Vitor de Oliveira Ferreira

Álgebra Linear

8 de outubro de 2020

Prefácio

Estas notas compreendem todo o conteúdo da sequência de Álgebra Linear para os alunos do primeiro ano dos cursos de Engenharia na Escola Politécnica da Universidade de São Paulo.

A sequência de Álgebra Linear é constituída de duas disciplinas semestrais: MAT3457 - Álgebra Linear I e MAT3458 - Álgebra Linear II. Aqui, encontra-se o conteúdo unificado delas.

São parte integrante destas notas as listas de exercícios elaboradas pelos docentes responsáveis pelas turmas e disponibilizadas aos alunos.

O símbolo $\diamond$ marca o final da solução de um exercício, o símbolo $\square$, o de uma demonstração.

Sumário

1 Sistemas lineares, matrizes e determinantes 1
1.1 Sistemas lineares 1
1.2 Matrizes 8
1.3 Determinantes 12

2 Vetores 17
2.1 Vetores e operações entre vetores 17
2.2 Dependência linear 20
2.3 Bases e coordenadas 23
2.4 Produto escalar 28
2.5 Projeção ortogonal 30
2.6 Mudança de base 31
2.7 Produto vetorial 33
2.8 Produto misto 37

3 Geometria analítica 43
3.1 Sistemas de coordenadas 43
3.2 Retas 45
3.3 Planos 48
3.4 Posições relativas 53
3.5 Perpendicularidade e distâncias 57

4 Espaços vetoriais 63
4.1 Definição, exemplos e propriedades básicas 63
4.2 Subespaços vetoriais 67
4.3 Combinações lineares e dependência linear 69
4.4 Bases e dimensão 75
4.5 Coordenadas 79
4.6 Base e dimensão de subespaços 81
4.7 Soma e interseção de subespaços 86
4.8 Soma direta de subespaços 92

5 Espaços vetoriais com produto interno 95
5.1 Produto interno 95
5.2 Ortogonalidade 99
5.3 O complemento ortogonal 108

6 Transformações lineares . . . 111
6.1 Definição e exemplos . . . 111
6.2 Núcleo e imagem . . . 115
6.3 Teorema do núcleo e da imagem . . . 119
6.4 Operações com transformações lineares . . . 123
6.5 Matriz de uma transformação linear . . . 123
6.6 Matriz da transformação composta . . . 130
6.7 Mudança de base . . . 132

7 Diagonalização de operadores . . . 139
7.1 Autovalores e autovetores . . . 140
7.2 O polinômio característico . . . 143
7.3 Diagonalização . . . 147
7.4 Operadores diagonalizáveis . . . 151

A O posto de uma matriz . . . 159
A.1 Posto-linha e posto-coluna . . . 159
A.2 Extração de bases . . . 160

B Um pouco sobre funções . . . 163
B.1 Definições . . . 163
B.2 Composição de funções . . . 164

C Polinômios e suas raízes . . . 167
C.1 Polinômios com coeficientes reais . . . 167

Referências . . . 169

Índice Remissivo . . . 171

Capítulo 1
Sistemas lineares, matrizes e determinantes

Neste capítulo, e tem todos os capítulos subsequentes destas notas, o conjunto dos números reais será denotado por $\mathbb{R}$.

1.1 Sistemas lineares

Uma *equação linear* nas n incógnitas $x_1, x_2, \dots, x_n$ é uma equação da forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{1.1}$$

em que $a_1, a_2, \dots, a_n, b$ são números reais, chamados *coeficientes* da equação linear. Uma *solução* para (1.1) é uma sequência de n números reais $(s_1, s_2, \dots, s_n)$ tal que a igualdade envolvendo números reais

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = b$$

esteja satisfeita.

Um *sistema linear* de m equações nas n incógnitas $x_1, x_2, \dots, x_n$ é uma sequência de m equações lineares nas incógnitas $x_1, x_2, \dots, x_n$:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{1.2}$$

Uma solução para (1.2) é uma sequência $(s_1, s_2, \dots, s_n)$ de números reais que é solução para cada uma das m equações que compõem (1.2).

Exemplo 1.1.1 O par $(2, 1)$ é uma solução para o sistema linear

$$\begin{cases} 2x_1 + (-2)x_2 = 2 \\ 2x_1 + 2x_2 = 6, \end{cases} \tag{1.3}$$

como pode ser diretamente verificado. Ainda, $(2, 1)$ é a única solução desse sistema, pois se (s_1, s_2) é uma solução, então da primeira equação obtemos que $s_1 = s_2 + 1$. Como $(s_1, s_2) = (s_2 + 1, s_2)$ é solução também da segunda equação, segue $2(s_2 + 1) + 2s_2 = 6$, donde obtém-se $s_2 = 1$ e, portanto, $s_1 = 2$.

Exemplo 1.1.2 O sistema linear

$$\begin{cases} 2x_1 + 2x_2 = 4 \\ 2x_1 + 2x_2 = 2 \end{cases} \tag{1.4}$$

não tem soluções, uma vez que se (s_1, s_2) é uma solução da primeira equação do sistema, então $2s_1 + 2s_2 = 4$; assim, (s_1, s_2) não pode ser uma solução da segunda equação do sistema.

Exemplo 1.1.3 O sistema linear

$$\begin{cases} -6x_1 + 2x_2 = -8 \\ \frac{3}{2}x_1 + \left(-\frac{1}{2}\right)x_2 = 2 \end{cases} \tag{1.5}$$

tem infinitas soluções. De fato, para qualquer $t \in \mathbb{R}$, o par $(t, 3t - 4)$ é uma solução do sistema, como pode ser diretamente verificado.

Adiante, veremos um método para concluir, dado um sistema linear, se ele possui solução ou não e, no caso de possuir, uma maneira de encontrar todas as soluções do sistema.

Observação Adotaremos algumas simplificações na notação para sistemas lineares:

- Equações que envolvem coeficientes negativos terão um termo da forma

  $$\cdots + (-a)y + \ldots,$$

  com $a > 0$, reescritos na forma

  $$\cdots - ay + \ldots.$$

  Por exemplo, no Exemplo 1.1.1, o sistema (1.3) pode ser reescrito na forma

  $$\begin{cases} 2x_1 - 2x_2 = 2 \\ 2x_1 + 2x_2 = 6. \end{cases}$$

- Quando o número de incógnitas é pequeno, podemos rotulá-las por $x, y, z, \ldots$. Assim, o sistema (1.3) pode ser escrito, ainda, como

  $$\begin{cases} 2x - 2y = 2 \\ 2x + 2y = 6. \end{cases}$$

- Quando o número 1 aparecer como coeficiente de alguma incógnita, esse coeficiente será omitido, e quando o número 0 aparecer como coeficiente de alguma incógnita, o termo correspondente será omitido. Por exemplo, será mais comum denotarmos o sistema

  $$\begin{cases} 2x - 1y = 3 \\ 0x + 2y = 1 \end{cases}$$

  por

  $$\begin{cases} 2x - y = 3 \\ 2y = 1. \end{cases}$$

Também vale registrar que os termos *variável* ou *indeterminada* são utilizados como sinônimo de incógnita.

Sistemas lineares serão categorizados em três classes. Dizemos que um sistema linear é

- *impossível* (ou *incompatível*, ou, ainda, *inconsistente*) se não possuir soluções;
- *possível* (ou *compatível*, ou, ainda, *consistente*) *determinado* se possuir apenas uma solução;
- *possível* (ou *compatível*, ou, ainda, *consistente*) *indeterminado* se possuir pelo menos duas soluções.

Assim, nos três exemplos acima, o sistema (1.3) é possível determinado, o sistema (1.4), impossível, e o sistema (1.5), possível indeterminado.

Veremos que todos sistema linear possível indeterminado possui, de fato, infinitas soluções.

Notação matricial

O sistema linear

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

pode ser mais compactamente denotado por

$$AX = B,$$

em que

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

é uma *matriz* $m \times n$ — o que será doravante denotado por $A \in M_{m\times n}(\mathbb{R})$ —,

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

é uma matriz $m \times 1$ (ou $B \in M_{m\times 1}(\mathbb{R})$), e

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

é uma "matriz" $n \times 1$ cujas entradas são incógnitas.

A matriz

$$[A \mid B] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right] \in M_{m\times(n+1)}(\mathbb{R})$$

é chamada *matriz aumentada* do sistema.

Operações elementares

Para obter um método de resolução de sistemas, dado um sistema linear, em notação matricial,

$$AX = B \qquad (\mathcal{S})$$

consideraremos as seguintes operações sobre as equações que compõem $(\mathcal{S})$:

I. Permutar duas equações de $(\mathcal{S})$.
II. Multiplicar todos os coeficientes de uma equação de $(\mathcal{S})$ por um mesmo número real não nulo.

III. Somar a uma das equações de $(\mathcal{S})$ uma outra equação cujos coeficientes foram todos multiplicados por um mesmo número real.

As operações I, II e III são chamadas *operações elementares* sobre as equações de $(\mathcal{S})$.

Assim, quando se efetua uma operação elementar sobre as equações de um sistema linear, obtém-se um novo sistema linear com o mesmo número de incógnitas e o mesmo número de equações. (Especialmente no caso de operações elementares do tipo III, observe que o que se faz é substituir uma equação, a j-ésima, digamos, por uma nova equação: a j-ésima equação original somada a uma outra equação multiplicada, por sua vez, por um número real. Veremos, abaixo, um exemplo concreto.)

Operações elementares sobre as equações de um sistema linear não alteram suas soluções. Em outras palavras, se $(\mathcal{S}_1)$ é um sistema linear obtido a partir de $(\mathcal{S})$ pela aplicação de uma operação elementar sobre suas equações, então, dada uma sequência de números reais $(s_1, s_2, \dots, s_n)$, temos que $(s_1, s_2, \dots, s_n)$ é solução de $(\mathcal{S})$ se, e somente se, $(s_1, s_2, \dots, s_n)$ é solução de $(\mathcal{S}_1)$. É fácil ver que toda solução de $(\mathcal{S})$ é também uma solução de $(\mathcal{S}_1)$. Isso pode ser verificado considerando, um por vez, cada um dos três tipos de operações elementares sobre equações. Por outro lado, o mesmo argumento mostra que toda solução de $(\mathcal{S}_1)$ é solução de $(\mathcal{S})$. Isso decorre do fato de que operações elementares sobre equações podem ser "desfeitas" e, assim, $(\mathcal{S})$ pode ser obtido a partir de $(\mathcal{S}_1)$ por meio da aplicação de uma operação elementar sobre as equações de $(\mathcal{S}_1)$. Por exemplo, digamos que $(\mathcal{S}_1)$ tenha sido obtido a partir de $(\mathcal{S})$ por meio de uma operação elementar do tipo III, em que a j-ésima equação de $(\mathcal{S}_1)$ é o resultado da soma da j-ésima equação de $(\mathcal{S})$ com a i-ésima equação de $(\mathcal{S})$ multiplicada pelo número λ. Então, como o leitor pode facilmente verificar, $(\mathcal{S})$ pode ser obtido a partir de $(\mathcal{S}_1)$ por meio de uma operação elementar do tipo III: a j-ésima equação de $(\mathcal{S})$ é o resultado da soma da j-ésima equação de $(\mathcal{S}_1)$ com a i-ésima equação de $(\mathcal{S}_1)$ multiplicada por $(-\lambda)$. Deixamos a cargo do leitor verificar que um raciocínio similar se aplica a operações elementares dos tipos I e II.

Dizemos que dois sistemas lineares com o mesmo número de incógnitas são *equivalentes* se tiverem exatamente as mesmas soluções. Assim, o que vimos acima é que se $(\mathcal{S}_1)$ foi obtido por meio da aplicação de uma operação elementar sobre as equações de $(\mathcal{S})$, então $(\mathcal{S})$ e $(\mathcal{S}_1)$ são equivalentes.

Nosso objetivo será, dado um sistema linear $(\mathcal{S})$, obter um sistema linear $(\mathcal{S}')$ que é equivalente a ele, cujas soluções são fáceis (num sentido a ser tornado preciso adiante) de ser encontradas. Esse sistema $(\mathcal{S}')$ será obtido por meio de uma sequência finita de operações elementares sobre equações a começar pelo sistema $(\mathcal{S})$.

Antes, porém, de descrever o método, consideremos um exemplo.

Exemplo 1.1.4 Considere o sistema linear

$$\begin{cases} 2x - y + z = 4 \\ x - y + z = 1 \\ 3x - 6y + 6z = 0. \end{cases} \tag{1.6}$$

Se aplicarmos uma operação elementar de tipo I a esse sistema, permutando a primeira e a segunda equação, obtemos o sistema

$$\begin{cases} x - y + z = 1 \\ 2x - y + z = 4 \\ 3x - 6y + 6z = 0, \end{cases}$$

que é equivalente a (1.6). Agora, aplique a esse sistema uma operação elementar do tipo III, somando à segunda equação a primeira multiplicada por -2:

$$\begin{cases} x - y + z = 1 \\ y - z = 2 \\ 3x - 6y + 6z = 0. \end{cases}$$

Esse novo sistema é equivalente ao anterior, que, por sua vez, era equivalente a (1.6); portanto é ele, também, equivalente a (1.6). Agora, a esse último sistema aplicamos uma operação elementar do tipo III, somando à terceira equação a primeira multiplicada por -3:

$$\begin{cases} x - y + z = 1 \\ y - z = 2 \\ -3y + 3z = -3. \end{cases}$$

Mais uma vez, obtemos um sistema que é equivalente a (1.6). Finalmente, aplique uma operação elementar do tipo III, somando à terceira equação a segunda multiplicada por 3 e obtenha o sistema

$$\begin{cases} x - y + z = 1 \\ y - z = 2 \\ 0 = 3. \end{cases} \tag{1.7}$$

O sistema (1.7) é obviamente impossível (uma vez que sua terceira equação não tem solução). Como ele é equivalente a (1.6), pois foi obtido a partir dele por uma sequência de operações elementares sobre equações, segue que o sistema (1.6) é, também, impossível.

A cada operação elementar sobre as equações de um sistema $AX = B$ corresponde uma operação elementar sobre as linhas da matriz aumentada $[A \mid B]$ do sistema:

I. Permutar duas linhas de $[A \mid B]$.
II. Multiplicar todas as entradas de uma linha de $[A \mid B]$ por um mesmo número real não nulo.
III. Somar a uma das linhas de $[A \mid B]$ uma outra linha cujas entradas foram todas multiplicadas por um mesmo número real.

Assim, a sequência de operações elementares realizada no Exemplo 1.1.4 pode ser resgistrada como uma sequência de operações elementares sobre as linhas da matriz aumentada de (1.6):

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

Para descrever o método que aplicamos no exemplo acima, vamos introduzir uma definição.

Definição Uma matriz é dita ser *escalonada* se ambas as seguintes condições estiverem satisfeitas:

(i) todas suas linhas nulas (se existirem) estão abaixo das linhas não nulas, e
(ii) em cada linha não nula, o primeiro elemento não nulo (lendo da esquerda para a direita), chamado *pivô*, ocorre mais à direita do que o pivô da linha imediatamente acima dela.

Por exemplo, as matrizes $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ não são escalonadas, mas $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ é.

Observação Se $R \in M_{m \times n}(\mathbb{R})$ é uma matriz escalonada e R tem p pivôs, então $p \leq m$ e $p \leq n$. A primeira desigualdade segue do fato de que o número de pivôs em uma matriz escalonada é igual ao número de linhas não nulas dela. A segunda desigualdade é consequência do fato de que em cada coluna de uma matriz escalonada, há, no máximo, um pivô.

O processo de escalonamento

A seguir, apresentamos o resultado que está por trás do método de resolução de sistemas que adotaremos nesse curso. O teorema garante que o método se aplica a qualquer sistema. Sua demonstração é algorítmica, isto é, é apresentada em forma de uma sequência de passos que, uma vez seguidos, conduzem à tese do teorema. O procedimento a ser apresentado na demonstração do teorema é conhecido como *algoritmo de eliminação gaussiana* ou *processo de escalonamento*.

Teorema 1.1.5 *Para toda matriz, existe uma sequência de operações elementares sobre linhas que resulta em uma matriz escalonada.*

Demonstração Dada uma matriz A, submeta-a à seguinte sequência de operações elementares sobre linhas:

Passo 1: Se A é a matriz nula, pare. (Ela já é escalonada.)

Passo 2: Se não, encontre a primeira coluna (da esquerda para a direita) de A que contém uma entrada não nula — digamos a — e permute a linha que contém essa entrada com a primeira linha, se necessário.

Passo 3: Por adição de múltiplos adequados da primeira linha nas linhas abaixo dela, zere todas as entradas abaixo do elemento a.

Passo 4: Repita os passos 1–4 na matriz que consiste das linhas abaixo da primeira.

O processo termina quando não houver mais linhas no passo 4 ou quando as linhas restantes forem nulas. Em qualquer dos casos, a matriz resultante será escalonada. □

Vejamos um exemplo do processo de escalonamento de uma matriz.

No que segue, uma operação elementar do tipo I que permuta as linha i e j será denotada por $L_i \leftrightarrow L_j$; uma operação elementar do tipo II que multiplica a linha i por $\lambda \neq 0$ será denotada por $L_i \leftarrow \lambda L_i$; e uma operação elementar do tipo III que soma à linha j a linha i multiplicada por λ será denotada por $L_j \leftarrow L_j + \lambda L_i$. Os pivôs serão destacados em vermelho para facilitar sua identificação.

Exemplo 1.1.6 Apliquemos o processo de escalonamento à seguinte matriz 4×6:

$$\begin{bmatrix} 0 & 0 & 0 & 2 & 1 & 9 \\ 0 & -2 & -6 & 2 & 0 & 2 \\ 0 & 2 & 6 & -2 & 2 & 0 \\ 0 & 3 & 9 & 2 & 2 & 19 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_2} \begin{bmatrix} 0 & -2 & -6 & 2 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1 & 9 \\ 0 & 2 & 6 & -2 & 2 & 0 \\ 0 & 3 & 9 & 2 & 2 & 19 \end{bmatrix} \xrightarrow{L_3 \leftarrow L_3 + L_1} \begin{bmatrix} 0 & -2 & -6 & 2 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1 & 9 \\ 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 3 & 9 & 2 & 2 & 19 \end{bmatrix}$$

$$\xrightarrow{L_4 \leftarrow L_4 + \frac{3}{2} L_1} \begin{bmatrix} 0 & -2 & -6 & 2 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1 & 9 \\ 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 5 & 2 & 22 \end{bmatrix} \xrightarrow{L_4 \leftarrow L_4 + (-\frac{5}{2}) L_2} \begin{bmatrix} 0 & -2 & -6 & 2 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1 & 9 \\ 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \xrightarrow{L_4 \leftarrow L_4 + \frac{1}{4} L_3} \begin{bmatrix} 0 & -2 & -6 & 2 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1 & 9 \\ 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

A última matriz na sequência acima é uma matriz escalonada.

Observação No processo de escalonamento de uma matriz, é comum haver escolhas diferentes que podem ser feitas no passo 2. No exemplo acima, por exemplo, poderíamos ter realizado a operação $L_1 \leftrightarrow L_3$ ou $L_1 \leftrightarrow L_4$ no lugar de $L_1 \leftrightarrow L_2$, logo na primeira etapa. Essas escolhas diferentes conduziriam, ao final do processo, a uma matriz escalonada diferente da que obtivemos.

Em resumo, a matriz escalonada obtida ao final do processo de escalonamento não está univocamente determinada pela matriz com que começamos.

Método para resolução de sistemas lineares

Podemos aplicar o Teorema 1.1.5 para desenvolver um método de resolução de sistemas lineares.

Seja $(\mathcal{S})$ um sistema linear de m equações e n incógnitas, com matriz aumentada $M \in M_{m\times(n+1)}(\mathbb{R})$. Pelo Teorema 1.1.5, existe uma matriz escalonada R que pode ser obtida a partir de M por meio de operações elementares sobre linhas. Chamemos de $(\mathcal{S}')$ o sistema linear que tem R como matriz aumentada. Como $(\mathcal{S}')$ foi obtido a partir de $(\mathcal{S})$ por meio de operações elementares sobre suas equações (exatamente a mesma sequência que produziu R a partir de M), esses sistemas são equivalentes. Vejamos como encontrar as soluções de $(\mathcal{S}')$.

Há duas possibilidades:

Caso 1: A matriz R tem um pivô na coluna $n + 1$. Neste caso, $(\mathcal{S}')$ tem uma equação da forma $0 = b$, com $b \neq 0$. Logo $(\mathcal{S}')$, e, portanto, também, $(\mathcal{S})$, não tem solução.

Caso 2: A matriz R não tem pivô na coluna $n+1$. Neste caso, se p denota o número de pivôs de R, então $p \leq n$ (pois, como vimos, em uma matriz escalonada há, no máximo um pivô por coluna). Há dois subcasos a considerar:

(a) Se $p = n$, então $(\mathcal{S}')$, e, portanto, também, $(\mathcal{S})$, tem uma única solução, que é obtida resolvendo a primeira equação não nula, de baixo para cima, de $(\mathcal{S}')$, determinando-se, assim, o único valor possível para a variável x_n. Esse valor é substituído na equação exatamente acima dela, da qual resultará o único valor possível para x_{n-1}, e, assim, por diante, sempre substituindo-se em uma equação os únicos valores obtidos para as variáveis nas equações abaixo dela.

(b) Se $p < n$, então $(\mathcal{S}')$, e, portanto, também, $(\mathcal{S})$, tem infinitas soluções. Neste caso, a variáveis de $(\mathcal{S}')$ correspondentes às colunas de R, exceto a última, que não têm pivô são chamadas *variáveis livres*. As variáveis livres de $(\mathcal{S}')$ são em número de $n - p$. Para cada escolha independente de valores reais para as variáveis livres, uma solução para $(\mathcal{S}')$ pode ser obtida pelo mesmo método utilizado no caso (a). Assim, atribuem-se parâmetros independentes às variáveis livres e escrevem-se as demais (chamadas *variáveis pivô*) em termos das livres.

O método para encontrar soluções descrito no caso 2, acima, começando pela primeira equação não nula, de baixo para cima, e substituindo os valores encontrados nas equações acima dela é chamado *retrossubstituição*.

Observação Note que segue da análise acima que se um sistema é possível e indeterminado, isto é, se tem pelo menos duas soluções, então, necessariamente ele está no caso 2(b) e, portanto, tem, de fato, infinitas soluções.

Exemplo 1.1.7 Encontre todas as soluções do sistema linear

$$\begin{cases} x - 2y - z = 1 \\ 2x + y - 3z = 0 \\ x - 7y = 3. \end{cases}$$

Solução: Vamos escalonar a matriz aumentada do sistema:

$$\left[\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ 2 & 1 & -3 & 0 \\ 1 & -7 & 0 & 3 \end{array}\right] \xrightarrow[L_3 \leftarrow L_3 + (-1)L_1]{L_2 \leftarrow L_2 + (-2)L_1} \left[\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ 0 & 5 & -1 & -2 \\ 0 & -5 & 1 & 2 \end{array}\right] \xrightarrow{L_3 \leftarrow L_3 + L_2} \left[\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ 0 & 5 & -1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Estamos no caso 2(b), em que não há pivô na última coluna e existem 2 pivôs, destacados em vermelho. Os pivôs ocorrem nas colunas correspondentes às variáveis x e y, que são, portanto, variáveis pivô. A coluna correspondente à variável z não contém pivô; assim, z é uma variável livre. Atribuímos um parâmetro a ela, digamos $z = t$ e resolvemos o sistema resultante,

$$\begin{cases} x - 2y - t = 1 \\ 5y - t = -2, \end{cases}$$

por retrossubstituição. Da última equação, resolvendo para y, obtemos $y = -\frac{2}{5} + \frac{1}{5}t$. Substituindo esse valor na primeira equação e resolvendo para x, obtemos $x = \frac{1}{5} + \frac{7}{5}t$. Logo, as soluções do sistema original são dadas por

$$\left(\frac{1}{5} + \frac{7}{5}t,\ -\frac{2}{5} + \frac{1}{5}t,\ t\right), \quad t \in \mathbb{R}.$$

Soluções particulares são obtidas escolhendo valores para o parâmetro t. Por exemplo, tomando $t = 1$, encontramos a solução $(\frac{8}{5}, -\frac{1}{5}, 1)$. $\diamond$

Sistemas lineares homogêneos

Um sistema linear da forma $AX = 0$, em que 0 denota uma matriz de uma única coluna cujas entradas são todas iguais ao número 0, é dito *homogêneo*.

Todo sistema homogêneo é possível, uma vez que tem, pelo menos, a *solução trivial*: $(0, 0, \dots, 0)$.

Teorema 1.1.8 *Um sistema linear homogêneo com mais incógnitas do que equações é possível indeterminado.*

Demonstração Se p denota o número de pivôs da matriz escalonada obtida pelo processo de escalonamento da matriz aumentada de um sistema homogêneo com m equações e incógnitas, então, como vimos $p \leq m$. Mas, por hipótese, $m < n$. Isso resulta em $p < n$ e, portanto, estamos no caso 2(b) (pois, certamente, essa matriz escalonada tem como última coluna uma coluna de zeros, sem pivôs, portanto). □

É importante observar que o teorema nada afirma sobre sistemas homogêneos com número de incógnitas menor ou igual ao número de equações. Um sistema desses pode ser ou determinado ou indeterminado. (Você consegue produzir exemplos para os dois casos?)

Outra ressalva é a de que o teorema apenas se aplica a sistemas homogêneos. É muito fácil dar exemplos de sistemas não homogêneos com mais incógnitas do que equações e que são impossíveis.

Exercícios Lista 1 - Álgebra Linear I: Exs. 1–12.

1.2 Matrizes

Vimos, na seção anterior. que matrizes podem ser utilizadas para denotar sistemas lineares de modo mais compacto e, mais importante, que podem ser sujeitadas ao processo de escalonamento.

