Introdução à Álgebra Linear

Caio Tomás

January 24, 2020

Conteúdo

1 Introdução

Esse arquivo consiste das notas de aula tomadas durante o curso de Introdução à Álgebra Linear realizado em 2019/2 na UnB e ministrado por Pavel Zalesski[1]. Foi utilizado o livro *Introdução à Álgebra Linear, Editora PHB - C.H. Edwards, Jr. e D.E. Penney.*

2 Matrizes, sistemas lineares e determinantes

Definição. *Um sistema de m equações lineares com n incógnitas tem a seguinte forma:*

$$\begin{cases} a_{11}x_1 + a_{12}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}, \text{ onde } a_{ij}, b \in \mathbb{R} \text{ e } x_1, x_2, \ldots, x_n \text{ são incógnitas.} \tag{1}$$

A matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

é dita **matriz associada** de (1).

Exemplo. *A matriz associada de*

$$\begin{cases} 2x_1 + 3x_2 - 7x_3 + 4x_4 = 6 \\ x_2 + 3x_3 - 5x_4 = 0 \\ -x_1 + 2x_2 - 9x_4 = 17 \end{cases}$$

é

$$\begin{bmatrix} 2 & 3 & -7 & 4 & 6 \\ 0 & 1 & 3 & -5 & 0 \\ -1 & 2 & 0 & -9 & 17 \end{bmatrix}$$

Para resolver sistemas lineares, usamos operações de linha, que são:

1. Multiplicar uma linha por um número
2. Trocar duas linhas
3. Adicionar um múltiplo de uma linha a um múltiplo de outra linha

Fato. *Operações de linha não alteram o conjunto-solução do sistema.*

Além disso, podemos efetuar as operações de linha na matriz associada do sistema.

Exemplo. *Vamos resolver o sistema*

$$\begin{cases} x_1 + 2x_2 + x_3 = 6 \\ 3x_1 + 8x_2 + 7x_3 = 20 \\ 2x_1 + 7x_2 + 9x_3 = 32 \end{cases}$$

[1]Contato: p.zalesski@mat.unb.br (61) - 3107-6438 Sala: A1-414/06

Escrevendo a matriz associada e realizando as operações de linha, temos:

$$\begin{bmatrix} 1 & 2 & 1 & 4 \\ 3 & 8 & 7 & 20 \\ 2 & 7 & 9 & 23 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 4 \\ 0 & 2 & 4 & 8 \\ 0 & 3 & 7 & 15 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 4 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 7 & 15 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 4 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Logo, nosso sistema se torna

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

Definição. *Uma matriz é dita* ***escalonada*** *se:*

1. *Todas as linhas nulas ficam abaixo das não nulas*
2. *O primeiro elemento não nulo de cada linha tem somente zeros abaixo dele*

Os primeiros elementos não nulos de cada linha são ditos **líderes**. As incógnitas correspondentes também são ditas **líderes**. As outras incógnitas são ditas **livres**.

Exemplo. *A matriz*

$$E = \begin{bmatrix} 2 & -1 & 0 & 4 & 7 \\ 0 & 1 & 2 & 0 & -5 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

está escalonada, e seus elementos líderes são 2, 1 e 3.

Com isso, podemos formalizar um algoritmo para resolver um sistema na **forma escalonada**:

1. Atribuir parâmetros às incógnitas livres
2. Resolver a última equação para a incógnita líder
3. Substituir o resultado na penúltima equação e resolver para a incógnita líder
4. Repetir o procedimento até a primeira equação

Exemplo. *O sistema*

$$\begin{cases} x_1 - 2x_2 + 3x_3 + 2x_4 + x_5 = 10 \\ x_3 + 2x_5 = -3 \\ x_4 - 4x_5 = 7 \end{cases}$$

tem matriz associada

$$A = \begin{bmatrix} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 1 & 0 & 2 & -3 \\ 0 & 0 & 0 & 1 & -4 & 7 \end{bmatrix} \text{ (} x_2 \text{ e } x_5 \text{ são as incógnitas livres)}$$

que já está escalonada. Fazendo $x_2 = s$ e $x_5 = t$, com $s, t \in \mathbb{R}$, obtemos:

$$\begin{cases} x_1 = 5 + 2s - 3t \\ x_2 = s \\ x_3 = -3 - 2t \\ x_4 = 7 + 4t \\ x_5 = t \end{cases}$$

Mas como escalonamos uma matriz? Para isso, usamos o **Algoritmo de Eliminação de Gauss**:

1. Encontre a primeira coluna não nula
2. Encontre o primeiro elemento não nulo dessa coluna e troque a linha desse elemento com a primeira linha
3. Agora, use esse elemento para zerar todos os elementos abaixo dele

$$A = \begin{bmatrix} 0 & \alpha & \star & \star & \star & \cdots & \star \\ 0 & 0 & & & & & \\ \vdots & \vdots & & A_1 & & & \\ 0 & 0 & & & & & \\ 0 & 0 & & & & & \end{bmatrix}$$

4. Execute os passos anteriores para A_1
5. Continue até escalonar

Exemplo. *Vamos escalonar a matriz associada A de*

$$\begin{cases} x_1 - 2x_2 + 3x_3 + 2x_4 + x_5 = 10 \\ 2x_1 - 4x_2 + 8x_3 + 3x_4 + 10x_5 = 7 \\ 3x_1 - 6x_2 + 10x_3 + 6x_4 + 5x_5 = 27 \end{cases}$$

Temos que

$$A = \begin{bmatrix} 1 & -2 & 3 & 2 & 1 & 10 \\ 2 & -4 & 8 & 3 & 10 & 7 \\ 3 & -6 & 10 & 6 & 5 & 27 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 2 & -1 & 8 & -13 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 2 & -1 & 8 & -13 \\ 0 & 0 & 0 & \frac{1}{2} & -2 & \frac{7}{2} \end{bmatrix}$$

Definição. *Uma matriz escalonada é dita escalonada **reduzida** se:*

1. *Cada elemento líder é 1*
2. *O elemento líder é o único elemento não nulo na coluna*

Exemplo. *As matrizes*

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix}$$

estão escalonadas reduzidas, enquanto que as matrizes

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

estão escalonadas, mas não reduzidas.

Para escalonar e reduzir uma matriz, utilizamos o **Algoritmo de Eliminação de Gauss-Jordan**:

1. Execute o **Algoritmo de Eliminação de Gauss** para escalonar a matriz
2. Divida cada linha não nula pelo elemento líder
3. Use combinações lineares de linhas para zerar tudo acima dos elementos líderes

Exemplo. *Vamos escalonar e reduzir A.*

$$A = \begin{bmatrix} 1 & 2 & 1 & 4 \\ 3 & 8 & 7 & 20 \\ 2 & 7 & 9 & 23 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 4 \\ 0 & 2 & 4 & 8 \\ 0 & 3 & 7 & 15 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 4 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 7 & 15 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 4 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Teorema 2.1. *Um sistema linear de equações ou tem uma única solução, ou não tem nenhuma solução, ou tem infinitas soluções.*

Demonstração. Aplicando a eliminação de Gauss-Jordan ao sistema

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

e chamando as variáveis líderes de $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$, obtemos

$$\begin{cases} x_{j_1} + \sum c_{1j}x_i = d_1 \\ x_{j_2} + \sum c_{2j}x_i = d_2 \\ \qquad \vdots \\ x_{j_r} + \sum c_{rj}x_j = d_r \\ 0 = d_{r+1} \\ \ \vdots \\ 0 = d_m \end{cases}$$

em que os somatórios envolvem apenas variáveis livres. Note que o sistema só tem solução se $d_r, d_{r+1}, \ldots, d_m$ são todos nulos. Nesse caso, se $r < n$, então podemos atribuir parâmetros a algumas variáveis livres e, por isso, o sistema é possível e indeterminado (tem infinitas soluções). Se $r = n$, então o sistema é possível e determinado, isto é, há apenas uma solução, a saber, $x_i = d_i$ com $i = 1, 2, \ldots, n$. ☐

Definição. *Um sistema linear é dito* ***homogêneo*** *se tem a forma*

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

Note que (2) sempre tem **pelo menos** uma solução, a saber, a trivial ($x_i = 0, \forall i$). Se (2) tem mais incógnitas que equações, então há infinitas soluções.

Exemplo. *O sistema*

$$\begin{cases} 47x_1 - 73x_2 + 56x_3 + 21x_4 = 0 \\ 19x_1 + 81x_2 - 17x_3 - 99x_4 = 0 \\ 53x_1 + 62x_2 + 39x_3 + 25x_4 = 0 \end{cases}$$

tem infinitas soluções, pois há mais incógnitas que equações.

Definição. *A matriz $n \times n$*

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

é dita ***matriz identidade*** *de ordem n e denotada por I_n.*

Supondo, em (2), que $m = n$, a matriz associada é quadrada $n \times n$. Nesse caso, a única solução é a trivial e aplicando o **Algoritmo de Eliminação de Gauss-Jordan** à matriz associada (com exceção da última coluna), obtemos a matriz

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

ou seja, a matriz identidade de ordem n.

Observação. *Uma matriz A tem uma* ***única*** *matriz escalonada reduzida associada a ela.*

Exemplo. *A matriz associada ao sistema*

$$\begin{cases} x_1 + 2x_2 + x_3 = 0 \\ 3x_1 + 8x_2 + 7x_3 = 0 \\ 2x_1 + 7x_2 + 9x_3 = 0 \end{cases}$$

é (retirando a última coluna, pois ela não altera o escalonamento)

$$B = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 7 \\ 2 & 7 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 4 \\ 0 & 3 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 3 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Acrescentando a coluna de zeros, obtemos

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

logo a única solução do sistema é a trivial.

2.1 Operações com matrizes

As matrizes $A = (a_{ij}), m \times n$ e $B = (b_{ij}), k \times e$ são iguais se têm o mesmo tamanho, i.e., $m = k$ e $n = e$ e se $a_{ij} = b_{ij}, \forall i, j \in \mathbb{N}$.

A soma $A + B$ está definida se A e B têm mesmo tamanho e, neste caso, $A + B = (a_{ij} + b_{ij})$.

Exemplo. *Se*

$$A = \begin{bmatrix} 3 & 0 & -1 \\ 2 & -7 & 5 \end{bmatrix} \text{ e } B = \begin{bmatrix} 4 & -3 & 6 \\ 9 & 0 & -2 \end{bmatrix}$$

então

$$A + B = \begin{bmatrix} 7 & -3 & 5 \\ 11 & -7 & 3 \end{bmatrix}$$

Definição. *Para $c \in \mathbb{R}$, podemos definir $c \cdot A = (c \cdot a_{ij})$. Se $c = -1$, escrevemos $-A$ e definimos $B - A = B + (-A)$.*

Exemplo. *Com as matrizes do exemplo anterior, obtemos*

$$3A = \begin{bmatrix} 9 & 0 & -3 \\ 6 & -21 & 15 \end{bmatrix} \text{ e } 3A - B = \begin{bmatrix} 5 & 3 & -9 \\ -3 & -21 & 17 \end{bmatrix}$$

Definição. *Sejam $A = (a_{ij})$ $m \times p$ e $B = (b_{ij})$ $p \times n$. Então, $\underset{(m \times n)}{AB} = C = (c_{ij})$, com $c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$.*

Exemplo. *Note que*

$$A = \begin{bmatrix} 2 & -1 \\ -4 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & 5 \\ 3 & 7 \end{bmatrix} \Rightarrow AB = \begin{bmatrix} -1 & 3 \\ 5 & 1 \end{bmatrix}, BA = \begin{bmatrix} -18 & 14 \\ -22 & 18 \end{bmatrix}$$

ou seja, o produto de matrizes, em geral, ***não é comutativo****.*

Exemplo. *Usando a definição de produto de matrizes, podemos escrever o sistema (1) na seguinte forma:*

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}}_{B}$$

$AX = B$ *é dita* ***forma matricial*** *do sistema*

Observação. *É comum escrever a matriz das incógnitas com uma seta em cima ($\vec{X}$), representando um vetor.*

Exemplo. *A forma matricial do sistema*

$$\begin{cases} 3x_1 - 4x_2 + x_3 + 7x_4 = 10 \\ 4x_1 - 5x_3 + 2x_4 = 0 \\ x_1 + 9x_2 + 2x_3 - 6x_4 = 5 \end{cases} \quad \text{é} \quad \begin{bmatrix} 3 & -4 & 1 & 7 \\ 4 & 0 & -5 & 2 \\ 1 & 9 & 2 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 10 \\ 0 \\ 5 \end{bmatrix}$$

Observação. *Sejam*

$$A = \begin{bmatrix} 4 & 1 & -2 & 7 \\ 3 & 1 & -1 & 5 \end{bmatrix}, B = \begin{bmatrix} 1 & 5 \\ 3 & -1 \\ -2 & 4 \\ 2 & 3 \end{bmatrix}, C = \begin{bmatrix} 3 & 4 \\ 2 & 1 \\ -2 & 3 \\ 1 & -3 \end{bmatrix}$$

Então, temos

$$AB = \begin{bmatrix} 25 & -10 \\ 18 & -5 \end{bmatrix}, BA = \begin{bmatrix} 25 & -10 \\ 18 & -5 \end{bmatrix}$$

Percebemos então que $AB = AC \not\Rightarrow B = C$, em geral.

Definição. *A matriz cujas entradas são todas nulas é dita* ***matriz nula*** *e é denotada por* ***0*** *ou simplesmente* 0.

2.2 Propriedades de operações com matrizes

1. $A + B = B + A$ (comutatividade da soma);
2. $(A + B) + C = A + (B + C)$ (associatividade da soma);
3. $(AB)C = A(BC)$ (associatividade do produto);
4. $(A+B)C = AC+BC$ e $A(B+C) = AB+AC$ (distributividade do produto à direita e à esquerda);
5. $0 + A = A$, para qualquer matriz A (0 é o elemento neutro da soma);
6. Seja $A = (a_{ij})$ uma matriz $p \times q$. Então, $I_pA = AI_q = A$. (I_n é o elemento neutro da multiplicação).

Demonstração. **1.** Sejam $A = (a_{ij}), B = (b_{ij})$ tais que a soma está definida. Da definição, segue que $A + B = (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = B + A$.

