Álgebra Linear - Notas de aula (em construção)

Leonardo D. Secchin

UFES

São Mateus, 2020

Sumário

1 Matrizes **3**

1.1 Operações usuais com matrizes 4
1.2 Matrizes inversíveis 6
1.3 Demonstrações 7
1.4 Exercícios 7

2 Sistemas Lineares **10**

2.1 Operações e matrizes elementares 11
2.2 Resolução de sistemas lineares 15
2.3 Um processo para inversão de matrizes 17
2.4 Demonstrações 18
2.5 Exercícios 19
2.6 Notas 21

3 Determinantes **22**

3.1 Desenvolvimento de Laplace 23
3.1.1 Casos particulares 26
3.2 Determinantes e matrizes elementares 27
3.3 Propriedades adicionais dos determinantes 29
3.4 Demonstrações de resultados 30
3.5 Exercícios 32
3.6 Notas 34

4 Espaços Vetoriais **35**

4.1 Motivação 35
4.2 Espaços Vetoriais 35
4.3 Subespaços vetoriais 38
4.4 Combinação linear 41
4.5 Dependência e independência linear 43
4.6 Base e dimensão de um espaço vetorial 44
4.7 Mudança de base 50
4.8 Demonstrações 53
4.9 Exercícios 54

5 Transformações Lineares **58**

5.1 Motivação 58
5.2 Transformações Lineares – definição 59
5.3 Exemplos 59
5.3.1 Transformações de $\mathbb{R}^2$ em $\mathbb{R}^2$ (do "plano no plano") 59
5.3.2 Outros exemplos 63
5.4 Conceitos e teoremas 64

5.4.1 Imagem e Núcleo . . . 65
5.4.2 Injetividade e sobrejetividade . . . 66
5.5 A matriz de uma transformação linear . . . 68
5.6 Demonstrações . . . 74
5.7 Exercícios . . . 77

6 Autovalores e Autovetores 81
6.1 Matrizes . . . 83
6.2 Cálculo de autovalores – o polinômio característico . . . 85
6.2.1 Casos particulares . . . 88
6.3 Autovalores de operadores via matrizes . . . 89
6.4 Multiplicidades algébrica e geométrica . . . 90
6.5 Autopares e cônicas rotacionadas . . . 91
6.6 Exercícios . . . 95

7 Diagonalização 96
7.1 Diagonalização de operadores . . . 96
7.1.1 Bases formadas por autovetores . . . 96
7.1.2 Matrizes . . . 100
7.1.3 O polinômio minimal . . . 100
7.1.4 Diagonalização e multiplicidade geométrica – uma visão geométrica . . . 106
7.2 Diagonalização de matrizes simétricas . . . 106
7.3 Matrizes semelhantes . . . 107
7.4 Demonstrações . . . 109
7.5 Exercícios . . . 112

8 Produto Interno 114

A Dicas e respostas dos exercícios 115

Capítulo 1

Matrizes

Uma *matriz* $\mathbf{A}_{m\times n}$ *de ordem* $m \times n$ é uma tabela de números reais dispostos em m linhas e n colunas

$$\mathbf{A}_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]_{m\times n}$$

Definição 1.1. *Duas matrizes* $\mathbf{A} = [a_{ij}]_{m\times n}$ *e* $\mathbf{B} = [b_{ij}]_{r\times s}$ *são* ***iguais*** *se* $m = r$*,* $n = s$ *e* $a_{ij} = b_{ij}$ *para todos* $i = 1, \cdots m$ *e* $j = 1, \cdots, n$*.*

Dada uma matriz $\mathbf{A} = [a_{ij}]_{m\times n}$, dizemos que $\mathbf{A}$ é

- *quadrada* se $m = n$. Neste caso, podemos dizer simplesmente que $\mathbf{A}$ tem ordem n.
- *matriz nula* se $a_{ij} = 0$ para todos i, j.
- *matriz linha* se $\mathbf{A}$ tem ordem $1 \times n$ (possui apenas uma linha).
- *matriz coluna* se $\mathbf{A}$ tem ordem $m \times 1$ (possui apenas uma coluna).
- *matriz diagonal* se $\mathbf{A}$ é quadrada e $a_{ij} = 0$ sempre que $i \neq j$. Os elementos $a_{11}, a_{22}, \cdots, a_{nn}$ constituem a *diagonal principal* de $\mathbf{A}$. Por exemplo, a matriz $\mathbf{A}$ abaixo é matriz diagonal:

  $$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

  Em particular, a matriz diagonal

  $$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

  cuja diagonal principal é formada de 1's é chamada *matriz identidade (de ordem n).*

- *matriz triangular superior* se $\mathbf{A}$ é quadrada e $a_{ij} = 0$ sempre que $i > j$. Por exemplo,

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

  é matriz triangular superior.

- *matriz triangular inferior* se $\mathbf{A}$ é quadrada e $a_{ij} = 0$ sempre que $i < j$. Por exemplo,

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

  é matriz triangular inferior.

- *matriz simétrica* se $\mathbf{A}$ é quadrada e $a_{ij} = a_{ji}$ para todos i, j. Por exemplo,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 2 & -1 & -1 \end{bmatrix}$$

  é matriz simétrica.

1.1 Operações usuais com matrizes

Dadas $\mathbf{A} = [a_{ij}]_{m\times n}$ e $\mathbf{B} = [b_{ij}]_{m\times n}$ matrizes de mesma ordem $m \times n$ e $k \in \mathbb{R}$, definimos as operações com matrizes:

- a *soma* $\mathbf{A} + \mathbf{B}$ é a matriz $\mathbf{C} = [c_{ij}]_{m\times n}$ de ordem $m \times n$ tal que $c_{ij} = a_{ij} + b_{ij}$ para todos i, j.

- a *multiplicação por escalar* $k\mathbf{A}$ é a matriz $\mathbf{C} = [c_{ij}]_{m\times n}$ de ordem $m \times n$ tal que $c_{ij} = ka_{ij}$ para todos i, j.

Exemplo 1.1. Seja $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$ e $\mathbf{B} = \begin{bmatrix} 1 & -1 \\ 4 & 2 \end{bmatrix}$. Então

$$\mathbf{A} + 2\mathbf{B} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} + 2\begin{bmatrix} 1 & -1 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} + \begin{bmatrix} 2 & -2 \\ 8 & 4 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 7 & 6 \end{bmatrix}$$

□

> As operações de soma e multiplicação por escalar seguem as mesmas regras que a de números reais.

Teorema 1.1 (Propriedades da soma e da multiplicação por escalar)**.** *Dadas matrizes* $\mathbf{A}$, $\mathbf{B}$ *e* $\mathbf{C}$ *de ordem* $m \times n$ *e* $a, b \in \mathbb{R}$, *vale:*

(i) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ *(comutatividade)*

(ii) $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$

(iii) $a(b\mathbf{A}) = (ab)\mathbf{A}$

(iv) $\mathbf{A} + \mathbf{0}_{m\times n} = \mathbf{A}$, *onde* $\mathbf{0}_{m\times n}$ *é matriz nula*[1].

(v) $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$

(vi) $(a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$

[1]Denotaremos a matriz nula por **0** (em **negrito**) e o número real zero por 0.

(vii) $0\mathbf{A} = \mathbf{0}_{m\times n}$ *(a multiplicação da matriz* $\mathbf{A}$ *pelo escalar* 0 *é matriz nula)*

A *transposta* de uma matriz $\mathbf{A} = [a_{ij}]_{m\times n}$ é a matriz $\mathbf{A}^t = [a_{ji}]_{n\times m}$.

Exemplo 1.2.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \end{bmatrix}_{2\times 3} \Rightarrow \mathbf{A}^t = \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 2 \end{bmatrix}_{3\times 2}.$$

□

Teorema 1.2 (Propriedades da transposição de matrizes)**.** *Seja* $\mathbf{A}$ *uma matriz e* $a \in \mathbb{R}$*. Vale:*

(i) $\mathbf{A}$ *é simétrica se, e somente se* $\mathbf{A}^t = \mathbf{A}$.

(ii) $(\mathbf{A}^t)^t = \mathbf{A}$.

(iii) $(\mathbf{A}+\mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t$.

(iv) $(a\mathbf{A})^t = a\mathbf{A}^t$.

Agora, dadas $\mathbf{A} = [a_{ij}]_{m\times p}$ e $\mathbf{B} = [b_{ij}]_{p\times n}$ definimos a *multiplicação* $\mathbf{AB}$ (de matrizes) como a matriz $\mathbf{C} = [c_{ij}]_{m\times n}$ tal que

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}.$$

Muita atenção nas ordens das matrizes: o produto $\mathbf{AB}$ **só** é possível se o número de colunas de $\mathbf{A}$ é igual ao número de linhas de $\mathbf{B}$ (veja que a soma acima só faz sentido neste caso). Então ao multiplicar matrizes, observe as ordens:

$$\Big[m\times \mathbf{p}\Big]\ \Big[\mathbf{p}\times n\Big] = \Big[m\times n\Big]$$

Não confunda a multiplicação de uma matriz por um escalar com a multiplicação de duas matrizes. São operações diferentes!

A expressão de c_{ij} acima significa que a entrada (i,j) do produto $\mathbf{AB}$ é obtida somando os "produtos correspondentes entre a linha i de $\mathbf{A}$ e a coluna j de $\mathbf{B}$". De fato, observe na soma que aparecem os índices a_{i*} e b_{*j} (o índice k percorre a linha de $\mathbf{A}$ e a coluna de $\mathbf{B}$).

No exemplo a seguir, imagine a multplicação da linha 1 de $\mathbf{A}$ contra a coluna 1 de $\mathbf{B}$; isso fornecerá o elemento da primeira linha, primeira coluna de $\mathbf{AB}$. Faça o mesmo raciocínio para os outros elementos de $\mathbf{AB}$.

Exemplo 1.3.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \end{bmatrix}_{2\times \mathbf{3}},\quad \mathbf{B} = \begin{bmatrix} 1 & 0 & -1 & 4 \\ 2 & 2 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}_{\mathbf{3}\times 4} \Rightarrow \mathbf{AB} = \begin{bmatrix} 5 & 7 & 1 & 1 \\ -1 & 2 & 1 & -6 \end{bmatrix}_{2\times 4}.$$

□

Ao contrário das outras operações com matrizes, a multiplicação entre matrizes não se comporta exatamente como a multiplicação entre números reais. O exemplos a seguir mostra que

- **nem sempre** é verdade que $\mathbf{AB} = \mathbf{BA}$;
- **nem sempre** é verdade que se $\mathbf{AB} = \mathbf{0}$ então $\mathbf{A} = \mathbf{0}$ ou $\mathbf{B} = \mathbf{0}$.

Exemplo 1.4.

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{AB} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{BA} = \begin{bmatrix} 2 & -2 \\ 2 & -2 \end{bmatrix}.$$

□

Teorema 1.3 (Propriedades da multiplicação entre matrizes)**.** *Sejam* $\mathbf{A}$*,* $\mathbf{B}$ *e* $\mathbf{C}$ *matrizes. Desde que as operações sejam possíveis, vale:*

(*i*) $\mathbf{AI} = \mathbf{A}$ *e* $\mathbf{IA} = \mathbf{A}$

(*ii*) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ *(distributividade)*

(*iii*) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ *(distributividade)*

(*iv*) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ *(associatividade)*

(*v*) $(\mathbf{AB})^t = \mathbf{B}^t\mathbf{A}^t$

(*vi*) $\mathbf{A0} = \mathbf{0}$ *(a multiplicação da matriz* $\mathbf{A}$ *pela matriz nula é matriz nula)*

Atividade 1.1. Verifique as propriedades (*ii*) e (*v*) para matrizes $\mathbf{A}$ de ordem 2×3 e $\mathbf{B}$, $\mathbf{C}$ de ordem 3×3. Tente se convencer que vale para quaisquer ordens.

1.2 Matrizes inversíveis

Dada uma matriz quadrada $\mathbf{A}$ de ordem n, dizemos que $\mathbf{B}$ é uma *inversa* de $\mathbf{A}$ se $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$ ($\mathbf{B}$ também deve ser quadrada de ordem n). Neste caso diremos que $\mathbf{A}$ é *inversível.*

Teorema 1.4 (Unicidade da inversa)**.** *Se* $\mathbf{A}$ *admite uma inversa, ela é única.*

Devido à unicidade da inversa, denotaremos a inversa de $\mathbf{A}$ por

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

O próximo resultado é útil para verificar se uma matriz é a inversa de outra. Ele diz que para verificar que $\mathbf{B}$ é a inversa de $\mathbf{A}$, basta verificar **se alguma** das expressões vale: $\mathbf{AB} = \mathbf{I}$ ou $\mathbf{BA} = \mathbf{I}$; ou seja, não é necessário verificar as **duas**. Não daremos uma demonstração neste momento.

Teorema 1.5. *Se* $\mathbf{A}$ *é matriz quadrada e* $\mathbf{B}$ *é tal que* $\mathbf{BA} = \mathbf{I}$ *(ou* $\mathbf{AB} = \mathbf{I}$*), então* $\mathbf{A}$ *é inversível e* $\mathbf{A}^{-1} = \mathbf{B}$.

Exemplo 1.5. A matriz $\mathbf{B} = \frac{1}{2}\begin{bmatrix} -4 & 2 \\ 3 & -1 \end{bmatrix}$ é a inversa de $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. De fato,

$$\mathbf{AB} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\left(\frac{1}{2}\begin{bmatrix} -4 & 2 \\ 3 & -1 \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -4 & 2 \\ 3 & -1 \end{bmatrix} = \mathbf{I}_2.$$

□

Teorema 1.6 (Inversa do produto)**.** *Sejam* $\mathbf{A}$ *e* $\mathbf{B}$ *matrizes inversíveis de mesma ordem. Então* $\mathbf{AB}$ *é inversível, com*

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

1.3 Demonstrações

Demonstração do Teorema 1.4. Se $\mathbf{B}$ e $\mathbf{C}$ são inversas de $\mathbf{A}$ então $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$ e $\mathbf{AC} = \mathbf{CA} = \mathbf{I}$. Assim,

$$\mathbf{B} = \mathbf{BI} = \mathbf{B}(\mathbf{AC}) = (\mathbf{BA})\mathbf{C} = \mathbf{IC} = \mathbf{C}.$$

□

Demonstração do Teorema 1.6. Como $\mathbf{A}^{-1}$ e $\mathbf{B}^{-1}$ são matrizes quadradas de mesma ordem, o produto $\mathbf{B}^{-1}\mathbf{A}^{-1}$ pode ser realizado. Neste caso,

$$(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}(\mathbf{IB}) = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I},$$

e pelo Teorema 1.5 concluímos que $\mathbf{AB}$ é inversível, e sua inversa é $\mathbf{B}^{-1}\mathbf{A}^{-1}$. □

1.4 Exercícios

1. Dadas as matrizes $\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 3 & 0 \\ 2 & -2 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 2 \end{bmatrix}$ e $\mathbf{E} = \begin{bmatrix} 3 & 2 & -1 \\ -1 & 3 & 0 \\ 0 & 1 & -1 \end{bmatrix}$, faça:

   (a) $\mathbf{A} + \mathbf{B}$

   (b) $-3\mathbf{B} + 4\mathbf{B}$

   (c) $\mathbf{B}^3 = \mathbf{BBB}$

   (d) $2\mathbf{A}(\mathbf{C} - 2\mathbf{B}^t)$

   (e) $\mathbf{AB}(\mathbf{DE})$

   (f) $\mathbf{E}^t\mathbf{D}^t$

   (g) Mostre que $\mathbf{A}$ é *idempotente*, isto é, que $\mathbf{A}^2 = \mathbf{AA} = \mathbf{0}$

   (h) Mostre que $\mathbf{CB} \neq \mathbf{BC}$

2. Se $\mathbf{A}$ e $\mathbf{B}$ são matrizes $m \times n$ quaisquer, mostre que $(\mathbf{A}+\mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t$ e $(\alpha\mathbf{A})^t = \alpha\mathbf{A}^t$, onde $\alpha \in \mathbb{R}$.

3. Mostre que para qualquer matriz $\mathbf{A}$ (quadrada ou não), $\mathbf{AA}^t$ e $\mathbf{A}^t\mathbf{A}$ são matrizes simétricas.

4. Sejam $\mathbf{A}$, $\mathbf{B}$ matrizes simétricas.

   (a) Mostre que $\mathbf{A} + \mathbf{B}$ é simétrica.

   (b) Mostre que $\mathbf{AB}$ é simétrica se, e somente se, $\mathbf{BA} = \mathbf{AB}$.

   (c) Mostre que $\mathbf{A}^2 = \mathbf{AA}$ é simétrica.

(d) Mostre que $2\mathbf{A}^2 - 3\mathbf{A} + I_n$ é simétrica.

(e) Indicamos por $\mathbf{X}^m$ o produto $\mathbf{XX}\cdots\mathbf{X}$, onde X é multiplicada m vezes. Se $p(\mathbf{X}) = a_k\mathbf{X}^k + a_{k-1}\mathbf{X}^{k-1} + \cdots + a_1\mathbf{X} + a_0\mathbf{I}_n$ é um polinômio, onde $\mathbf{X}$ é uma matriz de ordem n, então $p(\mathbf{A})$ é simétrica? Justifique.

5. Seja $\mathbf{A}$ uma matriz de ordem n.

(a) Mostre que $\mathbf{A} + \mathbf{A}^t$ é simétrica.

(b) Uma matriz $\mathbf{B} = [b_{ij}]$ de ordem n é dita *anti-simétrica* se $b_{ij} = -b_{ji}, \forall i, j$ (nesse caso, quem é a diagonal principal?). Mostre que $\mathbf{A} - \mathbf{A}^t$ é anti-simétrica.

(c) Mostre que $\mathbf{A}$ é escrita, de maneira única, como $\mathbf{A} = \mathbf{S} + \mathbf{K}$, onde $\mathbf{S}$ é simétrica e $\mathbf{K}$ é anti-simétrica.

(d) Mostre que se $\mathbf{A}^2 - 3\mathbf{A} + \mathbf{I} = \mathbf{0}$ então $\mathbf{A}^{-1} = 3\mathbf{I} - \mathbf{A}$.

(e) Mostre que se $\mathbf{A}^2 = \mathbf{0}$ então $\mathbf{I} - \mathbf{A}$ é inversível.

(f) Mostre que se $\mathbf{A}^3 = \mathbf{0}$ então $\mathbf{I} - \mathbf{A}$ é inversível.

(g) Generalizando os dois itens anteriores, mostre que se $\mathbf{A}^{k+1} = \mathbf{0}$ para algum inteiro $k \geq 0$, então $(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \cdots + \mathbf{A}^k$.

(h) Mostre que se $\mathbf{A}^2 + 2\mathbf{A} + \mathbf{I} = \mathbf{0}$ então $\mathbf{A}$ é inversível.

(i) Mostre que se $\mathbf{A}^3 - \mathbf{A} + \mathbf{I} = \mathbf{0}$ então $\mathbf{A}$ é inversível.

(j) Se $\mathbf{A}$ é inversível e $\mathbf{AB} = \mathbf{AC}$, mostre que $\mathbf{B} = \mathbf{C}$.

(k) Dê um exemplo em que $\mathbf{A}$ é não nula e $\mathbf{AB} = \mathbf{AC}$, mas $\mathbf{B} \neq \mathbf{C}$
(e portanto você não pode "cortar" a matriz $\mathbf{A}$ como faz com números reais).

6. Mostre que se $\mathbf{A}$ ou $\mathbf{B}$ não for inversível então $\mathbf{AB}$ também não é (ou equivalentemente, se $\mathbf{AB}$ for inversível então $\mathbf{A}$ e $\mathbf{B}$ também são).

7. Seja $\mathbf{A}$ uma matriz inversível. Mostre que

(a) $(\mathbf{A}^t)^{-1} = (\mathbf{A}^{-1})^t$.

(b) $\mathbf{BA}^{-1} = \mathbf{A}^{-1}\mathbf{B}$ se, e somente se, $\mathbf{AB} = \mathbf{BA}$.

(c) $\mathbf{A} + \mathbf{B}$ e $\mathbf{I} + \mathbf{BA}^{-1}$ são ambas inversíveis ou ambas não inversíveis.

8. Mostre que se uma matriz quadrada $\mathbf{A}$ satisfaz $\mathbf{A}^3 + 4\mathbf{A}^2 - 2\mathbf{A} + 7\mathbf{I} = \mathbf{0}$ então $\mathbf{A}^t$ também satisfaz.

9. Mostre que se $\mathbf{A}^t\mathbf{A} = \mathbf{A}$ então $\mathbf{A}$ é simétrica e $\mathbf{A} = \mathbf{A}^2$.

10. Seja $\mathbf{A} = [a_{ij}]$ uma matriz de ordem n. O *traço* de $\mathbf{A}$, $tr(\mathbf{A})$, é a soma de todos os elementos da diagonal principal, isto é, $tr(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn}$. Mostre que

(a) $tr(\alpha\mathbf{A}) = \alpha tr(\mathbf{A})$, onde $\alpha \in \mathbb{R}$.

(b) $tr(\mathbf{A} + \mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B})$.

(c) $tr(\mathbf{AB}) = tr(\mathbf{BA})$.

(d) $tr(\mathbf{A}^t) = tr(\mathbf{A})$.

(e) $tr(\mathbf{A}^t\mathbf{A}) \geq 0$.

Dica: Faça as contas para uma matriz $\mathbf{A}$ *de ordem 3. Talvez você consiga visualizar melhor o processo.*

11. Mostre que não existem matrizes $\mathbf{A}$ e $\mathbf{B}$ de ordem n tais que $\mathbf{AB} - \mathbf{BA} = \mathbf{I}_n$.

12. Considere $\mathbf{A}$ e $\mathbf{B}$ como matrizes quadradas de mesma ordem, e suponha que $\mathbf{AB} = \mathbf{BA}$. Mostre que $(\mathbf{A}+\mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$ e $(\mathbf{A}+\mathbf{B})(\mathbf{A}-\mathbf{B}) = \mathbf{A}^2 - \mathbf{B}^2$.

13. (a) Seja $\mathbf{A}$ a matriz de ordem 3 cujas entradas da diagonal principal e abaixo dela são nulas, enquanto as entradas acima da diagonal principal são todas iguais a 1:
$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$
Encontre $\mathbf{A}^2, \mathbf{A}^3, \mathbf{A}^4, \ldots$. Faça o mesmo para matrizes de ordem 4.

    (b) Seja $\mathbf{B} = \mathbf{A} + \mathbf{I}_3$. Calcule $\mathbf{A}^2, \mathbf{A}^3$ e $\mathbf{A}^4$.

14. Seja $\mathbf{A} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$. Mostre que $\mathbf{A}^2 = \begin{bmatrix} \cos 2\theta & -\operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & \cos 2\theta \end{bmatrix}$. Por indução, determine $\mathbf{A}^k$ para qualquer inteiro positivo k.

15. Mostre que se $\mathbf{A}$ e $\mathbf{B}$ são matrizes quadradas de mesma ordem tais que $\mathbf{A}^k = \mathbf{0}$ para algum inteiro $k \geq 1$ e $\mathbf{AB} = \mathbf{BA}$, então $(\mathbf{AB})^k = \mathbf{0}$.

16. Diz-se que uma matriz quadrada $\mathbf{A}$ é *nilpotente* quando existe um inteiro $k \geq 1$ tal que $\mathbf{A}^k = \mathbf{0}$. Mostre que qualquer matriz nilpotente não é inversível.

17. Mostre que uma matriz diagonal $\mathbf{A} = [a_{ij}]$ de ordem n tal que $a_{11}a_{22}\cdots a_{nn} \neq 0$ é inversível, e calcule sua inversa.

18. Mostre que se a matriz $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ é inversível então $ad - bc \neq 0$. Neste caso calcule $\mathbf{A}^{-1}$.

19. Sejam $\mathbf{A}$ e $\mathbf{B}$ matrizes quadradas de ordem n. Mostre que se $\mathbf{AX} = \mathbf{BX}$ para todas as matrizes $\mathbf{X}$ de ordem $n \times 1$, então $\mathbf{A} = \mathbf{B}$.

20. Mostre que se $\mathbf{AB} = \mathbf{A}$ e $\mathbf{BA} = \mathbf{B}$ então $\mathbf{A}^2 = \mathbf{A}$ e $\mathbf{B}^2 = \mathbf{B}$.

Capítulo 2

Sistemas Lineares

Uma *equação linear* nas variáveis $x_1, \ldots, x_n$ é uma equação do tipo $a_1x_1 + \cdots + a_nx_n = b$ onde os a_i's e b são escalares. Um *sistema de equações lineares* (ou simplesmente um *sistema linear*) com m equações e n incógnitas é dado por

$$\left\{ \begin{array}{ccccc} a_{11}x_1 & +a_{12}x_2 & \cdots & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & \cdots & a_{2n}x_n & = & b_2 \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & \cdots & a_{mn}x_n & = & b_m \end{array} \right. \tag{2.1}$$

Uma *solução* do sistema linear (2.1) é uma lista de n números $(x_1, \ldots, x_n)$ que satisfaz cada uma de suas m equações.

Tomando $\mathbf{A} = [a_{ij}]_{m \times n}$, $\mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ e $\mathbf{B} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$, podemos escrever o sistema (2.1) na forma matricial

$$\mathbf{AX} = \mathbf{B}.$$

Neste caso, $\mathbf{A}$ é dita *matriz dos coeficientes* de (2.1); $\mathbf{X}$ é dita *matriz das incógnitas* de (2.1) e $\mathbf{B}$ é dita *matriz dos termos independentes* de (2.1). Em particular, se $\mathbf{B} = \mathbf{0}$ então o sistema $\mathbf{AX} = \mathbf{0}$ é dito *sistema homogêneo*. Considerando a forma matricial de um sistema linear, diremos também que a matriz $\mathbf{X}_0$ é *solução* do sistema $\mathbf{AX} = \mathbf{B}$ se $\mathbf{AX}_0 = \mathbf{B}$.

A matriz

$$[\mathbf{A}|\mathbf{B}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}_{m \times (n+1)}$$

é a *matriz ampliada* do sistema (2.1).

2.1 Operações e matrizes elementares

Numa matriz $\mathbf{A}$ de ordem $m \times n$, consideramos três operações sobre suas **linhas**:

(i) troca da linha i com a linha j $(i \neq j)$. Indicaremos essa operação por $L_i \leftrightarrow L_j$.

(ii) multiplicação da linha i por um número real $k \neq 0$ $(L_i \rightarrow kL_i)$.

(iii) substituição da linha i pela linha i somada ao múltiplo k da linha j $(i \neq j)$. Neste caso pode-se ter $k = 0$. Indicaremos essa operação por $L_i \rightarrow L_i + kL_j$.

As três operações acima são chamadas *operações elementares.*

Exemplo 2.1. Seja $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 4 & -1 \\ -3 & 2 \end{bmatrix}$.

- realizemos a operação elementar $L_1 \leftrightarrow L_2$ sobre $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 4 & -1 \\ -3 & 2 \end{bmatrix} \rightarrow \mathbf{A}_1 = \begin{bmatrix} 4 & -1 \\ 1 & 0 \\ -3 & 2 \end{bmatrix}.$$

- realizemos a operação elementar $L_3 \rightarrow 2L_3$ sobre $\mathbf{A}_1$:

$$\mathbf{A}_1 = \begin{bmatrix} 4 & -1 \\ 1 & 0 \\ -3 & 2 \end{bmatrix} \rightarrow \mathbf{A}_2 = \begin{bmatrix} 4 & -1 \\ 1 & 0 \\ -6 & 4 \end{bmatrix}.$$

- realizemos a operação elementar $L_2 \rightarrow L_2 + 3L_1$ sobre $\mathbf{A}_2$:

$$\mathbf{A}_2 = \begin{bmatrix} 4 & -1 \\ 1 & 0 \\ -6 & 4 \end{bmatrix} \rightarrow \mathbf{C} = \begin{bmatrix} 4 & -1 \\ 13 & -3 \\ -6 & 4 \end{bmatrix}.$$

$\square$

Definição 2.1. *Sejam* $\mathbf{A}$ *e* $\mathbf{C}$ *matrizes de mesma ordem. Dizemos que* $\mathbf{C}$ *é* ***linha equivalente*** *à* $\mathbf{A}$ *se* $\mathbf{C}$ *pode ser obtida de* $\mathbf{A}$ *pela aplicação de finitas operações elementares.*

Assim, no exemplo anterior $\mathbf{C}$ é linha equivalente à matriz $\mathbf{A}$.

O estudo de matrizes equivalentes é útil na resolução de sistemas lineares. Diremos que dois sistemas lineares $\mathbf{AX} = \mathbf{B}$ e $\mathbf{CX} = \mathbf{D}$ que possuem as mesmas soluções são *equivalentes*. Nesse contexto, temos o

Teorema 2.1. *Dois sistemas lineares cujas matrizes ampliadas são linha equivalentes são equivalentes.*

Em outras palavras, o Teorema anterior afirma que ao aplicarmos operações elementares sobre a matriz ampliada de um sistema, mantemos soluções do sistema. Logo, podemos resolver um sistema linear transformando-o em um sistema mais fácil, como no exemplo a seguir. **Essa é a justificativa para o processo de escalonamento.**

Exemplo 2.2. Considere o sistema linear

$$S : \begin{cases} 2x_1 & +4x_2 & +2x_3 & = & 6 \\ -x_1 & & +x_3 & = & -2 \\ x_1 & +x_2 & -x_3 & = & 3 \end{cases}.$$

Apliquemos operações elementares em sua matriz ampliada:

$$\left[\begin{array}{rrr|r} 2 & 4 & 2 & 6 \\ -1 & 0 & 1 & -2 \\ \boxed{1} & 1 & -1 & 3 \end{array}\right] \xrightarrow[L_1 \to 1/2 L_1]{L_3 \to L_3 + L_2} \left[\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ \boxed{-1} & 0 & 1 & -2 \\ 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow{L_2 \to L_2 + L_1}$$

$$\left[\begin{array}{rrr|r} 1 & \boxed{2} & 1 & 3 \\ 0 & \boxed{2} & 2 & 1 \\ 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow[L_2 \to L_2 - 2L_3]{L_1 \to L_1 - 2L_3} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 1 \\ 0 & 0 & 2 & -1 \\ 0 & \boxed{1} & 0 & 1 \end{array}\right] \xrightarrow[L_2 \leftrightarrow L_3]{L_2 \to 1/2 L_2}$$

$$\left[\begin{array}{rrr|r} 1 & 0 & \boxed{1} & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1/2 \end{array}\right] \xrightarrow{L_1 \to L_1 - L_3} \left[\begin{array}{rrr|r} 1 & 0 & 0 & 3/2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1/2 \end{array}\right] = [\mathbf{C}|\mathbf{D}].$$

O sistema $\mathbf{CX} = \mathbf{D}$ dado por

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

é equivalente ao sistema original S, e claramente tem única solução $(3/2, 1, -1/2)$. Portanto o sistema inicial S tem esse terno como única solução. $\square$

A última matriz $[\mathbf{C}|\mathbf{D}]$ do exemplo anterior tem uma forma interessante pois o sistema associado é de fácil resolução. A fim de resolver sistemas lineares de uma forma geral, procuraremos formalizar a estrutura dessa matriz.

Definição 2.2. *Uma matriz $\mathbf{A}$ de ordem $m \times n$ é* ***matriz linha reduzida à forma escada*** *(MLRFE) satisfaz as seguintes propriedades:*

(i) O primeiro elemento não nulo (da esquerda para a direita) de uma linha não nula é 1 (esse é o elemento líder *da linha);*

(ii) Cada coluna que contém o líder de alguma linha tem todos os seus outros elementos nulos;

(iii) Toda linha nula ocorre abaixo das linhas não nulas;

(iv) Se as linhas $1, 2, \ldots, r$ são as linhas não nulas de $\mathbf{A}$, e se o líder da linha i ocorre na coluna k_i, $i = 1, \ldots, r$, então $k_1 < k_2 < \cdots < k_r$ (forma escada).

Exemplo 2.3. São matrizes linha reduzidas à forma escada:

- $\begin{bmatrix} 1 & 0 & 0 & 3/2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1/2 \end{bmatrix}$
- $\begin{bmatrix} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$
- $\mathbf{I}_n$
- $\mathbf{0}_{m \times n}$

$\square$

Teorema 2.2. *Toda matriz $\mathbf{A}$ é linha equivalente a uma única matriz linha reduzida à forma escada.*

O Teorema acima diz que o processo de escalonamento feito no exemplo anterior é universal: ele **sempre** é possível, para **qualquer** sistema linear.

Relacionaremos agora operações elementares com produtos de matrizes.

Definição 2.3. *Uma matriz* $\mathbf{A}$ *quadrada de ordem* n *é dita* ***elementar*** *se pode ser obtida da identidade* I_n *por uma única operação elementar.*

Exemplo 2.4. São exemplos de matrizes elementares:

- $\mathbf{E} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $(L_1 \leftrightarrow L_2 \text{ em } I_3)$

- $\mathbf{E} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ $(L_1 \to 2L_1 \text{ em } I_2)$

- $\mathbf{E} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $(L_1 \leftrightarrow L_1 + 2L_2 \text{ em } I_3)$

□

Quando convenienete, denotaremos por $e(\mathbf{A})$ a matriz resultante da aplicação da operação elementar e sobre a matriz $\mathbf{A}$.

Teorema 2.3. *Seja* e *uma operação elementar e* $\mathbf{E} = e(\mathbf{I}_m)$ *a matriz elementar correspondente. Então para toda matriz* $\mathbf{A}$ *de ordem* $m \times n$ *temos*

$$e(\mathbf{A}) = \mathbf{EA}.$$

Em outras palavras, o Teorema 2.3 diz que aplicar uma operação elementar em $\mathbf{A}$ é o mesmo que multiplicar $\mathbf{A}$ a esquerda pela matriz elementar correspondente.

Cada matriz elementar é inversível, e sua inversa é a **matriz elementar correspondente à operação que desfaz** a original:

- a operação $L_i \to \frac{1}{k}L_i$ desfaz a operação $L_i \to kL_i$. Assim por exemplo, se $\mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$ então $\mathbf{E}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1/k \end{bmatrix}$ (verifique este fato constatando que $\mathbf{EE}^{-1} = \mathbf{I}_2$).

- a operação $L_i \to L_j$ desfaz a própria operação $L_i \to L_j$. Assim por exemplo, se $\mathbf{E} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ então $\mathbf{E}^{-1} = \mathbf{E}$ (verifique!).

- a operação $L_i \to L_i - kL_j$ desfaz a operação $L_i \to L_i + kL_j$. Assim por exemplo, se $\mathbf{E} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ então $\mathbf{E}^{-1} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$ (verifique!).

Atividade 2.1. Faça exemplos de inversas de matrizes elementares de ordem 3. Mostre que em geral matrizes elementares tem inversas descritas anteriormente.

Uma consequência imediata do Teorema 2.3 é o seguinte:

Corolário 2.1. *Sejam* $\mathbf{A}$ *e* $\mathbf{B}$ *matrizes de mesma ordem. Então* $\mathbf{B}$ *é linha equivalente a* $\mathbf{A}$ *se, e somente se* $\mathbf{B} = \mathbf{E}_k\mathbf{E}_{k-1}\cdots\mathbf{E}_2\mathbf{E}_1\mathbf{A}$ *para certas matrizes elementares* $\mathbf{E}_1, \dots, \mathbf{E}_k$.

Exemplo 2.5. Mostre que são linha equivalentes as matrizes $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ e $\mathbf{B} = \begin{bmatrix} 9 & 18 \\ 6 & 12 \end{bmatrix}$.

Vamos calcular a MLRFE de $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \xrightarrow{L_2 \to 1/3L_2} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \xrightarrow{L_2 \to L_2 - L_1} \mathbf{C} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

Observe que, sendo $\mathbf{E}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix}$ e $\mathbf{E}_2 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ as matrizes elementares correspondentes às operações realizadas, temos

$$\mathbf{C} = \mathbf{E}_2\mathbf{E}_1\mathbf{A}.$$

Agora, calculemos a MLRFE de $\mathbf{B}$:

$$\mathbf{B} = \begin{bmatrix} 9 & 18 \\ 6 & 12 \end{bmatrix} \xrightarrow{L_1 \to 1/9L_1} \begin{bmatrix} 1 & 2 \\ 6 & 12 \end{bmatrix} \xrightarrow{L_2 \to 1/6L_2} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \xrightarrow{L_2 \to L_2 - L_1} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = \mathbf{C}.$$

Sendo $\mathbf{E}_3 = \begin{bmatrix} 1/9 & 0 \\ 0 & 1 \end{bmatrix}$ e $\mathbf{E}_4 = \begin{bmatrix} 1 & 0 \\ 0 & 1/6 \end{bmatrix}$ temos

$$\mathbf{C} = \mathbf{E}_2\mathbf{E}_4\mathbf{E}_3\mathbf{B},$$

e assim $\mathbf{E}_2\mathbf{E}_4\mathbf{E}_3\mathbf{B} = \mathbf{E}_2\mathbf{E}_1\mathbf{A} \Rightarrow \mathbf{E}_4\mathbf{E}_3\mathbf{B} = \mathbf{E}_2^{-1}\mathbf{E}_2\mathbf{E}_1\mathbf{A} \Rightarrow \cdots \Rightarrow \mathbf{B} = \mathbf{E}_3^{-1}\mathbf{E}_4^{-1}\mathbf{E}_1\mathbf{A}$. Pelo Corolário anterior, $\mathbf{B}$ é linha equivalente à $\mathbf{A}$. $\square$

Em particular, se $\mathbf{B}$ é matriz quadrada de ordem n, linha equivalente à $\mathbf{I}_n$, então $\mathbf{B} = \mathbf{E}_k\cdots\mathbf{E}_1\mathbf{I}_n$. O produto $\mathbf{E}_k\cdots\mathbf{E}_1 = \mathbf{A}$ é inversível com inversa $\mathbf{A}^{-1} = \mathbf{E}_1^{-1}\cdots\mathbf{E}_k^{-1}$ (verifique!). Assim, $\mathbf{B} = \mathbf{A}\mathbf{I}_n \Rightarrow \mathbf{A}^{-1}\mathbf{B} = \mathbf{I}_n$, e $\mathbf{B}$ é inversível com inversa $\mathbf{B}^{-1} = \mathbf{A}^{-1} = \mathbf{E}_1^{-1}\cdots\mathbf{E}_k^{-1}$. **Concluímos então que se B é linha equivalente à $\mathbf{I}_n$** (ou equivalentemente, se $\mathbf{B}$ é produto de matrizes elementares) **então B é inversível.**

A recíproca deste fato também é verdadeira, isto é, se $\mathbf{B}$ é inversível então é linha equivalente a $\mathbf{I}_n$ (Exercício 9). Resumindo esse fato e considerando o Corolário 2.1, temos o

Teorema 2.4. *Seja* $\mathbf{A}$ *uma matriz quadrada de ordem* n. *São equivalentes as afirmações:*

(*i*) $\mathbf{A}$ *é inversível.*

(*ii*) $\mathbf{A}$ *é linha equivalente à* $\mathbf{I}_n$.

(*iii*) $\mathbf{A} = \mathbf{E}_k\cdots\mathbf{E}_1$, *para certas matrizes elementares* $\mathbf{E}_1, \dots, \mathbf{E}_k$.

Com as operações e matrizes elementares, estabeleceremos uma maneira de

1. Resolver um sistema linear;
2. Inverter uma matriz, ou constatar que não há inversa.

Veremos isso nas seções seguintes.

2.2 Resolução de sistemas lineares

A possibilidade de simplificação da matriz ampliada de um sistema linear para a forma reduzida, garantida pelo Teorema 2.2, resulta em um processo sistemático para resolução de qualquer sistema linear.

Este processo (de escalonamento) nos fornecerá:

- as soluções do sistema, caso existam;
- se o sistema possui única ou várias soluções;
- se o sistema não possui solução.

O estudo da quantidade/existência de soluções está relacionado com a noção de *posto* e *nulidade* das matrizes associadas ao sistema linear.

Definição 2.4. *Dada uma matriz* $\mathbf{A}$ *de ordem* $m \times n$*, seja* $\mathbf{B}$ *sua MLRFE. Então o* ***posto*** *de* $\mathbf{A}$ *é o número de linhas não nulas de* $\mathbf{B}$*. A* ***nulidade*** *de* $\mathbf{A}$ *é o número* $n - p$*, onde* p *é o posto de* $\mathbf{A}$*.*

Exemplo 2.6. Seja $\mathbf{A} = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & 2 \\ 1 & -5 & 1 \\ 4 & 16 & 8 \end{bmatrix}$. A MLRFE de $\mathbf{A}$ é a matriz $\mathbf{B} = \begin{bmatrix} 1 & 0 & 14/9 \\ 0 & 1 & 1/9 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, e portanto o posto de $\mathbf{A}$ é 2, e a nulidade de $\mathbf{A}$ é $3 - 2 = 1$. □

Exemplo 2.7. Seja $\mathbf{A} = \begin{bmatrix} 2 & 1 & 4 \\ 1 & 2 & 3 \\ 1 & -4 & -1 \end{bmatrix}$. A MLRFE de $\mathbf{A}$ é a matriz $\mathbf{B} = \begin{bmatrix} 1 & 0 & 5/3 \\ 0 & 1 & 2/3 \\ 0 & 0 & 0 \end{bmatrix}$, e portanto o posto de $\mathbf{A}$ é 2, e a nulidade de $\mathbf{A}$ é $3 - 2 = 1$. □

Teorema 2.5. *Seja* $\mathbf{AX} = \mathbf{B}$ *um sistema linear, com* m *equações e* n *incógnitas (*$\mathbf{A}$ *tem ordem* $m \times n$*). Então*

(*i*) $\mathbf{AX} = \mathbf{B}$ *admite solução se, e somente se o posto da matriz ampliada* $[\mathbf{A}|\mathbf{B}]$ *é igual ao posto da matriz dos coeficientes* $\mathbf{A}$*.*

(*ii*) *Se* $\mathbf{A}$ *e* $[\mathbf{A}|\mathbf{B}]$ *têm mesmo posto* $p = n$ *então* $\mathbf{AX} = \mathbf{B}$ *tem única solução.*

(*iii*) *Se* $\mathbf{A}$ *e* $[\mathbf{A}|\mathbf{B}]$ *têm mesmo posto* $p < n$ *então* $\mathbf{AX} = \mathbf{B}$ *tem infinitas soluções. Dizemos neste caso que a nulidade de* $\mathbf{A}$ *é o* grau de liberdade *de* $\mathbf{AX} = \mathbf{B}$*.*

Exemplo 2.8. Para cada sistema linear abaixo, diga se há ou não solução e, caso possua, se é única, infinitas; neste caso, calcule-as (faremos durante a aula):

(a) $S : \begin{cases} 2x_1 & +4x_2 & & = & 2 \\ & x_2 & & = & 3 \\ & x_2 & +x_3 & = & 4 \end{cases}$

Calculemos a MLRFE da matriz ampliada do sistema, escalonando-a:

$$\left[\begin{array}{ccc|c} 2 & 4 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 1 & 1 & 4 \end{array}\right] \xrightarrow{L_1 \to \frac{1}{2}L_1} \left[\begin{array}{ccc|c} 1 & \boxed{2} & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & \boxed{1} & 1 & 4 \end{array}\right] \xrightarrow[L_1 \to L_1 - 2L_2]{L_3 \to L_3 - L_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

Vemos que o número de linhas não nulas da MLRFE da matriz ampliada e da MLRFE da matriz dos coeficientes (as três primeiras colunas da última matriz acima) são iguais a 3. Ou seja,

$$p = \text{posto}\,[\mathbf{A}|\mathbf{B}] = \text{posto}\,\mathbf{A} = 3 = n.$$

Pelo Teorema 2.5(ii), o sistema S admite única solução. O sistema equivalente, associado à MLRFE da ampliada, é

$$S' : \begin{cases} x_1 & & & = & -5 \\ & x_2 & & = & 3 \\ & & x_3 & = & 1 \end{cases},$$

cuja solução é $\mathbf{X} = (-5, 3, 1)$. Geometricamente, o sistema original S é a interseção de três planos 2 a 2 não paralelos.

(b) $S : \begin{cases} x_1 & +x_2 & & +x_4 & = & 0 \\ & x_2 & & +2x_4 & = & 1 \\ & +x_2 & +x_3 & & = & 2 \end{cases}$

Primeiro escalonamos a matriz ampliada do sistema:

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & \boxed{1} & 1 & 0 & 2 \end{array}\right] \xrightarrow{L_3 \to L_3 - L_2} \left[\begin{array}{cccc|c} 1 & \boxed{1} & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 \end{array}\right] \xrightarrow{L_1 \to L_1 - L_2} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 \end{array}\right]$$

Vemos que o número de linhas não nulas na matriz ampliada escalonada é 3, igual ao número de linhas não nulas da matriz dos coeficientes escalonada (as quatro primeiras colunas da última matriz acima). Ou seja,

$$p = \text{posto}\,[\mathbf{A}|\mathbf{B}] = \text{posto}\,\mathbf{A} = 3 < 4 = n.$$

Pelo Teorema 2.5(iii), o sistema S admite infinitas soluções. Para encontrá-las, consideramos o sistema escalonado, proveniente da forma reduzida da matriz ampliada,

$$S' : \begin{cases} x_1 & & & -x_4 & = & 0 \\ & x_2 & & +2x_4 & = & 1 \\ & & +x_3 & -2x_4 & = & 1 \end{cases},$$

cujo *conjunto solução* é

$$\{\mathbf{X} \in \mathbb{R}^4 \mid X = (\, -1 + x_4 \,,\, 1 - 2x_4 \,,\, 1 + 2x_4 \,,\, x_4 \,),\ x_4 \in \mathbb{R}\}.$$

Veja que a nulidade desse sistema é $n - p = 1$, o que indica que temos “1 grau de liberdade”, isto é, podemos escolher *qualquer* valor para uma das variáveis (no conjunto acima, x_4), obtendo várias soluções. Assim, x_4 faz o papel de variável livre, enquanto as demais estão em função de x_4.

É interessante observar a geometria do conjunto solução: trata-se de uma reta em $\mathbb{R}^4$, passando pelo ponto $(-1, 1, 1, 0)$ na direção do vetor $(1, -2, 2, 1)$.

(c) $S : \begin{cases} x_1 & +x_2 & = & 1 \\ 2x_1 & -x_2 & = & 2 \\ x_1 & -2x_2 & = & 0 \end{cases}$

Escalonemos a matriz ampliada do sistema:

$$\left[\begin{array}{rr|r} 1 & 1 & 1 \\ 2 & -1 & 2 \\ \boxed{1} & -2 & 0 \end{array}\right] \xrightarrow{L_3 \to L_3 - L_1} \left[\begin{array}{rr|r} 1 & 1 & 1 \\ \boxed{2} & -1 & 2 \\ 0 & -3 & -1 \end{array}\right] \xrightarrow{L_2 \to L_2 - 2L_1} \left[\begin{array}{rr|r} 1 & 1 & 1 \\ 0 & -3 & 0 \\ 0 & \boxed{-3} & -1 \end{array}\right] \xrightarrow[L_2 \to \frac{1}{2} L_2]{L_3 \to L_3 - L_2}$$

$$\left[\begin{array}{rr|r} 1 & \boxed{1} & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{array}\right] \xrightarrow[L_3 \to -L_3]{L_1 \to L_1 - L_2} \left[\begin{array}{rr|r} 1 & 0 & \boxed{1} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow{L_1 \to L_1 - L_3} \left[\begin{array}{rr|r} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right].$$

Assim, as matrizes reduzidas à forma escada da ampliada e da matriz dos coeficientes são, respectivamente,

$$\left[\begin{array}{rr|r} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \quad \text{e} \quad \left[\begin{array}{rr} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{array}\right].$$

Essas matrizes têm posto diferente: a primeira tem posto 3, e a segunda, 2. Do Teorema 2.5(i), concluímos que o sistema S não possui solução.

