, *se* $u \neq 0$;

para quaisquer $u, v, w \in V$ *e* $\lambda \in K$.

A seguir, algumas observações a cerca da definição acima.

Observação 5.2. *Seja V um espaço vetorial sobre K com produto interno $\langle\ ,\ \rangle$.*

1) *Segue da definição de produto interno que*

$$\langle 0, v\rangle = \langle v, 0\rangle = 0, \quad \forall v \in V$$

$$e$$

$$\langle v, v\rangle = 0 \;\Leftrightarrow\; v = 0.$$

2) *De* (P1) *e* (P3) *segue*

(P5) $\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle, \;\; \forall u, v, w \in V$.

De fato,

$\langle u, v+w\rangle \overset{(P3)}{=} \overline{\langle v+w, u\rangle} \overset{(P1)}{=} \overline{\langle v, u\rangle + \langle w, u\rangle} = \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} \overset{(P3)}{=} \langle u, v\rangle + \langle u, w\rangle.$

3) *De* (P2) *e* (P3) *segue*

(P6) $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle, \;\; \forall u, v \in V$ *e* $\lambda \in K$.

De fato,

$\langle u, \lambda v\rangle \overset{(P3)}{=} \overline{\langle \lambda v, u\rangle} \overset{(P2)}{=} \overline{\lambda\langle v, u\rangle} = \overline{\lambda}\ \overline{\langle v, u\rangle} \overset{(P3)}{=} \overline{\lambda}\langle u, v\rangle.$

4) *Das observações* 2) *e* 3) *acima, obtemos*

$$\langle \sum_{i=1}^{n} a_i u_i \ , \ \sum_{j=1}^{m} b_j v_j \ \rangle = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i \overline{b_j} \langle u_i, v_j \rangle$$

para quaisquer $u_i, v_j \in V$ *e* $a_i, b_j \in K$.

5) *Se* $K = \mathbb{R}$, *a condição* (P3) *é*

$$\langle u, v\rangle = \langle v, u\rangle, \quad \forall u, v \in V. \qquad (*)$$

Essa simetria se perde quando $K = \mathbb{C}$. *De fato, se* V *é um espaço vetorial sobre* $\mathbb{C}$ *com produto interno* $\langle \ , \ \rangle$, *temos que*

$$\langle v, v\rangle > 0, \quad \forall v \in V, v \neq 0.$$

Então

$$\langle iv, iv\rangle > 0, \quad \forall v \in V, v \neq 0.$$

Por outro lado, se valesse a simetria (*), *chegaríamos a*

$$\langle iv, iv\rangle = i\langle v, iv\rangle \overset{(*)}{=} i\langle iv, v\rangle = i^2\langle v, v\rangle = -\langle v, v\rangle < 0, \quad \forall v \in V, v \neq 0,$$

uma contradição.

Exemplo 5.3. *Seja* $V = K^n$ ($K = \mathbb{R}$ *ou* $K = \mathbb{C}$). *Temos que*

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle := x_1\overline{y}_1 + x_2\overline{y}_2 + \cdots + x_n\overline{y}_n.$$

é um produto interno sobre K^n (*Verifique que* (P1)-(P4) *valem!*). *Este produto interno é chamado de* ***produto interno canônico***.

Em geral, se $\alpha_1, \alpha_2, \ldots, \alpha_n$ *são números reais positivos, então*

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle := \alpha_1 x_1\overline{y}_1 + \alpha_2 x_2\overline{y}_2 + \cdots + \alpha_n x_n\overline{y}_n.$$

é um produto interno em K^n. *Verifique que* (P1)-(P4) *valem e entenda porque os* $\alpha_i's$ *são números reais positivos.*

Exemplo 5.4. *Considere* $V = \mathcal{C}([a, b], \mathbb{R})$, *o espaço vetorial das funções contínuas do intervalo* $[a, b] \subset \mathbb{R}$ *em* $\mathbb{R}$. *Defina*

$$\langle f, g\rangle := \int_a^b f(t)g(t)dt, \quad \forall f, g \in V.$$

As regras de integração garantem que $\langle \ , \ \rangle$ *como definido acima é um produto interno sobre* V. *(Verifique!)*

Exemplo 5.5. *Seja* $V = \mathcal{M}_n(K)$, *o espaço vetorial das matrizes* $n \times n$ *com entradas em* K, *onde* $K = \mathbb{R}$ *ou* $\mathbb{C}$. *O* ***produto interno canônico sobre*** $\mathcal{M}_n(K)$ *é definido por*

$$\langle A, B\rangle := \sum_{i,j=1}^{n} a_{ij}\overline{b_{ij}},$$

onde $A = (a_{ij})$ *e* $B = (b_{ij})$ *são matrizes em* V. *Verifique que* (P1)-(P4) *valem.*

Exemplo 5.6. *Sejam* V *e* W *dois espaços vetoriais sobre* K *e* $\langle\ ,\ \rangle$ *um produto interno sobre* V. *Se* $T : W \to V$ *é uma transformação linear* ***injetora***, *então podemos definir um produto interno sobre* W *da seguinte maneira:*

$$\langle u, w\rangle_T := \langle T(u), T(w)\rangle, \qquad \forall u, w \in W.$$

$\langle\ ,\ \rangle_T$ *é, de fato, um produto interno: para quaisquer* $u, v, w \in W$ *e* $\lambda \in K$, *temos*

(P1) $\langle u + v, w\rangle_T = \langle T(u + v), T(w)\rangle = \langle T(u) + T(v), T(w)\rangle$
$$\overset{\text{(P1)}}{=} \langle T(u), T(w)\rangle + \langle T(v), T(w)\rangle = \langle u, w\rangle_T + \langle v, w\rangle_T$$

(P2) $\langle \lambda u, w\rangle_T = \langle T(\lambda u), T(w)\rangle = \langle \lambda T(u), T(w)\rangle \overset{\text{(P2)}}{=} \lambda\langle T(u), T(w)\rangle\rangle = \lambda\langle u, w\rangle_T$

(P3) $\langle u, w\rangle_T = \langle T(u), T(w)\rangle \overset{\text{(P3)}}{=} \overline{\langle T(w), T(u)\rangle} = \overline{\langle w, u\rangle_T}$

(P4) *Para cada* $u \neq 0$, *como* T *é injetora, temos* $T(u) \neq 0$. *Então*

$$\langle u, u\rangle_T = \langle T(u), T(u)\rangle \overset{(P4)}{>} 0, \quad \forall u \in W,\ u \neq 0.$$

Caso especial do Exemplo 5.6: *Seja* V *um espaço vetorial de dimensão* $n \geq 1$ *sobre* K *e sejam* $\mathcal{B} = \{v_1 v_2, \ldots, v_n\}$ *uma base de* V *e* $\mathcal{C} = \{e_1, e_2, \ldots, e_n\}$ *a base canônica de* K^n.

Consideremos K^n *com seu produto interno canônico e seja* $T : V \to K^n$ *a transformação linear tal que* $T(v_i) = e_i$, $i = 1, 2, \ldots, n$ *(que existe e é unica, pelo Teorema 3.9), ou seja,*

$$T(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i T(v_i) = \sum_{i=1}^{n} \alpha_i e_i = (\alpha_1, \alpha_2, \ldots, \alpha_n).$$

Como T *leva base em base, temos que* T *é um isomorfismo e, portanto, injetora. Então* T *e o produto interno canônico em* K^n *induzem um produto interno em* V *dado por:*

dados $u = \sum_{i=1}^{n} \alpha_i v_i$ *e* $v = \sum_{i=1}^{n} \beta_i v_i$,

$$\begin{aligned}\langle u,v\rangle_T &= \langle T(u),T(v)\rangle = \langle \textstyle\sum_{i=1}^n \alpha_i T(v_i), \sum_{i=1}^n \beta_i T(v_i)\rangle \\ &= \langle \textstyle\sum_{i=1}^n \alpha_i e_i, \sum_{i=1}^n \beta_i e_i\rangle \\ &= \langle(\alpha_1,\alpha_2,\dots,\alpha_n),(\beta_1,\beta_2,\dots,\beta_n)\rangle \\ &= \alpha_1\overline{\beta_1}+\alpha_2\overline{\beta_2}+\cdots+\alpha_n\overline{\beta_n}\end{aligned}$$

Em particular, temos que $\langle v_i,v_j\rangle_T = \delta_{ij}$, *para* $i,j=1,2,\dots,n$, *onde* $\delta_{ii}=1$ *e* $\delta_{ij}=0$ *para* $i\neq j$.

Agora, note que temos

$$[u]_{\mathcal{B}} = \begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix} \qquad e \qquad [v]_{\mathcal{B}} = \begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_n\end{pmatrix}.$$

Com isso, vemos que

$$\langle u,v\rangle_T = ([u]_{\mathcal{B}})^t\cdot\overline{[v]}_{\mathcal{B}}.$$

Definição 5.7. *Seja* V *um espaço vetorial sobre* K, *munido de um produto interno* $\langle\ ,\rangle$. *Para cada* $v\in V$, *chamamos de* ***norma*** *de* v *o número real dado por*

$$||v|| = \sqrt{\langle v,v\rangle}.$$

Observe que a definição de norma de um vetor depende do produto interno dado. Além disso, da definição segue que

(a) $||v||\geq 0$, $\forall v\in V$, e $||v||=0$ se, e somente se, $v=0$;

(b) $||\alpha v|| = \sqrt{\langle\alpha v,\alpha v\rangle} = \sqrt{\alpha\overline{\alpha}\langle v,v\rangle} \overset{(*)}{=} |\alpha|\sqrt{\langle v,v\rangle} = |\alpha|\cdot||v||$, para todo $\alpha\in K$.