Nesta seção, veremos que existem operações que nos permitirão falar na álgebra de matrizes.

Se $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \in M_{m\times n}(\mathbb{R})$, escreveremos, a título de abreviação, $A = (a_{ij})_{1\leq i\leq m, 1\leq j\leq n}$, ou, simplesmente, $A = (a_{ij})$, quando não houver dúvida a respeito do tamanho de A.

Dizemos que duas matrizes são iguais se tiverem o mesmo tamanho e suas entradas correspondentes forem iguais. Mais detalhadamente, dadas $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$ e $B = (b_{ij}) \in M_{r\times s}(\mathbb{R})$, então $A = B$ se, e somente se, $m = r$, $n = s$ e, para cada $i = 1, \dots, m$ e $j = 1, \dots, n$, tivermos $a_{ij} = b_{ij}$.

Dadas matrizes $A = (a_{ij}), B = (b_{ij}) \in M_{m\times n}(\mathbb{R})$, a soma $A + B \in M_{m\times n}(\mathbb{R})$ é definida como sendo a matriz dada por $A + B = (c_{ij}) \in M_{m\times n}(\mathbb{R})$, em que $c_{ij} = a_{ij} + b_{ij}$, para todos $i = 1, \dots, m$, $j = 1, \dots n$. Por exemplo,

$$\begin{bmatrix} 2 & 3 & 4 \\ 1 & 0 & 2 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 2 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 6 \\ 1 & 0 & 4 \end{bmatrix}.$$

Dada uma matriz $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$ e dado um número real λ, define-se a multiplicação de λ por A como sendo a matriz dada por $\lambda A = (d_{ij}) \in M_{m\times n}(\mathbb{R})$, em que $d_{ij} = \lambda a_{ij}$, para todos $i = 1, \dots, m$, $j = 1, \dots n$. Por exemplo,

$$\frac{1}{3}\begin{bmatrix} 2 & 3 & 9 \\ -1 & 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & 1 & 3 \\ -\frac{1}{3} & 0 & \frac{1}{6} \end{bmatrix}.$$

O produto de duas matrizes, a ser definido em seguida, não envolve matrizes de mesmo tamanho. Dada uma matriz $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$ e uma matriz $B = (b_{ij}) \in M_{n\times r}(\mathbb{R})$, define-se o produto $AB \in M_{m\times r}(\mathbb{R})$, como sendo a matriz dada por $AB = (e_{ij})$, em que

$$e_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj},$$

para todos $i = 1, \ldots, m$ e $j = 1, \ldots, r$. Em poucas palavras, o produto de uma matrix $m \times n$ por uma matriz $n \times r$ é uma matriz $m \times r$ cuja entrada na posição (i, j) é dada pelo produto da linha i de A pela coluna j de B. Por exemplo,

$$\begin{bmatrix} 2 & 1 & -1 \\ 3 & 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} \\ 1 & 2 & 1 & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 3 & -2 & 0 & 1 \\ 10 & 8 & 6 & -1 \end{bmatrix}.$$

As operações definidas entre matrizes satisfazem as seguintes propriedades, que podem ser diretamente demonstradas a partir das definições.

Proposição 1.2.1 *Sejam A, B, C, M, N, L matrizes e sejam $\lambda, \mu \in \mathbb{R}$. Então, se estão definidas as matrizes no lado esquerdo das igualdades, as matrizes do lado direito também estão definidas e elas são iguais:*

(i) $A + B = B + A$;
(ii) $(A + B) + C = A + (B + C)$;
(iii) $(AM)N = A(MN)$;
(iv) $A(M + L) = AM + AL$;
(v) $(A + B)M = AM + BM$;
(vi) $\lambda(A + B) = \lambda A + \lambda B$;
(vii) $(\lambda + \mu)A = \lambda A + \mu A$;
(viii) $\lambda(AM) = (\lambda A)M = A(\lambda M)$;
(ix) $(\lambda\mu)A = \lambda(\mu A)$.

Ou seja, muitas das propriedades das operações entre números reais continuam válidas para matrizes. Porém, ao contrário do produto entre números, o produto entre matrizes não é comutativo. Isto é, se A e B são matrizes e ambos os produtos AB e BA estão definidos, nem sempre eles são iguais. Por exemplo,

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 6 & 4 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

Uma matriz que tem o mesmo número de linhas que o númereo de colunas é chamada *matriz quadrada*. O conjunto de todas as matrizes quadradas $n \times n$ será denotado, simplesmente, por $M_n(\mathbb{R})$.

A *matriz identidade* de tamanho $n \times n$ é definida como sendo a matriz quadrada $I_n = (e_{ij}) \in M_n(\mathbb{R})$, em que $e_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases}$. Por exemplo, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ e $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. É fácil concluir que para qualquer matriz $A \in M_{m\times n}(\mathbb{R})$, tem-se $I_m A = A = A I_n$.

Definição Dizemos que uma matriz quadrada $A \in M_n(\mathbb{R})$ é *inversível* se existir uma matriz $B \in M_n(\mathbb{R})$ tal que $AB = I_n$ e $BA = I_n$.

Por exemplo, a matriz $A = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}$ é inversível, pois tomando-se $B = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$, vale $AB = BA = I_2$. Já a matriz $C = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$ não é inversível, pois se $D = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R})$ é uma matriz arbitrária, temos $CD = \begin{bmatrix} 2a + c & 2b + d \\ 0 & 0 \end{bmatrix}$, que não é a matriz identidade, quaiquer que sejam $a, b, c, d \in \mathbb{R}$.

Proposição 1.2.2 *Seja $A \in M_n(\mathbb{R})$ uma matriz inversível. Então, existe uma única matriz $B \in M_n(\mathbb{R})$ tal que $AB = BA = I_n$. Essa matriz será denominada* matriz inversa *de A e será denotada por A^{-1}.*

Demonstração Sejam $B, C \in M_n(\mathbb{R})$ tais que $AB = BA = I_n$ e $AC = CA = I_n$. Então, $B = I_n B = (CA)B = C(AB) = C I_n = C$. □

As demonstrações das propriedades enunciadas na próxima proposição são simples e ficam a cargo do leitor.

Proposição 1.2.3 *Sejam $A, B \in M_n(\mathbb{R})$ matrizes inversíveis. Então,*

(i) a matriz AB é inversível e $(AB)^{-1} = B^{-1}A^{-1}$;
(ii) a matriz A^{-1} é inversível e $(A^{-1})^{-1} = A$;
(iii) a matriz I_n é inversível e $I_n^{-1} = I_n$.

Veremos, na próxima seção, que se uma matriz tiver um inverso de um lado, então ela será automaticamente inversível. Mais precisamente, veremos, após a demonstração do Teorema 1.3.6, que se $A, B \in M_n(\mathbb{R})$ são tais que $AB = I_n$, então, A é inversível e $B = A^{-1}$. (Em particular, $BA = I_n$.)

Observação Vale ressaltar que o conceito de invertibilidade só se aplica a matrizes quadradas. Para matrizes não quadradas, fenômenos "patológicos" podem ocorrer; por exemplo, considere as matrizes $A = \begin{bmatrix} 2 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Então, $AB = I_1$, mas $BA \neq I_2$.

Método prático para inversão de matrizes

Uma matriz $n \times n$ é chamada *matriz elementar* se pode ser obtida a partir da matriz identidade I_n por meio de uma operação elementar sobre as linhas de I_n. Vejamos alguns exemplos:

- A matriz $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ é uma matriz elementar, pois foi obtida a partir de I_3 pela permutação das linhas 2 e 3.
- A matriz $F = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ é uma matriz elementar, pois foi obtida a partir de I_3 por meio da multiplicação da primeira linha por -2.
- A matriz $G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$ é uma matriz elementar, pois foi obtida a partir de I_3 por meio da soma da primeira linha multiplicada por 3 à terceira linha.

O efeito da multiplicação de uma matriz elementar pela esquerda é bem familiar:

Proposição 1.2.4 *Seja $A \in M_{m\times n}(\mathbb{R})$ e seja $E \in M_m(\mathbb{R})$ uma matriz elementar. Então EA é a matriz obtida a partir de A pela mesma aplicação da operação elementar sobre linhas que produziu E a partir de I_m.*

Demonstração Este resultado pode ser demonstrado considerando-se, um por vez, cada um dos três tipos de operação elementar sobre linhas de uma matriz. □

Por exemplo, considere as matrizes

$$A = \begin{bmatrix} 1 & 2 & 1 \\ -2 & 3 & 7 \\ \frac{1}{2} & -1 & \frac{3}{2} \end{bmatrix} \xrightarrow{L_3 \leftarrow L_3 + (-\frac{1}{2})L_1} \begin{bmatrix} 1 & 2 & 1 \\ -2 & 3 & 7 \\ 0 & -2 & 1 \end{bmatrix} = B.$$

Então, $B = EA$ e E é a matriz obtida por

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{L_3 \leftarrow L_3 + (-\frac{1}{2})L_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} = E.$$

Vimos que toda operação elementar sobre as linhas de uma matriz pode ser desfeita por uma operação elementar do mesmo tipo. Logo, toda matriz elementar é inversível, e sua inversa é também elementar: se E é obtida a partir de I_n por uma operação elementar, seja F a matriz obtida a partir de I_n pela operação elementar que a desfaz. Então $FE = I_n = EF$.

A seguir, introduziremos um resultado que dará origem a um método prático para encontrar a inversa de uma matriz, quando existir.

Teorema 1.2.5 *Se existe uma sequência de operações elementares sobre linhas que, a partir de uma matriz $A \in M_n(\mathbb{R})$, produz a matriz identidade I_n, então A é inversível. Além disso, a mesma sequência de operações elementares sobre linhas que produziu I_n a partir de A produz A^{-1} a partir de I_n.*

Demonstração Como toda operação elementar sobre linhas pode ser realizada por multiplicação à esquerda por matrizes elementares, dizer que existe uma sequência de operações elementares que produz I_n a partir de A é dizer que existem matrizes elementares $E_1, E_2, \dots, E_r \in M_n(\mathbb{R})$ tais que

$$E_r E_{r-1} \dots E_2 E_1 A = I_n.$$

Como produto de matrizes inversíveis é inversível, a matriz $E_r E_{r-1} \dots E_2 E_1$ é inversível. Multiplicando-se a igualdade acima pela inversa dessa matriz à esquerda, obtemos

$$A = (E_r E_{r-1} \dots E_2 E_1)^{-1}(E_r E_{r-1} \dots E_2 E_1 A) = (E_r E_{r-1} \dots E_2 E_1)^{-1} I_n = (E_r E_{r-1} \dots E_2 E_1)^{-1}.$$

Portanto, A é inversível (pois é a inversa de uma matriz inversível) e

$$A^{-1} = \left((E_r E_{r-1} \dots E_2 E_1)^{-1}\right)^{-1} = E_r E_{r-1} \dots E_2 E_1 = E_r E_{r-1} \dots E_2 E_1 I_n,$$

que é a matriz obtida a partir de I_n pela mesma aplicação da sequência de operações elementares que produziu I_n a partir de A. □

Costumamos registrar o método obtido na demonstração do teorema acima da seguinte maneira:

$$[A \mid I_n] \longrightarrow [I_n \mid A^{-1}].$$

Vejamos um exemplo.

Exemplo 1.2.6 Encontre a inversa da matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$.

Solução: Realizamos operações elementares simultâneas sobre as linhas de A e da matriz identidade I_3 com o intuito de obter, a partir de A, a matriz identidade I_3:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right] \xrightarrow[L_3 \leftarrow L_3 + (-1)L_1]{L_2 \leftarrow L_2 + (-2)L_1} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right] \xrightarrow{L_3 \leftarrow L_3 + 2L_2} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right]$$

$$\xrightarrow{L_3 \leftarrow (-1)L_3} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \xrightarrow[L_1 \leftarrow L_1 + (-3)L_3]{L_2 \leftarrow L_2 + 3L_3} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \xrightarrow{L_1 \leftarrow L_1 + (-2)L_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

Segue que $A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$. ◇

Dois comentários são devidos. O primeiro é que a sequência de operações elementares à qual a matriz A deve ser submetida a fim de produzir a matriz identidade pode, sempre, ser realizada na seguinte ordem:

1. Por escalonamento, produz-se uma matriz escalonada a partir de A.
2. Em seguida, por operações elementares do tipo II, transformam-se todos os pivôs em 1.
3. Finalmente, por "retroescalonamento", isto é, "escalonamento de baixo para cima", zeram-se todas as entradas acima dos pivôs.

Essa foi a ordem adotada no exemplo acima.

O segundo comentário diz respeito ao fato de só podermos aplicar o método se soubermos, de antemão, que a matriz A é inversível. Veremos, na próxima seção, que há um critério simples para determinar se uma matriz é ou não inversível sem necessariamente exibir a inversa. Além disso, também veremos que, no caso de ser inversível, o processo de escalonamento sempre conduzirá a uma matriz escalonada com pivôs em todas as linhas, o que garante que o procedimento acima sempre poderá ser aplicado.

Exercícios Lista 1 - Álgebra Linear I: Exs. 13–17.

1.3 Determinantes

Nesta seção, seremos menos rigorosos em nossos argumentos. Muitos dos resultados serão enunciados sem demonstração. Recomendamos, ao leitor interessado em maiores detalhes, a leitura de [5, Seções 2.1 e 2.2] ou [1, Capítulo 2]. Para um tratamento rigoroso da teoria de determinantes, recomendamos [4, Capítulo 7].

O *determinante* de uma matriz quadrada A, denotado por $\det(A)$, é um número real associado a ela que contém informação relevante a respeito de sua invertibilidade.

Para uma matriz 1×1, seu determinante será definido como sendo o número real que ocupa sua única entrada: $\det[a] = a$. Para o cálculo do determinante de matrizes de tamanho maior, usaremos os fatos que serão listados abaixo.

Necessitaremos dos seguintes conceitos a respeito de matrizes:

- Uma matriz quadrada $A = (a_{ij}) \in M_n(\mathbb{R})$ é chamada *triangular superior* se $a_{ij} = 0$ sempre que $i > j$, ou seja, se A tiver o formato
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}$$
- Dada uma matriz (não necessariamente quadrada) $B = (b_{ij}) \in M_{m\times n}(\mathbb{R})$, define-se a matriz *transposta* de B como sendo a matriz $B^\intercal = (c_{ij}) \in M_{n\times m}(\mathbb{R})$, em que $c_{ij} = b_{ji}$, para todos $i = 1, \dots, m$, $j = 1, \dots, n$, ou seja, a i-ésima linha de $B^\intercal$ é a i-ésima coluna de B. Por exemplo,
$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^\intercal = \begin{bmatrix} 1 & 4 \\ 2 & 3 \\ 3 & 6 \end{bmatrix}.$$

  Não é difícil demonstrar que

  (i) ${I_n}^\intercal = I_n$;
  (ii) se A e B são matrizes tais que o produto AB esteja definido, então $(AB)^\intercal = B^\intercal A^\intercal$;
  (iii) se D é uma matriz inversível, então $D^\intercal$ é inversível, e $(D^\intercal)^{-1} = (D^{-1})^\intercal$.

Os fatos (assumidos sem demonstração) que nos permitirão calcular determinantes de matrizes de tamanhos arbitrários são os seguintes:

Fato 1. Se $A = (a_{ij}) \in M_n(\mathbb{R})$ é uma matriz triangular superior, então
$$\det(A) = a_{11}a_{22}a_{33}\dots a_{nn}.$$
(Em particular, $\det(I_n) = 1$.)

Fato 2. Para qualquer matriz quadrada A, vale
$$\det(A) = \det(A^\intercal).$$

Fato 3. Os efeitos das operações elementares sobre linhas de uma matriz quadrada A em seu determinante são os seguintes:

I. Se B é obtida de A por meio da permutação de duas linhas de A, então
$$\det(B) = -\det(A).$$

II. Se B é obtida a partir de A por meio da multiplicação de uma das linhas de A pelo número λ (não necessariamente não nulo), então
$$\det(B) = \lambda \det(A).$$

III. Se B é obtida a partir de A por meio da adição de uma linha multiplicada por um número a uma outra linha, então
$$\det(B) = \det(A).$$

De posse dos fatos acima, é possível calcular o determinante de qualquer matriz quadrada fazendo uso do processo de escalonamento, uma vez que toda matriz quadrada escalonada é triangular superior.

Ainda, por causa do Fato 2, segue que o determinante de matrizes triangulares inferiores (aquelas em que as entradas acima da diagonal principal são nulas) também é dado pelo produto dos elementos na diagonal principal. Além disso, o Fato 2 também nos proporciona um método de cálculo do determinante fazendo uso de "operações elementares sobre colunas" (ou seja, valem as afirmações do Fato 3 com cada ocorrência de "linha" trocada por "coluna").

Vejamos dois exemplos. Um primeiro, numérico, e outro contendo uma fórmula familiar.

Exemplo 1.3.1 Calcule o determinante da matriz $A = \begin{bmatrix} 3 & 5 & -2 & 6 \\ 1 & 2 & -1 & 1 \\ 2 & 4 & 1 & 5 \\ 3 & 7 & 5 & 3 \end{bmatrix}$.

Solução: Vamos escalonar A:
$$A = \begin{bmatrix} 3 & 5 & -2 & 6 \\ 1 & 2 & -1 & 1 \\ 2 & 4 & 1 & 5 \\ 3 & 7 & 5 & 3 \end{bmatrix} \to \begin{bmatrix} 0 & -1 & 1 & 3 \\ 1 & 2 & -1 & 1 \\ 0 & 0 & 3 & 3 \\ 0 & 1 & 8 & 0 \end{bmatrix} = B.$$

Aqui, fizemos uso de operações elementares do tipo III utilizando a segunda linha para zerar três entradas na primeira coluna. Como foram usadas apenas operações do tipo III, segue $\det(B) = \det(A)$. Agora,
$$B = \begin{bmatrix} 0 & -1 & 1 & 3 \\ 1 & 2 & -1 & 1 \\ 0 & 0 & 3 & 3 \\ 0 & 1 & 8 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & 3 & 3 \\ 0 & 1 & 8 & 0 \end{bmatrix} = C.$$

Nessa passagem, realizamos uma operação elementar do tipo I, permutando a primeira e segunda linhas. Assim, $\det(C) = -\det(B)$; portanto, como $\det(B) = \det(A)$, segue que $\det(A) = -\det(C)$. Aplicando mais uma operação elementar do tipo III, que não altera o determinante, obtemos:
$$C = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & 3 & 3 \\ 0 & 1 & 8 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 9 & 3 \end{bmatrix} = D.$$

Logo, $\det(A) = -\det(C) = -\det(D)$. Finalmente, um último uso de uma operação elementar do tipo III:

$$D = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 9 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & -6 \end{bmatrix} = E.$$

Como $\det(E) = 1 \cdot (-1) \cdot 3 \cdot (-6) = 18$, uma vez que E é triangular superior, concluímos que $\det(A) = -\det(D) = -\det(E) = -18$. $\diamond$

Exemplo 1.3.2 Mostre, usando o que vimos sobre o efeito no determiante de operações elementares sobre linhas de uma matriz, que

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

Solução: Consideremos, primeiramente, o caso em que $a = 0$. Temos

$$\det \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} = -\det \begin{bmatrix} c & d \\ 0 & b \end{bmatrix} = -cb,$$

o que comprova a fórmula neste caso.

Suponha, agora, que $a \neq 0$. Neste caso,

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \det \begin{bmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{bmatrix} = ad - bc.$$

Seguem dos fatos a respeito de determinantes que citamos acima, as seguintes propriedades.

Proposição 1.3.3 *Seja A uma matriz quadrada.*

(i) Se A tem uma linha nula, então $\det(A) = 0$.
(ii) Se A tem duas linhas iguais, então $\det(A) = 0$.
(iii) Se A tem tamanho $n \times n$ e $\lambda \in \mathbb{R}$, então $\det(\lambda A) = \lambda^n \det(A)$.

Demonstração Para (i), observe que A é obtida a partir de si mesma pela multiplicação de sua linha nula por 0. Logo, $\det(A) = 0 \det(A) = 0$. Para (ii), utilize as linhas iguais de A para produzir, por meio de uma operação elementar do tipo III, uma matriz que tem uma linha nula. Finalmente, (iii) segue da aplicação do Fato 3.II a cada uma das n linhas de A. $\square$

Uma das propriedades mais importantes do determinante de uma matriz será enunciada no resultado a seguir, apresentado sem demonstração.

Teorema 1.3.4 *Sejam A e B matrizes quadradas de mesmo tamanho. Então,*

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

Esse teorema tem a consequência imediata de que se A é uma matriz inversível, então $\det(A) \neq 0$. Com efeito, suponha que A tem tamanho $n \times n$ e que B seja uma matriz tal que $AB = I_n$. Segue do Teorema 1.3.4 que $\det(A)\det(B) = \det(AB) = \det(I_n) = 1$. Logo, $\det(A) \neq 0$. Na realidade, essa condição sobre o determinante é também suficiente para garantir a invertibilidade de A.

Corolário 1.3.5 *Seja A uma matriz quadrada. Então, A é inversível se, e somente se,* $\det(A) \neq 0$.

Demonstração Já vimos que a condição é necessária. Reciprocamente, suponha que $\det(A) \neq 0$, e seja R uma matriz escalonada obtida a partir de A por meio do processo de escalonamento. Como vimos, temos que $\det(A) \neq 0$ se, e somente se, $\det(R) \neq 0$, uma vez que operações elementares do tipo II só são permitidas, no escalonamento, com constantes não nulas. Agora, o fato de $\det(R)$ não ser nulo implica que R tem pivô em todas suas linhas. Fazendo, agora, retroescalonamento, podemos obter a matriz identidade a partir de A por meio de operações elementares sobre linhas, o que é equivalente a escrever A como um produto de matrizes elementares, que são todas inversíveis, como vimos. Logo A é inversível. $\square$

Finalmente, podemos enunciar um resultado que relaciona invertibilidade de matrizes com sistemas lineares. Para tanto, observemos, primeiramente, que se $AX = B$ é um sistema linear (em notação matricial) com n variáveis e se $(s_1, s_2, \dots, s_n)$ é uma sequência de n números reais, então $(s_1, s_2, \dots, s_n)$ é solução do sistema $AX = B$ se, e somente se, $AS = B$, em que $S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$ e AS denota o produto matricial. (É esse fato, aliás, que justifica a adoção da notação matricial para sistemas lineares.) Se, de fato, $AS = B$, será comum, deste ponto em diante, referirmo-nos à matriz S como uma solução do sistema linear $AX = B$.

Teorema 1.3.6 *Seja $A \in M_n(\mathbb{R})$. São equivalentes:*

(a) A é inversível.
(b) Para qualquer $B \in M_{n\times 1}(\mathbb{R})$, o sistema linear $AX = B$ é possível determinado.
(c) O sistema linear homogêneo $AX = 0$ é possível determinado.
(d) $\det(A) \neq 0$.

Demonstração Para ver que (a) implica (b), suponha que A é inversível e seja $S = A^{-1}B$. Mostremos que S é a única solução de $AX = B$. Primeiramente, S é, de fato, solução de $AX = B$, uma vez que $AS = A(A^{-1}B) = (AA^{-1})B = I_nB = B$. Suponha, agora, que $S' \in M_{n\times 1}(\mathbb{R})$ seja uma solução de $AX = B$. Então, $AS' = B$. Multiplicando essa igualdade por A^{-1} à esquerda em ambos os lados, obtemos $A^{-1}(AS') = A^{-1}B$, o que resulta em $S' = S$, provando que S é a única solução de $AX = B$.

É claro que (b) implica (c), pois (c) é apenas o caso particular de (b) em que $B = 0$.

Vejamos, agora, por que (c) implica (d). Se o sistema $AX = 0$ é possível determinado e $R \in M_{n\times(n+1)}(\mathbb{R})$ é uma matriz obtida no processo de escalonamento da matriz aumentada $[A \mid 0]$ de $AX = 0$, então R tem exatamente n pivôs, ou seja, R é da forma $R = [T \mid 0]$, em que T é uma matriz quadrada de tamanho $n \times n$ que é triangular superior e cujas entradas na diagonal principal são todas não nulas. Assim, $\det(T) \neq 0$. Como T foi obtida a partir de A por operações elementares sobre linhas, segue que $\det(A) \neq 0$.

O Corolário 1.3.5 dá conta do fato que (d) implica (a).

Logo, as três condições são equivalentes. □

Note que segue do resultado acima que se $A \in M_n(\mathbb{R})$ é uma matriz para a qual existe uma matriz $B \in M_n(\mathbb{R})$ tal que $AB = I_n$, então A é inversível e $A^{-1} = B$. Isso, pois de $AB = I_n$, segue que $\det(A)\det(B) = \det(AB) = \det(I_n) = 1$. Portanto, $\det(A) \neq 0$, o que implica, pela equivalência entre (a) e (d) do Teorema 1.3.6, que A é inversível. Multiplicando a igualdade $AB = I_n$ por A^{-1} à esquerda, obtemos $B = A^{-1}$. Ou seja, para demonstrar que uma matriz é inversível, basta exibir uma inversa à direita. Um argumento análogo mostra que uma matriz quadrada que tem uma inversa à esquerda é inversível.

Terminamos esta seção com uma ressalva importante. Apesar de a função determinante ser multiplicativa, de acordo com o Teorema 1.3.4, ela não é uma função aditiva, ou seja se A e B são matrizes quadradas de mesmo tamanho, então não vale sempre que $\det(A + B)$ será igual à soma dos números $\det(A)$ e $\det(B)$, como mostra, por exemplo, a conta a seguir:

$$\det\left(\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix} = 2,$$

ao passo que

$$\det\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} + \det\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = -1 + 1 = 0.$$

Observação Há outros métodos usuais para o cálculo de determinantes, como as conhecidas fórmulas de Sarrus (para determinantes de matrizes 3×3) e de Laplace (também conhecida como expansão em cofatores). Ainda, no que tange ao uso de determinates para resolução de sistemas, costuma também ser adotada a fórmula de Cramer. O leitor encontra referência a esses métodos na bibliografia indicada no início da seção.