□

2.3 Um processo para inversão de matrizes

Sabemos que uma matriz quadrada $\mathbf{A}$ é inversível se, e somente se $\mathbf{A} = \mathbf{E}_1^{-1} \cdots \mathbf{E}_k^{-1}$, onde $\mathbf{E}_1^{-1}, \ldots, \mathbf{E}_k^{-1}$ são matrizes elementares. Neste caso,

$$\mathbf{E}_k \cdots \mathbf{E}_1 \mathbf{A} = \mathbf{I} \quad \text{e} \quad \mathbf{A}^{-1} = \mathbf{E}_k \cdots \mathbf{E}_1 \mathbf{I}.$$

Assim, aplicando as operações elementares relativas às matrizes elementares $\mathbf{E}_1, \ldots, \mathbf{E}_k$ sobre $[\mathbf{A}|\mathbf{I}]$, obtemos a sequência

$$[\mathbf{A}|\mathbf{I}] \xrightarrow{\mathbf{E}_1} [\mathbf{E}_1\mathbf{A}|\mathbf{E}_1\mathbf{I}] \xrightarrow{\mathbf{E}_2} \cdots \xrightarrow{\mathbf{E}_k} [\mathbf{E}_k \cdots \mathbf{E}_1\mathbf{A}|\mathbf{E}_k \cdots \mathbf{E}_1\mathbf{I}] = [\mathbf{I}|\mathbf{A}^{-1}].$$

Em outras palavras, $\mathbf{A}$ é inversível se, e somente se $[\mathbf{A}|\mathbf{I}]$ é linha equivalente a uma matriz $[\mathbf{I}|\mathbf{S}]$, e neste caso $\mathbf{A}^{-1} = \mathbf{S}$.

Exemplo 2.9. Calcular a inversa de cada matriz abaixo, se existir.

(a) $\mathbf{A} = \left[\begin{array}{rrr} 1 & 0 & 1 \\ -1 & 2 & 1 \\ 0 & 2 & 0 \end{array}\right].$

$$[\mathbf{A}|\mathbf{I}_3] = \left[\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[L_1 \to L_1 + L_2]{L_2 \to L_2 - L_3} \left[\begin{array}{rrr|rrr} 0 & 0 & 2 & 1 & 1 & -1 \\ -1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 2 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[\substack{L_1 \leftrightarrow L_2 \\ L_2 \leftrightarrow L_3}]{L_2 \to -L_2} \left[\begin{array}{rrr|rrr} 1 & 0 & -1 & 0 & -1 & 1 \\ 0 & 2 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 & -1 \end{array}\right] \xrightarrow[\substack{L_3 \to 1/2 L_3 \\ L_1 \to L_1 + L_3}]{L_2 \to 1/2 L_2} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1/2 & -1/2 & 1/2 \\ 0 & 1 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 1/2 & 1/2 & -1/2 \end{array}\right].$$

Logo $\mathbf{A}^{-1} = \dfrac{1}{2} \left[\begin{array}{rrr} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & -1 \end{array}\right].$

(b) $\mathbf{B} = \begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix}$.

$$[\mathbf{B}|\mathbf{I}_2] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ -2 & -4 & 0 & 1 \end{array}\right] \xrightarrow{L_2 \to L_2 + 2L_1} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{array}\right].$$

Observe que a MLRFE de $\mathbf{B}$ é a matriz $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \neq I_2$, e daí $\mathbf{B}$ não é inversível.

□

2.4 Demonstrações

Demonstração do Teorema 2.1. Seja $\mathbf{AX} = \mathbf{B}$ um sistema linear. É suficiente mostrar que cada uma das três operações elementares sobre $[\mathbf{A}|\mathbf{B}]$ resulta num sistema linear $\mathbf{CX} = \mathbf{D}$ com as mesmas soluções de $\mathbf{AX} = \mathbf{B}$. Façamos a prova para a operação $L_i \to L_i + kL_j$ $(i \neq j)$ e deixamos as outras duas como exercício. Supomos que ao realizar a operação $L_i \to L_i + kL_j$ sobre $[\mathbf{A}|\mathbf{B}]$, obtemos a matriz linha equivalente $[\mathbf{C}|\mathbf{D}]$. Comparando os sistemas $\mathbf{AX} = \mathbf{B}$ e $\mathbf{CX} = \mathbf{D}$, vemos que a única diferença está na linha i:

- $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$ é a linha i de $\mathbf{AX} = \mathbf{B}$;
- $(a_{i1} + ka_{j1})x_1 + (a_{i2} + a_{j2})x_2 + \cdots + (a_{in} + a_{jn})x_n = (b_i + kb_j)$ é a linha i de $\mathbf{CX} = \mathbf{D}$;

Então,

$$\begin{gathered} (x_1, \ldots, x_n) \text{ é solução de } \mathbf{CX} = \mathbf{D} \\ \Updownarrow \\ (a_{i1} + ka_{j1})x_1 + (a_{i2} + a_{j2})x_2 + \cdots + (a_{in} + a_{jn})x_n = (b_i + kb_j), \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = b_k, \forall k \neq i \\ \Updownarrow \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i - k\underbrace{(a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n - b_j)}_{0}, \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = b_k, \forall k \neq i \\ \Updownarrow \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = b_k, \forall k \\ \Updownarrow \\ (x_1, \ldots, x_n) \text{ é solução de } \mathbf{AX} = \mathbf{B}, \end{gathered}$$

isto é, $\mathbf{CX} = \mathbf{D}$ e $\mathbf{AX} = \mathbf{B}$ têm as mesmas soluções. □

Atividade 2.2. Complete a prova do teorema anterior fazendo a demonstração para as operações $L_i \leftrightarrow L_j$ e $L_i \to kL_i$.

Demonstração do Teorema 2.3. Deixamos a prova do resultado para as operações $L_i \leftrightarrow L_j$ e $L_i \to kL_i$ para o leitor. Seja e a operação $L_i \to L_i + kL_j$ $(i \neq j)$. Sem perda de generalidade, vamos supor que $i = 1$ e $j = 2$. Assim

$$\mathbf{EA} = \begin{bmatrix} 1 & k & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} + ka_{21} & a_{12} + ka_{22} & a_{13} + ka_{23} & \cdots & a_{1n} + ka_{2n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} = e(\mathbf{A}).$$

□

Atividade 2.3. Complete a prova do teorema anterior.

2.5 Exercícios

1. Faça os exercícios 1, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 21, 22 do Capítulo 2 do livro texto [2].

2. Mostre que $\mathbf{A}$ e $\mathbf{B}$ são matrizes linha equivalentes.

   (a) $\mathbf{A} = \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

   (b) $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$

   (c) $\mathbf{A} = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 4 & 0 \\ 2 & 3 & 4 \end{bmatrix}$

3. Encontre a inversa de cada matriz se possível.

   (a) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$

   (b) $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$

   (c) $\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 1 & 3 & 4 \end{bmatrix}$

   (d) $\begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix}$

   (e) $\begin{bmatrix} 1 & 2 & -3 & 1 \\ -1 & 3 & -3 & -2 \\ 2 & 0 & 1 & 1 \\ 3 & 1 & -2 & 3 \end{bmatrix}$

4. Mostre que se $\mathbf{X}_0$ é uma solução do sistema $\mathbf{AX} = \mathbf{0}$ e $\mathbf{X}_1$ é uma solução de $\mathbf{AX} = \mathbf{B}$ então $\mathbf{X}_0 + \mathbf{X}_1$ é solução de $\mathbf{AX} = \mathbf{B}$. O sistema $\mathbf{AX} = \mathbf{0}$ cujos termos livres são todos nulos é chamado *sistema homogêneo*.

5. Mostre que se $\mathbf{X}_1$ e $\mathbf{X}_2$ são soluções de $\mathbf{AX} = \mathbf{B}$ então $\mathbf{X}_1 - \mathbf{X}_2$ é solução de $\mathbf{AX} = \mathbf{0}$. Conclua que toda solução $\overline{\mathbf{X}}$ de $\mathbf{AX} = \mathbf{B}$ é a soma de uma solução do sistema homogêneo $\mathbf{AX} = \mathbf{0}$ com uma solução particular $\mathbf{X}_1$ de $\mathbf{AX} = \mathbf{B}$.

6. Diga se os sistemas abaixo tem única, infinitas ou nenhuma solução analizando os postos de suas matrizes. Então resolva-os, se for o caso.

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

(b) $\begin{cases} -x & -y & +z & = & 0 \\ 2x & +y & +z & = & 1 \\ 5x & +4y & -2z & = & 1 \end{cases}$

(c) $\begin{cases} x & +y & +2z & = & 8 \\ -x & -2y & +3z & = & 1 \\ 3x & -7y & +4z & = & 10 \end{cases}$

(d) $\begin{cases} 3x & +2y & -z & = & -15 \\ 5x & +3y & +2z & = & 0 \\ 3x & +y & +3z & = & 11 \\ 11x & +7y & & = & -30 \end{cases}$

(e) $\begin{cases} 3x_1 & +4x_2 & -x_3 & +2x_4 & = & 2 \\ 2x_1 & -2x_2 & +x_3 & -3x_4 & = & 5 \\ -x_1 & +3x_2 & +2x_3 & -x_4 & = & 3 \\ 2x_1 & +7x_2 & +x_3 & +x_4 & = & -1 \end{cases}$

(f) $\begin{cases} x & +y & +z & = & 2 \\ x & -y & +z & = & 2 \\ x & +2y & +z & = & -1 \end{cases}$

(g) $\begin{cases} -2x & +y & +z & = & 1 \\ x & -2y & +z & = & 1 \\ x & +y & -2z & = & 1 \end{cases}$

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

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

(j) $\begin{cases} x_1 & +2x_2 & -x_3 & +2x_4 & = & -1 \\ 3x_1 & -x_2 & +2x_3 & -5x_4 & = & 5 \\ 4x_1 & +x_2 & +x_3 & -3x_4 & = & 4 \\ 7x_1 & & +3x_3 & -8x_4 & = & 9 \end{cases}$

(k) $\begin{cases} x_1 & +3x_2 & -2x_3 & & +2x_5 & & = & 0 \\ 2x_1 & +6x_2 & -5x_3 & -2x_4 & +4x_5 & -3x_6 & = & -1 \\ & & 5x_3 & +10x_4 & & +15x_6 & = & 5 \\ 2x_1 & +6x_2 & & +8x_4 & +4x_5 & +18x_6 & = & 6 \end{cases}$

7. Encontre um polinômio $p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$ que passe pelos pontos $(-2, 24)$, $(-1, 24)$, $(1, 60)$ e $(2, 12)$.

8. Considere a matriz
$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -5 & 2 \end{bmatrix}$$
   (a) Encontre matrizes elementares $\mathbf{E}_1$ e $\mathbf{E}_2$ tais que $\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}_2$.
   (b) A partir do item anterior, escreva $\mathbf{A}$ e $\mathbf{A}^{-1}$ como produto de matrizes elementares.

9. Mostre que se $\mathbf{A}$ é uma matriz inversível, então $\mathbf{A}$ é produto de matrizes elementares.

10. Usando matrizes elementares, mostre que se $\mathbf{A}$ é inversível e $\mathbf{B}$ é linha equivalente à $\mathbf{A}$ então $\mathbf{B}$ também é inversível.

    Dica: Você pode usar a seguinte propriedade: se duas matrizes quadradas $\mathbf{C}$ *e* $\mathbf{D}$ *de ordem* n *são inversíveis então* $(\mathbf{CD})^{-1} = \mathbf{D}^{-1}\mathbf{C}^{-1}$.

11. **(Álgebra básica do método simplex para resolução de problemas de otimização linear)** Considere um sistema linear $\mathbf{Ax} = \mathbf{b}$, onde $\mathbf{A}$ é matriz $m \times n$ com $m < n$. Escrevemos $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{N} \end{bmatrix}$, onde $\mathbf{B}$ é a matriz quadrada de ordem n formada pelas n primeiras colunas de $\mathbf{A}$, e $\mathbf{N}$ a matriz das $n - m$ últimas colunas de $\mathbf{A}$. Tomemos $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix}$, onde $\mathbf{x}_B$ é formado pelas n primeiras entradas de $\mathbf{x}$ e $\mathbf{x}_N$ formado pelas $n - m$ últimas entradas de $\mathbf{x}$. O sistema $\mathbf{Ax} = \mathbf{b}$ pode ser então escrito como
$$\begin{bmatrix} \mathbf{B} & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \mathbf{b}.$$

(a) Mostre que $\begin{bmatrix} \mathbf{B} & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N$, e que portanto o sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ pode ser escrito como

$$\mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}$$

(isto indica que a multiplicação de matrizes pode ser feita por blocos).

(b) Se $\mathbf{B}$ for inversível, mostre que $\mathbf{x} = \begin{bmatrix} \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N \\ \mathbf{x}_N \end{bmatrix}$ é solução do sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$, para qualquer $\mathbf{x}_N$ escolhido.

Em particular, conclua que, fazendo $\mathbf{x}_N = \mathbf{0}$, $\mathbf{x} = \begin{bmatrix} \mathbf{B}^{-1}\mathbf{b} \\ \mathbf{0} \end{bmatrix}$ é solução de $\mathbf{A}\mathbf{x} = \mathbf{b}$ (esta solução se chama *solução básica* de $\mathbf{A}\mathbf{x} = \mathbf{b}$ e é a base para o funcionamento do método simplex).

(c) Encontre uma solução básica do sistema abaixo identificando as matrizes $\mathbf{B}$, $\mathbf{N}$, usando o item anterior, e fazendo $\mathbf{x}_N = \mathbf{0}$:

$$\begin{cases} 2x_1 & & & +2x_4 & +25x_5 & = & 2 \\ & -x_2 & & -13x_4 & -11x_5 & = & 5 \\ & & x_3 & -x_4 & +8x_5 & = & 3 \end{cases}$$

2.6 Notas

O método de resolução de sistemas lineares visto nesta seção é conhecido como **método do escalonamento**, ou **método da eliminação de Gauss**, ou ainda **método da eliminação de Gauss-Jordan**, ainda que alguns autores os diferenciem entre zerar elementos superiores e inferiores dos elementos líderes, ou somente os inferiores.

Resolver sistemas lineares numericamente é uma necessidade frequente: aparece em computação gráfica, otimização, resolução de equações diferenciais etc.

A ideia de escalonar um sistema é muito antiga, data, pelo menos, do século 18. A eliminação de Gauss consiste em uma maneira sistemática (implementável em computador) para resolução de sistemas lineares quaisquer usando as 3 operações elementares sobre linhas da matriz ampliada do sistema, como visto aqui. As operações são aplicadas de forma ordenada, de modo que os zeros na matriz ampliada apareçam primeiro nas colunas à esquerda, depois nas colunas à direita. Veja o Wikipedia para uma explicação "super-rápida"(ou o Wikipedia em inglês, que está melhor escrito).

Mas, e daí?

Daí que, por incrível que pareça, a ideia deste método resiste até hoje em pacotes computacionais modernos. Talvez o exemplo mais famoso seja a "rotina MA57". Esta rotina é usada em inúmeros métodos computacionais de hoje em dia, e é considerada como uma espécie de "padrão de qualidade" na resolução de sistemas. Evidentemente, os métodos implementados nas rotinas modernas são versões melhoradas daquele exposto aqui (eles envolvem, por exemplo, técnicas "espertas" para escolha de operações elementares, e os chamados pré-condicionadores).

Você pode consultar a implementação em Fortran da MA57 de um grande grupo de pesquisa do Reino Unido em `http://www.hsl.rl.ac.uk/catalogue/ma57.html` (sim, Fortran ainda é usado fortemente em computação de alto desempenho!). A descrição do pacote diz que ele implementa uma variante do método da eliminação de Gauss. A última versão deste pacote é de 2016, e baseia-se em um artigo científico de 1983 (1983 parece antigo, mas quando comparado à ideia de escalonamento, "é ontem"). E tenha certeza que artigos científicos foram submetidos à revistas especializadas em 2020 usando esta rotina!!! Ou seja, a ideia de escalonar uma matriz está presente na pesquisa de ponta até hoje.

Capítulo 3

Determinantes

Dada uma matriz $\mathbf{A} = [a_{ij}]$ quadrada, vamos associar um número real

$$\det \mathbf{A},$$

chamado *determinante de* $\mathbf{A}$. A seguir, exibimos como calcular o determinante de matrizes de ordem ≤ 3.

Matrizes de ordem 1: À matriz $\mathbf{A} = [a_{11}]$, definimos

$$\det \mathbf{A} = a_{11}.$$

Matrizes de ordem 2: À matriz $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, definimos

$$\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21}.$$

Matrizes de ordem 3: À matriz $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, definimos

$$\det \mathbf{A} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \tag{3.1}$$

Nosso objetivo é definir o determinante de uma matriz de ordem n qualquer ($n \geq 1$). Perceba que na definição do determinante de matrizes de ordem 2 e 3 ocorre uma certa regularidade: em cada produto de elementos da matriz, os termos das linhas aparecem ordenados, enquanto os termos das colunas não. Observe ainda que nos índices das colunas, todas as ordens possíveis dos números $1, 2, 3$ (ou $1, 2$ no caso de ordem 2) aparecem. O que não é muito visível por hora é como o sinal de cada produto é atribuído. O que faremos no caso geral é "estender" essa regularidade. Portanto, tenha os determinantes anteriores em mente!

Uma forma de definirmos o determinante de uma matriz qualquer é fazer um estudo minucioso de produtos do tipo

$$\pm\, a_{1j_1}a_{2j_2}\cdots a_{nj_n},$$

onde $j_1, \ldots, j_n$ são os índices das colunas em alguma ordem (uma *permutação* de $1, \ldots, n$). Essa é uma forma direta de estender a soma (3.1) para matrizes de ordem $n \times n$. O livro do Boldrini faz dessa maneira. **Porém** essa forma de apresentar o determinante não nos dá uma maneira prática de calculá-lo. Portanto, apresentaremos o determinante por meio de um método para calculá-lo: o **desenvolvimento de Laplace**.

3.1 Desenvolvimento de Laplace

Voltemos ao determinante de matrizes de ordem 2:

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Observe que podemos escrever

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}\big[(-1)^{1+1}\det \mathbf{A}_{11}\big] + a_{12}\big[(-1)^{1+2}\det \mathbf{A}_{12}\big]$$

(lembre-se que o determinante de uma matriz 1×1 é o próprio elemento da matriz), onde as matrizes $\mathbf{A}_{ij}$ são obtidas da matriz $\mathbf{A}$ eliminando a linha i e a coluna j:

- $\begin{bmatrix} \boxed{\mathbf{a_{11}}} & \mathbf{a_{12}} \\ \mathbf{a_{21}} & a_{22} \end{bmatrix} \longrightarrow \mathbf{A}_{11} = \begin{bmatrix} a_{22} \end{bmatrix}$
- $\begin{bmatrix} \mathbf{a_{11}} & \boxed{\mathbf{a_{12}}} \\ a_{21} & \mathbf{a_{22}} \end{bmatrix} \longrightarrow \mathbf{A}_{12} = \begin{bmatrix} a_{21} \end{bmatrix}$

O mesmo ocorre com o determinante de ordem 3

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

Podemos escrever

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

que por sua vez pode ser escrito como

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}\big[(-1)^{1+1}\det \mathbf{A}_{11}\big] + a_{12}\big[(-1)^{1+2}\det \mathbf{A}_{12}\big] + a_{13}\big[(-1)^{1+3}\det \mathbf{A}_{13}\big],$$

onde as matrizes $\mathbf{A}_{ij}$ são obtidas de maneira análoga:

- $\begin{bmatrix} \boxed{\mathbf{a_{11}}} & \mathbf{a_{12}} & \mathbf{a_{13}} \\ \mathbf{a_{21}} & a_{22} & a_{23} \\ \mathbf{a_{31}} & a_{32} & a_{33} \end{bmatrix} \longrightarrow \mathbf{A}_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$
- $\begin{bmatrix} \mathbf{a_{11}} & \boxed{\mathbf{a_{12}}} & \mathbf{a_{13}} \\ a_{21} & \mathbf{a_{22}} & a_{23} \\ a_{31} & \mathbf{a_{32}} & a_{33} \end{bmatrix} \longrightarrow \mathbf{A}_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$
- $\begin{bmatrix} \mathbf{a_{11}} & \mathbf{a_{12}} & \boxed{\mathbf{a_{13}}} \\ a_{21} & a_{22} & \mathbf{a_{23}} \\ a_{31} & a_{32} & \mathbf{a_{33}} \end{bmatrix} \longrightarrow \mathbf{A}_{13} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$

Observe que os sinais de cada parcela são corretamente capturados pelas potências de -1.

Note que reduzimos o cálculo de determinantes de ordem 3 em determinantes de ordem 2. Essa é uma maneira prática de calcular determinantes, já que determinantes de ordem 2 são fáceis.

Agora estenderemos essa ideia para o caso $n \times n$.

Dada uma matriz $\mathbf{A} = [a_{ij}]$ de ordem n, consideremos a matriz quadrada $\mathbf{A}_{ij}$ de ordem $n-1$ obtida de $\mathbf{A}$ pela retirada da linha i e da coluna j. Definimos o *cofator de* a_{ij} como o número

$$\Delta_{ij} = (-1)^{i+j} \det \mathbf{A}_{ij}.$$

Veja que o cofator é o que aparece no que fizemos anteriormente.

Exemplo 3.1. Dada $\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 3 & -1 & -2 \end{bmatrix}$ temos

$$\Delta_{11} = (-1)^{1+1} \det \begin{bmatrix} 1 & 2 \\ -1 & -2 \end{bmatrix} = 0,$$

$$\Delta_{12} = (-1)^{1+2} \det \begin{bmatrix} 0 & 2 \\ 3 & -2 \end{bmatrix} = 6,$$

$$\Delta_{13} = (-1)^{1+3} \det \begin{bmatrix} 0 & 1 \\ 3 & -1 \end{bmatrix} = -3.$$

□

Nesse contexto, definimos o determinante de matrizes quadradas quaisquer.

Definição 3.1. *Dada uma matriz* $\mathbf{A} = [a_{ij}]$ *de ordem* n*, definimos o* ***determinante de*** $\mathbf{A}$ *como o número real*

$$\det \mathbf{A} = a_{11}\Delta_{11} + a_{12}\Delta_{12} + \cdots + a_{1n}\Delta_{1n}.$$

Um fato importante sobre o desenvolvimento de Laplace é que o determinante por ser calculado escolhendo **qualquer** linha (e não só a primeira como da definição anterior). Pode-se ainda escolher **colunas** para desenvolver a soma de cofatores.

É o que o teorema a seguir diz (não o provaremos).

Teorema 3.1 (Desenvolvimento de Laplace)**.** *Dada uma matriz* $\mathbf{A} = [a_{ij}]$ *de ordem* n*, escolhemos uma linha qualquer* k *de* $\mathbf{A}$*. Então podemos desenvolver o determinante sobre esta linha:*

$$\det \mathbf{A} = a_{k1}\Delta_{k1} + a_{k2}\Delta_{k2} + \cdots + a_{kn}\Delta_{kn}.$$

Também, podemos escolher uma coluna qualquer p *de* $\mathbf{A}$*, e desenvolver o determinante sobre esta coluna:*

$$\det \mathbf{A} = a_{1p}\Delta_{1p} + a_{2p}\Delta_{2p} + \cdots + a_{np}\Delta_{np}.$$

Exemplo 3.2. Considere a matriz do exemplo anterior

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 3 & -1 & -2 \end{bmatrix}.$$

- Escolhendo a primeira linha, temos

$$\det \mathbf{A} = 1\Delta_{11} + 2\Delta_{12} + (-1)\Delta_{13} = 0 + 12 + 3 = 15$$

(esses cofatores foram calculados no exemplo anterior).

- Escolhendo a linha 2, temos

$$\det \mathbf{A} = 0\Delta_{21} + 1\Delta_{22} + 2\Delta_{23}.$$

Como

$$\Delta_{22} = (-1)^{2+2} \det \begin{bmatrix} 1 & -1 \\ 3 & -2 \end{bmatrix} = 1, \qquad \Delta_{23} = (-1)^{2+3} \det \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} = 7,$$

temos

$$\det \mathbf{A} = 1 + 14 = 15.$$

- Escolhendo a coluna 1, temos

$$\det \mathbf{A} = 1\Delta_{11} + 0\Delta_{21} + 3\Delta_{31},$$

com

$$\Delta_{11} = (-1)^{1+1} \det \begin{bmatrix} 1 & 2 \\ -1 & -2 \end{bmatrix} = 0, \qquad \Delta_{31} = (-1)^{3+1} \det \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} = 5,$$

e assim

$$\det \mathbf{A} = 0 + 15 = 15.$$

□

O leitor atento deve ter percebido o seguinte:

se podemos escolher qualquer linha ou coluna para desenvolver o determinante, é melhor escolhermos a linha/coluna que contenha mais zeros, pois desta forma, teremos que calcular menos (sub)determinantes!

É o que o exemplo a seguir ilustra.

Exemplo 3.3. Vamos calcular $\det \mathbf{A}$ onde $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 2 & 1 & 2 \\ 1 & 0 & 3 & 4 \\ 2 & 0 & 1 & -1 \end{bmatrix}$. É interessante aplicar o desenvolvimento de Laplace sobre uma linha ou coluna com mais zeros. Podemos escolher então a linha 1, que contém dois zeros. Assim,

$$\det \mathbf{A} = 1\Delta_{11} + 0\Delta_{12} + 0\Delta_{13} + (-1)\Delta_{14} = \Delta_{11} - \Delta_{14}.$$

Temos

$$\Delta_{11} = (-1)^{1+1} \det \begin{bmatrix} 2 & 1 & 2 \\ 0 & 3 & 4 \\ 0 & 1 & -1 \end{bmatrix} = 1 \cdot 2 \cdot (-1)^{1+1} \det \begin{bmatrix} 3 & 4 \\ 1 & -1 \end{bmatrix} = -14,$$

$$\Delta_{14} = (-1)^{1+4} \det \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 3 \\ 2 & 0 & 1 \end{bmatrix} = (-1) \cdot 2 \cdot (-1)^{1+2} \det \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} = -10,$$

e logo $\det \mathbf{A} = -14 + 10 = -4$.

Também podemos escolher a coluna 2, que contém 3 zeros, para escrever

$$\det \mathbf{A} = 2\Delta_{22} = 2 \cdot (-1)^{2+2} \cdot \det \begin{bmatrix} 1 & 0 & -1 \\ 1 & 3 & 4 \\ 2 & 1 & -1 \end{bmatrix} = \cdots = -4.$$

□

Atividade 3.1. Faça a conta pendente no exemplo anterior. Veja que ao escolher a linha 1 de **A**, devemos calcular dois determinantes de ordem 3, mas com colunas contendo dois zeros... por outro lado escolhendo a coluna 2 de **A**, devemos calcular apenas um determinante de ordem 3, porém com menos zeros. Qual você prefere?
Ou seja, **a escolha de qual linha/coluna usar é uma questão de bom senso!**

3.1.1 Casos particulares

O determinante em alguns casos pode ser calculado com mínimo esforço. Vejamos alguns exemplos:

Determinantes de matrizes com uma linha/coluna nula. Se uma matriz quadrada **R** de ordem n tem uma linha nula então $\det \mathbf{R} = 0$. De fato, digamos que tal linha seja a k. Daí,

$$\det \mathbf{R} = 0\Delta_{k1} + 0\Delta_{k2} + \cdots + 0\Delta_{kn} = 0.$$

O mesmo ocorre se **R** tem uma coluna nula: basta aplicar o desenvolvimento de Laplace sobre a tal coluna. Em particular, o determinante de uma matriz nula é zero.

Exemplo 3.4. $\det \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 1 & 0 & 3 & 4 \\ 2 & 0 & 1 & -1 \end{bmatrix} = 0$, dado que a segunda coluna é nula. □

Determinantes de matrizes triangulares. Seja

$$\mathbf{T} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & \cdots & t_{1n} \\ 0 & t_{22} & t_{23} & \dots & t_{2n} \\ 0 & 0 & t_{33} & \dots & t_{3n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \cdots & t_{nn} \end{bmatrix}$$

uma matriz triangular superior de ordem n. Afirmamos que seu determinante é o produto $t_{11}t_{22}t_{33}\cdots t_{nn}$ dos elementos de sua diagonal. De fato, escolhendo a primeira coluna obtemos

$$\det \mathbf{T} = t_{11}\Delta_{11} = t_{11} \det \begin{bmatrix} t_{22} & t_{23} & \dots & t_{2n} \\ 0 & t_{33} & \dots & t_{3n} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{bmatrix}.$$

Agora, escolhendo a primeira coluna da submatriz de ordem $n-1$, obtemos

$$\det \mathbf{T} = t_{11} t_{22} \det \begin{bmatrix} t_{33} & \dots & t_{3n} \\ \vdots & \dots & \vdots \\ 0 & \cdots & t_{nn} \end{bmatrix}.$$

Continuando a escolher sucessivamente as primeiras colunas das submatrizes triangulares, obteremos, depois de n passos,

$$\det \mathbf{T} = t_{11} t_{22} t_{33} \cdots t_{nn}.$$

O mesmo ocorre vale para matrizes triangulares inferiores: basta aplicar o desenvolvimento de Laplace sobre as primeiras linhas.

Exemplo 3.5. $\det \begin{bmatrix} 1 & 56 & 23 & 11 \\ 0 & -2 & \frac{3}{5} & 2 \\ 0 & 0 & 3 & -7 \\ 0 & 0 & 0 & -1 \end{bmatrix} = 1 \cdot (-2) \cdot 3 \cdot (-1) = 6.$ □

Determinantes de matrizes diagonais. Uma matriz diagonal é, em particular, uma matriz triangular. Portanto seu determinante é igual ao produto dos elementos da diagonal. Em particular, o determinante da matriz identidade de ordem n é 1:

$$\det \mathbf{I} = 1 \cdot 1 \cdots 1 = 1.$$

3.2 Determinantes e matrizes elementares

Uma outra forma de calcular determinantes é através de matrizes elementares. Iniciamos estabelecendo o determinante de cada tipo matriz elementar.

Teorema 3.2 (Determinantes de matrizes elementares)**.** *Temos*

(*i*) $\det \mathbf{E} = -1$ *onde* $\mathbf{E}$ *é a matriz elementar relativa à operação* $L_i \leftrightarrow L_j$ $(i \neq j)$*;*

(*ii*) $\det \mathbf{E} = k$ *onde* $\mathbf{E}$ *é a matriz elementar relativa à operação* $L_i \to kL_i$*;*

(*iii*) $\det \mathbf{E} = 1$ *onde* $\mathbf{E}$ *é a matriz elementar relativa à operação* $L_i \to L_i + kL_j$ $(i \neq j)$*.*

O Teorema anterior estabelece o determinante de **qualquer** matriz elementar, independentemente de sua ordem ou da escolha da operação.
Assim,

- **qualquer** matriz relativa à troca de linhas sempre terá determinante igual a -1;
- **qualquer** matriz relativa à multiplicação de uma linha por k sempre terá determinante igual a k;
- e **qualquer** matriz relativa à operações do tipo $L_i \to L_i + kL_j$ sempre terá determinante igual a 1.

Vimos anteriormente que o processo de escalonamento e a inversão de matrizes envolvem produtos por matrizes elementares. Portanto, é razoável que olhemos para o determinante do produto de matrizes.

Teorema 3.3. *Seja* $\mathbf{A}$ *uma matriz de ordem* n *e* $\mathbf{E}$ *uma matriz elementar também de ordem* n. *Então*

$$\det(\mathbf{EA}) = \det \mathbf{E} \cdot \det \mathbf{A}.$$

Ou seja, "o determinante do produto de uma matriz por uma elementar é o produto dos determinantes".

Atividade 3.2. Se convença do teorema anterior fazendo exemplos com matrizes 3×3.

Com os dois teoremas anteriores, podemos calcular o determinante de uma matriz após realizada operações elementares sobre suas linhas. Isso pode facilitar o cálculo de alguns determinantes.

Exemplo 3.6. Vamos calcular $\det \mathbf{A}$ onde $\mathbf{A} = \begin{bmatrix} 2 & 4 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix}$ usando o histórico de operações elementares da obtenção de sua MLRFE.

Primeiro, calculemos a MLRFE de $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 2 \\ -1 & 0 & 1 \\ \boxed{1} & 1 & -1 \end{bmatrix} \xrightarrow[L_1 \to 1/2 L_1]{L_3 \to L_3 + L_2} \begin{bmatrix} 1 & 2 & 1 \\ \boxed{-1} & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{L_2 \to L_2 + L_1}$$

$$\begin{bmatrix} 1 & \boxed{2} & 1 \\ 0 & \boxed{2} & 2 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow[L_2 \to L_2 - 2L_3]{L_1 \to L_1 - 2L_3} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 0 & \boxed{1} & 0 \end{bmatrix} \xrightarrow[L_2 \leftrightarrow L_3]{L_2 \to 1/2 L_2}$$

$$\begin{bmatrix} 1 & 0 & \boxed{1} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{L_1 \to L_1 - L_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}_3.$$

Interpretamos a aplicação das operações elementares como produto por matrizes elementares:

1ª oper.: $\mathbf{E}_{L_3 \to L_3 + L_2} \mathbf{A}$

2ª oper.: $\mathbf{E}_{L_1 \to \frac{1}{2} L_1} \mathbf{E}_{L_3 \to L_3 + L_2} \mathbf{A}$

3ª oper.: $\mathbf{E}_{L_2 \to L_2 + L_1} \mathbf{E}_{L_1 \to \frac{1}{2} L_1} \mathbf{E}_{L_3 \to L_3 + L_2} \mathbf{A}$

4ª oper.: $\mathbf{E}_{L_1 \to L_1 - 2L_3} \mathbf{E}_{L_2 \to L_2 + L_1} \mathbf{E}_{L_1 \to \frac{1}{2} L_1} \mathbf{E}_{L_3 \to L_3 + L_2} \mathbf{A}$

5ª oper.: $\mathbf{E}_{L_2 \to L_2 - 2L_3} \mathbf{E}_{L_1 \to L_1 - 2L_3} \mathbf{E}_{L_2 \to L_2 + L_1} \mathbf{E}_{L_1 \to \frac{1}{2} L_1} \mathbf{E}_{L_3 \to L_3 + L_2} \mathbf{A}$

6ª oper.: $\mathbf{E}_{L_2 \to \frac{1}{2} L_2} \mathbf{E}_{L_2 \to L_2 - 2L_3} \mathbf{E}_{L_1 \to L_1 - 2L_3} \mathbf{E}_{L_2 \to L_2 + L_1} \mathbf{E}_{L_1 \to \frac{1}{2} L_1} \mathbf{E}_{L_3 \to L_3 + L_2} \mathbf{A}$

7ª oper.: $\mathbf{E}_{L_2 \leftrightarrow L_3} \mathbf{E}_{L_2 \to \frac{1}{2} L_2} \mathbf{E}_{L_2 \to L_2 - 2L_3} \mathbf{E}_{L_1 \to L_1 - 2L_3} \mathbf{E}_{L_2 \to L_2 + L_1} \mathbf{E}_{L_1 \to \frac{1}{2} L_1} \mathbf{E}_{L_3 \to L_3 + L_2} \mathbf{A}$

8ª oper.: $\mathbf{E}_{L_1 \to L_1 - L_3} \mathbf{E}_{L_2 \leftrightarrow L_3} \mathbf{E}_{L_2 \to \frac{1}{2} L_2} \mathbf{E}_{L_2 \to L_2 - 2L_3} \mathbf{E}_{L_1 \to L_1 - 2L_3} \mathbf{E}_{L_2 \to L_2 + L_1} \mathbf{E}_{L_1 \to \frac{1}{2} L_1} \mathbf{E}_{L_3 \to L_3 + L_2} \mathbf{A} = \mathbf{I}_3$

Usando o Teorema 3.3 sucessivas vezes na última igualdade, obtemos

$$1 \cdot (-1) \cdot \frac{1}{2} \cdot 1 \cdot 1 \cdot 1 \cdot \frac{1}{2} \cdot 1 \cdot \det \mathbf{A} = \det \mathbf{I}_3 \quad \Rightarrow \quad \det \mathbf{A} = -4.$$

$\square$

3.3 Propriedades adicionais dos determinantes

Seja $\mathbf{A}$ uma matriz de ordem n. São fatos sobre determinantes:

1. $$\det \mathbf{A} = \det \mathbf{A}^t.$$

2. $$\det(\mathbf{AB}) = \det \mathbf{A} \cdot \det \mathbf{B}$$

   (o determinante do produto é o produto dos determinantes)

   Justificativa. Dividimos a prova em dois casos:

   CASO 1: AB não inversível.

   Neste caso, $\det(\mathbf{AB}) = 0$. Mais ainda, $\mathbf{A}$ ou $\mathbf{B}$ não é inversível, pois se $\mathbf{A}$ e $\mathbf{B}$ fossem inversíveis então $\mathbf{AB}$ também seria com $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. Assim, $\det \mathbf{A} = 0$ ou $\det \mathbf{B} = 0$, e temos $\det(\mathbf{AB}) = 0 = \det \mathbf{A} \cdot \det \mathbf{B}$.

   CASO 2: AB inversível.

   Sendo $\mathbf{AB}$ inversível, $\mathbf{A}$ também é (exercício 6). Logo $\mathbf{A}$ é um produto de matrizes elementares $\mathbf{A} = \mathbf{E}_k \cdots \mathbf{E}_1$ (exercício 9) e daí

   $$\begin{aligned}\det(\mathbf{AB}) &= \det(\mathbf{E}_k \cdots \mathbf{E}_1\mathbf{B}) = \det \mathbf{E}_k \cdot \det(\mathbf{E}_{k-1} \cdots \mathbf{E}_1\mathbf{B}) \\ &= \cdots = \det \mathbf{E}_k \cdots \det \mathbf{E}_1 \cdot \det \mathbf{B} = \det(\mathbf{E}_k E_{k-1}) \cdots \det \mathbf{E}_1 \cdot \det \mathbf{B} \\ &= \cdots = \det(\mathbf{E}_k \cdots \mathbf{E}_1) \cdot \det \mathbf{B} = \det \mathbf{A} \cdot \det \mathbf{B}.\end{aligned}$$

   □

3. Se $\mathbf{A}$ é inversível então $\det \mathbf{A}^{-1} = \dfrac{1}{\det \mathbf{A}}$.

   Justificativa. $1 = \det \mathbf{I} = \det(\mathbf{AA}^{-1}) = \det \mathbf{A} \cdot \det \mathbf{A}^{-1}$, donde segue o resultado. □

4. Se $\mathbf{A}$ tem duas linhas ou duas colunas iguais i e j $(i \neq j)$ então $\det \mathbf{A} = 0$.

   Justificativa. De fato, seja $\mathbf{E}$ a matriz elementar relativa à operação $L_i \to L_i - L_j$. Então a linha i de $\mathbf{EA}$ é nula, e logo

   $$\det \mathbf{A} = \underbrace{\det \mathbf{E}}_{1} \cdot \det \mathbf{A} = \det(\mathbf{EA}) = 0.$$

   O caso em que $\mathbf{A}$ tem duas colunas iguais segue do fato de $\mathbf{A}^t$ ter duas linhas iguais e $\det \mathbf{A} = \det \mathbf{A}^t$. □

A propriedade 3 acima tem relação com o resultado abaixo:

Teorema 3.4 (Inversibilidade é equivalente a determinante não nulo)**.** *Seja* $\mathbf{A}$ *uma matriz de ordem* n*. Então* $\det \mathbf{A} \neq 0$ *se, e somente se* $\mathbf{A}$ *é inversível.*

O Teorema acima diz que uma matriz tem inversa se seu determinante é diferente de zero, e vice-versa.

Exemplo 3.7. Vimos no Exemplo 3.6 que o determinante da matriz

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

é igual a -4. Então essa matriz tem inversa. De fato, no Exemplo 3.6 vimos que ela é linha equivalente à identidade (isto é, sua MLRFE é I). Podemos usar o processo de inversão de matrizes por operações elementares visto no Capítulo 2 para calcular sua inversa (exercício!!!). $\square$

3.4 Demonstrações de resultados

Demonstração do Teorema 3.2. (i): Suponha que $\mathbf{E}$ seja a matriz elementar relativa à $L_1 \leftrightarrow L_2$. A primeira linha de $\mathbf{E}$ é

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}. \tag{3.2}$$

Aplicando o desenvolvimento de Laplace nesta linha, obtemos

$$\det \mathbf{E} = e_{12}\Delta_{12} = (-1)^{1+2} \det \mathbf{I}_{n-1} = -1.$$

A fim de mostrar a afirmação para qualquer operação $L_i \leftrightarrow L_j$, basta aplicar sucessivamente o desenvolvimento de Laplace da primeira até a linha j. A submatriz resultante terá primeira linha no formato (3.2), recaindo no caso anterior.

(ii): Se $\mathbf{E}$ é a matriz elementar relativa à operação $L_i \to kL_i$, aplicamos o desenvolvimento de Laplace sobre a linha i, obtendo

$$\det \mathbf{E} = e_{ii}\Delta_{ii} = k \det \mathbf{I}_{n-1} = k.$$

(iii): Se $\mathbf{E}$ é a matriz elementar relativa à operação $L_i \to L_i + kL_j$ então a linha i de $\mathbf{E}$ é

$$\begin{bmatrix} 0 & \cdots & 0 & \underbrace{1}_{\text{coluna } i} & 0 & \cdots & 0 & \underbrace{k}_{\text{coluna } j} & 0 & \cdots & 0 \end{bmatrix}.$$

Aplicando o desenvolvimento de Laplace sobre esta linha i obtemos

$$\det \mathbf{E} = e_{ii}\Delta_{ii} + e_{ij}\Delta_{ij} = 1 \det \mathbf{I}_{n-1} + k \det \mathbf{E}_{ij},$$

onde a matriz $\mathbf{E}_{ij}$, de ordem $n-1$, tem a linha j nula (ao retirar a coluna j, elimina-se o 1 da linha j; faça um exemplo 4×4 e se convença disso). Portanto $\det \mathbf{E} = 1 + 0 = 1$. $\square$

Demonstração do Teorema 3.3. A prova é similar à do Teorema anterior. Analisamos cada tipo de operação elementar separadamente.

(i): Suponha que $\mathbf{E}$ seja a matriz elementar relativa à $L_1 \leftrightarrow L_2$. O produto $\mathbf{EA}$ troca as linhas 1 e 2 da matriz $\mathbf{A}$, ou seja, $\mathbf{EA}$ tem a forma

$$\mathbf{EA} = \begin{bmatrix} a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{12} & a_{12} & a_{13} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}. \tag{3.3}$$

Sejam $\tilde{\Delta}_{ij}$'s os cofatores relativos à matriz $\mathbf{EA}$ e Δ_{ij}'s os relativos à matriz original $\mathbf{A}$. Observe que

$$\tilde{\Delta}_{1j} = (-1)^{1+j} \det \mathbf{A}_{2j} = -\Big[(-1)^{2+j} \det \mathbf{A}_{2j}\Big] = -\Delta_{2j},$$

para todo j. Assim, aplicando o desenvolvimento de Laplace sobre a linha 1 de $\mathbf{EA}$ obtemos

$$\det(\mathbf{EA}) = \sum_{j=1}^{m} a_{2j}\tilde{\Delta}_{1j} = -\sum_{j=1}^{m} a_{2j}\Delta_{2j}.$$

A última soma é o desenvolvimento de Laplace sobre a linha 2 de $\mathbf{A}$ com sinal trocado. Como $\det \mathbf{E} = -1$ (Teorema 3.2(i)), segue que $\det(\mathbf{EA}) = \det \mathbf{E} \cdot \det \mathbf{A}$.

A prova para uma operação $L_i \leftrightarrow L_j$ qualquer segue a mesma ideia.

(ii): Se $\mathbf{E}$ é a matriz elementar relativa à operação $L_i \to kL_i$, a linha i de $\mathbf{EA}$ estará multiplicada por k. Aplicamos o desenvolvimento de Laplace sobre esta linha i e usamos o Teorema 3.2(ii), obtendo

$$\det(\mathbf{EA}) = \sum_{j=1}^{m} (ka_{ij})\Delta_{ij} = k\sum_{j=1}^{m} a_{ij}\Delta_{ij} = k \det \mathbf{A} = \det \mathbf{E} \cdot \det \mathbf{A}.$$

(iii): Se $\mathbf{E}$ é a matriz elementar relativa à operação $L_i \to L_i + kL_j$ então a linha i de $\mathbf{EA}$ é

$$\begin{bmatrix} a_{i1} + ka_{j1} & a_{i2} + ka_{j2} & \cdots & a_{in} + ka_{jn} \end{bmatrix}.$$

Aplicando o desenvolvimento de Laplace sobre esta linha i obtemos

$$\det(\mathbf{EA}) = \sum_{\ell=1}^{m} (a_{i\ell} + ka_{j\ell})\Delta_{i\ell} = \sum_{\ell=1}^{m} a_{i\ell}\Delta_{i\ell} + k\sum_{\ell=1}^{m} a_{j\ell}\Delta_{i\ell}.$$

A primeira soma do lado direito da última igualdade é o determinante de $\mathbf{A}$ desenvolvido sobre a linha i; já a segunda soma é o múltiplo k do determinante de uma matriz com duas linhas j de $\mathbf{A}$. Portanto essa soma é nula (veja a propriedade 4 no texto). Como neste caso temos $\det \mathbf{E} = 1$ (Teorema 3.2(iii)), concluímos que $\det(\mathbf{EA}) = \det \mathbf{A} = \det \mathbf{E} \cdot \det \mathbf{A}$. □

Demonstração do Teorema 3.4. Se $\mathbf{A}$ é inversível então $\mathbf{A} = \mathbf{E}_k \cdots \mathbf{E}_1$ para certas matrizes elementares $\mathbf{E}_1, \ldots, \mathbf{E}_k$ (exercício 9). Assim, $\det \mathbf{A} = \det(\mathbf{E}_k \cdots \mathbf{E}_1) = \det \mathbf{E}_k \cdots \det \mathbf{E}_1 \neq 0$. Reciprocamente, se $\mathbf{A}$ não é inversível, então $\mathbf{A} = \mathbf{E}_k \cdots \mathbf{E}_1\mathbf{R}$, onde $\mathbf{R}$ é a MLRFE de $\mathbf{A}$ com uma linha nula (ou seja, $\mathbf{R} \neq \mathbf{I}$). Assim, $\det \mathbf{A} = \det \mathbf{E}_k \cdots \det \mathbf{E}_1 \cdot \underbrace{\det \mathbf{R}}_{0} = 0$. □

3.5 Exercícios

1. Dadas $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ e $\mathbf{B} = \begin{bmatrix} 3 & -1 \\ 0 & 1 \end{bmatrix}$, calcule $\det \mathbf{A} + \det \mathbf{B}$ e $\det(\mathbf{A} + \mathbf{B})$, e conclua que, em geral, o determinante da soma de matrizes não é igual a soma dos determinantes.