2. Sejam $A = (a_{ij}), B = (b_{ij}), C = (c_{ij})$ tais que a soma está definida. Da definição, segue que $(A + B) + C = ((a_{ij} + b_{ij}) + c_{ij}) = (a_{ij} + (b_{ij} + c_{ij})) = A + (B + C)$.

3. Sejam $A_{m\times n} = (a_{ij}), B_{n\times p} = (b_{ij}), C_{p\times q} = (c_{ij})$. Daí, temos

$$\begin{aligned}
(AB)C &= \left(\sum_{j=1}^{p}\left[\left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)c_{jr}\right]\right)\\
&= \left(\sum_{j=1}^{p}\left[\sum_{k=1}^{n} a_{ik}b_{kj}c_{jr}\right]\right)\\
&= \left(\sum_{k=1}^{n}\left[\left(\sum_{j=1}^{p} b_{kj}c_{jr}\right)a_{ik}\right]\right)\\
&= A(BC)
\end{aligned}$$

4. Vamos demonstrar essa propriedade apenas à esquerda. A demonstração à direita é análoga. Sejam $A = (a_{ij}), B = (b_{ij}), C = (c_{ij})$ tais que a soma e o produto estão definidos e $D = A(B + C)$. Temos

$$\begin{aligned}
d_{ij} &= \sum_{k=1}^{p} a_{ik}(b_{kj} + c_{kj})\\
&= \sum_{k=1}^{p} a_{ik}b_{kj} + \sum_{k=1}^{p} a_{ik}c_{kj}\\
&= AB + AC
\end{aligned}$$

5. Segue da definição de soma.

6. Seja $A = (a_{ij})$ uma matriz $p \times q$. Da definição de produto, segue que $I_qA = (c_{ij})$ sendo

$$c_{ij} = \sum_{k=1}^{q} a_{ik}b_{kj}$$

Como b_{kj} são elementos da matriz identidade, eles são nulos para $k \neq j$ e 1 para $k = j$. Daí, para cada j, ou seja, para cada coluna, $c_{ij} = a_{ik} \cdot 1 = a_{ik}$, logo $I_qA = A$. A demonstração para AI_p é análoga. □

Observação. *Para todo $c \in \mathbb{R}, c \cdot A = C \cdot A$, sendo*

$$C = \begin{bmatrix} c & 0 & 0 & \cdots & 0\\ 0 & c & 0 & \cdots & 0\\ 0 & 0 & c & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & 0\\ 0 & 0 & 0 & \cdots & c \end{bmatrix} = c \cdot I_n$$

2.3 Matriz inversa

Considere que todas as matrizes a seguir são quadradas.

Definição. *Dizemos que $A_{n\times n}$ é **inversível** se $\exists B | AB = BA = I_n$. Nesse caso, B é a **inversa** de A e é denotada por A^{-1}.*

Proposição. *Se A é inversível, então A^{-1} é única.*

Demonstração. Suponha que B e C são inversas de A. Da definição, segue que

$$BAC = BI_n = B \ \wedge \ BAC = I_nC = C$$
$$\therefore B = C$$

□

Proposição. *Tome* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *A é inversível se, e só se,* $ad - bc \neq 0$ *e, neste caso,* $A^{-1} = \frac{1}{ad-bc} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Demonstração. ($\Rightarrow$) Se A é inversível, então existe $B = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ tal que

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = I_2 \Rightarrow \begin{cases} ax + bz = 1 \\ ay + bw = 0 \\ cx + dz = 0 \\ cy + dw = 1 \end{cases} \Rightarrow \begin{cases} x = \dfrac{d}{ad-bc} \\ y = -\dfrac{b}{ad-be} \\ z = -\dfrac{c}{ad-bc} \\ w = \dfrac{a}{ad-bc} \end{cases} \Rightarrow B = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

($\Leftarrow$) Note que

$$\frac{1}{ad-bc} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad-bc} \cdot \begin{bmatrix} da-bc & 0 \\ 0 & ad-bc \end{bmatrix} = I_2$$
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{ad-bc} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = I_2$$

□

Exemplo.

$$A = \begin{bmatrix} 4 & 6 \\ 5 & 9 \end{bmatrix} \Rightarrow A^{-1} = \frac{1}{6} \begin{bmatrix} 9 & -6 \\ -5 & 4 \end{bmatrix}$$

2.4 Propriedades da inversa

1. $(A^{-1})^{-1} = A$
2. $(AB)^{-1} = B^{-1}A^{-1}$ e, mais geralmente, $(A_1A_2 \cdots A_n)^{-1} = A_n^{-1}A_{n-1}^{-1} \cdots A_1^{-1}$

Demonstração. **1.** Da definição de inversa, sabemos que $AA^{-1} = A^{-1}A = I_n$, logo A é a inversa de A^{-1}.

2. Note que

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

logo $(AB)^{-1} = B^{-1}A^{-1}$. □

Definição. *Dizemos que a matriz E é* ***elementar*** *se E é obtida da matriz identidade através de uma operação de linha.*

Observação. *Realizar uma operação de linha em uma matriz é equivalente a multiplicar à esquerda pela matriz elementar correspondente.*

Exemplo.

$$\begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c^2 & cd \end{bmatrix}$$
$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ f & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ fa+c & fb+d \end{bmatrix}$$

Fato. *Se A é inversível, então $E_1E_2\cdots E_kA = I_n$. Daí, $A = E_k^{-1}E_{k-1}^{-1}\cdots E_1^{-1}$ e, consequentemente, $A^{-1} = E_1E_2\cdots E_kI_n$.*

Esse fato nos dá o seguinte método para inverter matrizes através de operações de linha:

Exemplo.

$$\left[\begin{array}{ccc|ccc} 4 & 3 & 2 & 1 & 0 & 0 \\ 5 & 6 & 3 & 0 & 1 & 0 \\ 3 & 5 & 2 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 4 & 3 & 2 & 1 & 0 & 0 \\ 1 & 3 & 1 & -1 & 1 & 0 \\ 3 & 5 & 2 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & -1 & 1 & 0 \\ 4 & 3 & 2 & 1 & 0 & 0 \\ 3 & 5 & 2 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & -1 & 1 & 0 \\ 0 & -9 & -2 & 5 & -4 & 0 \\ 0 & -4 & -1 & 3 & -3 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & -1 & 1 & 0 \\ 0 & -9 & -2 & 5 & -4 & 0 \\ 0 & -8 & -2 & 6 & -6 & 2 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & -1 & 1 & 0 \\ 0 & -9 & -2 & 5 & -4 & 0 \\ 0 & 1 & 0 & 1 & -2 & 2 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 & -2 & 2 \\ 0 & -9 & -2 & 5 & -4 & 0 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & -4 & 7 & -6 \\ 0 & 1 & 0 & 1 & -2 & 2 \\ 0 & 0 & -2 & 14 & -22 & 18 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & -4 & 7 & -6 \\ 0 & 1 & 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -7 & 11 & 9 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -4 & 3 \\ 0 & 1 & 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -7 & 11 & 9 \end{array}\right]$$

$$\therefore \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -4 & 3 \\ 0 & 1 & 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -7 & 11 & 9 \end{array}\right] \textit{ é a inversa de } \left[\begin{array}{ccc|ccc} 4 & 3 & 2 & 1 & 0 & 0 \\ 5 & 6 & 3 & 0 & 1 & 0 \\ 3 & 5 & 2 & 0 & 0 & 1 \end{array}\right]$$

2.5 Determinantes

Considere todas as matrizes a seguir quadradas.

1. Se $A = (a)$, $det(A) = |A| = a$;
2. Se $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $det(A) = ab - cd$.

Definição. *O* ***menor*** *M_{ij} da matriz $A_{n\times n}$ é o determinante da submatriz obtida de A através da eliminação da i-ésima linha e j-ésima coluna.*

Definição. *O determinante de $A = (a_{ij})$ é definido por*

$$|A| = a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{n+1}a_{1n}M_{1n} = \sum_{j=1}^{n}(-1)^{j+1}a_{1j}M_{1j} \textit{ (expansão pela primeira linha)}$$

Exemplo.

$$\begin{vmatrix} 5 & -2 & -3 \\ 4 & 0 & 1 \\ 3 & -1 & 2 \end{vmatrix} = 5\begin{vmatrix} 0 & 1 \\ -1 & 2 \end{vmatrix} - 2(-1)\begin{vmatrix} 4 & 1 \\ 3 & 2 \end{vmatrix} - 3\begin{vmatrix} 4 & 0 \\ 3 & -1 \end{vmatrix} = 5 + 2(5) - 3(-4) = 27$$

Exemplo.

$$\begin{vmatrix} 2 & 0 & 0 & -3 \\ 0 & -1 & 0 & 0 \\ 7 & 4 & 3 & 5 \\ -6 & 2 & 2 & 4 \end{vmatrix} = 2\begin{vmatrix} -1 & 0 & 0 \\ 4 & 3 & 5 \\ 2 & 2 & 4 \end{vmatrix} + 3\begin{vmatrix} 0 & -1 & 0 \\ 7 & 4 & 3 \\ -6 & 2 & 2 \end{vmatrix} = 2(-1)\begin{vmatrix} 3 & 5 \\ 2 & 4 \end{vmatrix} + 3(-1)(-1)\begin{vmatrix} 7 & 3 \\ -6 & 2 \end{vmatrix} = -2(2) + 3(32) = 92$$

2.6 Propriedades dos determinantes

1. Se $B = c \cdot A$, então $det(B) = c \cdot det(A)$;
2. Trocando duas linhas (ou colunas) de A, obtemos A' e temos $det(A') = -det(A)$;
3. Trocar uma linha de A por uma combinação linear de linhas de A não altera o determinante

4. $det(AB) = det(A)det(B)$

Corolário 2.1.1. *Segue da propriedade 2 que a expansão do determinante pode ser feita escolhendo-se qualquer linha ou coluna da matriz. Daí, temos*

$$|A| = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}M_{ij} = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}M_{ij}$$

Note que é mais vantajoso, para o cálculo do determinante, escolher uma linha ou coluna com o maior número de zeros possível. Segue do corolário 2.1.1 que se a matriz possui uma linha ou coluna de zeros, então $det(A) = 0$.

Definição. *Uma matriz é dita* ***triangular*** *se possui apenas zeros abaixo ou acima da diagonal principal (triangular superior e inferior).*

Exemplo.

$$\begin{vmatrix} 3 & 6 & 10 \\ 0 & 5 & 8 \\ 0 & 0 & 7 \end{vmatrix} = 7\begin{vmatrix} 3 & 6 \\ 0 & 5 \end{vmatrix} = 105$$

$$\begin{vmatrix} 1 & 0 & 0 \\ 3 & 7 & 0 \\ 4 & 6 & 5 \end{vmatrix} = 1\begin{vmatrix} 7 & 0 \\ 6 & 5 \end{vmatrix} = 35$$

Note que se A é triangular, $det(A)$ é o produto dos elementos da diagonal principal. Além disso, uma matriz com duas linhas ou colunar proporcionais tem determinante nulo, visto que podemos anular uma dessas linhas ou colunas através de uma combinação linear.

Definição. *A matriz* ***transposta*** A^T *de* A *é a matriz obtida trocando as linhas de* A *pelas colunas correspondentes.*

Exemplo.

$$\begin{bmatrix} 7 & -2 & 6 \\ 1 & 2 & 3 \\ 5 & 0 & 4 \end{bmatrix}^T = \begin{bmatrix} 7 & 1 & 5 \\ -2 & 2 & 0 \\ 6 & 3 & 4 \end{bmatrix}$$

2.7 Propriedades da transposta

1. $(A^T)^T = A$;
2. $(AB)^T = B^TA^T$;
3. $|A| = |A^T|$

Demonstração. **1.** Segue da definição de transposta;

2. Sejam A e B matrizes $n \times n$. Da definição de produto de matrizes, sabemos que:

$$(AB)^T = \left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)^T = \left(\sum_{k=1}^{n} a_{jk}b_{ki}\right)$$

Por outro lado, também sabemos que:

$$B^TA^T = \left(\sum_{k=1}^{n} b'_{ik}a'_{kj}\right) = \left(\sum_{k=1}^{n} b_{ki}a_{jk}\right)$$

3. Pelo corolário 2.1.1, podemos escolher qualquer linha ou coluna de A para o cálculo do determinante. Como escolher uma linha de A equivale a escolher uma coluna de A^T e vice-versa, $|A| = |A^T|$. □

Exemplo.

$$\begin{vmatrix} 7 & 6 & 0 \\ 9 & -3 & 2 \\ 4 & 5 & 0 \end{vmatrix} = -2 \begin{vmatrix} 7 & 6 \\ 4 & 5 \end{vmatrix} = -22$$

Teorema 2.2. *A é inversível se, e só se, $det(A) \neq 0$.*

Demonstração. Sabemos que A é inversível se, e só se, $E_1 E_2 \cdots E_k A = I_n$. Note que as operações de linha, i.e., as matrizes elementares $E_1, E_2, \ldots, E_k$ alteram, na pior das hipóteses, o sinal de $det(A)$. Logo, A é inversível $\Leftrightarrow E_1 E_2 \cdots E_k A = I_n \Leftrightarrow \det(A) \neq 0$. □

Para matrizes 3×3, podemos calcular o determinante da seguinte forma:

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

Exemplo.

$$\begin{vmatrix} 2 & -3 & -4 \\ -1 & 4 & 2 \\ 3 & 10 & 1 \end{vmatrix} = 8 - 18 + 40 + 48 - 3 - 40 = 35$$

Para o cálculo do determinante matrizes de ordem superior, vamos tentar triangular a matriz através de operações de linha.

Exemplo.