Exercícios Lista 1 - Álgebra Linear I: Exs. 18–35.

Capítulo 2
Vetores

Neste capítulo será definido o conjunto dos vetores no espaço tridimensional e serão introduzidas operações envolvendo os elementos desse conjunto.

Admitem-se, no que segue, conhecimentos de geometria euclidiana no plano e no espaço.

O conjunto formado por todos os pontos do espaço tridimensional será denotado por $\mathbb{E}^3$.

A referência principal para este capítulo é o livro [2].

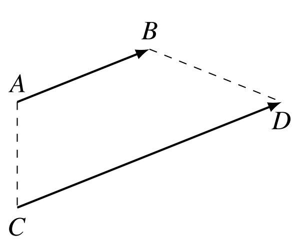
2.1 Vetores e operações entre vetores

Dados dois pontos $A, B \in \mathbb{E}^3$, o par ordenado (A, B) será denominado *segmento orientado* de extremidade inicial A e extremidade final B.

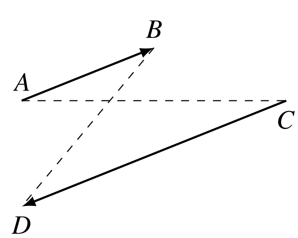
Ao segmento orientado (A, B) de $\mathbb{E}^3$ associamos um vetor, denotado por $\overrightarrow{AB}$. No conjunto dos vetores, está definida a seguinte relação de igualdade: se $A, B, C, D \in \mathbb{E}^3$, então $\overrightarrow{AB} = \overrightarrow{CD}$ se, e somente se, as seguinte três condições estiverem satisfeitas:

(i) Os segmentos de reta AB e CD têm o mesmo comprimento.
(ii) Os segmentos de reta AB e CD são paralelos.
(iii) Os segmentos orientados (A, B) e (C, D) têm o mesmo sentido.

Sejam (A, B) e (C, D) segmentos orientados tais que os segmentos de reta AB e CD sejam pararelos. Dizemos que (A, B) e (C, D) têm o *mesmo sentido* se $AC \cap BD = \varnothing$. Caso contrário, dizemos que (A, B) e (C, D) têm *sentidos opostos*.[1]



mesmo sentido



sentidos opostos

[1] Essa é a definição de segmentos orientados de mesmo sentido quando os segmentos de reta suporte são paralelos, mas não estão contidos em uma mesma reta. Se (A, B) e (C, D) são segmentos orientados com AB e CD contidos em uma mesma reta, tome pontos C', D' fora dessa reta de forma que CD e $C'D'$ sejam segmentos de reta paralelos e (C, D) e (C', D') sejam segmentos orientados de mesmo sentido. Assim (A, B) e (C, D) terão mesmo sentido se (A, B) e (C', D') tiverem mesmo sentido.

Por convenção, os segmentos de reta de comprimento nulo, isto é, aqueles da forma AA, são paralelos a qualquer outro segmento de reta. Ainda, convenciona-se também que os segmentos orientados (A, A) e (C, D) têm o mesmo sentido, quaisquer que sejam $A, C, D \in \mathbb{E}^3$.

Se $\vec{v}$ é um vetor e $A, B \in \mathbb{E}^3$ são tais que $\vec{v} = \overrightarrow{AB}$, dizemos que o segmento orientado (A, B) é um *representante* do vetor $\vec{v}$.

Para todos $A, B \in \mathbb{E}^3$, $\overrightarrow{AA} = \overrightarrow{BB}$. Esse vetor será chamado *vetor nulo* e será denotado por $\vec{0}$.

O conjunto de todos os vetores será denotado por $\mathbb{V}^3$.

Dado $\vec{v} \in \mathbb{V}^3$, sejam $A, B \in \mathbb{E}^3$ tais que $\vec{v} = \overrightarrow{AB}$. Define-se o *comprimento* (ou *norma*) de $\vec{v}$ como sendo o comprimento do segmento de reta AB. O comprimento de $\vec{v}$ será denotado por $\|\vec{v}\|$. É claro que $\left\|\vec{0}\right\| = 0$ e que o vetor nulo é o único vetor de $\mathbb{V}^3$ com comprimento igual a 0.

Dado $\vec{v} \in \mathbb{V}^3$ tal que $\vec{v} = \overrightarrow{AB}$, com $A, B \in \mathbb{E}^3$, define-se o *vetor oposto* de $\vec{v}$ como sendo o vetor $\overrightarrow{BA}$ e o denotamos por $-\vec{v}$. (Isto é, se $\vec{v} = \overrightarrow{AB}$, então $-\vec{v} = \overrightarrow{BA}$).

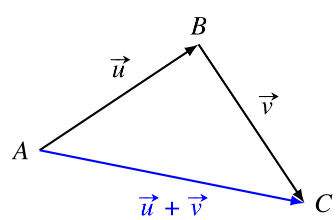
Dados dois vetores $\vec{u}, \vec{v}$, dizemos que eles são *paralelos* se tiverem representantes paralelos, isto é, se, escrevendo $\vec{u} = \overrightarrow{AB}$ e $\vec{v} = \overrightarrow{CD}$, com $A, B, C, D \in \mathbb{E}^3$, os segmentos de reta AB e CD forem paralelos. Por convenção, o vetor nulo é paralelo a qualquer vetor.

A próxima observação será crucial no que segue.

Observação Dados $A \in \mathbb{E}^3$ e $\vec{v} \in \mathbb{V}^3$, existe um único $B \in \mathbb{E}^3$ tal que $\vec{v} = \overrightarrow{AB}$. (Ou seja, fixado um ponto e um vetor, existe um representante desse vetor com extremidade inicial naquele ponto.)

Soma de vetores

Dados $\vec{u}, \vec{v} \in \mathbb{V}^3$, se $\vec{u} = \overrightarrow{AB}$, com $A, B \in \mathbb{E}^3$, seja $C \in \mathbb{E}^3$ tal que $\vec{v} = BC$. Define-se a *soma* $\vec{u} + \vec{v}$ como sendo o vetor dado por $\vec{u} + \vec{v} = \overrightarrow{AC}$.

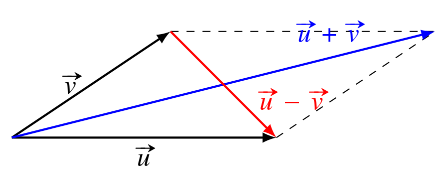


A proprosição a seguir reúne propriedades da soma entre vetores, que podem ser verificadas por meio de argumentos geométricos no plano.

Proposição 2.1.1 *Para quaisquer $\vec{u}, \vec{v}, \vec{w} \in \mathbb{V}^3$, valem:*

(i) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$*;*
(ii) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$*;*
(iii) $\vec{u} + \vec{0} = \vec{u}$*;*
(iv) $\vec{u} + (-\vec{u}) = \vec{0}$.

Uma questão notacional: se $\vec{u}, \vec{v} \in \mathbb{V}^3$, utilizaremos a abreviação $\vec{u} - \vec{v}$ para significar o vetor $\vec{u} + (-\vec{v})$. Assim, dados os vetores $\vec{u}$ e $\vec{v}$, podemos identificar os vetores $\vec{u} + \vec{v}$ e $\vec{u} - \vec{v}$ como as diagonais do paralelogramo definido por $\vec{u}$ e $\vec{v}$, como na figura:

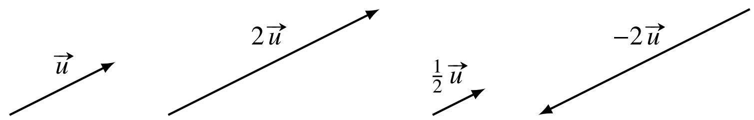


Para ver isso, chame de $\vec{x}$ o vetor destacado em vermelho. Esse vetor satistaz $\vec{v} + \vec{x} = \vec{u}$. Somando $-\vec{v}$ a ambos os lados dessa igualdade, obtemos $\vec{x} = \vec{u} - \vec{v}$.

Multiplicação de um escalar por um vetor

A partir deste ponto, usaremos o termo *escalar* para nos referirmos a um número real.

Dados um vetor $\vec{u} \in \mathbb{V}^3$ e um escalar $\lambda \in \mathbb{R}$, define-se a *multiplicação do escalar λ pelo vetor $\vec{u}$* como sendo o vetor $\lambda\vec{u}$ que tem comprimento $|\lambda|\,\|\vec{u}\|$, é pararelo a $\vec{u}$ e é tal que $\vec{u}$ e $\lambda\vec{u}$ têm o mesmo sentido se $\lambda > 0$ e sentidos opostos se $\lambda < 0$. Se $\lambda = 0$, define-se $\lambda\vec{u} = \vec{0}$.



Segue, imediatamente da definição, que para qualquer vetor $\vec{u}$, temos $(-1)\vec{u} = -\vec{u}$.

Aqui, também, argumentos de natureza geométrica permitem a demonstração das seguintes propriedades.

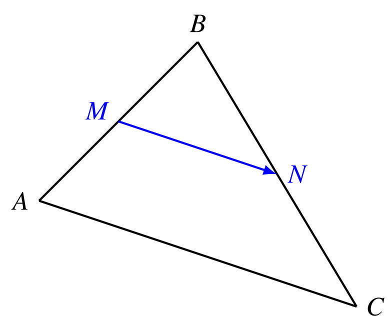
Proposição 2.1.2 *Para quaisquer $\vec{u}, \vec{v} \in \mathbb{V}^3$ e $\lambda, \mu \in \mathbb{R}$, valem:*

(i) $\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}$*;*
(ii) $(\lambda + \mu)\vec{u} = \lambda\vec{u} + \mu\vec{u}$*;*
(iii) $1\vec{u} = \vec{u}$*;*
(iv) $(\lambda\mu)\vec{u} = \lambda(\mu\vec{u})$.

Vejamos, em um exemplo, como utilizar a linguagem de vetores para fazer argumentos sobre problemas geométricos.

Exemplo 2.1.3 (Lista 1 - Álgebra Linear I, ex. 41) Mostre que o segmento que une os pontos médios de dois lados de um triângulo é paralelo ao terceiro lado e tem metade de sua medida.

Solução: Sejam A, B, C os vértices do triângulo. Seja M o ponto médio do lado AB e seja N o ponto médio do lado BC. Consideremos o vetor $\overrightarrow{MN}$.



Como M é o ponto médio de AB, segue que $\overrightarrow{MB} = \frac{1}{2}\overrightarrow{AB}$, e, como N é o ponto médio de BC, segue que $\overrightarrow{BN} = \frac{1}{2}\overrightarrow{BC}$. Assim,

$$\overrightarrow{MN} = \overrightarrow{MB} + \overrightarrow{BN} = \frac{1}{2}\overrightarrow{AB} + \frac{1}{2}\overrightarrow{BC} = \frac{1}{2}\left(\overrightarrow{AB} + \overrightarrow{BC}\right) = \frac{1}{2}\overrightarrow{AC}.$$

Segue, portanto, que o segmento MN é paralelo ao segmento AC (uma vez que, por definição, o vetor $\frac{1}{2}\overrightarrow{AC}$ é paralelo ao vetor $\overrightarrow{AC}$) e a medida do segmento MN é igual a $\left\|\overrightarrow{MN}\right\| = \left\|\frac{1}{2}\overrightarrow{AC}\right\| = \left|\frac{1}{2}\right|\left\|\overrightarrow{AC}\right\| = \frac{1}{2}\left\|\overrightarrow{AC}\right\|$, que é a metade da medida do segmento AC. $\diamond$

Observação Podemos utilizar o conceito de multiplicação de um escalar por um vetor para representar a noção de paralelismo entre vetores. Mais precisamente, dados $\vec{u}, \vec{v} \in \mathbb{V}^3$, $\vec{u} \neq \vec{0}$, temos que $\vec{u}$ e $\vec{v}$ são paralelos se, e somente se, existir $\lambda \in \mathbb{R}$ tal que $\vec{v} = \lambda \vec{u}$. (Com efeito, se $\vec{v} = \lambda \vec{u}$, então $\vec{v}$ é paralelo a $\vec{u}$, por definição. Para a recíproca, suponha que $\vec{u}$ e $\vec{v}$ sejam vetores paralelos. Então, $\vec{v} = \lambda \vec{u}$, com $\lambda = \frac{\|\vec{v}\|}{\|\vec{u}\|}$, se $\vec{u}$ e $\vec{v}$ tiverem o mesmo sentido, e $\lambda = -\frac{\|\vec{v}\|}{\|\vec{u}\|}$, se $\vec{u}$ e $\vec{v}$ tiverem sentidos opostos, como o leitor pode facilmente verificar.)

A construção do conjunto $\mathbb{V}^3$ e a definição das operações de soma e multiplicação por escalar pode ser replicada para dar origem ao conjunto $\mathbb{V}^2$, dos vetores *bidimensionais*, partindo de segmentos orientados cujas extermidades são elementos do conjunto $\mathbb{E}^2$, formado por todos os pontos do espaço bidimensional.

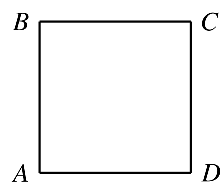
Exercícios Lista 1 - Álgebra Linear I: Exs. 36–38.

2.2 Dependência linear

Dizemos que um vetor $\vec{u}$ é *combinação linear* dos vetores $\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}$ se existirem escalares $\lambda_1, \lambda_2, \dots, \lambda_n$ tais que

$$\vec{u} = \lambda_1 \vec{v_1} + \lambda_2 \vec{v_2} + \cdots + \lambda_n \vec{v_n}.$$

Por exemplo, se A, B, C, D são os vértices de um quadrado, conforme a figura



então o vetor $\overrightarrow{AC}$ é combinação linear de $\overrightarrow{CB}$ e $\overrightarrow{DC}$, uma vez que

$$\overrightarrow{AC} = \overrightarrow{AD} + \overrightarrow{DC} = \overrightarrow{BC} + \overrightarrow{DC} = -\overrightarrow{CB} + \overrightarrow{DC} = (-1)\overrightarrow{CB} + 1\overrightarrow{DC}.$$

Por outro lado, $\overrightarrow{AC}$ não é combinação linear de $\overrightarrow{CB}$ e $\overrightarrow{AD}$, pois qualquer combinação linear de $\overrightarrow{CB}$ e $\overrightarrow{AD}$ será um vetor paralelo a $\overrightarrow{AD}$, mas $\overrightarrow{AC}$ não é paralelo a $\overrightarrow{AD}$.

A definição central nesta seção utiliza o conceito de combinação linear.

Definição Um conjunto de vetores $\{\vec{u_1}, \vec{u_2}, \dots, \vec{u_n}\}$ é dito *linearmente dependente* (doravante abrevidado por *LD*) se algum $\vec{u_i}$ for combinação dos demais vetores do conjunto. Caso contrário, dizemos que o conjunto é *linearmente independente* (ou, simplesmente, *LI*).

No exemplo do quadrado acima, $\{\overrightarrow{AC}, \overrightarrow{CB}, \overrightarrow{DC}\}$ é LD, e $\{\overrightarrow{AC}, \overrightarrow{BC}\}$ é LI.

Observação Seguem imediatamente da definição:

- Qualquer conjunto que contenha o vetor nulo é LD. (Pois o vetor nulo se escreve como combinação linear dos demais utilizando-se apenas o número 0 como coeficiente em cada termo da combinação linear.)
- Qualquer subconjunto de um conjunto LI é LI. (Isso segue do fato de que se $\{\vec{u_1}, \vec{u_2}, \dots, \vec{u_n}\}$ é LD, então $\{\vec{u_1}, \vec{u_2}, \dots, \vec{u_n}, \vec{v}\}$ é LD, qualquer que seja $\vec{v} \in \mathbb{V}^3$, bastando usar o escalar 0 como coeficiente de $\vec{v}$.)

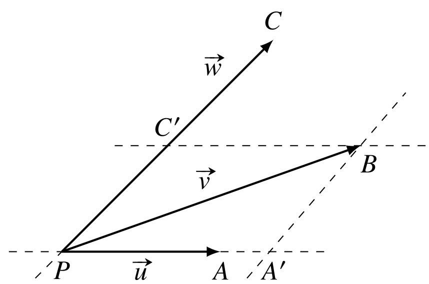
Interpretação geométrica da dependência linear

Vejamos, geometricamente, o significa dizer que um conjunto de vetores é LD. Vamos fazer um tratamento em casos por tamanho do conjunto.

Comecemos lembrando que dois vetores são paralelos se tiverem representantes paralelos. Diremos que três vetores são coplanares se tiverem representantes coplanares. Mais precisamente, dados $\vec{u}, \vec{v}, \vec{w} \in \mathbb{V}^3$, dizemos que $\vec{u}, \vec{v}, \vec{w}$ são *coplanares* se existirem $P, A, B, C \in \mathbb{E}^3$ em um mesmo plano tais que $\vec{u} = \overrightarrow{PA}$, $\vec{v} = \overrightarrow{PB}$, $\vec{w} = \overrightarrow{PC}$.

A interpretação geométrica da dependência linear segue abaixo.

1. *Por convenção, um conjunto com um único vetor $\{\vec{u}\}$ é LD se, e somente se, $\vec{u} = \vec{0}$.*
2. *$\{\vec{u}, \vec{v}\}$ é LD se, e somente se, $\vec{u}, \vec{v}$ são paralelos.*
   Com efeito, se $\{\vec{u}, \vec{v}\}$ é LD, então, um deles é um múltiplo escalar do outro e, portanto, eles são paralelos. Reciprocamente, se são paralelos e $\vec{u} \neq \vec{0}$, por exemplo, então, como vimos, existe $\lambda \in \mathbb{R}$ tal que $\vec{v} = \lambda\vec{u}$, o que garante que $\{\vec{u}, \vec{v}\}$ é LD. Se $\vec{u} = \vec{0}$, o conjunto $\{\vec{u}, \vec{v}\}$ é trivialmente LD: $\vec{u} = 0\vec{v}$.
3. *$\{\vec{u}, \vec{v}, \vec{w}\}$ é LD se, e somente se, $\vec{u}, \vec{v}, \vec{w}$ são coplanares.*
   Com efeito, suponha, por exemplo, que existem escalares $\lambda, \mu \in \mathbb{R}$ tais que $\vec{u} = \lambda\vec{v} + \mu\vec{w}$. Tome $P, A, B, C \in \mathbb{E}^3$ tais que $\vec{u} = \overrightarrow{PA}$, $\vec{v} = \overrightarrow{PB}$, $\vec{w} = \overrightarrow{PC}$. Se são P, B, C colineares, P, A, B, C estão em um mesmo plano e, portanto, $\vec{u}, \vec{v}, \vec{w}$ são coplanares. Resta considerar o caso em que P, B, C não são colineares. Neste caso, seja $M \in \mathbb{E}^3$ tal que $\lambda\vec{v} = \overrightarrow{PM}$ e seja $N \in \mathbb{E}^3$ tal que $\mu\vec{w} = \overrightarrow{PN}$. Então, M está na reta PB e N está na reta PC. Como $\overrightarrow{PA} = \vec{u} = \lambda\vec{v} + \mu\vec{w} = \overrightarrow{PM} + \overrightarrow{PN}$, segue que P, A, B, C estão no plano definido por P, B, C. Logo, $\vec{u}, \vec{v}, \vec{w}$ são coplanares.
   Reciprocamente, suponha que $\vec{u} = \overrightarrow{PA}$, $\vec{v} = \overrightarrow{PB}$, $\vec{w} = \overrightarrow{PC}$, em que os pontos P, A, B, C estão em um mesmo plano. Podemos supor que $\vec{u}, \vec{w}$ não são paralelos, pois se são, $\{\vec{u}, \vec{w}\}$ é LD e, *a fortiori*, $\{\vec{u}, \vec{v}, \vec{w}\}$ é LD. Considere o ponto A' dado pela interseção da reta paralela a PC que passa por B com a reta PA e o ponto C' dado pela interseção da reta paralela a PA que passa por B com a reta PC, como na figura

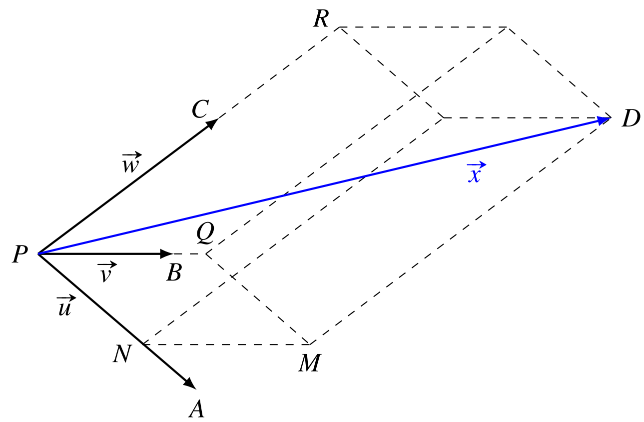
   Como $\overrightarrow{PA'}$ é paralelo a $\overrightarrow{PA}$, existe $\lambda \in \mathbb{R}$ tal que $\overrightarrow{PA'} = \lambda\overrightarrow{PA} = \lambda\vec{u}$. Como $\overrightarrow{PC'}$ é paralelo a $\overrightarrow{PC}$, existe $\mu \in \mathbb{R}$ tal que $\overrightarrow{PC'} = \mu\overrightarrow{PC} = \mu\vec{w}$. Assim, $\vec{v} = \overrightarrow{PB} = \overrightarrow{PA'} + \overrightarrow{A'B} = \overrightarrow{PA'} + \overrightarrow{PC'} = \lambda\vec{u} + \mu\vec{w}$, e $\{\vec{u}, \vec{v}, \vec{w}\}$ é LD.
4. *Qualquer conjunto com quatro ou mais vetores é LD.*
   Isso é consequência do resultado a seguir.

Teorema 2.2.1 *Seja $\{\vec{u}, \vec{v}, \vec{w}\}$ um conjunto LI contendo três vetores. Seja $\vec{x} \in \mathbb{V}^3$. Então $\vec{x}$ é combinação linear de $\vec{u}, \vec{v}, \vec{w}$.*

Demonstração Tome $P, A, B, C \in \mathbb{E}^3$ tais que $\vec{u} = \overrightarrow{PA}$, $\vec{v} = \overrightarrow{PB}$, $\vec{w} = \overrightarrow{PC}$. Como $\{\vec{u}, \vec{v}, \vec{w}\}$ é LI, esses quatro pontos não estão em um mesmo plano. Seja $D \in \mathbb{E}^3$ tal que $\vec{x} = \overrightarrow{PD}$. Considere os seguintes quatro pontos:

- M, dado pela interseção do plano PAB com a reta paralela a PC que passa por D;
- N, dado pela interseção da reta PA com a reta paralela a PB que passa por M;
- Q, dado pela interseção da reta PB com a reta paralela a PA que passa por M;
- R, dado pela interseção da reta PC com o plano paralelo ao plano PAB que passa por D;

conforme o paralelepípedo na figura.



Como $\overrightarrow{PA}$ e $\overrightarrow{PN}$ são paralelos, $\overrightarrow{PN} = \alpha\overrightarrow{PA} = \alpha\vec{u}$, para algum $\alpha \in \mathbb{R}$. Similarmente, $\overrightarrow{PQ} = \beta\overrightarrow{PB} = \beta\vec{v}$, para algum $\beta \in \mathbb{R}$, e $\overrightarrow{PR} = \gamma\overrightarrow{PC} = \gamma\vec{w}$, para algum $\gamma \in \mathbb{R}$.

Logo, $\vec{x} = \overrightarrow{PD} = \overrightarrow{PM} + \overrightarrow{MD} = \overrightarrow{PN} + \overrightarrow{NM} + \overrightarrow{MD} = \overrightarrow{PN} + \overrightarrow{PQ} + \overrightarrow{PR} = \alpha\vec{u} + \beta\vec{v} + \gamma\vec{w}$. □

Corolário 2.2.2 *Qualquer conjunto com quatro ou mais vetores é LD.*

Demonstração Seja $\{\vec{u_1}, \vec{u_2}, \vec{u_3}, \vec{u_4}, \dots\}$ um conjunto com pelo menos quatro vetores. Se $\{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ é LD, então o conjunto maior é LD. Se $\{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ é LI, então pelo Teorema 2.2.1, $\vec{u_4}$ é combinação linear de $\vec{u_1}, \vec{u_2}, \vec{u_3}$ e, portanto, também é combinação linear de todos os outros vetores do conjunto, utilizando o escalar 0, se necessário, como coeficiente dos demais. □

Resumindo, podemos caracterizar a dependência linear em termos geométricos por:

1. $\{\vec{u}\}$ é LD se, e somente se, $\vec{u} = \vec{0}$.
2. $\{\vec{u}, \vec{v}\}$ é LD se, e somente se, $\vec{u}, \vec{v}$ são paralelos.
3. $\{\vec{u}, \vec{v}, \vec{w}\}$ é LD se, e somente se, $\vec{u}, \vec{v}, \vec{w}$ são coplanares.
4. $\{\vec{u}, \vec{v}, \vec{w}, \vec{z}, \dots\}$ é sempre LD.

Dependência e independência linear podem ser caracterizadas por uma condição equivalente à da definição e que é frequentemente conveniente.