(*) $\alpha = x+iy$ e $\overline{\alpha} = x-iy \Rightarrow \alpha\overline{\alpha} = x^2+y^2 = |\alpha|^2$.

Exemplo 5.8. *Considere* $V=\mathbb{C}^3$ *com o produto interno canônico e* $w=(-\frac{2i}{5},1-2i,0)\in\mathbb{C}^3$.

Então

$\langle w,w\rangle = \langle(-\frac{2i}{5},1-2i,0),(-\frac{2i}{5},1-2i,0)\rangle = -\frac{2i}{5}\cdot\overline{-\frac{2i}{5}}+(1-2i)(\overline{1-2i}) = \frac{4}{25}+5 = \frac{129}{25}$

Assim, $||w|| = \sqrt{\langle w,w\rangle} = \frac{\sqrt{129}}{5}$.

A seguir apresentamos algumas desigualdades envolvendo produto interno e a norma associada.

Proposição 5.9. *Seja* V *um espaço vetorial sobre* K *com produto interno* $\langle\ ,\rangle$. *Então*

i) Desigualdade de Cauchy-Schwarz: $|\langle u,v\rangle| \leq ||u|| \cdot ||v||, \quad \forall u,v \in V.$
E a igualdade $|\langle u,v\rangle| = ||u|| \cdot ||v||$ *vale se, e somente se,* $\{u,v\}$ *é L.D..*

ii) Desigualdade triangular: $||u+v|| \leq ||u|| + ||v||, \quad \forall u,v \in V.$

Demonstração. i) Para quaisquer $\alpha, \beta \in K$ e $u, v \in V$, temos (Verifique!)

$$\langle \alpha u - \beta v, \alpha u - \beta v\rangle = |\alpha|^2||u||^2 - 2Re(\alpha\overline{\beta}\langle u,v\rangle) + |\beta|^2||v||^2 \qquad (*)$$

onde, dado $z \in \mathbb{C}$, $Re(z)$ denota a parte real de z.

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

$$\alpha\overline{\beta}\langle u,v\rangle = ||v||^2\overline{\beta}\beta = ||v||^2|\beta|^2 \quad \in \mathbb{R}.$$

Substituindo em (*), obtemos

$$\begin{aligned} 0 \leq \langle \alpha u - \beta v, \alpha u - \beta v\rangle &= ||u||^2||v||^4 - 2||v||^2|\beta|^2 + |\langle u,v\rangle|^2||v||^2 \\ &= ||v||^2(||u||^2||v||^2 - |\langle u,v\rangle|^2). \qquad (**) \end{aligned}$$

Portanto

$$||u||^2||v||^2 - |\langle u,v\rangle|^2 \geq 0,$$

ou seja,

$$|\langle u,v\rangle| \leq ||u|| \cdot ||v||.$$

Agora, se vale a igualdade $|\langle u,v\rangle| = ||u|| \cdot ||v||$, obtemos em (**) que

$$\langle \alpha u - \beta v, \alpha u - \beta v\rangle = 0.$$

Assim

$$\alpha u - \beta v = 0$$

Se $v = 0$, $\{u,v\}$ é L.D. independentemente da igualdade. E se $v \neq 0$, então $\alpha = ||v||^2 \neq 0$. Portanto, segue da combinação linear acima que $\{u,v\}$ é L.D..

Reciprocamente, se $\{u,v\}$ é L.D., então $u = \lambda v$ ou $v = \lambda u$, para algum $\lambda \in K$, e uma conta simples mostra que vale a igualdade $|\langle u,v\rangle| = ||u||^2||v||^2$. (Verifique!)

ii) Temos que (Verifique!)

$$||u+v||^2 = ||u||^2 + 2Re(\langle u,v\rangle) + ||v||^2, \quad \forall u,v \in V.$$

Como, para todo $z \in \mathbb{C}$, temos $Re(z) \leq |z|$, obtemos

$$||u+v||^2 \leq ||u||^2 + 2|\langle u,v\rangle| + ||v||^2 \overset{i)}{\leq} ||u||^2 + 2||u||.||v|| + ||v||^2 = (||u|| + ||v||)^2.$$

■

5.2 Ortogonalidade

Definição 5.10. *Seja V um espaço vetorial sobre K (lembre que aqui sempre $K = \mathbb{R}$ ou $K = \mathbb{C}$), com produto interno $\langle\ ,\rangle$. Dizemos que:*

i) *$u, v \in V$ são **ortogonais** quando $\langle u, v\rangle = 0$;*

ii) *$\mathcal{B} \subset V$ é um **subconjunto ortogonal** quando seus elementos são dois-a-dois ortogonais;*

iii) *$\mathcal{B} \subset V$ é um **subconjunto ortonormal** quando $\mathcal{B}$ é ortogonal e $||u|| = 1$, $\forall u \in \mathcal{B}$.*

Notação: Escrevemos $u \perp v$ para indicar que u e v são ortogonais.

Observação 5.11. *O vetor nulo é ortogonal a todos os vetores de V, pois*

$$\langle 0, v\rangle = 0, \quad \forall v \in V.$$

Além disso, o vetor nulo é o único com esta propriedade. (Por quê?)

Exemplo 5.12. *As bases canônicas de $\mathbb{R}^n$, $\mathbb{C}^n$, $\mathcal{M}_n(\mathbb{R})$ e $\mathcal{M}_n(\mathbb{C})$ com os respectivos produtos internos canônicos são conjuntos ortogonais.*

Proposição 5.13. *Seja V um espaço vetorial sobre K com produto interno e seja $\mathcal{B}$ um subconjunto ortogonal de V formado por vetores não nulos. Então*

i) *$\mathcal{B}$ é L.I.;*

ii) *se $u_1, u_2, \dots, u_n \in \mathcal{B}$ e $v \in [u_1, u_2, \dots, u_n]$, então*

$$v = \frac{\langle v, u_1\rangle}{||u_1||^2} \cdot u_1 + \frac{\langle v, u_2\rangle}{||u_2||^2} \cdot u_2 + \dots + \frac{\langle v, u_n\rangle}{||u_n||^2} \cdot u_n.$$

Demonstração. Primeiro vamos mostrar que $\mathcal{B}$ é L.I.. De fato, sejam $u_1, u_2, \dots, u_n \in \mathcal{B}$ e suponha que

$$k_1u_1 + k_2u_2 + \dots + k_nu_n = 0, \quad \text{com} \quad k_1, k_2, \dots, k_n \in K.$$

Fazendo o produto interno com u_1, obtemos

$$\langle k_1u_1 + k_2u_2 + \dots + k_nu_n, u_1\rangle = \langle 0, u_1\rangle = 0.$$

Assim,

$$k_1\langle u_1, u_1\rangle + k_2\langle u_2, u_1\rangle + \dots + k_n\langle u_n, u_1\rangle = 0.$$

Como $\mathcal{B}$ é ortogonal, temos que $\langle u_i, u_j\rangle = 0$, $\forall u_i, u_j \in \mathcal{B}$ e $i \neq j$. Então obtemos da igualdade acima que $k_1\langle u_1, u_1\rangle = 0$. Como os vetores em $\mathcal{B}$ são não nulos, temos $\langle u_1, u_1\rangle > 0$ e portanto $k_1 = 0$. Repetindo o mesmo argumento com o produto interno com u_i, para $i = 2, 3, \dots, n$, obtemos $k_i = 0$ para $i = 1, 2, \dots, n$. Portanto, os vetores $u_1, u_2, \dots, u_n$ são L.I., para quaisquer $u_1, u_2, \dots, u_n \in \mathcal{B}$. Logo $\mathcal{B}$ é L.I..

Agora vamos mostrar a segunda parte do enunciado. Se $v \in [u_1, u_2, \dots, u_n]$, então existem escalares $b_1, b_2, \dots, b_n \in K$ tais que

$$v = b_1 u_1 + b_2 u_2 + \cdots + b_n u_n.$$

Vamos mostrar que os escalares $b_1, b_2, \dots, b_m$ são iguais aos respectivos escalares da expressão em ii). De fato, considere o produto interno $\langle v, u_1 \rangle$. Substituindo a expessão de v acima, obtemos

$$\begin{aligned} \langle v, u_1 \rangle &= \langle b_1 u_1 + b_2 u_2 + \cdots + b_n u_n, u_1 \rangle \\ &= b_1 \langle u_1, u_1 \rangle + b_2 \langle u_2, u_1 \rangle + \cdots + b_n \langle u_n, u_1 \rangle \\ &= b_1 ||u_1||^2, \end{aligned}$$

pois $\mathcal{B}$ é ortogonal e $u_1, u_2, \dots, u_n \in \mathcal{B}$. Como u_1 é não nulo, obtemos

$$b_1 = \frac{\langle v, u_1 \rangle}{||u_1||^2}.$$

Repetindo o mesmo argumento com o produto interno $\langle v, u_i \rangle$, para $i = 2, 3, \dots, n$ obtemos

$$b_i = \frac{\langle v, u_i \rangle}{||u_i||^2}, \quad \text{para } i = 1, 2, \dots, n,$$

como queríamos mostrar.