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

$$= -2 \begin{vmatrix} 1 & 1 & -1 & 0 \\ 0 & -2 & 6 & -1 \\ 0 & 4 & -1 & 4 \\ 3 & 1 & 3 & -1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 1 & -1 & 0 \\ 0 & -2 & 6 & -1 \\ 0 & 0 & 11 & 2 \\ 0 & 0 & 17 & -2 \end{vmatrix} = -2 \cdot 1 \cdot (-2) \cdot \begin{vmatrix} 11 & 2 \\ 17 & -2 \end{vmatrix} = 4(-22 - 34) = -224$$

3 Espaços vetoriais

Definição. *Um conjunto V com duas operações, soma $(+)$ e multiplicação por escalar $(\cdot)$, chama-se espaço vetorial se satisfaz as seguintes propriedades, dados $u, v, w \in V$ e $a, b, c \in \mathbb{R}$:*

1. $w + v = v + w$ *(comutatividade da soma)*
2. $(u + v) + w = u + (v + w)$ *(associatividade da soma)*
3. $\exists 0 : 0 + v = v + 0 = v, \forall v \in V$ *(existência do elemento neutro da soma)*
4. $\forall v \in V, \exists w \in V : w + v = 0$ $(w = -v)$ *(existência do elemento oposto)*
5. $c(v + w) = cv + cw$ *(distributividade)*
6. $(a + b)v = av + bv$ *(distributividade)*
7. $(ab)v = a(bv)$ *(associatividade do produto)*
8. $1 \cdot v = v, \forall v \in V$ *(elemento neutro da multiplicação)*

*Os elementos de V são chamados **vetores**.*

Exemplo. *O conjunto* $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) | x_i \in \mathbb{R}\}$ *é um espaço vetorial. Note que*

$$\begin{aligned} (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) &\coloneqq (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \\ c(x_1, x_2, \ldots, x_n) &\coloneqq (cx_1, cx_2, \ldots, cx_n) \\ 0 &\coloneqq (0, 0, \ldots, 0) \\ (-x_1, -x_2, \ldots, -x_n) &\text{ é o oposto de } (x_1, x_2, \ldots, x_n) \end{aligned}$$

Exemplo. *O conjunto* $M_{m\times n}$ *de todas as matrizes* $m \times n$ *é um espaço vetorial*

Exemplo. *O conjunto* τ *de funções* $f : \mathbb{R} \to \mathbb{R}$*, sendo*

$$\begin{aligned} (f + g)(x) &\coloneqq f(x) + g(x) \\ (c \cdot f)(x) &\coloneqq c \cdot f(x) \\ 0 : f(x) &\equiv 0 \end{aligned}$$

Definição. *Um subconjunto* W *do espaço vetorial* V *chama-se* ***subespaço*** *se* W *é um espaço vetorial com as operações de* V. W *é subespaço de* V *se, e só se,* W *é fechado sob as operações de* V*, i.e.:*

1. $v, w \in W \Rightarrow v + w \in W$
2. $v \in W \Rightarrow cv \in W, \forall c \in \mathbb{R}$

Observação. *Note que o vetor nulo sempre está em um subespaço* W *de* V*, pois se* $w \in W$ *então* $-w \in W$ *(propriedade 4) e, portanto,* $w + (-w) = 0 \in W$ *(propriedade 1).*

Além disso, V *é um subespaço de* V*. Os subespaços de* V *diferentes de* $\{0\}$ *e* V *são ditos* ***subespaços próprios****.*

Exemplo. *Seja* A *uma matriz* $m \times n$*. Então, o conjunto solução* W *do sistema* $AX = 0$ *é um subespaço de* $\mathbb{R}^n$*.*

Demonstração. Sejam $x, y \in W$. Note que $A(x + y) = Ax + Ay = 0 + 0 = 0$, logo $x + y$ é solução do sistema e, portanto, $x + y \in W$.

Agora, sejam $x \in W$ e $c \in \mathbb{R}$. Note que $A(cx) = c(Ax) = c \cdot 0 = 0, \forall c \in \mathbb{R}$. Logo, cx é solução do sistema e, portanto, $cx \in W$.

$$\therefore W \text{ é subespaço de } \mathbb{R}^n$$

□

W *é dito* ***espaço-solução*** *do sistema.*

Exemplo. *Seja* $V = M_{m\times n}$ *e* W *o conjunto das matrizes de* V *com* $a_{ij} = 0$ *para* i, j *fixos.* W *é subespaço de* V*? Para isso, vamos verificar se as propriedades valem.*

Demonstração. Sejam $A, B \in W$. Note que a matriz $A + B$ tem o $i \times j$-ésimo elemento dado por $a_{ij} + b_{ij} = 0 + 0 = 0$. Logo, o $i \times j$-ésimo elemento de $A + B$ é nulo, ou seja, $A + B \in W$.

Agora, sejam $A \in W$ e $c \in \mathbb{R}$. Note que o $i \times j$-ésimo elemento de $c \cdot A$ é $c \cdot a_{ij} = 0$. Logo, $c \cdot A \in W$.

$$\therefore W \text{ é subespaço.}$$

□

Exemplo. $W = \{(x_1, x_2, x_3, x_4) | x_i \geq 0\}$ *não é subespaço de* $\mathbb{R}^4$*, pois apesar da propriedade 1 ser imediata, a propriedade 2 não é válida, uma vez que* $-1 \cdot (1, 1, 1, 1) \notin W$ *mas* $(1, 1, 1, 1) \in W$*.*

Exemplo. $W = \{(x_1, x_2, x_3, x_4) | x_1 \cdot x_4 = 0\}$ *não é subespaço de* $\mathbb{R}^4$*, pois tanto* $(0, 1, 1, 1)$ *quanto* $(1, 1, 1, 0)$ *estão em* W*, mas* $(0, 1, 1, 1) + (1, 1, 1, 0) = (1, 2, 2, 1) \notin W$*, ou seja, a propriedade 1 não vale.*

Definição. *Uma expressão* $w = c_1v_1 + c_2v_2 + \cdots + c_nv_n$, *sendo* $c_i \in \mathbb{R}$ *e* v_i *vetores do espaço vetorial* V, *chama-se* ***combinação linear***.

Exemplo. *Queremos saber se* $w = (2, -6, 3)$ *é combinação linear de* $v_1 = (1, -2, -1)$ *e* $v_2 = (3, -5, 4)$. *Para isso, devemos encontrar constantes* c_1 *e* c_2 *tais que* $c_1v_1 + c_2v_2 = w$. *Daí, temos:*

$$\begin{cases} c_1 + 3c_2 = 2 \\ -2c_1 - 5c_2 = -6 \\ -c_1 + 4c_2 = 3 \end{cases} \Leftrightarrow \begin{cases} c_1 + 3c_2 = 2 \\ c_2 = 2 \\ c_2 = \dfrac{5}{7} \end{cases}$$

o que é absurdo. Logo, w *não é combinação linear de* v_1 *e* v_2.

Exemplo. *Queremos saber se* $w = (-7, 7, 11)$ *é combinação linear de* $v_1 = (1, 2, 1)$, $v_2 = (-4, -1, 2)$ *e* $v_3 = (-3, 1, 3)$. *Para tal, devemos encontrar constantes* c_1, c_2 *e* c_3 *tais que*

$$\begin{aligned} & c_1v_1 + c_2v_2 + c_3v_3 = w \\ \Rightarrow & \begin{cases} c_1 - 4c_2 - 3c_3 = -7 \\ 2c_1 - c_2 + c_3 = 7 \\ c_1 + 2c_2 + 3c_3 = 11 \end{cases} \\ \Rightarrow & \begin{cases} c_1 = 5 - s \\ c_2 = 3 - s \\ c_3 = s \end{cases} , s \in \mathbb{R} \end{aligned}$$

Tomando $s = 0$, *obtemos* $w = 3v_2 + 5v_1$. *Note que essa combinação não é única.*

Exemplo. *Tome o espaço vetorial* $V = \mathbb{R}^3$ *e os vetores* $i = (1, 0, 0)$, $j = (0, 1, 0)$ *e* $k = (0, 0, 1)$. *Note que todo vetor* $v = (x, y, z)$ *de* V *pode ser escrito como* $v = x \cdot i + y \cdot j + z \cdot k$, *ou seja, todo vetor de* V *é combinação linear de* i, j, *e* k.

Teorema 3.1. *O conjunto de todas as combinações lineares de* $v_1, v_2, \ldots, v_n \in V$ *forma um subespaço de* V *chamado* ***subespaço gerado*** *por* $v_1, v_2, \ldots, v_n$.

Demonstração. Vamos mostrar que tal subconjunto de fato é subespaço de V.

Sejam $c_1v_1 + c_2v_2 + \cdots + c_nv_n$ e $c'_1v_1 + c'_2v_2 + \cdots + c'_nv_n$ duas combinações lineares. Note que, somando as duas, obtemos

$$(c_1 + c'_1)v_1 + (c_2 + c'_2)v_2 + \cdots + (c_n + c'_n)v_n \text{ que também é uma combinação linear.}$$

Além disso, seja $\alpha \in \mathbb{R}$. Então, multiplicando a primeira combinação linear por α, obtemos

$$(\alpha c_1)v_1 + (\alpha c_2)v_2 + \cdots + (\alpha c_n)v_n \text{ que também é uma combinação linear, basta notar que } \alpha c_i = d_i \in \mathbb{R}.$$

□

Definição. *Dizemos que* $v_1, v_2, \ldots, v_n$ *são* ***linearmente independentes (L.I.)*** *se* $c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0$ *implica* $c_1 = c_2 = \cdots = c_n = 0$. *Do contrário, dizemos que* $v_1, v_2, \ldots, v_n$ *são* ***linearmente dependentes (L.D.)***.

Exemplo. *Sejam* $V = \mathbb{R}^n$ *e* $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$, $\ldots$, $e_n = (0, 0, \ldots, 1)$. *Queremos verificar se* $e_1, e_2, \ldots, e_n$ *são L.I. Para isso, note que*

$$
\begin{aligned}
& c_1e_1 + c_2e_2 + \cdots + c_ne_n = 0 \\
\Rightarrow & \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \\
& \Rightarrow c_1 = c_2 = \cdots = c_n = 0
\end{aligned}
$$

Logo, $e_1, e_2, \ldots, e_n$ são, de fato, L.I.

Exemplo. *Sejam $v_1 = (1,2,2,2)$, $v_2 = (2,3,4,1)$ e $v_3 = (3,8,7,5)$. Note que*

$$
\begin{aligned}
& c_1v_1 + c_2v_2 + c_3v_3 = 0 \\
& \Rightarrow c_1 \begin{bmatrix} 1 \\ 2 \\ 2 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 3 \\ 8 \\ 7 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\
& \Rightarrow \begin{cases} c_1 + 2c_2 + 3c_3 = 0 \\ 2c_1 + 3c_2 + 8c_3 = 0 \\ 2c_1 + 4c_2 + 7c_3 = 0 \\ 2c_1 + c_2 + 5c_3 = 0 \end{cases} \\
& \Rightarrow \begin{cases} c_1 = 0 \\ c_2 = 0 \\ c_3 = 0 \end{cases}
\end{aligned}
$$

Logo, v_2, v_2, v_3 são L.I.

Fato. *Se $v_1, v_2, \ldots, v_n$ são L.I., então $w = c_1v_1 + c_2v_2 + \cdots + c_nv_n$ é único, isto é, w pode ser escrito como combinação linear de $v_1, v_2, \ldots, v_n$ de forma única.*

Demonstração. Suponha que $w = c_1'v_1 + c_2'v_2 + \cdots + c_n'v_n$. Então, subtraindo w de si mesmo, temos

$$
(c_1' - c_1)v_1 + (c_2' - c_2)v_2 + \cdots + (c_n' - c_n)v_n = 0
$$

Como $v_1, v_2, \ldots, v_n$ são L.I., devemos ter $c_i' - c_i = 0$, ou seja, $c_i' = c_i, i = 1, 2, \ldots, n$.

$\square$

Exemplo. *Sejam $v_1 = (2,1,3)$, $v_2 = (5,-2,4)$, $v_3 = (3,8,-6)$ e $v_4 = (2,7,-4)$. Note que $c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 = 0$ implica*

$$
\begin{cases} 2c_1 + 5c_2 + 3c_3 + 2c_4 = 0 \\ c_1 - 2c_2 + 8c_3 + 7c_4 = 0 \\ 3c_1 + 4c_2 - 6c_3 - 4c_4 = 0 \end{cases}
$$

Como o sistema acima é homogêneo e possui mais incógnitas que equações, ele possui infinitas soluções. Logo, pela contrapositiva do fato anterior, sabemos que $v_1, v_2, \ldots, v_n$ são L.D.

Generalizando o exemplo anterior, sabemos que m vetores em $\mathbb{R}^n$, $m > n$, são L.D. Se $m = n$, temos

$$
\begin{aligned}
&c_1v_1 + \cdots + c_nv_n = 0\\
&\Rightarrow c_1 \begin{bmatrix} a_{11}\\ a_{21}\\ \vdots\\ a_{n1}\end{bmatrix} + \cdots + c_n \begin{bmatrix} a_{n1}\\ a_{n2}\\ \vdots\\ a_{nn}\end{bmatrix} = \begin{bmatrix} 0\\ 0\\ \vdots\\ 0\end{bmatrix}\\
&\Rightarrow A \begin{bmatrix} c_1\\ c_2\\ \vdots\\ c_n\end{bmatrix} = \begin{bmatrix} 0\\ 0\\ \vdots\\ 0\end{bmatrix}
\end{aligned}
$$

Esse sistema tem solução única (a trivial) se, e só se, A é inversível. Mas A é inversível se, e só se, $det(A) \neq 0$, logo $v_1, v_2, \ldots, v_n$ são L.I. se, e só se, a matriz cuja i-ésima coluna são as entradas do i-ésimo vetor tem determinante não nulo.

Definição. *Uma* ***base*** *de um espaço vetorial V é um conjunto de vetores L.I $\{v_1, v_2, \ldots, v_n\}$ que geram V, isto é, todo vetor de V pode ser escrito como combinação linear de $v_1, v_2, \ldots, v_n$, de forma única.*

Exemplo. *Os vetores $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)$ geram $V = \mathbb{R}^n$, pois $e_1, e_2, \ldots, e_n$ são L.I., como mostramos anteriormente e, para todo $v = (x_1, x_2, \ldots, x_n)$ em V, $v = x_1 \cdot e_1 + \cdots + x_n \cdot e_n$. A base $\{e_1, \ldots, e_n\}$ é dita* ***base canônica*** *de $\mathbb{R}^n$.*

Fato. *Sejam $v_1, v_2, \cdots, v_n$ vetores L.I. de $\mathbb{R}^n$. Então, $\{v_1, \ldots, v_n\}$ é uma base de $\mathbb{R}^n$.*

Demonstração. Seja $w \in \mathbb{R}^n$ tal que $\{w, v_1, \ldots, v_n\}$ é L.D., ou seja, existem constantes $c_0, c_1, \ldots, c_n \in \mathbb{R}$ não todas nulas tais que

$$c_0w + c_1v_1 + \cdots + c_nv_n = 0$$

Note que devemos ter $c_0 \neq 0$, pois do contrário $c_0 = c_1 = \cdots = c_n = 0$. Logo, isolando w, temos

$$w = -\frac{c_1}{c_0}v_1 + \cdots + \Big(-\frac{c_n}{c_0}v_n\Big)$$

Logo, fazendo $d_i = -\dfrac{c_i}{c_0}$, vemos que w é combinação linear de $v_1, \ldots, v_n$. Portanto, $\{v_1, \ldots, v_n\}$ é base de $\mathbb{R}^n$.