Proposição 2.2.3 *Sejam $\vec{u_1}, \vec{u_2}, \dots, \vec{u_n} \in \mathbb{V}^3$. Então, o conjunto $\{\vec{u_1}, \vec{u_2}, \dots, \vec{u_n}\}$ é LD se, e somente se, existem $\lambda_1, \lambda_2, \dots, \lambda_n \in R$, não todos nulos, tais que*

$$\lambda_1\vec{u_1} + \lambda_2\vec{u_2} + \cdots + \lambda_n\vec{u_n} = \vec{0}. \tag{2.1}$$

Demonstração Se $\{\vec{u_1}, \vec{u_2}, \dots, \vec{u_n}\}$ é LD, então, digamos (para efeito da argumentação), $\vec{u_1}$ é combinação linear de $\vec{u_2}, \dots, \vec{u_n}$, ou seja, existem $\beta_2, \dots, \beta_n \in \mathbb{R}$ tais que

$$\vec{u_1} = \beta_2\vec{u_2} + \cdots + \beta_n\vec{u_n}.$$

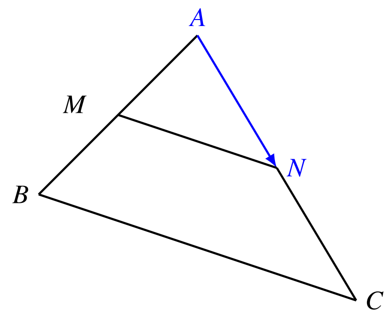
Isso implica que vale a igualdade (2.1), em que $\lambda_1 = 1$ e $\lambda_i = -\beta_i$, para todo $i = 2, \dots, n$. Em particular, o coeficiente λ_1 não é nulo.

Reciprocamente, suponha que existem $\lambda_1, \lambda_2, \dots, \lambda_n \in R$, não todos nulos, tais que (2.1) seja válida. Suponha, para efeitos da argumentação, que $\lambda_1 \neq 0$. Então, $\vec{u_1} = \beta_2\vec{u_2} + \cdots + \beta_n\vec{u_n}$, em que $\beta_i = -\frac{\lambda_i}{\lambda_1}$, para todo $i = 2, \dots, n$. Portanto, $\{\vec{u_1}, \vec{u_2}, \dots, \vec{u_n}\}$ é LD. □

Essa condição equivalente para dependência linear implica a seguinte condição equivalente para a independência linear: um conjunto de vetores $\{\vec{u_1}, \vec{u_2}, \dots, \vec{u_n}\}$ é LI se, e somente se, a única combinação linear de $\vec{u_1}, \vec{u_2}, \dots, \vec{u_n}$ que resulta no vetor nulo é aquela em que todos os escalares aparecendo como coeficientes são nulos.

Exemplo 2.2.4 No triângulo de vértices A, B, C, seja M o ponto médio do lado AB e seja N o ponto no lado AC tal que o segmento MN seja paralelo ao lado BC. Mostre que N é o ponto médio do lado AC.

Solução: Para mostrar que N é o ponto médio de AC, é suficiente mostrar que $\overrightarrow{AN} = \frac{1}{2}\overrightarrow{AC}$.



Como $\overrightarrow{AN}$ e $\overrightarrow{AC}$ são paralelos, existe $\lambda \in \mathbb{R}$ tal que $\overrightarrow{AN} = \lambda\overrightarrow{AC}$. Precisamos mostrar que $\lambda = \frac{1}{2}$. Como $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$, temos, por um lado,

$$\overrightarrow{AN} = \lambda\overrightarrow{AC} = \lambda(\overrightarrow{AB} + \overrightarrow{BC}) = \lambda\overrightarrow{AB} + \lambda\overrightarrow{BC}. \tag{2.2}$$

Por outro lado, sabemos que $\overrightarrow{AN} = \overrightarrow{AM} + \overrightarrow{MN}$. Como M é o ponto médio de AB, temos $\overrightarrow{AM} = \frac{1}{2}\overrightarrow{AB}$. Ainda, como $\overrightarrow{MN}$ e $\overrightarrow{BC}$ são paralelos (por hipótese), existe $\mu \in \mathbb{R}$ tal que $\overrightarrow{MN} = \mu\overrightarrow{BC}$. Logo,

$$\overrightarrow{AN} = \overrightarrow{AM} + \overrightarrow{MN} = \frac{1}{2}\overrightarrow{AB} + \mu\overrightarrow{BC}. \tag{2.3}$$

Comparando (2.2) com (2.3), obtemos

$$\lambda\overrightarrow{AB} + \lambda\overrightarrow{BC} = \frac{1}{2}\overrightarrow{AB} + \mu\overrightarrow{BC},$$

que, por sua vez, implica

$$\left(\lambda - \frac{1}{2}\right)\overrightarrow{AB} + (\lambda - \mu)\overrightarrow{BC} = \vec{0}.$$

Como $\{\overrightarrow{AB}, \overrightarrow{BC}\}$ é LI, pois $\overrightarrow{AB}$ e $\overrightarrow{BC}$ não são paralelos, segue que $\lambda - \frac{1}{2} = 0$ e que $\lambda - \mu = 0$. Em particular, $\lambda = \frac{1}{2}$, que é o que desejávamos provar. (Observe que obtemos, também, a informação de que $\mu = \frac{1}{2}$, ou seja, de que a medida do segmento MN é a metada da medida do lado BC.)

Exercícios Lista 1 - Álgebra Linear I: Exs. 39–45.

2.3 Bases e coordenadas

Vimos, na seção anterior, que dado um conjunto LI de três vetores, qualquer vetor se decompõe como combinação linear desses três vetores (Teorema 2.2.1). Nesta seção usaremos esse fato para definir bases e coordenadas em $\mathbb{V}^3$.

Definição Um conjunto LI ordenado formado por três vetores em $\mathbb{V}^3$ será chamado *base*.

Dada uma base $\mathcal{B} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ de $\mathbb{V}^3$ e dado um vetor qualquer $\vec{u}$, pelo Teorema 2.2.1, existem escalares $\lambda_1, \lambda_2, \lambda_3$ tais que

$$\vec{u} = \lambda_1\vec{e_1} + \lambda_2\vec{e_2} + \lambda_3\vec{e_3}. \tag{2.4}$$

Mostremos que os escalares na decomposição (2.4) estão univocamente determinados por $\vec{u}$. Com efeito, suponha que $\mu_1, \mu_2, \mu_3 \in \mathbb{R}$ sejam tais que $\vec{u} = \mu_1\vec{e_1} + \mu_2\vec{e_2} + \mu_3\vec{e_3}$. Comparando as duas decomposições de $\vec{u}$ como combinação linear de $\vec{e_1}, \vec{e_2}, \vec{e_3}$, obtemos

$$\lambda_1\vec{e_1} + \lambda_2\vec{e_2} + \lambda_3\vec{e_3} = \mu_1\vec{e_1} + \mu_2\vec{e_2} + \mu_3\vec{e_3},$$

donde segue que

$$(\lambda_1 - \mu_1)\vec{e_1} + (\lambda_2 - \mu_2)\vec{e_2} + (\lambda_3 - \mu_3)\vec{e_3} = \vec{0}.$$

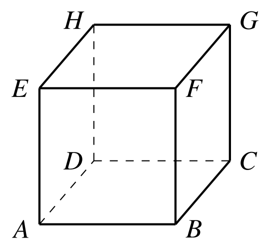
Como o conjunto $\{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ é LI (pois é uma base), segue que $\lambda_1 - \mu_1 = 0$, $\lambda_2 - \mu_2 = 0$ e $\lambda_3 - \mu_3 = 0$, ou, ainda, que

$$\lambda_1 = \mu_1, \quad \lambda_2 = \mu_2, \quad \lambda_3 = \mu_3.$$

Por causa da unicidade dos escalares utilizados na decomposição (2.4) de $\vec{u}$ como combinação linear de $\{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$, podemos introduzir a definição seguinte.

Definição Seja $\mathcal{B} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ uma base de $\mathbb{V}^3$ e seja $\vec{u} \in \mathbb{V}^3$. Os escalares $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tais que $\vec{u} = \lambda_1\vec{e_1} + \lambda_2\vec{e_2} + \lambda_3\vec{e_3}$ são chamados *coordenadas* de $\vec{u}$ em relação à base $\mathcal{B}$. Usamos a notação $\vec{u} = (\lambda_1, \lambda_2, \lambda_3)_{\mathcal{B}}$.

Exemplo 2.3.1 Considere, em $\mathbb{E}^3$, o cubo de vértices A, B, C, D, E, F, G, H, conforme a figura.



Os vetores $\overrightarrow{AC}, \overrightarrow{AH}, \overrightarrow{EF}$ são LI, uma vez que não são coplanares. Então, o conjunto ordenado $\mathcal{B} = \{\overrightarrow{AC}, \overrightarrow{AH}, \overrightarrow{EF}\}$ é uma base de $\mathbb{V}^3$. Encontre as coordenadas de $\overrightarrow{AF}$ em relação à base $\mathcal{B}$.

Solução: Precisamos encontrar a decomposição de $\overrightarrow{AF}$ como combinação linear de $\overrightarrow{AC}, \overrightarrow{AH}, \overrightarrow{EF}$. Temos

$$\begin{aligned}\overrightarrow{AF} &= \overrightarrow{AE} + \overrightarrow{EF} = \overrightarrow{AH} + \overrightarrow{HE} + \overrightarrow{EF} = \overrightarrow{AH} + \overrightarrow{CB} + \overrightarrow{EF}\\ &= \overrightarrow{AH} + \overrightarrow{CA} + \overrightarrow{AB} + \overrightarrow{EF} = \overrightarrow{AH} - \overrightarrow{AC} + \overrightarrow{EF} + \overrightarrow{EF}\\ &= (-1)\overrightarrow{AC} + 1\overrightarrow{AH} + 2\overrightarrow{EF}.\end{aligned}$$

Assim, $\overrightarrow{AF} = (-1, 1, 2)_{\mathcal{B}}$. $\diamond$

Observação O Exemplo 2.3.1 nos será útil para algumas observações:

- A ordem em que os vetores de uma base são exibidos é fundamental. Por exemplo, em relação à base $\mathcal{C} = \{\overrightarrow{AH}, \overrightarrow{EF}, \overrightarrow{AC}\}$, que é formada pelos mesmos vetores que formam $\mathcal{B}$, mas em outra ordem, as coordenadas de $\overrightarrow{AF}$ são outras: $\overrightarrow{AF} = (1, 2, -1)_{\mathcal{C}}$.
- A unicidade dos escalares na decomposição de um vetor como combinação de vetores não está garantida se os vetores utilizados na decomposição não forem LI. Por exemplo, o conjunto $\{\overrightarrow{AB}, \overrightarrow{EG}, \overrightarrow{HE}\}$ não é LI e o vetor $\overrightarrow{DB}$ pode ser decomposto como combinação linear de $\{\overrightarrow{AB}, \overrightarrow{EG}, \overrightarrow{HE}\}$, por exemplo, das seguintes duas maneiras:

$$\begin{aligned}\overrightarrow{DB} &= \overrightarrow{DA} + \overrightarrow{AB} = \overrightarrow{HE} + \overrightarrow{AB}\\ &= 1\overrightarrow{AB} + 0\overrightarrow{EG} + 1\overrightarrow{HE}\end{aligned}$$

e

$$\begin{aligned}\overrightarrow{DB} &= \overrightarrow{DA} + \overrightarrow{AB} = \overrightarrow{HE} + \overrightarrow{EF} = \overrightarrow{HE} + \overrightarrow{EG} + \overrightarrow{GF} = \overrightarrow{HE} + \overrightarrow{EG} + \overrightarrow{HE}\\ &= 0\overrightarrow{AB} + 1\overrightarrow{EG} + 2\overrightarrow{HE}.\end{aligned}$$

(Há uma infinidade de outras decomposições de $\overrightarrow{DB}$ como combinação linear de $\{\overrightarrow{AB}, \overrightarrow{EG}, \overrightarrow{HE}\}$, uma vez que $\overrightarrow{AB} = \overrightarrow{HE} + \overrightarrow{EG}$. Assim, por exemplo, para qualquer $\lambda \in \mathbb{R}$, temos $\overrightarrow{DB} = (1+\lambda)\overrightarrow{AB} + (-\lambda)\overrightarrow{EG} + (1-\lambda)\overrightarrow{HE}$.)

Se $\mathcal{B}$ é uma base de $\mathbb{V}^3$, como os vetores que a compõem formam um conjunto LI, segue que $\vec{0} = (0,0,0)_{\mathcal{B}}$. Além disso, $\vec{0}$ é o único vetor de $\mathbb{V}^3$ cujas coordenadas são todas nulas.

A principal vantagem de se trabalhar com coordenadas é que elas se comportam bem com respeito às operações de soma e de multiplicação por escalar definidas em $\mathbb{V}^3$, como mostra próximo resultado.

Proposição 2.3.2 *Seja $\mathcal{B}$ uma base de $\mathbb{V}^3$ e sejam $\vec{u}, \vec{v} \in \mathbb{V}^3$, com $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)_{\mathcal{B}}$ e $\vec{v} = (\beta_1, \beta_2, \beta_3)_{\mathcal{B}}$. Então,*

(i) $\vec{u} + \vec{v} = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \alpha_3 + \beta_3)_{\mathcal{B}}$;
(ii) $\lambda \vec{u} = (\lambda\alpha_1, \lambda\alpha_2, \lambda\alpha_3)_{\mathcal{B}}$, qualquer que seja $\lambda \in \mathbb{R}$.

Demonstração Se $\mathcal{B} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$, então as hipóteses dizem que $\vec{u} = \alpha_1\vec{e_1} + \alpha_2\vec{e_2} + \alpha_3\vec{e_3}$ e que $\vec{v} = \beta_1\vec{e_1} + \beta_2\vec{e_2} + \beta_3\vec{e_3}$. Assim,

$$\vec{u} + \vec{v} = (\alpha_1\vec{e_1} + \alpha_2\vec{e_2} + \alpha_3\vec{e_3}) + (\beta_1\vec{e_1} + \beta_2\vec{e_2} + \beta_3\vec{e_3}) = (\alpha_1 + \beta_1)\vec{e_1} + (\alpha_2 + \beta_2)\vec{e_2} + (\alpha_3 + \beta_3)\vec{e_3}$$

e

$$\lambda\vec{u} = \lambda(\alpha_1\vec{e_1} + \alpha_2\vec{e_2} + \alpha_3\vec{e_3}) = (\lambda\alpha_1)\vec{e_1} + (\lambda\alpha_2)\vec{e_2} + (\lambda\alpha_3)\vec{e_3},$$

provando o que se queria. □

Exemplo 2.3.3 Considerando coordenadas em relação a uma base fixada de $\mathbb{V}^3$, em cada um dos itens abaixo, verifique se os dois vetores apresentados formam um conjunto LD ou LI.

(i) $\vec{u} = (3, 10, 11)$ e $\vec{v} = (4, 7, -1)$
(ii) $\vec{u} = (1, -7, 2)$ e $\vec{v} = (-\frac{1}{2}, \frac{7}{2}, -1)$

Solução: (i) Como $\vec{u} \neq \vec{0}$, segue que $\{\vec{u}, \vec{v}\}$ é LD se, e somente se, existir $\lambda \in \mathbb{R}$ tal que $\vec{v} = \lambda\vec{u}$. Isso implicaria, pela Proposição 2.3.2, que $(4, 7, -1) = (3\lambda, 10\lambda, 11\lambda)$. Como estamos comparando coordenadas em relação a uma base, isso só seria possível se $4 = 3\lambda$, $7 = 10\lambda$ e $-1 = 11\lambda$. Claramente, não existe $\lambda \in \mathbb{R}$ que satisfaça essas três equações. Portanto, $\{\vec{u}, \vec{v}\}$ é LI.

(ii) Como $\vec{u} = (1, -7, 2) = -2(\frac{-1}{2}, \frac{7}{2}, -1) = -2\vec{v}$, segue que $\{\vec{u}, \vec{v}\}$ é LD. ◇

Para verificar dependência linear entre três vetores, coordenadas proporcionam um teste rápido.

Proposição 2.3.4 *Seja $\mathcal{B}$ uma base de $\mathbb{V}^3$ e sejam $\vec{u}, \vec{v}, \vec{w} \in \mathbb{V}^3$, com $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)_{\mathcal{B}}$, $\vec{v} = (\beta_1, \beta_2, \beta_3)_{\mathcal{B}}$ e $\vec{w} = (\gamma_1, \gamma_2, \gamma_3)_{\mathcal{B}}$. Então,*

$$\{\vec{u}, \vec{v}, \vec{w}\} \text{ é LD} \quad \text{se, e somente se,} \quad \det\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix} = 0.$$

Demonstração Um conjunto de vetores ser LD significa não ser LI. Sabemos que $\{\vec{u}, \vec{v}, \vec{w}\}$ é LI se, e somente se, os únicos escalares λ, μ, ρ tais que

$$\lambda\vec{u} + \mu\vec{v} + \rho\vec{w} = \vec{0} \tag{2.5}$$

forem $\lambda = \mu = \rho = 0$. Em coordenadas, usando a Proposição 2.3.2, a combinação linear (2.5) lê-se

$$\left(\lambda\alpha_1 + \mu\beta_1 + \rho\gamma_1,\ \lambda\alpha_2 + \mu\beta_2 + \rho\gamma_2,\ \lambda\alpha_3 + \mu\beta_3 + \rho\gamma_3\right) = (0,0,0),$$

isto é, (λ, μ, ρ) é solução do sistema linear

$$\begin{cases} \alpha_1 x + \beta_1 y + \gamma_1 z = 0 \\ \alpha_2 x + \beta_2 y + \gamma_2 z = 0 \\ \alpha_3 x + \beta_3 y + \gamma_3 z = 0 \end{cases}$$

Pelo Teorema 1.3.6, esse sistema tem apenas a solução trivial se, e somente se, $\det \begin{bmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{bmatrix} \neq 0$. Logo, $\{\vec{u}, \vec{v}, \vec{w}\}$ é LI se, e somente se, $\det \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix} \neq 0$, usando o fato de que o determinante de uma matriz coincide com o de sua transposta. □

Exemplo 2.3.5 Seja $\mathcal{B} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ uma base de $\mathbb{V}^3$ e considere os vetores

$$\vec{f_1} = 2\vec{e_1} - \vec{e_2}, \quad \vec{f_2} = \vec{e_1} - \vec{e_2} + 2\vec{e_3}, \quad \vec{f_3} = \vec{e_1} + 2\vec{e_3}.$$

(i) Mostre que $C = \{\vec{f_1}, \vec{f_2}, \vec{f_3}\}$ é uma base de $\mathbb{V}^3$.
(ii) Encontre as coordenadas de $\vec{v} = (1, 2, -2)_C$ em relação à base $\mathcal{B}$.

(Adiante, veremos um método eficiente para, dadas coordenadas em uma base, encontrá-las em outra.)

Solução: (i) Temos $\vec{f_1} = (2, -1, 0)_{\mathcal{B}}$, $\vec{f_2} = (1, -1, 2)_{\mathcal{B}}$ e $\vec{f_3} = (1, 0, 2)_{\mathcal{B}}$. Como $\det \begin{bmatrix} 2 & -1 & 0 \\ 1 & -1 & 2 \\ 1 & 0 & 2 \end{bmatrix} = -4 \neq 0$, segue, da Proposição 2.3.4, que $\{\vec{f_1}, \vec{f_2}, \vec{f_3}\}$ é LI, e, portanto, C é uma base de $\mathbb{V}^3$.
(ii) $\vec{v} = (1, 2, -2)_C = \vec{f_1} + 2\vec{f_2} - 2\vec{f_3} = (2, -1, 0)_{\mathcal{B}} + 2(1, -1, 2)_{\mathcal{B}} - 2(1, 0, 2)_{\mathcal{B}} = (2, -3, 0)_{\mathcal{B}}$. ◇

Como vimos no Exemplo 2.3.3, a verificação da existência de uma relação de dependência linear entre dois vetores, em termos de coordenadas, é bastante imediata. Apesar disso, há um teste geral que nos será útil no próximo capítulo.

Proposição 2.3.6 *Seja $\mathcal{B}$ uma base de $\mathbb{V}^3$ e sejam $\vec{u}, \vec{v} \in \mathbb{V}^3$, com $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)_{\mathcal{B}}$ e $\vec{v} = (\beta_1, \beta_2, \beta_3)_{\mathcal{B}}$. Então,*

$$\{\vec{u}, \vec{v}\} \text{ é LD} \quad \text{se, e somente se,} \quad \det \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix} = \det \begin{bmatrix} \alpha_1 & \alpha_3 \\ \beta_1 & \beta_3 \end{bmatrix} = \det \begin{bmatrix} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{bmatrix} = 0.$$

Demonstração Se $\vec{u} = \vec{0}$, então $\{\vec{u}, \vec{v}\}$ é LD e os três determinantes são nulos. Podemos supor, então, que $\vec{u} \neq \vec{0}$. Se $\{\vec{u}, \vec{v}\}$ é LD, existem $\lambda, \mu \in \mathbb{R}$, não ambos nulos, tais que $\lambda\vec{u} + \mu\vec{v} = \vec{0}$. Se $\mu = 0$, teríamos $\lambda\vec{u} = \vec{0}$, o que implicaria $\lambda = 0$, uma vez que $\vec{u} \neq \vec{0}$. Mas isso contradiz o fato de λ e μ não serem ambos nulos. Logo, $\mu \neq 0$ e, portanto, $\vec{v} = -\frac{\lambda}{\mu}\vec{u}$. Daí segue que $\vec{v} = \left(-\frac{\lambda}{\mu}\alpha_1, -\frac{\lambda}{\mu}\alpha_2, -\frac{\lambda}{\mu}\alpha_3\right)_{\mathcal{B}}$, o que implica que os três determinantes são nulos (pois são determinantes de matrizes que têm linhas proporcionais).

Reciprocamente, suponha que os três determinantes sejam nulos. Como $\vec{u} \neq \vec{0}$, alguma de suas coordenadas deve ser necessariamente diferente de zero. Suponha que $\alpha_1 \neq 0$. Assim, do primeiro determinante, obtemos $\beta_2 = \frac{\beta_1}{\alpha_1}\alpha_2$, e, do segundo, $\beta_3 = \frac{\beta_1}{\alpha_1}\alpha_3$. Portanto, $\vec{v} = \left(\frac{\beta_1}{\alpha_1}\alpha_1, \frac{\beta_1}{\alpha_1}\alpha_2, \frac{\beta_1}{\alpha_1}\alpha_3\right)_{\mathcal{B}}$, isto é, $\vec{v} = \frac{\beta_1}{\alpha_1}\vec{u}$, donde segue que $\{\vec{u}, \vec{v}\}$ é LD. Se $\alpha_2 \neq 0$ ou $\alpha_3 \neq 0$, o argumento é análogo. □

Bases ortonormais

Dizemos que dois vetores não nulos $\vec{u}$ e $\vec{v}$ são *ortogonais* se tiverem representantes ortogonais, isto é, se existirem $A, B, C \in \mathbb{E}^3$ tais que $\vec{u} = \overrightarrow{AB}$ e $\vec{v} = \overrightarrow{AC}$ com os segmentos de reta AB e AC ortogonais. Por convenção, o vetor nulo é ortogonal a qualquer vetor. Denotaremos a ortogonalidade entre $\vec{u}$ e $\vec{v}$ por $\vec{u} \perp \vec{v}$.

Definição Uma base $\{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ de $\mathbb{V}^3$ é *ortonormal* se os vetores $\vec{e_1}, \vec{e_2}, \vec{e_3}$ forem, dois a dois, ortogonais e $\|\vec{e_1}\| = \|\vec{e_2}\| = \|\vec{e_3}\| = 1$.

Trabalhar com bases ortonormais é conveniente, pois existe, por exemplo, uma maneira simples de expressar norma em termos de coordenadas, como mostra a próxima proposição. Veremos, adiante, que bases ortonormais têm outras propriedades boas.

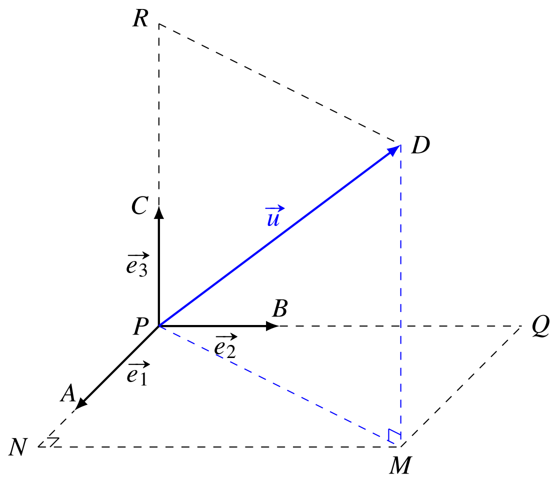
Proposição 2.3.7 *Seja $\mathcal{B}$ uma base ortonormal de $\mathbb{V}^3$ e seja $\vec{u} = (\alpha, \beta, \gamma)_{\mathcal{B}} \in \mathbb{V}^3$. Então,*

$$\|\vec{u}\| = \sqrt{\alpha^2 + \beta^2 + \gamma^2}.$$

Demonstração Suponha que $\mathcal{B} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$. Como vimos na demonstração do Teorema 2.2.1, podemos obter as coordenadas de $\vec{u}$ em relação à base $\mathcal{B}$ por meio de uma construção geométrica. Sejam $P, A, B, C, D \in \mathbb{E}^3$ tais que

$$\vec{e_1} = \overrightarrow{PA}, \quad \vec{e_2} = \overrightarrow{PB}, \quad \vec{e_3} = \overrightarrow{PC}, \quad \vec{u} = \overrightarrow{PD}.$$

Sejam N o ponto da reta PA tal que $\overrightarrow{PN} = \alpha\vec{e_1}$; seja Q o ponto da reta PB tal que $\overrightarrow{PQ} = \beta\vec{e_2}$; seja R o ponto da reta PC tal que $\overrightarrow{PR} = \gamma\vec{e_3}$; e seja M o ponto do plano PAB tal que $\overrightarrow{PM} = \overrightarrow{PN} + \overrightarrow{PQ}$, conforme a figura:



Então, $\vec{u} = \alpha\vec{e_1} + \beta\vec{e_2} + \gamma\vec{e_3} = \overrightarrow{PN} + \overrightarrow{PQ} + \overrightarrow{PR} = \overrightarrow{PM} + \overrightarrow{PR}$. Como o triângulo PMD é reto, com ângulo reto em M (pois $\vec{e_3}$ é ortogonal a $\vec{e_1}$ e a $\vec{e_2}$), segue, do Teorema de Pitágoras, que

$$\|\vec{u}\|^2 = PD^2 = PM^2 + MD^2. \tag{2.6}$$

Agora, como $\overrightarrow{MD} = \overrightarrow{PR}$, segue

$$MD^2 = \left\|\overrightarrow{PR}\right\|^2 = \|\gamma\vec{e_3}\|^2 = |\gamma|^2 \|\vec{e_1}\|^2. \tag{2.7}$$

Finalmente, a ortogonalidade entre $\vec{e_1}$ e $\vec{e_2}$ garante que o triângulo PNM é reto com ângulo reto em N. Portanto,

$$PM^2 = PN^2 + NM^2 = \left\|\overrightarrow{PN}\right\|^2 + \left\|\overrightarrow{PQ}\right\|^2 = \|\alpha\vec{e_1}\|^2 + \|\beta\vec{e_2}\|^2 = |\alpha|^2 \|\vec{e_1}\|^2 + |\beta|^2 \|\vec{e_2}\|^2. \tag{2.8}$$

Substituindo (2.7) e (2.8) em (2.6) e usando que os três vetores da base têm norma 1, obtemos

$$\|\vec{u}\|^2 = \alpha^2 + \beta^2 + \gamma^2,$$

como desejávamos. □

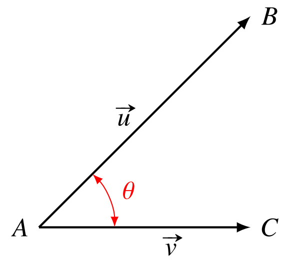
Observação Existem infinitas bases ortonormais em $\mathbb{V}^3$. Para escolher uma, por exemplo, basta escolher dois pontos $A, B \in \mathbb{E}^3$ tais que o segmento AB tenha comprimento 1. Agora, no plano perpendicular ao segmento AB que contém A, tome um ponto C de modo que o segmento AC tenha comprimento 1. Finalmente, na reta perpendicular ao plano ABC, tome um ponto D tal que AD tenha comprimento 1. Então, $\{\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}\}$ é uma base ortonormal de $\mathbb{V}^3$.