■

Corolário 5.14. *Seja V um espaço vetorial sobre K com produto interno e seja $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ uma base ortonormal de V. Então, $\forall\ v \in V$,*

$$v = \langle v, u_1 \rangle\ u_1 + \langle v, u_2 \rangle\ u_2 + \cdots + \langle v, u_n \rangle\ u_n.$$

Nosso próximo objetivo é mostrar que todo espaço vetorial de dimensão finita sobre K, onde $K = \mathbb{R}$ ou $K = \mathbb{C}$, possui bases ortonormais. A seguir apresentamos o primeiro passo para demonstrarmos este resultado, que é ortogonalizar uma base qualquer do espaço.

Processo de Ortogonalização de Gram-Schmidt: Seja V um espaço vetorial sobre K com produto interno $\langle\ ,\rangle$ e seja $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ um subconjunto L.I. de V. A partir de $\mathcal{B}$ vamos construir um conjunto $\mathcal{B}' = \{w_1, w_2, \dots, w_n\} \subset V$ que seja ortogonal e tal que o subespaço gerado por $\mathcal{B}'$ seja igual ao subespaço gerado por $\mathcal{B}$, ou seja, $[w_1, w_2, \dots, w_n] = [v_1, v_2, \dots, v_n]$ e $w_i \perp w_j$, $\forall i \neq j$. Esta construção é feita indutivamente como segue. Defina

- $w_1 := v_1$
- $w_2 := v_2 - \frac{\langle v_2, w_1 \rangle}{||w_1||^2} w_1$

Note que $w_2 \neq 0$, pois v_1 e v_2 são L.I.. E $w_1 \perp w_2$: de fato,

$$\begin{aligned} \langle w_2, w_1 \rangle = \langle v_2 - \tfrac{\langle v_2, w_1 \rangle}{||w_1||^2} w_1, w_1 \rangle &= \langle v_2, w_1 \rangle - \tfrac{\langle v_2, w_1 \rangle}{||w_1||^2} \langle w_1, w_1 \rangle \\ &= \langle v_2, w_1 \rangle - \tfrac{\langle v_2, w_1 \rangle}{||w_1||^2} ||w_1||^2 \\ &= \langle v_2, w_1 \rangle - \langle v_2, w_1 \rangle = 0 \end{aligned}$$

- definidos $w_1, w_2, \ldots, w_k$, com $1 < k < n$, definimos w_n por

$$w_n := v_n - \frac{\langle v_n, w_1 \rangle}{||w_1||^2} w_1 - \frac{\langle v_n, w_2 \rangle}{||w_2||^2} w_2 - \cdots - \frac{\langle v_n, w_{n-1} \rangle}{||w_{n-1}||^2} w_{n-1} = v_n - \sum_{i=1}^{n-1} \frac{\langle v_n, w_i \rangle}{||w_i||^2} w_i.$$

Exercício: Verifique que w_n é ortogonal a $w_1, w_2, \ldots, w_k$, com $1 \leq k < n$.

Com isso, $\mathcal{B}' = \{w_1, w_2, \ldots, w_n\}$ é um conjunto ortogonal, em particular, linearmente independente. Como cada w_i é combinação linear dos v_i's, segue que $\mathcal{B}'$ gera o mesmo subespaço que $\mathcal{B}$.

Exemplo 5.15. *Seja $V = \mathbb{C}^3$, com o produto interno canônico. Vamos encontrar uma base ortogonal de $\mathbb{C}^3$ contendo o vetor $(1, 2i, 0)$.*

Primeiro vamos tomar uma base qualquer de $\mathbb{C}^3$ contendo o vetor $(1, 2i, 0)$, como por exemplo

$$\mathcal{B} = \{v_1 = (1, 2i, 0)\ ,\ v_2 = (0, 1, 0)\ ,\ v_3 = (0, 0, 1)\}.$$

Agora vamos aplicar Gram-Schmidt para ortogonalizar $\mathcal{B}$ e obtermos a base desejada. Defina

$$w_1 := v_1 = (1, 2i, 0)$$

Então $||w_1||^2 = \langle w_1, w_1 \rangle = \langle (1, 2i, 0), (1, 2i, 0) \rangle = 1 + 2i \cdot \overline{2i} = 1 + 4 = 5$

$$w_2 := v_2 - \frac{\langle v_2, w_1 \rangle}{||w_1||^2} w_1 = (0, 1, 0) + \frac{2i}{5}(1, 2i, 0) = (\tfrac{2i}{5}, \tfrac{1}{5}, 0)$$

Então $||w_2||^2 = \frac{4}{25} + \frac{1}{25} = \frac{1}{5}$.

$$w_3 := v_3 - \frac{\langle v_3, w_1 \rangle}{||w_1||^2} w_1 - \frac{\langle v_3, w_2 \rangle}{||w_2||^2} w_2 = (0, 0, 1) + 0(1, 2i, 0) + 0(\tfrac{2i}{5}, \tfrac{1}{5}, 0) = (0, 0, 1)$$

Portanto $\mathcal{B}' = \{w_1 = (1, 2i, 0)\ ,\ w_2 = (\frac{2i}{5}, \frac{1}{5}, 0)\ ,\ w_3 = (0, 0, 1)\}$ é uma base ortogonal de $\mathbb{C}^3$ contendo o vetor $(1, 2i, 0)$.

Com o Processo de Ortogonalização de Gram-Schmidt em mãos, estamos prontos para provar que todo espaço vetorial não nulo de dimensão finita possui base ortonormal.

Teorema 5.16. *Todo espaço vetorial de dimensão finita $n \geq 1$ sobre K com produto interno possui uma base ortonormal.*

Demonstração. Seja $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ uma base qualquer de V. Aplicando Gram-Schmidt, obtemos uma base ortogonal $\mathcal{B}' = \{w_1, w_2, \ldots, w_n\}$ de V. Por fim, tomamos $\mathcal{B}'' = \{\frac{w_1}{||w_1||}, \frac{w_2}{||w_2||}, \ldots, \frac{w_n}{||w_n||}\}$, que é uma base ortonormal de V.

■

Exemplo 5.17. *Considerando a base ortonormal* $\mathcal{B}' = \{(1, 2i, 0), (\frac{2i}{5}, \frac{1}{5}, 0), (0, 0, 1)\}$ *de* $\mathbb{C}^3$, *obtida no exemplo anterior, temos que, normalizando os vetores de* $\mathcal{B}'$,

$$\mathcal{B}'' = \{(\tfrac{1}{\sqrt{5}}, \tfrac{2i}{\sqrt{5}})\ ,\ (\tfrac{2\sqrt{5}i}{5}, \tfrac{\sqrt{5}}{5}, 0)\ ,\ (0, 0, 1)\}$$

é uma base ortonormal de $\mathbb{C}^3$.

Observação 5.18. *É interessante observar que trabalhar com bases ortonormais faz com que o produto interno possa ser descrito em termos das coordenadas dos vetores.*

Mais precisamente, seja $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ *uma base ortonormal de um espaço vetorial* V *com produto interno. Então, dados* $u, v \in V$, *temos que*

$$u = \sum_{i=1}^{n} \alpha_i u_i \qquad \text{e} \qquad v = \sum_{i=1}^{n} \beta_i u_i, \qquad \alpha_i, \beta_i \in K.$$

Assim,

$$\langle u, v\rangle = \langle \sum_{i=1}^{n} \alpha_i u_i\ ,\ \sum_{j=1}^{n} \beta_j u_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \overline{\beta_j} \langle u_i, u_j\rangle \overset{(*)}{=} \sum_{i=1}^{n} \alpha_i \overline{\beta_i},$$

onde em (*) estamos usando o fato de que $\mathcal{B}$ *é ortonormal.*

Logo, para um produto interno qualquer $\langle\ ,\rangle$ *em* V,

$$\langle u, v\rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}$$

é dado em termos das coordenadas de u *e de* v *na base ortonormal* $\mathcal{B}$. *Além disso,*

$$\alpha_i = \langle u, u_i\rangle \qquad \text{e} \qquad \beta_i = \langle v, u_i\rangle.$$

Uma outra consequência interessante da existência de bases ortonormais é a seguinte.

Corolário 5.19. *Seja* V *um espaço vetorial sobre* K *com produto interno. Sejam* $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ *e* $\mathcal{B}' = \{v_1, v_2, \ldots, v_n\}$ *duas bases ortonormais de* V. *Então, se* $M = [I]^{\mathcal{B}}_{\mathcal{B}'}$ *é a matriz de mudança de base, temos que*

$$M \cdot \overline{M^t} = \overline{M^t} \cdot M = I_n,$$

ou seja,

$$M^{-1} = \overline{M^t},$$

a inversa de M *é sua trasposta conjugada.*

Demonstração. Seja $M = [I]^{\mathcal{B}}_{\mathcal{B}'} = (a_{ij}) \in \mathcal{M}_n(K)$. Como cada coluna j de M é formada pelas coordenadas de u_j na base $\mathcal{B}'$, temos que

$$u_j = \sum_{k=1}^{n} a_{kj} v_k, \qquad j = 1, 2, \ldots, n.$$

Como $\mathcal{B}'$ é ortonormal, temos que $\langle u_i, u_j\rangle = \delta_{ij}$. Então, como $\mathcal{B}'$ também é ortonormal, da Observação 5.18 obtemos que

$$\delta_{ij} = \langle u_i, u_j\rangle = \langle\sum_{k=1}^{n} a_{ki}v_k \ , \ \sum_{k=1}^{n} a_{kj}v_k\rangle = \sum_{k=1}^{n} a_{ki}\overline{a_{kj}} = \sum_{k=1}^{n} \overline{a_{kj}}a_{ki}$$

para cada $i, j = 1, 2, \ldots, n$. O somatório da direita é exatamente a entrada ji da matriz produto $\overline{M^t} \cdot M$. Logo, $\overline{M^t} \cdot M = I_n$.