2. Calcule o posto e a nulidade das matrizes abaixo, encontrando a matriz linha reduzida à forma escada de cada matriz. Use o histórico de operações elementares para calcular o determinante de cada uma das matrizes.

   (a) $\begin{bmatrix} 1 & -2 & 3 \\ 2 & 2 & 1 \\ 2 & -4 & 6 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 1 & 2 \\ -1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 3 & 1 & -1 & 3 \end{bmatrix}$

3. Admitindo que $\det \mathbf{A} = 5$ onde $\mathbf{A} = [a_{ij}]_{3\times 3}$, calcule:

   (a) $\det(3\mathbf{A})$

   (b) $\det(2\mathbf{A}^{-1})$

   (c) $\det((2\mathbf{A})^{-1})$

   (d) $\det \begin{bmatrix} a_{11} & a_{31} & a_{21} \\ a_{12} & a_{32} & a_{22} \\ a_{13} & a_{33} & a_{23} \end{bmatrix}$

4. Calcule os determinantes:

   (a) $\det \begin{bmatrix} a+b & b+d \\ a+c & c+d \end{bmatrix}$

   (b) $\det \begin{bmatrix} a+b & a-b \\ a-b & a+b \end{bmatrix}$

   (c) $\det \begin{bmatrix} x-1 & 1 \\ x^3 & x^2+x+1 \end{bmatrix}$

   (d) $\det \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$

   (e) $\det \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}$

   (f) $\det \begin{bmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix}$

   (g) $\det \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix}$

   (h) $\det \begin{bmatrix} 1 & 2 & 3 & 4 \\ -2 & 1 & -4 & 3 \\ 3 & -4 & -1 & 2 \\ 4 & 3 & -2 & -1 \end{bmatrix}$

   (i) $\det \begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 & 1 \\ 1 & 1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 5 & 1 \\ 1 & 1 & 1 & 1 & 6 \end{bmatrix}$

   (j) $\det \begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 4 & -1 & 3 \\ 2 & 1 & 2 & 13 & 7 \\ 2 & 1 & 1 & 2 & 9 \\ 2 & 1 & 1 & 1 & 2 \end{bmatrix}$

5. Sejam $\mathbf{A}$ e $\mathbf{P}$ matrizes quadradas de mesma ordem com P inversível. Mostre que $\det(\mathbf{P}\mathbf{A}\mathbf{P}^{-1}) = \det \mathbf{A}$ (as matrizes $\mathbf{P}\mathbf{A}\mathbf{P}^{-1}$ e $\mathbf{A}$ são chamadas *matrizes semelhantes*).

6. Seja $\mathbf{B}$ uma matriz de ordem n tal que $\mathbf{B}^t\mathbf{B} = \mathbf{I}_n$. Mostre que $\det \mathbf{B} = 1$ ou $\det \mathbf{B} = -1$.

7. Diz-se que uma matriz quadrada $\mathbf{A}$ é *idempotente* quando $\mathbf{A}^2 = \mathbf{A}$. Mostre que se $\mathbf{A}$ é idempotente então $\det \mathbf{A} = 1$ ou $\det \mathbf{A} = 0$.

8. Dada uma matriz quadrada $\mathbf{A}$ de ordem n, definimos a *matriz ajdunta* de $\mathbf{A}$ como sendo a transposta da matriz dos cofatores dos elementos a_{ij} de $\mathbf{A}$. Denotaremos a adjunta de $\mathbf{A}$ por $\operatorname{adj} \mathbf{A}$. Em outras palavras,

$$\operatorname{adj} \mathbf{A} = \begin{bmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{bmatrix}.$$

(a) Considere a matriz $\mathbf{C} = \mathbf{A}\,(\operatorname{adj}\mathbf{A})$. Mostre que os elementos c_{ij} de $\mathbf{C}$ são tais que

$$c_{ii} = \det\mathbf{A},\ \forall i \quad \text{e} \quad c_{ij} = 0,\ i \neq j,$$

isto é, que $\mathbf{A}\,(\operatorname{adj}\mathbf{A}) = (\det\mathbf{A})\,\mathbf{I}_n$.

Dica: os elementos c_{ij} são da forma $c_{ij} = a_{i1}\Delta_{j1} + a_{i2}\Delta_{j2} + \cdots + a_{in}\Delta_{jn}$. Quando $i = j$ então c_{ii} é o desenvolvimento de Laplace da matriz $\mathbf{A}$ escolhendo a linha i, e logo $c_{ii} = \det\mathbf{A}$. Quando $i \neq j$, c_{ij} é o desenvolvimento de Laplace de uma matriz com duas linhas iguais a linha i de $\mathbf{A}$ (e logo cujo determinante é 0).

(b) Usando o item anterior, mostre que se $\mathbf{A}$ é inversível então $\mathbf{A}^{-1} = \dfrac{1}{\det\mathbf{A}}(\operatorname{adj}\mathbf{A})$.

(c) Calcule a inversa de

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

usando a adjunta de $\mathbf{A}$.

9. **(Regra de Cramer)** Considere um sistema linear $\mathbf{AX} = \mathbf{B}$, onde $\mathbf{A}$ é matriz quadrada de ordem n (o sistema tem n equações e n incógnitas). Suponha que $\det\mathbf{A} \neq 0$ (ou equivalentemente, que $\mathbf{A}$ é inversível).

(a) Mostre que $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$ é a única solução do sistema $\mathbf{AX} = \mathbf{B}$.

(b) Usando o fato de que $\mathbf{A}^{-1} = \dfrac{1}{\det\mathbf{A}}(\operatorname{adj}\mathbf{A})$ (exercício anterior), mostre que o termo i da solução $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$ é dado por

$$x_i = \frac{b_1\Delta_{1i} + b_2\Delta_{2i} + \cdots + b_n\Delta_{ni}}{\det\mathbf{A}} = \frac{\det\mathbf{C}_i}{\det\mathbf{A}},$$

onde $\mathbf{C}_i$ é a matriz obtida de $\mathbf{A}$ pela substituição da coluna i pela matriz coluna $\mathbf{B}$, isto é,

$$\mathbf{C}_i = \begin{bmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{bmatrix}.$$

Essa técnica de resolução de sistemas lineares $n \times n$ é conhecida como *Regra de Cramer.*

(c) Resolva o sistema $\left\{\begin{array}{rrrcl} 2x & -3y & +7z & = & 1 \\ x & & +3z & = & 5 \\ & 2y & -z & = & 0 \end{array}\right.$ pela Regra de Cramer.

OBSERVAÇÃO:

A regra de Cramer está presente em vários livros de Álgebra Linear, inclusive no livro do Boldrini. Porém este método de resolução de sistemas lineares não é usado na prática computacional. Há pelo menos os seguintes motivos:

- só se aplica a sistemas quadrados (número de equações igual ao número de incógnitas);
- não se aplica a sistemas sem solução ou com infinitas soluções, somente àqueles com única solução;

- tem custo computacional exagerado. Veja que a regra de Cramer envolve o cálculo de vários determinantes, e cada determinante consiste em um escalonamento (pense no determinante como produto de matrizes elementares). Por outro lado, a resolução do sistema linear original pode ser feita realizando apenas um processo de escalonamento sobre a matriz ampliada;

- é instável numericamente, dado que erros de arredondamento se acumulam no cálculo dos vários determinantes.

Por esses motivos, não discutimos a regra de Cramer. Este método tem apenas caráter histórico e teórico.

3.6 Notas

Nos livros de geometria analítica, determinantes são usados para calcular áreas e volumes de certas figuras geométricas. Para além do uso descrito nos livros básicos, determinantes aparecem em teorias mais sofisticadas. Um exemplo é o estudo de problemas de programação linear inteira (PLI). Um PLI consiste em encontrar valores mínimos de uma função $f(\mathbf{X}) = \mathbf{C}^t\mathbf{X}$, onde $\mathbf{C}$ é uma matriz coluna, restrito à $\mathbf{X}$ ser solução de um sistema $\mathbf{AX} = \mathbf{B}$ com $\mathbf{X} \geq \mathbf{0}$, $\mathbf{X} \in \mathbb{Z}^n$. Escrevemos o PLI como

$$\min \mathbf{C}^t\mathbf{X} \quad \text{sujeito a} \quad \mathbf{AX} = \mathbf{B},\ \mathbf{X} \geq \mathbf{0},\ \mathbf{X} \in \mathbb{Z}^n$$

(supomos $\mathbf{A}$ matriz $m \times n$ com posto $\mathbf{A} \leq m < n$). A maior dificuldade em resolver este problema é garantir que as componentes de uma solução x sejam todas inteiras. Costuma-se então relaxar o problema retirando essa exigência:

$$\min \mathbf{C}^t\mathbf{X} \quad \text{sujeito a} \quad \mathbf{AX} = \mathbf{B},\ \mathbf{X} \geq \mathbf{0}.$$

O problema acima é muito mais fácil de resolver!!! Porém, em geral uma solução sua pode não ser toda inteira... Bem, o fato é que isso não é uma preocupação se $\mathbf{A}$ for **totalmente unimodular**. Uma matriz é totalmente unimodular se toda submatriz quadrada sua tem determinante 0, 1 ou -1. É possível provar que se $\mathbf{A}$ é totalmente unimodular, então o problema relaxado admite soluções com todas as componentes inteiras! Ou seja, neste caso resolvemos um PLI com esforço computacional relativamente baixo. Esse tipo de situação aparece em problemas de "fluxo em rede", importantes em aplicações como projeto de redes de comunicação e otimização de rotas urbanas.

O artigo a seguir (um pouco antigo, é verdade) traz outros usos interessantes dos determinantes:

P.R.Vein. A short survey of some recent applications of determinants. Linear Algebra and its Applications, 42, p 287-297, 1982.

Capítulo 4

Espaços Vetoriais

4.1 Motivação

Espaços vetoriais podem ser vistos como uma generalização da geometria de vetores no plano e no espaço, geralmente vistos em cursos de Geometria Analítica.

- Veja a introdução do Capítulo 4 do livro do Boldrini;
- Assista aos vídeos indicados no AVA;
- Recapitule a primeira aula síncrona sobre o tema.

4.2 Espaços Vetoriais

Definição 4.1. *Um* ***espaço vetorial*** *é constituído de um conjunto V, cujos elementos são chamados* vetores*, no qual estão definidas duas operações:*

- *a* ***soma*** *(ou* ***adição****) que leva um par de vetores $(u, w) \in V \times V$ a um vetor $u + w \in V$;*
- *a* ***multiplicação por escalar****, que leva um par $(a, u) \in \mathbb{R} \times V$ a um vetor $au \in V$,*

que devem satisfazer as seguintes propriedades para quaisquer $a, b \in \mathbb{R}$ e $u, v, w \in V$:

(i) $u + w = w + u$ *(comutatividade)*

(ii) $(u + v) + w = u + (v + w)$ *e* $(ab)u = a(bu)$ *(associatividade)*

(iii) existe um vetor $\mathbf{0} \in V$, chamado vetor nulo tal que $u + \mathbf{0} = \mathbf{0} + u = u$, para todo $u \in V$ (existência de vetor nulo)

(iv) $(a + b)u = au + bu$ *e* $a(u + w) = au + aw$ *(distributividade)*

(v) $1 \cdot u = u$ *(multiplicação por 1)*

(vi) para todo $u \in V$, existe um vetor $-u \in V$ tal que $u + (-u) = \mathbf{0}$ (inverso aditivo)

Essas propriedades “imitam” a geometria de vetores no plano. No entanto, observe que o conjunto V pode ser formado de elementos muito mais gerais, tais como listas ordenadas de n escalares (vetores no $\mathbb{R}^n$), matrizes, ou até mesmo polinômios e funções!

Exemplo 4.1. 1. O conjunto $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$ com as operações usuais

$$(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2) \qquad \text{e} \qquad a(u_1, u_2) = (au_1, au_2).$$

2. O conjunto $\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$ com as operações usuais

$$(u_1, u_2, u_3)+(v_1, v_2, v_3) = (u_1+v_1, u_2+v_2, u_3+v_3) \qquad \text{e} \qquad a(u_1, u_2, u_3) = (au_1, au_2, au_3).$$

3. O conjunto $\mathbb{R}^n = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R}, \forall i\}$ com as operações usuais

$$(x_1, \ldots, x_n)+(y_1, \ldots, y_n) = (x_1+y_1, \ldots, x_n+y_n) \qquad \text{e} \qquad a(x_1, \ldots, x_n) = (ax_1, \ldots, ax_n).$$

Note que essas operações coincidem com a geometria de vetores no plano e no espaço.

4. O conjunto $M(m, n)$ das matrizes de ordem $m \times n$ com as operações usuais definidas anteriormente, a saber, a soma de matrizes e a multiplicação por escalar:

$$(A + B)_{ij} = a_{ij} + b_{ij}, \qquad (\alpha A)_{ij} = \alpha a_{ij}.$$

5. O conjunto P_n dos polinômios de grau $\leq n$ mais o polinômio nulo, munido das operações

$$(p + q)(x) = p(x) + q(x) \qquad \text{e} \qquad (ap)(x) = ap(x).$$

Note que $p + q$ e ap são polinômios. Por exemplo, se $p(x) = x^3 - 2x^2 + 4$ e $q(x) = 3x^3 + x^2 - 4x$, então

$$(p+q)(x) = 4x^3 - x^2 - 4x + 4, \quad (2p)(x) = 2x^3 - 4x^2 + 8, \quad (3p - q)(x) = -7x^2 + 4x + 12.$$

$\square$

Atividade 4.1. Verifique que os conjuntos do exemplo anterior com as suas operações são espaços vetoriais.

Neste texto, $\mathbf{0}$ (em negrito) será usado para o *vetor nulo*, para distinguir do número zero. Por exemplo,

- em $\mathbb{R}^2$, $\mathbf{0} = (0, 0)$
- em $\mathbb{R}^n$, $\mathbf{0} = (0, \cdots, 0)$
- em $M(m, n)$, $\mathbf{0} =$"matriz nula de ordem $m \times n$"
- em P_n, $\mathbf{0}$ é o polinômio identicamente nulo $p(x) = 0$ para todo x.

São fatos acerca de qualquer espaço vetorial:

1. Vale a "lei do corte", tal como para números reais:

$$\textbf{se} \quad w + u = w + v \quad \textbf{então} \quad u = v.$$

Justificativa: Temos

$$\begin{aligned} u &= \mathbf{0} + u = (-w + w) + u \\ &= -w + (w + u) = -w + (w + v) \\ &= (-w + w) + v = \mathbf{0} + v \\ &= v. \end{aligned}$$

2. **Se $w + u = w$ então $u = \mathbf{0}$.**

   Justificativa: Utilizando a lei do corte vem

$$w + u = w \Rightarrow w + u = w + \mathbf{0} \Rightarrow u = \mathbf{0}.$$

Isso mostra que o vetor nulo é único.

3. **Se $w + u = \mathbf{0}$ então $u = -w$.**

   Justificativa: Utilizando a lei do corte vem

$$w + u = \mathbf{0} \Rightarrow w + u = w + (-w) \Rightarrow u = -w.$$

Isso mostra que o inverso aditivo de um vetor $w \in V$ é único. Em particular, esse fato implica que a propriedade (vi) da definição de espaço vetorial decorre das outras.

4. **Dado $w \in V$, temos $0w = \mathbf{0}$** (a multiplicação de um vetor por zero é o vetor nulo).

   Justificativa: De fato, $w + 0w = 1w + 0w = (1 + 0)w = 1w = w = w + \mathbf{0} \Rightarrow 0w = \mathbf{0}$.

5. **Dado $a \in \mathbb{R}$, temos $a\mathbf{0} = \mathbf{0}$** (qualquer múltiplo do vetor nulo é o próprio vetor nulo).

   Justificativa: De fato, $a\mathbf{0} = a(-w + w) = a(-w) + aw = (-a)w + aw = (-a + a)w = 0w = \mathbf{0}$.

6. **Se $a \neq 0$ e $w \neq \mathbf{0}$ então $aw \neq \mathbf{0}$** (múltiplos não nulos de vetores não nulos são vetores não nulos).

   Justificativa. Suponha por absurdo que $a \neq 0$, $w \neq \mathbf{0}$ e $aw = \mathbf{0}$. Então $w = 1w = \left(\frac{1}{a} \cdot a\right) w = \frac{1}{a}(aw) = \frac{1}{a}\mathbf{0} = \mathbf{0}$, isto é, $w = \mathbf{0}$, um absurdo. $\square$

7. **Dado $w \in V$, temos $(-1)w = -w$.**

   Justificativa. $w + (-1)w = 1w + (-1)w = (1 - 1)w = 0w = \mathbf{0} \Rightarrow (-1)w = -w$. $\square$

4.3 Subespaços vetoriais

Definição 4.2. *Dado um espaço vetorial V, um subconjunto $W \subset V$ é um* ***subespaço vetorial*** *de V se*

(*i*) $u + v \in W$ *para todos* $u, v \in W$ *(W é fechado para a soma)*

(*ii*) $au \in W$ *para todos* $a \in \mathbb{R}$ *e* $u \in W$ *(W é fechado para multiplicação por escalar)*

Naturalmente, as operações em W são as mesmas de V. Assim, um subespaço vetorial W de V é ele mesmo um espaço vetorial com as operações herdadas de V. Observe ainda que o vetor nulo de W é o vetor nulo de V.

Exemplo 4.2. Dado um espaço vetorial V, são subespaços vetoriais *triviais* de V os subconjuntos $\{\mathbf{0}\}$ e o próprio V. □

Exemplo 4.3. O subconjunto

$$W = \{(at, bt, ct) \in \mathbb{R}^3 \mid t \in \mathbb{R}\}$$

é subespaço vetorial de $\mathbb{R}^3$, com as mesmas operações usuais de $\mathbb{R}^3$. Geometricamente, se $(a, b, c) \neq \mathbf{0}$ então W é a reta do espaço na direção do vetor (a, b, c) e que passa pela origem. No caso em que $(a, b, c) = \mathbf{0}$, $W = \{\mathbf{0}\}$. □

Exemplo 4.4. O subconjunto

$$W = \{(x, y, z) \in \mathbb{R}^3 \mid z = 3x, x = 2y\}$$

é subespaço vetorial de $\mathbb{R}^3$. De fato, podemos escrever $W = \{(2y, y, 6y) \in \mathbb{R}^3 \mid y \in \mathbb{R}\}$, e W é a reta na direção de $(2, 1, 6)$ e que passa pela origem. □

Exemplo 4.5. Considere um sistema linear homogêneo $\mathbf{AX} = \mathbf{0}$, onde $\mathbf{A}$ tem ordem $m \times n$. Então o conjunto das soluções $\mathbf{X}$ desse sistema,

$$W = \{\mathbf{X} \in M(n, 1) \mid \mathbf{AX} = \mathbf{0}\},$$

munido das operações usuais sobre matrizes é um subespaço vetorial de $M(n, 1)$. De fato, dadas duas soluções $\mathbf{U}$ e $\mathbf{V}$ de $\mathbf{AX} = \mathbf{0}$ e um escalar a, temos

$$\mathbf{A}(\mathbf{U} + \mathbf{V}) = \mathbf{AU} + \mathbf{AV} = \mathbf{0} + \mathbf{0} = \mathbf{0} \qquad \text{e} \qquad \mathbf{A}(a\mathbf{U}) = a\mathbf{AU} = a\mathbf{0} = \mathbf{0},$$

e logo $\mathbf{U} + \mathbf{V} \in W$ e $a\mathbf{U} \in W$.

Por exemplo, considere o sistema homogêneo

$$\left\{\begin{array}{rrrcl} 2x & +4y & +z & = & 0 \\ x & +y & +2z & = & 0 \\ x & +3y & -z & = & 0 \end{array}\right.,$$

que em forma matricial fica

$$\begin{bmatrix} 2 & 4 & 1 \\ 1 & 1 & 2 \\ 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

As soluções do sistema são matrizes colunas em $M(3,1)$ que satisfazem a equação acima. Sejam

$$\mathbf{U} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \quad \text{e} \quad \mathbf{V} = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$$

duas soluções, e $a \in \mathbb{R}$. Você pode verificar que

$$\begin{bmatrix} 2 & 4 & 1 \\ 1 & 1 & 2 \\ 1 & 3 & -1 \end{bmatrix} \left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 2 & 4 & 1 \\ 1 & 1 & 2 \\ 1 & 3 & -1 \end{bmatrix} \left(a \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

□

Exemplo 4.6. O conjunto das soluções de um sistema **não** homogêneo nem sempre é subespaço vetorial. Para ver isso, considere por exemplo o sistema

$$\begin{cases} x & +4y & +z & = & 1 \\ x & +y & +z & = & 1 \\ & 3y & & = & 0 \end{cases},$$

e suas soluções $\mathbf{U} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^t$ e $\mathbf{V} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^t$ e **verifique que $\mathbf{U}+\mathbf{V}$ não é solução.**
□

Exemplo 4.7. O conjunto $T_S(n)$ das matrizes triangulares superiores de ordem n é subespaço vetorial de $M(n,n)$.

De fato, dadas $\mathbf{U} = [u_{ij}]$ e $\mathbf{V} = [v_{ij}]$ matrizes em $T_S(n)$ temos $u_{ij} = 0$ e $v_{ij} = 0$ sempre que $i > j$. Assim, os elementos c_{ij} de $\mathbf{U}+\mathbf{V}$ são tais que $c_{ij} = u_{ij} + v_{ij} = 0$ sempre que $i > j$, isto é, $\mathbf{U}+\mathbf{V} \in T_S(n)$. Isso mostra que $T_S(n)$ é fechado para a soma.

Agora, seja $a \in \mathbb{R}$. Os elementos d_{ij} de $\alpha\mathbf{U}$ são tais que $d_{ij} = au_{ij} = 0$ sempre que $i > j$, ou seja, $a\mathbf{U} \in T_S(n)$, e $T_S(n)$ é fechado para a multiplicação por escalar.

Podemos ver isso "escrevendo" as matrizes: dadas matrizes $\mathbf{U}, \mathbf{V} \in T_S(n)$

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}, \qquad \mathbf{V} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ 0 & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & v_{nn} \end{bmatrix}$$

e $a \in \mathbb{R}$ quaisquer, você pode notar que as matrizes

$$\mathbf{U}+\mathbf{V} = \begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} & \cdots & u_{1n}+v_{1n} \\ 0 & u_{22}+v_{22} & \cdots & u_{2n}+v_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & u_{nn}+v_{nn} \end{bmatrix}, \qquad a\mathbf{U} = \begin{bmatrix} au_{11} & au_{12} & \cdots & au_{1n} \\ 0 & au_{22} & \cdots & au_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & au_{nn} \end{bmatrix}$$

são também triangulares superiores. □

Exemplo 4.8. O conjunto $T_I(n)$ das matrizes traingulares inferiores de ordem n é subespaço vetorial de $M(n,n)$. A prova é análoga ao exemplo anterior. □

Exemplo 4.9. O subconjunto de P_2 dado por $W = \{p \in P_2 \mid p(0) = 0, p'(1) = 0\}$ é subespaço vetorial dos polinômios de grau ≤ 2 (p' é a derivada de p).

De fato, para que um polinômio

$$p(x) = ax^2 + bx + c$$

esteja em W, devemos ter $p(0) = 0$ e $p'(1) = 0$. A primeira condição implica $c = 0$, e com a segunda condição, chegamos a $2a + b = 0$. Assim concluímos que

$$W = \{ax^2 + bx \mid 2a + b = 0\}.$$

Mostremos agora que W é fechado para as operações em P_2:

(*i*) Se $p(x) = ax^2 + bx$ e $q(x) = \overline{a}x^2 + \overline{b}x$ estão em W, então

$$(p + q)(x) = p(x) + q(x) = (ax^2 + bx) + (\overline{a}x^2 + \overline{b}x) = (a + \overline{a})x^2 + (b + \overline{b})x.$$

Mas $2(a + \overline{a}) + (b + \overline{b}) = (2a + b) + (2\overline{a} + \overline{b}) = 0$, e logo $p + q \in W$.

(*ii*) Se $p(x) = ax^2 + bx$ está em W e $\alpha \in \mathbb{R}$, então

$$(\alpha p)(x) = \alpha p(x) = \alpha(ax^2 + bx) = (\alpha a)x^2 + (\alpha b)x.$$

Mas $2(\alpha a) + \alpha b = \alpha(2a + b) = 0$, e portanto $\alpha p \in W$.

□

Exemplo 4.10. O vetor nulo $\mathbf{0} \in V$ pertence a qualquer subespaço W de V. De fato, W é fechado para multiplicação por escalar (propriedade (ii) da Definição 4.2), e logo $\mathbf{0} = 0u \in W$.

Assim,

$$W = \{(x, y) \in \mathbb{R}^2 \mid x + 2y = 1\}$$

não é subespaço vetorial de $\mathbb{R}^2$, pois $\mathbf{0} = (0, 0) \notin W$. Geometricamente, W representa uma reta que **não** passa pela origem. □

Exemplo 4.11. O conjunto

$$W = \{(x, x^2) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$$

não é subespaço de $\mathbb{R}^2$, apesar de $\mathbf{0} \in W$. De fato, W não é fechado para a soma: tome por exemplo $u = (1, 1)$ e $v = (2, 4)$. Temos $u, v \in W$, mas $u + v = (3, 5) \notin W$. □

Teorema 4.1 (Interseção de subespaços). *Sejam V um espaço vetorial e W_1, W_2 subespaços seus. Então $W_1 \cap W_2$ é subespaço vetorial de V.*

Exemplo 4.12. O conjunto $D(n) = T_S(n) \cap T_I(n)$ é o subespaço vetorial de $M(n, n)$ das matrizes diagonais de ordem n. □

Contrariando as expectativas, a união $W_1 \cup W_2$ de subespaços vetoriais W_1 e W_2 **nem sempre é um subespaço vetorial**.
Por exemplo, $W_1 = \{(t, 0) \mid t \in \mathbb{R}\}$ e $W_2 = \{(0, s) \mid s \in \mathbb{R}\}$ são subespaços de $\mathbb{R}^2$, mas

$$(1, 0), (0, 1) \in W_1 \cup W_2$$

e

$$(1, 0) + (0, 1) = (1, 1) \notin W_1 \cup W_2,$$

o que mostra que $W_1 \cup W_2$ não é fechado para a soma, e portanto não é subespaço.

No entanto, se "unirmos" subespaços **somando vetores** seus, geramos um novo subespaço vetorial. Mais especificamente, temos o

Teorema 4.2 (Soma de subespaços)**.** *Sejam V um espaço vetorial e W_1, W_2 subespaços seus. Então o conjunto*

$$W_1 + W_2 = \{v \in V \mid v = w_1 + w_2,\ w_1 \in W_1,\ w_2 \in W_2\}$$

é subespaço vetorial de V.

O subespaço $W_1 + W_2$ é chamado a *soma* dos subespaços W_1 e W_2. Quando $W_1 \cap W_2 = \{\mathbf{0}\}$ a soma $W_1 + W_2$ é chamada *soma direta*, e escrevemos $W_1 \oplus W_2$.

Exemplo 4.13. Considere os subespaços de $\mathbb{R}^2$ dados por

$$W_1 = \{(t, 0) \mid t \in \mathbb{R}\} \qquad \text{e} \qquad W_2 = \{(0, s) \mid s \in \mathbb{R}\}.$$

Então

$$W_1 + W_2 = \{v \in \mathbb{R}^2 \mid v = (t, 0) + (0, s), t, s \in \mathbb{R}\} = \{(t, s) \mid t, s \in \mathbb{R}\} = \mathbb{R}^2.$$

Ou seja, $\mathbb{R}^2 = W_1 + W_2$. Observe ainda que $W_1 \cap W_2 = \{(0, 0)\}$, e logo $\mathbb{R}^2 = W_1 \oplus W_2$. □

Atividade 4.2. Reveja TODOS os exemplos da seção 4.3, tópico ”Subespaços Vetoriais”, do livro texto.

4.4 Combinação linear

Definição 4.3. *Seja V um espaço vetorial e vetores $v_1, \ldots, v_n \in V$. Dados $a_1, \ldots, a_n \in \mathbb{R}$, o vetor de V*

$$v = a_1 v_1 + \cdots + a_n v_n$$

*é uma **combinação linear** de $v_1, \ldots, v_n$.*

Fixados $v_1, \ldots, v_n \in V$, o conjunto W de todas as combinações lineares de $v_1, \ldots, v_n$,

$$W = [v_1, \ldots, v_n] = \{v \in V \mid\ v = a_1 v_1 + \cdots + a_n v_n,\ a_1, \ldots, a_n \in \mathbb{R}\}$$

é um subespaço vetorial de V, chamado *subespaço gerado* por $v_1, \ldots, v_n$. Dizemos também que $v_1, \ldots, v_n$ *geram* W.

Vamos mostrar que $W = [v_1, \ldots, v_n]$ é subespaço: sejam dados $u, v \in W$ e $\alpha \in \mathbb{R}$. Existem $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$ tais que $u = \sum_{i=1}^n a_i v_i$ e $v = \sum_{i=1}^n b_i v_i$. Assim

(i)

$$u + v = \sum_{i=1}^n a_i v_i + \sum_{i=1}^n b_i v_i = \sum_{i=1}^n (a_i + b_i) v_i \in W$$

e W é fechado para a soma;

(ii) $\alpha u = \alpha \sum_{i=1}^n a_i v_i = \sum_{i=1}^n (\alpha a_i) v_i \in W$, e W é fechado para a multiplicação por escalar.

Concluímos portanto que W é subespaço.

Exemplo 4.14. Considere um vetor $v \in \mathbb{R}^3$ com $v \neq \mathbf{0}$. Então o subespaço gerado por v,

$$[v] = \{av \in \mathbb{R}^3 \mid\ a \in \mathbb{R}\},$$

é a reta que passa pela origem $(0, 0, 0)$ na direção do vetor v. □

Atividade 4.3. Faça uma figura que represente esta geometria e veja que a soma de vetores sobre a reta está contida na reta.
Agora, imagine uma reta que não passe pela origem. Porque essa reta **não** constitui um subespaço vetorial? Quais propriedades de subespaço vetorial não são satisfeitas?

Exemplo 4.15. Considere dois vetores $u, v \in \mathbb{R}^3$ não colineares (isto é, $u \neq av$ para todo $a \in \mathbb{R}$ e $v \neq \mathbf{0}$). Então

$$[u, v] = \{au + bv \in \mathbb{R}^3 \mid a.b \in \mathbb{R}\}$$

é o plano que passa pela origem e é paralelo aos vetores u e v. □

Atividade 4.4. Faça uma figura que represente esta geometria e veja que a soma de vetores sobre o plano está contida no plano.

Exemplo 4.16. O subespaço de $\mathbb{R}^3$ gerado por $(1, 0, 1)$ e $(0, 1, -1)$ é

$$W = [(1, 0, 1); (0, 1, -1)] = \{a(1, 0, 1) + b(0, 1, -1) \mid a, b \in \mathbb{R}\} = \{(a, b, a - b) \mid a, b \in \mathbb{R}\}.$$

Geometricamente, W é o plano que passa pela origem e é paralelo aos vetores $(1, 0, 1)$ e $(0, 1, -1)$. **Faça uma figura que represente esta geometria**. □

Exemplo 4.17. O subespaço de $\mathbb{R}^2$ gerado por $(1, 0)$ e $(0, 1)$ é

$$[(1, 0); (0, 1)] = \{a(1, 0) + b(0, 1) \mid a, b \in \mathbb{R}\} = \{(a, b) \mid a, b \in \mathbb{R}\} = \mathbb{R}^2.$$

O subespaço de $\mathbb{R}^2$ gerado por $(1, 0)$, $(0, 1)$ e $(1, 1)$ é

$$\begin{aligned} [(1, 0); (0, 1); (1, 1)] &= \{a(1, 0) + b(0, 1) + c(1, 1) \mid a, b, c \in \mathbb{R}\} \\ &= \{(a + c, b + c) \mid a, b, c \in \mathbb{R}\} = \{(a', b') \mid a', b' \in \mathbb{R}\} = \mathbb{R}^2. \end{aligned}$$

Assim, $[(1, 0); (0, 1); (1, 1)] = [(1, 0); (0, 1)]$. □

Exemplo 4.18. Generalizando o exemplo anterior, vale

$$u \in [v_1, \ldots, v_n] \Leftrightarrow [v_1, \ldots, v_n, u] = [v_1, \ldots, v_n].$$

Ou seja, vetores que são combinações lineares de outros não fazem diferença no subespaço que eles geram. Isso será útil no estudo de base de um espaço vetorial.

Vamos mostrar tal afirmação. Se $u \in [v_1, \ldots, v_n]$ então $u = \sum_{i=1}^{n} a_i v_i$ para certos $a_1, \ldots, a_n \in \mathbb{R}$. Assim

$$\begin{aligned} [v_1, \ldots, v_n, u] &= [v_1, \ldots, v_n, \sum_{i=1}^{n} a_i v_i] \\ &= \left\{ v \in V \mid v = b_1 v_1 + \cdots + b_n v_n + b \left(\sum_{i=1}^{n} a_i v_i \right),\ b_1, \ldots, b_n, b \in \mathbb{R} \right\} \\ &= \{v \in V \mid v = (b_1 + b a_1) v_1 + \cdots + (b_n + b a_n) v_n,\ b_1, \ldots, b_n, b \in \mathbb{R}\} \\ &= \{v \in V \mid v = c_1 v_1 + \cdots + c_n v_n,\ c_1, \ldots, c_n \in \mathbb{R}\} \\ &= [v_1, \ldots, v_n]. \end{aligned}$$

Reciprocamente, $u = 0v_1 + \cdots + 0v_n + 1u \in [v_1, \ldots, v_n, u] = [v_1, \ldots, v_n]$, e logo $u \in [v_1, \ldots, v_n]$. □

Exemplo 4.19. Vamos encontrar um conjunto de vetores que geram o subespaço

$$W = \{(x, y, 2x, x + y) \mid x, y \in \mathbb{R}\}$$

de $\mathbb{R}^4$. Observe que

$$(x, y, 2x, x + y) = (x, 0, 2x, x) + (0, y, 0, y) = x(1, 0, 2, 1) + y(0, 1, 0, 1),$$

e logo $W = [(1, 0, 2, 1); (0, 1, 0, 1)]$. □

4.5 Dependência e independência linear

Definição 4.4. *Sejam V um espaço vetorial e $v_1, \ldots, v_n \in V$. Dizemos que o conjunto $\{v_1, \ldots, v_n\}$ é **linearmente independente (LI)** ou que os vetores $v_1, \ldots, v_n$ são LI se a equação*

$$a_1 v_1 + \cdots + a_n v_n = \mathbf{0}$$

*admitir somente a solução trivial $a_1 = a_2 = \cdots = a_n = 0$. Se $\{v_1, \ldots, v_n\}$ não for LI, então dizemos que este conjunto é **linearmente dependente (LD)**, ou que os vetores $v_1, \ldots, v_n$ são LD.*

Observamos que o subconjunto $\{\mathbf{0}\}$ de um espaço vetorial V qualquer não é LI, pois a equação da Definição 4.4 é satisfeita para todo $a_1 \in \mathbb{R}$.

Exemplo 4.20. Os vetores

$$(1, 0, 0), \quad (0, 1, 0), \quad (0, 0, 1)$$

de $\mathbb{R}^3$ são LI. De fato, da equação $a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) = (0, 0, 0)$ segue que $a_1 = a_2 = a_3 = 0$. □

Exemplo 4.21. O subconjunto de $M(2, 2)$

$$\left\{ \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ 0 & -1 \end{bmatrix} \right\}$$

é LI pois de

$$\begin{bmatrix} a_1 + a_2 & 2a_1 + 3a_2 \\ 0 & a_1 - a_2 \end{bmatrix} = a_1 \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 & 3 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

chegamos ao sistema

$$\begin{cases} a_1 & +a_2 & = & 0 \\ 2a_1 & +3a_2 & = & 0 \\ a_1 & -a_2 & = & 0 \end{cases},$$

que possui somente a solução trivial $a_1 = a_2 = 0$. □

Exemplo 4.22. O subconjunto $\{(1, 2); (2, 2); (3, 7)\}$ de $\mathbb{R}^2$ é LD pois

$$(3, 7) = 4(1, 2) - \frac{1}{2}(2, 2).$$

□

Exemplo 4.23. O subconjunto $\{p, q, r\}$ de P_2, onde

$$p(x) = x^2, \quad q(x) = x + 1, \quad r(x) = x^2 + x - 1$$

é LI. De fato, a equação $a_1p + a_2q + a_3r = \mathbf{0}$ diz que $a_1p(x) + a_2q(x) + a_3r(x) = 0$ para todo $x \in \mathbb{R}$. Assim,

$$a_1(x^2) + a_2(x + 1) + a_3(x^2 + x - 1) = (a_1 + a_3)x^2 + (a_2 + a_3)x + (a_2 - a_3) = 0$$

para todo $x \in \mathbb{R}$, donde segue o sistema

$$\left\{\begin{array}{ccccc} a_1 & & +a_3 & = & 0 \\ & a_2 & +a_3 & = & 0 \\ & a_2 & -a_3 & = & 0 \end{array}\right. .$$

Esse sistema possui somente a solução trivial $a_1 = a_2 = a_3 = 0$ (verifique!), e portanto $\{p, q, r\}$ é LI. $\square$

Você pode notar no Exemplo 4.22 que um dos vetores foi escrito como combinação linear dos outros. Quando os vetores são LD, isso sempre ocorre. De fato, quando $a_1v_1+a_2v_2+\cdots+a_nv_n = \mathbf{0}$ admite solução não trivial, digamos $a_1 \neq 0$, então podemos escrever

$$v_1 = -\frac{a_2}{a_1}v_2 - \cdots - \frac{a_n}{a_1}v_n.$$

A volta também vale: se um vetores é combinação dos outros, então os vetores são LD.

Teorema 4.3. *$\{v_1, \ldots, v_n\}$ é LD se, e somente se, um dos vetores desse conjunto é combinação linear dos outros.*

Uma forma equivalente de escrever o Teorema anterior é a seguinte:

"$\{v_1, \ldots, v_n\}$ é LI se, e somente se, nenhum dos vetores desse conjunto é combinação linear dos outros".

4.6 Base e dimensão de um espaço vetorial

Definição 4.5. *Dado um espaço vetorial V, um conjunto $\beta = \{v_1, \ldots, v_n\} \subset V$ é uma **base** de V se:*

(i) β é LI;

(ii) β gera V, isto é, $[\beta] = [v_1, \ldots, v_n] = V$.

Exemplo 4.24. O conjunto

$$\{e_1, e_2\} \subset \mathbb{R}^2, \quad \text{onde} \quad e_1 = (1, 0),\ e_2 = (0, 1)$$

é base de $\mathbb{R}^2$. De fato, temos

$$(a_1, a_2) = a_1(1, 0) + a_2(0, 1) = (0, 0) \ \Rightarrow\ a_1 = a_2 = 0$$

e qualquer (x, y) em $\mathbb{R}^2$ pode ser escrito como a combinação linear

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

Portanto $\{e_1, e_2\}$ é LI e gera $\mathbb{R}^2$, ou seja, é base. Esta é chamada a *base canônica de* $\mathbb{R}^2$. $\square$

Exemplo 4.25. $\{e_1, e_2, \ldots, e_n\} \subset \mathbb{R}^n$ onde

$$e_i = (0, \ldots, 0, \underbrace{1}_{\text{posição } i}, 0, \ldots, 0), \quad i = 1, \ldots, n,$$

é o vetor de $\mathbb{R}^n$ constituído de zeros, exceto na posição i igual a 1, é a *base canônica* de $\mathbb{R}^n$. □

Atividade 4.5. Adapte o argumento do Exemplo 4.24 para mostrar que $\{e_1, e_2, \ldots, e_n\}$ é base de $\mathbb{R}^n$.

Exemplo 4.26. $\{(1,1);(0,1)\}$ é base de $\mathbb{R}^2$ pois é LI e qualquer $(x,y) \in \mathbb{R}^2$ pode ser escrito como $(x,y) = x(1,1) + (y-x)(0,1)$. □

Exemplo 4.27. $\{(0,1);(0,2)\}$ não é base de $\mathbb{R}^2$ pois não é LI (um dos vetores é claramente combinação do outro). Também, este conjunto não gera $\mathbb{R}^2$. Por exemplo, o vetor $(1,0)$ não é combinação linear de $(0,1)$ e $(0,2)$ (tente escrever a combinação e veja que não dá certo). □

Exemplo 4.28. $\{(1,0,0);(0,1,0)\}$ não é base de $\mathbb{R}^3$ pois não gera $\mathbb{R}^3$ (por exemplo, $(0,0,1) \notin [(1,0,0);(0,1,0)]$). □

Exemplo 4.29. O conjunto

$$\beta = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

é *base canônica* de $M(2,2)$. De fato, β é LI pois

$$a_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow a_1 = a_2 = a_3 = a_4 = 0$$

e gera $M(2,2)$, pois qualquer matriz 2×2 pode ser escrita como

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_{21} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

□

Exemplo 4.30. Generalizando o exemplo anterior, a *base canônica* de $M(m,n)$ é o conjunto das matrizes $\mathbf{A}^{ij} = [a_{ij}]_{m \times n}$ tais que $a_{ij} = 1$ e $a_{kl} = 0$ sempre que $(k,l) \neq (i,j)$. □

Atividade 4.6. Mostre que o conjunto das matrizes $\mathbf{A}^{ij}$ do exemplo anterior é base de $M(m,n)$.

Exemplo 4.31. O conjunto de polinômios $\beta = \{1, x, x^2, \ldots, x^n\}$ é a *base canônica* de P_n. □

Atividade 4.7. Mostre que o conjunto β do exemplo anterior é base de P_n.

Observamos que, de acordo com o primeiro item da Definição 4.5, o espaço vetorial trivial $\{\mathbf{0}\}$ não admite base.

Vimos no Exemplo 4.17 que $\mathbb{R}^2$ pode ser gerado por dois vetores LI, e que adicionando um terceiro vetor, o subespaço gerado não muda. De outro modo, dos três vetores usados, podemos extrair dois que geram todo o $\mathbb{R}^2$ (uma base de $\mathbb{R}^2$). Em geral, sempre é possível extrair bases de conjunto de vetores que geram todo o espaço vetorial.

Teorema 4.4. *Sejam $v_1, \ldots, v_n$ vetores não nulos que geram um espaço vetorial V. Então, dentre esses vetores podemos extrair uma base de V.*

Exemplo 4.32. Afirmamos que

$$[(1,0,0);(0,1,-1);(1,1,0);(1,1,1)] = \mathbb{R}^3.$$

De fato, você pode verificar que $(x,y,z) = x(1,0,1) + y(0,1,-1) + (x-y-z)(1,1,0) + (-x+y+z)(1,1,1)$.

Agora, observe que

$$(1,1,0) = 1(1,0,1) + 1(0,1,-1) + 0(1,1,1),$$

isto é, $\alpha = \{(1,0,0);(0,1,-1);(1,1,0);(1,1,1)\}$ é LD. Mais ainda, isso mostra que

$$(1,1,0) \in [(1,0,1);(0,1,-1);(1,1,1)]$$

e logo $[(1,0,1);(0,1,-1);(1,1,1)] = \mathbb{R}^3$. Como $\beta = \{(1,0,1);(0,1,-1);(1,1,1)\}$ é LI (verifique!), segue que β é base de $\mathbb{R}^3$. O que fizemos portanto foi extrair do conjunto α de geradores de $\mathbb{R}^3$ uma base β de $\mathbb{R}^3$. □

Pelo que vimos acima, três vetores LI geram $\mathbb{R}^3$. O exemplo indica ainda que ao agregar um quarto vetor, este será LD com os anteriores: é como que a base de $\mathbb{R}^3$ não comporta mais que 3 vetores. O mesmo ocorre com $\mathbb{R}^2$: uma base tem no máximo 2 vetores (geometricamente, tente desenhar 3 vetores LI no plano e se convença que é impossível). Esse fato vale para qualquer espaço vetorial: se um conjunto de vetores já gera todo o espaço, qualquer vetor adicional será LD.

Teorema 4.5. *Seja V um espaço vetorial. Se $V = [v_1, \ldots, v_n]$ (n vetores) então qualquer subconjunto de V com mais de n vetores é LD.*

Exemplo 4.33. Sabemos que a base canônica de $\mathbb{R}^3$ tem três elementos, e que gera $\mathbb{R}^3$. Com isso já sabemos de antemão que o conjunto α do Exemplo 4.32 é LD, pois possui mais de três vetores de $\mathbb{R}^3$. □

Em outras palavras, o Teorema anterior diz o seguinte: se n vetores geram V, qualquer vetor adicional será LD com os anteriores. Então o número mínimo de geradores de todo o espaço é sempre o mesmo, independente de quais geradores escolhamos. Isso nos leva ao conceito de dimensão de um espaço vetorial.

Corolário 4.1. *Qualquer base de um espaço vetorial V tem sempre o mesmo número de elementos. Este número é chamado* **dimensão de** *V, e denotado por* $\dim V$.

Reforçando o que o resultado anterior diz: a **dimensão** de um espaço vetorial é a quantidade de vetores de qualquer base sua.

Exemplo 4.34. Como $\{(1,0);(0,1)\}$ é base de $\mathbb{R}^2$, temos $\dim \mathbb{R}^2 = 2$. Mais geralmente, considerando a base canônica de $\mathbb{R}^n$, segue que $\dim \mathbb{R}^n = n$. □

Exemplo 4.35. Do Exemplo 4.29, segue que $\dim M(2,2) = 4$. Mais geralmente, do exemplo 4.30 concluímos que $\dim M(m,n) = mn$. □

Exemplo 4.36. Do Exemplo 4.31 segue que $\dim P_n = n+1$, já que $\beta = \{1, x, x^2, \ldots, x^n\}$ é uma base de P_n. □

Atividade 4.8. Seja W um subespaço vetorial de V. Mostre que $\dim W = \dim V$ se, e somente se $W = V$.

O espaço vetorial trivial $\{\mathbf{0}\}$ não admite base, pois não há subconjunto seu que seja LI. Convencionaremos no entanto que $\dim\{\mathbf{0}\} = 0$.

Como dito no Exemplo 4.28, os vetores $(1, 0, 0)$ e $(0, 1, 0)$ são LI, não geram $\mathbb{R}^3$. Porém, agregando um terceiro vetor LI, temos uma base de $\mathbb{R}^3$. Por exemplo,

$$\{(1,0,0);(0,1,0);(0,0,1)\}, \qquad \{(1,0,0);(0,1,0);(0,0,4)\}, \qquad \{(1,0,0);(0,1,0);(1,1,1)\}$$

são bases de $\mathbb{R}^3$. O teorema abaixo diz que **sempre** é possível completar um conjunto LI a fim de obter bases de um espaço vetorial.