Exercícios Lista 1 - Álgebra Linear I: Exs. 46–53.

2.4 Produto escalar

Nesta seção, introduziremos uma nova operação no espaço do vetores, o produto escalar. Para tanto, necessitaremos do conceito de ângulo entre vetores.

Sejam $\vec{u}, \vec{v} \in \mathbb{V}^3$ não nulos. Tome $A, B, C \in \mathbb{E}^3$ tais que $\vec{u} = \overrightarrow{AB}$ e $\vec{v} = \overrightarrow{AC}$. Definimos o *ângulo* entre $\vec{u}$ e $\vec{v}$ como sendo o ângulo entre os segmentos de reta AB e AC.



A medida do ângulo entre dois vetores será sempre feita em radianos e é um valor θ satisfazendo $0 \leq \theta \leq \pi$.

Definição Sejam $\vec{u}, \vec{v} \in \mathbb{V}^3$. Definimos o *produto escalar* $\vec{u} \cdot \vec{v}$ como sendo o número real dado por

$$\vec{u} \cdot \vec{v} = \begin{cases} 0 & \text{se } \vec{u} = \vec{0} \text{ ou } \vec{v} = \vec{0} \\ \|\vec{u}\|\,\|\vec{v}\| \cos(\theta) & \text{se } \vec{u} \neq \vec{0} \text{ e } \vec{v} \neq \vec{0}, \end{cases}$$

em que θ denota a medida do ângulo entre $\vec{u}$ e $\vec{v}$.

O produto escalar provê um critério rápido para detecção de ortogonalidade: dados dois vetores $\vec{u}, \vec{v}$, temos que

$$\vec{u} \perp \vec{v} \quad \text{se, e somente se,} \quad \vec{u} \cdot \vec{v} = 0.$$

De fato, $\vec{u}$ e $\vec{v}$ são não nulos e ortogonais se, e somente se, o ângulo entre eles medir $\frac{\pi}{2}$ radianos, o que, por sua vez, é equivalente a $\cos(\theta) = 0$. Essa última condição ocorre precisamente quando $\vec{u} \cdot \vec{v} = 0$. Quando um dos dois vetores é nulo, a equivalência é imediata.

Lembrando que um ângulo de medida θ é dito *agudo* se $0 \leq \theta < \frac{\pi}{2}$, e *obtuso* se $\frac{\pi}{2} < \theta \leq \pi$, dados $\vec{u}, \vec{v} \in \mathbb{V}^3$ não nulos, o ângulo entre eles é agudo se, e somente se, $\vec{u} \cdot \vec{v} < 0$, e obstuso se, e somente se, $\vec{u} \cdot \vec{v} > 0$.

Observação Tanto a norma de um vetor quanto a medida do ângulo entre dois vetores não nulos podem ser obtidos por meio do produto escalar:

- Para qualquer $\vec{u} \in \mathbb{V}^3$, vale $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}$.
- Se $\vec{u}, \vec{v} \in \mathbb{V}^3$ são não nulos, então a medida do ângulo entre eles é dada por $\arccos\left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|}\right)$.

Em função de coordenadas em relação a uma base ortonormal, o produto escalar adquire uma forma bastante simples.

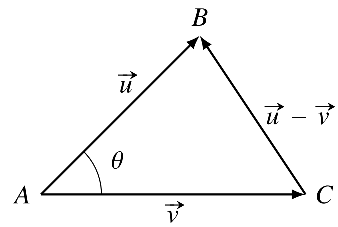
Proposição 2.4.1 *Seja $\mathcal{B}$ uma base ortonormal de $\mathbb{V}^3$ e sejam $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)_{\mathcal{B}}$, $\vec{v} = (\beta_1, \beta_2, \beta_3)_{\mathcal{B}}$. Então,*

$$\vec{u} \cdot \vec{v} = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3.$$

Demonstração Se $\vec{u} = \vec{0}$ ou $\vec{v} = \vec{0}$, então a fórmula é trivialmente válida. Suponha, portanto, que $\vec{u} \neq \vec{0}$ e $\vec{v} \neq \vec{0}$, e seja θ a medida do ângulo entre eles. Pela Proposição 2.3.2, $\vec{u} - \vec{v} = (\alpha_1 - \beta_1, \alpha_2 - \beta_2, \alpha_3 - \beta_3)_{\mathcal{B}}$, e, pela Proposição 2.3.7, temos

$$\begin{aligned} \|\vec{u}\|^2 &= \alpha_1^2 + \alpha_2^2 + \alpha_3^2 \\ \|\vec{v}\|^2 &= \beta_1^2 + \beta_2^2 + \beta_3^2 \\ \|\vec{u} - \vec{v}\|^2 &= (\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2 + (\alpha_3 - \beta_3)^2. \end{aligned}$$

Se $A, B, C \in \mathbb{E}^3$ são tais que $\vec{u} = \overrightarrow{AB}$ e $\vec{v} = \overrightarrow{AC}$, então $\overrightarrow{CB} = \vec{u} - \vec{v}$. E, pela lei dos cossenos aplicada ao triângulo ABC,



obtemos

$$\|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\|\|\vec{u}\|\cos(\theta) = \|\vec{u} - \vec{v}\|^2 .$$

Substituindo, no lado esquerdo, $\|\vec{u}\|\|\vec{v}\|\cos(\theta)$ por $\vec{u} \cdot \vec{v}$ e utilizando os valores dos quadrados das normas em termos das coordenadas obtidos acima, segue a igualdade desejada.[2] □

Em relação às operações de soma e multiplicação por escalar, o produto escalar tem as seguintes propriedades.

Proposição 2.4.2 *Sejam $\vec{u}, \vec{u}, \vec{w} \in \mathbb{V}^3$ e seja $\lambda \in \mathbb{R}$. Então,*

(i) $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$;
(ii) $\vec{u} \cdot (\lambda\vec{v}) = (\lambda\vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v})$;
(iii) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$;
(iv) $\vec{u} \cdot \vec{u} \geq 0$, e $\vec{u} \cdot \vec{u} = 0$ se, e somente se, $\vec{u} = \vec{0}$.

Demonstração (iii) é consequência imediata da definição. (iv) segue de $\vec{u} \cdot \vec{u} = \|\vec{u}\|^2$. Para (i) e (ii), fixe uma base ortonormal em $\mathbb{V}^3$ e tome coordenadas em relação a essa base. Por exemplo, se $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)$, $\vec{v} = (\beta_1, \beta_2, \beta_3)$ e $\vec{w} = (\gamma_1, \gamma_2, \gamma_3)$, em relação a uma base ortonormal de $\mathbb{V}^3$, então

$$\vec{u} \cdot (\vec{v} + \vec{w}) = \alpha_1(\beta_1 + \gamma_1) + \alpha_2(\beta_2 + \gamma_2) + \alpha_3(\beta_3 + \gamma_3),$$

ao passo que

$$\vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w} = (\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3) + (\alpha_1\gamma_1 + \alpha_2\gamma_2 + \alpha_3\gamma_3).$$

A igualdade segue da propriedade distributiva do produto em relação à soma de números reais. □

Exemplo 2.4.3 (Prova 1, Álgebra Linear I, 2015) Sejam $a, b, c \in \mathbb{R}$ e seja $\mathcal{E}$ uma base ortonormal de $\mathbb{V}^3$. Considere os vetores $\vec{z} = (1, 0, 1)_{\mathcal{E}}$, $\vec{v} = (-2, 1, 0)_{\mathcal{E}}$, $\vec{w} = (0, -1, 1)_{\mathcal{E}}$ e $\vec{x} = (a, b, c)_{\mathcal{E}}$. Se $\|\vec{x}\| = 3$, $\vec{x}$ é ortogonal a $\vec{z}$ e $\{\vec{v}, \vec{w}, \vec{x}\}$ é linearmente dependente, então $|a + b + c|$ é igual a

(A) 2 (B) 3 (C) 1 (D) 7 (E) 5

Solução: Como $\vec{z} \perp \vec{x}$, segue $\vec{z} \cdot \vec{x} = 0$. Mas, $\vec{z} \cdot \vec{x} = (1, 0, 1)_{\mathcal{E}} \cdot (a, b, c)_{\mathcal{E}} = 1a + 0b + 1c = a + c$, uma vez que $\mathcal{E}$ é ortonormal. Logo, $a + c = 0$, donde segue que $c = -a$, e, portanto, $\vec{x} = (a, b, -a)_{\mathcal{E}}$. Como $\{\vec{v}, \vec{w}, \vec{x}\}$ é LD, segue da Proposição 2.3.4, que $\det \begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & 1 \\ a & b & -a \end{bmatrix} = 0$. Calculando o determinante, obtemos a relação $-a + 2b = 0$. Logo, $a = 2b$. Assim, $\vec{x} = (2b, b, -2b)_{\mathcal{E}}$. Portanto, usando a Proposição 2.3.7, obtemos $\|\vec{x}\|^2 = (2b)^2 + b^2 + (-2b)^2 = 9b^2$. Como, por hipótese, $\|\vec{x}\| = 3$, segue que $b = 1$ ou $b = -1$. Assim, $\vec{x} = (2, 1, -2)_{\mathcal{E}}$ ou $\vec{x} = (-2, -1, 2)_{\mathcal{E}}$. Em ambos os casos, $|a + b + c| = 1$. *Resposta*: (C). ◇

[2] A rigor, a demonstração que fizemos só se aplica no caso em que $\{\vec{u}, \vec{v}\}$ é LI, pois somente neste caso teremos, de fato, um triângulo. Porém, o mesmo argumento se aplica quando $\vec{u}$ e $\vec{v}$ são paralelos, pois, ainda nesse caso, a lei dos cossenos continua válida.

Exemplo 2.4.4 Sejam $\vec{e_1}, \vec{e_2}, \vec{e_3} \in \mathbb{V}^3$. Mostre que $\{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ é uma base ortonormal de $\mathbb{V}^3$ se, e somente se, para todos $i, j = 1, 2, 3$, tivermos $\vec{e_i} \cdot \vec{e_j} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j. \end{cases}$

Solução: A condição é equivalente ao fato de $\vec{e_1}, \vec{e_2}, \vec{e_3}$ serem vetores de norma 1 que são, dois a dois, ortogonais. Assim, a única coisa que resta a ser demonstrada é que se $\vec{e_1}, \vec{e_2}, \vec{e_3}$ satisfazem $\vec{e_i} \cdot \vec{e_j} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j, \end{cases}$ então $\{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ é LI. Sejam, então $\alpha, \beta, \gamma \in \mathbb{R}$ tais que $\alpha\vec{e_1} + \beta\vec{e_2} + \gamma\vec{e_3} = \vec{0}$. Considerando o produto escalar com o vetor $\vec{e_1}$, obtemos

$$\vec{e_1} \cdot (\alpha\vec{e_1} + \beta\vec{e_2} + \gamma\vec{e_3}) = \vec{e_1} \cdot \vec{0} = 0.$$

Por outro lado,

$$\vec{e_1} \cdot (\alpha\vec{e_1} + \beta\vec{e_2} + \gamma\vec{e_3}) = \alpha\vec{e_1} \cdot \vec{e_1} + \beta\vec{e_1} \cdot \vec{e_2} + \gamma\vec{e_1} \cdot \vec{e_3} = \alpha + 0 + 0 = \alpha.$$

Logo, $\alpha = 0$. Tomando o produto escalar com $\vec{e_2}$ e $\vec{e_3}$, obtemos, analogamente, $\beta = 0$ e $\gamma = 0$. Portanto, $\{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ é, de fato, LI. $\diamond$

Exercícios Lista 1 - Álgebra Linear I: Exs. 54–68.

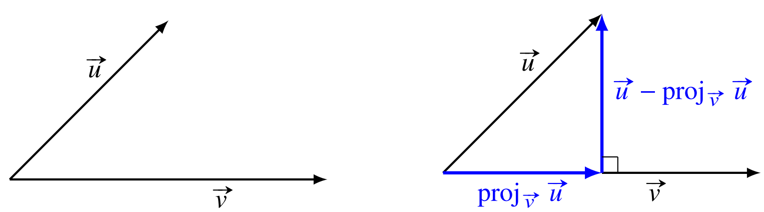
2.5 Projeção ortogonal

Nesta seção, veremos como utilizar o produto escalar para encontrar a projeção ortogonal de um vetor sobre outro.

Definição Seja $\vec{v} \in \mathbb{V}^3$, $\vec{v} \neq \vec{0}$. Dado $\vec{u} \in \mathbb{V}^3$, chama-se *projeção ortogonal* de $\vec{u}$ sobre $\vec{v}$ o vetor $\vec{w}$ que satisfaz as seguintes duas condições:

(i) $\vec{w}$ é paralelo a $\vec{v}$, e
(ii) $\vec{u} - \vec{w}$ é ortogonal a $\vec{v}$.

Veremos, em seguida, que a projeção ortogonal sempre existe e é única. Assim, utilizaremos a notação $\mathrm{proj}_{\vec{v}}\,\vec{u}$ para denotá-la.



Proposição 2.5.1 *Seja $\vec{v} \in \mathbb{V}^3$, $\vec{v} \neq \vec{0}$ e seja $\vec{u} \in \mathbb{V}^3$. Então, a projeção ortogonal de $\vec{u}$ sobre $\vec{v}$ é dada por*

$$\mathrm{proj}_{\vec{v}}\,\vec{u} = \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v}.$$

(Mais precisamene, o vetor $\frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v}$ satisfaz as condições de projeção ortogonal de $\vec{u}$ sobre $\vec{v}$ e é o único vetor a satisfazê-las.)

Demonstração Considere o vetor $\vec{w} = \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v}$. Mostremos que esse vetor satisfaz as condições na definição de projeção ortogonal de $\vec{u}$ sobre $\vec{v}$. A condição (i) está claramente satisfeita, uma vez que $\vec{w}$ é um múltiplo escalar de $\vec{v}$. Para a condição (ii), calculemos o produto escalar $(\vec{u} - \vec{w}) \cdot \vec{v}$:

$$(\vec{u} - \vec{w}) \cdot \vec{v} = \left(\vec{u} - \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v}\right) \cdot \vec{v} = \vec{u} \cdot \vec{v} - \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}\vec{v} \cdot \vec{v} = 0,$$

uma vez que $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$. Portanto, a condição (ii) também está satisfeita.

Finalmente, verifiquemos que $\vec{w}$ é o único vetor que tem as propriedades (i) e (ii). Seja $\vec{z} \in \mathbb{V}^3$ um vetor que é paralelo a $\vec{v}$ e que satisfaz $(\vec{u} - \vec{z}) \perp \vec{v}$. Como $\vec{v} \neq \vec{0}$ e $\vec{z}$ é paralelo a $\vec{v}$, existe $\lambda \in \mathbb{R}$ tal que $\vec{z} = \lambda\vec{v}$. Agora, de $(\vec{u} - \vec{z}) \perp \vec{v}$ segue que $(\vec{u} - \vec{z}) \cdot \vec{v} = 0$. Portanto, $(\vec{u} - \lambda\vec{v}) \cdot \vec{v} = 0$. O lado esquerdo dessa última igualdade coincide com $\vec{u} \cdot \vec{v} - \lambda\|\vec{v}\|^2$. Assim, $\lambda = \frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|^2}$. Logo, $\vec{z} = \frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v} = \vec{w}$. □

Exemplo 2.5.2 Seja $\mathcal{B}$ uma base ortonormal de $\mathbb{V}^3$. Considere os vetores $\vec{u} = (3, -6, 0)_{\mathcal{B}}$ e $\vec{v} = (2, -2, 1)_{\mathcal{B}}$.

(i) Encontre a projeção ortogonal de $\vec{u}$ sobre $\vec{v}$.
(ii) Encontre um vetor $\vec{p}$ paralelo a $\vec{v}$ e um vetor $\vec{q}$ ortogonal a $\vec{v}$ tais que $\vec{u} = \vec{p} + \vec{q}$.

Solução: (i) Sabemos, pela Proposição 2.5.1, que $\operatorname{proj}_{\vec{v}} \vec{u} = \frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v}$. Como a base $\mathcal{B}$ é ortonormal, podemos utilizar a Proposição 2.4.1 para calcular numerador e denominador do escalar que aparece na fórmula:

$$\vec{u} \cdot \vec{v} = (3, -6, 0)_{\mathcal{B}} \cdot (2, -2, 1)_{\mathcal{B}} = 6 + 12 + 0 = 18$$

e

$$\|\vec{v}\|^2 = \vec{v} \cdot \vec{v} = (2, -2, 1)_{\mathcal{B}} \cdot (2, -2, 1)_{\mathcal{B}} = 4 + 4 + 1 = 9.$$

Logo, $\operatorname{proj}_{\vec{v}} \vec{u} = \frac{18}{9}(2, -2, 1)_{\mathcal{B}} = (4, -4, 2)_{\mathcal{B}}$.

(ii) Sabemos que $\vec{u} - \operatorname{proj}_{\vec{v}} \vec{u}$ é ortogonal a $\vec{v}$. Como também vale que $\operatorname{proj}_{\vec{v}} \vec{u}$ é paralelo a $\vec{v}$, basta tomar $\vec{p} = \operatorname{proj}_{\vec{v}} \vec{u} = (4, -2, 2)_{\mathcal{B}}$ e $\vec{q} = \vec{u} - \operatorname{proj}_{\vec{v}} \vec{u} = (3, -6, 0)_{\mathcal{B}} - (4, -2, 2)_{\mathcal{B}} = (-1, -2, -2)_{\mathcal{B}}$. ◇

No exemplo anterior, como $\vec{u} \neq \vec{0}$, também é possível considerar a projeção ortogonal de $\vec{v}$ sobre $\vec{u}$. O leitor pode verificar que $\operatorname{proj}_{\vec{u}} \vec{v} = (\frac{6}{5}, -\frac{12}{5}, 0)_{\mathcal{B}}$.

Exercícios Lista 1 - Álgebra Linear I: Exs. 69–78.

2.6 Mudança de base

Vimos que, fixada uma base de $\mathbb{V}^3$, podemos associar a cada vetor de $\mathbb{V}^3$ suas coordenadas em relação a essa base. Se tomarmos uma segunda base, um mesmo vetor terá, agora, duas sequências de coordenadas: uma em relação à primeira e outra, à segunda. Qual é a relação entre elas? Essa será a questão a ser abordada nesta seção.

Sejam $\mathcal{B} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ e $\mathcal{C} = \{\vec{f_1}, \vec{f_2}, \vec{f_3}\}$ bases de $\mathbb{V}^3$ e seja $\vec{u} \in \mathbb{V}^3$. Suponha que

$$\vec{u} = \big(x_1, x_2, x_3\big)_{\mathcal{B}} \quad \text{e} \quad \vec{u} = \big(y_1, y_2, y_3\big)_{\mathcal{C}}.$$

Como $\mathcal{B}$ é uma base de $\mathbb{V}^3$, os vetores que compõem a base $\mathcal{C}$ também se escrevem como combinação linear dos elementos de $\mathcal{B}$, ou seja, existem escalares $\alpha_{ij} \in \mathbb{R}$, com $i, j = 1, 2, 3$ tais que

$$\begin{aligned}
\vec{f_1} &= \big(\alpha_{11}, \alpha_{21}, \alpha_{31}\big)_{\mathcal{B}}\\
\vec{f_2} &= \big(\alpha_{12}, \alpha_{22}, \alpha_{32}\big)_{\mathcal{B}}\\
\vec{f_3} &= \big(\alpha_{13}, \alpha_{23}, \alpha_{33}\big)_{\mathcal{B}}.
\end{aligned}$$

Vamos procurar uma relação entre as coordenadas de $\vec{u}$ em relação à base $\mathcal{B}$ e à base $\mathcal{C}$.

Usando as coordenadas de $\vec{u}$ em relação à base $\mathcal{C}$ e as coordenadas de $\vec{f_1}, \vec{f_2}, \vec{f_3}$ em relação à base $\mathcal{B}$, podemos escrever

$$\begin{aligned}
\vec{u} &= y_1\vec{f_1} + y_2\vec{f_2} + \vec{f_3}\\
&= y_1(\alpha_{11}\vec{e_1} + \alpha_{21}\vec{e_2} + \alpha_{31}\vec{e_3}) + y_2(\alpha_{12}\vec{e_1} + \alpha_{22}\vec{e_2} + \alpha_{32}\vec{e_3}) + y_3(\alpha_{13}\vec{e_1} + \alpha_{23}\vec{e_2} + \alpha_{33}\vec{e_3})\\
&= (\alpha_{11}y_1 + \alpha_{12}y_2 + \alpha_{13}y_3)\vec{e_1} + (\alpha_{21}y_1 + \alpha_{22}y_2 + \alpha_{23}y_3)\vec{e_2} + (\alpha_{31}y_1 + \alpha_{32}y_2 + \alpha_{33}y_3)\vec{e_3}.
\end{aligned}$$

Os escalares obtidos, acima, na decomposição de $\vec{u}$ como combinação linear dos elementos da base $\mathcal{B}$, fornecem-nos as coordenadas de $\vec{u}$ em relação à base $\mathcal{B}$, e, portanto, segue que

$$\begin{aligned} x_1 &= \alpha_{11}y_1 + \alpha_{12}y_2 + \alpha_{13}y_3 \\ x_2 &= \alpha_{21}y_1 + \alpha_{22}y_2 + \alpha_{23}y_3 \\ x_3 &= \alpha_{31}y_1 + \alpha_{32}y_2 + \alpha_{33}y_3, \end{aligned}$$

ou ainda, de modo mais compacto, usando o produto entre matrizes,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

Neste ponto, introduziremos as seguintes notações:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [\vec{u}]_{\mathcal{B}}, \quad \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = [\vec{u}]_{\mathcal{C}} \quad \text{e} \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} = M_{\mathcal{C}\mathcal{B}}.$$

A matriz 3×1 $[\vec{u}]_{\mathcal{B}}$ (respectivamente, $[\vec{u}]_{\mathcal{C}}$) é chamada *vetor de coordenadas* de $\vec{u}$ em relação à base $\mathcal{B}$ (respectivamente, $\mathcal{C}$).

Assim, demonstramos o seguinte resultado fundamental.

Teorema 2.6.1 *Sejam $\mathcal{B}$ e $\mathcal{C}$ bases de $\mathbb{V}^3$ e seja $\vec{u} \in \mathbb{V}^3$. Então,*

$$[\vec{u}]_{\mathcal{B}} = M_{\mathcal{C}\mathcal{B}}[\vec{u}]_{\mathcal{C}}. \tag{2.9}$$

A matriz $M_{\mathcal{C}\mathcal{B}}$ é o que se chama de uma *matriz de mudança de base*.

Observe que na matriz de mudança de base $M_{\mathcal{C}\mathcal{B}}$, a primeira coluna é formada pelas coordenadas do primeiro vetor da base $\mathcal{C}$ em relação à base $\mathcal{B}$. Na segunda e terceira colunas estão as coordenadas do segundo e terceiro vetores, respectivamente, da base $\mathcal{C}$ em relação à base $\mathcal{B}$:

$$\text{se} \quad \begin{aligned} \vec{f_1} &= (\alpha_{11}, \alpha_{21}, \alpha_{31})_{\mathcal{B}} \\ \vec{f_2} &= (\alpha_{12}, \alpha_{22}, \alpha_{32})_{\mathcal{B}}, \\ \vec{f_3} &= (\alpha_{13}, \alpha_{23}, \alpha_{33})_{\mathcal{B}} \end{aligned} \qquad \text{então} \quad M_{\mathcal{C}\mathcal{B}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix},$$

sendo $\mathcal{C} = \{\vec{f_1}, \vec{f_2}, \vec{f_3}\}$.