■

Complemento Ortogonal: Seja V um espaço vetorial com produto interno e W um subespaço de V. Considere

$$W^{\perp} = \{v \in V \mid \langle v, w\rangle = 0, \ \forall w \in W\}.$$

Temos que $W^{\perp}$ é subespaço de V. De fato, claramente $0 \in W^{\perp}$; e se $u, v \in W^{\perp}$ e $k \in K$, então $\langle ku + v, w\rangle = k\langle u, w\rangle + \langle v, w\rangle = k0 + 0 = 0$. Logo $ku + v \in W^{\perp}$. Além disso, temos o seguinte.

Proposição 5.20. *Seja V um espaço vetorial de dimensão finita $n \geq 1$ sobre K com produto interno e seja W um subespaço próprio de V. Então*

$$V = W \oplus W^{\perp}.$$

Em particular, $\dim V = \dim W + \dim W^{\perp}$.

Demonstração. Se W=$\{0\}$, não há o que provar. Vamos assumir então que $W \neq \{0\}$. Temos que mostrar que

$$W \cap W^{\perp} = \{0\} \quad \text{e} \quad V = W + W^{\perp}.$$

Primeiro, $W \cap W^{\perp} = \{0\}$, pois se $w \in W \cap W^{\perp}$ então w é ortogonal a si mesmo, ou seja, $\langle w, w\rangle = 0$; logo $w = 0$.

Agora, para mostrar que $V = W + W^{\perp}$, seja $\mathcal{B} = \{w_1, w_2, \ldots, w_m\}$ uma base ortonormal de W, que existe pelo Teorema 5.16. E considere uma base ortonormal $\mathcal{B}' = \{w_1, w_2, \ldots, w_m, v_{m+1}, \ldots, v_n\}$ de V contendo $\mathcal{B}$ (basta complementar $\mathcal{B}$ até uma base de V e depois ortonormalizar os vetores aplicando Gram-Schmidt). Então, pelo Corolário 5.14, cada $v \in V$ é escrito como

$$v = \sum_{i=1}^{m}\langle v, w_i\rangle w_i + \sum_{i=m+1}^{n}\langle v, v_i\rangle v_i. \tag{14}$$

Temos que $\sum_{i=1}^{m}\langle v, w_i\rangle w_i \in W$, pois é combinação linear de $\mathcal{B}$.

E temos que $\sum_{i=m+1}^{n}\langle v, v_i\rangle v_i \in W^{\perp}$. De fato, dado $w \in W$, temos

$$\langle \sum_{i=m+1}^{n}\langle v, v_i\rangle v_i \ , \ w \ \rangle = \sum_{i=m+1}^{n}\langle v, v_i\rangle\langle v_i, w\rangle = 0,$$

pois, como $w \in W$, temos que w é uma combinação linear de $w_1, w_2, \ldots, w_m$, que são ortogonais a $v_{m+1}, \ldots, v_n$, logo w é ortogonal a $v_{m+1}, \ldots, v_n$ e portanto cada $\langle v_i, w\rangle$ é igual a 0.

Então, de (14), temos que cada $v \in V$ pertence a $W + W^{\perp}$, ou seja, $V = W + W^{\perp}$.

■

5.3 Funcionais Lineares e Adjuntos

Definição 5.21. *Seja V um espaço vetorial sobre K. Um **funcional linear em** V é qualquer transformação linear $f : V \to K$ de V no corpo K.*

Seja V um espaço vetorial V sobre K. Vimos na Seção 4.2 que o conjunto $L(V, K)$ de todos os funcionais lineares em V forma um espaço vetorial sobre K.

Notação: Denotamos o espaço $L(V, K)$ por V^* e o chamamos de ***espaço dual a*** V.

Pelo Teorema 4.20, temos que $\dim_K L(V, K) = \dim_K V \cdot \dim_K K = \dim_K V$. Assim,

$$\dim_K V^* = \dim_K V.$$

Agora seja V um espaço vetorial sobre K com produto interno $\langle\ ,\rangle$. Dado $w \in V$, temos que w define um funcional linear em V^* da seguinte maneira:

$$\begin{array}{rccl} f_w : & V & \to & K \\ & v & \mapsto & f_w(v) = \langle v, w\rangle \end{array}$$

Pelas propriedades (P1) e (P2) da definição e produto interno, mostra-se facilmente que f_w é, de fato, linear. O interessante é que, em espaços vetoriais com produto interno, todos os funcionais lineares são desse tipo, como veremos a seguir. Antes disso, vamos observar o seguinte: Dados um espaço vetorial V sobre K com produto interno $\langle\ ,\rangle$ e uma base ortonormal $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ de V, cada $w \in V$ pode ser escrito como $w = \sum_{i=1}^{n} \alpha_i u_i$, com $\alpha_i \in K$. Considerando, agora, o funcional linear f_w determinado por w, obtemos que

$$f_w(u_j) = \langle u_j, w\rangle = \langle u_j, \sum_{i=1}^{n} \alpha_i u_i\rangle = \sum_{i=1}^{n} \overline{\alpha_i}\langle u_j, u_i\rangle = \overline{\alpha_j}.$$

Com isso, obtemos que

$$w = \sum_{j=1}^{n} \overline{f_w(u_j)}\, u_j \tag{15}$$

Essa expressão será usada na demonstração do resultado a seguir.

Proposição 5.22. *Seja V um espaço vetorial sobre K com produto interno e de dimensão finita $n \geq 1$. Se $f \in V^*$, então existe um único $w \in V$ tal que $f(v) = \langle v, w\rangle$, para todo $v \in V$.*

Demonstração. (Existência) Pelo Teorema 5.16, podemos tomar uma base ortonormal $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ de V. Queremos mostrar que existe $w \in V$ tal que $f = f_w$, ou seja, que $f(v) = f_w(v)$, $\forall v \in V$. Para isso, tome

$$w = \sum_{j=1}^{n} \overline{f(u_j)} u_j.$$

Então, para $k = 1, 2, \ldots, n$, temos que

$$f_w(u_k) = \langle u_k, w \rangle = \langle u_k, \sum_{j=1}^{n} \overline{f(u_j)} u_j \rangle = \sum_{j=1}^{n} \overline{\overline{f(u_j)}} \langle u_k, u_j \rangle = f(u_k).$$

Como f e f_w coincidem nos elementos de uma base, seque que $f = f_w$, como queríamos.

(Unicidade) Se $w_1, w_2 \in V$ são tais que $f = f_{w_1} = f_{w_2}$, ou seja, $f(v) = \langle v, w_1 \rangle$ e $f(v) = \langle v, w_2 \rangle$, para todo $v \in V$, então $\langle v, w_1 \rangle = \langle v, w_2 \rangle$, para todo $v \in V$. Assim, $\langle v, w_1 - w_2 \rangle = 0$, para todo $v \in V$. Mas isso implica que $w_1 - w_2 = 0$, ou seja, $w_1 = w_2$.

∎

Nota: Quando V tem dimensão infinita, proposição acima não é verdadeira em geral. Entretanto, ela é verdadeira para funcionais lineares contínuos sobre Espaços de Hilbert de dimensão infinita, resultado conhecido como Teorema de Riesz.

A seguir apresentamos uma importante consequência da proposição acima.

Teorema 5.23. *Seja V um espaço vetorial de dimensão finita sobre K com produto interno. Se $T \in L(V, V)$, então existe um único operador linear $T^* \in L(V, V)$ tal que*

$$\langle T(u), v \rangle = \langle u, T^*(v) \rangle, \qquad \textit{para todos } u, v \in V.$$

Demonstração. Para cada $v \in V$, queremos construir sua imagem $T^*(v) \in V$. Para isso, considere

$$\begin{array}{rccl} f: & V & \rightarrow & K \\ & u & \mapsto & f(u) = \langle T(u), v \rangle. \end{array} \tag{16}$$

Temos que f é um funcional linear, pois para quaisquer $u_1, u_2 \in V$ e $k \in K$ temos que

$$\begin{aligned} f(ku_1 + u_2) = \langle T(ku_1 + u_2), v \rangle &= \langle kT(u_1) + T(u_2), v \rangle \\ &= k\langle T(u_1), v \rangle + \langle T(u_2), v \rangle = kf(u_1) + f(u_2). \end{aligned}$$

Então, pela Proposição 5.22, existe um único $w \in V$ tal que $f(u) = \langle u, w \rangle$, para todo $u \in V$, ou seja, tal que

$$\langle T(u), v \rangle = \langle u, w \rangle, \qquad \text{para todo } \ u \in V.$$

Como w é determinado de modo único por v, definimos $T^*(v) = w$. Com isso, temos pela construção que

$$\langle T(u), v \rangle = \langle u, T^*(v) \rangle, \qquad \text{para todos } \ u, v \in V.$$

Resta mostrar que T^* assim definida é linear. De fato, dados $u, v_1, v_2 \in V$ e $k \in K$, temos que

$$\begin{aligned}\langle u, T^*(kv_1 + v_2)\rangle = \langle T(u), kv_1 + v_2\rangle &= \overline{k}\langle T(u), v_1\rangle + \langle T(u), v_2\rangle \\ &= \overline{k}\langle u, T^*(v_1)\rangle + \langle u, T^*(v_2)\rangle = \langle u, kT^*(v_1) + T^*(v_2)\rangle.\end{aligned}$$

Segue então que a diferença entre os dois lados da igualdade acima é igual a zero, ou seja,

$$\langle u, kT^*(v_1) + T^*(v_2) - T^*(kv_1 + v_2)\rangle = 0, \quad \text{para todo } \ u \in V.$$

Assim, segue que $kT^*(v_1)+T^*(v_2)-T^*(kv_1+v_2) = 0$, ou seja, $T^*(kv_1+v_2) = kT^*(v_1)+T^*(v_2)$. Portanto T^* é linear.