Teorema 4.6. *Qualquer conjunto de vetores LI de um espaço vetorial V pode ser completado de modo a obter-se uma base de V.*

Corolário 4.2. *Se $n = \dim V$, qualquer subconjunto de V com n vetores LI é uma base de V.*

Exemplo 4.37. Sabemos que $\dim \mathbb{R}^3 = 3$. Então o conjunto LI $\{(1, 0, -1); (0, 1, 2)\}$ não é base de $\mathbb{R}^3$. A demonstração do Teorema 4.6 nos dá uma maneira de completar esse conjunto a uma base de $\mathbb{R}^3$: basta escolher um vetor $(x, y, z) \notin [(1, 0, -1); (0, 1, 2)]$ e uni-lo ao conjunto. Observe que

$$[(1,0,-1);(0,1,2)] = \{a(1,0,-1) + b(0,1,2) \mid a, b \in \mathbb{R}\} = \{(a, b, 2b - a) + b(0,1,2) \mid a, b \in \mathbb{R}\},$$

e logo $(1, 1, 0) \notin [(1, 0, -1); (0, 1, 2)]$. Com isso $\beta = \{(1, 0, -1); (0, 1, 2); (1, 1, 0)\}$ é LI, e como possui $3 = \dim \mathbb{R}^3$ vetores, é uma base de $\mathbb{R}^3$. □

Exemplo 4.38. Vamos completar o conjunto LI $\{(1, 0, 1, 2); (2, 1, 0, 0)\}$ a uma base de $\mathbb{R}^4$. Primeiro, temos

$$W_1 = [(1,0,1,2);(2,1,0,0)] = \{(a + 2b, b, a, 2a) \mid a, b \in \mathbb{R}\}.$$

Escolhemos $(1, 1, 0, 0) \notin W_1$. Agora,

$$W_2 = [(1,0,1,2);(2,1,0,0);(1,1,0,0)] = \{(a + 2b + c, b + c, a, 2a) \mid a, b, c \in \mathbb{R}\}.$$

Escolhendo então $(1, 0, 2, 0) \notin W_2$, obtemos a base

$$\beta = \{(1,0,1,2);(2,1,0,0);(1,1,0,0);(1,0,2,0)\}$$

de $\mathbb{R}^4$. □

Teorema 4.7. *Se U e W são subespaços de V então $\dim U \leq \dim V$, $\dim W \leq \dim V$ e*

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

Exemplo 4.39. Considere os subespaços de $\mathbb{R}^3$

$$U = \{(x, y, z) \mid x = y\}, \quad W = \{(x, y, z) \mid x + y - z = 0\}.$$

Como

$$U = \{(x, x, z) \mid x, z \in \mathbb{R}\} \quad \text{e} \quad W = \{(x, y, x + y) \mid x, y \in \mathbb{R}\}$$

segue que $\beta_U = \{(1,1,0);(0,0,1)\}$ e $\beta_W = \{(1,0,1);(0,1,1)\}$ são bases de U e W, respectivamente. Assim, $\dim U = \dim W = 2$, e

$$3 = \dim \mathbb{R}^3 = \dim U + \dim W - \dim(U \cap W) \;\Rightarrow\; \dim(U \cap W) = 1.$$

Vamos verificar que de fato $\dim(U \cap W) = 1$. Temos

$$U \cap W = \{(x,y,z) \mid x+y-z=0,\ x=y\} = \{(x,x,2x) \mid x \in \mathbb{R}\} = [(1,1,2)],$$

ou seja, $\dim(U \cap W) = 1$. □

Pela definição de base, qualquer vetor v de um espaço vetorial V pode ser escrito em uma base β. O resultado a seguir diz que essa escrita só pode ser feita de uma única forma. Você poderá notar isso nos exemplos e exercícios!

Teorema 4.8. *Dada uma base $\beta = \{v_1, \ldots, v_n\}$ de V, cada vetor v de V se escreve de maneira **única** como combinação linear dos vetores de β.*

Já que a escrita em uma base é única, vamos chamar os coeficientes $a_1, \ldots, a_n$ da escrita

$$v = a_1 v_1 + \cdots + a_n v_n$$

de v na base $\beta = \{v_1, \ldots, v_n\}$ de *coordenadas* de v em relação à β. Se levarmos em consideração a ordem dos vetores em β (que neste caso referimos a β como *base ordenada*), escrevemos

$$[\,v\,]_\beta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Poderemos omitir o termo “ordenada” quando o contexto deixar claro que se trata de uma base ordenada.

Recorde que em Geometria Analítica, as coordenadas de um vetor $v = (x,y,z)$ são os números x, y e z. De fato, esses são as **coordenadas** da escrita de v na base canônica de $\mathbb{R}^3$:

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

Em Geometria Analítica, a base canônica é largamente utilizada (ela dá as direções dos eixos coordenados).

Exemplo 4.40. Em $\mathbb{R}^2$, consideremos a base (ordenada) $\beta = \{(1,1);(-1,2)\}$. Como

$$(-1,8) = 2(1,1) + 3(-1,2)$$

segue que

$$[\,(-1,8)\,]_\beta = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Se considerarmos a base $\beta' = \{(-1,2);(1,1)\}$ proveniente da mudança de ordem em β, vemos que

$$[\,(-1,8)\,]_{\beta'} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

□

Daqui para frente, vamos denotar a base canônica de um espaço por "*can*". Assim,

- em $\mathbb{R}^2$, $can = \{(1,0);(0,1)\}$
- em $\mathbb{R}^n$, $can = \{e_1, \ldots, e_n\}$
- em $M(2,2)$, $can = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$
- em $M(m,n)$, $can = \{A^{ij} \mid A^{ij}_{ij} = 1,\ A^{ij}_{kl} = 0, \forall (k,l) \neq (i,j),\ \text{para cada } i,j\}$
- em P_n, $can = \{1, x, x^2, \ldots, x^n\}$

Exemplo 4.41. Em $M(2,2)$, considere a base canônica *can*. Como

$$v = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} = 1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + (-1) \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

segue que

$$[\, v\,]_{can} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \end{bmatrix}.$$

□

Exemplo 4.42. Em $\mathbb{R}^n$, considere a base canônica $can = \{e_1, \ldots, e_n\}$. Qualquer vetor $v = (x_1, \ldots, v_n) \in \mathbb{R}^n$ pode ser facilmente escrito na base *can*. De fato, temos

$$v = x_1 e_1 + \cdots + x_n e_n$$

e logo

$$[\, v\,]_{can} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

□

Exemplo 4.43. Um vetor de um espaço vetorial pode ser facilmente descrito em certas bases. Por exemplo, o exemplo anterior mostra que é fácil descrever um vetor de $\mathbb{R}^n$ em sua base canônica. Outro exemplo é o seguinte: considere a base canônica $can = \{1, x, x^2, \ldots, x^n\}$ de P_n. Um vetor $p(x) = a_n x^n + \cdots + a_1 x + a_0$ qualquer de P_n tem escrita na base canônica

$$[\, p\,]_{can} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

□

Exemplo 4.44. Qualquer matriz $\mathbf{A} = [a_{ij}]_{m \times n}$ do espaço vetorial $M(m,n)$ tem escrita na base canônica de $M(m,n)$

$$[\, \mathbf{A}\,]_{can} = \begin{bmatrix} a_{11} & \cdots & a_{1n} & a_{21} & \cdots & a_{2n} & \cdots & a_{m1} & \cdots & a_{mn} \end{bmatrix}^t.$$

□

4.7 Mudança de base

Seja V um espaço vetorial e $\beta = \{u_1, \ldots, u_n\}$, $\beta' = \{w_1, \ldots, w_n\}$ bases ordenadas suas. Nosso objetivo é descobrir qual a relação entre as escritas de um vetor $v \in V$ nessas bases, isto é, qual a relação entre $[\,v\,]_\beta$ e $[\,v\,]_{\beta'}$. Fixado $v \in V$, escrevamos

$$v = \sum_1^n x_i u_i \quad \text{e} \quad v = \sum_1^n y_i w_i,$$

isto é,

$$[\,v\,]_\beta = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad [\,v\,]_{\beta'} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Como β é base de V, podemos escrever cada vetor de β' como cominação linear dos vetores de β, digamos

$$\begin{cases} w_1 &= a_{11}u_1 + a_{21}u_2 + \cdots a_{n1}u_n \\ w_2 &= a_{12}u_1 + a_{22}u_2 + \cdots a_{n2}u_n \\ \quad\vdots & \\ w_n &= a_{1n}u_1 + a_{2n}u_2 + \cdots a_{nn}u_n \end{cases}.$$

Assim,

$$\begin{aligned} v = \sum_1^n y_i w_i &= \sum_1^n y_i(a_{1i}u_1 + a_{2i}u_2 + \cdots a_{ni}u_n) \\ &= (a_{11}y_1 + a_{12}y_2 + \cdots a_{1n}y_n)u_1 + \cdots + (a_{n1}y_1 + a_{n2}y_2 + \cdots a_{nn}y_n)u_n \\ &= \sum_1^n (a_{i1}y_1 + a_{i2}y_2 + \cdots a_{in}y_n)u_i \end{aligned}$$

Mas a escrita de v na base $\beta = \{u_1, \ldots, u_n\}$ é única, de modo que os coeficientes dos u_i's que aparecem na soma anterior são os x_i's:

$$\begin{cases} x_1 &= a_{11}y_1 + a_{12}y_2 + \cdots a_{1n}y_n \\ x_2 &= a_{21}y_1 + a_{22}y_2 + \cdots a_{2n}y_n \\ \quad\vdots & \\ x_n &= a_{n1}y_1 + a_{n2}y_2 + \cdots a_{nn}y_n \end{cases}$$

ou ainda,

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Fazendo

$$[\,I\,]_\beta^{\beta'} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix},$$

chegamos à expressão

$$[\,v\,]_\beta = [\,I\,]_\beta^{\beta'}\,[\,v\,]_{\beta'}.$$

A matriz $[\,I\,]_\beta^{\beta'}$ é chamada *matriz de mudança da base β' para a base β*.

A notação é adequada: você pode ver $[\,I\,]_\beta^{\beta'}\,[\,v\,]_{\beta'}$ como o "produto $v = Iv$" onde as bases β' são "canceladas".

Observe que as colunas de $[\,I\,]_\beta^{\beta'}$ são os coeficientes da escrita dos vetores da base inicial $\beta' = \{w_1, \ldots, w_n\}$ na base final $\beta = \{u_1, \ldots, u_n\}$. Para indicar tal fato, podemos escrever

$$[\,I\,]_\beta^{\beta'} = \begin{bmatrix} | & & | \\ [\,w_1\,]_\beta & \cdots & [\,w_n\,]_\beta \\ | & & | \end{bmatrix}.$$

Dessa maneira, fica fácil saber o que fazer para calcular uma matriz de mudança de base.

Exemplo 4.45. Sejam $\beta = \{(2,-1);(3,4)\}$ e $can = \{(1,0);(0,1)\}$ bases de $\mathbb{R}^2$. Vamos calcular a matriz de mudança da base can para β:

$$[\,I\,]_\beta^{can} = \begin{bmatrix} | & | \\ [\,(1,0)\,]_\beta & [\,(0,1)\,]_\beta \\ | & | \end{bmatrix}.$$

Devemos então calcular a escrita dos vetores de can na base β:

- $(1,0) = a(2,-1) + b(3,4) \Rightarrow \left\{ \begin{array}{rrcl} 2a & +3b & = & 1 \\ -a & +4b & = & 0 \end{array} \right. \Rightarrow a = 4/11, b = 1/11.$

  Logo, $[\,(1,0)\,]_\beta = \begin{bmatrix} 4/11 \\ 1/11 \end{bmatrix}$.

- Fazendo as contas, conclui-se que $[\,(0,1)\,]_\beta = \begin{bmatrix} -3/11 \\ 2/11 \end{bmatrix}$.

Portanto

$$[\,I\,]_\beta^{can} = \begin{bmatrix} 4/11 & -3/11 \\ 1/11 & 2/11 \end{bmatrix}.$$

Com isso podemos encontrar qualquer vetor na base β. Por exemplo, tomando $v = (5,-8)$, vemos que $[\,v\,]_{can} = \begin{bmatrix} 5 \\ -8 \end{bmatrix}$. Assim,

$$[\,v\,]_\beta = [\,I\,]_\beta^{can}\,[\,v\,]_{can} = \begin{bmatrix} 4/11 & -3/11 \\ 1/11 & 2/11 \end{bmatrix} \begin{bmatrix} 5 \\ -8 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}.$$

Você pode verificar que realmente

$$(5,-8) = 4(2,-1) - 1(3,4).$$

□

Exemplo 4.46. Seja $\beta = \{v_1, \ldots, v_n\}$ base de um espaço vetorial V. Então

$$[\,I\,]_\beta^{\beta} = \begin{bmatrix} | & | & & | \\ [\,v_1\,]_\beta & [\,v_2\,]_\beta & \cdots & [\,v_n\,]_\beta \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{I}_n.$$

De fato, a escrita de cada vetor v_i na base β é

$$v_i = 0v_1 + 0v_2 + \cdots + 1v_i + \cdots + 0v_n,$$

ou seja, $[\,v_i\,]_\beta$ é a coluna i da matriz identidade $\mathbf{I}_n$. $\square$

Uma pergunta surge: se conhecemos a mudança de β' para β ($[\,I\,]_\beta^{\beta'}$), como obter a mudança inversa, de β para β' ($[\,I\,]_{\beta'}^{\beta}$)?

Vejamos: dado $v \in V$ qualquer, temos

$$[\,v\,]_\beta = [\,I\,]_\beta^{\beta'}\,[\,v\,]_{\beta'} \quad \text{e} \quad [\,v\,]_{\beta'} = [\,I\,]_{\beta'}^{\beta}\,[\,v\,]_\beta.$$

Assim,

$$[\,v\,]_{\beta'} = [\,I\,]_{\beta'}^{\beta}\,[\,I\,]_\beta^{\beta'}\,[\,v\,]_{\beta'}$$

(veja como as bases se "cancelam" no produto de matrizes). Agora, se $\beta' = \{w_1, \ldots, w_n\}$ então

$$[\,w_1\,]_{\beta'} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad [\,w_2\,]_{\beta'} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad [\,w_n\,]_{\beta'} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

ou seja, $[\,w_i\,]_{\beta'}$ é a coluna i da identidade I_n. Então a expressão

$$[\,w_i\,]_{\beta'} = \left([\,I\,]_{\beta'}^{\beta}\,[\,I\,]_\beta^{\beta'}\right)[\,w_i\,]_{\beta'}$$

diz que a coluna i de $[\,I\,]_{\beta'}^{\beta}\,[\,I\,]_\beta^{\beta'}$ é a coluna i de $\mathbf{I}_n$, para todo i. Portanto,

$$[\,I\,]_{\beta'}^{\beta}\,[\,I\,]_\beta^{\beta'} = \mathbf{I}_n,$$

isto é, $[\,I\,]_\beta^{\beta'}$ é inversível e

$$[\,I\,]_{\beta'}^{\beta} = \left([\,I\,]_\beta^{\beta'}\right)^{-1}.$$

Exemplo 4.47. Considere as bases *can* canônica e $\beta = \{(1,0,1);(0,0,2);(2,2,0)\}$ de $\mathbb{R}^3$. Vamos encontrar $[\,I\,]_\beta^{can}$. Observe que nessa matriz as colunas são as escritas dos vetores canônicos na base β. É conveniente portanto calculá-la invertendo a matriz

$$[\,I\,]_{can}^{\beta} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}:$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[L_2 \leftrightarrow L_3]{\substack{L_1 \to L_1 - 2L_2 \\ L_2 \to 1/2 L_2}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1/2 & 0 \end{array}\right]$$

$$\xrightarrow{\substack{L_2 \to L_2 - L_1 \\ L_2 \to 1/2 L_2}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1/2 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 & 1/2 & 0 \end{array}\right]$$

e logo

$$[\,I\,]_\beta^{can} = \left([\,I\,]_{can}^{\beta}\right)^{-1} = \frac{1}{2}\begin{bmatrix} 2 & -2 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

$\square$

4.8 Demonstrações

Demonstração do Teorema 4.1. $u, w \in W_1 \cap W_2 \Rightarrow u, w \in W_1$ e $u, w \in W_2$. Como W_1 e W_2 são subespaços, temos $u + w \in W_1$ e $u + w \in W_2$. Assim $u + w \in W_1 \cap W_2$, ou seja, $W_1 \cap W_2$ é fechado para a soma. □

Atividade 4.9. Mostre que a interseção $W_1 \cap W_2$ no teorema anterior é fechada para a multiplicação por escalar, e conclua a prova desse teorema.

Demonstração do Teorema 4.2. $u, v \in W_1 + W_2 \Rightarrow u = w_1 + w_2$, $v = \overline{w}_1 + \overline{w}_2$, $w_1, \overline{w}_1 \in W_1, w_2, \overline{w}_2 \in W_2 \Rightarrow u + v = (w_1 + w_2) + (\overline{w}_1 + \overline{w}_2) = (w_1 + \overline{w}_1) + (w_2 + \overline{w}_2) \in W_1 + W_2$ pois W_1 e W_2 são subespaços. Isso mostra que $W_1 + W_2$ é fechado para a soma. □

Atividade 4.10. Mostre que $W_1 + W_2$ no teorema anterior é fechado para a multiplicação por escalar, e conclua a prova desse teorema.

Demonstração do Teorema 4.3. Se $\{v_1, \ldots, v_n\}$ é LD então existem $a_1, \ldots, a_n \in \mathbb{R}$ não todos nulos (digamos que $a_i \neq 0$) tais que $a_1 v_1 + \cdots + a_n v_n = \mathbf{0}$. Assim,

$$\begin{aligned} v_i &= -\frac{1}{a_i}(a_1 v_1 + \cdots + a_{i-1} v_{i-1} + a_{i+1} v_{i+1} + \cdots + a_n v_n) \\ &= \left(-\frac{a_1}{a_i}\right) v_1 + \cdots + \left(-\frac{a_{i-1}}{a_i}\right) v_{i-1} + \left(-\frac{a_{i+1}}{a_i}\right) v_{i+1} + \cdots + \left(-\frac{a_n}{a_i}\right) v_n, \end{aligned}$$

ou seja, v_i é combinação linear de $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$.

Reciprocamente, se v_i é combinação dos outros vetores, então

$$\begin{aligned} v_i &= a_1 v_1 + \cdots + a_{i-1} v_{i-1} + a_{i+1} v_{i+1} + \cdots + a_n v_n \\ &\Rightarrow a_1 v_1 + \cdots + a_{i-1} v_{i-1} + (-1) v_i + a_{i+1} v_{i+1} + \cdots + a_n v_n = \mathbf{0}. \end{aligned}$$

Com isso, a equação $a_1 v_1 + \cdots + a_n v_n = \mathbf{0}$ admite uma solução não trivial. Portanto $\{v_1, \ldots, v_n\}$ é LD. □

Demonstração do Corolário 4.1. Sejam $\alpha = \{v_1, \ldots, v_n\}$ e $\beta = \{w_1, \ldots, w_k\}$ bases de V. Como $V = [\alpha]$ e β é LI, o Teorema 4.5 diz que $k \leq n$. Por outro lado, como $V = [\beta]$ e α é LI, também $k \geq n$. Assim, $k = n$, como queríamos. □

Demonstração do Teorema 4.6. Seja $n = \dim V$ e $v_1, \ldots, v_r \in V$ vetores LI. Se $[v_1, \ldots, v_r] = V$ então $\{v_1, \ldots, v_r\}$ já é base de V $(r = n)$. Caso contrário, se $[v_1, \ldots, v_r] \subsetneq V$, existe $v_{r+1} \in V$ tal que $v_{r+1} \notin [v_1, \ldots, v_r]$ $(r < n)$. Neste caso, $\{v_1, \ldots, v_r, v_{r+1}\}$ é LI. Com isso, se $[v_1, \ldots, v_r, v_{r+1}] = V$, então $\{v_1, \ldots, v_r, v_{r+1}\}$ é base de V. Caso contrário, existe $v_{r+2} \in V$ tal que

$$[v_1, \ldots, v_r, v_{r+1}] \subsetneq [v_1, \ldots, v_r, v_{r+1}, v_{r+2}].$$

Prosseguindo se necessário, existem então $k = n - r$ vetores $v_{r+1}, \ldots, v_{r+k} \in V$ tais que $\beta = \{v_1, \ldots, v_r, v_{r+1}, v_{r+k}\}$ é LI e $[\beta] = V$, ou seja, β é base de V. Note que este processo pára pois qualquer conjunto com n vetores de V é LD. □

Demonstração do Corolário 4.2. Suponha por absurdo que o conjunto β com n vetores LI não é base de V. Assim, completamos β a uma base de V, obtendo uma base de V com mais de n vetores, um absurdo. □

Demonstração do Teorema 4.8. Seja $v \in V$. Como $[v_1, \dots, v_n] = V$, v é combinação linear de $v_1, \dots, v_n]$. Escrevamos $v = \sum_1^n a_i v_i$ e $v = \sum_1^n b_i v_i$. Devemos mostrar que $a_i = b_i$ para todo i. Ora, como

$$\mathbf{0} = v - v = \sum_1^n a_i v_i - \sum_1^n b_i v_i = \sum_1^n (a_i - b_i) v_i$$

e β é LI, segue que $a_i - b_i = 0$ para todo i, como queríamos demonstrar. □

4.9 Exercícios

1. Verifique que os conjuntos abaixo são subespaços de $\mathbb{R}^2$ ou $\mathbb{R}^3$.

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

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

   (c) $D = \{(x, y, z) \in \mathbb{R}^3 \mid x = y,\ 2y = z\}$

2. Diga quais dos subconjuntos a seguir são subespaços vetoriais. Se não for, diga o porquê. Considere como operações de soma e multiplicação por escalar as usuais de $\mathbb{R}^n$ ou $M(m, n)$.

   (a) $X = \{(x, y, z) \in \mathbb{R}^3 \mid xy = 0\} \subset \mathbb{R}^3$.

   (b) O conjunto $Z \subset M(2, 3)$ nas quais alguma coluna é formada por elementos iguais.

   (c) O conjunto $L \subset \mathbb{R}^n$ dos vetores $v = (x, 2x, 3x, \dots, nx)$, onde $x \in \mathbb{R}$ é arbitrário.

   (d) O conjunto dos vetores de $\mathbb{R}^5$ que têm duas ou mais coordenadas nulas.

   (e) O conjunto dos vetores de $\mathbb{R}^3$ que têm pelo menos uma coordenada ≥ 0.

   (f) O conjunto dos vetores de $\mathbb{R}^n$ cujas $k \geq 1$ primeiras coordenadas são iguais.

   (g) $Y = \{(x, y) \in \mathbb{R}^2 \mid x^2 + 3x = y^2 + 3y\} \subset \mathbb{R}^2$.

3. Exiba uma base para cada um dos subespaços de $\mathbb{R}^4$.

   (a) $F = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 = x_2 = x_3 = x_4\}$

   (b) $H = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 = x_2 = x_3\}$

   (c) $K = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + x_2 + x_3 + x_4 = 0\}$

4. Verifique que os vetores $u = (1, 1, 1)$, $v = (1, 2, 1)$ e $w = (2, 1, 2)$ são linearmente dependentes.

5. Verifique que os vetores $u = (1, 1, 1)$, $v = (1, 2, 3)$ e $w = (1, 4, 9)$ formam uma base de $\mathbb{R}^3$. Exprima cada um dos vetores $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ e $e_3 = (0, 0, 1)$ como combinação linear dos elementos dessa base.

6. Dados $u = (1, 2)$ e $w = (-1, 2)$, sejam F e G as retas que passam pela origem em $\mathbb{R}^2$ nas direções de u e w, respectivamente.

   (a) Verifique que F e G são subespaços de $\mathbb{R}^2$.

(b) Mostre que $\mathbb{R}^2 = F \oplus G$.

7. Considere os subespaços $W_1, W_2 \in \mathbb{R}^3$ assim definidos: W_1 é o conjunto de todos os vetores $v = (x, x, x)$ e W_2 é o conjunto de todos os vetores $w = (x, y, 0)$. Mostre que $\mathbb{R}^3 = W_1 \oplus W_2$.

8. (a) Mostre que para todo subespaço vetorial W de $\mathbb{R}^n$, existe um subespaço vetorial W_0 de $\mathbb{R}^n$ tal que $\mathbb{R}^n = W \oplus W_0$.

   (b) Dado o subespaço $W = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + z = 0\}$ de $\mathbb{R}^3$, encontre um subespaço W_0 tal que $\mathbb{R}^3 = W \oplus W_0$.

9. Diga se cada afirmação é verdadeira ou falsa, justificando sua resposta.

   (a) Se $V = W_1 \oplus W_2$ então $\dim V = \dim W_1 + \dim W_2$.

   (b) Se $\dim V = \dim W_1 + \dim W_2$, então $V = W_1 \oplus W_2$.

   (c) Se W_1 e W_1 são subespaços de V, $\dim W_1 > \dfrac{\dim V}{2}$ e $\dim W_2 > \dfrac{\dim V}{2}$, então $W_1 \cap W_2 \neq \{\mathbf{0}\}$.

   (d) Se W_1 e W_2 são subespaços de V então $[W_1 \cup W_2] = W_1 + W_2$.

10. Mostre que todo espaço vetorial V com $\dim V = n$ é soma direta de n subespaços de dimensão 1.

11. Mostre que $\beta = \{(1, 1, 0, 0), (-1, 0, 2, 3), (0, 0, 1, 1), (1, 0, 0, 3)\}$ forma uma base de $\mathbb{R}^4$ e obtenha as matrizes mudança de base $[I]^{\beta}_{can}$ e $[I]^{can}_{\beta}$, onde *can* é a base canônica de $\mathbb{R}^4$.

12. Sejam α e β bases de $\mathbb{R}^2$. Se $\alpha = \{(1, 2), (2, 3)\}$ e $[I]^{\beta}_{\alpha} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$, encontre β.

13. Considere um subconjunto LD $\alpha = \{v_1, v_2, \cdots, v_n\}$ de $\mathbb{R}^n$ com n vetores. Justifique porque o determinante da matriz cujas colunas são os vetores $v_1, v_2, \cdots, v_n$ do conjunto α é nulo.

14. Mostre que, dado um subconjunto de $\mathbb{R}^3$ $\alpha = \{(a_{11}, a_{21}, a_{31}), (a_{12}, a_{22}, a_{32}), (a_{13}, a_{23}, a_{33})\}$, temos que α é base de $\mathbb{R}^3$ se, e somente se, $\det[a_{ij}]_{3\times 3} \neq 0$. Tente generalizar o resultado para $\mathbb{R}^n$. Volte ao exercício 11 e veja que o determinante da matriz 4×4 cujas colunas são os vetores de β é não nulo, e portanto β é base.
    Dica: no caso de $\mathbb{R}^3$, escreva a matriz $[a_{ij}]_{3\times 3}$ e observe que ela é justamente a matriz de mudança de base $[I]^{\alpha}_{can}$. Sabendo que essa matriz é inversível, conclua o resultado usando o exercício 13 *e lembrando do teorema que diz "$\mathbf{A}$ é inversível se, e somente se, $\det \mathbf{A} \neq 0$".*

15. Considere o plano $\pi : 2x - y + 3z = 0$ que passa pela origem de $\mathbb{R}^3$ e a reta $r : (x, y, z) = (0, 0, 0) + t(1, 1, -3)$. Seja P e R os subespaços de $\mathbb{R}^3$ correspondentes ao plano e à reta, respectivamente.

    (a) Encontre bases para P e R.

    (b) Mostre que $\mathbb{R}^3 = P \oplus R$.

    (c) Dê uma interpretação geométrica para as bases de P e R, considerando o seu conhecimento em geometria analítica.

16. Considere a afirmação

"A união de dois subconjuntos LI X e Y do espaço vetorial V é ainda um conjunto LI"

Diga se cada alternativa abaixo é correta ou incorreta, justificando suas respostas.

(a) A afirmação é sempre verdadeira.

(b) A afirmação nunca é verdadeira.

(c) A afirmação é verdadeira quando X e Y são disjuntos, isto é, quando $X \cap Y = \emptyset$.

(d) A afirmação é verdadeira quando $X \subset Y$ ou $Y \subset X$.

(e) A afirmação é verdadeira quando um dos conjuntos X ou Y é disjunto do subespaço gerado pelo outro.

(f) A afirmação é verdadeira quando o número de elementos de X somado ao número de elementos de Y é igual à dimensão de V.

17. Se os coeficientes $a_1, a_2, \ldots, a_n$ não são todos iguais a zero (isto é, pelo menos um deles é não nulo), mostre que o *hiperplano*

$$H = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\}$$

é um subespaço vetorial de dimensão $n-1$ em $\mathbb{R}^n$ (o caso $n = 3$ remete à geometria vista em GA).

EXERCÍCIOS EXTRAS

18. Seja $X = \{x_1, x_2, \ldots, x_n\}$ um conjunto de n elementos, e considere o conjunto V de todas as funções $f : X \to \mathbb{R}$. Definimos, sobre os elementos de V, as operações de soma e multiplicação por um escalar $\alpha \in \mathbb{R}$ fazendo $(f+g)(x) = f(x) + g(x)$ e $(\alpha f)(x) = \alpha f(x)$, respectivamente.

(a) Mostre que V com as operações acima é um espaço vetorial.

(b) Para cada $i = 1, \ldots n$, definimos as funções $f_i : X \to \mathbb{R}$ pondo $f_i(x_i) = 1$ e $f_i(x_j) = 0$, para todo $j \neq i$. Mostre que $\{f_1, f_2, \ldots, f_n\}$ é uma base para V. Conclua então que $\dim V = n$. Você vê alguma semelhança entre V e o espaço vetorial $\mathbb{R}^n$?

(c) Escreva a função $f : X \to \mathbb{R}$ definida por $f(x_i) = i$ como uma combinação linear das funções $f_1, \ldots f_n$.

19. Considere P_n o espaço vetorial dos polinômios de grau $\leq n$, mais o polinômio nulo, com as operações usuais $(p+q)(x) = p(x) + q(x)$ e $(\alpha p)(x) = \alpha p(x)$.

(a) Exiba uma base para P_n, e encontre $\dim P_n$. Você vê alguma semelhança entre P_n o espaço vetorial $\mathbb{R}^{n+1}$?

(b) Escreva o polinômio $p(x) = 2x^6 - 4x^5 + x^3 - 7x + 2$ na base que você encontrou.

20. Seja $M(n,n)$ o espaço vetorial das matrizes reais $n \times n$ com as operações usuais.

(a) Mostre que

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

é uma base de $M(2,2)$.

(b) Generalize o item anterior e exiba uma base de $M(n,n)$. Com isto, encontre $\dim M(n,n)$.

(c) O conjunto $T_{\text{sim}}(n)$ das matrizes simétricas de ordem n munido das operações usuais é um subespaço de $M(n,n)$. Exiba uma base de $T_{\text{sim}}(n)$ e encontre $\dim T_{\text{sim}}(n)$.

(d) Sejam $T_S(n)$ e $T_I(n)$ os subespaços das matrizes triangulares superiores e inferiores, respectivamente. Mostre que $M(n,n) = T_S(n) + T_I(n)$, e que NÃO se tem $M(n,n) = T_S(n) \oplus T_I(n)$. Qual a dimensão dos subespaços $T_S(n)$ e $T_I(n)$?

21. Mostre que os elementos são linearmente independentes.

(a) $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ e $\mathbf{C} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

(b) $p(x) = x^3 - 5x^2 + 1$, $q(x) = 2x^4 + 5x - 6$ e $r(x) = x^2 - 5x + 2$

22. Seja V um espaço vetorial e $u, v, w \in V$. Mostre que $[u,v] \subset [u,v,w]$. Quando se tem $[u,v] = [u,v,w]$? Quando se tem $[u,v] \subsetneq [u,v,w]$?

Capítulo 5

Transformações Lineares

5.1 Motivação

Transformações lineares são funções entre espaços vetoriais que satisfazem certas propriedades. Essas propriedades, que constituem o que coloquialmente chamamos de *linearidade*, são um dos conceitos mais fundamentais da Álgebra Linear (não à toa dá o nome à disciplina!).

Lidar com funções lineares no computador é particularmente interessante, pois, como veremos, a avaliação de tais funções pode ser vista como um produto "matriz-vetor" $\mathbf{Ax}$. De fato, linguagens de programação para computação de alta eficiência são pensadas em termos matriciais, como por exemplo, o Fortran. Python, Matlab, Octave e Julia são outros exemplos de linguagens que dedicam especial atenção à operações com matrizes.

A facilidade computacional com funções lineares permite-nos resolver vários problemas reais. Por exemplo, ao querer encontrar o *valor mínimo* que uma função f qualquer atinge, pode-se conceber um algoritmo iterativo que a cada passo *aproxima* f por uma função linear; a sequência de problemas aproximados são mais fáceis de resolver justamente por envolverem funções lineares. Em particular, problemas de otimização onde os dados são descritos por funções lineares são amplamente estudados, com aplicações em diversos ocasiões (cálculo de melhores rotas, problemas de logística em geral, minimização de perdas em cortes de chapas metálicas etc – provavelmente você estudará problemas deste tipo na disciplina "Pesquisa Operacional").

Enfim, funções lineares aparecem em inúmeras situações. Veja a seção "Curiosidades - aplicações da Álgebra Linear" no sítio da disciplina.

Como ponto de partida, sugiro que você

- veja a introdução do Capítulo 5 do livro do Boldrini;
- assista aos vídeos indicados no sítio da disciplina.

5.2 Transformações Lineares – definição

Iniciamos com o estudo geral de transformações lineares. Logo após estudaremos transformações do plano no plano, que lembrará conteúdos vistos em Geometria Analítica.

Definição 5.1. *Sejam V e W espaços vetoriais. Uma* ***transformação linear*** *é uma aplicação (função)*

$$T : V \to W$$

que, para quaisquer $u, w \in V$ e $a \in \mathbb{R}$, satisfaz

(*i*) $T(u + w) = T(u) + T(w)$

(*ii*) $T(au) = aT(u)$

Nas condições (*i*) e (*ii*) aparecem as duas operações de espaços vetoriais.
(*i*) diz que a imagem da soma é a soma das imagens;
(*ii*) diz que a imagem da multiplicação por escalar é multiplicação da imagem pelo escalar.
Note que as operações à esquerda das igualdades (antes da aplicação de T) são as operações do espaço vetorial de partida V, enquanto que as operações entre as imagens são aquelas de W, o espaço de chegada.

Se $T : V \to W$ é uma transformação linear, a segunda condição da Definição 5.1 garante que

$$T(\mathbf{0}_V) = \mathbf{0}_W$$

(basta tomar $a = 0$).

Neste texto podemos escrever $\mathbf{0}_V$ para indicar o vetor nulo do espaço vetorial V quando houver risco de confusão.

Também, veja que

$$T(-u) = -T(u)$$

(tome $a = -1$) e

$$T(u - w) = T(u) - T(w)$$

pois $T(u - w) = T(u + (-w)) = T(u) + T(-w) = T(u) - T(w)$.

5.3 Exemplos

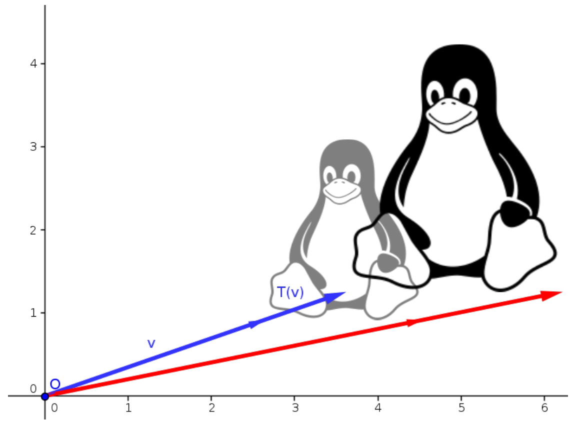
5.3.1 Transformações de $\mathbb{R}^2$ em $\mathbb{R}^2$ (do "plano no plano")

Vamos ver exemplos de transformações quando $V = W = \mathbb{R}^2$. Esses exemplos são instrutivos pois ajudam a intuição geométrica. Vamos visualizar também cada transformação na forma matricial, motivando para o estudo geral à frente.

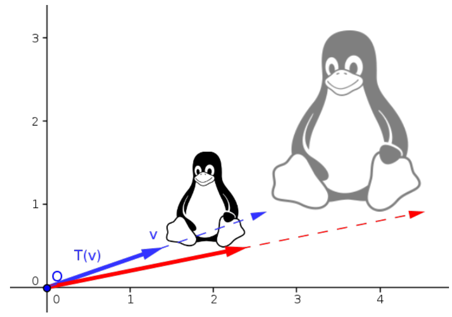
Exemplo 5.1. – Expansão (ou contração) uniforme.

$$\begin{aligned} T : \mathbb{R}^2 &\to \mathbb{R}^2 \\ v &\to av \end{aligned}$$

onde $a \in \mathbb{R}$ é fixado. Quando $0 < a < 1$, T é chamada *contração*, e quando $a > 1$, *expansão*.



$a > 1$



$0 < a < 1$

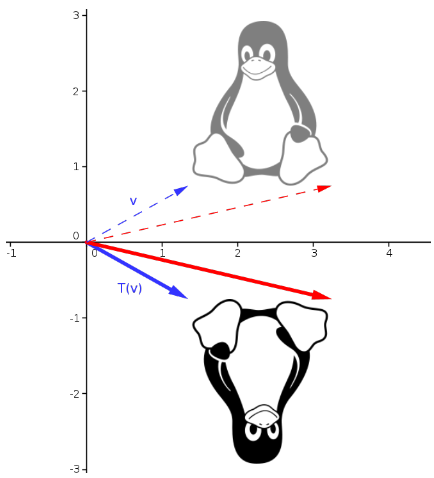
Em termos de matrizes, T pode ser representada por

$$\begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

□

Exemplo 5.2. – Reflexão em torno do eixo x.

$$\begin{aligned} T : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\to (x, -y). \end{aligned}$$



$$\begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

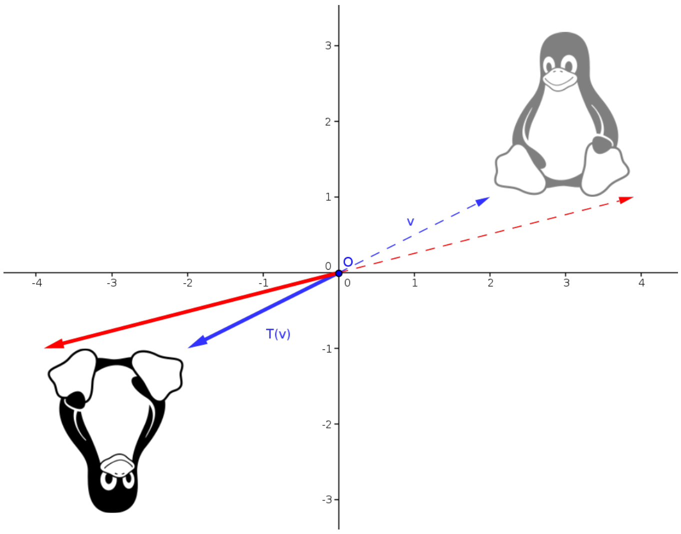
□

Atividade 5.1. Escreva a expressão da reflexão em torno do eixo y e sua forma matricial.

Exemplo 5.3. – Reflexão ao redor da origem.

$$\begin{aligned} T : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\to (-x, -y). \end{aligned}$$

Essa transformação corresponde ao Exemplo 5.1 com $a = -1$.



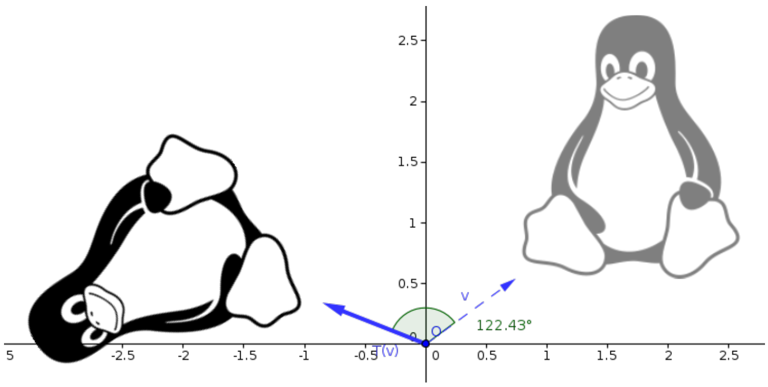
$$\begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

$\square$

Exemplo 5.4. – Rotação ao redor da origem (no sentido anti-horário).

$$\begin{aligned} T : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\to (\ x\cos\theta - y\,\mathrm{sen}\,\theta\ ,\ y\cos\theta + x\,\mathrm{sen}\,\theta\ ) \end{aligned}$$

onde $\theta \in \mathbb{R}$ é o ângulo de rotação em radianos.

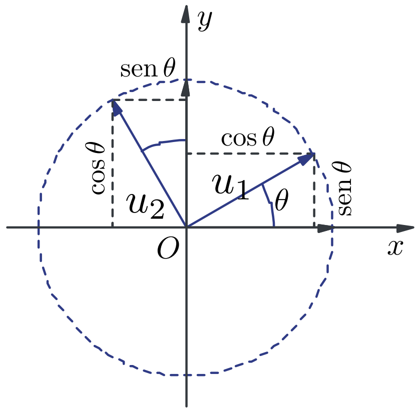


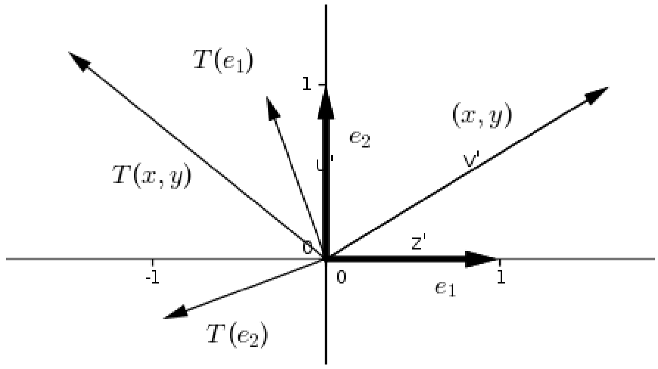
Uma justificativa para a expressão de $T(x, y)$ é a seguinte: uma trigonometria simples nos leva à concluir que

$$u_1 = T(e_1) = T(1, 0) = (\cos\theta,\ \mathrm{sen}\,\theta) \qquad \text{e} \qquad u_2 = T(e_2) = T(0, 1) = (-\mathrm{sen}\,\theta,\ \cos\theta)$$

(veja figura à esquerda). Ademais, (x, y) escrito na base $can = \{(1, 0); (0, 1)\}$ mantém os mesmos coeficientes após a rotação (os três vetores são igualmente rotacionados – veja figura à direita). Portanto

$$T(x, y) = T\big(x(1, 0) + y(0, 1)\big) = xT(1, 0) + yT(0, 1) = (\ x\cos\theta - y\,\mathrm{sen}\,\theta\ ,\ y\cos\theta + x\,\mathrm{sen}\,\theta\ ).$$





Em termos de matrizes, T pode ser representada por

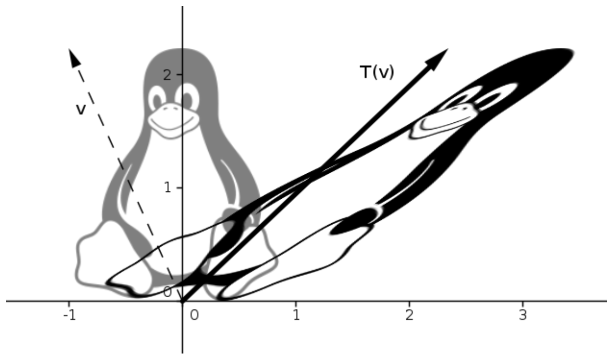
$$\begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

□

Exemplo 5.5. – Cisalhamento horizontal.

$$\begin{aligned} T : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\to (x + ay, y) \end{aligned}$$

onde $a \in \mathbb{R}$ é fixado.



$$\begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

□

Atividade 5.2. Escreva a expressão do cisalhamento vertical e interprete geometricamente fazendo uma figura. Escreva sua forma matricial.

Atividade 5.3. Mostre que as aplicações desta subseção são lineares (ou seja, verifique as duas condições da Definição 5.1).

5.3.2 Outros exemplos

Exemplo 5.6. – Aplicação nula.
Sejam V, W espaços vetoriais quaisquer. A aplicação

$$\begin{aligned} T : V &\to W \\ v &\to \mathbf{0}_W \end{aligned}$$

que leva cada vetor v de V ao vetor nulo de W é linear. A matriz associada à essa transformação é a matriz nula (olhe para sua forma matricial). □

Exemplo 5.7. – Aplicação identidade em V.
Sejam V um espaço vetorial qualquer. A aplicação

$$\begin{aligned} Id_V : V &\to V \\ v &\to v \end{aligned}$$

que leva cada vetor v de V à ele mesmo é linear. A matriz associada à essa transformação é a matriz identidade $\mathbf{I}_n$, onde $n = \dim V$. □

Exemplo 5.8. São transformações lineares:

1. $T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por $T(x, y, z) = (x, y + z)$.
2. $D : P_n \to P_{n-1}$ definida por $D(p) = p'$, onde p' é a derivada do polinômio p.
3. $S : P_n \to P_{n+1}$ definida por $S(p) = \int_0^1 p(x)dx$.
4. $T : M(n, n) \to M(n, n)$ dada por $T(A) = A^t$.
5. $T : M(2, 2) \to \mathbb{R}^4$ definida por $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a, b, c, d)$.
6. $T : T_S(2) \to \mathbb{R}^5$ dada por $T\left(\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}\right) = (\ a + b\ ,\ c\ ,\ 0\ ,\ b + c\ ,\ 2a\ )$.

□

Note que as transformações do exemplos acima (exceto o item 4), são de um espaço vetorial em outro diferente.
Observe inclusive que as dimensões dos espaços do domínio e contra-domínio podem ser diferentes – veja por exemplo o item 1.

Atividade 5.4. Mostre que as funções dos exemplos acima são transformações lineares.

Exemplo 5.9. As seguintes aplicações **NÃO** são lineares:

1. $T : M(n, n) \to \mathbb{R}$, com $n > 1$, dada por $T(\mathbf{A}) = \det \mathbf{A}$.

   De fato,
   $$T(2\mathbf{I}_n) = \det(2\mathbf{I}_n) = 2^n \neq 2 = 2 \det \mathbf{I}_n = 2T(\mathbf{I}_n),$$
   o que contradiz o item (ii) da Definição 5.1. Também **não** temos em geral o item (i).

   Atividade 5.5. Dê um exemplo de matrizes $\mathbf{A}$ e $\mathbf{B}$ de ordem 2×2 tais que $T(\mathbf{A} + \mathbf{B}) = \det(\mathbf{A} + \mathbf{B}) \neq \det \mathbf{A} + \det \mathbf{B} = T(\mathbf{A}) + T(\mathbf{B})$.

2. $T : \mathbb{R} \to \mathbb{R}$ definida por $T(x) = x + 1$

   De fato, $T(0) = 1 \neq 0$, o que contradiz o item (ii) da Definição 5.1 para $a = 0$.

$\square$

O item 2 do exemplo anterior indica que retas que **não** passam pela origem **não** são gráficos de funções lineares. De fato, o item (ii) da Definição 5.1 exige que $T(0) = 0$.

5.4 Conceitos e teoremas

Teorema 5.1. *Sejam V e W espaços vetoriais e $\beta = \{v_1, \ldots, v_n\}$ uma base de V. Dados $w_1, \ldots, w_n \in W$ quaisquer, existe uma única transformação linear $T : V \to W$ tal que $T(v_i) = w_i$ para todo $i = 1, \ldots, n$.*

Esse teorema diz que uma transformação linear fica bem definida
dizendo apenas seu valor nos elementos de uma base.