Observação Se $\mathcal{B}$ e $\mathcal{C}$ são bases de $\mathbb{V}^3$, então $\det(M_{\mathcal{C}\mathcal{B}}) \neq 0$, uma vez que $\mathcal{C}$ é um conjunto LI (veja a Proposição 2.3.4). Portanto, $M_{\mathcal{C}\mathcal{B}}$ é uma matriz inversível. Assim, se $\vec{u} \in \mathbb{V}^3$, a identidade (2.9) lê-se também na forma

$$[\vec{u}]_{\mathcal{C}} = (M_{\mathcal{C}\mathcal{B}})^{-1}[\vec{u}]_{\mathcal{B}}. \tag{2.10}$$

Exemplo 2.6.2 No Exemplo 2.3.5, o item (ii) poderia ter sido mais rapidamente resolvido usando a matriz de mudança de base. Lembre que, como $\vec{f_1} = (2, -1, 0)_{\mathcal{B}}$, $\vec{f_2} = (1, -1, 2)_{\mathcal{B}}$ e $\vec{f_3} = (1, 0, 2)_{\mathcal{B}}$, temos

$$M_{\mathcal{C}\mathcal{B}} = \begin{bmatrix} 2 & 1 & 1 \\ -1 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix},$$

e, portanto, usando (2.9),

$$[\vec{v}]_{\mathcal{B}} = M_{\mathcal{C}\mathcal{B}}[\vec{v}]_{\mathcal{C}} = \begin{bmatrix} 2 & 1 & 1 \\ -1 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}.$$

Agora,

$$(M_{\mathcal{C}\mathcal{B}})^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{4} \\ -\frac{1}{2} & -1 & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{4} \end{bmatrix}.$$

Assim, neste exemplo, se $\vec{w} = (4, 1, 2)_{\mathcal{B}}$, usando (2.10), obtemos

$$[\vec{w}]_{\mathcal{C}} = (M_{\mathcal{C}\mathcal{B}})^{-1}[\vec{w}]_{\mathcal{B}} = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{4} \\ -\frac{1}{2} & -1 & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix}.$$

Proposição 2.6.3 *Sejam $\mathcal{B}$, $\mathcal{C}$ e $\mathcal{D}$ bases de $\mathbb{V}^3$. Então, $M_{\mathcal{D}\mathcal{B}} = M_{\mathcal{C}\mathcal{B}}M_{\mathcal{D}\mathcal{C}}$.*

Demonstração Suponha que $\mathcal{D} = \{\vec{g_1}, \vec{g_2}, \vec{g_3}\}$. Então, para todo $j = 1, 2, 3$, segue do Teorema 2.6.1, que

$$[\vec{g_j}]_{\mathcal{B}} = M_{\mathcal{C}\mathcal{B}}[\vec{g_j}]_{\mathcal{C}},$$

isto é, a j-ésima coluna da matriz $M_{\mathcal{D}\mathcal{B}}$ coincide com o produto de $M_{\mathcal{C}\mathcal{B}}$ pela j-ésima coluna de $M_{\mathcal{D}\mathcal{C}}$. Logo, a igualdade desejada segue. □

Corolário 2.6.4 *Sejam $\mathcal{B}$ e $\mathcal{C}$ bases de $\mathbb{V}^3$. Então, $M_{\mathcal{B}\mathcal{C}} = (M_{\mathcal{C}\mathcal{B}})^{-1}$.*

Demonstração É claro que $M_{\mathcal{B}\mathcal{B}} = I_3$. Da Proposição 2.6.3 obtemos $I_3 = M_{\mathcal{B}\mathcal{B}} = M_{\mathcal{C}\mathcal{B}}M_{\mathcal{B}\mathcal{C}}$. Portanto, $M_{\mathcal{B}\mathcal{C}} = (M_{\mathcal{C}\mathcal{B}})^{-1}$. □

Exercícios Lista 2 - Álgebra Linear I: Exs. 1–6.

2.7 Produto vetorial

Nesta seção, introduziremos uma última operação envolvendo vetores, o produto vetorial. Para tanto, será preciso fixar uma orientação no espaço $\mathbb{V}^3$.

Dadas bases $\mathcal{B}$ e $\mathcal{C}$ de $\mathbb{V}^3$, dizemos que $\mathcal{B}$ e $\mathcal{C}$ têm a *mesma orientação*, se $\det(M_{\mathcal{C}\mathcal{B}}) > 0$. Caso contrário, isto é, se $\det(M_{\mathcal{C}\mathcal{B}}) < 0$, dizemos que $\mathcal{B}$ e $\mathcal{C}$ têm *orientações opostas*.

Observação É relevante notar as seguintes consequências da definição acima.

- $\det(M_{\mathcal{C}\mathcal{B}}) > 0$ se, e somente se, $\det(M_{\mathcal{B}\mathcal{C}}) > 0$. Ou seja, a noção de bases de mesma orientação, ou de orientações opostas, não depende da ordem em que as bases foram apresentadas.
- A propriedade de ter a mesma orientação é transitiva, no sentido de que se $\mathcal{B}$, $\mathcal{C}$ e $\mathcal{D}$ são bases de $\mathbb{V}^3$ tais que $(\mathcal{B}, \mathcal{C})$ é um par de bases de mesma orientação e $(\mathcal{C}, \mathcal{D})$ também é um par de bases de mesma orientação, então $\mathcal{B}$ e $\mathcal{D}$ têm a mesma orientação. Isso segue, imediatamente, da Proposição 2.6.3.
- Também é consequência da Proposição 2.6.3 que se $\mathcal{B}$, $\mathcal{C}$ e $\mathcal{D}$ são bases de $\mathbb{V}^3$ tais que $\mathcal{B}$ e $\mathcal{C}$ têm orientações opostas, e $\mathcal{B}$ e $\mathcal{D}$ também têm orientações opostas, então $\mathcal{C}$ e $\mathcal{D}$ têm mesma orientação.

Diremos que $\mathbb{V}^3$ está *orientado* se estiver fixada uma base $\mathcal{B}$ em $\mathbb{V}^3$. Neste caso, dada uma outra base $\mathcal{C}$ de $\mathbb{V}^3$, diremos que $\mathcal{C}$ é uma *base positiva* se $\mathcal{B}$ e $\mathcal{C}$ tiverem a mesma orientação. Caso contrário, diremos que $\mathcal{C}$ é uma *base negativa*.

Por exemplo, se $\mathcal{B} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ é a base que dá a orientação de $\mathbb{V}^3$, então $\mathcal{C} = \{\vec{e_3}, \vec{e_2}, \vec{e_1}\}$ é uma base negativa (pois $M_{\mathcal{C}\mathcal{B}} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, que tem determinante igual a -1), e $\mathcal{D} = \{\vec{e_3}, \vec{e_1}, \vec{e_2}\}$ é uma base positiva (pois $M_{\mathcal{D}\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, que tem determinante igual a 1).

Observação Se o espaço $\mathbb{V}^3$ estiver orientado, existem infinitas bases ortonormais positivas (e infinitas bases ortonormais negativas). De fato, como vimos na Seção 2.3, existem infinitas bases ortonormais em $\mathbb{V}^3$. Se $C = \{\vec{u}, \vec{v}, \vec{w}\}$ é uma delas e C não é positiva, então $\{\vec{v}, \vec{u}, \vec{w}\}$ será.

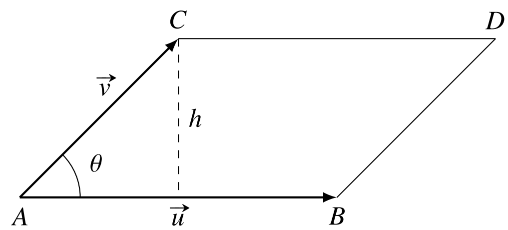
Definição Considere $\mathbb{V}^3$ com uma orientação fixada. Dados $\vec{u}, \vec{v} \in \mathbb{V}^3$, definimos o *produto vetorial* de $\vec{u}$ e $\vec{v}$ como sendo o vetor $\vec{u} \wedge \vec{v}$ que satisfaz as seguintes propriedades:

(i) se $\{\vec{u}, \vec{v}\}$ é LD, então $\vec{u} \wedge \vec{v} = \vec{0}$;
(ii) se $\{\vec{u}, \vec{v}\}$ é LI, então

(a) $\|\vec{u} \wedge \vec{v}\| = \|\vec{u}\|\,\|\vec{v}\| \operatorname{sen}(\theta)$, em que θ denota a medida do ângulo entre $\vec{u}$ e $\vec{v}$;
(b) $\vec{u} \wedge \vec{v}$ é ortogonal a $\vec{u}$ e a $\vec{v}$;
(c) $\{\vec{u}, \vec{v}, \vec{u} \wedge \vec{v}\}$ é uma base positiva de $\mathbb{V}^3$.

Observação Essa definição do produto vetorial em termos geométricos é difícil de ser aplicada. Veremos, adiante, que, em termos de coordenadas em relação a uma base ortonormal positiva, existe uma fórmula útil para o produto vetorial. No entanto, algumas consequências podem ser obtidas já a partir da definição em termos geométricos:

- $\|\vec{u} \wedge \vec{v}\|$ é igual à área do paralelogramo definido por $\vec{u}$ e $\vec{v}$. Mais precisamente, suponha $\{\vec{u}, \vec{v}\}$ LI (caso contrário, eles não definem um paralelogramo) e sejam $A, B, C \in \mathbb{E}^3$ tais que $\vec{u} = \overrightarrow{AB}$ e $\vec{v} = \overrightarrow{AC}$. Seja $D \in \mathbb{E}^3$ tal que $\overrightarrow{AD} = \vec{u} + \vec{v}$. Então, A, B, C, D são os vértices de um paralelogramo, como na figura:



  Denote por h a altura do paralelogramo em relação ao lado AB. Então, $h = (AC)\operatorname{sen}(\theta)$ e a área do paralelogramo é dada por $(AB)h = (AB)(AC)\operatorname{sen}(\theta) = \|\vec{u}\|\,\|\vec{v}\|\operatorname{sen}(\theta) = \|\vec{u} \wedge \vec{v}\|$.
- No caso em que $\{\vec{u}, \vec{v}\}$ é LI, o produto vetorial $\vec{u} \wedge \vec{v}$ não é o vetor nulo, pois, neste caso, $\theta \neq 0$ e $\theta \neq \pi$, o que garante que $\operatorname{sen}(\theta) \neq 0$ e, portanto, $\|\vec{u} \wedge \vec{v}\| \neq 0$. Assim, o produto vetorial fornece um teste de dependência linear:
$$\left\{\vec{u}, \vec{v}\right\} \text{ é LD} \quad \text{se, e somente se,} \quad \vec{u} \wedge \vec{v} = \vec{0}.$$
  Em particular, segue que $\vec{u} \wedge \vec{u} = \vec{0}$, qualquer que seja $\vec{u} \in \mathbb{V}^3$.

Assim, temos, em $\mathbb{V}^3$ dois produtos definidos: o produto escalar, que é um número real, e o produto vetorial, que é um vetor.

Vejamos como encontrar as coordenadas do produto vetorial.

Proposição 2.7.1 *Considere $\mathbb{V}^3$ com uma orientação fixada. Seja $\mathcal{B}$ uma base ortonormal positiva de $\mathbb{V}^3$ e sejam $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)_{\mathcal{B}}$, $\vec{v} = (\beta_1, \beta_2, \beta_3)_{\mathcal{B}} \in \mathbb{V}^3$. Então,*
$$\vec{u} \wedge \vec{v} = \left(\alpha_2\beta_3 - \alpha_3\beta_2,\ \alpha_3\beta_1 - \alpha_1\beta_3,\ \alpha_1\beta_2 - \alpha_2\beta_1\right)_{\mathcal{B}}. \tag{2.11}$$

Demonstração Seja $\vec{w} = (\gamma_1, \gamma_2, \gamma_3)_{\mathcal{B}} \in \mathbb{V}^3$, o vetor cujas coordenadas são dadas por
$$\gamma_1 = \alpha_2\beta_3 - \alpha_3\beta_2, \quad \gamma_2 = \alpha_3\beta_1 - \alpha_1\beta_3, \quad \gamma_3 = \alpha_1\beta_2 - \alpha_2\beta_1.$$
Mostremos que $\vec{w} = \vec{u} \wedge \vec{v}$.

Primeiramente, note que se $\{\vec{u}, \vec{v}\}$ é LD, então $\vec{u} = \lambda\vec{v}$ ou $\vec{v} = \lambda\vec{u}$, para algum $\lambda \in \mathbb{R}$. Em qualquer caso, teremos $\gamma_1 = \gamma_2 = \gamma_3 = 0$ e, portanto, $\vec{w} = \vec{0}$.

Suponha, então, que $\{\vec{u}, \vec{v}\}$ seja LI. Por um lado, temos

$$\|\vec{w}\|^2 = \gamma_1^2 + \gamma_2^2 + \gamma_3^2 = (\alpha_2\beta_3 - \alpha_3\beta_2)^2 + (\alpha_3\beta_1 - \alpha_1\beta_3)^2 + (\alpha_1\beta_2 - \alpha_2\beta_1)^2. \tag{2.12}$$

Por outro lado,

$$\begin{aligned} \|\vec{u} \wedge \vec{v}\|^2 &= \|\vec{u}\|^2\,\|\vec{v}\|^2 \operatorname{sen}^2(\theta) = \|\vec{u}\|^2\,\|\vec{v}\|^2\,(1 - \cos^2(\theta)) \\ &= \|\vec{u}\|^2\,\|\vec{v}\|^2 - \|\vec{u}\|^2\,\|\vec{v}\|^2 \cos^2(\theta) = \|\vec{u}\|^2\,\|\vec{v}\|^2 - (\vec{u} \cdot \vec{v})^2 \\ &= (\alpha_1^2 + \alpha_2^2 + \alpha_3^2)(\beta_1^2 + \beta_2^2 + \beta_3^2) - (\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3)^2. \end{aligned} \tag{2.13}$$

Comparando os lados direitos de (2.12) e (2.13), obtemos que $\|\vec{w}\| = \|\vec{u} \wedge \vec{v}\|$.

Agora,

$$\vec{u} \cdot \vec{w} = \alpha_1\gamma_1 + \alpha_2\gamma_2 + \alpha_3\gamma_3 = \det \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{bmatrix} = 0.$$

A segunda igualdade acima pode ser verificada expandindo-se o determinante em cofatores ao longo da primeira linha. Segue que $\vec{u}$ e $\vec{w}$ são ortogonais. De maneira análoga, prova-se que $\vec{v}$ e $\vec{w}$ também são ortogonais.

Finalmente, mostremos que $\mathcal{C} = \{\vec{u}, \vec{v}, \vec{w}\}$ é uma base positiva de $\mathbb{V}^3$. Para tanto, basta verificar que esse conjunto é LI e que a matriz de mudança de base $M_{\mathcal{CB}}$ tem determinante positivo. Para mostrar que $\mathcal{C}$ é LI, considere a matriz

$$A = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix}.$$

Calculando o determinante de A por expansão em cofatores ao longo da terceira linha, obtemos

$$\det(A) = \gamma_1(\alpha_2\beta_3 - \alpha_3\beta_2) - \gamma_2(\alpha_1\beta_3 - \alpha_3\beta_1) + \gamma_3(\alpha_1\beta_2 - \alpha_2\beta_1) = \gamma_1^2 + \gamma_2^2 + \gamma_3^2 = \|\vec{w}\|^2.$$

Como vimos acima que $\|\vec{w}\| = \|\vec{u} \wedge \vec{w}\|$ e $\{\vec{u}, \vec{v}\}$ é LI, segue que $\det(A) = \|\vec{u} \wedge \vec{v}\|^2 \neq 0$. Assim, $\mathcal{C}$ é, de fato, uma base de $\mathbb{V}^3$. Para ver que $\mathcal{C}$ é positiva, é preciso mostrar que $\det(M_{\mathcal{CB}}) > 0$. Agora, $M_{\mathcal{CB}} = A^\intercal$. Portanto, $\det(M_{\mathcal{CB}}) = \det(A^\intercal) = \det(A) = \|\vec{u} \wedge \vec{v}\|^2 > 0$. Logo, $\mathcal{C}$ é, com efeito, uma base positiva de $\mathbb{V}^3$.

Conclui-se que $\vec{w} = \vec{u} \wedge \vec{v}$, pois esses dois vetores têm o mesmo comprimento, são paralelos e têm o mesmo sentido. □

Notação para o produto vetorial em forma de determinante

A fim de facilitar a memorização da expressão (2.11), utilizamos a seguinte notação expandida para o determinante. Considere uma orientação fixada em $\mathbb{V}^3$, e seja $\mathcal{B} = \{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ uma base ortonormal positiva. Então, se $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)_{\mathcal{B}}$ e $\vec{v} = (\beta_1, \beta_2, \beta_3)_{\mathcal{B}}$, segue que

$$\vec{u} \wedge \vec{v} = \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{bmatrix},$$

em que esse "determinante" é calculado por "expansão em cofatores ao longo da primeira linha":

$$\vec{u} \wedge \vec{v} = \det \begin{bmatrix} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{bmatrix} \vec{\imath} - \det \begin{bmatrix} \alpha_1 & \alpha_3 \\ \beta_1 & \beta_3 \end{bmatrix} \vec{\jmath} + \det \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix} \vec{k}.$$

Exemplo 2.7.2 Considere uma orientação fixada em $\mathbb{V}^3$ e coordenadas dadas em relação a uma base ortonormal positiva. Calcule $\vec{u} \wedge \vec{v}$, em que $\vec{u} = (1, 2, 3)$ e $\vec{v} = (-1, 1, 2)$.

Solução: Usando a notação em forma de determinante,

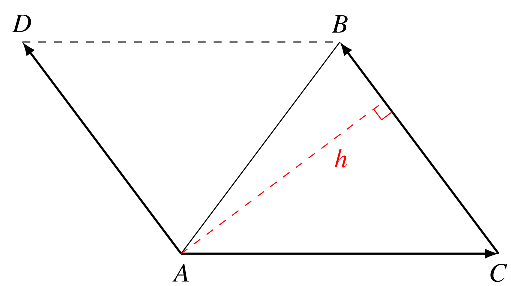
$$\vec{u} \wedge \vec{v} = \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix} = \det \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \vec{\imath} - \det \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix} \vec{\jmath} + \det \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \vec{k} = (1, -5, 3). \quad \diamond$$

Observação É fácil verificar, usando, por exemplo, a notação em forma de determinante para o produto vetorial, que se $\mathbb{V}^3$ tem uma orientação fixada e $\{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ é uma base ortonormal positiva, então

$$\begin{aligned} \vec{\imath} \wedge \vec{\jmath} &= \vec{k}, & \qquad \vec{\jmath} \wedge \vec{\imath} &= -\vec{k}, \\ \vec{\jmath} \wedge \vec{k} &= \vec{\imath}, & \qquad \vec{k} \wedge \vec{\jmath} &= -\vec{\imath}, \\ \vec{k} \wedge \vec{\imath} &= \vec{\jmath}, & \qquad \vec{\imath} \wedge \vec{k} &= -\vec{\jmath}. \end{aligned}$$

Exemplo 2.7.3 Calcule a área do triângulo ABC, em que $\overrightarrow{AC} = (1, 1, 3)$ e $\overrightarrow{CB} = (-1, 1, 0)$ e as coordenadas estão dadas em relação a uma base ortonormal positiva de $\mathbb{V}^3$. Determine a altura do triângulo ABC em relação ao lado BC.

Solução: Seja D o ponto tal que $\overrightarrow{AD} = \overrightarrow{CB}$, conforme a figura:



Sabemos que a área do paralelogramo $ACBD$ é dada por $\left\|\overrightarrow{AC} \wedge \overrightarrow{AD}\right\| = \left\|\overrightarrow{AC} \wedge \overrightarrow{CB}\right\|$. Logo, a área do triângulo ABC é dada por $\frac{1}{2}\left\|\overrightarrow{AC} \wedge \overrightarrow{CB}\right\|$. Agora,

$$\overrightarrow{AC} \wedge \overrightarrow{CB} = \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 1 & 3 \\ -1 & 1 & 0 \end{bmatrix} = (-3, -3, 2).$$

Logo, a área do triângulo ABC é igual a $\frac{1}{2}\sqrt{(-3)^2 + (-3)^2 + 2^2} = \frac{1}{2}\sqrt{22}$. Se h denota a altura do triângulo ABC em relação ao lado BC, então a área de ABC também pode ser dada por $\frac{1}{2}\left\|\overrightarrow{BC}\right\| h$. Como $\left\|\overrightarrow{BC}\right\| = \left\|\overrightarrow{CB}\right\| = \sqrt{(-1)^2 + 1^2 + 0^2} = \sqrt{2}$, segue que $\frac{1}{2}\sqrt{22} = \frac{1}{2}\sqrt{2}h$. Portanto, $h = \sqrt{11}$. $\diamond$

A seguir vemos como o produto vetorial comporta-se em relação a outras operações definidas no espaço do vetores.

Proposição 2.7.4 *Considere fixada uma orientação em $\mathbb{V}^3$. Sejam $\vec{u}, \vec{u_1}, \vec{u_2}, \vec{v}, \vec{v_1}, \vec{v_2} \in \mathbb{V}^3$ e $\lambda \in \mathbb{R}$. Então,*

(i) $\vec{u} \wedge (\vec{v_1} + \vec{v_2}) = \vec{u} \wedge \vec{v_1} + \vec{u} \wedge \vec{v_2}$;
(ii) $(\vec{u_1} + \vec{u_2}) \wedge \vec{v} = \vec{u_1} \wedge \vec{v} + \vec{u_2} \wedge \vec{v}$;
(iii) $(\lambda\vec{u}) \wedge \vec{v} = \vec{u} \wedge (\lambda\vec{v}) = \lambda(\vec{u} \wedge \vec{v})$;
(iv) $\vec{u} \wedge \vec{v} = -(\vec{v} \wedge \vec{u})$.

Demonstração Para a demonstração dessas propriedades, basta considerar uma base ortonormal positiva $\mathcal{B}$ e coordenadas em relação a essa base. Assim, por exemplo, se $\vec{u} = (x, y, z)_{\mathcal{B}}$, $\vec{v_1} = (x_1, y_1, z_1)_{\mathcal{B}}$ e $\vec{v_2} = (x_2, y_2, z_2)_{\mathcal{B}}$, então

$$\vec{u} \wedge (\vec{v_1} + \vec{v_2}) = \big(y(z_1 + z_2) - z(y_1 + y_2),\ z(x_1 + x_2) - x(z_1 + z_2),\ x(y_1 + y_2) - y(x_1 + x_2)\big)$$
$$= (yz_1 - zy_1, zx_1 - xz_1, xy_1 - yx_1) + (yz_2 - zy_2, zx_2 - xz_2, xy_2 - yx_2),$$

o que demonstra (i). As demais propriedades podem ser mostradas de modo análogo. □

Observação Como vimos, no item (iv) proposição acima, o produto vetorial não é comutativo, isto é, a ordem em que os vetores aparecem importa.

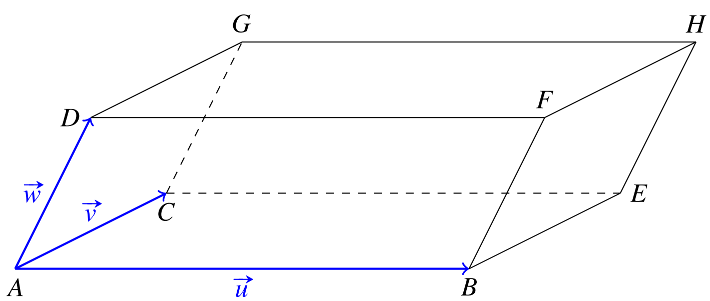
Também vale notar que o produto vetorial não é uma operação associativa, isto é, em uma sequência de três vetores, distribuições diferentes de parênteses podem resultar em vetores diferentes. Por exemplo, se $\{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ é uma base ortonormal positiva, então $(\vec{\jmath} \wedge \vec{\jmath}) \wedge \vec{\imath} = \vec{0} \wedge \vec{\imath} = \vec{0}$, ao passo que $\vec{\jmath} \wedge (\vec{\jmath} \wedge \vec{\imath}) = \vec{\jmath} \wedge (-\vec{k}) = -(\vec{\jmath} \wedge \vec{k}) = -\vec{\imath}$.

Exercícios Lista 2 - Álgebra Linear I: Exs. 7–23.

2.8 Produto misto

Ao procurar calcular o volume de um paralelepípedo, deparamo-nos com uma expressão em que aparecem tanto o produto vetorial quando o produto escalar. Chamaremos essa expressão de produto misto, como definido adiante.

Sejam $\vec{u}, \vec{v}, \vec{w} \in \mathbb{V}^3$ com $\{\vec{u}, \vec{v}, \vec{w}\}$ LI. Se $A, B, C, D \in \mathbb{E}^3$ são tais que $\vec{u} = \overrightarrow{AB}$, $\vec{v} = \overrightarrow{AC}$ e $\vec{w} = \overrightarrow{AD}$, então os pontos A, B, C, D não estão em um mesmo plano e portanto definem um paralelepípedo, conforme a figura



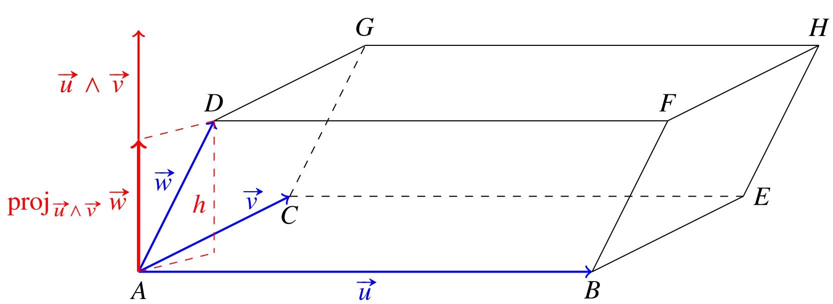
Vamos calcular o volume **V** do paralelepípedo $ABCDEFGH$. Sabemos que esse volume pode ser expresso por

$$\mathbf{V} = \mathbf{A}_\mathrm{b} h,$$

em que $\mathbf{A}_\mathrm{b}$ denota a área da base, formada pelo paralelogramo $ABCE$, e h denota a altura do paralelepípedo em relação a essa face. Já vimos que, se $\mathbb{V}^3$ estiver orientado (e, portanto, o produto vetorial entre quaisquer dois vetores estiver definido), então

$$\mathbf{A}_\mathrm{b} = \|\vec{u} \wedge \vec{v}\|.$$

Resta-nos determinar a altura h. Sabemos que $\vec{u} \wedge \vec{v}$ é um vetor ortogonal a $\vec{u}$ e a $\vec{v}$. Assim, seu representante com extremidade inicial no ponto A é perpendicular ao plano ABC, que contém o paralelogramo $ABEC$, como na figura.