A unicidade de T^* segue da construção. De fato, se $S \in L(V, V)$ é tal que $\langle T(u), v\rangle = \langle u, S(v)\rangle$, para todos $u, v \in V$, então

$$\langle T(u), v\rangle = \langle u, T^*(v)\rangle \quad \text{e} \quad \langle T(u), v\rangle = \langle u, S(v)\rangle, \quad \forall u, v \in V.$$

Logo

$$\langle u, T^*(v)\rangle = \langle u, S(v)\rangle, \quad \forall u, v \in V,$$

de onde obtemos $S(v) = T^*(v)$, para todo $v \in V$, ou seja, $S = T^*$.

■

O operador linear T^* do teorema acima recebe um nome especial.

Definição 5.24. *Seja V um espaço vetorial sobre K com produto interno e $T \in L(V, V)$. Dizemos que T* ***possui um operador adjunto*** *quando existe um operador linear $T^* \in L(V, V)$ tal que*

$$\langle T(u), v\rangle = \langle u, T^*(v)\rangle, \quad \textit{para todos } u, v \in V.$$

Neste caso, dizemos que T^* é ***o adjunto*** *de T.*

Observação 5.25. a) *O adjunto de T, quando existe, depende do produto interno considerado.*

b) *Fixados uma base ortonormal $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ de V e $T \in L(V, V)$, podemos dar a seguinte fórmula explícita para T^*:*

$T^*(v)$ construído na demostração do Teorema 5.23 é $w \in V$ tal que o funcional linear f em (16) é f_w. Em (15) vimos que

$$w = \sum_{j=1}^{n} \overline{f_w(u_j)} u_j,$$

assim, como $f_w(u_j) = f(u_j) = \langle T(u_j), v\rangle$, obtemos

$$T^*(v) = \sum_{j=1}^{n} \overline{\langle T(u_j), v\rangle} u_j, \quad \forall v \in V.$$

c) *No Teorema 5.23 garantimos que se V é de dimensão finita, então todo operador linear $T \in L(V,V)$ possui adjunto. Isso não é verdade, em geral, para espaços de dimensão infinita. Neste caso, o resultado vale para operadores lineares contínuos em espaços de Hilbert.*

Exemplo 5.26. *Considere o espaço vetorial $V = \mathcal{M}_n(\mathbb{C})$ com o produto interno (verifique!) dado por*

$$\langle A, B\rangle = tr(\overline{B}^{\,t} A).$$

Dada uma matriz $M \in \mathcal{M}_n(\mathbb{C})$, defina o operador linear

$$\begin{array}{rccl} T_M : & V & \to & V \\ & A & \mapsto & T_M(A) = MA \end{array}$$

Queremos descrever o adjunto $(T_M)^*$, que sabemos que existe pelo Teorema 5.23. Para isso, vamos calcular $\langle T_M(A), B\rangle$:

$$\langle T_M(A), B\rangle = \langle MA, B\rangle = tr(\overline{B}^{\,t}(MA)) = tr((\overline{B}^{\,t}M)A) = tr(\overline{(\overline{M}^{\,t}B)^t}A) = \langle A, \overline{M}^{\,t}B\rangle$$

Como $\langle T_M(A), B\rangle = \langle A, \overline{M}^{\,t}B\rangle$, para todo $A \in V$, temos que, pela unicidade do operador adjunto, $(T_M)^*(B) = \overline{M}^{\,t}B$.

A seguir algumas propriedades algébricas envolvendo operadores adjuntos.

Proposição 5.27. *Seja V um espaço vetorial sobre K com produto interno. Sejam $T, S \in L(V,V)$ com operadores adjuntos T^* e S^*, respectivamente, e $\lambda \in K$. Então*

a) *$T + S$ tem adjunto e $(T+S)^* = T^* + S^*$;*

b) *λT tem adjunto e $(\lambda T)^* = \overline{\lambda}\, T^*$;*

c) *$T \circ S$ tem adjunto e $(T \circ S)^* = S^* \circ T^*$;*

d) *T^* tem adjunto e $(T^*)^* = T$.*

Demonstração. Sejam $u, v \in V$. Então

$$\begin{aligned} \langle (T+S)(u), v\rangle = \langle T(u) + S(u), v\rangle &= \langle T(u), v\rangle + \langle S(u), v\rangle \\ &= \langle u, T^*(v)\rangle + \langle u, S^*(v)\rangle = \langle u, (T^* + S^*)(v)\rangle \end{aligned}$$

de onde obtemos a);

$$\langle (\lambda T)(u), v\rangle = \langle \lambda T(u), v\rangle = \lambda\langle T(u), v\rangle = \lambda\langle u, T^*(v)\rangle = \langle u, \overline{\lambda}T^*(v)\rangle = \langle u, (\overline{\lambda}T^*)(v)\rangle$$

de onde obtemos b);

$$\langle (T \circ S)(u), v \rangle = \langle T(S(u)), v \rangle = \langle S(u), T^*(v) \rangle = \langle u, S^*(T^*(v)) \rangle = \langle u, (S^* \circ T^*)(v) \rangle$$

de onde obtemos c); e por fim

$$\langle T^*(u), v \rangle = \overline{\langle v, T^*(u) \rangle} = \overline{\langle T(v), u \rangle} = \langle u, T(v) \rangle$$

de onde obtemos d).

■

Sabemos (mostramos) que, quando V é um espaço vetorial de dimensão finita sobre K, qualquer operador linear $T \in L(V, V)$ tem adjunto. E, quando V tem produto interno, vimos na Observação 5.25 b) uma fórmula explícita para T^* quando fixamos uma base ortonormal de V. Nos dois próximos resultados veremos como são as matrizes de T e T^* em uma base ortonormal.

Proposição 5.28. *Seja V um espaço vetorial de dimensão finita sobre K com produto interno. Sejam $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ uma base ortonormal de V e $T \in L(V, V)$. Então*

$$[T]_{\mathcal{B}} = (a_{ij}), \quad \textit{onde} \;\; a_{ij} = \langle T(u_j), u_i \rangle, \quad \forall i, j.$$

Demonstração. Por definição, a coluna j de $[T]_{\mathcal{B}} = (a_{ij})$ é formada pelas coordenadas de $T(u_j)$ na base $\mathcal{B}$.

Agora, como $\mathcal{B}$ é uma base ortonormal, pelo Corolário 5.14, para todo $v \in V$,

$$v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i.$$

Em particular,

$$T(u_j) = \sum_{i=1}^{n} \langle T(u_j), u_i \rangle u_i, \quad \text{para} \;\; j = 1, 2, \ldots, n.$$

Portanto, temos que $a_{ij} = \langle T(u_j), u_i \rangle$. ■

Teorema 5.29. *Seja V um espaço vetorial de dimensão finita sobre K com produto interno e $T \in L(V, V)$. Em relação a qualquer base ortonormal de V,* ***a matriz de T^* é a matriz transposta conjugada da matriz de T****.*

Demonstração. Seja $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ uma base ortonormal de V e sejam

$$[T]_{\mathcal{B}} = (a_{ij}) \qquad \text{e} \qquad [T^*]_{\mathcal{B}} = (b_{ij}).$$

Sabemos da Proposição 5.28 que

$$a_{ij} = \langle T(u_j), u_i \rangle \quad \text{e} \quad b_{ij} = \langle T^*(u_j), u_i \rangle,$$

para todos $i, j = 1, 2, \ldots, n$. Agora, pela definição de T^*, obtemos que

$$b_{ij} = \langle T^*(u_j), u_i \rangle = \overline{\langle u_i, T^*(u_j) \rangle} = \overline{\langle T(u_i), u_j \rangle} = \overline{a_{ji}}.$$

Portanto

$$[T^*]_{\mathcal{B}} = \overline{[T]}^t_{\mathcal{B}}.$$

■

Exemplo 5.30. *Considere $T : \mathbb{C}^3 \to \mathbb{C}^3$, definido por*

$$T(z_1, z_2, z_3) = (z_1 + 2z_2 \ , \ iz_3 \ , \ z_2 - iz_3).$$

Na base canônica $\mathcal{C}$ de $\mathbb{C}^3$, temos

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

Como a base canônica é ortonormal, pelo Teorema 5.29, temos que

$$[T^*]_{\mathcal{C}} = \overline{[T]}^t_{\mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 0 & -i & i \end{pmatrix},$$

de onde obtemos que $T^*(z_1, z_2, z_3) = (z_1 \ , \ 2z_1 + z_3 \ , -iz_2 + iz_3)$.

Agora observe que, na base $\mathcal{B} = \{(1,1,1), (0,1,1), (0,0,1)\}$, que não é ortonormal, temos

$$[T]_{\mathcal{B}} = \begin{pmatrix} 3 & 2 & 0 \\ i-3 & i-2 & i \\ 1-2i & 1-2i & -2i \end{pmatrix} \quad e \quad [T^*]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 1 \\ -3 & -1 & i-1 \end{pmatrix} \neq \overline{[T]}^t_{\mathcal{B}}.$$

Com isso, gostaríamos de enfatizar a importância de se ter uma base ortonormal ao aplicarmos o Teorema 5.29.