□

Exemplo. *Sejam $v_1 = (1, -1, -2, -3), v_2 = (1, -1, 2, 3), v_3 = (1, -1, -3, -2), v_4 = (0, 3, -1, 2)$. Note que*

$$\begin{vmatrix} 1 & 1 & 1 & 0\\ -1 & -1 & -1 & 3\\ -2 & 2 & -3 & -1\\ -3 & 3 & -2 & 2\end{vmatrix} = 30 \neq 0$$

logo v_1, v_2, v_3, v_4 são L.I. Além disso, $\{v_1, v_2, v_3, v_4\}$ é base de $\mathbb{R}^4$, ou seja, gera $\mathbb{R}^4$.

Teorema 3.2. *Seja $\{v_1, \ldots, v_n\}$ uma base do espaço vetorial V. Então, qualquer conjunto $\{w_1, \ldots, w_m\}$, com $m > n$, é L.D.*

Demonstração. Podemos escrever

$$
\begin{aligned}
w_1 &= a_{11}v_1 + \cdots + a_{n1}v_n \\
&\vdots \\
w_m &= a_{1m}v_1 + \cdots + a_{nm}v_n
\end{aligned}
$$

Queremos mostrar que $c_1w_1 + \cdots + c_mw_m = 0$ tem solução não trivial. Substituindo $w_1, \ldots, w_m$, temos

$$
\begin{aligned}
& c_1(a_{11}v_1 + \cdots + a_{n1}v_n) + \cdots + c_m(a_{1m}v_1 + \cdots + a_{nm}v_n) = 0 \\
\Leftrightarrow\ & (c_1a_{11} + \cdots + c_ma_{1m})v_1 + \cdots + (c_1a_{n1} + \cdots + c_ma_{mn})v_n = 0
\end{aligned}
$$

Como $v_1, \ldots, v_n$ são L.I., segue que

$$
\begin{cases}
a_{11}c_1 + a_{12}c_2 + \cdots + a_{1m}c_m = 0 \\
\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad \vdots \\
a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nm}c_m = 0
\end{cases}
$$

Como esse sistema é homogêneo e possui mais incógnitas que equações, pois $m > n$, então há infinitas soluções. Portanto, $\{w_1, \ldots, w_m\}$ é L.D.

□

Corolário 3.2.1. *O número de vetores de uma base não depende da escolha dos vetores, ou seja, toda base de um espaço vetorial tem a mesma quantidade de vetores. Essa quantidade é chamada* ***dimensão*** *do espaço vetorial e denotada por dim(V).*

Observação. *Note que o vetor nulo nunca está em uma base, pois, do contrário, os vetores da base seriam L.D.*

Exemplo. *Seja $\mathcal{P}^n$ o espaço vetorial dos polinômios de grau menor ou igual a n. Queremos encontrar dim$(\mathcal{P}^n)$. Para isso, seja $p(x) \in \mathcal{P}^n$. Note que $p(x) = a_0 + a_1x + \cdots + a_nx^n, a_i \in \mathbb{R}, i = 0, 1, \ldots, n$, ou seja, $p(x)$ é combinação linear de $1, x, \ldots, x^n$. Logo, $\{1, x, \ldots, x^n\}$ gera $\mathcal{P}^n$. Resta ver se tal conjunto é base (pois é possível que um desses vetores seja L.D. aos outros, ou seja, podemos ter vetores demais). Ora, basta notar que*

$$c_0 + c_1x + \cdots + c_nx^n = 0 \Rightarrow c_0 = c_1 = \cdots = c_n = 0$$

Logo, $1, x, \ldots, x^n$ são L.I. e geram $\mathcal{P}^n$. Portanto, o conjunto $\{1, x, \ldots, x^n\}$ é base de $\mathcal{P}^n$ e dim$(\mathcal{P}^n) = n + 1$.

Exemplo. *Seja $M_{2\times 3} = \left\{ \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \middle| a_{ij} \in \mathbb{R} \right\}$. Note que*

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

é base de $M_{2\times 3}$, logo dim$(M_{2\times 3}) = 6$.

Em geral, $\dim(M_{m\times n}) = mn$.

Teorema 3.3. *Sejam V um espaço vetorial com dim$(V) = n$ e S um subconjunto de vetores de V. Então*

1. *Se S é L.I. e tem n vetores, então S é base de V;*

2. Se S gera V e tem n vetores, então S é base de V;

3. Se S tem mais que n vetores e gera V, então S contém uma base de V;

4. Se S tem menos que n vetores L.I., então S está contido em uma base.

Demonstração. **1.** Toda base de V tem n vetores. Como S tem n vetores L.I., S gera V e, portanto, é base.

2. Segue da definição de base;

3. Se S gera V e tem mais de n vetores, então S é, necessariamente, L.D. (pois se fosse L.I. seria base, mas toda base de V tem n vetores). Mas, como S gera V, então existe um subconjunto L.I. de S que gera V sendo, portanto, base. Logo, S contém uma base de V;

4. Para toda base W de V, qualquer subconjunto de vetores de W é L.I., logo todo subconjunto de V com menos de n vetores L.I. está contido em uma base. □

3.1 Algoritmo para obter uma base do espaço-solução de $AX = 0$

1. Escalonar A;
2. Atribuir parâmetros às incógnitas livres e resolver o sistema para as incógnitas líderes;
3. Faça um dos parâmetros igual a 1 e todos os outro 0. A solução correspondente é um vetor da base. Faça isso até que todos os parâmetros tenham assumido o valor 1. O conjunto das soluções correspondentes forma uma base do espaço-solução.

Após a execução do algoritmo, obtemos

$$\begin{aligned} x_1 &= \sum_{i=1}^{n} c_{1i}t_i \\ x_2 &= \sum_{i=1}^{n} c_{2i}t_i \\ &\vdots \\ x_n &= \sum_{i=1}^{n} c_{ni}t_i \\ x_{n+1} &= t_1 \\ x_{n+2} &= t_2 \\ &\vdots \\ x_m &= t_{m-n} \end{aligned}$$

A solução geral do sistema é $(x_1, x_2, \ldots, x_n, t_1, t_2, \ldots, t_{m-n})$. Tomando $t_1 = 1, t_2 = 0, \ldots, t_{m-n} = 0$, obtemos $w_1 = (x_1, x_2, \ldots, x_n, 1, 0, \ldots, 0)$. Fazendo $t_1 = 0, t_2 = 1, t_3 = 0, \ldots, t_{m-n} = 0$, obtemos $w_2 = (x_1, x_2, \ldots, x_n, 0, 1, \ldots, 0)$. Seguindo o mesmo processo, obtemos $\{w_1, \ldots, w_{m-n}\}$, que é um conjunto de vetores L.I. e que geram o espaço-solução, logo é base.

Exemplo. *Vamos encontrar uma base do espaço-solução de*

$$\begin{cases} 3x_1 + 6x_2 - x_3 - 5x_4 + 5x_5 = 0 \\ 2x_1 + 4x_2 - x_3 - 3x_4 + 2x_5 = 0 \\ 3x_1 + 6x_2 - 2x_3 - 4x_4 + x_5 = 0 \end{cases}$$

Escalonando a matriz dos coeficientes, obtemos

$$E = \begin{bmatrix} 1 & 2 & 0 & -2 & 3 \\ 0 & 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Daí

$$\begin{cases} x_1 = -2t_1 + 2t_2 - 3t_3 \\ x_2 = t_1 \\ x_3 = t_2 - 4t_3 \\ x_4 = t_2 \\ x_5 = t_3 \end{cases}$$

Fazendo $t_1 = 1, t_2 = 0 = t_3$, *temos* $(-2, 1, 0, 0, 0)$. *Fazendo* $t_1 = 0 = t_3, t_2 = 1$, *temos* $(2, 0, 1, 1, 0)$. *Fazendo* $t_1 = 0 = t_2, t_3 = 1$, *temos* $(-3, 0, -4, 0, 1)$. *Portanto,* $\{(-2, 1, 0, 0, 0), (2, 0, 1, 1, 0), (-3, 0, -4, 0, 1)\}$ *é base do espaço-solução.*

3.2 Espaço-linha

Seja A uma matriz $m \times n$.

Definição. *O subespaço de* $\mathbb{R}^n$ *gerado pelas linhas de* A *chama-se espaço-linha e é denotado por linha*(A).

Para obter uma base do espaço-linha de uma matriz, temos o seguinte algoritmo:

1. Escalone A;
2. Linhas não nulas da matriz escalonada E formam uma base de linha(A).

Exemplo. *Vamos encontrar uma base de linha*(A).

$$A = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 3 & 4 & 9 & 0 & 7 \\ 2 & 3 & 5 & 1 & 8 \\ 2 & 2 & 8 & -3 & 5 \end{bmatrix} \therefore E = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 1 & -3 & 5 & -4 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Vemos que as três primeiras linhas de E *são não nulas. Logo,* $\{(1, 2, 1, 3, 2), (0, 1, -3, 5, -4), (0, 0, 0, 1, -7)\}$ *é base de linha*(A).

3.3 Espaço-coluna

Seja A uma matriz $m \times n$.

Definição. *O subespaço de* $\mathbb{R}^m$ *gerado pelas colunas de* A *chama-se espaço-coluna de* A *e é denotado por col*(A).

Observação. *Colunas linearmente independentes são aquelas que têm elemento líder.*

Para obter uma base do espaço-coluna de uma matriz, temos o seguinte algoritmo:

1. Escalone A;
2. Encontre colunas com elementos líderes da matriz escalonada E;
3. As colunas correspondentes da matriz original A formam uma base de col(A).

Observação. *Podemos enxergar esse algoritmo como um método para "descartar"vetores. Isso porque esse algoritmo nos diz quais colunas (vetores) da matriz (conjunto) são linearmente independentes entre si.*

Usando o exemplo anterior, vemos que a 1°, 2° e 4° colunas têm elementos líderes.

Logo, $\{(1,3,2,2),(2,4,3,2),(3,0,1,-3)\}$ é base de $\text{col}(A)$.

Exemplo. *Encontre um subconjunto dos vetores* $v_1 = (1,-1,2,2), v_2 = (-3,4,1,-2), v_3 = (0,1,7,4), v_4 = (-5,7,4,-2)$ *que forme uma base do espaço* W *gerado por* $\{v_1,v_2,v_3,v_4\}$*. Para isso, vamos usar o algoritmo do espaço-coluna (pois ele nos dá uma base com as colunas da matriz original). Fazendo*

$$A = \begin{bmatrix} 1 & -3 & 0 & -5 \\ -1 & 4 & 1 & 7 \\ 2 & 1 & 7 & 4 \\ 2 & -2 & 4 & -2 \end{bmatrix} \Rightarrow E = \begin{bmatrix} 1 & -3 & 0 & -5 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Note que as duas primeiras colunas contêm variáveis líderes. Logo, $\{(1,-1,2,2),(-3,4,1,-2)\}$ *é base de* W*.*

Note que $\dim(\text{linha}(A)) = \dim(\text{col}(A))$, visto que o número de linhas não nulas da matriz escalonada nos dá o número de variáveis líderes, ou seja, a dimensão de $\text{col}(A)$ e, ao mesmo tempo, a dimensão de $\text{linha}(A)$.

Como o número de incógnitas livres nos dá a dimensão do espaço-soluçao de $AX = 0$, sabemos que

$$\dim(\text{col}(A)) + \dim(\text{sol}(A)) = n$$

3.4 Produto escalar

Seja $V = \mathbb{R}^n$.

Definição. *O produto escalar* $v \cdot w$ *de vetores* $v = (a_1, a_2, \ldots, a_n)$ *e* $w = (b_1, b_2, \ldots, b_n)$ *é o número* $v \cdot w = \sum_{i=1}^{n} a_i \cdot b_i = \sum_{i=1}^{n} b_i \cdot a_i = w \cdot v$.

Definição. *O* ***comprimento*** *(módulo) de um vetor é* $|v| = \sqrt{v \cdot v} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$.

O produto escalar possui as seguintes propriedades:

1. $v \cdot w = w \cdot v$;
2. $u(v + w) = u \cdot v + u \cdot w$;
3. $c(v \cdot w) = (cv) \cdot w, \forall c \in \mathbb{R}$;
4. $v \cdot v \geq 0 \forall v \in V$ e $v \cdot v = 0 \Leftrightarrow v = 0$;
5. $|v \cdot w| \leq |v| \cdot |w|, \forall v, w \in V$ (desigualdade de Cauchy-Schwartz).

Demonstração. **1.** Segue da definição.

2. Seja $u = (c_1, \ldots, c_n)$. Daí, temos

$$u \cdot (v + w) = \sum_{i=1}^{n} c_i(a_i + b_i) = \sum_{i=1}^{n} c_i \cdot a_i + \sum_{i=1}^{n} c_i \cdot b_i = u \cdot v + u \cdot w$$

3. Basta notar que

$$c(v \cdot w) = c \sum_{i=1}^{n} a_i \cdot b_i = \sum_{i=1}^{n} (ca_i) \cdot b_i = (cv) \cdot w$$

4. Primeiro, note que

$$v \cdot v = \sum_{i=1}^{n} a_i^2 \geq 0$$

Em particular, perceba que

$$v \cdot v = 0 \Leftrightarrow \sum_{i=1}^{n} a_i^2 = 0 \Leftrightarrow a_i^2 = 0, \forall i \Leftrightarrow a_i = 0, \forall i \Leftrightarrow v = (0, \ldots, 0)$$

5. Seja $x \in \mathbb{R}$. Note que

$$(xv + w) \cdot (xv + w) = x^2 v \cdot v + 2xv \cdot w + w \cdot w = ax^2 + bx + c \geq 0, \forall x \in \mathbb{R}$$

Portanto, devemos ter

$$b^2 - 4ac \leq 0 \Leftrightarrow b^2 \leq 4ac \Leftrightarrow 4(v \cdot w)^2 \leq 4v^2 \cdot w^2 \Leftrightarrow |v \cdot w| \leq |v| \cdot |w|$$

□

Daí, podemos definir $\cos(\theta) = \dfrac{v \cdot w}{|v| \cdot |w|}$, sendo θ o ângulo entre v e w. Com isso, vemos que dois vetores v e w são ortogonais se, e só se, $v \cdot w = 0$.