A altura h é dada pela distância do vértice D ao plano ABC, e, portanto, coincide com a norma da projeção ortogonal de $\vec{w}$ sobre $\vec{u} \wedge \vec{v}$. [3] Logo,

$$\begin{aligned} h &= \left\| \operatorname{proj}_{\vec{u}\wedge\vec{v}} \vec{w} \right\| = \left\| \frac{\vec{w} \cdot (\vec{u} \wedge \vec{v})}{\| \vec{u} \wedge \vec{v} \|^2} \vec{u} \wedge \vec{v} \right\| \\ &= \frac{|\vec{w} \cdot (\vec{u} \wedge \vec{v})|}{\| \vec{u} \wedge \vec{v} \|^2} \| \vec{u} \wedge \vec{v} \| = \frac{|(\vec{u} \wedge \vec{v}) \cdot \vec{w}|}{\| \vec{u} \wedge \vec{v} \|}. \end{aligned}$$

Assim, o volume do paralelepípedo $ABCDEFGH$ é dado por

$$\mathbf{V} = \mathbf{A}_b h = \| \vec{u} \wedge \vec{v} \| \frac{|(\vec{u} \wedge \vec{v}) \cdot \vec{w}|}{\| \vec{u} \wedge \vec{v} \|} = |(\vec{u} \wedge \vec{v}) \cdot \vec{w}|.$$

Esse cálculo motiva a seguinte definição.

Definição Considere uma orientação fixa em $\mathbb{V}^3$. Dados $\vec{u}, \vec{v}, \vec{w} \in \mathbb{V}^3$, definimos o *produto misto* de $\vec{u}, \vec{v}, \vec{w}$ (nesta ordem) como sendo o número real dado por $(\vec{u} \wedge \vec{v}) \cdot \vec{w}$ e denotado por $[\vec{u}, \vec{v}, \vec{w}]$.

Conforme vimos acima, o volume do paralelepípedo definido pelos vetores $\vec{u}, \vec{v}, \vec{w}$ é igual a $|[\vec{u}, \vec{v}, \vec{w}]|$.

O próximo resultado exibe uma fórmula para o produto misto em termos de coordenadas.

Proposição 2.8.1 *Considere fixada uma orientação em $\mathbb{V}^3$, e seja $\mathcal{B}$ uma base ortonormal positiva de $\mathbb{V}^3$. Se $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)_{\mathcal{B}}$, $\vec{v} = (\beta_1, \beta_2, \beta_3)_{\mathcal{B}}$, $\vec{w} = (\gamma_1, \gamma_2, \gamma_3)_{\mathcal{B}} \in \mathbb{V}^3$, então*

$$[\vec{u}, \vec{v}, \vec{w}] = \det \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix}.$$

Demonstração Sabemos, da Proposição 2.7.1, que

$$\vec{u} \wedge \vec{v} = \left(\det \begin{bmatrix} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{bmatrix}, -\det \begin{bmatrix} \alpha_1 & \alpha_3 \\ \beta_1 & \beta_3 \end{bmatrix}, \det \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix} \right)_{\mathcal{B}}.$$

Assim,

$$\begin{aligned} [\vec{u}, \vec{v}, \vec{w}] = (\vec{u} \wedge \vec{v}) \cdot \vec{w} &= \det \begin{bmatrix} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{bmatrix} \gamma_1 - \det \begin{bmatrix} \alpha_1 & \alpha_3 \\ \beta_1 & \beta_3 \end{bmatrix} \gamma_2 + \det \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix} \gamma_3 \\ &= \det \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix}, \end{aligned}$$

expandido em cofatores ao longo da terceira linha. □

Assim como o produto vetorial fornecia um teste para dependência linear entre dois vetores, o produto misto fornece um teste para dependência linear entre três vetores:

Corolário 2.8.2 *Considere fixada uma orientação em $\mathbb{V}^3$, e seja $\mathcal{B}$ uma base ortonormal positiva de $\mathbb{V}^3$. Sejam $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)_{\mathcal{B}}$, $\vec{v} = (\beta_1, \beta_2, \beta_3)_{\mathcal{B}}$, $\vec{w} = (\gamma_1, \gamma_2, \gamma_3)_{\mathcal{B}} \in \mathbb{V}^3$. Então,*

$$\left\{ \vec{u}, \vec{v}, \vec{w} \right\} \text{ é LD} \quad \text{se, e somente se,} \quad \left[\vec{u}, \vec{v}, \vec{w} \right] = 0.$$

[3] A figura sugere que $\{\vec{u}, \vec{v}, \vec{w}\}$ e $\{\vec{u}, \vec{v}, \vec{u} \wedge \vec{v}\}$ sejam bases de mesma orientação. Mesmo que não sejam, a altura do parelelepípedo coincide com $\left\| \operatorname{proj}_{\vec{u}\wedge\vec{v}} \vec{w} \right\|$.

Demonstração Vimos, na Proposição 2.3.4 que $\{\vec{u}, \vec{v}, \vec{w}\}$ é LD se, e somente se, $\det \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix} = 0$. Como vimos na proposição acima, o determinante coincide com o produto misto $[\vec{u}, \vec{v}, \vec{w}]$. □

Em particular, se há uma repetição de vetores em um produto misto, ele é nulo: $[\vec{u}, \vec{u}, \vec{v}] = [\vec{u}, \vec{v}, \vec{u}] = [\vec{v}, \vec{u}, \vec{u}] = 0$. Ao final desta seção, veremos como o produto misto se comporta em relação a permutações de vetores e em relação às operações de soma e multiplicação por escalar.

Exemplo 2.8.3 Calcule, usando o produto misto, o volume do tetraedro de vértices A, B, C, D.

Solução: Sabemos que o volume $\mathbf{V}$ do tetraedo de vértices A, B, C, D é dado por

$$\mathbf{V} = \frac{1}{3}\mathbf{A}_\mathrm{b} h,$$

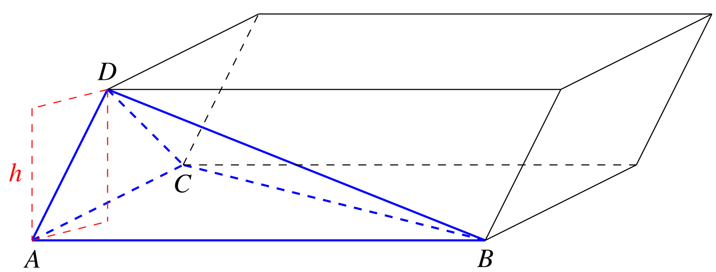
em que $\mathbf{A}_\mathrm{b}$ denota a área da base triangular ABC e h denota a altura do tetraedo em relação a essa face. Como a área do triângulo ABC é metade da área do paralelogramo definido por $\overrightarrow{AB}$ e $\overrightarrow{AC}$, segue

$$\mathbf{A}_\mathrm{b} = \frac{1}{2}\left\|\overrightarrow{AB} \wedge \overrightarrow{AC}\right\|.$$

A altura h pode ser expressa por

$$h = \frac{|[\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}]|}{\left\|\overrightarrow{AB} \wedge \overrightarrow{AC}\right\|},$$

uma vez que ela coincide com a altura do paralelepípedo definido pelos vetores $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}$.

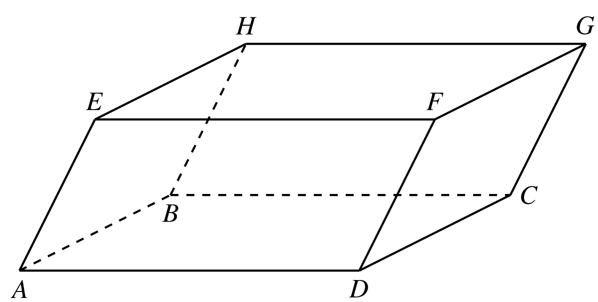


Daí, conclui-se que

$$\mathbf{V} = \frac{1}{3}\frac{1}{2}\left\|\overrightarrow{AB} \wedge \overrightarrow{AC}\right\| \frac{|[\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}]|}{\left\|\overrightarrow{AB} \wedge \overrightarrow{AC}\right\|} = \frac{1}{6}|[\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}]|.$$

(Aqui, supusemos, implicitamente, que $\mathbb{V}^3$ estava orientado, para que pudéssemos falar no produto vetorial e misto de vetores. Isso será feito sempre que pertinente.) ◇

Exemplo 2.8.4 Considere o paralelepípedo $ABCDEFGH$ da figura:



Suponha que esteja fixada uma orientação em $\mathbb{V}^3$ e que $\overrightarrow{AB} = (1, 0, 1)$, $\overrightarrow{BE} = (1, 1, 1)$ e $\overrightarrow{AD} = (0, 3, 3)$, e que essas coordenadas estejam expressas em relação a uma base ortonormal positiva de $\mathbb{V}^3$.

(i) Encontre o volume do paralelepípedo $ABCDEFGH$.
(ii) Encontre o volume do tetraedro $EABD$.
(iii) Determine a altura do tetraedro $EABD$ em relação à face DEB.

Solução: Temos $\overrightarrow{AE} = \overrightarrow{AB} + \overrightarrow{BE} = (1, 0, 1) + (1, 1, 1) = (2, 1, 2)$. Logo, $[\overrightarrow{AB}, \overrightarrow{AD}, \overrightarrow{AE}] = \det \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 3 \\ 2 & 1 & 2 \end{bmatrix} = -3.$

(i) O volume do paralelepípedo é igual a $|-3| = 3$.

(ii) Conforme vimos no exemplo anterior, o volume do tetraedro $EABD$ é igual a $\frac{1}{6}|-3| = \frac{1}{2}$.

(iii) Sabemos que $\frac{1}{2}$ = volume de $EABD = \frac{1}{3}\mathbf{A}h$, em que $\mathbf{A}$ denota a área do triângulo DEB e h denota a altura do tetraedo $EADB$ em relação à face DEB. Como $\mathbf{A} = \frac{1}{2}\left\|\overrightarrow{DE} \wedge \overrightarrow{DB}\right\|$, segue que

$$h = \frac{3}{\left\|\overrightarrow{DE} \wedge \overrightarrow{DB}\right\|}.$$

Resta, portanto, encontrar $\overrightarrow{DE} \wedge \overrightarrow{DB}$. Temos

$$\overrightarrow{DE} = \overrightarrow{DA} + \overrightarrow{AE} = -\overrightarrow{AD} + \overrightarrow{AE} = -(0, 3, 3) + (2, 1, 1) = (2, -2, -1)$$

e

$$\overrightarrow{DB} = \overrightarrow{DA} + \overrightarrow{AB} = -\overrightarrow{AD} + \overrightarrow{AB} = -(0, 3, 3) + (1, 0, 1) = (1, -3, -2).$$

Logo,

$$\overrightarrow{DE} \wedge \overrightarrow{DB} = \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 2 & -2 & -1 \\ 1 & -3 & -2 \end{bmatrix} = (1, 3, -4).$$

Assim, $\left\|\overrightarrow{DE} \wedge \overrightarrow{DB}\right\| = \sqrt{26}$. Portanto, $h = \frac{3}{\sqrt{26}}$. $\diamond$

Proposição 2.8.5 *Considere fixada uma orientação em $\mathbb{V}^3$, seja $\mathcal{B}$ uma base ortormal positiva de $\mathbb{V}^3$ e seja $C = \{\vec{u}, \vec{v}, \vec{w}\}$ uma outra base de $\mathbb{V}^3$. Então,*

$$\det(M_{C\mathcal{B}}) = [\vec{u}, \vec{v}, \vec{w}].$$

Demonstração Suponha que $\vec{u} = (\alpha_1, \alpha_2, \alpha_3)_{\mathcal{B}}$, $\vec{v} = (\beta_1, \beta_2, \beta_3)_{\mathcal{B}}$ e $\vec{w} = (\gamma_1, \gamma_2, \gamma_3)_{\mathcal{B}}$. Então, pela Proposição 2.8.1, temos

$$[\vec{u}, \vec{v}, \vec{w}] = \det \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix} = \det({M_{C\mathcal{B}}}^{\intercal}) = \det(M_{C\mathcal{B}}),$$

como queríamos demonstrar. □

Podemos juntar o Corolário 2.8.2 com a conclusão da Proposição 2.8.5 para obter o seguinte critério:

Corolário 2.8.6 *Considere fixada uma orientação em $\mathbb{V}^3$, e sejam $\vec{u}, \vec{v}, \vec{w} \in \mathbb{V}^3$.*

(i) Se $[\vec{u}, \vec{v}, \vec{w}] = 0$, então $\{\vec{u}, \vec{v}, \vec{w}\}$ não é base de $\mathbb{V}^3$.
(ii) Se $[\vec{u}, \vec{v}, \vec{w}] > 0$, então $\{\vec{u}, \vec{v}, \vec{w}\}$ é uma base positiva de $\mathbb{V}^3$.
(iii) Se $[\vec{u}, \vec{v}, \vec{w}] < 0$, então $\{\vec{u}, \vec{v}, \vec{w}\}$ é uma base negativa de $\mathbb{V}^3$.

Uma das propriedades mais relevantes do produto misto é que ele é o que se chama de um produto *alternado*, isto é, se dois dos vetores em um produto misto forem permutados, o produto misto muda de sinal:

$$[\vec{u}, \vec{v}, \vec{w}] = -[\vec{v}, \vec{u}, \vec{w}] = [\vec{v}, \vec{w}, \vec{u}] = -[\vec{u}, \vec{w}, \vec{v}] = [\vec{w}, \vec{u}, \vec{v}] = -[\vec{w}, \vec{v}, \vec{u}].$$

Isso pode ser demonstrando, por exemplo, por meio da Proposição 2.8.1, usando o fato de que uma permutações entre duas linhas no cálculo de um determinante muda seu sinal.

Como consequência, segue que $(\vec{u} \wedge \vec{v}) \cdot \vec{w} = \vec{u} \cdot (\vec{v} \wedge \vec{w})$, pois o primeiro número é igual a $[\vec{u}, \vec{v}, \vec{w}]$ e o segundo, a $(\vec{v} \wedge \vec{w}) \cdot \vec{u} = [\vec{v}, \vec{w}, \vec{u}]$.

Por fim, como consequência das Proposições 2.4.2 e 2.7.4, pode-se demonstrar que o produto misto é *trilinear*, isto é, que se em uma de suas três entradas se colocar uma combinação linear de vetores, isso resultará na combinação linear correspondente de produtos mistos. Por exemplo,

$$[\lambda\vec{u_1} + \mu\vec{u_2}, \vec{v}, \vec{w}] = \lambda[\vec{u_1}, \vec{v}, \vec{w}] + \mu[\vec{u_2}, \vec{v}, \vec{w}],$$

e analogamente para combinações na segunda ou terceira posições do produto misto.

Como consequência dessa última propriedade segue que o produto misto não se altera se somarmos a um de seus vetores uma combinação linear dos outros dois.

Exemplo 2.8.7 (Prova 2, Álgebra Linear I, 2015) Suponha fixada uma orientação no espaço $\mathbb{V}^3$. Considere as seguintes afirmações sobre vetores $\vec{v}, \vec{w}, \vec{z}, \vec{z_1}, \vec{z_2} \in \mathbb{V}^3$.

I. Se $[\vec{v}, \vec{w}, \vec{z}] = 5$, então $[\vec{v}, \vec{w} + 2\vec{z}, \vec{z} - 3\vec{v}] = 5$.
II. Se $[\vec{v}, \vec{w}, \vec{z_1}] = 2$ e $[\vec{v}, \vec{w}, \vec{z_2}] = 3$, então $[\vec{v}, \vec{w}, \vec{z_1} + 3\vec{z_2}] = 11$.
III. $[\vec{v}, \vec{w}, \vec{z}] = [\vec{w}, \vec{z}, \vec{v}]$.

Está correto o que se afirma em

(A) I, II e III. (B) II e III, apenas. (C) I e III, apenas. (D) III, apenas. (E) I, apenas.

Solução: Vejamos I:

$$\begin{aligned}
[\vec{v}, \vec{w} + 2\vec{z}, \vec{z} - 3\vec{v}] &= [\vec{v}, \vec{w}, \vec{z} - 3\vec{v}] + 2[\vec{v}, \vec{z}, \vec{z} - 3\vec{v}] \\
&= [\vec{v}, \vec{w}, \vec{z}] - 3\underbrace{[\vec{v}, \vec{w}, \vec{v}]}_{=0} + 2(\underbrace{[\vec{v}, \vec{z}, \vec{z}]}_{=0} - 3\underbrace{[\vec{v}, \vec{z}, \vec{v}]}_{=0}) \\
&= [\vec{v}, \vec{w}, \vec{z}] = 5.
\end{aligned}$$

Logo, I está correta. Consideremos, agora, II:

$$[\vec{v}, \vec{w}, \vec{z_1} + 3\vec{z_2}] = [\vec{v}, \vec{w}, \vec{z_1}] + 3[\vec{v}, \vec{w}, \vec{z_2}] = 2 + 9 = 11.$$

Assim, II também está correta. Finalmente, III:

$$[\vec{v}, \vec{w}, \vec{z}] = -[\vec{w}, \vec{v}, \vec{z}] = [\vec{w}, \vec{z}, \vec{v}].$$

Ou seja, III está correta. *Reposta:* (A). ◇

Exercícios Lista 2 - Álgebra Linear I: Exs. 24–27.

Capítulo 3
Geometria analítica

Neste capítulo, utilizaremos o que vimos sobre vetores para resolver problemas geométricos tridimensionais, envolvendo pontos, retas e planos.

3.1 Sistemas de coordenadas

Definição Um *sistema de coordenadas* em $\mathbb{E}^3$ é um par $\Sigma = (O, \mathcal{B})$, em que $O \in \mathbb{E}^3$ é um ponto, chamado *origem* de Σ, e $\mathcal{B}$ é uma base de $\mathbb{V}^3$. Se $\mathcal{B}$ for uma base ortonormal, dizemos que o sistema de coordenadas Σ é *ortogonal*.

Dado um sistema de coordenadas $\Sigma = (O, \mathcal{B})$ em $\mathbb{E}^3$ e dado um ponto $P \in \mathbb{E}^3$, se $\overrightarrow{OP} = (x, y, z)_{\mathcal{B}}$, dizemos que x, y, z são as *coordenadas* de P em relação ao sistema de coordenadas Σ, e escrevemos $P = (x, y, z)_\Sigma$.

Assim como coordenadas de um vetor em relação a uma base são univocamente determinadas pelo vetor, coordenadas de um ponto em relação a um sistema de coordenadas também são únicas. Mais precisamente, seja $\Sigma = (O, \mathcal{B})$ um sistema de coordenadas em $\mathbb{E}^3$ e sejam $P, Q \in \mathbb{E}^3$ tais que $P = (x, y, z)_\Sigma$ e $Q = (x, y, z)_\Sigma$. Então, $\overrightarrow{OP} = (x, y, z)_{\mathcal{B}} = \overrightarrow{OQ}$, e isso implica $P = Q$.

Nosso primeiro resultado relaciona as coordenadas de dois pontos com as coordenadas do vetor que tem um representante cujas extremidades são esses pontos.

Proposição 3.1.1 *Seja $\Sigma = (O, \mathcal{B})$ um sistema de coordenadas em $\mathbb{E}^3$. Se $P, Q \in \mathbb{E}^3$ são tais que $P = (x_1, y_1, z_1)_\Sigma$ e $Q = (x_2, y_2, z_2)_\Sigma$, então $\overrightarrow{PQ} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)_{\mathcal{B}}$.*

Demonstração Pela definição de coordenadas de um ponto, temos que $\overrightarrow{OP} = (x_1, y_1, z_1)_{\mathcal{B}}$ e $\overrightarrow{OQ} = (x_2, y_2, z_2)_{\mathcal{B}}$. Então,

$$\begin{aligned}\overrightarrow{PQ} &= \overrightarrow{PO} + \overrightarrow{OQ} = -\overrightarrow{OP} + \overrightarrow{OQ} = -(x_1, y_1, z_1)_{\mathcal{B}} + (x_2, y_2, z_2)_{\mathcal{B}} \\ &= (x_2 - x_1, y_2 - y_1, z_2 - z_1)_{\mathcal{B}},\end{aligned}$$

que é a expressão desejada. □

Soma de ponto com vetor

Vimos, no início do Capítulo 2, que dado um vetor $\vec{u} \in \mathbb{V}^3$ e um ponto $P \in \mathbb{E}^3$, existe um único ponto $Q \in \mathbb{E}^3$ tal que $\vec{u} = \overrightarrow{PQ}$. Esse fato será, por conveniência, denotado por $Q = P + \vec{u}$. (Podemos pensar que Q é o ponto obtido a partir de P por translação, ao longo da direção e sentido de $\vec{u}$, de distância $\|\vec{u}\|$ de P.)

Dados $P \in \mathbb{E}^3$ e $\vec{u} \in \mathbb{V}^3$, a título de abreviação, escreveremos $P - \vec{u}$ para denotar o ponto $P + (-\vec{u})$.

São de demonstração imediata, a partir da definição, as propridades listadas a seguir.

Proposição 3.1.2 *Sejam $P, Q \in \mathbb{E}^3$ e $\vec{u}, \vec{v} \in \mathbb{V}^3$. Então, valem:*

(i) $P + \vec{0} = P$;
(ii) $(P + \vec{u}) + \vec{v} = P + (\vec{u} + \vec{v})$;
(iii) se $P + \vec{u} = P + \vec{v}$, então $\vec{u} = \vec{v}$;
(iv) se $P + \vec{u} = Q + \vec{u}$, então $P = Q$.

As coordenadas da soma de um ponto com um vetor podem ser obtidas a partir das coordenadas do ponto e das coordenadas do vetor, conforme a seguinte proposição.

Proposição 3.1.3 *Seja $\Sigma = (O, \mathcal{B})$ um sistema de coordenadas em $\mathbb{E}^3$. Se $P = (x, y, z)_\Sigma$ e $\vec{u} = (a, b, c)_\mathcal{B}$, então*

$$P + \vec{u} = \big(x + a, y + b, z + c\big)_\Sigma.$$

Demonstração Seja $Q = P + \vec{u}$. Então, $\overrightarrow{PQ} = \vec{u}$. Se $Q = (x', y', z')_\Sigma$, pela Proposição 3.1.1, temos

$$(x' - x, y' - y, z' - z)_\mathcal{B} = \overrightarrow{PQ} = \vec{u} = (a, b, c)_\mathcal{B}.$$

Logo, $x' - x = a$, $y' - y = b$, $z' - z = c$. Portanto, $Q = (x + a, y + b, z + c)_\Sigma$. □

Vejamos alguns exemplos.

Exemplo 3.1.4 Suponha fixado o sistema de coordenadas $\Sigma = (O, \mathcal{B})$ em $\mathbb{E}^3$. Dados $P = (1, 3, -3)_\Sigma$, $Q = (0, -1, 4)_\Sigma \in \mathbb{E}^3$ e $\vec{v} = (-1, 4, 0)_\mathcal{B} \in \mathbb{V}^3$, determine as coordenadas de

(i) $\overrightarrow{QP}$;
(ii) $P + \vec{v}$;
(iii) $Q + 2\overrightarrow{PQ}$.

Solução: (i) $\overrightarrow{QP} = (1 - 0, 3 - (-1), -3 - 4)_\mathcal{B} = (1, 4, -7)_\mathcal{B}$; (ii) $P + \vec{v} = (1 + (-1), 3 + 4, -3 + 0)_\Sigma = (0, 7, -3)_\Sigma$; (iii) sabemos que $\overrightarrow{PQ} = -\overrightarrow{QP} = -(1, 4, -7)_\mathcal{B} = (-1, -4, 7)_\mathcal{B}$. Portanto, $2\overrightarrow{PQ} = 2(-1, -4, 7)_\mathcal{B} = (-2, -8, 14)_\mathcal{B}$. Logo, $Q + 2\overrightarrow{PQ} = (0 + (-2), -1 + (-8), 4 + 14)_\Sigma = (-2, -9, 18)_\Sigma$. ◇

Exemplo 3.1.5 Suponha fixado o sistema de coordenadas $\Sigma = (O, \mathcal{B})$ em $\mathbb{E}^3$. Determine as coordenadas do ponto médio do segmento de extremidades $A = (-1, 4, 7)_\Sigma$ e $B = (0, 1, 1)_\Sigma$.