5.4 Operadores Autoadjuntos

Definição 5.31. *Seja $T \in L(V, V)$, onde V é um espaço vetorial sobre K com produto interno. Dizemos que T é* ***autoadjunto*** *se T tem operador adjunto e $T = T^*$. No caso em que $K = \mathbb{C}$, usamos também o termo* ***hermitiano****; e no caso em que $K = \mathbb{R}$, usamos também o termo* ***simétrico****.*

Vimos no Teorema 5.29 que, em dimensão finita, todo T tem adjunto e a matriz de T^* numa base ortonormal é a transposta conjugada da matriz de T. Assim, os termos hermitiano e simétrico para $K = \mathbb{C}$ e $K = \mathbb{R}$, respectivamente, na definição acima ficam bem motivados.

Proposição 5.32. *Seja V um espaço vetorial de dimensão finita sobre K com produto interno e $T \in L(V, V)$. Então as seguintes afirmações são equivalentes:*

i) *T é autoadjunto.*

ii) $\overline{[T]}_{\mathcal{B}}^{t} = [T]_{\mathcal{B}}$ *para toda base ortonormal de V.*

iii) *Existe uma base ortonormal $\mathcal{B}$ de V tal que $\overline{[T]}_{\mathcal{B}}^{t} = [T]_{\mathcal{B}}$.*

Demonstração. Seja $\mathcal{B}$ uma base ortonormal de V. Pelo Teorema 5.29, para qualquer operador linear T em V, temos que $[T^*]_{\mathcal{B}} = \overline{[T]}_{\mathcal{B}}^{t}$.

i) $\Rightarrow$ ii): Se T é autoadjunto, então $T^* = T$. Logo, $[T]_{\mathcal{B}} = [T^*]_{\mathcal{B}} = \overline{[T]}_{\mathcal{B}}^{t}$.

ii) $\Rightarrow$ iii): imediato.

iii) $\Rightarrow$ i): Se $\overline{[T]}_{\mathcal{B}}^{t} = [T]_{\mathcal{B}}$, para alguma base ortonormal $\mathcal{B}$ de V, então $[T^*]_{\mathcal{B}} = \overline{[T]}_{\mathcal{B}}^{t} = [T]_{\mathcal{B}}$, logo $T^* = T$, ou seja, T é autoadjunto.

∎

Uma osbervação imediata da proposição acima, nos dá o seguinte.

Corolário 5.33. *Seja V um espaço vetorial de dimensão finita sobre K com produto interno, $T \in L(V,V)$ e $\mathcal{B}$ uma base ortonormal de V. Então T é autoadjunto se, e somente se, $[T]_{\mathcal{B}} = (a_{ij})$, com $a_{ij} = \overline{a_{ji}}$, $\forall i,j$. Neste caso, em particular, os elementos da diagonal de $[T]_{\mathcal{B}}$ são números reais.*

Exemplo 5.34. *Considere $\mathbb{C}^2$ com o produto interno canônico e seja $T : \mathbb{C}^2 \to \mathbb{C}^2$ dado por*

$$T(z,w) = (\ 2z + (1+i)w\ ,\ (1-i)z + 3w\ ).$$

Na base canônica $\mathcal{C}$ de $\mathbb{C}^2$,

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

Como $\mathcal{C}$ é base ortonormal e $[T]_{\mathcal{C}} = \overline{[T]}_{\mathcal{C}}^{t}$, temos que T é autoadjunto.

Exemplo 5.35. *Considere $\mathbb{C}^2$ com o produto interno canônico, $\mathcal{B} = \{(1,0),(1,i)\}$ e $T : \mathbb{C}^2 \to \mathbb{C}^2$ tal que*

$$[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 2-i \\ 2+i & -2 \end{pmatrix}.$$

Temos que $[T]_{\mathcal{B}} = \overline{[T]}_{\mathcal{B}}^{t}$, mas $\mathcal{B}$ não é ortonormal, então não podemos afirmar que T é autoadjunto. E, de fato, T não é autoadjunto: a matriz de T na base canônica $\mathcal{C}$ é (mostre!)

$$[T]_{\mathcal{C}} = \begin{pmatrix} 3+i & -2+3i \\ -1+2i & -4-i \end{pmatrix}.$$

Como $\mathcal{C}$ é base ortonormal e $[T]_{\mathcal{C}} \neq \overline{[T]}_{\mathcal{C}}^{t}$, temos que T não é autoadjunto.

Para espaços vetoriais sobre $\mathbb{C}$, temos as seguintes propriedades.

Lema 5.36. *Seja V um espaço vetorial sobre $\mathbb{C}$ com produto interno e $T \in L(V,V)$. Então são equivalente:*

i) $T = 0$;

ii) $\langle T(u), u\rangle = 0$, *para todo* $u \in V$;

iii) $\langle T(u), v\rangle = 0$, *para todos* $u, v \in V$.

Demonstração. Que i) implica em ii) é óbvio. Agora, suponha que vale ii). Então, para quaisquer $u, v \in V$, temos

$$\langle T(u+v), u+v\rangle = 0.$$

Desenvolvendo o produto interno acima, obtemos

$$\begin{aligned} \langle T(u), u\rangle + \langle T(u), v\rangle + \langle T(v), u\rangle + \langle T(v), v\rangle = 0 \\ \Rightarrow \langle T(u), v\rangle + \langle T(v), u\rangle = 0 \qquad (*) \end{aligned}$$

Analogamente, de $\langle T(iu+v), iu+v\rangle = 0$ obtemos

$$\langle T(iu), v\rangle + \langle T(v), iu\rangle = 0$$

$$\Rightarrow i\langle T(u), v\rangle - i\langle T(v), u\rangle = 0 \qquad (**)$$

Multiplicando (*) por i e somando com (**), segue que $2i\langle T(u), v\rangle = 0$, para todos $u, v \in V$. Portanto iii) vale.

Por fim, supondo que iii) vale, tome $u \in V$ qualquer e $v = T(u)$. Assim, por hipótese, temos $\langle T(u), T(u)\rangle = 0$. Logo, $T(u) = 0$ para todo $u \in V$. Portanto, i) vale.

∎

Observação 5.37. *Para espaços vetoriais sobre* $\mathbb{R}$, *a equivalência de* i) *e* iii) *no Lema acima ainda é verdadeira. De fato, que* i) *implica em* iii) *é óbvio e note que, na demonstração, o mesmo argumento mostra que* iii) *implica em* i). *E* iii) *implica em* ii) *trivialmente.*

Mas a implicação ii)⇒iii) *não é verdadeira. Por exemplo, considere* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *dado por* $T(x, y) = (-y, x)$. *Considerando o produto interno canônico em* $\mathbb{R}^2$, *temos*

$$\langle T(x,y), (x,y)\rangle = \langle (-y,x), (x,y)\rangle = -yx + xy = 0, \quad \forall (x,y) \in \mathbb{R}^2,$$

mas $\langle T(1,0), (0,1)\rangle = \langle (0,1), (0,1)\rangle = 1$.

Proposição 5.38. *Seja V um espaço vetorial sobre $\mathbb{C}$ com produto interno e $T \in L(V,V)$. Então T é um operador hermitiano (auto-adjunto) se, e somente se,*

$$\langle T(v), v\rangle \in \mathbb{R}, \quad \textit{para todo} \;\; v \in V.$$

Demonstração. Se T é hermitiano, então $T^* = T$. Com isso,

$$\langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, T(v)\rangle = \overline{\langle T(v), v\rangle},$$

logo, $\langle T(v), v\rangle \in \mathbb{R}$.

Reciprocamente, se $\langle T(v), v\rangle \in \mathbb{R}$, para todo $v \in V$, então

$$\langle T(v), v\rangle = \langle v, T(v)\rangle, \quad \text{para todo } \; v \in V.$$

Então T satizfaz a definição de adjunto de T, ou seja, T possui adjunto e $T^* = T$.

■

Note que a proposição acima não é verdadeira para espaços vetoriais sobre $\mathbb{R}$, pois neste caso $\langle T(v), v\rangle$ sempre é um número real e nem todo operador linear é simétrico.

5.5 Operadores Unitários

Definição 5.39. *Seja V um espaço vetorial sobre K com produto interno e $T \in L(V, V)$. Dizemos que T é um operador unitário quando T é um isomorfismo que preserva produto interno, ou seja,*

$$\langle u, v\rangle = \langle T(u), T(v)\rangle, \quad \forall u, v \in V.$$

Exemplo 5.40. *Seja $V = \mathcal{M}_{n\times 1}(\mathbb{C})$ com produto interno $\langle M, N\rangle = \overline{N}^t M$ e seja*

$$T : V \to V, \quad \textit{dado por} \quad T(X) = AX,$$

onde $A \in \mathcal{M}_n(\mathbb{C})$ é dada. Temos que

$$\langle X, Y\rangle = \overline{Y}^t X.$$

Calculando $\langle T(X), T(Y)\rangle$, temos que

$$\langle T(X), T(Y)\rangle = \langle AX, AY\rangle = (\overline{AY})^t (AX) = \overline{Y}^t \overline{A}^t AX$$

para qualquer $X, Y \in V$. Com isso vemos que T é unitário se, e somente se, $\overline{A}^t A = I_n$.

Em geral, temos a seguinte caracterização de operadores unitários.