Definição. *A* ***distância*** *entre dois vetores v e w, denotada por $d(v, w)$, é definida por $d(v, w) = |v - w| = |w - v| = d(w, v)$.*

Teorema 3.4 (Desigualdade triangular)**.** *Para quaisquer vetores v e w em V, $|v + w| \leq |v| + |w|$.*

Demonstração. Note que

$$(v + w)^2 = v \cdot v + 2v \cdot w + w \cdot w \overset{\mathbf{5}}{\leq} v \cdot v + 2|v||w| + w \cdot w = |v||v| + 2|v||w| + |w||w| = (|v| + |w|)^2$$

Logo $|v + w| \leq |v| + |w|$

□

Note que a igualdade vale somente para $v \cdot w = |v||w|$, ou seja, quando v e w são paralelos. Além disso, se v e w são perpendiculares, temos $|v + w|^2 = |v|^2 + |w|^2$ (teorema de Pitágoras).

Teorema 3.5. *Se $v_1, v_2, \ldots, v_k$ são não nulos e mutuamente ortogonais, então eles são L.I.*

Demonstração. Queremos mostrar que $c_1 v_1 + \cdots + c_k v_k = 0$ implica $c_1 = \cdots = c_k = 0$. Multiplicando essa igualdade por v_i, temos

$$v_i(c_1 v_1 + \cdots + c_i v_i + \cdots + c_k v_k) = 0 \Leftrightarrow c_i \underset{\geq 0}{v_i^2} = 0 \Leftrightarrow c_i = 0, i = 1, 2, \ldots, k.$$

□

Definição. *Dizemos que $v \in \mathbb{R}^n$ é* ***ortogonal ao subespaço*** *W se v é ortogonal a* ***todo vetor*** *de W. O conjunto de todos os vetores ortogonais a W é chamado* ***complemento ortogonal*** *de W e denotado por $W^{\perp}$.*

Algumas propriedades do complemento ortogonal:

1. $W^{\perp}$ é subespaço de $\mathbb{R}^n$;

2. $(W^{\perp})^{\perp} = W$;

3. $W \cap W^{\perp} = \{0\}$;

4. Se $S = \{v_1, \ldots, v_k\}$ gera W, então v é ortogonal a W se, e só se, $v \cdot v_i = 0, i = 1, 2, \ldots, k$.

Demonstração. **1.** Sejam $v_1, v_2 \in W$, $v = v_1 + v_2$ e $cv_1 = a, c \in \mathbb{R}$. Seja ainda $w \in W$. Daí, $v \cdot w = (v_1 + v_2) \cdot w = v_1 \cdot w + v_2 \cdot w = 0 + 0 = 0$. Portanto, $v \perp w$, ou seja ,$v \in W$. Além disso, $w \cdot a = c(v_1 \cdot w) = 0$. Portanto, $a \perp w$, ou seja, $a \in W$.

2. O conjunto de todos os vetores que são perpendiculares a $W^{\perp}$ é, por definição, W.

3. Seja $w \in W \cap W^{\perp}$. Então, $w \cdot w = 0 \Leftrightarrow w = 0$.

4. $(\Rightarrow)$ Se v é ortogonal a W, então $v \cdot v_i = 0, i = 1, 2, \ldots, k$, pois S gera W.
$(\Leftarrow)$ Seja $w \in W$. Note que $w = c_1 v_1 + \cdots + c_k v_k$. Multiplicando por v, temos

$$v \cdot w = v(c_1 v_1 + \cdots + c_k v_k) = c_1(v \cdot v_1) + \cdots + c_k(v \cdot v_k) = 0$$

Logo, $v \perp w, \forall w \in W$, ou seja, v é ortogonal a W.

□

Seja A uma matriz $m \times n$. O sistema

$$A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

pode ser escrito como

$$\begin{cases} \vec{l_1} \cdot \vec{x} = 0 \\ \vec{l_2} \cdot \vec{x} = 0 \\ \vdots \\ \vec{l_m} \cdot \vec{x} = 0 \end{cases}$$

sendo $\vec{l_i}$ as linhas de A. Note que $\vec{x}$ é solução do sistema se, e só se, $\vec{x}$ é ortogonal ao espaço linha de A, ou seja, $(\text{linha}(A))^{\perp} = \text{sol}(A)$.

Seja W um subespaço gerado por $v_1, v_2, \ldots, v_k$. Para encontrarmos uma base de $W^{\perp}$, fazemos:

1. Formar a matriz A cujas linhas são $v_1, v_2, \ldots, v_k$;

2. Considere o sistema homogêneo $AX = 0$;

3. A base do espaço-solução será a base desejada.

Exemplo. *Suponha que* $v = (1, -3, 5)$ *gere* W*. Queremos uma base de* $W^{\perp}$*. Nesse caso,* $A = (1, -3, 5)$ *e o sistema* $AX = 0$ *tem solução*

$$\begin{cases} x = 3s - 5t \\ y = s \\ z = t \end{cases}$$

Daí, podemos escrever

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3s - 5t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}$$

Logo, $\{(3, 1, 0), (-5, 0, 1)\}$ *é base de* $W^{\perp}$*.*

Exemplo. *Sejam* $v_1 = (1, 2, 1, -3, -3), v_2 = (2, 5, 6, -10, -12)$ *vetores que geram* W*. Queremos encontrar uma base de* $W^{\perp}$*. Nesse caso, temos*

$$A = \begin{bmatrix} 1 & 2 & 1 & -3 & -3 \\ 2 & 5 & 6 & -10 & -12 \end{bmatrix} \Rightarrow E = \begin{bmatrix} 1 & 2 & 1 & -3 & -3 \\ 0 & 1 & 4 & -4 & -6 \end{bmatrix}$$

Daí, o sistema $AX = 0$ *tem solução*

$$\begin{cases} x_1 = 7r - 5s - 9t \\ x_2 = -4r + 4s + 6t \\ x_3 = r \\ x_4 = s \\ x_5 = t \end{cases}$$

Daí, as soluções do sistema podem ser escritas como

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 7r - 5s - 9t \\ -4r + 4s + 6t \\ r \\ s \\ t \end{bmatrix} = r \begin{bmatrix} 7 \\ -4 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -5 \\ 4 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -9 \\ 6 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Logo, $\{(-7, 4, 1, 0, 0), (-5, 4, 0, 1, 0), (-9, 6, 0, 0, 1)\}$ *é base de* $W^{\perp}$*.*

Observação. *Note que* $dim(W)$ *nada mais é que a dimensão do espaço coluna da matriz* A *cujas colunas são os vetores que geram* W*, ou seja,* $dim(W) = dim(col(A))$*. Por outro lado,* $dim(W^{\perp})$ *nada mais é que a dimensão do espaço solução de* $AX = 0$*, ou seja,* $dim(W^{\perp}) = dim(sol(A))$*. Portanto, vemos que*

$$dim(W) + dim(W^{\perp}) = n$$

Seja $V = \mathbb{R}^n$. Dado um subespaço W de V e $b \in \mathbb{R}^n$, queremos encontrar $p \in W, q \in W^{\perp}$ tais que $b = p + q$. Note que se $\{v_1, \ldots, v_k\}$ é base de W e $\{u_1, \ldots, u_{n-k}\}$ é base de $W^{\perp}$, então

$$b = \underbrace{c_1 v_1 + \cdots + c_k v_k}_{p} + \underbrace{c_{k+1} u_1 + \cdots + c_n u_{n-k}}_{q}$$

Definição. *Uma base* $\{v_1, \ldots, v_k\}$ *é dita* ***ortogonal*** *se os vetores dessa base são mutuamente ortogonais.*

Exemplo. *A base canônica* $\{e_1, \ldots, e_n\}$ *é ortogonal, pois* $e_i \cdot e_j = 0, \forall i \neq j$.

Se $\{v_1, \ldots, v_k\}$ é uma base **ortogonal** de W, então

$$p = \frac{v_1 \cdot b}{v_1 \cdot v_1} \cdot v_1 + \frac{v_2 \cdot b}{v_2 \cdot v_2} \cdot v_2 + \cdots + \frac{v_k \cdot b}{v_k \cdot v_k} \cdot v_k \tag{3}$$

O vetor p é a **projeção ortogonal** de b em W.

Exemplo. *Os vetores* $v_1 = (1,1,0,1), v_2 = (1,-2,3,1), v_3 = (-4,3,3,1)$ *são ortogonais. Seja* $b = (0,7,0,7)$. *Queremos encontrar a projeção ortogonal* p *de* b *em* W *gerado por* $\{v_1, v_2, v_3\}$. *Como a nossa base já é ortogonal, podemos usar a equação 3 diretamente para obter*

$$\begin{aligned} p &= \frac{14}{3}(1,1,0,1) - \frac{7}{15}(1,-2,3,1) + \frac{4}{5}(-4,3,3,1) \\ &= \frac{1}{15}\Big[(70,70,0,70) - (7,-14,21,7) + (-48,36,36,12)\Big] \\ &= \frac{1}{15}(15,120,15,75) \\ &= (1,8,1,5) \end{aligned}$$

Contudo, nossa base nem sempre é ortogonal, e a equação 3 só se aplica para bases ortogonais. Por isso, utilizamos um algoritmo que, dada uma base de W, nos dá uma base ortogonal de W.

3.5 Algoritmo de Gram-Schmidt

Dada uma base $\{v_1, \ldots, v_k\}$ de W, tome

$$\begin{cases} u_1 = v_1 \\ u_2 = v_2 - \dfrac{v_1 \cdot v_2}{v_1 \cdot v_1} \cdot v_1 = v_2 - \dfrac{u_1 \cdot v_2}{u_1 \cdot u_1} \cdot u_1 \\ u_3 = v_3 - \dfrac{u_1 \cdot v_3}{u_1 \cdot u_1} \cdot u_1 - \dfrac{u_2 \cdot v_3}{u_2 \cdot u_2} \cdot u_2 \\ \vdots \\ u_k = v_k - \dfrac{u_1 \cdot v_k}{u_1 \cdot u_1} \cdot u_1 - \cdots - \dfrac{u_{k-1} \cdot v_k}{u_{k-1} \cdot u_{k-1}} \cdot u_{k-1} \end{cases}$$

O conjunto $\{u_1, \ldots, u_k\}$ é base ortogonal de W. Além disso, podemos alterar o comprimento dos u_i, pois essa mudança não interfere na ortogonalidade.

Exemplo. *Sejam* $v_1 = (3,1,1), v_2 = (1,3,1), v_3 = (1,1,3)$ *base de* W. *Note que essa base não é ortogonal. Usando o alagoritmo de Gram-Schmidt, temos*

$$u_1 = (3,1,1)$$
$$u_2 = (1,3,1) - \frac{7}{11}(3,1,1) = \frac{1}{11}(-10,26,4) \approx (-5,13,2)$$
$$u_3 = (1,1,3) - \frac{7}{11}(3,1,1) - \frac{7}{99}(-5,13,2) = \frac{1}{99}(-55,-55,22-) \approx (1,1,-4)$$

Logo, $\{(1,1,3),(-5,13,2),(1,1,-4)\}$ *é base ortogonal de* W.

4 Autovalores e autovetores

Admita que todas as matrizes a seguir são quadradas.

Definição. *Um vetor $v \neq 0 \in \mathbb{R}^n$ chama-se* ***autovetor*** *de uma matriz $A_{n\times n}$ se existe $\lambda \in \mathbb{R}$ tal que $A \cdot v = \lambda v$. Nesse caso, λ é dito* ***autovalor*** *de A associado a v. Note que um mesmo v pode ter mais de um autovalor.*

Exemplo. *Tome $A = \begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix}$ e $v = (2, 1)$. Queremos verificar se v é autovetor de A e, se for, encontrar o(s) autovalor(es) de A associados a v. Para isso, fazemos*

$$A \cdot v = \begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Logo, v é, de fato, autovetor de A e $\lambda = 2$ é o seu autovalor associado.

Proposição. *Seja λ um autovalor de A. O conjunto de todos os autovetores associados a λ, junto com o vetor nulo, é um subespaço de $\mathbb{R}^n$, chamado* ***autoespaço*** *associado a λ e denotado por E_λ.*

Demonstração. Seja W_λ tal conjunto. Sejam ainda $v, w \in W_\lambda$ e $c \in \mathbb{R}$. Note que

$$A \cdot (v + w) = A \cdot v + A \cdot w = \lambda v + \lambda w = \lambda(v + w)$$

Logo, se w e v são autovetores de A associados a λ, então $w + v$ é autovetor de A associado a λ.

Além disso, temos

$$A(cv) = c(A \cdot v) = c(\lambda v) = \lambda(cv)$$

Logo, se v é autovetor de A associado a λ, então cv é autovetor de A associado a λ.

$\square$

Note que, da definição de autovetor e autovalor, temos:

$$A \cdot v - \lambda v = 0 \Leftrightarrow A \cdot v - \lambda I \cdot v = 0 \Leftrightarrow (A - \lambda I)v = 0$$

ou seja, para obter os autovetores de A, devemos encontrar v não trivial que seja solução do sistema acima.

Logo, vemos que λ é autovalor de A se, e só se

$$|A - \lambda I| = 0$$

Essa equação, chamada **equação característica** ou **polinômio característico** gera um polinômio em λ. Encontrando os valores de λ que satisfazem essa equação (ou seja, encontrando as raízes do polinômio), resolvemos o sistema para obter os autovetores associados a λ. Além disso, para encontrar uma base do autoespaço associado a λ, basta encontrar uma base do espaço-solução do sistema.

Com essas observações, podemos utilizar o seguinte algoritmo para encontrar uma base do autoespaço E_λ.