Esta é uma das facilidades das transformações lineares: é possível calcular $T(v)$ sabendo **somente** a imagem por T em cada vetor de uma base!!! Isso não é possível em funções não lineares: pense por exemplo na função não linear $f(x) = x^2$.

Os exemplos a seguir ilustram como podemos calcular a transformação linear somente a partir dos seus valores sobre uma base do espaço vetorial do domínio.

Exemplo 5.10. Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ uma transformação linear tal que

$$T(1,0) = (2,-1,0) \qquad \text{e} \qquad T(0,1) = (0,0,1)$$

(imagens por T sobre a base canônica de $\mathbb{R}^2$). Dado um vetor $(x, y) \in \mathbb{R}^2$ qualquer, queremos determinar $T(x, y)$. Ora, como $can = \{(1,0); (0,1)\}$ é base de $\mathbb{R}^2$, o Teorema 5.1 diz que T está bem definida. Escrevemos (x, y) na base *can*:

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

Aplicando T e usando a lineridade obtemos

$$T(x, y) = T(x(1,0) + y(0,1)) = xT(1,0) + yT(0,1) = x(2,-1,0) + y(0,0,1)$$

donde conclue-se que

$$T(x, y) = (2x, -x, y).$$

$\square$

Exemplo 5.11. Dada $T : \mathbb{R}^2 \to P_2$ onde

$$T(1,1)(x) = 2 - 3x + x^2 \qquad \text{e} \qquad T(2,3)(x) = 1 - x^2,$$

vamos encontrar $T(a, b)$. Sendo $\beta = \{(1,1); (2,3)\}$ base de $\mathbb{R}^2$, T está bem definida. Fixado $v = (a, b)$ qualquer, queremos encontrar $[\,v\,]_\beta$. Temos

$$[\,I\,]^{\beta}_{can\mathbb{R}^2} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$$

e, calculando a inversa dessa matriz, obtemos

$$[\,I\,]_{\beta}^{can\mathbb{R}^2} = \left([\,I\,]_{can\mathbb{R}^2}^{\beta}\right)^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}.$$

Assim,

$$[\,v\,]_{\beta} = [\,I\,]_{\beta}^{can\mathbb{R}^2}[\,v\,]_{can\mathbb{R}^2} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix},$$

ou seja,

$$v = (3a-2b)(1,1) + (-a+b)(2,3).$$

Aplicando T e usando a linearidade obtemos

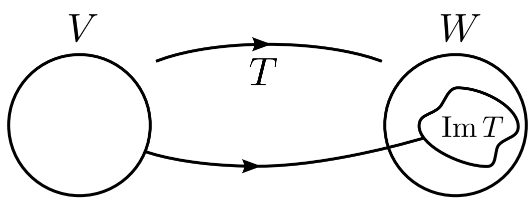
$$\begin{aligned} T(a,b)(x) &= (3a-2b)T(1,1)(x) + (-a+b)T(2,3)(x) \\ &= (3a-2b)(2-3x+x^2) + (-a+b)(1-x^2) \\ &= (5a-3b) + (-9a+6b)x + (4a-3b)x^2. \end{aligned}$$

Assim, por exemplo, $T(-1,2)(x) = -11 + 21x - 10x^2$. □

5.4.1 Imagem e Núcleo

Definição 5.2. *Seja $T : V \to W$ uma transformação linear. A **imagem** de T é conjunto*

$$\operatorname{Im} T = \{w \in W \mid T(v) = w \text{ para algum } v \in V\} = \{T(v) \mid v \in V\}.$$



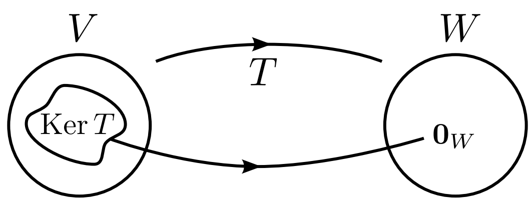
A imagem de T é o conjunto de todos os vetores do espaço de chegada, o contra-domínio W, que são imagem de algum vetor por T. Pode-se ainda escrever "$T(V)$"para denotar a imagem de T.

A imagem de $T : V \to W$ é um subespaço vetorial de W.

De fato, se $w_1, w_2 \in \operatorname{Im} T$ então existem $v_1, v_2 \in V$ tais que $T(v_1) = w_1$ e $T(v_2) = w_2$. Assim, $w_1 + w_2 = T(v_1) + T(v_2) = T(v_1 + v_2)$, donde segue que $w_1 + w_2 \in \operatorname{Im} T$. Também, se $w = T(v) \in \operatorname{Im} T$ e $a \in \mathbb{R}$ então $aw = aT(v) = T(av)$ e logo $aw \in \operatorname{Im} T$.

Definição 5.3. *Seja $T : V \to W$ uma transformação linear. O **núcleo** de T é conjunto*

$$\operatorname{Ker} T = \{v \in V \mid T(v) = \mathbf{0}\}.$$



O núcleo de T é o conjunto dos vetores do domínio V cuja imagem por T é o vetor nulo de W. O termo "Ker" vem do inglês *kernel.* Alguns livros escrevem "Nuc(T)" ou ainda "$\mathcal{N}(T)$".

O núcleo de $T : V \to W$ é um subespaço de V.

De fato,
$v_1, v_2 \in \operatorname{Ker} T \Rightarrow T(v_1) = T(v_2) = \mathbf{0} \Rightarrow T(v_1 + v_2) = T(v_1) + T(v_2) = \mathbf{0} \Rightarrow v_1 + v_2 \in \operatorname{Ker} T$ e
$v \in \operatorname{Ker} T, a \in \mathbb{R} \Rightarrow T(av) = aT(v) = a\mathbf{0} = \mathbf{0} \Rightarrow av \in \operatorname{Ker} T$.

5.4.2 Injetividade e sobrejetividade

Recapitulando a definição de funções injetoras e sobrejetoras, que vale para qualquer função, e geralmente é vista no ensino médio:

Definição 5.4. *A aplicação $T : V \to W$ é* ***injetora*** *se dados $u, v \in V$ com $T(u) = T(v)$ tivermos $u = v$.*

Equivalentemente, T é injetora se dados $u, v \in V$ com $u \neq v$, então $T(u) \neq T(v)$.

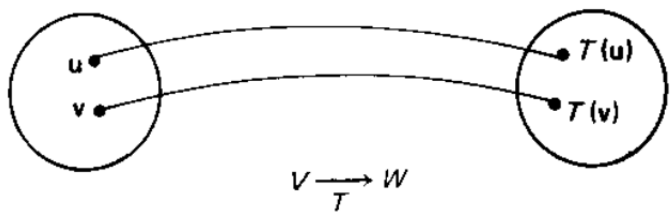


Imagem retirada de [2].

Definição 5.5. *A aplicação $T : V \to W$ é* ***sobrejetora*** *se sua imagem é todo o contradomínio, ou seja, se* $\operatorname{Im} T = W$.

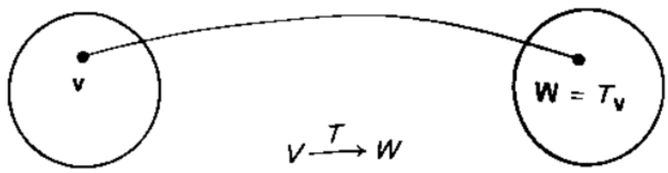


Imagem retirada de [2].

No caso de transformações lineares, podemos caracterizar injetividade através do núcleo:

Teorema 5.2. *Seja $T : V \to W$ uma transformação linear. Então* $\operatorname{Ker} T = \{\mathbf{0}\}$ *se, e somente se T é injetora.*

Exemplo 5.12. Considere a transformação $T : \mathbb{R}^3 \to \mathbb{R}^4$ dada por

$$T(x, y, z) = (\ x + y\ ,\ y\ ,\ 2x\ ,\ z\ ).$$

Vamos encontrar $\operatorname{Im} T$ e $\operatorname{Ker} T$. Temos

$$\begin{aligned}
\operatorname{Im} T &= \{T(u) \in \mathbb{R}^4 \mid v \in \mathbb{R}^3\} \\
&= \{(x + y, y, 2x, z) \mid x, y, z \in \mathbb{R}\} \\
&= \{x(1, 0, 2, 0) + y(1, 1, 0, 0) + z(0, 0, 0, 1) \mid x, y, z \in \mathbb{R}\} \\
&= [\ (1, 0, 2, 0)\ ;\ (1, 1, 0, 0)\ ;\ (0, 0, 0, 1)\ ].
\end{aligned}$$

Como $\beta_I = \{(1,0,2,0);(1,1,0,0);(0,0,0,1)\}$ é LI (verifique!), β_I é uma base de $\operatorname{Im} T$, e logo $\dim \operatorname{Im} T = 3$. Segue ainda que T não é sobrejetora.

Da mesma forma,

$$\operatorname{Ker} T = \{u \in \mathbb{R}^3 \mid T(u) = \mathbf{0}\} = \{(x,y,z) \mid (x+y, y, 2x, z) = (0,0,0,0)\} = \{(0,0,0)\},$$

$\dim \operatorname{Ker} T = 0$ e T é injetora, pelo Teorema 5.2. □

Toda transformação $T : V \to W$ injetora leva vetores LI $\{v_1, \ldots, v_n\}$ de V em vetores LI $\{T(v_1), \ldots, T(v_n)\}$ de W.

De fato,

$$\begin{aligned} a_1T(v_1) + \cdots + a_nT(v_n) = \mathbf{0} &\Rightarrow T(a_1v_1 + \cdots + a_nv_n) = \mathbf{0} \\ &\Rightarrow a_1v_1 + \cdots + a_nv_n \in \operatorname{Ker} T. \end{aligned}$$

Como T é injetora, do Teorema 5.2 segue que $\operatorname{Ker} T = \{\mathbf{0}\}$. Logo $a_1v_1 + \cdots + a_nv_n = \mathbf{0}$. Mas $\{v_1, \ldots, v_n\}$ é LI, e logo $a_1 = \cdots = a_n = 0$, donde conclue-se que $\{T(v_1), \ldots, T(v_n)\}$ é LI, como queríamos.

Um importante resultado é o Teorema do Núcleo e da Imagem. Ele relaciona as dimensões dos subespaços $\operatorname{Im} T$ e $\operatorname{Ker} T$ com a dimensão do domínio.

Teorema 5.3 (do núcleo e da imagem)**.** *Seja $T : V \to W$ uma transformação linear. Então*

$$\dim V = \dim \operatorname{Im} T + \dim \operatorname{Ker} T.$$

Uma consequência do Teorema do Núcleo e da Imagem é a seguinte correspondência entre injetividade e sobrejetividade:

Corolário 5.1. *Seja $T : V \to W$ uma transformação linear onde* $\dim V = \dim W$. *Então T é injetora se, e somente se T é sobrejetora.*

Uma transformação $T : V \to W$ bijetora (injetora e sobrejetora simultaneamente) é chamada **isomorfismo**.

Quando existe um isomorfismo entre dois espaços vetoriais V e W, dizemos que V e W são espaços vetoriais *isomorfos*. É comum também dizer que uma transformação linear $T : V \to V$ de V no próprio V é um *operador* (sobre V).

Exemplo 5.13. O seguinte fato é verdadeiro:

Isomorfismos levam bases de V em bases de W.

Vamos mostrar esse fato. Seja $T : V \to W$ um isomorfismo e $\beta_V = \{v_1, \ldots, v_n\}$ uma base de V. Devemos mostrar que $\beta_W = \{T(v_1), \ldots, T(v_n)\}$ é base de W. Primeiramente, como T é injetora, β_W é LI (observação anterior). Agora, afirmamos que β_W é base de $\operatorname{Im} T$. De fato, é LI e dado qualquer $w = T(v) \in \operatorname{Im} T$, segue que

$$v = a_1v_1 + \cdots + a_nv_n \quad \Rightarrow \quad w = T(v) = T(a_1v_1 + \cdots + a_nv_n) = a_1T(v_1) + \cdots + a_nT(v_n),$$

ou seja, β_W gera $\operatorname{Im} T$. Mas T é sobrejetora, e assim $W = \operatorname{Im} T = [\beta_W]$, isto é, β_W é base de W, como queríamos demonstrar. □

Todo isomorfismo $T : V \to W$ admite uma única inversa $T^{-1} : W \to V$ tal que $T^{-1} \circ T = Id_V$ e $T \circ T^{-1} = Id_W$. Neste caso, T^{-1} é também um isomorfismo.

Atividade 5.6. Mostre que um isomorfismo admite uma única inversa e que esta é também um isomorfismo.

Exemplo 5.14. Considere a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$T(x, y, z) = (\ x - 2y\ ,\ z\ ,\ x + y\ ).$$

Vamos mostrar que T é isomorfismo. Note que, tendo em vista o Corolário 5.1, basta mostrar que T é injetora pois domínio e contradomínio têm mesma dimensão. Ora,

$$\operatorname{Ker} T = \{(x, y, z) \mid (x - 2y, z, x + y) = (0, 0, 0)\},$$

e logo $(x, y, z) \in \operatorname{Ker} T$ se, e somente se

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

Esse sistema possui somente a solução trivial, e portanto $\operatorname{Ker} T = \{(0, 0, 0)\}$ ($\Rightarrow T$ é injetora). Concluímos então que T é isomorfismo.

Agora, calculemos T^{-1}. Sabemos que, sendo T isomorfismo, leva base em base. Em particular, o conjunto

$$\beta = \{T(1, 0, 0); T(0, 1, 0); T(0, 0, 1)\} = \{(1, 0, 1); (-2, 0, 1); (0, 1, 0)\}$$

proveniente da aplicação de T na base canônica de $\mathbb{R}^3$ é base de $\mathbb{R}^3$. Assim

$$\begin{aligned} T^{-1} \circ T(1, 0, 0) = (1, 0, 0) &\quad \Rightarrow \quad T^{-1}(1, 0, 1) = (1, 0, 0), \\ T^{-1} \circ T(0, 1, 0) = (0, 1, 0) &\quad \Rightarrow \quad T^{-1}(-2, 0, 1) = (0, 1, 0) \text{ e} \\ T^{-1} \circ T(0, 0, 1) = (0, 0, 1) &\quad \Rightarrow \quad T^{-1}(0, 1, 0) = (0, 0, 1). \end{aligned}$$

Ora, T^{-1} está definida sobre a base β de $\mathbb{R}^3$, e T^{-1} está bem definida em todo $\mathbb{R}^3$. Dado $(x, y, z) \in \mathbb{R}^3$, você pode verificar que

$$(x, y, z) = \frac{x + 2y}{3}(1, 0, 1) + \frac{z - x}{3}(-2, 0, 1) + y(0, 1, 0),$$

donde segue que

$$T^{-1}(x, y, z) = \left(\ \frac{x + 2y}{3}\ ,\ \frac{z - x}{3}\ ,\ y\ \right).$$

$\square$

5.5 A matriz de uma transformação linear

Recorde da Seção 5.3 que cada transformação tem uma escrita na forma matricial. Para exemplificar, considere a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x + y, 2x)$, e a matriz

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

Note que

$$\mathbf{AX} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ 2x \end{bmatrix}.$$

Assim, a imagem de (x, y) por T coincide, interpretada como uma matriz coluna, com o produto $\mathbf{AX}$. Isso nos motiva a associar transformações lineares a matrizes.

Dada uma transformação linear $T : V \to W$ e bases $\alpha = \{v_1, \ldots, v_n\}$, $\beta = \{w_1, \ldots, w_k\}$ de V e W, respectivamente, nosso objetivo é associar uma matriz à transformação T relativa às bases α e β. Como $T(v_1), \ldots, T(v_n) \in W$, escrevemos

$$\begin{cases} T(v_1) &= a_{11}w_1 + a_{21}w_2 + \cdots + a_{k1}w_k \\ T(v_2) &= a_{12}w_1 + a_{22}w_2 + \cdots + a_{k2}w_k \\ \quad\vdots & \\ T(v_n) &= a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{kn}w_k \end{cases}.$$

Agora, seja $v \in V$ e as escritas de v e $T(v)$ nas bases α e β, respectivamente, digamos $v = y_1v_1 + \cdots + y_nv_n$ e $T(v) = x_1w_1 + \cdots + x_kw_k$. Ou seja,

$$[\, v\,]_\alpha = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \text{e} \quad [\, T(v)\,]_\beta = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}.$$

Assim,

$$\begin{aligned} T(v) &= T(y_1v_1 + \cdots + y_nv_n) = y_1T(v_1) + \cdots + y_nT(v_n) \\ &= y_1(a_{11}w_1 + a_{21}w_2 + \cdots + a_{k1}w_k) + \cdots + y_n(a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{kn}w_k) \\ &= (y_1a_{11} + y_2a_{12} + \cdots + y_na_{1n})w_1 + \cdots + (y_1a_{k1} + y_2a_{k2} + \cdots + y_na_{kn})w_k. \end{aligned}$$

Mas a escrita de $T(v)$ na base β é única, e portanto

$$\begin{cases} x_1 &= a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ x_2 &= a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ \quad\vdots & \\ x_k &= a_{k1}y_1 + a_{k2}y_2 + \cdots + a_{kn}y_n \end{cases},$$

ou seja,

$$\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Fazendo

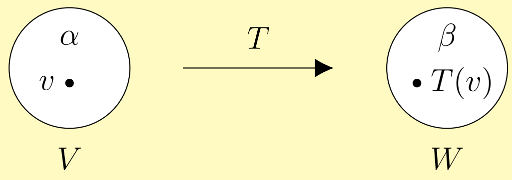
$$[\, T\,]^\alpha_\beta = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kn} \end{bmatrix},$$

chegamos à expressão para compacta para $[\, T(v)\,]_\beta$:

Teorema 5.4. *Sejam V e W espaços vetoriais, α base de V, β base de W e $T : V \to W$ uma transformação linear. Então*

$$[\,T(v)\,]_\beta = [\,T\,]_\beta^\alpha\,[\,v\,]_\alpha.$$

A matriz $[\,T\,]_\beta^\alpha$ é chamada *matriz de T em relação às bases α e β.*



Note que α é a base de partida (do domínio V) e β a base de chegada (do contradomínio W). Ou seja, a ordem dos índices coincide com àquela da escrita de matrizes de mudança de base.

Como fizemos com matrizes de mudança de base, podemos ver $[\,T\,]_\beta^\alpha$ em colunas. Observe que as colunas de $[\,T\,]_\beta^\alpha$ são os coeficientes das escritas das imagens $T(v_1), \ldots, T(v_n)$ na base β de W, isto é,

$$[\,T\,]_\beta^\alpha = \left[\begin{array}{ccc} | & & | \\ [\,T(v_1)\,]_\beta & \cdots & [\,T(v_n)\,]_\beta \\ | & & | \end{array}\right].$$

Com isso, uma transformação de V em W estará bem definida se dissermos sua matriz sobre bases de V e W. Observe também que se $n = \dim V$ e $k = \dim W$ então uma matriz da transformação $T : V \to W$ tem ordem $k \times n$.

Volte à seção de mudança de base e compare com as contas feitas aqui. Veja que a matriz de mudança de base é a matriz do operador linear identidade. Nesse sentido, a notação usada mostra sua conveniência: se $T = Id_V = I$ é o operador identidade sobre V, então a expressão $[\,T(v)\,]_\beta = [\,T\,]_\beta^\alpha\,[\,v\,]_\alpha$ fica $[\,v\,]_\beta = [\,I\,]_\beta^\alpha\,[\,v\,]_\alpha$, a mesma expressão obtida no estudo de matrizes de mudança de base.

Exemplo 5.15. Considere a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por

$$T(x, y, z) = (\ 2x - y + z\ ,\ 3x + y - 2z\ ).$$

a) A matriz de T, $[\,T\,]_\beta^\alpha$, nas bases $\alpha = \{(1,1,1); (0,1,1); (0,0,1)\}$ de $\mathbb{R}^3$ e $\beta = \{(2,1); (5,3)\}$ de $\mathbb{R}^2$ é

$$[\,T\,]_\beta^\alpha = \left[\begin{array}{ccc} | & | & | \\ [\,T(1,1,1)\,]_\beta & [\,T(0,1,1)\,]_\beta & [\,T(0,0,1)\,]_\beta \\ | & | & | \end{array}\right] = \left[\begin{array}{ccc} | & | & | \\ [\,(2,2)\,]_\beta & [\,(0,-1)\,]_\beta & [\,(1,-2)\,]_\beta \\ | & | & | \end{array}\right].$$

Para encontrar essa matriz, temos que encontrar as escritas dos vetores $(2,2), (0,-1)$ e $(1,-2)$ na base β. Fazendo as contas, chegamos à escrita do vetor (x,y) na base β:

$$[\,(x,y)\,]_\beta = \begin{bmatrix} 3x - 5y \\ -x + 2y \end{bmatrix}.$$

Assim,

$$[\,T\,]_\beta^\alpha = \begin{bmatrix} -4 & 5 & 13 \\ 2 & -2 & -5 \end{bmatrix}.$$

b) Usando a matriz $[\,T\,]_\beta^\alpha$ vamos encontrar $[\,T(3,-4,2)\,]_\beta$. Ora, temos

$$[\,T(3,-4,2)\,]_\beta = [\,T\,]_\beta^\alpha\,[\,(3,-4,2)\,]_\alpha.$$

Sabemos que

$$[\,(3,-4,2)\,]_\alpha = [\,I\,]_\alpha^{can\mathbb{R}^3}[\,(3,-4,2)\,]_{can\mathbb{R}^3} = ([\,I\,]_{can\mathbb{R}^3}^\alpha)^{-1}[\,(3,-4,2)\,]_{can\mathbb{R}^3}.$$

Você pode fazer as contas e descobrir que

$$[\,(3,-4,2)\,]_\alpha = \begin{bmatrix} 3 \\ -7 \\ 6 \end{bmatrix}.$$

Assim,

$$[\,T(3,-4,2)\,]_\beta = \begin{bmatrix} -4 & 5 & 13 \\ 2 & -2 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -7 \\ 6 \end{bmatrix} = \begin{bmatrix} 31 \\ -10 \end{bmatrix}$$

(você pode verificar que de fato $T(3,-4,2) = (12,1) = 31(2,1) - 10(5,3)$).

c) A matriz $[\,T\,]_{can\mathbb{R}^2}^\alpha$ é

$$[\,T\,]_{can\mathbb{R}^2}^\alpha = \begin{bmatrix} | & | & | \\ [\,T(1,1,1)\,]_{can} & [\,T(0,1,1)\,]_{can} & [\,T(0,0,1)\,]_{can} \\ | & | & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & | \\ [\,(2,2)\,]_{can} & [\,(0,-1)\,]_{can} & [\,(1,-2)\,]_{can} \\ | & | & | \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 \\ 2 & -1 & -2 \end{bmatrix}.$$

d) Você pode verificar que $[\,T\,]_\beta^\alpha = [\,I\,]_\beta^{can}\,[\,T\,]_{can}^\alpha$. Isso mostra mais uma vez a força da notação usada. Observe que a matriz $[\,T\,]_{can}^\alpha$ "vai" da base α (do domínio) para a base *can* (do contradomínio). Assim, para obter a matriz de T que vai de α para a base β, basta mudar a base do contradomínio da *can* para β, através da multiplicação da matriz de mudança de base $[\,I\,]_\beta^{can}$.

$\square$

Teorema 5.5. *Seja $T : V \to W$ uma transformação linear, α uma base de V e β uma base de W. Então*

$$\dim \operatorname{Im} T = \textit{posto de } [\,T\,]_\beta^\alpha \quad \textit{e} \quad \dim \operatorname{Ker} T = \textit{nulidade de } [\,T\,]_\beta^\alpha.$$

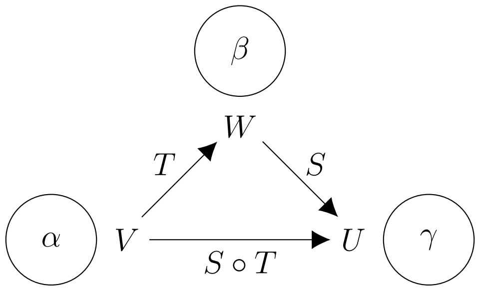
Teorema 5.6. *Sejam $T : V \to W$ e $S : W \to U$ transformações lineares, α, β e γ bases de V, W e U respectivamente. Então a composta*

$$\begin{aligned} S \circ T : V &\to U \\ v &\to (S \circ T)(v) = S(T(v)) \end{aligned}$$

é transformação linear, e

$$[\,S \circ T\,]_\gamma^\alpha = [\,S\,]_\gamma^\beta\,[\,T\,]_\beta^\alpha$$

(o produto matricial deve ser feito na mesma ordem da composição).



Exemplo 5.16. Considere os operadores lineares $T, S : \mathbb{R}^2 \to \mathbb{R}^2$ dados por $T(x,y) = (2x, 2y)$ e $S(x,y) = (x + 2y, y)$. Temos

$$[\,S \circ T\,]_{can}^{can} = [\,S\,]_{can}^{can}\,[\,T\,]_{can}^{can} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix}.$$

□

Exemplo 5.17. Sejam $T : \mathbb{R}^2 \to \mathbb{R}^3$ e $S : \mathbb{R}^3 \to \mathbb{R}^2$ definidas por

$$[\,T\,]_\beta^\alpha = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \quad \text{e} \quad [\,S\,]_\gamma^\beta = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

onde $\alpha = \{(1,0);(0,2)\}$ e $\gamma = \{(2,0);(1,1)\}$ são bases de $\mathbb{R}^2$ e β é uma base de $\mathbb{R}^3$. Vamos encontrar $S \circ T(x,y)$. Ora,

$$[\,S \circ T\,]_\gamma^\alpha = [\,S\,]_\gamma^\beta\,[\,T\,]_\beta^\alpha = \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}.$$

É fácil ver que

$$[\,(x,y)\,]_\alpha = \begin{bmatrix} x \\ y/2 \end{bmatrix}.$$

Assim,

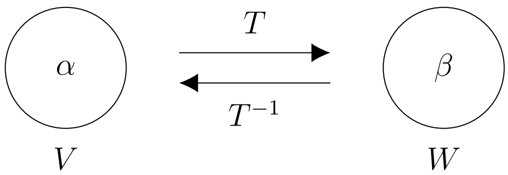
$$[\,S \circ T(x,y)\,]_\gamma = [\,S \circ T\,]_\gamma^\alpha\,[\,(x,y)\,]_\alpha = \begin{bmatrix} x - y \\ 0 \end{bmatrix},$$

ou seja,

$$S \circ T(x,y) = (x-y)(2,0) + 0(1,1) = (2x - 2y, 0).$$

□

Corolário 5.2. *Se $T : V \to W$ é um isomorfismo, α uma base de V e β uma base de W, então $[\,T^{-1}\,]_\alpha^\beta = \left([\,T\,]_\beta^\alpha\right)^{-1}$.*



Sabemos que matrizes quadradas possuem inversa se, e somente se, seu determinante é não nulo. Portanto, a mesma relação vale entre isomorfismos (isto é, a existência de T^{-1}) e determinantes.

Corolário 5.3. *Sejam V e W espaços vetoriais de mesma dimensão, $T : V \to W$ uma transformação linear e α, β bases de V e W respectivamente. Então T é isomorfismo se, e somente se* $\det[\,T\,]_\beta^\alpha \neq 0$.

Exemplo 5.18. Considere o operador $T : \mathbb{R}^2 \to \mathbb{R}^2$ dado por

$$T(x, y) = (3x + 4y, 2x + 3y).$$

a) T é isomorfismo pois o determinante da matriz $[\,T\,]_{can}^{can} = \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}$ é não nulo.

b) Vamos encontrar $T^{-1}(x, y)$. Ora,

$$[\,T^{-1}\,]_{can}^{can} = \left([\,T\,]_{can}^{can}\right)^{-1} = \begin{bmatrix} 3 & -4 \\ -2 & 3 \end{bmatrix} \Rightarrow [\,T^{-1}(x, y)\,]_{can} = [\,T^{-1}\,]_{can}^{can}[\,(x, y)\,]_{can} = \begin{bmatrix} 3x - 4y \\ -2x + 3y \end{bmatrix},$$

e assim $T^{-1}(x, y) = (3x - 4y)(1, 0) + (-2x + 3y)(0, 1) = (3x - 4y, -2x + 3y)$.

□

Corolário 5.4. *Sejam $T : V \to W$ uma transformação linear, α, α' bases de V e β, β' bases de W. Então*

$$[\,T\,]_{\beta'}^{\alpha'} = [\,I\,]_{\beta'}^{\beta}\,[\,T\,]_\beta^\alpha\,[\,I\,]_\alpha^{\alpha'}$$

(o encadeamento da primeira para a última base se lê da direita para a esquerda no produto das matrizes – veja a figura abaixo).

$$\begin{tikzcd}
 & & \alpha & \beta & & \\
 & & V \arrow[r, "T"] & W \arrow[dr, "I"] & & \\
\alpha' & V \arrow[ur, "I"] \arrow[rrr, "I \circ T \circ I = T"'] & & & W & \beta'
\end{tikzcd}$$

É importante ler e escrever a ordem dos índices nas matrizes de forma correta. Note que as bases no produto $[\,I\,]_{\beta'}^{\beta}\,[\,T\,]_\beta^\alpha\,[\,I\,]_\alpha^{\alpha'}$ "se cancelam", restando a base de partida α' e a de chegada β'.

Exemplo 5.19. Considere a transformação do exemplo anterior, ou seja, $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$T(x, y) = (\ 3x + 4y\ ,\ 2x + 3y\ ).$$

Dadas as bases $\alpha' = \{(1, 2); (-1, 3)\}$ e $\beta' = \{(1, 1); (2, 0)\}$, vamos calcular a matriz de T da base α' para a base β' utilizando o Corolário 5.4. Este resultado garante que

$$[\,T\,]_{\alpha'}^{\beta'} = [\,I\,]_{\alpha'}^{can}\,[\,T\,]_{can}^{can}\,[\,I\,]_{can}^{\beta'}.$$

Veja que as matrizes envolvendo a base canônica são fáceis de calcular. Temos

- $[\,I\,]_{\alpha'}^{can} = \left([\,I\,]_{can}^{\alpha'}\right)^{-1} = \begin{bmatrix} | & | \\ [\,(1, 2)\,]_{can} & [\,(-1, 3)\,]_{can} \\ | & | \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}^{-1} = \dfrac{1}{5}\begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}$

- $[\,T\,]_{can}^{can} = \left[\begin{array}{cc} | & | \\ [\,T(1,0)\,]_{can} & [\,T(0,1)\,]_{can} \\ | & | \end{array}\right] = \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}$

- $[\,I\,]_{can}^{\beta'} = \left[\begin{array}{cc} | & | \\ [\,(1,1)\,]_{can} & [\,(2,0)\,]_{can} \\ | & | \end{array}\right] = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}.$

Assim,

$$[\,T\,]_{\alpha'}^{\beta'} = \frac{1}{5}\begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 26 & 22 \\ -9 & -8 \end{bmatrix}.$$

□

5.6 Demonstrações

Demonstração do Teorema 5.1. Sejam T e S duas transformações de V em W com $T(v_i) = S(v_i) = w_i$ para todo i. Devemos mostrar que $T = S$, ou seja, que $T(v) = S(v)$ para qualquer $v \in V$. Fixado então $v \in V$ arbitrário, escrevemos

$$v = a_1v_1 + \cdots + a_nv_n.$$

Assim, pela linearidade de T segue que

$$T(v) = T\left(\sum_1^n a_iv_i\right) = \sum_1^n a_iT(v_i).$$

Da mesma forma, $S(v) = \sum_1^n a_iS(v_i)$. Conclue-se então que $S(v) = T(v)$, como queríamos. □

Demonstração do Teorema 5.2. Suponha que $\operatorname{Ker} T = \{\mathbf{0}\}$. Então dados $u, v \in V$ tais que $T(u) = T(v)$, temos

$$T(u - v) = T(u) - T(v) = \mathbf{0},$$

ou seja $u - v \in \operatorname{Ker} T$. Assim, $u = v$, e T é injetora.

Reciprocamente, suponha que T é injetora, e seja $u \in \operatorname{Ker} T$. Então, como $T(\mathbf{0}) = \mathbf{0} = T(u)$, segue que $u = \mathbf{0}$, ou seja, $\operatorname{Ker} T = \{\mathbf{0}\}$. □

Demonstração do Teorema 5.3. Tome uma base de $\operatorname{Ker} T$, digamos $\{v_1, \ldots, v_p\} \subset V$ (neste caso, $\dim \operatorname{Ker} T = p$). Completando esse conjunto a uma base de V, obtemos uma base $\{v_1, \ldots, v_p, v_{p+1}, \ldots, v_n\}$ de V formada, evidentemente, por $n = \dim V$ vetores. Vamos mostrar que os $n - p$ vetores $T(v_{p+1}), \ldots, T(v_n)$ formam uma base de $\operatorname{Im} T$. De fato, esses vetores são LI pois, pela linearidade de T, temos

$$a_{p+1}T(v_{p+1}) + \cdots + a_nT(v_n) = T(a_{p+1}v_{p+1} + \cdots + a_nv_n) = \mathbf{0} \;\Rightarrow\; a_{p+1}v_{p+1} + \cdots + a_nv_n \in \operatorname{Ker} T.$$

Daí, esta soma dos v's pode ser escrita na base de $\operatorname{Ker} T$, isto é,

$$a_{p+1}v_{p+1} + \cdots + a_nv_n = b_1v_1 + \cdots + b_pv_p,$$

o que fornece

$$-b_1v_1 - \cdots - b_pv_p + a_{p+1}v_{p+1} + \cdots + a_nv_n = \mathbf{0}.$$

Ora, mas os vetores $v_1, \ldots, v_n$ são LI, e logo devemos ter $a_{p+1} = \cdots = a_n = 0$. Portanto

$$a_{p+1}T(v_{p+1}) + \cdots + a_nT(v_n) = \mathbf{0} \Rightarrow a_{p+1} = \cdots = a_n = 0,$$

ou seja, $T(v_{p+1}), \ldots, T(v_n)$ são LI.

Agora, dado $T(u) \in \operatorname{Im} T$, escrevemos $u \in V$ na base de V:

$$u = c_1v_1 + \cdots + c_pv_p + c_{p+1}v_{p+1} + \cdots + c_nv_n.$$

Aplicando T e usando linearidade obtemos

$$T(u) = \underbrace{c_1T(v_1) + \cdots + c_pT(v_p)}_{=\mathbf{0} \text{ pois } v_1,\ldots,v_p \in \operatorname{Ker} T} + T(c_{p+1}v_{p+1} + \cdots + a_nv_n) = c_{p+1}T(v_{p+1}) + \cdots + c_nT(v_n).$$

Ou seja, $\{T(v_{p+1}), \ldots, T(v_n)\}$ gera $\operatorname{Im} T$, e logo é uma base de $\operatorname{Im} T$. Finalmente basta contar os elementos das bases de $\operatorname{Ker} T$ e $\operatorname{Im} T$ para concluir que

$$\dim V = n = (n - p) + p = \dim \operatorname{Im} T + \dim \operatorname{Ker} T.$$

$\square$

Demonstração do Corolário 5.1. Seja $T : V \to W$ transformação linear. Temos

$$\begin{aligned}
T \text{ injetora } &\Leftrightarrow \operatorname{Ker} T = \{\mathbf{0}\} && \text{(Teorema 5.2)}\\
&\Leftrightarrow \dim \operatorname{Ker} T = 0 &&\\
&\Leftrightarrow \dim \operatorname{Im} T = \dim V = \dim W && \text{(Teorema 5.3 [do Núcleo e da Imagem])}\\
&\Leftrightarrow \operatorname{Im} T = W && \text{(Atividade 4.8 de Espaços Vetoriais)}\\
&\Leftrightarrow T \text{ sobrejetora} && \text{(Definição 5.5).}
\end{aligned}$$

$\square$

Demonstração do Teorema 5.4. A demonstração está no próprio texto, antes do Teorema. $\square$

Demonstração do Teorema 5.5. Seja $\mathbf{R}$ a MLRFE de $[T]^{\alpha}_{\beta}$. Essa matriz não é necessariamente quadrada: tem ordem $m \times n$, onde $m = \dim W$ e $n = \dim V$. Consideremos o sistema homogêneo $\mathbf{Rv} = \mathbf{0}$. Como $\mathbf{R}$ está na forma escada reduzida, suas linhas nulas aparecem todas abaixo das não nulas; portanto podemos desconsiderá-las no sistema $\mathbf{Rv} = \mathbf{0}$. Reordenando as colunas se necessário, podemos então escrever o sistema $\mathbf{Rv} = \mathbf{0}$ de forma equivalente como

$$\begin{bmatrix} \mathbf{I}_p & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{v}_I \\ \mathbf{v}_N \end{bmatrix} = \mathbf{0},$$

onde $\mathbf{I}_p$ é matriz identidade de ordem $p =$ posto de $[T]^{\alpha}_{\beta}$ e $\mathbf{N}$ a matriz restante de ordem $p \times (p - n)$. Assim, $\mathbf{v}_I$ e $\mathbf{v}_N$ têm dimensões $p \times 1$ e $(n - p) \times 1$, respectivamente. O sistema anterior pode ser escrito simplesmente como

$$\mathbf{v}_I = -\mathbf{N}\mathbf{v}_n.$$

Note que a escolha de $\mathbf{v}_N$ é livre, mas a de $\mathbf{v}_I$ não. Ademais, $\operatorname{Ker} T$ é caracterizado pelo sistema homogêneo, e logo este subespaço é caracterizado pelas possíveis escolhas para $\mathbf{v}_N$. Ora, há exatamente $n-p$ escolhas para $\mathbf{v}_N$ de modo a obter vetores LI, por exemplo, escolhendo para $\mathbf{v}_N$ os vetores canônicos de $\mathbb{R}^{n-p}$. Assim, uma base de $\operatorname{Ker} T$ tem $n-p$ vetores, e logo

$$\dim \operatorname{Ker} T = n - p = \text{nulidade de } [\,T\,]^{\alpha}_{\beta}$$

(no caso extremo onde $n-p=0$, a matriz $\mathbf{N}$ não aparece, o sistema homogêneo é equivalente à $\mathbf{v}_I = \mathbf{0}$ e então, coerentemente, $\operatorname{Ker} T = \{\mathbf{0}\}$).

A relação envolvendo a imagem de T segue do Teorema do Núcleo e da Imagem:

$$\dim \operatorname{Im} T = \dim V - \dim \operatorname{Ker} T = n - (n-p) = p = \text{posto de } [\,T\,]^{\alpha}_{\beta}.$$

Isso finaliza a demonstração. □

Demonstração do Teorema 5.6. Primeiro mostramos que a composta $S \circ T : V \to U$ é linear. Dados vetores $v_1, v_2 \in V$ e $a \in \mathbb{R}$ quaisquer, temos

$$\begin{aligned}(S \circ T)(v_1 + v_2) &= S(T(v_1 + v_2)) \\ &= S(T(v_1) + T(v_2)) = S(T(v_1)) + S(T(v_2)) \\ &= (S \circ T)(v_1) + (S \circ T)(v_2).\end{aligned}$$

Da mesma forma prova-se que $(S \circ T)(av_1) = a(S \circ T)(v_1)$.

Vamos agora mostrar que a matriz de $S \circ T$ é o produto das matrizes de S e T. Usando o Teorema 5.4 temos, por um lado,

$$[\,S \circ T(v)\,]_{\gamma} = [\,S \circ T\,]^{\alpha}_{\gamma}\,[\,v\,]_{\alpha},$$

e, por outro lado,

$$[\,S \circ T(v)\,]_{\gamma} = [\,S(T(v))\,]_{\gamma} = [\,S\,]^{\beta}_{\gamma}\,[\,T(v)\,]_{\beta} = \Big([\,S\,]^{\beta}_{\gamma}\,[\,T\,]^{\alpha}_{\beta}\Big)[\,v\,]_{\alpha}.$$

Como as duas expressões acima valem para $v \in V$ arbitrário, só pode ser $[\,S \circ T\,]^{\alpha}_{\gamma} = [\,S\,]^{\beta}_{\gamma}\,[\,T\,]^{\alpha}_{\beta}$, como queríamos demonstrar. □

Demonstração do Corolário 5.2. Temos $T^{-1} \circ T = Id_V$. Se $\alpha = \{v_1, \ldots, v_n\}$ é base de V então

$$[\,Id_V\,]^{\alpha}_{\alpha} = \begin{bmatrix} | & & | \\ [\,v_1\,]_{\alpha} & \cdots & [\,v_n\,]_{\alpha} \\ | & & | \end{bmatrix} = \mathbf{I}_n.$$

Assim,

$$\mathbf{I}_n = [\,Id_V\,]^{\alpha}_{\alpha} = [\,T^{-1} \circ T\,]^{\alpha}_{\alpha} = [\,T^{-1}\,]^{\beta}_{\alpha}\,[\,T\,]^{\alpha}_{\beta},$$

e daí $[\,T^{-1}\,]^{\beta}_{\alpha} = \Big([\,T\,]^{\alpha}_{\beta}\Big)^{-1}$. □

Demonstração do Corolário 5.3. Se T é isomorfismo, então o Corolário 5.2 diz que $[\,T\,]_\beta^\alpha$ é inversível, e assim $\det[\,T\,]_\beta^\alpha \neq 0$. Reciprocamente, se $\det[\,T\,]_\beta^\alpha \neq 0$ então $[\,T\,]_\beta^\alpha$ é inversível, isto é, existe uma matriz quadrada $\mathbf{B}$ tal que $\mathbf{B}[\,T\,]_\beta^\alpha = \mathbf{I}$. Defina então a transformação $S : W \to V$ pondo $[\,S\,]_\alpha^\beta = \mathbf{B}$. Temos

$$[\,Id_V\,]_\alpha^\alpha = \mathbf{I} = \mathbf{B}\,[\,T\,]_\beta^\alpha = [\,S\,]_\alpha^\beta\,[\,T\,]_\beta^\alpha = [\,S \circ T\,]_\alpha^\alpha,$$

ou seja, $S \circ T = Id_V$. Assim, T é isomorfismo, com $T^{-1} = S$. □

Demonstração do Corolário 5.4. Basta observar que $T = Id_W \circ T \circ Id_V$ implica

$$[\,T\,]_{\beta'}^{\alpha'} = [\,Id_W \circ T \circ Id_V\,]_{\beta'}^{\alpha'} = [\,I\,]_{\beta'}^{\beta}\,[\,T\,]_\beta^\alpha\,[\,I\,]_\alpha^{\alpha'}.$$

□

5.7 Exercícios

Transformações Lineares

1. Determine quais das seguintes aplicações são lineares. Se não for, explique o porquê.

   (a) $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $T(x, y, z) = (x, z)$.

   (b) $T : \mathbb{R}^4 \to \mathbb{R}^4$ definida por $T(x) = -x$.

   (c) $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(x) = x + (0, -1, 0)$.

   (d) $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (2x + y, y)$.

   (e) $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (2x, y - x)$.

   (f) $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (y, x)$.

   (g) $T : \mathbb{R}^2 \to \mathbb{R}$ definida por $T(x, y) = xy$.

   (h) $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(x, y, z) = (x, 2^y, 2^z)$.

   (i) $T : \mathbb{R}^4 \to \mathbb{R}^3$ definida por $T(x, y, z, w) = (x - w, y - w, x + z)$.

   (j) $T : M(2, 2) \to \mathbb{R}$ definida por $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$.

   (k) $T : M(n, n) \to \mathbb{R}^n$ definida por $T(\mathbf{A}) = (a_{11}, a_{22}, \ldots, a_{nn})$.

   (l) $T : P_n \to P_n$ onde a imagem de um polinômio $p \in P_n$ é o polinômio definido por $T(p)(t) = p(t + 1)$ para todo $t \in \mathbb{R}$.

   (m) $T : P_n \to \mathbb{R}^n$ definida por $T(p) = p(1)$.

   (n) $T : P_2 \to \mathbb{R}^3$ definida por $T(p) = (p(-1), p(0), p(1))$.

   (o) $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(x, y, z) = (x + y, y + z, x + z)$.

   (p) $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x, y) = (x + y, y, x)$.

   (q) $T : M(n, n) \to \mathbb{R}$ definida por $T(\mathbf{A}) = tr\mathbf{A} = a_{11} + a_{22} + \cdots + a_{nn}$.

2. Para cada uma das transformações lineares do exercício anterior, faça:

   (a) Determine uma base da imagem de T, $\operatorname{Im} T$, caso não se reduza a $\{\mathbf{0}\}$.

   (b) Determine uma base do núcleo de T, $\operatorname{Ker} T$, caso não se reduza a $\{\mathbf{0}\}$.

(c) Diga se T é injetora, sobrejetora ou um isomorfismo, justificando sua resposta. Caso T for isomorfismo, calcule T^{-1}.

3. Seja V um espaço vetorial real e sejam $u, w \in V$. A *reta que passa por* u *e é paralela a* w é definida pelo conjunto de todos os elementos $u + tw$ com $t \in \mathbb{R}$. O *segmento de reta* entre os *pontos extremos* u e $u + w$ é definido pelo conjunto de todos os elementos $u + tw$ com $0 \leq t \leq 1$. Seja $L : V \to U$ uma transformação linear. Mostre que a imagem por L de um segmento de reta em V é um segmento de reta em U. Quais são os pontos extremos do segmento? Mostre que a imagem por L de uma reta é sempre uma reta ou um ponto.

4. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear. Em cada caso, determine $T(x, y)$.

   (a) $T(3, 1) = (1, 2)$ e $T(-1, 0) = (1, 1)$.

   (b) $T(4, 1) = (1, 1)$ e $T(1, 1) = (3, -2)$.

   (c) $T(1, 1) = (2, 1)$ e $T(-1, 1) = (6, 3)$.

5. Existe um operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $T(1, 1, 1) = (1, 2, 3)$, $T(1, 2, 3) = (1, 4, 9)$ e $T(2, 3, 4) = (1, 8, 25)$? (dizemos que uma transformação linear T é um *operador* se seus domínio e contradomínio são os mesmos) Justifique sua resposta. Se for o caso, encontre $T(x, y, z)$.

6. Diga cada sentença é verdadeira ou falsa. Se for verdadeira, prove-a, e se for falsa, dê um contra-exemplo.

   Seja dada uma transformação linear $T : V \to W$.

   (a) Se $v \in V$ é tal que $T(v) = \mathbf{0}$ então $v = \mathbf{0}$.

   (b) Se $T(w) = T(u) + T(v)$ então $w = u + v$.

   (c) Se v é combinação linear de $u_1, \ldots, u_m$ então $T(v)$ é combinação linear de $T(u_1), \ldots, T(u_m)$.

   Seja $n = \dim V$ e $k = \dim W$.

   (d) Se T é injetora então $\dim \operatorname{Im} T = n$.

   (e) Se T é sobrejetora então $\dim \operatorname{Ker} T = n - k$.

7. Sejam V e W espaços vetoriais. Se $\dim V < \dim W$, mostre que existem transformações lineares $T : V \to W$ e $S : W \to V$ tais que T é injetora e S é sobrejetora.

8. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ dado por $T(x, y) = (ax + by, cx + dy)$. Determine as constantes $a, b, c, d \in \mathbb{R}$ de modo que $T(1, 2) = (1, 1)$ e $T(3, 4) = (2, 2)$.

9. Seja $T : V \to W$ uma transformação linear. Mostre que:

   (a) Se os vetores $T(u_1), \ldots, T(u_m) \in W$ são LI então são também LI os vetores $u_1, \ldots, u_m \in V$.