Proposição 5.41. *Seja V um espaço vetorial sobre K com produto interno e $T \in L(V, V)$. Então T é unitário se, e somente se, T tem adjunto e $T^* \circ T = T \circ T^* = I$, onde I é o operador identidade.*

Demonstração. Se T é unitário, então T é isomorfismo e preserva produto interno. Então, para quaisquer $u, v \in V$, temos

$$\langle T(u), v\rangle = \langle T(u), T(T^{-1}(v))\rangle = \langle u, T^{-1}(v)\rangle.$$

Então T^{-1} é adjunto de T e, assim, $T^* \circ T = T \circ T^* = I$.

Reciprocamente, se T tem adjunto e $T^* \circ T = T \circ T^* = I$, então T é isomorfismo e $T^* = T^{-1}$. Assim, para quaisquer $u, v \in V$,

$$\langle T(u), T(v)\rangle = \langle u, T^*(T(v))\rangle = \langle u, T^{-1}(T(v))\rangle = \langle u, v\rangle.$$

Logo, T é unitário.

■

Definição 5.42. *Seja $A \in \mathcal{M}_n(K)$. Dizemos que A é* ***unitária*** *quando $\overline{A}^t A = I_n$. Quando $K = \mathbb{R}$, também dizemos que A é* ***ortogonal****.*

Note que a condição $\overline{A}^t A = I_n$ nos diz que os vetores coluna de A são ortonormais, considerando o produto interno canônico em K^n. Mais ainda, temos $A^{-1} = \overline{A}^t$ e

$$1 = \det I_n = \det(\overline{A}^t A) = \det \overline{A}^t \cdot \det A = \det \overline{A} \cdot \det A = \overline{\det A} \cdot \det A.$$

Então $|\det A| = 1$. Se $K = \mathbb{R}$, então $\det A = \pm 1$; se $K = \mathbb{C}$, então $\det A = e^{i\theta}$, θ real.

Exemplo 5.43. a) *Seja*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}).$$

Então A é ortogonal se, e somente se,

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} = A^t = A^{-1} = \frac{1}{ac - bd} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Como observado acima, temos que $\det A = \pm 1$. *Assim, A é ortogonal se, e somente se,*

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad ou \quad A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \quad onde \quad a, b \in \mathbb{R} \;\; e \;\; a^2 + b^2 = 1.$$

Agora, dados $a, b \in \mathbb{R}$ tais que $a^2 + b^2 = 1$, existe um θ ($0 \leq \theta \leq 2\pi$) tal que $a = \cos\theta$ e $b = \operatorname{sen}\theta$. Com isso, A é ortogonal se, e somente se,

$$A = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \quad ou \quad A = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix}.$$

b) *Seja*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{C}).$$

Então A é unitária se, e somente se,

$$\begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} = \overline{A}^t = A^{-1} = \frac{1}{ac - bd} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Como observado acima, temos que $\det A = e^{i\theta}$, *com θ real. Assim, A é unitária se, e somente se,*

$$A = \begin{pmatrix} a & b \\ -e^{i\theta}\bar{b} & e^{i\theta}\bar{a} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix} \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix},$$

onde $\theta \in \mathbb{R}$ e $a, b \in \mathbb{C}$, com $|a|^2 + |b|^2 = 1$.

5.6 Operadores Normais

Na Seção 4.1, estudamos condições necessárias e suficientes para um operador linear $T \in L(V,V)$ de espaço vetorial V de dimensão finita ser diagonalizável, ou seja, possuir uma base formada por autovetores de T. Mas até então não tínhamos levado em consideração informações sobre produtos internos em V.

Nesta seção, considerando que V possui produto interno, vamos discutir a existência de bases formadas por autovetores de T que são também ortonormais. Para isso, vamos observar o seguinte:

Seja V um espaço vetorial de dimensão finita sobre K ($K = \mathbb{R}$ ou $K = \mathbb{C}$) e com produto interno. Suponha que exista uma base ortonormal $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ de V formada por autovetores de T, isto é, $T(u_i) = \lambda_i u_i$, para certos $\lambda_i \in K$. Então, $[T]_\mathcal{B}$ é uma matriz diagonal com diagonal formada pelos autovalores $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Se $K = \mathbb{R}$, então $\lambda_i = \overline{\lambda}_i$, para todo i. Logo $\overline{[T]}_\mathcal{B}^t = [T]_\mathcal{B}$, portanto T é autoadjunto, pela Proposição 5.32.

Se $K = \mathbb{C}$, então T não é necessariamente autoadjunto. Mas, como $\mathcal{B}$ é uma base ortonormal, vimos no Teorema 5.29 que a matriz de T^* na base $\mathcal{B}$ é a transposta conjugada da matriz de T. Assim, vale a relação

$$[T]_\mathcal{B} \cdot [T^*]_\mathcal{B} = [T^*]_\mathcal{B} \cdot [T]_\mathcal{B},$$

ou seja, T comuta com T^*:

$$T \circ T^* = T^* \circ T.$$

Então, se V possui uma base ortonormal de autovetores de T, temos que T comuta com T^*. Vamos mostrar que a recíproca também é verdadeira. Para isso, vamos chamar operadores lineares que comutam com seu adjunto de normais.

Definição 5.44. *Seja V um espaço vetorial sobre K com produto interno e $T \in L(V,V)$. Dizemos que T é um **operador normal** se T possui adjunto e $T \circ T^* = T^* \circ T$.*

Observação 5.45. a) *Todo operador autoadjunto é normal.*

De fato, $T^* = T$ comuta com si mesmo.

b) *Todo operador unitário é normal.*

De fato, pela Proposição 5.41, operadores unitários T possuem adjunto e $T \circ T^* = T^* \circ T = I$.

c) *Todo múltiplo escalar de um operador normal também é normal.*

De fato, se T é normal e $\lambda \in K$, temos $T \circ T^* = T^* \circ T$ e, pela Proposição 5.27, temos que $(\lambda T)^* = \overline{\lambda} T^*$. Então

$$(\lambda T)^* \circ (\lambda T) = (\overline{\lambda} T^*) \circ (\lambda T) = \overline{\lambda}\lambda (T^* \circ T) = \overline{\lambda}\lambda (T \circ T^*) = (\lambda T) \circ (\overline{\lambda} T^*) = (\lambda T) \circ (\lambda T)^*.$$

d) *Se T e S são operadores normais de V, então*

$$(T+S)^* \circ (T+S) = (T^*+S^*) \circ (T+S) = T^* \circ T + T^* \circ S + S^* \circ T + S^* \circ S$$

e, por outro lado,

$$(T+S) \circ (T+S)^* = (T+S) \circ (T^*+S^*) = T \circ T^* + T \circ S^* + S \circ T^* + S \circ S^*.$$

Assim $T+S$ é normal se, e somente se, $T^* \circ S + S^* \circ T = T \circ S^* + S \circ T^*$. Com isso, temos que a soma de operadores normais não é necessariamente normal.

Proposição 5.46. *Seja V um espaço vetorial sobre K com produto interno e $T \in L(V,V)$ um operador normal. Então*

a) $||T(v)|| = ||T^*(v)||$, *para todo $v \in V$.*

b) *Se $T(v) = \lambda v$, para $v \in V$ e $\lambda \in K$, então $T^*(v) = \overline{\lambda} v$.*

c) *Se $T(v_1) = \lambda_1 v_1$ e $T(v_2) = \lambda_2 v_2$, para $v_1, v_2 \in V$ e $\lambda_1, \lambda_2 \in K$ com $\lambda_1 \neq \lambda_2$, então $\langle v_1, v_2 \rangle = 0$*

Demonstração. a) Seja $v \in V$. Pela definição de produto interno, temos que $\langle T(v), T(v) \rangle$ é um número real. Além disso, T é normal, então $T^* \circ T = T \circ T^*$. Assim,

$$\langle T(v), T(v) \rangle = \langle v, T^*(T(v)) \rangle = \langle v, T(T^*(v)) \rangle = \langle T(T^*(v)), v \rangle = \langle T^*(v), T^*(v) \rangle.$$

Logo $||T(v)|| = ||T^*(v)||$.

b) Se $T(v) = \lambda v$, então $(T - \lambda I)(v) = 0$ e assim $||(T - \lambda I)(v)|| = 0$. Pelo item a), segue que $||(T - \lambda I)^*(v)|| = 0$. Então $(T - \lambda I)^*(v) = 0$, de onde obtemos que $T^*(v) = \overline{\lambda} v$.

c) Pelo item b) temos que,

$$\langle T(v_1), v_2 \rangle = \langle v_1, T^*(v_2) \rangle = \langle v_1, \overline{\lambda}_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle.$$

Por outro lado,

$$\langle T(v_1), v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \lambda_1 \langle v_1, v_2 \rangle.$$

Então, obtemos que $\lambda_2 \langle v_1, v_2 \rangle = \lambda_1 \langle v_1, v_2 \rangle$. Como $\lambda_1 \neq \lambda_2$, segue que $\langle v_1, v_2 \rangle = 0$.