1. Resolva a equação característica $|A - \lambda I| = 0$ para obter os autovalores $\lambda_1, \ldots, \lambda_k$;
2. Para cada autovalor λ_i, encontre uma base do espaço-solução de $(A - \lambda_i I)v = 0$. A base encontrada é base do autoespaço correspondente, E_{λ_i}.

Exemplo. *Seja* $A = \begin{bmatrix} 5 & 7 \\ -2 & -4 \end{bmatrix}$. *Vamos encontrar bases para os autoespaços de* A.

Resolvendo a equação característica:

$$\begin{vmatrix} 5-\lambda & 7 \\ -2 & -4-\lambda \end{vmatrix} = 0 \Leftrightarrow (5-\lambda)(-4-\lambda)+14 = 0 \Leftrightarrow \lambda^2 - \lambda - 6 = 0 \Leftrightarrow \lambda_1 = 3, \lambda_2 = -2$$

Para λ_1, *temos o sistema*

$$\begin{bmatrix} 2 & 7 \\ -2 & -7 \end{bmatrix} \cdot v = 0 \Rightarrow 2x + 7y = 0 \Rightarrow \begin{cases} x = -\frac{7}{2}s \\ y = s \end{cases} , s \in \mathbb{R}$$

Logo, podemos escrever $v = s\begin{bmatrix} -\frac{7}{2} \\ 1 \end{bmatrix}$. *Consequentemente,* $\left\{\left(-\frac{7}{2}, 1\right)\right\}$ *é base de* E_{λ_1}.

Para λ_2, *temos o sistema*

$$\begin{bmatrix} 7 & 7 \\ -2 & -2 \end{bmatrix} \cdot v = 0 \Rightarrow x + y = 0 \Rightarrow \begin{cases} x = -s \\ y = s \end{cases} , s \in \mathbb{R}$$

Logo, $v = s\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. *Consequentemente,* $\{(-1, 1)\}$ *é base de* E_{λ_2}.

Exemplo. *Seja* $A = \begin{bmatrix} 3 & 5 \\ -5 & -3 \end{bmatrix}$. *Note que*

$$\begin{vmatrix} 3-\lambda & 5 \\ -5 & -3-\lambda \end{vmatrix} = 0 \Leftrightarrow (3-\lambda)(-3-\lambda)+25 = 0 \Leftrightarrow \lambda^2 + 16 = 0 \Leftrightarrow \lambda^2 = -16 \Leftrightarrow \lambda \notin \mathbb{R}$$

Logo, A *não possui autovalores nem autovetores (reais).*

Exemplo. *Seja* $A = I$. *Daí, temos*

$$\begin{vmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{vmatrix} = 0 \Leftrightarrow \lambda = 1$$

Note que $\lambda = 1$ *tem multiplicidade 2. Para esse autovalor, temos o sistema*

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \cdot v = 0$$

logo, qualquer vetor v *é autovetor de* I. *Consequentemente, podemos tomar como base de* E_λ *a base canônica de* $\mathbb{R}^2$, $\{(1,0), (0,1)\}$.

Exemplo. *Seja* $A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$. *A equação característica é*

$$(2-\lambda)^2 = 0 \Leftrightarrow \lambda = 2$$

Daí, temos o sistema

$$\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} \cdot v = 0 \Rightarrow \begin{cases} x = s \\ y = 0 \end{cases}$$

Logo, $v = s \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *e, consequentemente* $\{(1,0)\}$ *é base de* E_λ.

Exemplo. *Seja* $A = \begin{bmatrix} 3 & 0 & 0 \\ -4 & 6 & 0 \\ 16 & -15 & -5 \end{bmatrix}$. *A equação característica é*

$$\begin{vmatrix} 3-\lambda & 0 & 0 \\ -4 & 6-\lambda & 0 \\ 16 & -15 & -5-\lambda \end{vmatrix} = 0 \Leftrightarrow (3-\lambda)(6-\lambda)(-5-\lambda) = 0 \Leftrightarrow \lambda_1 = 3, \lambda_2 = 6, \lambda_3 = -5$$

Para λ_1*, temos o sistema*

$$\begin{bmatrix} 0 & 0 & 0 \\ -4 & 3 & 0 \\ 16 & -15 & -8 \end{bmatrix} \cdot v = 0 \Rightarrow \begin{cases} -4x + 3y = 0 \\ 16x - 15y - 8z = 0 \end{cases} \Rightarrow \begin{cases} x = -2s \\ y = -\dfrac{8}{3}s \\ z = s \end{cases}$$

Logo, $v_1 = s \begin{bmatrix} -2 \\ -\dfrac{8}{3} \\ 1 \end{bmatrix}$ *e, consequentemente,* $\left\{\left(-2, -\dfrac{8}{3}, 1\right)\right\}$ *é base de* E_{λ_1}.

Para λ_2*, temos o sistema*

$$\begin{bmatrix} -3 & 0 & 0 \\ -4 & 0 & 0 \\ 16 & -15 & -8 \end{bmatrix} \cdot v = 0 \Rightarrow \begin{cases} x = 0 \\ 15y + 11z = 0 \end{cases} \Rightarrow \begin{cases} x = 0 \\ y = -\dfrac{11}{15}s \\ z = s \end{cases}$$

Logo, $v_2 = s \begin{bmatrix} 0 \\ -\dfrac{11}{15} \\ 1 \end{bmatrix}$ *e, consequentemente,* $\left\{\left(0, -\dfrac{11}{15}, 1\right)\right\}$ *é base de* E_{λ_2}.

Para λ_3*, temos o sistema*

$$\begin{bmatrix} 8 & 0 & 0 \\ -4 & 0 & 0 \\ 16 & -15 & -11 \end{bmatrix} \cdot v = 0 \Rightarrow \begin{cases} x = 0 \\ y = 0 \\ z = s \end{cases}$$

Logo, $v = s \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ *e, consequentemenete,* $\{(0,0,1)\}$ *é base de* E_{λ_3}.

Exemplo. *Seja* $A = \begin{bmatrix} 4 & -2 & 1 \\ 2 & 0 & 1 \\ 2 & -2 & 3 \end{bmatrix}$. *A equação característica é*

$$\begin{vmatrix} 4-\lambda & -2 & 1 \\ 2 & -\lambda & 1 \\ 2 & -2 & 3-\lambda \end{vmatrix} = 0 \Rightarrow -\lambda^3 + 7\lambda^2 - 16\lambda + 12 = 0 \Rightarrow \lambda_1 = 2, \lambda_2 = 3$$

Para λ_1, temos o sistema

$$\begin{bmatrix} 2 & -2 & 1 \\ 2 & -2 & 1 \\ 2 & -2 & 1 \end{bmatrix} v = 0 \Leftrightarrow 2x - 2y + z \Leftrightarrow \begin{cases} x = s - \dfrac{1}{2}t \\ y = s \\ z = t \end{cases}$$

Logo, $v = s \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{bmatrix}$ *e, consequentemente,* $\left\{(1,1,0), \left(-\frac{1}{2}, 0, 1\right)\right\}$ *é base de* E_{λ_1}.

Para λ_2, temos o sistema

$$\begin{bmatrix} 1 & -2 & 1 \\ 2 & -3 & 1 \\ 2 & -2 & 0 \end{bmatrix} \cdot v = 0 \Rightarrow \begin{cases} x - 2y + z = 0 \\ 2x - 3y + z = 0 \\ x - y = 0 \end{cases} \Rightarrow \begin{cases} x = t \\ y = t \\ z = t \end{cases}$$

Logo, $v = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ *e, consequentemente,* $\{(1,1,1)\}$ *é base de* E_{λ_2}.

4.1 Diagonalização de matrizes

Considere $\mathbb{R}^n$ como nosso conjunto universo, suponha A uma matriz $n \times n$ e $v_1, \ldots, v_n$ autovetores L.I. de A. Considere as matrizes abaixo:

$$P = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}$$

$$AP = \begin{bmatrix} | & & | \\ Av_1 & \cdots & Av_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \lambda_1 v_1 & \cdots & \lambda_n v_n \\ | & & | \end{bmatrix}$$

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Note que

$$PD = \begin{bmatrix} | & & | \\ \lambda_1 v_1 & \cdots & \lambda_n v_n \\ | & & | \end{bmatrix} = AP \Rightarrow A = PDP^{-1}, \text{ pois } P \text{ é inversível}$$

Definição. *Dizemos que uma matriz A é **diagonalizável** se existe uma matriz inversível P tal que $P^{-1}AP = D$ ou, equivalentemente, $A = PDP^{-1}$, sendo D uma matriz diagonal. Nesse caso, P é dita matriz **diagonalizadora**.*

Teorema 4.1. *Uma matriz A é diagonalizável se, e só se, possui n autovetores L.I.*

Demonstração. ($\Leftarrow$) Foi provada no início da seção.

($\Rightarrow$) Se A é diagonalizável, então $AP = PD$, sendo P inversível a matriz formada pelos n autovetores de A. Logo, $v_1, \ldots, v_n$ são L.I. $\square$

Teorema 4.2. *Autovetores que correspondem a autovalores distintos são L.I.*

Demonstração. Sejam $v_1, \dots, v_k$ autovetores com autovalores $\lambda_1, \dots, \lambda_k$ distintos. Seja k o menor inteiro tal que o teorema não é válido. Temos:

$$c_1v_1 + \cdots + c_kv_k = 0 \Rightarrow (A - \lambda_1 I)(c_1v_1 + \cdots + c_kv_k) = 0 \Rightarrow (\lambda_2 - \lambda_1)c_2v_2 + \cdots + (\lambda_k - \lambda_1)c_kv_k = 0$$

Como k é o menor inteiro tal que o teorema não é válido, sabemos que $v_2, \dots, v_k$ são L.I., ou seja, $c_2 = \cdots = c_k = 0$. Mas isso implica $c_1 = 0$ que, por sua vez, implica $v_1, \dots, v_k$ L.I., o que é absurdo devido à nossa escolha de k. Logo, concluímos que $\nexists k \in \mathbb{Z}$ tal que o teorema não é válido, ou seja, nosso teorema sempre é verdadeiro.

□

Corolário 4.2.1. *Se A tem n autovalores distintos, então A é diagonalizável*

Demonstração. Do teorema 4.1 sabemos que A só é diagonalizável se possui n autovetores L.I. e, do teorema 4.2, sabemos que autovetores associados a autovalores distintos são L.I.; logo, se A tem n autovalores distintos, então os autovetores associados são L.I. e, consequentemente, A é diagonalizável. □

Observação. *Note que a recíproca nem sempre é verdadeira: A diagonalizável não implica n autovalores distintos.*

Corolário 4.2.2. *Sejam $S_1, S_2, \dots, S_k$ conjuntos L.I. de autovetores associados a autovalores distintos $\lambda_1, \dots, \lambda_k$. Então, o conjunto $S = \bigcup_{i=1}^{k} S_i$ é L.I.*

Demonstração. Sejam $v_1, \dots, v_n \in S$. Sabemos que $v_1, \dots, v_n$ são L.I., logo

$$c_1v_1 + \cdots + c_nv_n = 0 \Rightarrow w_1 + \cdots + w_k = 0$$

sendo w_i uma combinação linear dos vetores de S_i. Como $w_1, \dots, w_k$ são L.I. (pelo teorema 4.2), $w_1 = \cdots = w_k = 0$. Como S_i são L.I. para $i = 1, 2, \dots, k$, então $c_1 = \cdots = c_n = 0$ e, portanto, S é L.I.

□

Com esses fatos em mãos, podemos escrever o seguinte algoritmo para diagonalizar uma matriz $A_{n\times n}$:

1. Encontre os autovalores $\lambda_1, \dots, \lambda_k$ e uma base S_i para cada autoespaço associado a λ_i;
2. Considere $S = \bigcup_{i=1}^{n} S_i$. Se $|S| < n$, A não é diagonalizável. Se $|S| = n$, temos $P = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}$, $v_i \in S$ e $D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$.

Observação. *Note que a ordem **importa**. Se a primeira coluna de P contém o vetor v_i, a primeira coluna de D deverá conter o autovalor associado. Se, para um dado autovalor, tivermos dois ou mais autovetores associados, então a ordem desses autovetores não importa. Por exemplo, se λ_1 tem v_1 e v_2 como autovetores e λ_1 está na primeira coluna de D, então as duas primeiras colunas de P são v_1 e v_2 ou v_2 e v_1.*

Exemplo. *Seja $A = \begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix}$. Queremos diagonalizar A. Para isso, vamos usar o algoritmo. Resolvendo a equação característica:*

$$\begin{vmatrix} 5-\lambda & -6 \\ 2 & -2-\lambda \end{vmatrix} = 0 \Leftrightarrow (5-\lambda)(-2-\lambda) + 12 = 0 \Leftrightarrow \lambda^2 - 3\lambda + 2 = 0 \Leftrightarrow \lambda_1 = 2, \lambda_2 = 1$$

Para λ_1, temos o sistema

$$\begin{bmatrix} 3 & -6 \\ 2 & -4 \end{bmatrix} \cdot v = 0 \Rightarrow x - 2y = 0 \Rightarrow \begin{cases} x = 2s \\ y = s \end{cases}, s \in \mathbb{R}$$

Logo, as soluções do sistema têm a forma $s\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, *ou seja,* $v_1 = (2,1)$ *é o autovetor associado a* λ_1. *Consequentemente,* $\{(2,1)\}$ *é base de* E_{λ_1},

Para λ_2, temos o sistema

$$\begin{bmatrix} 4 & -6 \\ 2 & -3 \end{bmatrix} \cdot v = 0 \Rightarrow 2x - 3y = 0 \Rightarrow \begin{cases} x = \frac{3}{2}t \\ y = t \end{cases}, t \in \mathbb{R}$$

Logo, as soluções do sistema têm a forma $t\begin{bmatrix} \frac{3}{2} \\ 1 \end{bmatrix}$, *ou seja,* $v_2 = \left(\frac{3}{2}, 1\right)$ *é o autovetor associado a* λ_2. *Consequentemente,* $\left\{\left(\frac{3}{2}, 1\right)\right\}$ *é base de* E_{λ_2}.