   (b) Se $V = W$ e os vetores $T(u_1), \ldots, T(u_m)$ geram W, então $u_1, \ldots, u_m$ geram V.

10. Seja G um conjunto de geradores do espaço vetorial V. Mostre que se as transformações lineares $T, S : V \to W$ são tais que $T(u) = S(u)$ para todo $u \in G$, então $T(v) = S(v)$ para todo $v \in V$, isto é, T e S são iguais.

11. Seja $T : V \to V$ um operador linear e suponha que exista um operador $S : V \to V$ tal que $T \circ S = Id_V$, onde Id_V é o operador identidade em V. Mostre que T é um *automorfismo*, isto é, que T é um isomorfismo de V no próprio V.

12. Seja $T : V \to W$ um isomorfismo. Se $S : W \to V$ é tal que $T \circ S = Id_W$, mostre que S é transformação linear.

13. Sejam $T : V \to W$ e $S : W \to U$ transformações lineares tais que $\operatorname{Ker} T = \{\mathbf{0}_V\}$ e $\operatorname{Ker} S = \{\mathbf{0}_W\}$, onde $\mathbf{0}_V$ e $\mathbf{0}_W$ são os vetores nulos de V e W, respectivamente. Mostre que $\operatorname{Ker}(S \circ T) = \{\mathbf{0}_V\}$.

14. Seja $T : V \to V$ um operador linear tal que $T^2 = T \circ T = T$. Mostre que $V = \operatorname{Ker} T \oplus \operatorname{Im} T$.

15. Um conjunto C se chama *convexo* se, para quaisquer $x, y \in C$, tem-se $(1-t)x + ty \in C$ para todo $t \in [0,1]$ (isto é, se o segmento que liga dois pontos quaisquer de C ainda está em C). Mostre que uma transformação linear $T : V \to W$ transforma todo conjunto convexo $C \subset V$ num conjunto convexo $T(C) \subset W$.

16. (a) Sejam U e W espaços vetoriais. Considere o conjunto $U \times W$ de todos os pares ordenados (u, w) com $u \in U$ e $w \in W$. Se $(u_1, w_1), (u_2, w_2) \in U \times W$ e $\alpha \in \mathbb{R}$ definimos a soma em $U \times W$ por
$$(u_1, w_1) + (u_2, w_2) = (u_1 + u_2, w_1 + w_2)$$
e a multiplicação por escalar por $\alpha(u_1, w_1) = (\alpha u_1, \alpha w_1)$. Mostre que com essas operações, $U \times W$ é espaço vetorial.
    (b) Mostre que $\dim(U \times W) = \dim U + \dim W$.
    (c) Suponha que U e W sejam subespaços de V, e considere a aplicação $T : U \times W \to V$ dada por $T(u, w) = u - w$. Mostre que T é transformação linear e use-a para mostrar que
$$\dim U + \dim W = \dim(U + W) + \dim(U \cap W).$$

17. Seja $T : V \to V$ um operador linear.
    (a) Mostre que se $T(v)$ é múltiplo de v para todo $v \in V$, então $T = \alpha Id_V$ para algum $\alpha \in \mathbb{R}$, onde Id_V é o operador identidade em V.
    (b) Usando o item anterior, mostre que se T não é da forma αId_V então existe $w \in V$ tal que $\{w, T(w)\}$ é um conjunto LI.
    (c) Mostre que se T comuta com qualquer operador sobre V, isto é, se $T \circ S = S \circ T$ para qualquer operador $S : V \to V$, então $T = \alpha Id_V$ para algum $\alpha \in \mathbb{R}$.
    Dica: Se T não é da forma αId_V, pelo item anterior existe $w \in V$ tal que $\{w, T(w)\}$ é LI. Complete esse conjunto a uma base de V, digamos $\{w, T(w), v_3, \ldots, v_n\}$. Defina então o operador $P : V \to V$ pondo $P(w) = P(T(w)) = w$ e $P(v_3) = \cdots = P(v_n) = 0$. Mostre então que $P \circ T \neq T \circ P$, e conclua o resultado.

Matrizes de Transformações Lineares

1. Considere as tranformações lineares abaixo
    i) $T_1 : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $T_1(x, y, z) = (x, z)$.
    ii) $T_2 : \mathbb{R}^4 \to \mathbb{R}^4$ definida por $T_2(v) = -v$.

iii) $T_3 : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T_3(x, y) = (2x + y, y)$.

iv) $T_4 : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T_4(x, y) = (2x, y - x)$.

v) $T_5 : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T_5(x, y) = (y, x)$.

vi) $T_6 : \mathbb{R}^4 \to \mathbb{R}^3$ definida por $T_6(x, y, z, w) = (x - w, y - w, x + z)$.

vii) $T_7 : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T_7(x, y, z) = (x + y, y + z, x + z)$.

Faça:

(a) calcule suas matrizes nas bases canônicas de $\mathbb{R}^2, \mathbb{R}^3$ e $\mathbb{R}^4$.

(b) calcule suas matrizes nas bases $\alpha = \{(1, 1), (0, 1)\}$ de $\mathbb{R}^2$, $\beta = \{(1, 1, 1), (1, 0, 0), (0, 0, 2)\}$ de $\mathbb{R}^3$ e $\gamma = \{(1, 0, 0, 0), (0, 2, 0, 0), (1, 0, 3, 0), (0, 0, 0, -1)\}$ de $\mathbb{R}^4$.

(c) usando as matrizes que você encontrou, calcule $\dim \operatorname{Im} T_i$ e $\dim \operatorname{Ker} T_i$.

(d) a partir do item anterior, diga se T_i é isomorfismo. Caso for, encontre uma matriz para T_i^{-1} em certas bases.

(e) calcule $[T_3 \circ T_1]^\beta_\alpha$, $[T_6 \circ T_2]^{can}_\beta$, $[T_1 \circ T_6]^\gamma_{can}$, $[T_5 \circ T_3]^\alpha_\alpha$, $[T_5 \circ T_1 \circ T_6]^{can}_\alpha$.

(f) calcule $[T_5 \circ T_4 \circ T_3]^{can}_\alpha$ e a partir daí encontre $(T_5 \circ T_4 \circ T_3)(x, y)$.

2. Para cada matriz A abaixo, encontre uma trasnformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ tal que a matriz de T nas bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$ seja $\mathbf{A}$.

(a) $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$

(b) $\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 1 & 3 & 4 \end{bmatrix}$

(c) $\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix}$

(d) $\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 2 \end{bmatrix}$

(e) $\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 2 & 3 & -1 & 1 \\ 0 & 2 & 4 & 2 \end{bmatrix}$

(f) $\mathbf{A} = \begin{bmatrix} 1 & 2 & -3 & 1 \\ -1 & 3 & -3 & -2 \\ 2 & 0 & 1 & 1 \\ 3 & 1 & -2 & 3 \end{bmatrix}$

3. Seja $T : V \to V$ um operador linear e $\beta = \{v_1, \ldots, v_n\}$ uma base de V. Se existem $c_1, \ldots, c_n \in \mathbb{R}$ tais que $T(v_i) = c_i v_i$, para todo i, qual a matriz $[T]^\beta_\beta$ de T na base β?

4. Seja $T : V \to V$ um operador sobre V tal que $T^k = \mathbf{0}$ para algum $k \geq 1$. Mostre que T não é isomorfismo.

5. Mostre que os espaços vetoriais $\mathbb{R}^3$ e $T_S(2)$ são isomorfos ($T_S(2)$ é o espaço das matrizes triangulares superiores de ordem 2). Mais geralmente, argumente que espaços vetoriais de mesma dimensão são isomorfos.

Capítulo 6

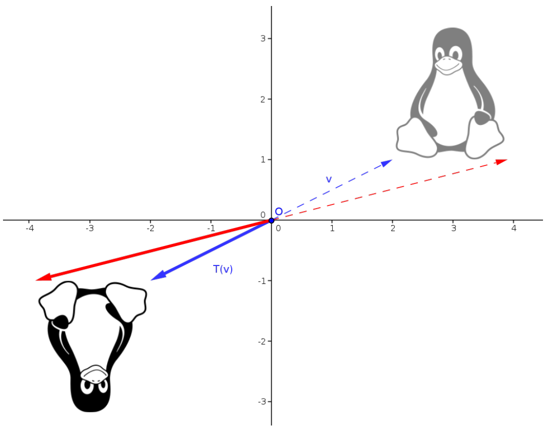
Autovalores e Autovetores

Considere a operador linear $S : \mathbb{R}^2 \to \mathbb{R}^2$ sobre $\mathbb{R}^2$. Podemos nos perguntar o seguinte: para quais vetores $v \in \mathbb{R}^2$, a imagem $S(v) \in \mathbb{R}^2$ tem a mesma direção de v? Isto é, queremos saber se para um dado $v \in \mathbb{R}^2$, existe $\lambda \in \mathbb{R}$ tal que

$$S(v) = \lambda v. \tag{6.1}$$

Exemplo 6.1. Considere a reflexão no plano ao redor da origem,

$$F(x, y) = (-x, -y).$$



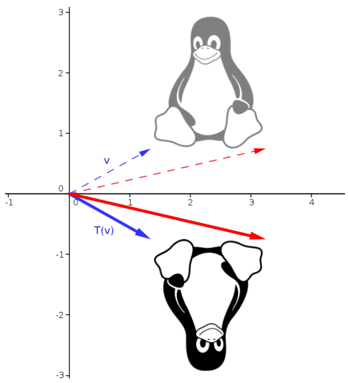
Como

$$F(x, y) = -1\,(x, y),$$

e logo (6.1) é satisfeita para todo (x, y) e $\lambda = -1$. □

Exemplo 6.2. Considere a reflexão em torno do eixo x, dada por

$$Q(x, y) = (x, -y).$$



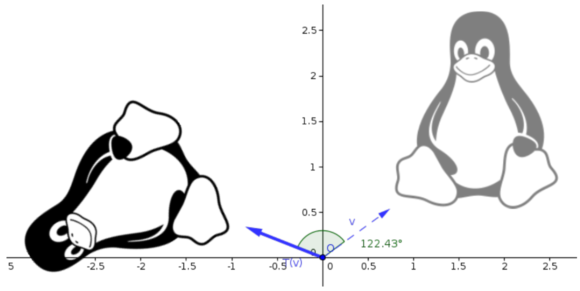
Os únicos vetores (x, y) que satisfazem (6.1) são da forma $(x, 0)$ ou $(0, y)$. Veja que

$$Q(x, 0) = 1\,(x, 0) \qquad \text{e} \qquad Q(0, y) = -1\,(0, y)$$

para $x, y \neq 0$, mas $Q(x, y) \neq \lambda(x, y)$ para qualquer $\lambda \in \mathbb{R}$ quando $y \neq 0$ e $x \neq 0$. Geometricamente, os únicos vetores que permanecem inalterados por reflexão ao redor do eixo x são aqueles na direção do eixo x, enquanto que os vetores na direção do eixo y mudam de sentido. □

Exemplo 6.3. Considere a rotação ao redor da origem por um ângulo θ no sentido anti-horário,

$$R(x, y) = (\ x\cos\theta - y\,\mathrm{sen}\,\theta\ ,\ y\cos\theta + x\,\mathrm{sen}\,\theta\ ).$$



- Se $\theta = 2k\pi$ para algum $k \in \mathbb{Z}$ (ângulo nulo + voltas completas), então para todo (x, y) temos
$$R(x, y) = (x, y).$$
- Se $\theta = \pi + 2k\pi$ para algum $k \in \mathbb{Z}$ (ângulo 180º + voltas completas), então para todo (x, y) temos
$$R(x, y) = -(x, y).$$
- Se $\theta \neq k\pi$ para todo $k \in \mathbb{Z}$ ($0 < \theta < \pi$ + meias voltas) então R não mantém a direção de vetor algum, isto é, após rotação por um ângulos desses, os vetores mudam de direção.

□

Atividade 6.1. Interprete geometricamente os três itens do exemplo anterior fazendo figuras.

Podemos levar esta ideia para um espaço vetorial qualquer: dada um operador linear $T : V \to V$, estamos interessados em determinar os pares (λ, v) para os quais

$$T(v) = \lambda v. \tag{6.2}$$

Neste caso, o escalar λ é chamado *autovalor* de T e v é chamado *autovetor* de T associado a λ. Podemos ainda dizer que (λ, v) é *autopar* de T.

Evidentemente, em qualquer caso, $v = \mathbf{0}$ satisfaz a equação (6.2) para qualquer λ, e logo o caso interessante é quando $v \neq \mathbf{0}$. Com isso, determinamos a noção de autovalor/autovetor de um operador qualquer.

Definição 6.1. *Seja $T : V \to V$ um operador linear sobre V. Se existirem $v \in V$, $v \neq \mathbf{0}$, e $\lambda \in \mathbb{R}$ tais que $T(v) = \lambda v$, então dizemos que λ é* ***autovalor*** *de T e v é* ***autovetor*** *de T associado a λ.*

Note que a definição exige que
$$T(v) = \lambda v \in V.$$
Logo, a imagem $T(v)$ deve ser elemento de V. Por isso **só definimos autovalores/autovetores para transformações entre um mesmo espaço vetorial** V (operador sobre V).

Um fato é que se $v \neq \mathbf{0}$ é autovetor de T associado ao autovalor λ, então av, $a \neq 0$ qualquer, também é autovetor de T associado ao mesmo autovalor λ. De fato, sendo $T(v) = \lambda v$ temos

$$T(av) = aT(v) = a(\lambda v) = \lambda(av).$$

Em resumo:

Teorema 6.1. *Se $v \neq \mathbf{0}$ é autovetor de T associado ao autovalor λ, então av, $a \neq 0$, é autovetor de T associado a λ.*

Dado um operador $T : V \to V$, o Teorema 6.1 motiva considerarmos o conjunto dos autovetores v associados a um autovalor λ, definido como

$$V_\lambda = \{v \in V \mid T(v) = \lambda v\}.$$

Atividade 6.2. Mostre que V_λ é subespaço vetorial de V.

O subespaço V_λ de V é chamado de *autoespaço associado ao autovalor* λ.

Uma visualização geométrica interessante dos autoespaços aparece um Geometria Analítica, da seguinte forma: dada uma cônica no plano rotacionada ao redor da origem (por exemplo, uma elipse rotacionada), os autoespaços associados à matriz dos termos de ordem 2 da cônica são as retas nas direções dos eixos da cônica. Veja a seção 6.5 para detalhes. Analogamente, autoespaços descrevem os eixos de quádricas rotacionadas no espaço.

6.1 Matrizes

Vimos anteriormente que podemos associar matrizes à transformações lineares. É natural, portanto, associar autovalores/autovetores à matrizes. Pela afirmação acima, somente **matrizes quadradas** são consideradas. Podemos pensar na matriz quadrada $\mathbf{A}$ $n \times n$ **como a matriz de** $T : \mathbb{R}^n \to \mathbb{R}^n$ **nas bases canônicas**. Logo, transitamos entre transformações e matrizes de forma natural:

$$[T]^{can}_{can} = \mathbf{A}$$

$$\begin{array}{rcl} T : \mathbb{R}^n \to \mathbb{R}^n & \longleftrightarrow & \mathbf{A} \text{ matriz } n \times n \\ v \in \mathbb{R}^n & \longleftrightarrow & \mathbf{v} \text{ matriz coluna } n \times 1 \\ T(v) = \lambda v & \longleftrightarrow & \mathbf{A}\mathbf{v} = \lambda \mathbf{v} \end{array}$$

Podemos então falar em autovalores/autovetores de matrizes quadradas:

$\mathbf{v} \neq \mathbf{0}$ é *autovetor* da matriz quadrada $\mathbf{A}$ associado ao *autovalor* λ se

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}.$$

Com isso, calculamos autovalores/autovetores de um operador olhando para sua matriz nas bases canônicas.

Exemplo 6.4. As matrizes nas bases canônicas associadas aos operadores lineares dos Exemplos 6.1 e 6.2 são, respectivamente,

$$\mathbf{B} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \text{e} \qquad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Temos

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -1 \begin{bmatrix} x \\ y \end{bmatrix}$$

e

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = 1 \begin{bmatrix} x \\ 0 \end{bmatrix} \qquad , \qquad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ y \end{bmatrix} = -1 \begin{bmatrix} 0 \\ y \end{bmatrix}.$$

□

Exemplo 6.5. Considere a matriz

$$\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix},$$

Seus autopares podem ser calculados resolvendo

$$\begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix},$$

isto é,

$$\begin{cases} 2x & +2y & = & \lambda x \\ & y & = & \lambda y \end{cases}.$$

- **Consideremos o caso em que $y \neq 0$.** Da segunda equação temos $\lambda = 1$, e logo a primeira equação torna-se $2x + 2y = x$, donde segue que $y = -\frac{1}{2}x$. Portanto

  $$\mathbf{v} = \begin{bmatrix} x \\ -\frac{1}{2}x \end{bmatrix}, \; x \neq 0, \quad \text{são autovetores associados ao autovalor} \quad \lambda_1 = 1.$$

  Você pode verificar que $\mathbf{A}\mathbf{v} = \mathbf{v}$.

- **Consideremos o caso em que $y = 0$.** Neste caso o sistema se reduz à equação $2x = \lambda x$. Para $x \neq 0$, devemos ter $\lambda = 2$, ou seja,

  $$\mathbf{u} = \begin{bmatrix} x \\ 0 \end{bmatrix}, \; x \neq 0, \quad \text{são autovetores associados ao autovalor} \quad \lambda_2 = 2.$$

  Você pode verificar que $\mathbf{A}\mathbf{u} = 2\mathbf{u}$.

□

6.2 Cálculo de autovalores – o polinômio característico

Como dissemos anteriormente, matrizes serão a ferramenta para o cálculo de autovalores. Veremos nesta seção uma forma sistemática de cálculo de autopares de matrizes (e consequentemente de operadores lineares).

Note que no Exemplo 6.5, calculamos os autopares $(\lambda, \mathbf{v})$ de uma matriz $\mathbf{A}$ resolvendo

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

para λ e $\mathbf{v}$ simultaneamente. Lembre que tivemos que separar em dois casos e fazer uma análise cuidadosa. Porém, se **soubéssemos antes quais eram os autovalores** o termo, recairíamos em um **sistema linear homogêneo** na variável $\mathbf{v}$, já que

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \quad \Leftrightarrow \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}. \tag{6.3}$$

Resolver $(\mathbf{A}-\lambda\mathbf{I})\mathbf{v} = \mathbf{0}$ pode ser feito por escalonamento. Então traçamos a seguinte estratégia:

1. Encontrar todos os autovalores $\lambda_1, \lambda_2, \ldots, \lambda_r$;
2. Para cada autovalor λ_i, resolver o sistema $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{v} = \mathbf{0}$ (ou $\mathbf{A}\mathbf{v} = \lambda_i\mathbf{v}$) para encontrar os autovetores associados.

Nossa primeira tarefa, portanto, é calcular os autovalores.

O sistema

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$$

tem matriz quadrada $\mathbf{A}-\lambda\mathbf{I}$, e portanto podemos calcular seu determinante. Se $\det(\mathbf{A}-\lambda\mathbf{I}) \neq 0$ então o sistema homogêneo acima possui única solução, que só pode ser $\mathbf{v} = \mathbf{0}$ ($\mathbf{0}$ sempre é solução). Porém, lembre-se que por definição, $\mathbf{v} = \mathbf{0}$ não pode ser autovetor. Logo, $\lambda \in \mathbb{R}$ só pode ser autovalor de $\mathbf{A}$ se

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

Deste modo, o sistema $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$ terá solução não trivial $\mathbf{v} \neq \mathbf{0}$, que será autovetor de $\mathbf{A}$ por (6.3). Ou seja, os autovalores são raízes do polinômio

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}).$$

$p(\lambda)$ é chamado *polinômio característico.*

Note que $p(\lambda)$ não envolve $\mathbf{v}$. Ou seja, ao resolver $p(\lambda) = 0$ estaremos calculando somente os autovalores de $\mathbf{A}$. Os autovetores são calculados em seguida, resolvendo o sistema linear associado, como no roteiro descrito anteriormente.

Abrindo a expressão de $p(\lambda)$, temos

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & c_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & c_{nn} - \lambda \end{bmatrix}.$$

Veja que de cada termo da diagonal de $\mathbf{A}$ é subtraído λ. O resultado dessa determinante é um polinômio na variável λ, o que explica seu nome. O roteiro para o cálculo de autovalores/autovetores torna-se:

1. Encontrar todos os autovalores $\lambda_1, \lambda_2, \ldots, \lambda_r$, calculando as raízes do polinômio característico
$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$$
2. Para cada autovalor λ_i, resolver o sistema $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{v} = \mathbf{0}$ (ou $\mathbf{A}\mathbf{v} = \lambda_i\mathbf{v}$) para encontrar os autovetores associados.

Exemplo 6.6. Vamos calcular os autovalores da matriz do Exemplo 6.5
$$\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}.$$
O polinômio característico associado é
$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \det\begin{bmatrix} 2-\lambda & 2 \\ 0 & 1-\lambda \end{bmatrix} = (2-\lambda)(1-\lambda) = \lambda^2 - 3\lambda + 2.$$
As raízes de p são $\lambda_1 = 1$ e $\lambda_2 = 2$, os mesmos autovalores encontrados no Exemplo 6.5. □

Exemplo 6.7. Vamos calcular os autovalores/autovetores da matriz
$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$
O polinômio característico é
$$p(\lambda) = \det(\mathbf{B} - \lambda\mathbf{I}_3) = \det\begin{bmatrix} 1-\lambda & 0 & 2 \\ -1 & -\lambda & 1 \\ 1 & 1 & 2-\lambda \end{bmatrix}.$$
Usando desenvolvimento de Laplace sobre a primeira linha, obtemos
$$\begin{aligned} p(\lambda) &= (-1)^{1+1}(1-\lambda)\det\begin{bmatrix} -\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} + (-1)^{1+3}\,2\det\begin{bmatrix} -1 & -\lambda \\ 1 & 1 \end{bmatrix} \\ &= (1-\lambda)\big[-\lambda(2-\lambda) - 1\big] + 2(-1+\lambda) \\ &= (1-\lambda)\big[\lambda^2 - 2\lambda - 3\big]. \end{aligned}$$
Temos $p(\lambda) = 0$ se, e somente se, $\lambda = 1$ ou $\lambda^2 - 2\lambda - 3 = 0$. Assim, os autovalores de $\mathbf{B}$ são
$$\lambda_1 = 1, \quad \lambda_2 = -1, \quad \lambda_3 = 3.$$

Autovetores associados a $\lambda_1 = 1$. Devemos resolver $\mathbf{B}\mathbf{v} = \mathbf{v}$ (onde $\mathbf{v} = \begin{bmatrix} x & y & z \end{bmatrix}^t$), ou seja, resolver o sistema
$$\left\{\begin{array}{rrrcl} x & & +2z & = & x \\ -x & & +z & = & y \\ x & +y & +2z & = & z \end{array}\right. \quad\Leftrightarrow\quad \left\{\begin{array}{rrrcl} & & 2z & = & 0 \\ -x & -y & +z & = & 0 \\ x & +y & +z & = & 0 \end{array}\right.$$
(o último sistema é $(\mathbf{B} - \mathbf{I}_3)\mathbf{v} = \mathbf{0}$, você pode escrevê-lo diretamente). O último sistema tem soluções não triviais
$$\begin{bmatrix} x & -x & 0 \end{bmatrix}^t, \quad x \neq 0$$
(verifique!), que são os autovetores associados à $\lambda_1 = 1$.

Autovetores associados a $\lambda_2 = -1$. Devemos resolver $\mathbf{B}\mathbf{v} = -\mathbf{v}$, ou equivalentemente, $(\mathbf{B} + \mathbf{I}_3)\mathbf{v} = \mathbf{0}$. Isto é,

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

Escalonando a matriz ampliada:

$$\left[\begin{array}{rrr|r} 2 & 0 & 2 & 0 \\ -1 & 1 & 1 & 0 \\ 1 & 1 & 3 & 0 \end{array}\right] \xrightarrow{L_1 \to \frac{1}{2}L_1} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ \boxed{-1} & 1 & 1 & 0 \\ \boxed{1} & 1 & 3 & 0 \end{array}\right] \xrightarrow[L_3 \to L_3 - L_1]{L_2 \to L_2 + L_1} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & \boxed{1} & 2 & 0 \end{array}\right]$$

$$\xrightarrow[L_3 \to L_3 - L_1]{L_3 \to L_3 - L_2} \left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Resolvendo o sistema associado à MLRFE (última matriz), concluímos que os autovetores de $\mathbf{B}$ associados à $\lambda_2 = -1$ são

$$\left[\begin{array}{ccc} -z & -2z & z \end{array}\right]^t, \quad z \neq 0.$$

Autovetores associados a $\lambda_3 = 3$. Devemos resolver $(\mathbf{B} - 3\,\mathbf{I}_3)\mathbf{v} = \mathbf{0}$, ou seja,

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

Você pode resolver este sistema e verificar que suas soluções não triviais são

$$\left[\begin{array}{ccc} z & 0 & z \end{array}\right]^t, \quad z \neq 0,$$

que são os autovetores de $\mathbf{B}$ associados a $\lambda_3 = 3$. □

No exemplo acima, obtemos todos os autovalores e autovetores da matriz $\mathbf{B}$. Note que, em particular,

- $\mathbf{v}_1 = \left[\begin{array}{ccc} 1 & -1 & 0 \end{array}\right]^t$ é autovetor associado à $\lambda_1 = 1$;
- $\mathbf{v}_2 = \left[\begin{array}{ccc} -1 & -2 & 1 \end{array}\right]^t$ é autovetor associado à $\lambda_2 = -1$;
- $\mathbf{v}_3 = \left[\begin{array}{ccc} 1 & 0 & 1 \end{array}\right]^t$ é autovetor associado à $\lambda_3 = 3$.

Observe que **todos** os autovetores são descritos pelos autoespaços V_λ associados a cada autovalor, e que cada um desses autoespaços é gerado pelos autovetores particulares acima (visto como vetores do $\mathbb{R}^3$):

$$V_1 = [\,v_1\,], \qquad V_{-1} = [\,v_2\,], \qquad V_3 = [\,v_3\,].$$

Atividade 6.3. Verifique que os três vetores acima v_1, v_2 e v_3 são LI. Assim, formam uma base do $\mathbb{R}^3$.

Veremos no próximo capítulo que o fato de v_1, v_2 e v_3 serem LI não é a toa: autovetores associados à autovalores distintos (é o caso do exemplo anterior) são sempre LI. Isso será importante na diagonalização de matrizes/operadores, quando buscaremos uma base do espaço vetorial do operador formada por autovetores seus.

É possível que o polinômio característico possua raízes complexas. Há teoria para lidar com esse tipo de situação, mas não faremos neste texto. Ou seja, vamos considerar apenas **autovalores reais**.

6.2.1 Casos particulares

Nesta seção apresentamos alguns casos frequentes em que o cálculo dos autovalores (ou parte deles) é direto.

Autovalores de matrizes não inversíveis. Se uma matriz $\mathbf{A}$ não possui inversa ($\det \mathbf{A} = 0$) então $\lambda = 0$ sempre será **um dos seus autovalores** (os outros devem ser calculados em cada caso!). De fato, quando $\mathbf{A}$ não é inversível, o sistema quadrado homogêneo

$$\mathbf{A}\mathbf{v} = \mathbf{0}$$

possui solução $\mathbf{v} \neq \mathbf{0}$. Logo, este $\mathbf{v}$ é autovetor associado ao autovalor nulo pois $\mathbf{A}\mathbf{v} = \mathbf{0} = 0\mathbf{v}$.

Autovalores de matrizes triangulares. Seja

$$\mathbf{T} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & \cdots & t_{1n} \\ 0 & t_{22} & t_{23} & \dots & t_{2n} \\ 0 & 0 & t_{33} & \dots & t_{3n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \cdots & t_{nn} \end{bmatrix}$$

uma matriz triangular superior de ordem n. Seu polinômio característico é

$$p(\lambda) = \det(\mathbf{T} - \lambda \mathbf{I}_n) = \det \begin{bmatrix} t_{11} - \lambda & t_{12} & t_{13} & \cdots & t_{1n} \\ 0 & t_{22} - \lambda & t_{23} & \dots & t_{2n} \\ 0 & 0 & t_{33} - \lambda & \dots & t_{3n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \cdots & t_{nn} - \lambda \end{bmatrix}.$$

Temos então um determinante de uma matriz triangular, que sabemos ser o produto dos elementos da sua diagonal (veja o capítulo sobre determinantes). Assim

$$p(\lambda) = (t_{11} - \lambda)(t_{22} - \lambda) \cdots (t_{nn} - \lambda),$$

cujas raízes são, claramente, t_{11}, $t_{22}, \dots, t_{nn}$.

O mesmo ocorre vale para matrizes triangulares inferiores: basta notar que no polinômio característico só aparecem os termos da diagonal multiplicados.

Exemplo 6.8. Os autovalores de $\mathbf{T} = \begin{bmatrix} 1 & 56 & 23 & 11 \\ 0 & -2 & \frac{3}{5} & 2 \\ 0 & 0 & 3 & -7 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ são 1, -2, 3 e -1. □

Autovalores de matrizes diagonais. Uma matriz diagonal é, em particular, uma matriz triangular. Portanto seus autovalores são os elementos da diagonal. Em particular, a matriz identidade de ordem n tem único autovalor $\lambda = 1$.

6.3 Autovalores de operadores via matrizes

Definimos em primeiro lugar autovalores/autovetores para operadores $T : V \to V$ sobre um espaço vetorial V qualquer. Vimos que a mesma definição vale para matrizes quadradas, e como podemos calcular seus autopares. Vamos agora conectar o cálculo de autovalores de matrizes com os operadores sobre V qualquer.

Seja β uma base de V e $T : V \to V$ um operador linear. Temos

$$T(v) = \lambda v \quad \Leftrightarrow \quad [\,T\,]_\beta^\beta [\,v\,]_\beta = \lambda [\,v\,]_\beta \quad \Leftrightarrow \quad \Big([\,T\,]_\beta^\beta - \lambda \mathbf{I}\Big)[\,v\,]_\beta = \mathbf{0} \quad \Leftrightarrow \quad \det([\,T\,]_\beta^\beta - \lambda \mathbf{I}) = 0.$$

Assim, calcular os autovalores de T é equivalente a encontrar as raízes do polinômio característico da matriz de T na base β, $[\,T\,]_\beta^\beta$.

Seja agora α outra base de V. Lembrando que $[\,I\,]_\alpha^\beta = \big([\,I\,]_\beta^\alpha\big)^{-1}$, veja que

$$\begin{aligned}
\det([\,T\,]_\alpha^\alpha - \lambda \mathbf{I}) &= \det([\,I\,]_\alpha^\beta [\,T\,]_\beta^\beta [\,I\,]_\beta^\alpha - \lambda [\,I\,]_\alpha^\beta [\,I\,]_\beta^\alpha) \\
&= \det\Big([\,I\,]_\alpha^\beta \left([\,T\,]_\beta^\beta - \lambda \mathbf{I}\right) [\,I\,]_\beta^\alpha\Big) \\
&= \det[\,I\,]_\alpha^\beta \det([\,T\,]_\beta^\beta - \lambda \mathbf{I}) \det[\,I\,]_\beta^\alpha \\
&= \underbrace{\det[\,I\,]_\alpha^\beta \det[\,I\,]_\beta^\alpha}_{=1} \det([\,T\,]_\beta^\beta - \lambda \mathbf{I}) = \det([\,T\,]_\beta^\beta - \lambda \mathbf{I}).
\end{aligned}$$

O que acabamos de provar é que

> O polinômio característico de $[\,T\,]_\beta^\beta$ é o mesmo para qualquer para qualquer base β de V. Assim, a fim de calcular os autovalores de T, podemos escolher qualquer base β e encontrar as raízes do polinômio característico
>
> $$p(\lambda) = \det\left([\,T\,]_\beta^\beta - \lambda \mathbf{I}\right).$$

Exemplo 6.9. Considere o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido por

$$T(x, y, z) = (\ x\ ,\ y\ ,\ -x + 2z\ ).$$

Escolhamos a base canônica do $\mathbb{R}^3$ para o cálculo dos autovalores,

$$[\,T\,]_{can}^{can} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 2 \end{bmatrix}.$$

Calculando o polinômio característico,

$$p(\lambda) = \det \begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 1/2 & 0 & 1/2-\lambda \end{bmatrix} = (1-\lambda)^2(2-\lambda) = 0 \quad \Rightarrow \quad \lambda_1 = 1, \quad \lambda_2 = 2.$$

Consideramos agora a matriz de T na base $\beta = \{(1,0,1); (0,2,0); (0,0,1)\}$. Calculando-a, obtemos

$$[\,T\,]_\beta^\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

uma matriz diagonal! Logo, o cálculo do polinômio característico é direto:

$$p(\lambda) = \det \begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{bmatrix} = (1-\lambda)^2(2-\lambda).$$

□

6.4 Multiplicidades algébrica e geométrica

Dois conceitos interessantes acerca de **autovalores** são as multiplicidades *algébrica* e *geométrica.*

- A **multiplicidade algébrica** de um autovalor λ é quantidade de vezes que ele aparece como raiz do polinômio característico;
- A **multiplicidade geométrica** de um autovalor λ é a dimensão do autoespaço associado V_λ.

O nome "algébrico" justifica-se pela relação/manipulação/operação com o polinômio característico, elementos próprios do que se entende por "álgebra". Já o nome "geométrico" justifica-se por fazer referência à forma do subespaço vetorial V_λ (dimensão 0, ponto; dimensão 1 reta; dimensão 2, plano etc).

Exemplo 6.10. No Exemplo 6.9, a multiplicidade algébrica do autovalor $\lambda_1 = 1$ é 2, e a de $\lambda_2 = 2$ é 1, visto que

$$p(\lambda) = (1-\lambda)^2(2-\lambda) = (1-\lambda)(1-\lambda)(2-\lambda).$$

Vamos calcular os autovetores do operador T daquele exemplo, cuja matriz na base

$$\beta = \{(1,0,1); (0,2,0); (0,0,1)\}$$

é

$$[\,T\,]_\beta^\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Autovetores associados a $\lambda_1 = 1$. Devemos resolver

$$\big([\,T\,]_\beta^\beta - \mathbf{I}_3\big)[\,v\,]_\beta = \mathbf{0}.$$

Como $[\,T\,]_\beta^\beta$ é diagonal, é fácil ver que

$$[\,v\,]_\beta = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}.$$

Expandindo v na base β, vemos que

$$v = x(1,0,1) + y(0,1,0) + 0(0,0,2) = (x,y,x) \neq \mathbf{0}$$

são os autovetores associados a 1. Veja que conseguimos no máximo dois vetores desse tipo que são LI entre si, por exemplo,

$$v_1 = (1,0,1) \qquad \text{e} \qquad v_2 = (0,1,0).$$

Logo $\{v_1, v_2\}$ é uma base para o autoespaço V_1, ou seja, λ_1 tem multiplicidade geométrica igual a 2 (V_1 é um plano).

Autovetores associados a $\lambda_2 = 2$. Resolvendo

$$\big([\,T\,]_\beta^\beta - 2\,\mathbf{I}_3\big)[\,v\,]_\beta = \mathbf{0}$$

obtemos

$$[\,v\,]_\beta = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$$

(verifique!). Expandindo v na base β, vemos que

$$v = 0(1,0,1) + 0(0,1,0) + z(0,0,2) = (0,0,2z) \neq \mathbf{0}$$

são os autovetores associados a 2. Conseguimos apenas 1 vetor desse tipo que forma um conjunto LI, por exemplo, $\{(0,0,2)\}$. Esse conjunto é uma base do autoespaço V_2, e portanto $\lambda_2 = 2$ tem multiplicidade geométrica igual a 1. $\square$

É possível mostrar que para qualquer autovalor,

$$\text{multiplicidade geométrica} \quad \leq \quad \text{multiplicidade algébrica.}$$

Há casos em que a multiplicidade geométrica é estritamente menor que a algébrica.

Atividade 6.4. Calcule as multiplicidades algébrica e geométrica do autovalor $\lambda = 1$ da matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

e verifique que a multiplicidade geométrica é menor que a algébrica.

A multiplicidade geométrica tem utilidade na diagonalização de operadores, que veremos no próximo capítulo. Diagonalizar um operador significa obter uma base para a qual a matriz do operador é diagonal, tal como ilustrado no Exemplo 6.9. O segredo é obter uma *base de autovetores*, que existirá se a soma das multiplicidades geométricas for igual à dimensão do espaço (neste caso, decompomos o espaço vetorial na soma direta dos autoespaços associados aos autovalores distintos: $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r}$).

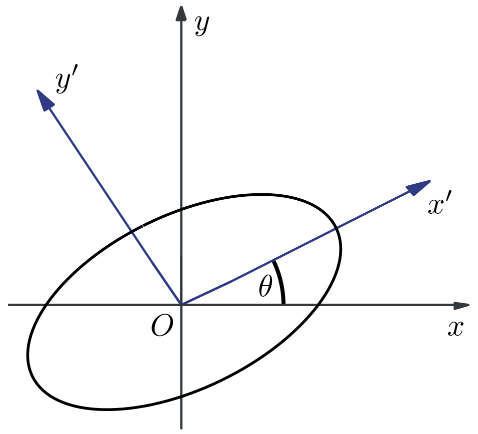
6.5 Autopares e cônicas rotacionadas

Nesta seção ilustramos o uso de autovalores/autovetores para identificação de cônicas rotacionadas no plano.

Seja dada a equação

$$ax^2 + bxy + cy^2 + f = 0.$$

Essa equação representa uma cônica rotacionada se $b \neq 0$.



No caso em que $b \neq 0$, veremos como encontrar um ângulo θ tal que, rotacionando eixos x e y, eliminemos o termo xy na equação, escrevendo-a no novo sistema $x'Oy'$ como

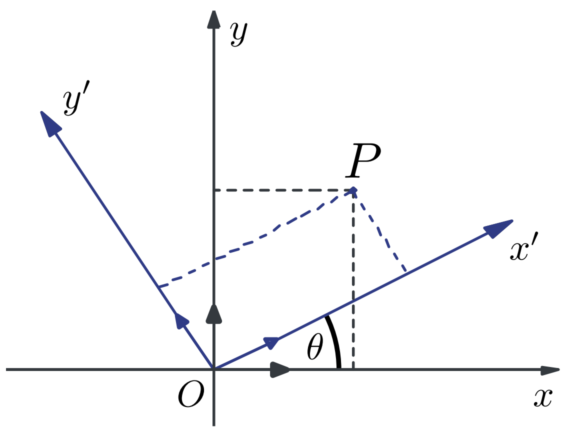
$$a'x'^2 + c'y'^2 + f' = 0 \quad (b' = 0).$$

A escrita acima é mais simples de interpretar, podemos facilmente capturar sua geometria (centro, focos, semi-eixos etc), como se faz em cursos de Geometria Analítica.

Vimos no início do capítulo sobre transformações lineares que rotacionar (x, y) no sentido anti-horário por um ângulo θ corresponde à multiplicação pela *matriz de rotação*

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

O que queremos é uma relação entre as coordenadas (x, y) de um ponto P do plano no sistema original xOy e suas coordenadas (x', y') no novo sistema rotacionado $x'Oy'$.



A figura acima ilustra que as coordenadas de P no sistema original xOy correspondem às coordenadas rotação de P no novo sistema $x'Oy'$ **após uma rotação** pelo ângulo θ (imagine a rotação de P e veja que as linhas pontilhadas azuis (novo sistema) terão os mesmos comprimentos das linhas pontilhadas pretas (sistema original)). Ou seja,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

Chamando $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}$ e $\mathbf{X}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ escrevemos

$$\mathbf{X} = \mathbf{R}_\theta \mathbf{X}'.$$

Voltemos a analisar a equação

$$ax^2 + bxy + cy^2 + f = 0. \tag{6.4}$$

Fazendo $\mathbf{A} = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$, podemos escrever (6.4) como

$$\mathbf{X}^t\mathbf{A}\mathbf{X} + f = 0.$$

Atividade 6.5. Verifique que a $\mathbf{X}^t\mathbf{A}\mathbf{X} + f = 0$ é a equação (6.4).

A fim de eliminar o termo xy, fazemos a rotação nos eixos $\mathbf{X} = \mathbf{R}_\theta\mathbf{X}'$. Lembrando que $(\mathbf{R}_\theta\mathbf{X}')^t = \mathbf{X}'^t\mathbf{R}_\theta^t$, temos

$$(\mathbf{R}_\theta\mathbf{X}')^t\mathbf{A}(\mathbf{R}_\theta\mathbf{X}') + f = 0 \quad \Rightarrow \quad \mathbf{X}'^t(\mathbf{R}_\theta^t\mathbf{A}\mathbf{R}_\theta)\mathbf{X}' + f = 0.$$

Chamando $\mathbf{B} = \mathbf{R}_\theta^t\mathbf{A}\mathbf{R}_\theta$ temos

$$\mathbf{X}'^t\mathbf{B}\mathbf{X}' + f = 0.$$

Por analogia, essa é exatamente a equação

$$a'x'^2 + b'x'y' + c'y'^2 + f = 0,$$

onde $\mathbf{B} = \begin{bmatrix} a' & b'/2 \\ b'/2 & c' \end{bmatrix} = \mathbf{R}_\theta^t\mathbf{A}\mathbf{R}_\theta$.

Portanto, nosso objetivo é encontrar θ tal que $b' = 0$.

Antes disso, é necessário calcular a' e c'. Precisamos do resultado seguinte:

Atividade 6.6. Mostre por um cálculo direto que $\mathbf{R}_\theta^t\mathbf{R}_\theta = \mathbf{R}_\theta\mathbf{R}_\theta^t = \mathbf{I}_2$. Mostre também que $\det\mathbf{R}_\theta = 1$.

Usando a atividade acima, para qualquer $\lambda \in \mathbb{R}$ temos

$$\begin{aligned} \det(\mathbf{B} - \lambda\mathbf{I}_2) &= \det(\mathbf{R}_\theta^t\mathbf{A}\mathbf{R}_\theta - \lambda\mathbf{R}_\theta^t\mathbf{R}_\theta) \\ &= \det(\mathbf{R}_\theta^t(\mathbf{A} - \lambda\mathbf{I}_2)\mathbf{R}_\theta) \\ &= \underbrace{\det\mathbf{R}_\theta^t}_{1}\det(\mathbf{A} - \lambda\mathbf{I}_2)\underbrace{\det\mathbf{R}_\theta}_{1} = \det(\mathbf{A} - \lambda\mathbf{I}_2). \end{aligned}$$

Logo, se $b' = 0$ então

$$\det(\mathbf{A} - \lambda\mathbf{I}_2) = \det(\mathbf{B} - \lambda\mathbf{I}_2) = \det\begin{bmatrix} a' - \lambda & 0 \\ 0 & c' - \lambda \end{bmatrix} = (a' - \lambda)(c' - \lambda).$$

Daí, a' e c' são raízes do polinômio característico de $\mathbf{A}$,

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}_2) = \det\begin{bmatrix} a - \lambda & b/2 \\ b/2 & c - \lambda \end{bmatrix}.$$

Ou seja, a' e c' são os **autovalores da matriz** $\mathbf{A} = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$.

Em resumo, passamos da equação original

$$ax^2 + bxy + cy^2 + f = 0$$

para a equação sem o termo misto xy,

$$a'x'^2 + c'y'^2 + f = 0,$$

calculando a' e c' como os autovalores da matriz

$$\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}.$$

Note que o termo livre f é o mesmo nas duas equações.

Finalmente, você pode verificar que os autovetores associados à a' e c' (os autovalores de $\mathbf{A}$) têm as direções dos eixos rotacionados x' e y'. Reflita sobre as figuras desta seção.

Atividade 6.7. Identifique cada cônica, eliminando o termo misto xy.

(*i*) $2x^2 + 5xy + 2y^2 = -1$

(*ii*) $x^2 + xy + y^2 - 1 = 0$

(*iii*) $5x^2 - 4xy + 8y^2 - 36 = 0$

(*iv*) $xy = 1$

Atividade 6.8. Mostre que $\mathbf{R}_\theta \mathbf{R}_\gamma = \mathbf{R}_{\theta+\gamma}$, isto é, rotacionar eixos por um ângulo θ e daí rotacioná-los por um ângulo γ é o mesmo que rotacionar os eixos por um ângulo $\theta + \gamma$.

Atividade 6.9. Mostre que sempre é possível eliminar o termo misto xy, isto é, mostre que existe um ângulo θ tal que $b' = 0$.

Solução: Observe que

$$\begin{aligned}
\mathbf{B} = \begin{bmatrix} a' & b'/2 \\ b'/2 & c' \end{bmatrix} = \mathbf{R}_\theta^t \mathbf{A} \mathbf{R}_\theta &= \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \\
\Rightarrow \frac{b'}{2} &= -a \operatorname{sen}\theta\cos\theta - \frac{b}{2}\operatorname{sen}^2\theta + \frac{b}{2}\cos^2\theta + c\operatorname{sen}\theta\cos\theta \\
\Rightarrow b' &= (c-a)\operatorname{sen}(2\theta) + b\cos(2\theta).
\end{aligned}$$

Assim, $b' = 0$ se, e somente se $(a - c)\operatorname{sen}(2\theta) = b\cos(2\theta)$. Logo, se $a - c = 0$, basta fazer $\theta = 45^o$ $(\theta = \pi/4)$ (pois com isso $\cos(2\theta) = 0$). Por outro lado, se $a - c \neq 0$, $b' = 0$ se

$$\operatorname{tg}(2\theta) = \frac{\operatorname{sen}(2\theta)}{\cos(2\theta)} = \frac{b}{a-c}.$$

Basta tomar então $\theta \in (-\pi/4, \pi/4)$ tal que a relação anterior valha. Isto mostra que sempre conseguimos θ tal que $b' = 0$.

6.6 Exercícios

1. Seja $T : V \to V$ um operador linear. Mostre que:

   (a) Se v é um autovetor de T associado ao autovalor α então v é autovetor de T^n associado ao autovalor α^n, para qualquer $n \geq 1$.

   (b) T é isomorfismo se, e somente se, $\mathbf{0}$ não é autovalor de T.

   (c) Se T é isomorfismo então α é autovalor de T se, e somente se $\frac{1}{\alpha}$ é autovalor de T^{-1}.

2. Sejam $T, S : V \to V$ operadores lineares tais que $T \circ S = S \circ T$. Mostre que se v é autovetor de T associado ao autovalor λ e $S(v) \neq \mathbf{0}$ então $S(v)$ é autovetor de T associado a λ.

3. Determine os autovalores e autovetores de cada matriz abaixo.

   (a) $\mathbf{A} = \begin{bmatrix} 1 & 3 \\ -2 & 6 \end{bmatrix}$

   (b) $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

   (c) $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$

   (d) $\mathbf{A} = \begin{bmatrix} 3 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

   (Dica: $\det \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} = \det \mathbf{B} \cdot \det \mathbf{C}$, onde $\mathbf{B}$ e $\mathbf{C}$ são matrizes quadradas)

4. Se $\mathbf{A}$, $\mathbf{B}$ são matrizes de ordem n, mostre que $\mathbf{AB}$ e $\mathbf{BA}$ têm os mesmos autovalores.