■

Corolário 5.47. *Seja V um espaço vetorial sobre K com produto interno, $T \in L(V,V)$ e $\lambda \in K$ um autovalor de T.*

a) *Se T é autoadjunto, então $\lambda \in \mathbb{R}$.*

b) *Se T é unitário, então $|\lambda| = 1$.*

Demonstração. Seja $v \in V$ um autovetor associado a λ.

a) Como $T = T^*$, segue da Proposição 5.46 b), que $\lambda = \overline{\lambda}$. Logo $\lambda \in \mathbb{R}$.

b) Como $T^* \circ T = I$, temos que $T^*(T(v)) = v$. Então,

$$T^*(T(v)) = v \Rightarrow T^*(\lambda v) = v \Rightarrow \lambda T^*(v) = v \Rightarrow \lambda\overline{\lambda}v = v.$$

Como v é autovetor, logo não nulo, segue que $\lambda\overline{\lambda} = 1$, ou seja, $|\lambda| = 1$. ∎

Teorema 5.48. *Seja V um espaço vetorial de dimensão finita $n \geq 1$ sobre K com produto interno e $T \in L(V, V)$. Se T é autoadjunto, então T possui um autovetor.*

Demonstração. Primeiramente observe que se $K = \mathbb{C}$, pelo Teorema Fundamental de Álgebra, o polinômio característico $p_T(x)$ tem raízes e elas são os autovalores de T. Logo T tem autovetor. Vamos assumir, então, que $K = \mathbb{R}$.

Seja $\mathcal{B}$ uma base ortonormal de V e $A = [T]_{\mathcal{B}}$. Como T é autoadjunto, temos que $A = \overline{A}^t = A^t$. Considere o espaço $W = \mathcal{M}_{n\times 1}(\mathbb{C})$ com produto interno $\langle X, Y\rangle = \overline{Y}^t X$ e o operador linear $S : W \to W$ dado por $S(X) = AX$. Então $\langle S(X), Y\rangle = \langle AX, Y\rangle = \overline{Y}^t(AX) = (\overline{A^tY})^t X = \langle X, A^tY\rangle = \langle X, AY\rangle$. Logo $S^* = S$, ou seja, S é autoadjunto.

Agora, temos que $p_T(x) = \det(xI - A) = p_S(x)$. Como W é um espaço vetorial sobre $\mathbb{C}$, temos que $p_S(x)$ tem uma raiz complexa λ. Mas como S é autoadjunto, pelo Corolário 5.47, $\lambda \in \mathbb{R}$. Então λ é uma raiz de $p_T(x)$ e, assim, T possui um autovetor. ∎

A hipótese de V ter dimensão finita é necessária para o teorema acima. Um operador autoadjunto sobre um espaço de dimensão infinita pode não ter autovetores, como mostra o exemplo a seguir.

Exemplo 5.49. *Seja $V = \mathcal{C}([0,1], \mathbb{C})$, o espaço das funções complexas (ou reais) contínuas, definidas no intervalo fechado $[0,1]$, com o produto interno*

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}dt.$$

O operador "multiplicação por t", dado por $T(f)(t) = tf(t)$, é autoadjunto. De fato,

$$\langle T(f), g\rangle = \int_0^1 tf(t)\overline{g(t)}dt = \int_0^1 f(t)\overline{tg(t)}dt = \langle f, T(g)\rangle, \qquad \forall f, g \in V.$$

Logo $T^* = T$.

Agora, suponha que $T(f) = \lambda f$, para algum $\lambda \in \mathbb{C}$. Então

$$(t - \lambda)f(t) = 0, \qquad t \in [0,1],$$

e, então, $f(t) = 0$ para $t \neq \lambda$. Como f é contínua, seque que $f = 0$. Logo T não possui autovetores.

Vamos usar o Teorema 5.48 acima para mostrar que se T é um operador autoadjunto em um espaço de dimensão finita V, então V possui base ortonormal de autovetores. Para isso, precisamos ainda do seguinte resultado auxiliar.

Lema 5.50. *Seja V um espaço vetorial de dimensão finita sobre K com produto interno e $T \in L(V,V)$. Se W é um subespaço T-invariante de V, então seu complemento ortogonal $W^\perp$ é T^*-invariante.*

Demonstração. Temos que mostrar que $T^*(w) \in W^\perp$, $\forall w \in W^\perp$. Para isso, para cada $w \in W^\perp$, vamos mostrar que $\langle v, T^*(w)\rangle = 0$, $\forall v \in W$. De fato, com W é T-invariante, temos $T(v) \in W$, $\forall v \in W$. Assim $\langle T(v), w\rangle = 0$. Com isso,

$$\langle v, T^*(w)\rangle = \langle T(v), w\rangle = 0.$$

■

Teorema 5.51. *Seja V um espaço vetorial de dimensão finita $n \geq 1$ sobre K com produto interno. Se $T \in L(V,V)$ é autoadjunto, então existe uma base ortonormal de V formada por autovetores de T.*

Demonstração. Pelo Teorema 5.48, T possui um autovetor $v_1 \in V$. Se $\dim V = 1$, então $\{\frac{v_1}{||v_1||}\}$ é a base desejada. Vamos supor agora que $n > 1$ e que o resultado é válido para todo espaço vetorial de dimensão $n-1$. Seja $W = [v_1]$, onde v_1 é um autovetor dado pelo Teorema 5.48. Então W é T-invariante. Logo, pelo Lema 5.50, $W^\perp$ é T^*-invariante. Como T é autoadjunto, temos $T^* = T$, assim $W^\perp$ é T-invariante e portanto a restrição $T|_{W^\perp}$ é um operador linear de $W^\perp$. E como $\dim W^\perp = n-1$ (Proposição 5.20), da hipótese de indução temos que $W^\perp$ possui uma base ortonormal $\{v_2, v_3, \ldots, v_n\}$ formada por autovetores de T. Como $V = W \oplus W^\perp$, segue que $\mathcal{B} := \{\frac{v_1}{||v_1||}, v_2, v_3, \ldots, v_n\}$ é uma base ortonormal de V formada por autovetores de T, como queríamos.

■

Corolário 5.52. i) *Seja $A \in \mathcal{M}_n(\mathbb{C})$ uma matriz hermitiana. Então existe uma matriz unitária $M \in \mathcal{M}_n(\mathbb{C})$ tal que $\overline{M}^t AM$ é diagonal.*

ii) *Seja $A \in \mathcal{M}_n(\mathbb{R})$ uma matriz simétrica. Então existe uma matriz ortogonal $M \in \mathcal{M}_n(\mathbb{R})$ tal que $M^t AM$ é diagonal.*

Demonstração. Seja $T : K^n \to K^n$ um operador linear tal que $[T]_\mathcal{C} = A$, onde $\mathcal{C}$ é a base canônica. Como A é hermitiana/simétrica e a base canônica é ortonormal, segue que T é autoadjunto (Proposição 5.32). Então, pelo Teorema 5.51, existe uma base ortonormal $\mathcal{B}$ de K^n formada por autovetores de T. Logo $[T]_\mathcal{B}$ é diagonal e

$$[T]_\mathcal{B} = [I]^\mathcal{C}_\mathcal{B}\ A\ [I]^\mathcal{B}_\mathcal{C},$$

onde $[I]^\mathcal{C}_\mathcal{B}$ e $[I]^\mathcal{B}_\mathcal{C}$ são as respectivas matrizes de mudança de base.

Considere $M = [I]^\mathcal{B}_\mathcal{C}$. Então $M^{-1} = [I]^\mathcal{C}_\mathcal{B}$ e

$$[T]_\mathcal{B} = M^{-1}\ A\ M.$$

Agora, como $\mathcal{C}$ e $\mathcal{B}$ são ortonormais, pelo Corolário 5.19, temos $M^{-1} = \overline{M}^t$. Portanto

$$[T]_{\mathcal{B}} = \overline{M}^t\ A\ M$$

e $\overline{M}^t\ A\ M$ é diagonal. ■

Observamos no início desta seção que, para espaço vetoriais V de dimensão finita sobre $\mathbb{R}$, se existe uma base ortonormal de V formada por autovetores de T, então T é autoadjunto. Isso, juntamente com o Teorema 5.51, nos diz que, para espaços vetoriais reais de dimensão finita, um operador linear T é autoadjunto se, e somente se, se existe uma base ortonormal de V formada por autovetores de T.

No contexto de espaços vetoriais de dimensão finita sobre $\mathbb{C}$, observamos no início da seção que se existe uma base ortonormal de V formada por autovetores de T, então T é normal. Finalizamos este texto com o seguinte resultado, que mostra que temos, na verdade, a equivalência destas duas condições.

Teorema 5.53. *Seja V um espaço vetorial de dimensão finita $n \geq 1$ sobre $\mathbb{C}$ com produto interno e $T \in L(V, V)$. Então T é um operador normal se, e somente se, existe uma base ortonormal de V formada por autovetores de T.*

Demonstração. Já observamos no início da seção que se existe uma base ortonormal de V formada por autovetores de T, então T é normal.

Reciprocamente, suponha que T é normal. Como V é um espaço vetorial de dimensão finita $n \geq 1$ sobre $\mathbb{C}$, T possui pelo menos um autovalor. Seja, então, $v_1 \in V$ um autovetor de T, que já vamos supor tal que $||v_1|| = 1$. E considere o subespaço $W = [v_1]$. Então W é T-invariante. Agora, pela Proposição 5.46, v_1 é também autovetor de T^*, logo W é também T^*-invariante. Com isso, pelo Lema 5.50, $W^\perp$ é T^{**}-invariante. Como $T^{**} = T$, temos que $W^\perp$ é T-invariante.

Como T é normal, a restrição $T|_{W^\perp}$ é um operador normal de $W^\perp$. Com isso, um argumento por indução sobre a dimensão, mostra que existe uma base ortonormal formada por autovetores de T.

■

Referências

[1] Flávio U. Coelho e Mary L. Lourenço, *Um curso de Álgebra Linear*, Editora EdUSP, 2a edição, 2013.

[2] Kenneth Hoffman e Ray Kunze, traduzido por Adalberto P. Bergamasco, *Álgebra Linear*, Livros Técnicos e Científicos Editora, Rio de Janeiro, 1976.