Logo, podemos escrever:

$$P = \begin{bmatrix} 2 & \frac{3}{2} \\ 1 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

Exemplo. *Seja* $A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$. *A equação característica é*

$$\begin{vmatrix} 2-\lambda & 3 \\ 0 & 2-\lambda \end{vmatrix} = 0 \Leftrightarrow \lambda_{1,2} = 2$$

Daí, temos o sistema

$$\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} \cdot v = 0 \Rightarrow \begin{cases} x = s \\ y = 0 \end{cases}, s \in \mathbb{R}$$

Logo, toda solução do sistema tem a forma $s\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, *ou seja,* $v = \begin{vmatrix} 1 \\ 0 \end{vmatrix}$ *é o único autovetor de* A. *Como temos apenas um autovetor, e não dois,* A *não é diagonalizável.*

Exemplo. *Seja* $A = \begin{bmatrix} 3 & 0 & 0 \\ -4 & 6 & 0 \\ 16 & -15 & -5 \end{bmatrix}$, *como em um exemplo anterior. Desse exemplo, sabemos que os autovalores de* A *são* $\lambda_1 = 3, \lambda_2 = 6, \lambda_3 = -5$, *e os autovetores associados são* $v_1 = \left(-2, -\frac{8}{3}, 1\right), v_2 = \left(0, -\frac{11}{15}, 1\right)$ *e* $v_3 = (0,0,1)$. *Daí, podemos escrever*

$$P = \begin{bmatrix} -2 & 0 & 0 \\ -\dfrac{8}{3} & -\dfrac{11}{15} & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

Exemplo. *Seja* $A = \begin{bmatrix} 4 & -2 & 1 \\ 2 & 0 & 1 \\ 2 & -2 & 3 \end{bmatrix}$*, como num exemplo anterior. Desse exemplo, sabemos que os autovalores são* $\lambda_1 = 3, \lambda_2 = 2$ *(com multiplicidade 2) e os autovetores associados são* $v_1 = (1,1,0), v_2 = (-1,0,2), v_3 = (1,1,0)$*. Daí, podemos escrever*

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

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

4.2 Potência de matrizes

Seja A uma matriz diagonalizável. Então, sabemos que

$$A^n = (PDP^{-1})^n = PDP^{-1}PDP^{-1} \cdots PDP^{-1} = PD^nP^{-1}$$

Além disso, note que

$$D^n = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_k^n \end{bmatrix}$$

Exemplo. *Usando o exemplo anterior, vamos calcular* A^5*:*

$$A^5 = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}^5 \begin{bmatrix} 2 & -2 & 1 \\ -2 & 3 & -1 \\ -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 243 & 32 & 32 \\ 243 & 32 & 0 \\ 243 & 0 & 64 \end{bmatrix} \begin{bmatrix} 2 & -2 & 1 \\ -2 & 3 & 1 \\ -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 454 & -422 & 211 \\ 422 & -390 & 211 \\ 422 & -422 & 243 \end{bmatrix}$$

Sejam $V = \mathbb{R}^2$ e W um subespaço próprio de V. Seja ainda $0 \neq v \in W$. Então, como $\dim(W) = 1$, $W = \{cv | c \in \mathbb{R}\}$, ou seja, W gera uma reta que passa pela origem. Perceba então que subespaços próprios de $\mathbb{R}^2$ nada mais são do que retas que passam pela origem.

Agora, sejam $V = \mathbb{R}^3$ e W um subespaço próprio de V. Seja $0 \neq v \in W$. Então, se $\dim(W) = 1$, $W = \{cv | c \in \mathbb{R}\}$, ou seja, uma reta pela origem. Se $\dim(W) = 2$, $W = \{c_1 w + c_2 v | c_1, c_2 \in \mathbb{R}\}$, sendo $w \in W$ tal que $\{v, w\}$ é L.I., ou seja, W é um plano que contém a origem. Perceba então que subespaços próprios de $\mathbb{R}^3$ nada mais são do que retas ou planos pela origem.

Exemplo. $W = \{(x,y,z) | x = 0\}$ *é um plano pela origem, logo* W *é subespaço próprio de* $\mathbb{R}^3$*.*

Exemplo. $W = \{(x,y,z) | y = 1\}$ *não é subespaço de* $\mathbb{R}^3$*, pois* W *é um plano que não passa pela origem.*

Definição. *Uma matriz* A *é dita* ***simétrica*** *se* $A^T = A$*.*

Teorema 4.3. *Autovetores de uma matriz simétrica A associados a autovalores distintos λ_1, λ_2 são ortogonais*

Demonstração. Sejam v_1, v_2 autovetores de A associados a λ_1 e λ_2, respectivamente. Como $Av_1 = \lambda_1 v_1$ e A é simétrica, então $v_1^T A = v_1^T \lambda_1$. Com isso, temos

$$v_1^T A v_2 = v_1^T \lambda_1 v_2 = \lambda_1 v_1^T v_2 = \lambda_1 v_1 \cdot v_2$$

Por outro lado, temos

$$v_1^T A v_2 = v_1^T \lambda_2 v_2 = \lambda_2 v_1^T v_2 = \lambda_2 v_1 \cdot v_2$$

Logo

$$\lambda_1 v_1 \cdot v_2 = \lambda_2 v_1 \cdot v_2 \Leftrightarrow \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} v_1 \cdot v_2 = 0 \Leftrightarrow v_1 \cdot v_2 = 0 \Leftrightarrow v_1 \perp v_2$$

□

Definição. *Uma base é dita **ortonormal** se os vetores dela têm tamanho 1 e formam uma base ortogonal.*

Definição. *Uma matriz quadrada é **ortogonal** se $A^T = A^{-1}$.*

Teorema 4.4. *As seguintes afirmações são equivalentes:*

1. *A é ortogonal;*
2. *A^T é ortogonal;*
3. *As colunas de A formam uma base ortonormal de $\mathbb{R}^n$;*
4. *As linhas de A formam uma base ortonormal de $\mathbb{R}^n$.*

Demonstração. **2.** Note que $A^T = A^{-1} \Leftrightarrow A = (A^{-1})^T \Leftrightarrow A = (A^T)^{-1}$.

3. Seja $A = (a_{ij})$. Se A é ortogonal, então

$$A^T A = I_n \Rightarrow \sum_{k=1}^{n} a_{ik}^2 = 1, i = 1, 2, \ldots, n$$

Ou seja, os vetores coluna de A têm módulo 1. Além disso, como A é inversível, então os vetores coluna de A são L.I., isto é, formam base de $\mathbb{R}^n$. Mas essa base é ortogonal, pois

$$\sum_{k=1, i\neq j}^{n} a_{ik} a_{jk} = 0, i, j = 1, 2, \ldots, n$$

ou seja, os vetores coluna de A são ortogonais. Portanto, se A é ortogonal, suas colunas formam base ortonormal de $\mathbb{R}^n$.

Por outro lado, se as colunas de A formam base ortonormal de $\mathbb{R}^n$, então sabemos que os vetores coluna de A têm módulo 1 e são mutuamente ortogonais. Logo, se $A = (a_{ij})$, então

$$\sum_{k=1}^{n} a_{ik}^2 = 1, i = 1, 2, \ldots, n$$

$$\sum_{k=1, i\neq j}^{n} a_{ik} a_{jk} = 0, i, j = 1, 2, \ldots, n$$

Daí, sabemos que $A^T A = I_n$, ou seja, $A^T = A^{-1}$.

4. A demonstração é análoga à anterior, basta substituir A por A^T. (afinal, as linhas de A são as colunas de A^T)

□

Definição. *Dizemos que $A_{n\times n}$ é **ortogonalmente diagonalizável** se existe uma matriz diagonalizadora ortogonal P tal que*

$$A = PD\underbrace{P^T}_{P^{-1}}$$

Teorema 4.5. *A é ortogonalmente diagonalizável se, e somente se, seus autovetores formam uma base ortonormal de $\mathbb{R}^n$.*

Demonstração. Do teorema 4.4, sabemos que uma matriz ser ortogonal implica que suas colunas formam uma base ortonormal de $\mathbb{R}^n$ e vice-versa. A matriz diagonalizadora P tem como colunas os autovetores de A. Daí, se A é ortogonalmente diagonalizável, então $P^T = P^{-1}$, ou seja, P é ortogonal e, consequentemente, suas colunas (os autovetores de A) formam base ortonormal de $\mathbb{R}^n$. Por outro lado, se os autovetores de A formam base ortonormal de $\mathbb{R}^n$, então as colunas de P formam uma base ortonormal de $\mathbb{R}^n$, ou seja, P é ortogonal e, portanto, A é ortogonalmente diagonalizável. □

Teorema 4.6. *A matriz A é ortogonalmente diagonalizável se, e somente se, A é simétrica.*

Demonstração. Se A é ortogonalmente diagonalizável, então

$$A^T = (PDP^T)^T = PD^TP^T = PDP^T = A$$

ou seja, A é simétrica. Por outro lado, se A é simétrica, então

$$A^T = A \Rightarrow (PDP^{-1})^T = PDP^{-1} \Rightarrow (P^T)^{-1}DP^T = PDP^{-1} \Rightarrow (P^TP)^{-1}DP^TP = D \Rightarrow DP^TP = P^TPD$$

$$\therefore P^TP = I \Rightarrow P^T = P^{-1}$$

□

Teorema 4.7. *A equação característica de uma matriz simétrica possui apenas soluções reais.*

Demonstração. Uma matriz simétrica é ortogonalmente diagonalizável, ou seja, para cada autovalor temos apenas um autovetor associado. Isso implica que todos os autovalores tenham multiplicidade 1 não podendo ser, portanto, imaginários. □

Com isso, podemos escrever um algoritmo para diagonalizar ortogonalmente uma matriz. Dada uma matriz simétrica A:

1. Resolva a equação característica $|A - \lambda I| = 0$ para encontrar os autovalores $\lambda_1, \dots, \lambda_k$;
2. Para cada autovalor encontrado, resolva $(A - \lambda I)v = 0$ para encontrar uma base de cada autoespaço E_{λ_i};
3. Use o algoritmo de Gram-Schmidt para ortonormalizar cada base S_i do autoespaço E_{λ_i};
4. Divida cada vetor da base ortogonal obtida pelo comprimento dele;
5. As colunas de P são os vetores de tal base e D é a matriz diagonal com os autovalores correspondentes na diagonal principal.

Exemplo. *Seja* $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 6 \end{bmatrix}$. *Daí, os autovalores de* A *são* $\lambda_1 = 9, \lambda_2 = 4$ *e* $\lambda_3 = 0$ *e os autovetores associados são* $v_1 = (1, 2, 2), v_2 = (0, 1, -1)$ *e* $v_3 = (4, -1, -1)$*, respectivamente. Note que* v_1, v_2 *e* v_3 *já são ortogonais. Além disso, dim(*E_{λ_i}*)* $= 1$*, logo as bases de* E_{λ_i} *já são ortogonais.*

Dividindo os autovetores por seus comprimentos respectivos, obtemos:

$$\begin{aligned} u_1 &= \frac{1}{3}(1, 2, 2) \\ u_2 &= \frac{1}{\sqrt{2}}(0, 1, -1) \\ u_3 &= \frac{1}{\sqrt{18}}(4, -1, -11) \end{aligned}$$

Daí, $\{u_1\}$, $\{u_2\}$ *e* $\{u_3\}$ *são bases ortonormais de* E_{λ_1}, E_{λ_2} *e* E_{λ_3}*, respectivamente. Consequentemente, podemos escrever*

$$\begin{aligned} P &= \begin{bmatrix} \frac{1}{3} & 0 & \frac{4}{\sqrt{18}} \\ \frac{2}{3} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{18}} \\ \frac{2}{3} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{18}} \end{bmatrix} \\ D &= \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Exemplo. *Seja* $A = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & 4 & 0 \\ 0 & 1 & 0 & 3 \end{bmatrix}$. *A equação característica é*

$$\begin{vmatrix} 4-\lambda & 0 & 0 & 0 \\ 0 & 3-\lambda & 0 & 1 \\ 0 & 0 & 4-\lambda & 0 \\ 0 & 1 & 0 & 3-\lambda \end{vmatrix} = 0 \Leftrightarrow \lambda_1 = 4 \text{ (tripla) }, \lambda_2 = 2$$

Para λ_1*, temos o sistema*

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \cdot v = 0 \Rightarrow x_2 = x_4 \Rightarrow \begin{cases} x_1 = r \\ x_2 = t \\ x_3 = s \\ x_4 = t \end{cases} , s, r, t \in \mathbb{R}$$

Logo, as soluções do sistema têm a forma $r \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$. *Consequentemente, os autovetores associados a* λ_1 *são* $v_1 = (1, 0, 0, 0), v_2 = (0, 0, 1, 0)$ *e* $v_3 = (0, 1, 0, 1)$ *e* $\{v_1, v_2, v_3\}$ *é base ortogonal de* E_{λ_1}.

Para λ_2*, temos o sistema*

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \cdot v = 0 \Rightarrow \begin{cases} x_1 = 0 \\ x_2 = -s \\ x_3 = 0 \\ x_4 = s \end{cases}, s \in \mathbb{R}$$

Logo, as soluções do sistema têm a forma $s\begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$. *Consequentemente, o autovetor associado a* λ_2 *é* $v_4 = (0, -1, 0, 1)$. *Note que* $\{v_4\}$ *é base ortogonal de* E_{λ_2}. *Dividindo* v_4 *por seu comprimento, temos*

$$u_4 = \frac{1}{\sqrt{2}}(0, -1, 0, 1)$$

Daí, $\{u_1, u_2, u_3\}$ *e* $\{u_4\}$ *são bases ortonormais de* E_{λ_1} *e* E_{λ_2}, *respectivamente. Portanto, podemos escrever*

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$
$$D = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

5 Transformações lineares

Definição. *Dizemos que uma aplicação* $f : V \to W$ *entre espaços vetoriais é uma* ***transformação linear*** *se*

1. $f(v_1 + v_2) = f(v_1) + f(v_2)$
2. $f(cv_1) = cf(v_1), \forall c \in \mathbb{R}$.

Exemplo. *Seja* $f : \mathbb{R}^n \to \mathbb{R}^m$ *dada por* $f(v) = A \cdot v$ *sendo* $A_{m\times n}$. *Então,* f *é uma transformação linear.*

Demonstração. Note que

$$f(v_1 + v_2) = A \cdot (v_1 + v_2) = A \cdot v_1 + A \cdot v_2 = f(v_1) + f(v_2)$$
$$f(cv_1) = A \cdot (cv_1) = c(A\cdot)v_1 = cf(v_1), \forall c \in \mathbb{R}$$

□

Exemplo. $F : \mathbb{R}^3 \to \mathbb{R}^2$ *dada por* $F(v) = (2x - y + z, x + 3y - z)$ *é transformação linear.*

Demonstração. Note que

$$F(v) = A \cdot v, \text{ sendo } A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 3 & -1 \end{bmatrix}$$

Logo, pelo exemplo anterior, F é transformação linear.