5. Mostre que as matrizes quadradas $\mathbf{A}$ e $\mathbf{A}^t$ têm os mesmos autovalores.

6. Duas matrizes $\mathbf{A}$ e $\mathbf{B}$ quadradas de ordem n são ditas *semelhantes* se existe uma matriz inversível $\mathbf{P}$ tal que $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$. Mostre que se $\mathbf{A}$ é semelhante a uma matriz diagonal $\mathbf{D}$, digamos $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$, então a diagonal de $\mathbf{D}$ é formada pelos autovalores de $\mathbf{A}$.

   Mostre também que as colunas de $\mathbf{P}$ são autovetores de $\mathbf{A}$, colocados na ordem em que os autovalores em $\mathbf{D}$ aparecem.

7. Seja $\mathbf{A}$ uma matriz quadrada de ordem n. Mostre que se 1 não é autovalor de $\mathbf{A}$, então $\mathbf{A} - \mathbf{I}$ é inversível.

Capítulo 7

Diagonalização

7.1 Diagonalização de operadores

Dado um operador linear $T : V \to V$, nosso objetivo é encontrar, se existir, uma base β de V tal que a matriz de T nessa base, $[\,T\,]_\beta^\beta$, seja diagonal. É a isso que nos referimos por *diagonalizar um operador*, ou ainda, dizer que um operador é *diagonalizável*. É evidente que matrizes diagonais são mais simples de lidar, por exemplo, quando queremos aplicar o operador a um vetor: $[\,T(v)\,]_\beta = [\,T\,]_\beta^\beta\,[\,v\,]_\beta$.

Antes de apresentar a teoria, vamos retomar um exemplo do capítulo anterior.

Considere o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido por

$$T(x, y, z) = (\ x\ ,\ y\ ,\ -x + 2z\ ).$$

Escolhendo a base

$$\beta = \{(1, 0, 1); (0, 2, 0); (0, 0, 1)\}$$

de $\mathbb{R}^3$ e fazendo as contas chegamos a

$$[\,T\,]_\beta^\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Note que $[\,T\,]_\beta^\beta$ é diagonal, e que cada vetor da base β é autovetor de T. De fato, você pode verificar que

$$T(1, 0, 1) = 1(1, 0, 1), \qquad T(0, 2, 0) = 1(0, 2, 0), \qquad T(0, 0, 1) = 2(0, 0, 1).$$

Este exemplo indica que a procura pela base β passa por estudar os autovetores de T.

7.1.1 Bases formadas por autovetores

Dado um operador linear $T : V \to V$ suponha que

$$\beta = \{v_1, \dots, v_n\}$$

seja uma base de V para a qual

$$[\,T\,]_\beta^\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Sabemos que a cada coluna i dessa matriz é exatamente $[\,T(v_i)\,]_\beta$, e assim

$$T(v_i) = 0v_1 + \cdots + \lambda_i v_i + \cdots + 0v_n \quad \Rightarrow \quad T(v_i) = \lambda_i v_i$$

para todo i. Ou seja,

Se $[\,T\,]_\beta^\beta$ é diagonal, então a base β é formada por autovetores de T.

Po outro lado, se

$$\beta = \{v_1, \ldots, v_n\}$$

é base V formada por autovetores de T, digamos $T(v_i) = \lambda_i v_i$ para todo i (autovetor v_i associado ao autovalor λ_i), então

$$T(v_i) = \lambda_i v_i \quad \Rightarrow \quad T(v_i) = 0v_1 + \cdots + \lambda_i v_i + \cdots + 0v_n$$

para todo i, e logo

$$[\,T\,]_\beta^\beta = \begin{bmatrix} | & | & & | \\ [\,T(v_1)\,]_\beta & [\,T(v_2)\,]_\beta & \cdots & [\,T(v_n)\,]_\beta \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Ou seja,

Se β é base formada por autovetores de T, então $[\,T\,]_\beta^\beta$ é diagonal. Neste caso, **a diagonal de $[\,T\,]_\beta^\beta$ será formada pelos autovalores de T na ordem em que a base de autovetores é construída**, como mostra a expressão acima.

Isso nos leva à concluir que T admite uma base β que torna sua matriz diagonal se, e somente se, essa base β é formada por autovetores T. Isso motiva a definição de *operador diagonalizável*:

Definição 7.1. *Um operador $T : V \to V$ é **diagonalizável** se existe uma base de V formada por autovetores de T.*

Exemplo 7.1. O operador T sobre $\mathbb{R}^3$ do início do capítulo,

$$T(x, y, z) = (\ x\ ,\ y\ ,\ -x + 2z\ ).$$

é diagonalizável pois, como vimos, $\beta = \{(1, 0, 1); (0, 2, 0); (0, 0, 1)\}$ é uma base de $\mathbb{R}^3$ onde todos os elementos são autovetores de T. Neste caso, vimos que $[\,T\,]_\beta^\beta$ é diagonal. □

Exemplo 7.2. Considere o operador $S : \mathbb{R}^2 \to \mathbb{R}^2$ definido por

$$S(x, y) = (\ x + y\ ,\ y\ ).$$

O polinômio característico associado é

$$p(\lambda) = \det\big([\,S\,]_{can}^{can} - \lambda \mathbf{I}_2\big) = \det \begin{bmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{bmatrix} = (1 - \lambda)^2,$$

cuja única raíz (autovalor) é $\lambda = 1$. Resolvendo o sistema $([\,S\,]_{can}^{can} - \mathbf{I}_2)[\,v\,]_{can} = \mathbf{0}$, obtemos

$$[\,v\,]_{can} = \begin{bmatrix} x \\ 0 \end{bmatrix}$$

(verifique!). Assim, conseguimos apenas um autovetor LI, por exemplo, $v = (1, 0)$. Logo, **não é possível** construir uma base de $\mathbb{R}^2$ formada por autovetores de S. Assim, S **não** é diagonalizável. Você pode tentar conseguir uma base para a qual a matriz de S seja diagonal, e não terá sucesso! □

O exemplo anterior mostra que **nem sempre** é possível diagonalizar um operador. **Então, como saber se um operador é diagonalizável?**

Passamos agora a estudar quando conseguimos diagonalizar um operador. Ou seja, **quando conseguimos autovetores LI suficientes para formar uma base**.

Teorema 7.1. *Autovetores associados a autovalores distintos são LI.*

Mais especificamente se $\lambda_1, \ldots, \lambda_r$ são autovalores de T, distintos entre si, e $v_1, \ldots, v_r$ são autovetores correspondentes então $\{v_1, \ldots, v_r\}$ é um conjunto LI.

Imediatamente, se o operador $T : V \to V$ possui n autovalores distintos ($n = \dim V$) então existem n autovetores LI, e logo formam uma base de V.

Corolário 7.1. *Se V é um espaço vetorial de dimensão n e o operador linear $T : V \to V$ possui n autovalores distintos, então T é diagonalizável.*

Exemplo 7.3. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ o operador linear definido por

$$T(x, y, z) = (\ x - y + z\ ,\ z\ ,\ 2x + y + 2z\ ).$$

Sua matriz na base canônica do $\mathbb{R}^3$ é

$$[\,T\,]_{can}^{can} = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix},$$

donde segue que o polinômio característico é

$$p(\lambda) = \det\big([\,T\,]_{can}^{can} - \lambda \mathbf{I}_3\big) = \det \begin{bmatrix} 1-\lambda & 0 & 2 \\ -1 & -\lambda & 1 \\ 1 & 1 & 2-\lambda \end{bmatrix}.$$

Usando desenvolvimento de Laplace sobre a primeira linha, obtemos

$$\begin{aligned} p(\lambda) &= (-1)^{1+1}(1-\lambda)\det\begin{bmatrix} -\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} + (-1)^{1+3}\, 2 \det \begin{bmatrix} -1 & -\lambda \\ 1 & 1 \end{bmatrix} \\ &= (1-\lambda)\big[-\lambda(2-\lambda) - 1\big] + 2(-1+\lambda) \\ &= (1-\lambda)\big[\lambda^2 - 2\lambda - 3\big]. \end{aligned}$$

Assim, os autovalores de T (raízes de p) são

$$\lambda_1 = 1, \quad \lambda_2 = -1, \quad \lambda_3 = 3.$$

Como os três autovalores são distintos entre si, T é diagonalizável. Podemos então calcular uma base de autovetores que diagonaliza T.

Autovetor associado a $\lambda_1 = 1$. Devemos encontrar uma solução não trivial do sistema $\big([\,T\,]_{can}^{can} - \mathbf{I}_3\big)[\,v\,]_{can} = \mathbf{0}$ (onde $[\,v\,]_{can} = \begin{bmatrix} x & y & z \end{bmatrix}^t$), ou seja,

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

Uma solução não trivial é

$$[\,v_1\,]_{can} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad \Rightarrow \qquad v_1 = (1, -1, 0).$$

Autovetor associado a $\lambda_2 = -1$. O sistema associado $\big([\,T\,]_{can}^{can} + \mathbf{I}_3\big)[\,v\,]_{can} = \mathbf{0}$ toma a forma

$$\begin{cases} 2x & & +2z & = & 0 \\ -x & +y & +z & = & 0 \\ x & +y & +3z & = & 0 \end{cases}.$$

Uma solução não trivial é (verifique)

$$[\,v_2\,]_{can} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \quad \Rightarrow \quad v_2 = (1, 2, -1).$$

Autovetor associado a $\lambda_3 = 3$. O sistema associado é $\big([\,T\,]_{can}^{can} - 3\,\mathbf{I}_3\big)[\,v\,]_{can} = \mathbf{0}$, ou seja,

$$\begin{cases} -2x & & +2z & = & 0 \\ -x & -3y & +z & = & 0 \\ x & +y & -z & = & 0 \end{cases}.$$

Uma solução não trivial é

$$[\,v_3\,]_{can} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \Rightarrow \quad v_3 = (1, 0, 1).$$

Portanto o conjunto

$$\beta = \{v_1, v_2, v_3\} = \{(1,-1,0)\ ;\ (1,2,-1)\ ;\ (1,0,1)\ \}$$

é uma base de $\mathbb{R}^3$ formanda por autovetores de T.

Finalmente, vimos que a diagonal de $[\,T\,]_\beta^\beta$ é formada pelos autovalores na mesma ordem em que os autovetores aparecem na base β, ou seja, **sem a necessidade de fazer contas** temos

$$[\,T\,]_\beta^\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

□

Exemplo 7.4. Considere o operador linear sobre $\mathbb{R}^4$

$$T(x,y,z,w) = (\ x + y - z + 2w\ ,\ 2y + w\ ,\ -z + 3w\ ,\ 5w\ ).$$

A matriz de T na base canônica de $\mathbb{R}^4$ é

$$[\,T\,]_{can}^{can} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ 2 & 1 & 3 & 5 \end{bmatrix},$$

cujos autovalores são

$$\lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = -1, \quad \lambda_4 = 5,$$

pois trata-se de uma matriz triangular inferior. Como todos os autovalores são distintos, T é diagonalizável. Isto é, é possível construir uma base β de $\mathbb{R}^4$ formada por autovetores de T, para a qual $[\,T\,]_\beta^\beta$ será diagonal. □

Atividade 7.1. Calcule uma base de $\mathbb{R}^4$ formada por autovetores de T no exercício anterior, e escreva a matriz de T nessa base.

7.1.2 Matrizes

Operadores lineares podem ser descritos por suas matrizes em determinadas bases; portanto, eles podemos falar em *matrizes diagonalizáveis* também.

Definição 7.2. *Uma matriz quadrada* $\mathbf{A}$ *de ordem* n *é dita* ***diagonalizável*** *se o operador associado* $T : \mathbb{R}^n \to \mathbb{R}^n$ *definido por* $[T]_{can}^{can} = \mathbf{A}$ *é diagonalizável. Isto é,* $\mathbf{A}$ $(n \times n)$ *é diagonalizável se admite* n *autovetores LI.*

Para saber se uma matriz é diagonalizável, procedemos da mesma forma que com operadores:

- Calculamos as raízes do polinômio característico de $\mathbf{A}$,
$$p(\lambda) = \det\big(\mathbf{A} - \lambda\mathbf{I}_n\big)$$
- Se as raízes (autovalores de $\mathbf{A}$) forem todas distintas entre si, então $\mathbf{A}$ é diagonalizável e podemos calcular os autovetores resolvendo os sistemas $\mathbf{A}\mathbf{v} = \lambda_i\mathbf{v}$ (ou $(\mathbf{A} - \lambda_i\mathbf{I}_n)\mathbf{v} = \mathbf{0}$).

Exemplo 7.5. A matriz
$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 8 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$
é diagonalizável pois seu polinômio característico é
$$p(\lambda) = \det\big(\mathbf{A} - \lambda\mathbf{I}_3\big) = \det\begin{bmatrix} 1-\lambda & 1 & 8 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & -1-\lambda \end{bmatrix} = (1-\lambda)(2-\lambda)(-1-\lambda),$$
cujas raízes são distintas entre si: $\lambda_1 = 1$, $\lambda_2 = 2$ e $\lambda_3 = -1$. □

Atividade 7.2. Calcule 3 autovetores LI da matriz $\mathbf{A}$ do exemplo anterior.

7.1.3 O polinômio minimal

É de interesse saber se um operador/matriz é diagonalizável antes de calcular seus autovetores, pois, caso não seja, evitamos buscar uma base de autovetores em vão. Nas seções anteriores vimos um caso onde isso é possível: quando T (ou $\mathbf{A}$) possui todos seus autovalores distintos entre si (Corolário 7.1). Mas quando T (ou $\mathbf{A}$) possui alguns autovalores iguais? Como decidir apenas pelos autovalores se T (ou $\mathbf{A}$) é diagonalizável?

A resposta passa por eliminar raízes repetidas do polinômio característico (autovalores com multiplicidade algébrica ≥ 1). Isso nos levará a um polinômio com mesmas raízes que o polinômio característico, mas com grau menor. Definiremos um polinômio de grau mínimo possível com essas características, que chamaremos de *polinômio minimal*.

Como de costume, utilizaremos matrizes do operador T em nossas contas. Iniciemos então o estudo em termos de matrizes.

Definição 7.3. *Seja*
$$q(x) = a_n x^n + \cdots + a_1 x + a_0$$
um polinômio e $\mathbf{A}$ *uma matriz quadrada. Então* $q(\mathbf{A})$ *é matriz*
$$q(\mathbf{A}) = a_n\mathbf{A}^n + \cdots + a_1\mathbf{A} + a_0\mathbf{I}.$$
Isto é, trocamos x *por* $\mathbf{A}$ *e* 1 *por* $\mathbf{I}$*, e operamos normalmente.*

Quando $q(\mathbf{A}) = \mathbf{0}$ *(matriz nula) dizemos que o polinômio* q ***anula*** *a matriz* $\mathbf{A}$.

Exemplo 7.6. Sejam $q(x) = x^2 - 9$, $t(x) = 2x + 3$ e $\mathbf{A} = \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix}$. Temos

$$q(\mathbf{A}) = \mathbf{A}^2 - 9\mathbf{I}_2 = \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix} - 9 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

e

$$t(\mathbf{A}) = 2\mathbf{A} - 3\mathbf{I}_2 = 2 \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix} + 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 8 \\ 4 & 5 \end{bmatrix}.$$

Assim, q anula $\mathbf{A}$ e t não anula $\mathbf{A}$. □

Definição 7.4. *Seja* $\mathbf{A}$ *uma matriz quadrada. O* ***polinômio minimal*** *de* $\mathbf{A}$ *é o polinômio da forma*

$$m(x) = x^r + a_{r-1}x^{r-1} + \cdots + a_1 x + a_0$$

(observe que $a_r = 1$*) tal que*

- $m(\mathbf{A}) = \mathbf{0}$*, isto é,* m *anula* $\mathbf{A}$*;*
- m *é o polinômio de menor grau que anula* $\mathbf{A}$*.*

Podemos responder se o operador linear $T : V \to V$ é diagonalizável ou não olhando para o polinômio minimal da matriz (quadrada) $[T]_\beta^\beta$ em qualquer base β de V.

Teorema 7.2. *Sejam* $T : V \to V$ *um operador linear e* β *uma base qualquer de* V*. Então* T *é diagonalizável se, e somente se, o polinômio minimal da matriz* $[T]_\beta^\beta$ *tem a forma*

$$m(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_r),$$

onde $\lambda_1, \ldots, \lambda_r$ *são os autovalores distintos de* T*.*

O teorema acima indica que precisamos apenas calcular so autovalores de T para saber se o operador é diagonalizável ou não. Isso envolve, evidentemente, calcular as raízes do polinômio característico.
O que falta é uma maneira de calcular o polinômio minimal! Veremos como fazer isso a partir do polinômio característico.

Teorema 7.3 (de Cayley-Hamilton)**.** *Sejam* $T : V \to V$ *um operador linear,* β *uma base de* V*, e*

$$p(\lambda) = \det\left([T]_\beta^\beta - \lambda \mathbf{I}\right)$$

o polinômio característico associado a T*. Então* p *anula* $[T]_\beta^\beta$*, isto é,*

$$p([T]_\beta^\beta) = \mathbf{0}.$$

NÃO se pode concluir que $p([T]_\beta^\beta) = \mathbf{0}$ pensando que "$\det\left([T]_\beta^\beta - [T]_\beta^\beta \mathbf{I}\right) = \det \mathbf{0}$"... **Esta conta está ERRADA e não faz sentido**, pois $\det \mathbf{0}$ é um número, não uma matriz nula! Deve-se primeiro calcular o polinômio $p(\lambda)$ desenvolvendo o determinante normalmente e **SÓ DEPOIS** substituir λ por $[T]_\beta^\beta$. A prova deste teorema é bem mais complicada...

O Teorema de Cayley-Hamilton diz que o polinômio característico é **candidato** a ser o polinômio minimal, pois anula a matriz de T. Agora, devemos cuidar das possíveis raízes repetidas!

Teorema 7.4. *Os polinômios minimal e característico de* $[T]_\beta^\beta$ *têm as mesmas raízes, exceto pelas multiplicidades.*

O teorema acima diz que se λ é autovalor de T, então ambos os polinômios característico e minimal possuem o fator (ou mais precisamente, são divisíveis por)

$$x - \lambda.$$

Assim, os Teoremas 7.2, 7.4 e 7.3 fornecem a seguinte maneira de calcular o polinômio minimal e decidir se T é diagonalizável ou não:

Cálculo do polinômio minimal

1. Calcule o polinômio característico $p(\lambda) = \det\big([T]_\beta^\beta - \lambda\mathbf{I}\big)$;
2. Encontre todas as raízes de p (autovalores); digamos que suas raízes distintas sejam $\lambda_1, \ldots, \lambda_r$;
3. Escreva $p(x)$ na forma

   $$p(x) = c\,(x - \lambda_1)^{d_1}(x - \lambda_2)^{d_2} \cdots (x - \lambda_r)^{d_r}$$

   (λ foi trocado por x por conveniência). Os expoentes $d_1, \ldots, d_r$ são o número de vezes que cada autovalor aparece como raiz de p, ou seja, são as multiplicidades algébricas de cada autovalor;
4. O polinômio minimal será o polinômio **de menor grau**, com $c = 1$, que contém pelo menos um de cada termo $x - \lambda_i$ $(i = 1, \ldots, r)$, e que **anule a matriz** $[T]_\beta^\beta$.

Decidindo se T é diagonalizável ou não

5. Com o polinômio minimal $m(x)$ em mãos, T será diagonalizável se, e somente se, se **cada termo** $x - \lambda_i$ aparece uma **única** vez em $m(x)$ (Teorema 7.2).

Exemplo 7.7. Considere o operador linear sobre $\mathbb{R}^3$ dado por

$$T(x, y, z) = (\ x\ ,\ y\ ,\ -x + 2z\ ).$$

Temos

$$[T]_{can}^{can} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix},$$

e o polinômio característico é

$$p(\lambda) = \det\big([T]_{can}^{can} - \lambda\mathbf{I}_3\big) = \det \begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ -1 & 0 & 2-\lambda \end{bmatrix} = (1-\lambda)^2(2-\lambda).$$

Como o polinômio minimal tem as mesmas raízes de p, ele deve ser um dos seguintes polinômios:

- $p_1(x) = (1 - x)(2 - x)$
- $p_2(x) = (1 - x)^2(2 - x)$ (o próprio polinômio característico p).

Por definição, o polinômio minimal $m(x)$ é aquele de **menor grau** que **anula** $[T]_{can}^{can}$. Então, testamos se cada candidato $p_1(x)$ e $p_2(x)$ anula $[T]_{can}^{can}$, em ordem crescente de grau. Ao encontrarmos o primeiro polinômio que anula esta matriz, paramos a busca!

Primeiro, verificamos se $p_1(x)$ anula $[T]_{can}^{can}$: expandindo $p_1(x)$, temos

$$p_1(x) = x^2 - 3x + 2.$$

Assim,

$$\begin{aligned} p_1([T]_{can}^{can}) &= \big([T]_{can}^{can}\big)^2 - 3[T]_{can}^{can} + 2\mathbf{I}_3 \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} - 3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} + 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 4 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 3 & 0 & -4 \end{bmatrix} = \mathbf{0} \end{aligned}$$

(você pode também calcular $(\mathbf{I}_3 - [T]_{can}^{can})(2\mathbf{I}_3 - [T]_{can}^{can})$ e chegar ao mesmo resultado). Ou seja, $p_1([T]_{can}^{can}) = \mathbf{0}$, e portanto p_1 anula $[T]_{can}^{can}$. Assim, o polinômio minimal é

$$m(x) = (1-x)(2-x).$$

Como cada um dos termos do tipo $(a-x)$ aparecem uma única vez $m(x)$, concluímos que T é diagonalizável. □

Exemplo 7.8. Considere o operador linear sobre $\mathbb{R}^4$ definido por

$$T(x, y, z, w) = (\ 3x - 4z\ ,\ 3y + 5z\ ,\ -z\ ,\ -w\ ).$$

Temos

$$[T]_{can}^{can} = \begin{bmatrix} 3 & 0 & -4 & 0 \\ 0 & 3 & 5 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

e, como esta matriz é triangular,

$$p(\lambda) = (3-\lambda)^2(-1-\lambda)^2.$$

Os candidatos à polinômio minimal são

- $p_1(x) = (3-x)(-1-x)$ (polinômio de grau 2)
- $p_2(x) = (3-x)^2(-1-x)$ (polinômio de grau 3)
- $p_3(x) = (3-x)(-1-x)^2$ (polinômio de grau 3)
- $p_4(x) = p(x) = (3-x)^2(-1-x)^2$ (polinômio de grau 4).

O operador T será diagonalizável somente se $p_1(x)$ for seu polinômio minimal, dado que em todos os outros candidatos há termos de grau 1 aparecendo mais de uma vez. Temos

$$\begin{aligned} p_1([T]_{can}^{can}) &= \big(3\mathbf{I}_4 - [T]_{can}^{can}\big)\big(-\mathbf{I}_4 - [T]_{can}^{can}\big) \\ &= \begin{bmatrix} 0 & 0 & 4 & 0 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} -4 & 0 & 4 & 0 \\ 0 & -4 & -5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{0}. \end{aligned}$$

Ou seja, $p_1(x)$ anula $[T]_{can}^{can}$. Sendo o candidato de menor grau, $p_1(x)$ é o polinômio minimal. Logo, T é diagonalizável. □

Exemplo 7.9. Considere o operador linear sobre $\mathbb{R}^5$ definido por

$$[T]_{can}^{can} = \begin{bmatrix} 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

O polinômio característico é

$$p(\lambda) = (2-\lambda)^3(1-\lambda)^2,$$

e os candidatos à polinômio minimal são

- $p_1(x) = (2-x)(1-x)$ (grau 2)
- $p_2(x) = (2-x)^2(1-x)$ (grau 3)
- $p_3(x) = (2-x)(1-x)^2$ (grau 3)
- $p_4(x) = (2-x)^3(1-x)$ (grau 4)
- $p_5(x) = (2-x)^2(1-x)^2$ (grau 4)
- $p_6(x) = p(x) = (2-x)^3(1-x)^2$ (grau 5).

Você pode verificar que

$$p_1([T]_{can}^{can}) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \neq \mathbf{0},$$

e logo T **não** é diagonalizável. Se calcularmos $p_2([T]_{can}^{can}), \ldots, p_5([T]_{can}^{can})$ em sequência, veremos que o primeiro polinômio que anula $[T]_{can}^{can}$ é $p_3(x)$. Ou seja, $p_3(x)$ é o polinômio minimal $m(x)$. $\square$

Exemplo 7.10. Considere o operador definido pela *matriz simétrica*

$$[T]_{can}^{can} = \begin{bmatrix} 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

que é uma pequena modificação do exemplo anterior. Calculando o polinômio característico, temos

$$p(\lambda) = \det\left([\,T\,]_{can}^{can} - \lambda \mathbf{I}_3\right) = \det \begin{bmatrix} 2-\lambda & 0 & 1 & 0 & 0 \\ 0 & 2-\lambda & 0 & 0 & 0 \\ 1 & 0 & 2-\lambda & 0 & 0 \\ 0 & 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 0 & 1-\lambda \end{bmatrix}$$

$$= (2-\lambda)(-1)^{1+1} \det \begin{bmatrix} 2-\lambda & 0 & 0 & 0 \\ 0 & 2-\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix}$$

$$+ (-1)^{1+3} \det \begin{bmatrix} 0 & 2-\lambda & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix}$$

$$= (2-\lambda)^3(1-\lambda)^2 - (2-\lambda)(1-\lambda)^2$$

$$= (2-\lambda)(1-\lambda)^2\left[(2-\lambda)^2 - 1\right]$$

$$= (2-\lambda)(1-\lambda)^2(\lambda^2 - 4\lambda + 3)$$

. As raízes da equação do segundo grau $\lambda^2 - 4\lambda + 3 = 0$ são 1 e 3. Portanto podemos escrever

$$p(\lambda) = (2-\lambda)\,(1-\lambda)^3\,(3-\lambda).$$

O candidato à polinômio minimal que não possui raízes repetidas é

$$p_1(x) = (2-x)\,(1-x)\,(3-x),$$

e, fazendo as contas, podemos verificar que

$$p_1([\,T\,]_{can}^{can}) = \mathbf{0}.$$

Ou seja, $m(x) = p_1(x)$ é o polinômio minimal. Logo T é diagonalizável. □

Os dois exemplos anteriores se parecem. Porém, o operador do primeiro exemplo não é diagonalizável, enquanto o operador do segundo exemplo sim. Isso não ocorre à toa: veremos adiante que toda matriz simétrica é diagonalizável!

Finalmente, a mesma estratégia feita para operadores pode ser aplicada à matrizes. Veja que em todos os exemplos anteriores olhamos para as matrizes dos operadores.

Exemplo 7.11. Pergunta-se se a matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

é diagonalizável. Seu polinômio característico é

$$p(\lambda) = \det\left(\mathbf{A} - \lambda \mathbf{I}_2\right) = \det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2.$$

Os candidatos à polinômio minimal são

- $p_1(x) = (1 - x)$
- $p_2(x) = p(x) = (1 - x)^2$.

Veja que

$$p_1(\mathbf{A}) = \mathbf{I}_2 - \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \mathbf{0}.$$

Daí, $p_1(x)$ não é o polinômio minimal de $\mathbf{A}$. Ou seja, o termo de grau 1 em $m(x)$ não aparece uma única vez, e logo $\mathbf{A}$ **não** é diagonalizável. □

7.1.4 Diagonalização e multiplicidade geométrica – uma visão geométrica

Existe uma relação entre diagonalização e as multiplicidades geométricas dos autovalores. Retomando, a multiplicidade geométrica de um autovalor λ do operador (ou matriz) $T : V \to V$ é a dimensão do autoespaço

$$V_\lambda = \{v \in V \mid T(v) = \lambda v\}.$$

Ora, dizer que T é diagonalizável é o mesmo que dizer que existe uma base de V formada por autovetores de T. É claro que tal base deve pertencer à união dos autoespaços; isto é, se $\lambda_1, \ldots, \lambda_r$ são os autovalores **distintos** de T, então uma base $\beta = \{v_1, \ldots, v_n\}$ de V formada por autovetores é tal que

$$\{v_1, \ldots, v_n\} \subset V_{\lambda_1} \cup V_{\lambda_2} \cup \cdots \cup V_{\lambda_r}.$$

Mas como a união $V_{\lambda_1} \cup V_{\lambda_2} \cup \cdots \cup V_{\lambda_r}$ está contida em V, devemos ter

$$V = V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_r}.$$

Do Teorema 7.1, esta soma é direta, isto é,

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r}.$$

Então podemos dizer que

"o operador $T : V \to V$ é diagonalizável se, e somente se, podemos decompor V como soma direta dos autoespaços associados aos autovalores de T"

ou ainda,

"o operador $T : V \to V$ é diagonalizável se, e somente se, a soma das multiplicidades geométricas dos autovalores distintos de T é igual à dimensão do espaço V".

7.2 Diagonalização de matrizes simétricas

Seja $\mathbf{A}$ uma matriz $n \times n$ simétrica, isto é, $\mathbf{A} = \mathbf{A}^t$. Considere o operador $T : \mathbb{R}^n \to \mathbb{R}^n$ definido por $[\,T\,]_{can}^{can} = \mathbf{A}$.

Operadores lineares sobre $\mathbb{R}^n$ cuja matriz na base canônica é simétrica são especiais, pois eles **sempre** são diagonalizáveis. Tais operadores são chamados *auto-adjuntos*. Não faremos o estudo de operadores auto-adjuntos aqui, pois isso depende de conceitos ainda não vistos. Nos restringiremos à linguagem das matrizes.

Lembre-se da Definição 7.2 que a matriz $\mathbf{A}$ de ordem $n \times n$ é diagonalizável se admite n autovetores LI. O primeiro fato sobre matrizes simétricas é que seu polinômio característico possui somente raízes reais.

Teorema 7.5. *Seja* $\mathbf{A}$ *matriz simétrica. Então o polinômio característico*

$$p(\lambda) = \det\big(\mathbf{A} - \lambda \mathbf{I}\big)$$

possui somente raízes reais.

O teorema anterior não vale em geral para uma matriz **não simétrica**. Por exemplo,

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

tem polinômio característico

$$p(\lambda) = (1-\lambda)^2 + 1 = \lambda^2 - 2\lambda + 2,$$

e a equação $p(\lambda) = 0$ possui raízes complexas $1 \pm i$.

Vimos anteriormente que há matrizes/operadores diagonalizáveis e outros não. A descoberta se uma matriz é ou não diagonalizável foi feita pelo polinômio minimal. Ocorre que para matrizes *simétricas*, temos a certeza que são diagonalizáveis!

Teorema 7.6. *Toda matriz simétrica* $\mathbf{A}$ *é diagonalizável.*

7.3 Matrizes semelhantes

Dado um operador $T : \mathbb{R}^n \to \mathbb{R}^n$ e bases α e β de $\mathbb{R}^n$, sabemos que

$$[\,T\,]_\beta^\beta = [\,I\,]_\beta^\alpha\, [\,T\,]_\alpha^\alpha\, [\,I\,]_\alpha^\beta \quad \Rightarrow \quad [\,T\,]_\beta^\beta = \big([\,I\,]_\alpha^\beta\big)^{-1} [\,T\,]_\alpha^\alpha\, [\,I\,]_\alpha^\beta. \tag{7.1}$$

Chamando $\mathbf{B} = [\,T\,]_\beta^\beta$, $\mathbf{A} = [\,T\,]_\alpha^\alpha$ e $\mathbf{P} = [\,I\,]_\alpha^\beta$, temos

$$\mathbf{B} = \mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P}.$$

Isso motiva a seguinte definição:

Definição 7.5. *Duas matrizes quadradas* $\mathbf{A}$ *e* $\mathbf{B}$ *de mesma ordem são* ***semelhantes*** *(ou* ***similares****) se existir uma matriz* $\mathbf{P}$ *inversível tal que*

$$\mathbf{B} = \mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P}.$$

Quando $\mathbf{A}$ é diagonalizável, é semelhante à uma matriz diagonal $\mathbf{D}$, pois neste caso $\mathbf{A}$ pode ser vista como a matriz de um operador diagonalizável $T : \mathbb{R}^n \to \mathbb{R}^n$, e $\mathbf{D}$ a matriz desse operador na base β que o diagonaliza.

Ou seja, no caso de $\mathbf{A}$ ser diagonalizável, podemos escrever

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P},$$

onde $\mathbf{D}$ é matriz diagonal. Mas quem são $\mathbf{D}$ e $\mathbf{P}$? Como calculá-las?

Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ o operador definido por $[\,T\,]_{can}^{can} = \mathbf{A}$. Da expressão (7.1), vemos que

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \qquad \text{e} \qquad \mathbf{P} = [\,I\,]_{can}^\beta = \begin{bmatrix} | & | & & | \\ [\,v_1\,]_{can} & [\,v_2\,]_{can} & \cdots & [\,v_n\,]_{can} \\ | & | & & | \end{bmatrix},$$

onde $\beta = \{v_1, \ldots, v_n\}$ é uma base de autovetores associados aos autovalores $\lambda_1, \ldots, \lambda_n$. Isto é,

Se $\mathbf{A}$ é uma matriz quadrada diagonalizável, então

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P},$$

onde $\mathbf{D}$ é a matriz diagonal formada pelos autovalores de $\mathbf{A}$ e $\mathbf{P}$ é a matriz cujas colunas são os autovetores associados (na mesma ordem).

Em particular, a escrita $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ **sempre** é possível para matrizes $\mathbf{A}$ **simétricas**.
Em outras palavras, qualquer matriz simétrica é semelhante à uma matriz diagonal.

Vale também a recíproca: se $\mathbf{A}$ é semelhante à uma matriz diagonal, digamos $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, então $\mathbf{D}$ é formada pelos autovalores e as colunas de $\mathbf{P}$ são os autovetores associados (veja o Exercício 6, Lista 7). Desta forma, toda matriz semelhante à uma diagonal é diagonalizável (isso dá ainda mais sentido ao termo "diagonalizável").

Exemplo 7.12. Considere a matriz simétrica

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

Sabemos que $\mathbf{A}$ é diagonalizável (Teorema 7.6). Vamos calcular $\mathbf{D}$ e $\mathbf{P}$ tais que $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

Autovalores de A: Temos

$$p(\lambda) = \det\big(\mathbf{A} - \lambda \mathbf{I}_4\big) = (-1-\lambda)(1-\lambda)^2(4-\lambda)$$

(verifique). Assim, os autovalores de $\mathbf{A}$ (com repetições) são $\lambda_1 = -1$, $\lambda_2 = 1$, $\lambda_3 = 1$ e $\lambda_4 = 4$, e logo

$$\mathbf{D} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

Autovetor associado a $\lambda_1 = -1$. Resolvendo $\mathbf{A}\mathbf{v} = -\mathbf{v}$, encontramos a solução particular

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

Autovetores associados a $\lambda_2 = \lambda_3 = 1$. Resolvendo $\mathbf{A}\mathbf{v} = \mathbf{v}$, encontramos as soluções LI particulares

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} \qquad \text{e} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Autovetor associado a $\lambda_4 = 4$. Resolvendo $\mathbf{Av} = 4\mathbf{v}$, encontramos a solução particular

$$\mathbf{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Assim, construímos **P seguindo a ordem dos autovalores/autovetores**:

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & -2 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Você pode verificar que, de fato,

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}.$$

□

7.4 Demonstrações

Demonstração do Teorema 7.1. Sejam $v_1, \ldots, v_r$ autovetores de T associados aos autovalores $\lambda_1, \ldots, \lambda_r$, distintos entre si. Considere a combinação linear

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r = \mathbf{0}. \tag{7.2}$$

Devemos mostrar que $a_i = 0$ para todo $i = 1, \ldots, r$. Para cada i, definimos a função $T - \lambda_i Id : V \to V$ por

$$(T - \lambda_i Id)(u) = T(u) - \lambda_i u.$$

É fácil mostrar que $T - \lambda_i Id$ é um operador linear sobre V. Veja que

$$(T - \lambda_i Id)(v_i) = T(v_i) - \lambda_i v_i = \mathbf{0}, \quad \forall i,$$

e que

$$(T - \lambda_i Id)(v_j) = T(v_j) - \lambda_i v_j = (\lambda_j - \lambda_i)v_j, \quad \forall i \neq j.$$

Assim, aplicando $T - \lambda_1 Id$ sobre (7.2) e usando sua linearidade, obtemos

$$\begin{aligned} a_1 \underbrace{(T - \lambda_1 Id)(v_1)}_{=\mathbf{0}} + a_2(T - \lambda_1 Id)(v_2) + \cdots + a_r(T - \lambda_1 Id)(v_r) &= \mathbf{0} \\ \Rightarrow \quad a_2(T - \lambda_1 Id)(v_2) + \cdots + a_r(T - \lambda_1 Id)(v_r) &= \mathbf{0} \\ \Rightarrow \quad a_2(\lambda_2 - \lambda_1)(v_2) + \cdots + a_r(\lambda_r - \lambda_1)(v_r) &= \mathbf{0} \end{aligned}$$

Da mesma forma, aplicando $T - \lambda_2 Id$ sobre a última equação acima obteremos

$$a_3(\lambda_3 - \lambda_2)(\lambda_3 - \lambda_1)(v_3) + \cdots + a_r(\lambda_r - \lambda_2)(\lambda_r - \lambda_1)(v_r) = \mathbf{0}.$$

Aplicando $T - \lambda_3 Id, \ldots, T - \lambda_r Id$ sucessivamente obteremos

$$a_r\big[(\lambda_r - \lambda_{r-1}) \cdots (\lambda_r - \lambda_1)\big] = \mathbf{0}.$$

Como todos os autovalores são distintos entre si, concluímos que $a_r = 0$. Substituindo $a_r = 0$ na equação imediatamente anterior, obtida pela aplicação sucessiva dos operadores $T - \lambda_i Id$, concluímos também que $a_{r-1} = 0$. Continuando a substituir esses coeficientes nulos nas equações anteriores, concluiremos que $a_1 = a_2 = \cdots = a_r = 0$, e logo os vetores $v_1, \ldots, v_r$ são LI, como queríamos demonstrar. □

Demonstração do Corolário 7.1. Consequência direta do Teorema 7.1 com $r = n$. □

Demonstração do Teorema 7.2. A prova é demasiadamente sofisticada para este texto. Veja, por exemplo, o livro de Álgebra Linear de Hoffman e Kunze. □

Demonstração do Teorema 7.3. Consulte [2] para uma demonstração no caso 2×2. Uma demonstração geral requer conceitos sofisticados, e pode ser encontrada no livro de Álgebra Linear de Hoffman e Kunze. □

Demonstração do Teorema 7.4. Sejam $p(x)$ e $m(x)$ os polinômios característico e minimal de $[T]_\beta^\beta$, respectivamente. Devemos mostrar que $p(\lambda) = 0$ se, e somente se, $m(\lambda) = 0$.

Primeiro, suponha que $m(\lambda) = 0$. Então podemos colocar o termo $x - \lambda$, referente à raiz λ, em evidência, obtendo

$$m(x) = (x - \lambda)q(x),$$

onde $q(x)$ é um polinômio de grau menor do que o grau de $m(x)$. Note ainda que

$$m([T]_\beta^\beta) = ([T]_\beta^\beta - \lambda\mathbf{I})\, q([T]_\beta^\beta).$$

Pela definição de polinômio minimal, $m(x)$ é o polinômio de menor grau que anula $[T]_\beta^\beta$, e logo $q(x)$ não pode anular esta matriz, isto é,

$$q([T]_\beta^\beta) \neq \mathbf{0}.$$

Assim, podemos escolher uma matriz coluna $\mathbf{w}$ tal que $\mathbf{v} = q([T]_\beta^\beta)\,\mathbf{w} \neq \mathbf{0}$. Como $m([T]_\beta^\beta) = \mathbf{0}$, temos

$$([T]_\beta^\beta - \lambda\mathbf{I})\,\mathbf{v} = ([T]_\beta^\beta - \lambda\mathbf{I})\, q([T]_\beta^\beta)\,\mathbf{w} = \mathbf{0} = m([T]_\beta^\beta)\,\mathbf{w} = \mathbf{0},$$

ou seja, $\mathbf{v}$ é autovetor da matriz $[T]_\beta^\beta$ associado à λ. Portanto, λ é autovalor de $[T]_\beta^\beta$, e logo $p(\lambda) = 0$.

Reciprocamente, suponha que $p(\lambda) = 0$. Vamos mostrar que $m(\lambda) = 0$. Sendo λ autovalor de $[T]_\beta^\beta$, temos

$$[T]_\beta^\beta\,\mathbf{v} = \lambda\mathbf{v} \tag{7.3}$$

para alguma matriz coluna $\mathbf{v} \neq \mathbf{0}$. Escrevamos o polinômio minimal m na forma

$$m(x) = x^r + a_{r-1}x^{r-1} + \cdots + a_1 x + a_0.$$

Por simplicidade, chamemos $\mathbf{A} = [T]_\beta^\beta$. Multiplicando a equação (7.3) por $\mathbf{A} = [T]_\beta^\beta$ à esquerda sucessivas vezes, obtemos

$$\begin{aligned}
&\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \\
\Rightarrow\ &\mathbf{A}^2\mathbf{v} = \lambda\mathbf{A}\mathbf{v} = \lambda^2\mathbf{v} \\
\Rightarrow\ &\mathbf{A}^3\mathbf{v} = \lambda\mathbf{A}^2\mathbf{v} = \lambda^3\mathbf{v} \\
&\vdots \\
\Rightarrow\ &\mathbf{A}^r\mathbf{v} = \lambda^r\mathbf{v}.
\end{aligned}$$

Assim,

$$\begin{aligned}
m(\mathbf{A})\,\mathbf{v} &= \left[\mathbf{A}^r + a_{r-1}\mathbf{A}^{r-1} + \cdots + a_1\mathbf{A} + a_0\mathbf{I}\right]\mathbf{v} \\
&= \mathbf{A}^r\mathbf{v} + a_{r-1}\mathbf{A}^{r-1}\mathbf{v} + \cdots + a_1\mathbf{A}\mathbf{v} + a_0\mathbf{I}\mathbf{v} \\
&= \lambda^r\mathbf{v} + a_{r-1}\lambda^{r-1}\mathbf{v} + \cdots + a_1\lambda\mathbf{v} + a_0\mathbf{v} \\
&= \left[\lambda^r + a_{r-1}\lambda^{r-1} + \cdots + a_1\lambda + a_0\right]\mathbf{v} = m(\lambda)\mathbf{v}.
\end{aligned}$$

Isto é, $m([\,T\,]_\beta^\beta)\,\mathbf{v} = m(\lambda)\mathbf{v}$. Agora, como $m([\,T\,]_\beta^\beta) = \mathbf{0}$ e $\mathbf{v} \neq \mathbf{0}$, concluímos que $m(\lambda) = 0$, isto é, λ é raiz do polinômio minimal. Isso finaliza a demonstração. □

Demonstração do Teorema 7.5. Olhando $\mathbf{A}$ como a matriz do operador linear $T : \mathbb{C}^n \to \mathbb{C}^n$ sobre o espaço vetorial das listas de números complexos $\mathbb{C}^n$, podemos interpretar as raízes de $p(\lambda)$ como autovalores associados à autovetores com componentes complexas $\mathbf{v} \in \mathbb{C}^n$. Assim, seja $\mathbf{v} \in \mathbb{C}^n$, $\mathbf{v} \neq \mathbf{0}$, autovetor de $\mathbf{A}$ associado ao autovalor λ. Multiplicando $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ à esquerda pelo conjugado de $\mathbf{v}^t$, a matriz linha $\overline{\mathbf{v}}^t$, obtemos

$$\overline{\mathbf{v}}^t\mathbf{A}\mathbf{v} = \lambda\,\overline{\mathbf{v}}^t\mathbf{v} \quad \Rightarrow \quad \lambda = \frac{\overline{\mathbf{v}}^t\mathbf{A}\mathbf{v}}{\overline{\mathbf{v}}^t\mathbf{v}}.$$

Dado um número complexo qualquer $a + bi$, temos $(a + bi)(a - bi) = a^2 + b^2 \in \mathbb{R}$. Isto é, o produto de um número complexo com seu conjugado é um número real. Assim,

$$\overline{\mathbf{v}}^t\mathbf{v} = \begin{bmatrix} \overline{v}_1 & \cdots & \overline{v}_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \overline{v}_1 v_1 + \cdots + \overline{v}_n v_n \in \mathbb{R}.$$

Da mesma forma, é fácil verificar que dados dois números complexos quaisquer $a + bi$ e $c + di$, vale

$$\overline{(a + bi)}(c + di) + (a + bi)\overline{(c + di)} = 2(ac + bd) \in \mathbb{R}.$$

Pela simetria de $\mathbf{A}$, temos $a_{ij} = a_{ji}$ para todos i, j, e portanto podemos escrever

$$\begin{aligned}
\overline{\mathbf{v}}^t\mathbf{A}\mathbf{v} &= \begin{bmatrix} \overline{v}_1 & \cdots & \overline{v}_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{v}_i v_j \\
&= 2\sum_{i=1}^{n} a_{ii}\overline{v}_i v_i + \sum_{i=1}^{n}\sum_{j=i+1}^{n} a_{ij}\left(\overline{v}_i v_j + v_i\overline{v}_j\right) \in \mathbb{R}.
\end{aligned}$$

Assim, concluímos que $\lambda = \frac{\overline{\mathbf{v}}^t\mathbf{A}\mathbf{v}}{\overline{\mathbf{v}}^t\mathbf{v}} \in \mathbb{R}$, como queríamos demonstrar. □

Demonstração do Teorema 7.6. A prova é feita por indução sobre n, a ordem da matriz simétrica $\mathbf{A}$. Para $n = 1$, a matriz $\mathbf{A}$ trivialmente admite 1 autovetor, que sozinho forma um conjunto LI. Suponha que $\mathbf{A}$ admita $n - 1$ ($n \geq 2$) autovetores LI e vamos mostrar que o teorema vale para n.

Sejam $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ autovetores LI associados à $\lambda_1, \ldots, \lambda_{n-1}$ respectivamente, e considere o conjunto

$$S^\perp = \{\mathbf{w} \in M(n,1) \mid \mathbf{v}_i^t\mathbf{w} = 0 \text{ para } i = 1, \ldots, n-1\}.$$

O sistema linear homogêneo

$$\left\{ \begin{array}{rcrcccrcl} \mathbf{v}_1^t\mathbf{w} & = & v_{11}w_1 & + & \cdots & + & v_{1n}w_n & = & 0 \\ \vdots & & & & & & & & \\ \mathbf{v}_{n-1}^t\mathbf{w} & = & v_{(n-1)1}w_1 & + & \cdots & + & v_{(n-1)n}w_n & = & 0 \end{array} \right.$$

tem $n-1$ equações e n incógnitas (as entradas de $\mathbf{w}$). Logo, possui solução não trivial, isto é, $S^\perp \neq \{\mathbf{0}\}$.

Tomemos $\mathbf{v}_n \in S^\perp$ não nulo. Neste caso, afirmamos que $\mathbf{v}_n$ não é combinação linear dos outros $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$. De fato, se fosse teríamos

$$\mathbf{v}_n = a_1\mathbf{v}_1 + \cdots + a_{n-1}\mathbf{v}_{n-1} \tag{7.4}$$

e multiplicando a expressão acima à esquerda pela transposta de $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_{n-1}\mathbf{v}_{n-1}$, teríamos

$$\sum_{i=1}^{n} w_i^2 = \mathbf{w}^t\mathbf{w} = \big(a_1\mathbf{v}_1 + \cdots + a_{n-1}\mathbf{v}_{n-1}\big)^t\mathbf{w} = a_1\mathbf{v}_1^t\mathbf{w} + \cdots + a_{n-1}\mathbf{v}_{n-1}^t\mathbf{w} = 0,$$

o que implica $\mathbf{w} = \mathbf{0}$. Mas como $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ são LI, isso implicaria $a_1 = \cdots = a_{n-1} = 0$, que por sua vez implicaria $\mathbf{v}_n = \mathbf{0}$, dado (7.4). Isso contraria o fato de $\mathbf{v}_n \neq \mathbf{0}$.

Agora, veja que

- o subespaço $[\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}]$ (subespaço gerado por $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$) tem dimensão $n-1$;
- o espaço das matrizes-coluna de tamanho n, $M(n,1)$, tem dimensão n ($M(n,1)$ é isomorfo a $\mathbb{R}^n$);
- da mesma forma como demonstrado anteriormente, $\mathbf{w} \notin [\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}]$ para todo $\mathbf{w} \in S^\perp$ não nulo. Dado que $S^\perp \subset M(n,1)$, temos $\dim S^\perp + \dim [\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}] = \dim M(n,1)$, ou seja
$$\dim S^\perp = \dim M(n,1) - \dim [\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}] = n - (n-1) = 1.$$
Assim $S^\perp = [\mathbf{v}_n]$.

Agora, veja que

$$(\mathbf{A}\mathbf{v}_n)^t\mathbf{v}_i = \mathbf{v}_n^t\mathbf{A}^t\mathbf{v}_i = \mathbf{v}_n^t\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_n^t\mathbf{v}_i = 0$$

para todo $i = 1, \ldots, n-1$. Isto é, $\mathbf{A}\mathbf{v}_n \in S^\perp$. Como $S^\perp = [\mathbf{v}_n]$, devemos ter $\mathbf{A}\mathbf{v}_n = \lambda_n\mathbf{v}_n$ para algum $\lambda_n \in \mathbb{R}$. Assim, construímos n autovetores $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{v}_n$ de $\mathbf{A}$ que são LI. Pelo principio da indução, a demonstração está concluída. □

7.5 Exercícios

1. Encontre os autovalores e autovetores dos operadores lineares abaixo e verifique se são diagonalizáveis. Caso afirmativo, encontre uma base β tal que $[T]_\beta^\beta$ seja diagonal.

   (a) $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $T(x,y,z) = (x+y, x-y+2z, 2x+y-z)$.

(b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $T(x, y, z) = (3x - 2y, -2x + 3y, 5z)$.

(c) $T : P_2 \to P_2$ dada por $T(p) = p$.

(d) $T : M(2,2) \to M(2,2)$ dada por $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2c & a+c \\ b-2c & d \end{bmatrix}$.

2. Diga se o operador é diagonalizável ou não, sem calcular seus autovetores.

   (a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (3x + y, -y)$.

   (b) $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $[T]^{\beta}_{\beta} = \begin{bmatrix} -1 & 1 \\ -6 & 4 \end{bmatrix}$, onde $\beta = \{(1, 1); (2, 3)\}$.

   (c) $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $T(x, y, z) = (x + 2z, x - y + z, 2x + z)$.

   (d) $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $[T]^{\alpha}_{\alpha} = \begin{bmatrix} -3 & -2 & 0 \\ 6 & 5 & 0 \\ 4 & 4 & -1 \end{bmatrix}$, onde $\alpha = \{(1, 0, 1); (2, 3, 0); (0, -1, 2)\}$.

3. Determine se cada matriz abaixo é diagonalizável ou não, sem calcular seus autovetores.

   (a) $\mathbf{A} = \begin{bmatrix} -3 & -4 \\ -1 & 1 \end{bmatrix}$

   (b) $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

   (c) $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 1 \\ 3 & 0 & 1 \end{bmatrix}$

   (d) $\mathbf{A} = \begin{bmatrix} 3 & 2 & -5 \\ 1 & 2 & -1 \\ 2 & 2 & -4 \end{bmatrix}$

   (e) $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & 4 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$

4. Para cada matriz **A diagonalizável** do exercício 3, encontre uma matriz inversível **P** que diagonaliza **A**, ou seja, tal que $\mathbf{P}^{-1}\mathbf{AP}$ seja diagonal.

5. Diagonalize as matrizes simétricas abaixo.

   (a) $\mathbf{A} = \begin{bmatrix} 1 & 4 & 2 \\ 4 & -5 & -4 \\ 2 & -4 & 1 \end{bmatrix}$

   (b) $\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & -3 \end{bmatrix}$

   (c) $\mathbf{A} = \begin{bmatrix} -1 & 3 \\ 3 & -1 \end{bmatrix}$

   (d) $\mathbf{A} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & c \\ 0 & c & b \end{bmatrix}$, onde $a, b, c \in \mathbb{R}$

6. Seja $T : V \to V$ um operador linear diagonalizável. Mostre que $V = \operatorname{Im} T \oplus \operatorname{Ker} T$.

Capítulo 8

Produto Interno

Veja as Seções 8.1 a 8.5, 9.2 e 9.3 de [2].

Apêndice A

Dicas e respostas dos exercícios

Capítulo 1 – Matrizes

1. (a) $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ (c) $\begin{bmatrix} -9 & 22 \\ -11 & 13 \end{bmatrix}$ (e) $\begin{bmatrix} 0 & 0 & 0 \\ 7 & 15 & -6 \end{bmatrix}$

   (b) $\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 0 \\ 2 & 4 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 3 \\ 7 & 4 \\ 0 & -3 \end{bmatrix}$

2. Sendo $\mathbf{A} = [a_{ij}]$ e $\mathbf{B} = [b_{ij}]$, temos $(\mathbf{A}+\mathbf{B})^t = [a_{ji}+b_{ji}] = [a_{ji}]+[b_{ji}] = \mathbf{A}^t+\mathbf{B}^t$. Também, $(\alpha\mathbf{A})^t = [\alpha a_{ji}] = \alpha[a_{ji}] = \alpha\mathbf{A}^t$.

3. Conclua a afirmação de $(\mathbf{A}^t\mathbf{A})^t = \mathbf{A}^t(\mathbf{A}^t)^t = \mathbf{A}^t\mathbf{A}$. O mesmo para $\mathbf{A}\mathbf{A}^t$.

4. (a) Como $\mathbf{A}^t = \mathbf{A}$ e $\mathbf{B}^t = \mathbf{B}$, temos $(\mathbf{A}+\mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t = \mathbf{A}+\mathbf{B}$.
   (b) $(\mathbf{AB})^t = \mathbf{AB} \Leftrightarrow \mathbf{B}^t\mathbf{A}^t = \mathbf{AB} \Leftrightarrow \mathbf{BA} = \mathbf{AB}$.
   (c) Mostre que $(\mathbf{A}^2)^t = \mathbf{A}^2$ ou use o item anterior.
   (e) Sim. Mostre que $(p(\mathbf{A}))^t = p(\mathbf{A})$.

5. (b) Note que $\mathbf{B}$ é anti-simétrica se, e somente se, $\mathbf{B} = -\mathbf{B}^t$. Mostre então que $\mathbf{A}-\mathbf{A}^t = -(\mathbf{A}-\mathbf{A}^t)^t$. A diagonal principal é de zeros.
   (c) Faça $\mathbf{S} = \frac{1}{2}(\mathbf{A}+\mathbf{A}^t)$ e $\mathbf{K} = \frac{1}{2}(\mathbf{A}-\mathbf{A}^t)$. Agora, se $\mathbf{S}'$ simétrica e $\mathbf{K}'$ anti-simétrica são matrizes tais que $\mathbf{A} = \mathbf{S}'+\mathbf{K}'$, então $\mathbf{S}' = \mathbf{A}-\mathbf{K}'$. Ora, mas $\mathbf{S}'^t = \mathbf{S}'$ e $\mathbf{K}'^t = -\mathbf{K}'$, e logo $\mathbf{S}' = \mathbf{S}'^t = (\mathbf{A}-\mathbf{K}')^t = \mathbf{A}^t+\mathbf{K}'$. Somando as relações chegamos a $\mathbf{S}' = \frac{1}{2}(\mathbf{A}+\mathbf{A}^t) = \mathbf{S}$. Isso mostra que $\mathbf{S}$ é única. Assim, $\mathbf{K}' = \mathbf{A}-\mathbf{S}$ também é única.
   (d) Basta notar que $\mathbf{A}(3\mathbf{I}-\mathbf{A}) = -(\mathbf{A}^2-3\mathbf{A}) = \mathbf{I}$.
   (g) Basta notar que $(\mathbf{I}-\mathbf{A})(\mathbf{I}+\mathbf{A}+\cdots+\mathbf{A}^k) = \mathbf{I}-\mathbf{A}^{k+1} = \mathbf{I}$ (verifique!).

6. Se $\mathbf{AB}$ for inversível, então existe $\mathbf{M}$ tal que $(\mathbf{AB})\mathbf{M} = \mathbf{I}$. Use então a associatividade $\mathbf{A}(\mathbf{BM}) = (\mathbf{AB})\mathbf{M}$ para exibir a inversa de $\mathbf{A}$. Use um raciocínio análogo para $\mathbf{B}$.

7. (c) Note que $\mathbf{A}+\mathbf{B} = (\mathbf{I}+\mathbf{BA}^{-1})\mathbf{A}$. Pelo exercício anterior, se $\mathbf{A}+\mathbf{B}$ é inversível, então $\mathbf{I}+\mathbf{BA}^{-1}$ também é. Agora, se $\mathbf{I}+B\mathbf{A}^{-1}$ for inversível então, como $\mathbf{A}$ é inversível, o produto $(\mathbf{I}+\mathbf{BA}^{-1})\mathbf{A}$ também é, com inversa $\mathbf{A}^{-1}(\mathbf{I}+\mathbf{BA}^{-1})^{-1}$. Logo $\mathbf{A}+\mathbf{B}$ é inversível.
   Conclusão: $\mathbf{I}+\mathbf{BA}^{-1}$ é inversível se, e somente se $\mathbf{A}+\mathbf{B}$ é inversível. Observe que é possível essas matrizes não serem inversíveis: por exemplo, tome $\mathbf{B} = -\mathbf{A}$. Neste caso, $\mathbf{A}+\mathbf{B} = \mathbf{I}+\mathbf{BA}^{-1} = 0$.

10. (c) Sendo $\mathbf{B} = [b_{ij}]$, $\mathbf{AB} = [c_{ij}]$ e $\mathbf{BA} = [d_{ij}]$, temos $c_{ii} = \sum_{k=1}^n a_{ik}b_{ki}$ e $d_{ii} = \sum_{k=1}^n a_{ki}b_{ik}$. Logo

$$tr(\mathbf{AB}) = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n \sum_{k=1}^n a_{ik}b_{ki} = \sum_{k=1}^n \sum_{i=1}^n a_{ik}b_{ki} = \sum_{i=1}^n \sum_{k=1}^n a_{ki}b_{ik} = \sum_{i=1}^n d_{ii} = tr(\mathbf{BA})$$

(e) Note que os elementos da diagonal de $\mathbf{A}^t\mathbf{A}$ são os elementos $c_{ii} = \sum_{k=1}^n a_{ki}a_{ki} = \sum_{k=1}^n a_{ki}^2 \geq 0$. Conclua.

11. Usando o exercício anterior, $tr(\mathbf{AB}) - tr(\mathbf{BA}) = 0 \neq n = tr\mathbf{I}_n$, donde conclui-se o resultado.

13. (a) $\mathbf{A}^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ e $\mathbf{A}^3 = \mathbf{A}^4 = \cdots = \mathbf{0}$. Se $\mathbf{B}$ é uma matriz de ordem 4 do mesmo tipo, então $\mathbf{B}^2 = \begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $\mathbf{B}^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ e $\mathbf{B}^4 = \mathbf{B}^5 = \cdots = \mathbf{0}$.

14. $\mathbf{A}^k = \begin{bmatrix} \cos k\theta & -\operatorname{sen} k\theta \\ \operatorname{sen} k\theta & \cos k\theta \end{bmatrix}$.

15. Primeiro observe que $\mathbf{AB} = \mathbf{BA} \Rightarrow \mathbf{A}^2\mathbf{B} = \mathbf{ABA} = \mathbf{BA}^2 \Rightarrow \cdots \Rightarrow \mathbf{A}^r\mathbf{B} = \mathbf{BA}^r$ para qualquer inteiro $r \geq 1$. Assim, $(\mathbf{AB})^k = \mathbf{AB}(\mathbf{AB})^{k-1} = \mathbf{ABAB}(\mathbf{AB})^{k-2} = \mathbf{BAAB}(A\mathbf{B})^{k-2} = \mathbf{BA}^2\mathbf{B}(\mathbf{AB})^{k-2} = \mathbf{B}^2\mathbf{A}^2(\mathbf{AB})^{k-2}$. Continuando,

$$(\mathbf{AB})^k = \mathbf{B}^2\mathbf{A}^2\mathbf{AB}(\mathbf{AB})^{k-3} = \mathbf{B}^2\mathbf{A}^3\mathbf{B}(\mathbf{AB})^{k-3} = \mathbf{B}^2\mathbf{BA}^3(\mathbf{AB})^{k-3} = \mathbf{B}^3\mathbf{A}^3(\mathbf{AB})^{k-3}.$$

Continuando o processo, concluímos que $(\mathbf{AB})^k = \mathbf{B}^r\mathbf{A}^r(\mathbf{AB})^{k-r}$ para todo inteiro r entre 1 e k. Assim, para $r = k$ temos $(\mathbf{AB})^k = \mathbf{B}^k\mathbf{A}^k = \mathbf{B}^k\mathbf{0} = \mathbf{0}$.

16. Suponha por absurdo que $\mathbf{A}$ seja inversível e use o exercício anterior.

17. Considere o produto $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ e use o exercício anterior.

18. Se $a_{11} \cdots a_{nn} \neq 0$ então $a_{ii} \neq 0$ para todo $1 \leq i \leq n$. Verifique então que a matriz diagonal cujos elementos da diagonal principal são $1/a_{ii}$ é a inversa de $\mathbf{A}$.

19. Considere as matrizes $\mathbf{X}_i$ de ordem $n \times 1$ onde a i-ésima entrada é 1 e as outras 0. Verifique então que $\mathbf{AX} = \mathbf{BX}$ diz que as i-ésimas colunas de $\mathbf{A}$ e $\mathbf{B}$ são iguais. Conclua então que $\mathbf{A} = \mathbf{B}$.

20. Multiplique as duas primeiras expressões por $\mathbf{A}$ e/ou $\mathbf{B}$ e manipule para chegar ao resultado.

Capítulo 2 – Sistemas Lineares

2. Obtenha as matrizes linha reduzidas à forma escada de $\mathbf{A}$ e $\mathbf{B}$ e veja que elas coincidem. Porque neste caso $\mathbf{A}$ e $\mathbf{B}$ são linha equivalentes?

3.

(a) $\dfrac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & -1 \\ -1 & 1 & 1 \end{bmatrix}$

(b) $\dfrac{1}{2}\begin{bmatrix} -1 & -1 & 3 \\ 2 & 0 & -2 \\ -1 & 1 & 1 \end{bmatrix}$

(c) não é inversível.

(d) $\dfrac{1}{8}\begin{bmatrix} 2 & 0 & 2 \\ 1 & 4 & -3 \\ -5 & 4 & -1 \end{bmatrix}$

(e) $\dfrac{1}{2}\begin{bmatrix} -14 & 6 & -4 & 10 \\ 18 & -7 & 7 & -13 \\ 12 & -5 & 5 & -9 \\ 16 & -7 & 5 & -11 \end{bmatrix}$

4. Sejam $\mathbf{X}_0$ e $\mathbf{X}_1$ soluções de $\mathbf{AX} = \mathbf{0}$ e $\mathbf{AX} = \mathbf{B}$, respectivamente. Temos $\mathbf{A}(\mathbf{X}_0 + \mathbf{X}_1) = \mathbf{AX}_0 + \mathbf{AX}_1 = \mathbf{0} + \mathbf{B} = \mathbf{B}$, ou seja, $\mathbf{X}_0 + \mathbf{X}_1$ é solução de $\mathbf{AX} = \mathbf{B}$.

5. Seja $\mathbf{X}_1$ solução qualquer de $\mathbf{AX} = \mathbf{B}$. Podemos escrever uma solução $\overline{\mathbf{X}}$ de $\mathbf{AX} = \mathbf{B}$ como $\overline{\mathbf{X}} = \mathbf{X}_1 + (\overline{\mathbf{X}} - \mathbf{X}_1)$. Observe que $\overline{\mathbf{X}} - \mathbf{X}_1$ é solução de $\mathbf{AX} = \mathbf{0}$ (verifique!), e que portanto $\overline{\mathbf{X}}$ pode ser escrito como soma da solução arbitrária $\mathbf{X}_1$ de $\mathbf{AX} = \mathbf{B}$ com a solução $\overline{\mathbf{X}} - \mathbf{X}_1$ do sistema homogêneo associado.

6. (a) Solução única $(-2, -1, 3)$.
   (b) Infinitas soluções $(x, y, z) = (1 - 2z, -1 + 3z, z), z \in \mathbb{R}$.
   (c) Solução única $(3, 1, 2)$.
   (d) Solução única $(-4, 2, 7)$.
   (e) Não tem solução.
   (f) Não tem solução.
   (g) Não tem solução.
   (h) Solução única $(1, 0, 1)$.
   (i) Infinitas soluções. Especifique o conjunto das soluções!
   (j) Infinitas soluções. Especifique o conjunto das soluções!
   (k) Infinitas soluções. Especifique o conjunto das soluções!

7. Arme um sistema linear que represente o polinômio passar pelos pontos dados. Isto é, use as expressões $p(-2) = 24$, $p(-1) = 24$, ... As variáveis serão os coeficientes $a_0, \ldots, a_3$. A resposta é $p(x) = -3x^3 - 16x^2 + 21x + 58$.

9. Considere o sistema $\mathbf{AX} = \mathbf{0}$. Multiplicando à esquerda por $\mathbf{A}^{-1}$, concluímos que o sistema tem como única solução $\mathbf{X} = \mathbf{0}$. Escalonando o sistema, obtemos o sistema equivalente $\mathbf{RX} = \mathbf{0}$ onde $\mathbf{R} = \mathbf{E}_n \cdots \mathbf{E}_1\mathbf{A}$ é a MLRFE de $\mathbf{A}$. Como $\mathbf{X} = \mathbf{0}$ também é a única solução de $\mathbf{RX} = \mathbf{0}$, só pode ser $\mathbf{R} = \mathbf{I}$ (pense na forma escada da matriz quadrada $\mathbf{R}$ de forma que $\mathbf{RX} = \mathbf{0}$ tem única solução). Assim, $\mathbf{E}_n \cdots \mathbf{E}_1\mathbf{A} = \mathbf{I}$, e multiplicando sucessivamente à esquerda pelas inversas das elementares $\mathbf{E}_i$ (que também são matrizes elementares), obtemos $\mathbf{A} = \mathbf{E}_n^{-1} \cdots \mathbf{E}_1^{-1}$.

10. Como $\mathbf{B}$ é linha equivalente a $\mathbf{A}$, temos $\mathbf{B} = \mathbf{E}_1 \cdots \mathbf{E}_m\mathbf{A}$. Como $\mathbf{A}$ e toda matriz elementar são inversíveis, o produto $\mathbf{E}_1 \cdots \mathbf{E}_m\mathbf{A}$ também é inversível. Daí, invertendo ambos os lados da equação, usando a dica, obtemos $\mathbf{B}^{-1} = \mathbf{A}^{-1}\mathbf{E}_m^{-1} \cdots \mathbf{E}_1^{-1}$.

11. (b) Veja que $\mathbf{Ax} = \begin{bmatrix} \mathbf{B} & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N \\ \mathbf{x}_N \end{bmatrix} = \mathbf{B}(\mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N) + \mathbf{N}\mathbf{x}_N = \mathbf{b} - \mathbf{N}\mathbf{x}_N + \mathbf{N}\mathbf{x}_N = \mathbf{b}$.
    (c) $\mathbf{x} = (1, -1/5, 3, 0, 0)$

Capítulo 3 – Determinantes

2.

(a) posto = 2, nulidade = 1, det. = 0 (b) posto = 4, nulidade = 0, det. = 8

4. (d) 1 (e) 1 (f) 48 (g) 1 (h) 900 (i) 394 (j) 2

6. $(\det \mathbf{B})^2 = \det(\mathbf{B}^t\mathbf{B}) = \det \mathbf{I}_n = 1 \Rightarrow |\det \mathbf{B}| = 1 \Rightarrow \det \mathbf{B} = 1$ ou $\det \mathbf{B} = -1$.

7. $\det \mathbf{A} = \det(\mathbf{A}\mathbf{A}) = (\det \mathbf{A})^2 \Rightarrow \det \mathbf{A} = 0$ ou $\det \mathbf{A} = 1$.

8. (a) Basta seguir a dica. Se não conseguir visualizar, mostre a igualdade para matrizes de ordem 3 primeiro.
   (b) $\mathbf{A}\,(\operatorname{adj}\mathbf{A}) = (\det \mathbf{A})\,\mathbf{I}_n \Rightarrow \mathbf{A}\,\left[\frac{1}{\det \mathbf{A}}(\operatorname{adj}\mathbf{A})\right] = \mathbf{I}_n \Rightarrow \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}}(\operatorname{adj}\mathbf{A})$
   (c) $\mathbf{A}^{-1} = \dfrac{1}{19}\begin{bmatrix} 19 & 5 & -4 \\ -19 & -10 & 8 \\ 19 & 11 & -5 \end{bmatrix}$

9. (b) $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B} = \dfrac{1}{\det \mathbf{A}}(\operatorname{adj}\mathbf{A})\mathbf{B} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \dfrac{1}{\det \mathbf{A}}\begin{bmatrix} \Delta_{11}b_1 + \Delta_{21}b_2 + \cdots + \Delta_{n1}b_n \\ \Delta_{12}b_1 + \Delta_{22}b_2 + \cdots + \Delta_{n2}b_n \\ \vdots \\ \Delta_{1n}b_1 + \Delta_{2n}b_2 + \cdots + \Delta_{nn}b_n \end{bmatrix}.$
   Pelo desenvolvimento de Laplace, conclue-se que o termo i dessa última matriz é $\det \mathbf{C}_i$.
   (c) $x = -49, y = 9, z = 18$

Capítulo 4 – Espaços Vetoriais

1. Basta verificar que os conjuntos são fechados para as operações de soma e multiplicação por escalar.

2. (a) Não pois não é fechado para soma (dê um exemplo!).
   (b) Não pois não é fechado para soma (dê um exemplo!). Note que dois elementos de Z não necessariamente têm a mesma coluna com elementos iguais.
   (c) Sim
   (d) Não pois não é fechado para soma. Por exemplo, $(0,0,1,1,1) + (1,1,0,0,0) = (1,1,1,1,1)$.
   (e) Não pois não é fechado para multiplicação por escalar. Por exemplo, $-1(1,2,3) = (-1,-2,-3)$.
   (f) Sim
   (g) Não pois não é fechado para soma (dê um exemplo!).

3. (a) $\{(1,1,1,1)\}$
   (b) $\{(1,1,1,0),(0,0,0,1)\}$
   (c) $\{(1,0,0,-1),(0,1,0,-1),(0,0,1,-1)\}$

6. (a) Qualquer reta passando pela origem $R = \{(x,y) \in \mathbb{R}^2; (x,y) = tv, t \in \mathbb{R}\}$ é subespaço vetorial, pois para quaisquer $u_1, u_2 \in R$ e $\alpha \in \mathbb{R}$ temos $u_1 + u_2 = t_1v + t_2v = (t_1 + t_2)v \in R$ e $\alpha u_1 = (\alpha t_1)v \in R$.

(b) É fácil mostrar que $\dim F = \dim G = 1$. Basta então mostrar que a soma $F + G$ é direta. Seja $v \in F \cap G$. Então $v = t_1u = t_2w$. Assim $t_1u - t_2w = 0 \Rightarrow t_1 = t_2 = 0$ pois u e w são LI (verifique!). Logo $v = \mathbf{0}$ e $F \cap G = \{\mathbf{0}\}$.

8. (a) Tome W_0 como o subespaço gerado por vetores que completam uma base de $\mathbb{R}^n$ a partir de uma base W.

   (b) Note que $W = [(1, 0, -1), (0, 1, -2)]$. Tome, por exemplo, $W_0 = [(0, 1, 0)]$.

9. (a) Verdadeiro. De fato, sejam β_1 e β_2 bases de W_1 e W_2, respectivamente. Como a soma $W_1 + W_2$ é direta, $\beta_2 \cap \beta_2 = \emptyset$, e logo $\dim V = \dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2$.

   (b) Falso. Só é verdadeiro se W_1 e W_2 forem subespaços de V.

   (c) Verdadeiro, pois se fosse $W_1 \cap W_2 = \{\mathbf{0}\}$ então, como uma base de $W_1 \oplus W_2$ possui $\dim W_1 + \dim W_2 \leq \dim V$ elementos, teríamos $\dim W_1 \leq \dim V/2$ ou $\dim W_2 \leq \dim V/2$.

   (d) Verdadeiro. De fato, sejam $\{w_1, \ldots w_m\}$ uma base de W_1 e $\{u_1, \ldots u_k\}$ uma base de W_2. Imediatamente, $v \in [W_1 \cup W_2] \Leftrightarrow v = (a_1w_1 + \ldots + a_mw_m) + (b_1u_1 + \ldots + b_ku_k) \Leftrightarrow v \in W_1 + W_2$.

10. Se $\{v_1, \ldots, v_n\}$ é uma base de V então podemos escrever $V = [v_1] \oplus \cdots \oplus [v_n]$.

11. É fácil mostrar que β é base. $[I]^{\beta}_{can} = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 3 & 1 & 3 \end{bmatrix}$, $[I]^{can}_{\beta} = \dfrac{1}{4}\begin{bmatrix} 0 & 4 & 0 & 0 \\ -3 & 3 & -1 & 1 \\ 6 & -6 & 6 & -2 \\ 1 & -1 & -1 & 1 \end{bmatrix}$.

12. $\beta = \{(-1, -1), (3, 4)\}$.

13. Porque uma linha é combinação das outras. Daí sua matriz linha reduzida à forma escada tem uma linha nula, e logo seu determinante é nulo.

15. (a) $\{(1, 2, 0), (0, 3, 1)\}$ é base para P e $\{(1, 1, -3)\}$ é base para R. Obviamente não são as únicas!

    (b) Pelo item anterior, $\dim P + \dim R = 3 = \dim \mathbb{R}^3$. Mostre que os vetores das duas bases são LI, e logo a soma $P + R$ é direta.

    (c) Uma base de P é formada por vetores não colineares paralelos ao plano π, e uma base de R é formada por um vetor não nulo na direção da reta r.

16. (a) Falso. (b) Falso. (c) Falso. (d) Verdadeiro. (e) Verdadeiro. (f) Falso.

17. H é subespaço de $\mathbb{R}^n$ pois é o conjunto solução de um sistema linear homogêneo. Podemos supor sem perda de generalidade que $a_n \neq 0$. Então $a_1x_1 + \cdots + a_nx_n = 0 \Rightarrow x_n = \frac{a_1}{a_n}x_1 + \cdots + \frac{a_{n-1}}{a_n}x_{n-1}$. Assim $\left\{(1, 0, \ldots, \frac{a_1}{a_n}), (0, 1, \ldots, \frac{a_2}{a_n}), \ldots, (0, 0, \ldots, 1, \frac{a_{n-1}}{a_n})\right\}$ é uma base de H, e logo $\dim H = n - 1$.

19. Um polinômio de grau $\leq n$ tem a forma $p(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots a_1x + a_01$. Uma base visível é formada pelos $n + 1$ polinômios $p_i(x) = x^i$, $i = 0, 1, \ldots, n$ (essa é a *base canônica* de P_n). Assim, $\dim P_n = n + 1 = \dim \mathbb{R}^{n+1}$. Escrever qualquer polinômio nesta base é trivial (por exemplo, aquele do item (b)).

20. (a) Verifique que a única solução para a combinação nula é a trivial.

(b) Uma base é formada por n^2 matrizes, cada uma com um 1 em uma posição diferente e o resto zero. Assim, $\dim M(n,n) = n^2$.

(c) Colete as matrizes simétricas da base anterior. Aí está uma base para $T_{\text{sim}}(n)$, o que implica $\dim T_{\text{sim}}(n) = n$.

(d) A soma é fácil de ver. Ela não é direta pois uma matriz diagonal é triangular superior e inferior simultaneamente. $T_S(n)$ e $T_I(n)$ têm ambos dimensão igual a $n(n+1)/2$ (conte os elementos de uma base para ver isso).

21. (a) Verifique que na combinação nula $\alpha\mathbf{A} + \beta\mathbf{B} + \gamma\mathbf{C} = \mathbf{0}$, só pode ser $\alpha = \beta = \gamma = 0$.

(b) Analise a combinação nula dos polinômios. Lembre-se que o elemento nulo do espaço vetorial dos polinômios é o polinômio definido por $\mathbf{0}(x) = 0$ para todo x.

22. Vale $[u,v] \subset [u,v,w]$ pois todo vetor $x = au + bv$ pode ser escrito como $x = au + bv + 0w$. Agora, $[u,v] = [u,v,w] \Leftrightarrow w \in [u,v]$. Assim, $[u,v] \subsetneq [u,v,w] \Leftrightarrow w$ não for combinação linear de u e v.

Capítulo 5 – Transformações Lineares

Transformações Lineares

1. (a) Sim

(b) Sim

(c) Não, pois $T(\mathbf{0}) \neq \mathbf{0}$

(d) Sim

(e) Sim

(f) Sim

(g) Não, pois, por exemplo, $T((1,0) + (0,1)) = 1 \neq 0 = T(1,0) + T(0,1)$.

(h) Não, pois $T(0,0,0) \neq (0,0,0)$.

(i) Sim

(j) Não, pois $T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = 1 \neq 0 = T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) + T\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right)$

(k) Sim

(l) Sim

(m) Sim

(n) Sim

(o) Sim

(p) Sim

(q) Sim. Veja exercício 10(a,b) de "Matrizes".

2. (a) sobrejetora

(b) isomofismo, $T^{-1} = T$

(d) isomofismo, $T^{-1} = (1/2(x-y), y)$

(e) isomofismo, $T^{-1} = (x/2, x/2 + y)$

(f) isomofismo, $T^{-1} = T$

(i) sobrejetora

(k) sobrejetora

(l) isomorfismo, $T^{-1}(p)(t) = p(t-1)$

(m) sobrejetora

(n) injetora

(q) sobrejetora

(n) isomorfismo, $T^{-1}(x,y,z) = \left(\frac{x}{2}(-t+t^2), y(1-t^2), \frac{z}{2}(t+t^2)\right)$

(o) isomorfismo, $T^{-1}(x,y,z) = 1/2(x-y+z, x+y-z, -x+y+z)$

3. Seja R o segmento de reta com pontos extremos u e $u+w$. Então $L(R) = \{L(u+tw) \mid 0 \leq t \leq 1\} = \{L(u) + tL(w) \mid 0 \leq t \leq 1\}$, isto é, $L(R)$ é um segmento com pontos extremos $L(u)$ e $L(u) + L(w) = L(u+w)$. O mesmo raciocínio é usado no caso de retas.

4. (a) $T(x,y) = (4y - x, 5y - x)$

   (b) $T(x,y) = (-\frac{4}{3}x + -\frac{13}{3}y, x - 3y)$
   (c) $T(x,y) = 2(2y - x, y)$

5. Não.

6. (a) Falso
   (b) Falso
   (c) Verdadeiro
   (d) Verdadeiro
   (e) Verdadeiro

7. Tome bases $\{v_1, \ldots, v_n\}$ de V e $\{w_1, \ldots, w_k\}$ de W. Temos $n < k$. Defina T pondo $T(v_i) = w_i$, $i = 1, \ldots, n$, e S pondo $S(w_j) = v_j$, $j = 1, \ldots, n$, $S(w_j) = \mathbf{0}$, $j = n+1, \ldots, k$. É fácil verificar que $\operatorname{Ker} T = \{\mathbf{0}\}$ (e logo T é injetora) e que $\operatorname{Im} S = V$ (e logo S é sobrejetora).

8. Impondo $T(1,2) = (1,1)$ e $T(3,4) = (2,2)$ você cairá em um sistema linear nas variáveis a, b, c, d. Esse sistema certamente tem única solução, pois T está definida sobre a base $\{(1,2);(3,4)\}$ de $\mathbb{R}^2$. Basta então resolvê-lo.

9. (a) $a_1u_1 + \cdots + a_mu_m = \mathbf{0} \Rightarrow T(a_1u_1 + \cdots + a_mu_m) = T(\mathbf{0}) \Rightarrow a_1T(u_1) + \cdots + a_mT(u_m) = \mathbf{0} \Rightarrow a_1 = a_2 = \cdots = a_m = \mathbf{0}$ pois $T(u_1), \ldots, T(u_m)$ são LI.
   (b) Digamos que $\dim V = \dim W = n$. Como $T(u_1), \ldots, T(u_m)$ geram W, podemos extrair uma base de W, digamos $\{T(u_1), \ldots, T(u_n)\}$. Pelo item anterior, $u_1, \ldots, u_n \in V$ são LI, e como $\dim V = n$, formam uma base para V. Portanto $u_1, \ldots, u_m$ geram V.

10. Como $G = \{u_1, \ldots, u_m\}$ gera V, qualquer $v \in V$ pode ser escrito uma combinação $v = a_1u_1 + \cdots + a_mu_m$. Assim, $T(v) = T(a_1u_1 + \cdots + a_mu_m) = a_1T(u_1) + \cdots + a_mT(u_m) = a_1S(u_1) + \cdots + a_mS(u_m) = S(a_1u_1 + \cdots + a_mu_m) = S(v)$.

11. Por um lado, $\operatorname{Im} T \subset V$. Por outro lado, $V = \operatorname{Im} Id_V = \operatorname{Im}(T \circ S) = T(S(V)) \subset T(V) = \operatorname{Im} T$, e logo $V \subset \operatorname{Im} T$. Assim, T é sobrejetora. Como T é uma transformação injetora de V em V, é também sobrejetora. Portanto T é automorfismo.

12. Você pode verificar que S é a transformação inversa de T: $T \circ S = ID_W \Rightarrow S = T^{-1} \circ Id_W = T^{-1}$. Logo é linear.

13. $S \circ T(v) = \mathbf{0}_V \Rightarrow S(T(v)) = \mathbf{0}_V \Rightarrow T(v) \in \operatorname{Ker} S = \{0_W\} \Rightarrow T(v) = \mathbf{0}_W \Rightarrow v \in \operatorname{Ker} T = \{0_V\} \Rightarrow v = \mathbf{0}_V \Rightarrow \operatorname{Ker}(S \circ T) = \{\mathbf{0}_V\}$.

14. Dado $v \in V$, podemos escrever $v = (v - T(v)) + T(v)$. Note que $T(v - T(v)) = T(v) - T^2(v) = 0$, e portanto v é soma de um elemento de $\operatorname{Ker} T$ com um elemento de $\operatorname{Im} T$. Segue que $V = \operatorname{Ker} T + \operatorname{Im} T$. Mostremos agora que a soma é direta. O caso em que $\operatorname{Im} T = \{\mathbf{0}\}$ ou $\operatorname{Ker} T = \{\mathbf{0}\}$ é óbvio. Considere o caso em que $\dim \operatorname{Im} T = k$, $0 < k < \dim V$. Seja então $v \in \operatorname{Ker} T \cap \operatorname{Im} T$, e suponha por absurdo que $v \neq \mathbf{0}$. Por um lado, $T(v) = \mathbf{0}$. Por outro lado $v = a_1T(w_1) + \cdots + a_kT(w_k)$ onde $T(w_1), \ldots, T(w_k)$ é uma base de $\operatorname{Im} T$. Como $v \neq \mathbf{0}$, pelo menos um a_i é não nulo, digamos que $a_1 \neq 0$. Assim, $\mathbf{0} = T(v) = a_1T^2(w_1) + \cdots + a_kT^2(w_k) = a_1T(w_1) + \cdots + a_kT(w_k) \Rightarrow a_1 = 0$, um absurdo.

15. Temos que mostrar que $T(C)$ é convexo. Sejam então $w_1, w_2 \in T(C)$. Pela definição de $T(C)$, existem $u_1, u_2 \in C$ tais que $T(u_1) = w_1$ e $T(u_2) = w_2$. Assim, para $t \in [0,1]$, $(1-t)w_1+tw_2 = (1-t)T(u_1)+tT(u_2) = T((1-t)u_1+tu_2) \in T(C)$ pois $(1-t)u_1+tu_2 \in C$, o que mostra que $T(C)$ é convexo.

16. (b) Se $\{u_1, \ldots, u_m\}$ é base de U e $\{w_1, \ldots, w_k\}$ é base de W, mostre que

$$\{(u_1, 0), \ldots, (u_m, 0), (0, w_1), \ldots, (0, w_k)\}$$

é base de $U \times W$. Conclua daí o resultado.

(c) Basta observar que a base acima tem $m + k$ elementos.

(d) É fácil mostrar que T é linear. Agora, construa bases para $\operatorname{Im} T$ e $\operatorname{Ker} T$ para mostrar que $\dim \operatorname{Im} T = \dim U + \dim W$ e $\dim \operatorname{Ker} T = \dim(U \cap W)$. Conclua então o resultado aplicando o Teorema do Núcleo e da Imagem.

17. (a) Primeiramente, note que se $\dim V = 1$ ou $\dim V = 0$, o resultado é claro. Suponha então que $\dim V > 1$. Considere uma base $\{v_1, \ldots, v_n\}$ de V e digamos que $T(v_i) = a_i v_i$ para todo $i = 1, \ldots, n$ (observe que $n > 1$). Fixemos i e j distintos e arbitrários. Temos $T(v_i - v_j) = a_{ij}(v_i - v_j)$. Por outro lado, $T(v_i - v_j) = T(v_i) - T(v_j) = a_i v_i - a_j v_j$. Portanto $a_i v_i - a_j v_j = a_{ij}(v_i - v_j)$ e logo $(a_i - a_{ij})v_i + (a_{ij} - a_j)v_j = 0$. Como v_i e v_j são LI, necessariamente $a_i - a_{ij} = a_{ij} - a_j = 0$, isto é, $a_i = a_j = a_{ij}$. Sendo i e j arbitrários, temos $a_i = a_j = \alpha$ para todos $i, j = 1, \ldots, n$, e logo $T = \alpha Id_V$ como queríamos.

(b) Se T não é da forma αId_V então, pelo item anterior, existe $w \in V$ que não é múltiplo de $T(w) \in V$. Portanto $\{w, T(w)\} \subset V$ é LI.

(c) Siga a dica. No fim, avalie $P \circ T(w)$ e $T \circ P(w)$.

Matrizes de Transformações Lineares

3. A matriz diagonal $[d_{ij}]$ com $d_{ii} = c_i$ para todo i.

4. Seja $[T]^\alpha_\beta$ a matriz de T em bases α e β de V quaisquer. Temos $(\det[T]^\alpha_\beta)^k = \det\left([T]^\alpha_\beta\right)^k = 0$, e logo $\det[T]^\alpha_\beta = 0$. Portanto T não é isomorfismo.

5. Basta definir um isomorfismo entre $\mathbb{R}^3$ e $T_S(2)$, levando uma base de $\mathbb{R}^3$ a uma base de $T_S(2)$ (por exemplo, as bases canônicas desses espaços).

   O mesmo argumento serve para quaisquer dois espaços de mesma dimensão: sejam V e W dois espaços vetoriais com $\dim V = \dim W$. Então suas bases têm mesmo número de elementos, digamos, $\alpha = \{v_1, \ldots, v_n\}$ e $\beta = \{w_1, \ldots, w_n\}$, respectivamente. Basta definir o isomorfismo $T : V \to W$ por $T(v_i) = w_i$ para todo $i = 1, \ldots, n$.

Capítulo 6 – Autovalores e autovetores

1. (a) $T(v) = \alpha v \Rightarrow T^2(v) = T(T(v)) = T(\alpha v) = \alpha^2 v \Rightarrow \cdots \Rightarrow T^n(v) = \alpha^n v$.

(b) 0 é autovalor de $T \Leftrightarrow$ existe $v \neq 0$ tal que $T(v) = 0v = 0 \Leftrightarrow 0 \neq v \in \operatorname{Ker} T \Leftrightarrow \operatorname{Ker} T \neq \{0\} \Leftrightarrow T$ não é injetiva $\Leftrightarrow T$ não é isomorfismo (observe que a última equivalência vale pois se T é injetora se, e somente se T é sobrejetora).

(c) $T(v) = \alpha v$ onde, pelo item anterior, $\alpha \neq 0$. Então, $T(v) = \alpha v \Leftrightarrow T^{-1} \circ T(v) = \alpha T^{-1}(v) \Leftrightarrow T^{-1}(v) = \frac{1}{\alpha} v$.

2. Basta notar que $T(S(v)) = T \circ S(v) = S \circ T(v) = S(T(v)) = S(\lambda v) = \lambda S(v)$.

4. Primeiramente,

$$p_{\mathbf{AB}}(0) = \det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B} = \det(\mathbf{BA}) = p_{\mathbf{BA}}(0),$$

o que mostra que 0 é autovalor de **AB** se, e somente se 0 é autovalor de **BA**. Suponha agora que $\alpha \neq 0$ é autovalor de **AB**, e mostremos que α também é autovalor de **BA**. Existe $\mathbf{v}$ tal que

$$\mathbf{ABv} = \alpha\mathbf{v} \quad \Rightarrow \quad \mathbf{BA}(\mathbf{Bv}) = \alpha\mathbf{Bv}.$$

Ora, $\mathbf{Bv} \neq \mathbf{0}$ pois se fosse $\mathbf{Bv} = \mathbf{0}$ teríamos $\mathbf{ABv} = \mathbf{0} = 0\mathbf{v} \Rightarrow \alpha = 0$. Segue então da expressão $\mathbf{BA}(\mathbf{Bv}) = \alpha\mathbf{Bv}$ que α é autovalor de **BA**. A recíproca se mostra de maneira análoga.

5. Observe que $(\mathbf{A} - \lambda\mathbf{I})^t = \mathbf{A}^t - \lambda\mathbf{I}^t = \mathbf{A}^t - \lambda\mathbf{I}$. Assim $\det(\mathbf{A} - \lambda\mathbf{I}) = \det(\mathbf{A} - \lambda\mathbf{I})^t = \det(\mathbf{A}^t - \lambda\mathbf{I})$.

6. Seja **A** e **D** tais que $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$ é matriz diagonal. Temos

$$\begin{aligned} p_{\mathbf{D}}(\lambda) &= \det(\mathbf{D} - \lambda\mathbf{I}) \\ &= \det(\mathbf{P}^{-1}\mathbf{AP} - \lambda\mathbf{P}^{-1}\mathbf{P} \\ &= \det(\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}) \\ &= \det \underbrace{\mathbf{P}^{-1} \det \mathbf{P}}_{=1} \det(\mathbf{A} - \lambda\mathbf{I}) = p_{\mathbf{A}}(\lambda). \end{aligned}$$

Ou seja, os polinômios característicos de **A** e **D** são iguais, e logo os autovalores dessas matrizes são os mesmos. Como **D** é matriz diagonal, sua diagonal é precisamente formada pelos autovalores.

Para mostrar que a coluna i de **P** é autovetor de **A** associado ao autovalor λ_i, primeiro note que se considerarmos a matriz coluna

$$\mathbf{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

com 1 na posição i e zero no resto, então

$$\lambda_i\mathbf{e}_i = \mathbf{De}_i \quad \text{e} \quad \mathbf{Pe}_i = \text{“coluna } i \text{ de } \mathbf{P}\text{”}.$$

Assim,

$$\lambda_i\mathbf{e}_i = \mathbf{De}_i = \mathbf{P}^{-1}\mathbf{APe}_i \quad \Rightarrow \quad \mathbf{A}(\mathbf{Pe}_i) = \lambda(\mathbf{Pe}_i),$$

isto é, a coluna i de **P** é autovetor de **A** associado ao autovalor λ_i.

7. Se 1 não é autovalor de **A** então $p(1) = \det(\mathbf{A} - \mathbf{I}) \neq 0$. Ou seja, $\mathbf{A} - \mathbf{I}$ possui inversa.

Capítulo 7 – Diagonalização de operadores

1.

(a) autoval: $-2, -1, 2$

(b) autoval: $5, 1$

(c) autoval: 1, autovet: qualquer $p \neq \mathbf{0}$

(d) autoval: $-2, -1, 1$

2. (a) Sim (b) Sim (c) Sim (d) Sim

3. (a) Sim (b) Sim (c) Não (d) Sim (e) Sim

6. Como o contra-domínio de T é V, temos $\operatorname{Im} T \subset V$. Pela definição de núcleo, vale $\operatorname{Ker} T \subset V$. Então $\operatorname{Im} T + \operatorname{Ker} T \subset V$. Vamos mostrar a inclusão oposta $V \subset \operatorname{Im} T + \operatorname{Ker} T$.

   Seja $v \in V$. Como T é diagonalizável, existe uma base $\{v_1, \ldots, v_n\}$ de V formada por autovetores de T, digamos, $T(v_i) = \lambda_i v_i$, $i = 1, \ldots, n$. Para simplificar, podemos supor que os autovalores não nulos são os r primeiros, $\lambda_1, \ldots, \lambda_r$ ($r \leq n$). Escrevemos v na base de autovetores, separando autovetores associados à autovalores não nulos daqueles associados à zero:

$$v = (a_1 v_1 + \cdots + a_r v_r) + (a_{r+1} v_{r+1} + \cdots + a_n v_n). \tag{A.1}$$

   Veja que $w \in \operatorname{Ker} T \Leftrightarrow T(w) = \mathbf{0} \Leftrightarrow T(w) = 0w \Leftrightarrow w$ é autovetor de T associado a zero. Por outro lado, cada autovetor u associado a autovalores não nulos são elementos de $\operatorname{Im} T$, pois $\lambda u = T(u) \Rightarrow u = \frac{1}{\lambda} T(u) = T(\frac{1}{\lambda} u) \in \operatorname{Im} T$. Assim, a escrita (A.1) diz que $v \in \operatorname{Im} T + \operatorname{Ker} T$, e logo $V \subset \operatorname{Im} T + \operatorname{Ker} T$. Concluímos então que $V = \operatorname{Im} T + \operatorname{Ker} T$.

   Para finalizar, aplicamos o Teorema do Núcleo e da Imagem ($\dim V = \dim \operatorname{Im} T + \dim \operatorname{Ker} T$) para concluir que só pode ser $V = \operatorname{Im} T \oplus \operatorname{Ker} T$.

Referências Bibliográficas

[1] Howard Anton e Chris Rorres. *Álgebra Linear com aplicações*. Bookman, 10 edition, 2010.

[2] José Luiz Boldrini e outros. *Álgebra Linear*. Harper & Row do Brasil, São Paulo, 3 edition, 1980.

[3] Alfredo Steinbruch e Paulo Winterle. *Álgebra Linear*. Pearson, São Paulo, 2 edition, 1987.

[4] David Lay. *Álgebra Linear*. LTC, Rio de Janeiro, 2 edition, 1999.

[5] David Poole. *Álgebra linear*. Thonsom Learning, São Paulo, 2006.