□

Exemplo. *Seja* $T : \mathbb{R}^2 \to \mathbb{R}^3$ *dado por* $T(x,y) = A\begin{bmatrix} x \\ y \end{bmatrix}$*, sendo* $A = \begin{bmatrix} 1 & 2 \\ 3 & -2 \\ 4 & 3 \end{bmatrix}$*. Consequentemente,* $A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+2y \\ 3x-2y \\ 4x+3y \end{bmatrix}$*, logo* $T(x,y) = (x+2y, 3x-2y, 4x+3y)$*.*

Exemplo. *Seja* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *dado por* $T(x,y) = (3x-5y, 4x+7y)$*. Queremos escrever* T *na forma* $A\begin{bmatrix} x \\ y \end{bmatrix}$*. Ora, é evidente que* $A = \begin{bmatrix} 3 & -5 \\ 4 & 7 \end{bmatrix}$*.*

Exemplo. *Seja* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *dado por* $T(x,y) = (xy, x+y+1)$*. Note que* $T((0,0)+(0,1)) = (0,2)$*, mas* $T(0,0)+T(0,1) = (0,1)+(0,2) = (0,3)$*. Logo,* T *não é transformação linear.*

Teorema 5.1. T *é transformação linear se, e somente se,* $T(c_1v_1 + c_2v_2) = c_1T(v_1) + c_2T(v_2)$*.*

Demonstração. ($\Rightarrow$) Se T é transformação linear, então $T(c_1v_1 + c_2v_2) = T(c_1v_1) + T(c_2v_2) = c_1T(v_1) + c_2T(v_2)$.

($\Leftarrow$) Se $T(c_1v_1 + c_2v_2) = c_1T(v_1) + c_2T(v_2)$, então tomando $c_1 = 1 = c_2$, temos $T(v_1+v_2) = T(v_1) + T(v_2)$ e, tomando $c_2 = 0$, temos $T(c_1v_1) = c_1T(v_1)$, logo T é transformação linear. □

Corolário 5.1.1. *Consequentemente, podemos afirmar que*

1. $T(0) = 0$ *(o vetor nulo do domínio é levado no vetor nulo do contradomínio);*
2. $T(-v) = -T(v)$*;*
3. $T(v-w) = T(v) - T(w)$*.*

Teorema 5.2. *A função* $T : \mathbb{R}^n \to \mathbb{R}^n$ *é uma transformação linear se, e somente se,* T *é uma transformação matricial com matriz* $\begin{bmatrix} | & | & \cdots & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & \cdots & | \end{bmatrix}$*, sendo* $e_1, \ldots, e_n$ *vetores da base canônica de* $\mathbb{R}^n$*.*

Demonstração. ($\Leftarrow$) Foi provada no primeiro exemplo após a definição de transformação linear.

($\Rightarrow$) Se $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$, podemos escrever $v = x_1e_1 + \cdots + x_ne_n$. Daí, temos

$$T(v) = x_1T(e_1) + \cdots + x_nT(e_n) = \begin{bmatrix} | & | & \cdots & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

□

Observação. *Seja* V *um espaço vetorial e* $\{v_1, \ldots, v_n\}$ *uma base de* V*. Seja* $T : V \to W$ *uma transformação linear. Então, como* $v = c_1v_1 + \cdots + c_nv_n \in V$*,* $T(v) = c_1T(v_1) + \cdots + c_nT(v_n)$*. Se* $T : \mathbb{R}^n \to \mathbb{R}^n$*, temos* $A\underbrace{\begin{bmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & \cdots & | \end{bmatrix}}_{B} = \begin{bmatrix} | & | & \cdots & | \\ w_1 & w_2 & \cdots & w_n \\ | & | & \cdots & | \end{bmatrix}$ *e* $A = \begin{bmatrix} | & | & \cdots & | \\ w_1 & w_2 & \cdots & w_n \\ | & | & \cdots & | \end{bmatrix} B^{-1}$*.*

Exemplo. *Sejam* $v_1 = (3,5)$ *e* $v_2 = (4,7)$*. Seja ainda* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *tal que* $T(v_1) = (2,4)$ *e* $T(v_2) = (-1,3)$*. Daí, a matriz de* T *é*

$$A = \begin{bmatrix} 2 & -1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 5 & 7 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 7 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 19 & -11 \\ 13 & -7 \end{bmatrix}$$

Definição. *Dizemos que uma transformação linear* T *é* ***injetora*** *se* $v_1 \neq v_2$ *implica* $T(v_1) \neq T(v_2)$*. Dizemos também que* T *é* ***sobrejetora*** *se* $T(V) = W$*.*

Exemplo. *Seja* $T_1 : \mathbb{R}^2 \to \mathbb{R}^3$ *dada por* $T_1(x, y) = (x, y, 0)$*. Note que* T_1 *é injetora, mas não é sobrejetora pois qualquer vetor da forma* (x, y, z) *com* $z \neq 0$ *não é imagem de nenhum vetor no domínio.*

Agora, seja $T_2 : \mathbb{R}^3 \to \mathbb{R}^2$ *dada por* $T_2(x, y, z) = (x, y)$*. Note que* T *é sobrejetora mas não é injetora, pois os vetores* $(1, 2, 3)$ *e* $(1, 2, 4)$ *têm mesma imagem.*

Definição. *Uma transformação linear* $T : V \to W$ *injetora e sobrejetora (ou seja, bijetora) chama-se* ***isomorfismo****. Nesse caso, dizemos que* V *e* W *são* ***isomorfos****.*

Exemplo. *Vamos mostrar que* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *dada por* $T(x, y) = \underbrace{\begin{bmatrix} 19 & -11 \\ 13 & -7 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \end{bmatrix}$ *é isomorfismo.*

Demonstração. Suponha $T(x_1, y_1) = T(x_2, y_2)$. Então, $A \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = A \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$, logo $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$, pois $|A| \neq 0$. Portanto, T é injetora. Como A é inversível, então todo elemento $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ do contradomínio tem um vetor associado $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ tal que $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = A^{-1} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$. Portanto, T é isomorfismo. □

Exemplo. *Seja* V *um espaço vetorial e* $\{v_1, \ldots, v_n\}$ *uma base de* V*. Seja* $T : \mathbb{R}^n \to V$ *dada por* $T(x_1, \ldots, x_n) = x_1 v_1 + \cdots + x_n v_n$*. Então,* T *é isomorfismo.*

Demonstração. Note que

$$\begin{aligned} T[c_1(x_1, \ldots, x_n) + c_2(x'_1, \ldots, x'_n)] &= (c_1 x_1 + c_2 x'_1)v_1 + \cdots + (c_1 x_n + c_2 x'_n)v_n \\ &= c_1(x_1 v_1 + \cdots + x_n v_n) + c_2(x'_1 v_1 + \cdots + x'_n v_n) \\ &= c_1 T(x_1, \ldots, x_n) + c_2 T(x'_1, \ldots, x'_n) \end{aligned}$$

Portanto, T é transformação linear.

Agora, note que se $v \in V$, então $v = c_1 v_1 + \cdots + c_n v_n = T(c_1, \ldots, c_n)$, logo T é sobrejetora.

Por fim, se $T(x_1, \ldots, x_n) = T(x'_1, \ldots, x'_n)$, então $(x_1 - x'_1)v_1 + \cdots + (x_n - x'_n)v_n = 0$. Como $v_1, \ldots, v_n$ formam base de V, devemos ter $x_i = x'_i, i = 1, 2, \ldots, n$. Logo, T é injetora e, portanto é isomorfismo.

Concluímos que qualquer espaço vetorial de dimensão n é isomorfo a $\mathbb{R}^n$.

□

Definição. *O* ***núcleo*** *de um transformação linear* $T : V \to W$*, denotado por* $ker(T)$*, é o conjunto* $ker(T) = \{v \in V | T(v) = 0\}$ *e a* ***imagem*** *de* T*, denotada por* $Im(T)$*, é o conjunto* $T(V) = \{T(v) | v \in V\}$*. Note que, por definição,* T *é sobrejetora se, e somente se,* $T(V) = W = Im(T)$*.*

Proposição. $ker(T) = \{0\} \Leftrightarrow T$ *é injetora.*

Demonstração. Note que

$$T(v_1) = T(v_2) \Leftrightarrow T(v_1 - v_2) = 0 \Leftrightarrow v_1 - v_2 \in \ker(T)$$

Se $\ker(T) = \{0\}$, então $v_1 = v_2$ e T é injetora.

Se $\ker(T) \neq \{0\}$, então existe v_1 não nulo em $\ker(T)$. Fazendo $v_2 = 0$, temos $T(v_1) = 0 = T(v_2)$, logo T não é injetora.

□

Sabemos que se $\dim(V) = n$, então V é isomorfo a $\mathbb{R}^n$. Logo, toda transformação linear é da forma $T : \mathbb{R}^n \to \mathbb{R}^m$ dada por $T(v) = A \cdot v$, sendo $A_{m \times n}$ e $v = (x_1, \ldots, x_n)$. Daí, $\text{Im}(T) = \text{col}(A)$ e $\ker(T) = \text{sol}(A)$.

Corolário 5.2.1. *$ker(T)$ é subespaço de $\mathbb{R}^n$, $Im(T)$ é subespaço de $\mathbb{R}^m$ e $dim(ker(T)) + dim(Im(T)) = n$.*

Exemplo. *Seja $T : \mathbb{R}^5 \to \mathbb{R}^4$ dada por $T(v) = Av = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 3 & 4 & 9 & 0 & 7 \\ 2 & 3 & 5 & 1 & 8 \\ 2 & 2 & 8 & -3 & 5 \end{bmatrix} v$. Queremos obter uma base de $ker(T)$ e uma base de $Im(T)$. Escalonando A, obtemos $E = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 1 & -3 & 5 & -4 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. Logo, $\{(1,3,2,2),(2,4,3,2),(3,0,1,-3)\}$ é base de $Im(T)$.*

Resolvedo $Ev = 0$, obtemos

$$\begin{cases} x_1 = -7s + 39t \\ x_2 = 3s - 31t \\ x_3 = s \\ x_4 = 7t \\ x_5 = t \end{cases}$$

Daí, as soluções do sistema têm a forma $s \begin{bmatrix} -7 \\ 3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 39 \\ -31 \\ 0 \\ 7 \\ 1 \end{bmatrix}$, ou seja, $\{(-7,3,1,0,0),(39,-31,0,7,1)\}$ é base de $ker(T)$.

6 Mudança de base

Seja $B = \{v_1, \ldots, v_n\}$ uma base de $\mathbb{R}^n$. Então, se $v \in \mathbb{R}^n$, podemos escrever $v = x_1v_1 + \cdots + x_nv_n$. Temos que:

$$v = (x_1, \ldots, x_n)_B = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}_B$$

Exemplo. *Sejam $v_1 = (1,1,-2), v_2 = (3,1,3)$ e $v_3 = (2,3,4)$ (que formam uma base B de $\mathbb{R}^3$) e seja $v = (5,10,5)$. Queremos encontrar x_1, x_2, x_3 tais que*

$$x_1 \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 5 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} 1 & 3 & 2 \\ 1 & 1 & 3 \\ -2 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 5 \end{bmatrix}$$
$$\Rightarrow \begin{cases} x_1 = 2 \\ x_2 = -1 \\ x_3 = 3 \end{cases}$$

Logo, nosso vetor v, que originalmente foi escrito na base canônica, se torna $v_B = (2,-1,3)$ na base B.

Sejam $v \in \mathbb{R}^n$, $B = \{v_1, \ldots, v_n\}$, $B' = \{v'_1, \ldots, v'_n\}$. Note que

$$
\begin{aligned}
v &= x_1v_1 + \cdots + x_nv_n = x_1'v_1' + \cdots + x_n'v_n' \\
&\Rightarrow \underbrace{\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}}_{M_B} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \underbrace{\begin{bmatrix} | & & | \\ v_1' & \cdots & v_n' \\ | & & | \end{bmatrix}}_{M_{B'}} \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} \\
&\Rightarrow M_B \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = M_{B'} \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} \\
&\Rightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \underbrace{M_B^{-1}M_{B'}}_{P} \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}
\end{aligned}
$$

Exemplo. *Rotacionando o plano xy por um ângulo θ em torno da origem, obtemos um novo sistema de coordenadas cujos vetores unitários são $u_1 = (\cos(\theta), \sin(\theta))$ e $u_2 = (-\sin(\theta), \cos(\theta))$. Logo, se $v = (x, y)$, temos que*

$$
\begin{aligned}
\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \\
&\Rightarrow \begin{cases} x = x'\cos\theta + y'\sin\theta \\ y = -x'\sin\theta + y'\cos\theta \end{cases}
\end{aligned}
$$

Suponha $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear e $B = \{v_1, \ldots, v_n\}$ base de $\mathbb{R}^n$ e $C = \{w_1, \ldots, w_m\}$ base de $\mathbb{R}^m$. Se $v \in \mathbb{R}^n$, temos

$$
\begin{aligned}
v = x_1v_1 + \cdots + x_nv_n \Rightarrow T(v) &= x_1T(v_1) + \cdots + x_nT(v_n) \in \mathbb{R}^n \\
&\Rightarrow \begin{bmatrix} | & & | \\ T(v_1)_C & \cdots & T(v_n)_C \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}
\end{aligned}
$$

Agora, suponha $T : \mathbb{R}^n \to \mathbb{R}^n$, com $B = \{v_1, \ldots, v_n\}$ e $\{v_1', \ldots, v_n'\}$ bases de $\mathbb{R}^n$. Sejam $v_B = X$ e $T(X_B) = Y_B$. Temos $X = PX'$ e $Y = PY'$, sendo $X' = v_{B'}$ e $Y' = T(v)_{B'}$. Além disso, sabemos que

$$
T(X) = AX = Y \Rightarrow APX' = PY' \Rightarrow P^{-1}APX' = Y'
$$

Logo, se A é diagonalizável, basta escolher uma base de autovetores de A para realizarmos a mudança de base.