Solução: Se M denota o ponto médio do segmento AB, então $AM = \frac{1}{2}\overrightarrow{AB}$, ou seja, $M = A + \frac{1}{2}\overrightarrow{AB}$. Como $\overrightarrow{AB} = (0 - (-1), 1 - 4, 1 - 7)_\mathcal{B} = (1, -3, -6)_\mathcal{B}$, segue que $M = (-1 + \frac{1}{2}, 4 - \frac{3}{2}, 7 - 3)_\Sigma = (-\frac{1}{2}, \frac{5}{2}, 4)_\Sigma$. ◇

Observe que o argumento utilizado no exemplo acima pode ser generalizado para obter que as coordenadas do ponto médio M do segmento de extremidades $A = (x_1, y_1, z_1)_\Sigma$ e $B = (x_2, y_2, z_2)_\Sigma$ são dadas por

$$M = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}, \frac{z_1 + z_2}{2}\right)_\Sigma,$$

isto é, são as médias das coordenadas das extremidades. (Verifique!)

Exemplo 3.1.6 Mostre que os pontos $A = (1, 0, 1)$, $B = (-1, 0, 2)$ e $C = (1, 1, 1)$ são vértices de um triângulo retângulo. Aqui, coordenadas estão dadas em relação a um sistema ortogonal.

Solução: Temos

$$\overrightarrow{AB} = (-2, 0, 1)_\mathcal{B} \quad \text{e} \quad \overrightarrow{AC} = (0, 1, 0)_\mathcal{B},$$

em que $\mathcal{B}$ denota a base do sistema de coordenadas. Como $\{\overrightarrow{AB}, \overrightarrow{AC}\}$ é LI, os três pontos não são colineares, e, portanto, são vértices de um triângulo. Esse triângulo é retângulo, pois $\overrightarrow{AB}$ e $\overrightarrow{AC}$ são vetores ortogonais, já que $\overrightarrow{AB} \cdot \overrightarrow{AC} = 0$. Para esse último cálculo, utilizamos a Proposição 2.4.1, uma vez que $\mathcal{B}$ é uma base ortonormal de $\mathbb{V}^3$. ◇

Observação Se $\Sigma = (O, \mathcal{B})$ é um sistema ortogonal de coordenadas em $\mathbb{E}^3$ e $P, Q \in \mathbb{E}^3$ são tais que $P = (x_1, y_1, z_1)_\Sigma$ e $Q = (x_2, y_2, z_2)_\Sigma$, então a *distância* $d(P, Q)$ entre P e Q é dada por

$$d(P,Q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

Com efeito, $d(P,Q) = \left\|\overrightarrow{PQ}\right\|$.

3.2 Retas

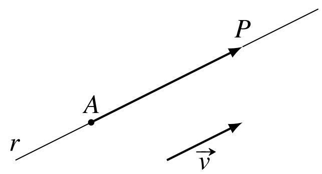
Nesta seção, assumiremos como conhecido o conceito de reta no espaço tridimensional $\mathbb{E}^3$.

Equação vetorial da reta

Seja r uma reta em $\mathbb{E}^3$. Tome um ponto $A \in r$ e um vetor $\vec{v} \neq \vec{0}$ paralelo[1] a r. Então, dado $P \in \mathbb{E}^3$, temos que

$$\begin{aligned} P \in r \quad &\Longleftrightarrow \quad \{\overrightarrow{AP}, \vec{v}\} \text{ é LD} \\ &\Longleftrightarrow \quad \text{existe } \lambda \in \mathbb{R} \text{ tal que } \overrightarrow{AP} = \lambda \vec{v} \ \ (\text{pois } \vec{v} \neq \vec{0}) \\ &\Longleftrightarrow \quad \text{existe } \lambda \in \mathbb{R} \text{ tal que } P = A + \lambda \vec{v}. \end{aligned}$$

A figura a seguir ilustra a situação descrita.



A equação

$$r: \quad X = A + \lambda \vec{v} \quad (\lambda \in \mathbb{R})$$

chama-se *equação vetorial* da reta r, e $\vec{v}$ é dito um *vetor diretor* de r.

O que acabamos de ver é que um ponto $P \in \mathbb{E}^3$ está na reta r se, e somente se, P satisfizer a equação vetorial de r, ou seja, se, e somente se, existir $\lambda \in \mathbb{R}$ tal que $P = A + \lambda \vec{v}$.

Observação Uma maneira de se pensar na equação de uma reta é a de que ela dá uma trajetória sobre a reta. Por exemplo, podemos pensar em

$$r: \quad X = A + \lambda \vec{v} \quad (\lambda \in \mathbb{R})$$

como uma trajetória sobre r que se inicia no ponto A — a origem da trajetória — e percorre r com velocidade $\vec{v}$: no instante λ está-se no ponto $A + \lambda \vec{v}$.

É claro que se $\vec{w} \in \mathbb{V}^3$ é um outro vetor não nulo paralelo à reta r e B é um outro ponto de r, então

$$r: \quad X = B + \mu \vec{w} \quad (\mu \in \mathbb{R})$$

é também uma equação vetorial de r. Aliás, se se conhecem dois pontos distintos A e B sobre uma reta r, então $\overrightarrow{AB}$ é um vetor diretor de r e

[1] O vetor $\vec{v}$ é paralelo à reta r se existirem $M, N \in \mathbb{E}^3$ tais que $\vec{v} = \overrightarrow{MN}$ e o segmento MN for paralelo à reta r

$$r : \quad X = A + \lambda\overrightarrow{AB} \quad (\lambda \in \mathbb{R})$$

é uma equação vetorial de r.

Equações paramétricas da reta

Suponha, agora, que $\Sigma = (O, \mathcal{B})$ seja um sistema de coordenadas em $\mathbb{E}^3$, e considere a reta r em $\mathbb{E}^3$ que passa pelo ponto $A = (x_0, y_0, z_0)_\Sigma$ e tem vetor diretor $\vec{v} = (a, b, c)_\mathcal{B}$. Então, dado um ponto $P = (x, y, z)_\Sigma \in \mathbb{E}^3$, conforme vimos acima, sabemos que $P \in r$ se, e somente se, existir $\lambda \in \mathbb{R}$ tal que $P = A + \lambda\vec{v}$, o que, por sua vez, em vista da Proposição 3.1.3 é equivalente a dizer que existe $\lambda \in \mathbb{R}$ tal que

$$x = x_0 + \lambda a, \quad y = y_0 + \lambda b, \quad z = z_0 + \lambda c.$$

As equações

$$\begin{cases} x = x_0 + \lambda a \\ y = y_0 + \lambda b \\ z = z_0 + \lambda c \end{cases} \qquad (\lambda \in \mathbb{R})$$

chamam-se *equações paramétricas* da reta que passa pelo ponto $A = (x_0, y_0, z_0)_\Sigma$ e é paralela ao vetor $\vec{v} = (a, b, c)_\mathcal{B}$.

Observe que, como $\vec{v}$ é um vetor diretor de r, em particular, $\vec{v} \neq \vec{0}$, o que implica $a^2 + b^2 + c^2 \neq 0$. No caso particular em que $a \neq 0$, $b \neq 0$ e $c \neq 0$, podemos resolver para λ e obter

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Essas duas igualdades chamam-se *equações de r na forma simétrica*.

Vejamos alguns exemplos de como descrever uma reta e suas propriedades a partir de equações que a definem, e vice-versa, como encontrar equações que definem uma reta satisfazendo determinadas propriedades.

Nos exemplos que se seguem, considera-se fixado um sistema de coordenadas em $\mathbb{E}^3$. Coordenadas de pontos são relativas a esse sistema de coordenadas, e coordenadas de vetores são relativas à base do sistema de coordenadas.

Exemplo 3.2.1 Encontre uma equação vetorial, equações paramétricas e, se existirem, equações na forma simétrica para a reta r que passa pelos pontos $A = (1, 0, 1)$ e $B = (0, 1, 0)$.

Solução: Sabemos que $\overrightarrow{AB} = (-1, 1, -1)$ é um vetor diretor de r. Assim, uma equação vetorial para r é

$$X = (1, 0, 1) + \lambda(-1, 1, -1) \quad (\lambda \in \mathbb{R}).$$

Equações paramétricas para r são:

$$\begin{cases} x = 1 - \lambda \\ y = \lambda \\ z = 1 - \lambda \end{cases} \qquad (\lambda \in \mathbb{R})$$

e equações na forma simétrica:

$$\frac{x - 1}{-1} = y = \frac{z - 1}{-1}.$$

Podemos utilizar as equações na forma simétrica, por exemplo, para verificar se um determinado ponto de $\mathbb{E}^3$ está ou não em r. Assim, $P = (2, -1, 2) \in r$, uma vez que suas coordenadas satisfazem as equações na forma simétrica, mas $Q = (2, 1, 0) \notin r$. $\diamond$

Exemplo 3.2.2 Determine uma equação vetorial da reta r que passa pelo ponto médio do segmento AB, em que $A = (1,1,3)$ e $B = (3,1,0)$, e tem vetor diretor $\vec{v} = \left(\frac{\sqrt{3}}{49}, \frac{3\sqrt{3}}{98}, -\frac{\sqrt{3}}{7}\right)$.

Solução: Seja M o ponto médio do segmento AB. Como vimos, $M = \left(\frac{1+3}{2}, \frac{1+1}{2}, \frac{3+0}{2}\right) = \left(2, 1, \frac{3}{2}\right)$. Agora, $\vec{u} = (2,3,-14) = \frac{98}{\sqrt{3}}\vec{v}$ é paralelo a $\vec{v}$ e, portanto, também é diretor de r. Logo, uma equação vetorial para r é

$$X = \left(2, 1, \frac{3}{2}\right) + \lambda(2,3,-14) \quad (\lambda \in \mathbb{R}). \quad \diamond$$

Exemplo 3.2.3 Encontre uma equação vetorial da reta r definida pelas equações paramétricas

$$\begin{cases} x = 1 + 3\lambda \\ y = 2\lambda \\ z = 6 - 5\lambda \end{cases} \qquad (\lambda \in \mathbb{R}).$$

Solução: Vamos reescrever as equações paramétricas de r explicitando seus componentes:

$$\begin{cases} x = 1 + \lambda 3 \\ y = 0 + \lambda 2 \\ z = 6 + \lambda(-5) \end{cases} \qquad (\lambda \in \mathbb{R}).$$

Assim, a reta r passa pelo ponto $A = (1,0,6)$ e tem vetor diretor $\vec{v} = (3,2,-5)$. Logo, uma equação vetorial de r é

$$X = (1,0,6) + \lambda(3,2,-5) \quad (\lambda \in \mathbb{R}). \quad \diamond$$

Exemplo 3.2.4 Verifique se o ponto $P = (4,1,-1)$ pertence à reta $r : \ X = (1,0,1) + \lambda(2,1,1)$ $(\lambda \in \mathbb{R})$.

Solução: É preciso decidir se existe $\lambda \in \mathbb{R}$ tal que $P = (1,0,1) + \lambda(2,1,1)$. Para tanto, λ deve satisfazer $(4,1,-1) = (1+2\lambda, \lambda, 1+\lambda)$. Ou seja, é preciso decidir se existe $\lambda \in \mathbb{R}$ tal que

$$\begin{cases} 4 = 1 + 2\lambda \\ 1 = \lambda \\ -1 = 1 + \lambda \end{cases}$$

Mas é claro que nenhum λ satisfaz essas três equações. Logo, $P \notin r$. $\diamond$

Exemplo 3.2.5 Exiba um ponto e um vetor diretor da reta definida por

$$\frac{2x-1}{3} = \frac{1-y}{2} = z + 1. \tag{3.1}$$

Solução: Comecemos por reescrever as equações acima a fim de identificá-las, de fato, com as equações de uma reta. Temos

$$\begin{aligned} \frac{2x-1}{3} &= \frac{2(x-\frac{1}{2})}{3} = \frac{x-\frac{1}{2}}{\frac{3}{2}}, \\ \frac{1-y}{2} &= \frac{-(y-1)}{2} = \frac{y-1}{-2}, \\ z+1 &= \frac{z-(-1)}{1}. \end{aligned}$$

Assim, as equações (3.1) podem ser escritas na seguinte forma, que mostra se tratarem de equações na forma simétrica de uma reta, em que seus componentes foram explicitados:

$$\frac{x - \frac{1}{2}}{\frac{3}{2}} = \frac{y - 1}{-2} = \frac{z - (-1)}{1}.$$

Ou seja, (3.1) definem uma reta que passa pelo ponto $A = (\frac{1}{2}, 1, -1)$ e tem vetor diretor $\vec{v} = (\frac{3}{2}, -2, 1)$. $\diamond$

Exercícios Lista 2 - Álgebra Linear I: Exs. 28–30.

3.3 Planos

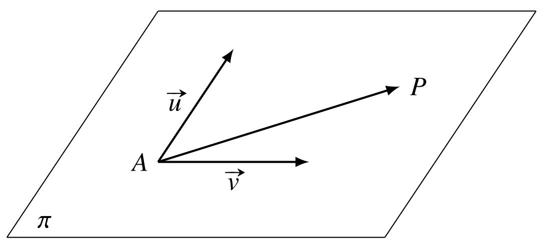
Nesta seção, assumiremos como conhecido o conceito de plano no espaço tridimensional $\mathbb{E}^3$.

Equação vetorial do plano

Seja π um plano em $\mathbb{E}^3$. Tome um ponto $A \in \pi$ e um par de vetores $\vec{u}, \vec{v} \in \mathbb{V}^3$ tais que $\vec{u}$ e $\vec{v}$ sejam paralelos[2] a π e $\{\vec{u}, \vec{v}\}$ seja LI. Então, dado $P \in \mathbb{E}^3$, temos que

$$\begin{aligned} P \in \pi \quad &\iff \quad \left\{\overrightarrow{AP}, \vec{u}, \vec{v}\right\} \text{ é LD} \\ &\iff \quad \text{existem } \lambda, \mu \in \mathbb{R} \text{ tais que } \overrightarrow{AP} = \lambda\vec{u} + \mu\vec{v} \ \ (\text{pois } \{\vec{u}, \vec{v}\} \text{ é LI}) \\ &\iff \quad \text{existem } \lambda, \mu \in \mathbb{R} \text{ tais que } P = A + \lambda\vec{u} + \mu\vec{v}. \end{aligned}$$

A figura a seguir ilustra a situação descrita.



A equação

$$r: \quad X = A + \lambda\vec{u} + \mu\vec{v} \quad (\lambda, \mu \in \mathbb{R})$$

chama-se *equação vetorial* do plano π, e $\vec{u}, \vec{v}$ são chamados *vetores diretores* de π.

Vimos que um ponto $P \in \mathbb{E}^3$ está no plano π se, e somente se, P satisfizer a equação vetorial de π, ou seja, se, e somente se, existirem $\lambda, \mu \in \mathbb{R}$ tais que $P = A + \lambda\vec{u} + \mu\vec{v}$.

Equações paramétricas do plano

Suponha, agora, que $\Sigma = (O, \mathcal{B})$ seja um sistema de coordenadas em $\mathbb{E}^3$, e considere o plano π em $\mathbb{E}^3$ que contém o ponto $A = (x_0, y_0, z_0)_\Sigma$ e tem vetores diretores $\vec{u} = (r, s, t)_\mathcal{B}$, $\vec{v} = (m, n, p)_\mathcal{B}$. Então, dado um ponto $P = (x, y, z)_\Sigma \in \mathbb{E}^3$, conforme vimos acima, sabemos que $P \in \pi$ se, e somente se, existirem $\lambda, \mu \in \mathbb{R}$ tais que $P = A + \lambda\vec{u} + \mu\vec{v}$, o que, por sua vez, é equivalente a

$$x = x_0 + \lambda r + \mu m, \quad y = y_0 + \lambda s + \mu n, \quad z = z_0 + \lambda t + \mu p.$$

As equações

[2] O vetor $\vec{u}$ é paralelo ao plano π se existirem $M, N \in \mathbb{E}^3$ tais que $\vec{u} = \overrightarrow{MN}$ e o segmento MN for paralelo ao plano π.

$$\begin{cases} x = x_0 + \lambda r + \mu m \\ y = y_0 + \lambda s + \mu n \\ z = z_0 + \lambda t + \mu p \end{cases} \qquad (\lambda, \mu \in \mathbb{R})$$

chamam-se *equações paramétricas* do plano que contém o ponto $A = (x_0, y_0, z_0)_\Sigma$ e é paralelo aos vetores $\vec{u} = (r, s, t)_\mathcal{B}$, $\vec{v} = (m, n, p)_\mathcal{B}$.

Observe que como $\vec{u}, \vec{v}$ são diretores de π, em particular $\{\vec{u}, \vec{v}\}$ é LI. Assim, $\vec{u} \wedge \vec{v} \neq \vec{0}$, qualquer que seja a orientação que se coloque em $\mathbb{V}^3$.

Nos exemplos a seguir, coordenadas de pontos são relativas a um sistema de coordenadas em $\mathbb{E}^3$ fixo e coordenadas de vetores são relativas à base desse sistema de coordenadas.

Exemplo 3.3.1 Encontre uma equação vetorial e equações paramétricas do plano que contém os pontos $A = (0, 1, 0)$, $B = (1, 0, 1)$ e $C = (0, 0, 1)$.

Solução: Os vetores $\overrightarrow{AB} = (1, -1, 1)$ e $\overrightarrow{AC} = (0, -1, 1)$ são paralelos ao plano e $\{\overrightarrow{AB}, \overrightarrow{AC}\}$ é LI. Logo, eles são diretores do plano. Assim, uma equação vetorial do plano é

$$X = (0, 1, 0) + \lambda(1, -1, 1) + \mu(0, -1, 1) \quad (\lambda, \mu \in \mathbb{R}),$$

e equações paramétricas são

$$\begin{cases} x = \lambda \\ y = 1 - \lambda - \mu \\ z = \lambda + \mu \end{cases} \qquad (\lambda, \mu \in \mathbb{R}) \quad \diamond$$

Equação geral do plano

Seja $\Sigma = (O, \mathcal{B})$ um sistema de coordenadas em $\mathbb{E}^3$ e seja π o plano de $\mathbb{E}^3$ que contém o ponto $A = (x_0, y_0, z_0)_\Sigma$ e tem vetores diretores $\vec{u} = (r, s, t)_\mathcal{B}$, $\vec{v} = (m, n, p)_\mathcal{B}$. Vimos que dado um ponto $P = (x_1, y_1, z_1)_\Sigma \in \mathbb{E}^3$, temos que $P \in \pi$ se, somente se, o conjunto $\{\overrightarrow{AP}, \vec{u}, \vec{v}\}$ for LD. Como $\overrightarrow{AP} = (x_1 - x_0, y_1 - y_0, z_1 - z_0)_\mathcal{B}$, segue, da Proposição 2.3.4, que $P \in \pi$ se, e somente se, $\det \begin{bmatrix} x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ r & s & t \\ m & n & p \end{bmatrix} = 0$. Desenvolvendo o determinante, por expansão em cofatores ao longo da primeira linha, por exemplo, obtemos que $P \in \pi$ se, e somente se, $ax_1 + by_1 + cz_1 + d = 0$, em que $a = \det \begin{bmatrix} s & t \\ n & p \end{bmatrix} = sp - tn$, $b = \det \begin{bmatrix} t & r \\ p & m \end{bmatrix} = tm - rp$, $c = \det \begin{bmatrix} r & s \\ m & n \end{bmatrix} = rn - sm$ e $d = -(ax_0 + by_0 + cz_0)$.

A equação

$$ax + by + cz + d = 0$$

chama-se *equação geral* do plano π. Vimos que um ponto de $\mathbb{E}^3$ está no plano π se, e somente se, suas coordenadas satisfizerem a equação geral de π. Note que, na equação geral do plano π, $a^2 + b^2 + c^2 \neq 0$. Isso segue da Proposição 2.3.6, uma vez que $\{\vec{u}, \vec{v}\}$ é LI.

Observação Existe uma espécie de recíproca do fato de, fixado um sistema de coordenadas, todo plano ter uma equação geral. Mais precisamente, seja $\Sigma = (O, \mathcal{B})$ um sistema de coordenadas em $\mathbb{E}^3$ e sejam $a, b, c, d \in \mathbb{R}$ tais que $a^2 + b^2 + c^2 \neq 0$. Então, existe um plano π em $\mathbb{E}^3$ que tem

$$ax + by + cz + d = 0 \tag{3.2}$$

como equação geral.

Com efeito, da condição $a^2 + b^2 + c^2 \neq 0$ segue que os escalares a, b, c não podem ser os três iguais a zero. Suponha, por exemplo, que $a \neq 0$. Na equação (3.2), fazendo

- $y = z = 0$, obtemos $x = -\frac{d}{a}$;

- $y = 0$ e $z = 1$, obtemos $x = -\frac{c+d}{a}$;
- $y = 1$ e $z = 0$, obtemos $x = -\frac{b+d}{a}$.

Considere os pontos $A = (-\frac{d}{a}, 0, 0)_\Sigma$, $B = (-\frac{c+d}{a}, 0, 1)_\Sigma$ e $C = (-\frac{b+d}{a}, 1, 0)_\Sigma$. Então os vetores

$$\overrightarrow{AB} = \left(-\frac{c}{a}, 0, 1\right)_{\mathcal{B}} \quad \text{e} \quad \overrightarrow{AC} = \left(-\frac{b}{a}, 1, 0\right)_{\mathcal{B}}$$

são LI e, portanto, existe um único plano π contendo os pontos A, B, C. Esse plano tem $(-a)\overrightarrow{AB}, \overrightarrow{AC}$ como vetores diretores e, assim, tem equação geral dada por

$$\det\begin{bmatrix} x - (-\frac{d}{a}) & y - 0 & z - 0 \\ c & 0 & -a \\ -\frac{b}{a} & 1 & 0 \end{bmatrix} = 0.$$

Desenvolvendo o determinante, obtemos $ax + by + cz + d = 0$. O argumento é análogo se $b \neq 0$ ou $c \neq 0$.

Para os exemplos, suponha fixado um sistema de coordenadas em $\mathbb{E}^3$.

Exemplo 3.3.2 Encontre uma equação geral do plano que contém o ponto $A = (9, -1, 0)$ e é paralelo aos vetores $\vec{u} = (0, 1, 0)$ e $\vec{v} = (1, 1, 1)$.

Solução: Comecemos por observar que, de fato, $\{\vec{u}, \vec{v}\}$ é LI e, portanto, definem com o ponto A um único plano. Uma equação geral desse plano é

$$\det\begin{bmatrix} x - 9 & y - (-1) & z - 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = 0,$$

isto é, $x - z - 9 = 0$. $\diamond$

Exemplo 3.3.3 Encontre uma equação geral do plano que passa pelos pontos $A = (1, 0, 1)$, $B = (-1, 0, 1)$ e $C = (2, 1, 2)$.

Solução: Os vetores $\overrightarrow{AB} = (-2, 0, 0)$ e $\overrightarrow{AC} = (1, 1, 1)$ são paralelos ao plano e $\{\overrightarrow{AB}, \overrightarrow{AC}\}$ é LI. Logo, uma equação geral desse plano é

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

ou, $2y - 2z + 2 = 0$. Como as soluções dessa equação coincidem com as soluções de $y - z + 1 = 0$, essa última equação também é uma equação geral desse mesmo plano. $\diamond$

Exemplo 3.3.4 Encontre uma equação geral do plano definido pelas equações paramétricas

$$\begin{cases} x = -1 + 2\lambda - 3\mu \\ y = 1 + \lambda + \mu \\ z = \lambda \end{cases} \qquad (\lambda, \mu \in \mathbb{R})$$

Solução: Das equações paramétricas, obtemos que o plano passa pelo ponto $A = (-1, 1, 0)$ e tem vetores diretores $\vec{u} = (2, 1, 1)$ e $\vec{v} = (-3, 1, 0)$. (Observe que $\{\vec{u}, \vec{v}\}$ é, de fato, LI.) Assim, uma equação geral desse plano é

$$0 = \det\begin{bmatrix} x - (-1) & y - 1 & z - 0 \\ 2 & 1 & 1 \\ -3 & 1 & 0 \end{bmatrix} = -x - 3y + 5z + 2,$$

ou, equivalentemente, $x + 3y - 5z - 2 = 0$. $\diamond$

Exemplo 3.3.5 Encontre uma equação vetorial para o plano definido pela equação geral

$$x + 2y - z - 1 = 0. \tag{3.3}$$

Solução: Basta encontrarmos três pontos do plano que não sejam colineares, ou seja, busquemos três soluções de (3.3) de modo que os pontos cujas coordenadas sejam essas soluções não estejam sobre uma mesma reta. Façamos, em (3.3), a seguinte tentativa:

- $x = y = 0$ e, portanto, $z = -1$;
- $x = z = 0$ e, portanto, $y = \frac{1}{2}$;
- $y = z = 0$ e, portanto, $x = 1$.

Temos, então, três pontos no plano: $A = (0, 0, -1)$, $B = (0, \frac{1}{2}, 0)$, $C = (1, 0, 0)$. Consideremos os vetores $\overrightarrow{AB} = (0, \frac{1}{2}, 1)$ e $\overrightarrow{AC} = (1, 0, 1)$. Como $2\overrightarrow{AB}$ e $\overrightarrow{AC}$ são paralelos ao plano e $\{2\overrightarrow{AB}, \overrightarrow{AC}\}$ é LI,

$$X = (0, 0, -1) + \lambda(0, 1, 2) + \mu(1, 0, 1) \quad (\lambda, \mu \in \mathbb{R})$$

é uma equação vetorial para esse plano.

Se as soluções de (3.3) que encontramos tivessem dado origem a três pontos colineares, bastaria trocar uma delas por uma outra que evitasse a colinearidade. Como o número de soluções de (3.3) é infinito, essa tarefa não seria difícil. $\diamond$

Exemplo 3.3.6 Encontre equações paramétricas para a reta r determinada pela interseção dos planos

$$\pi_1 : x + y + z - 1 = 0 \qquad \text{e} \qquad \pi_2 : x + y - z = 0.$$

Solução: A interseção de dois planos é uma reta precisamente quando esses planos não são paralelos. Veremos, adiante, nessas notas, como verificar se dois planos são paralelos ou não a partir de equações gerais deles (cf. Exercício 3.3.9). Por ora, aceitemos, neste exemplo, que π_1 e π_2 não são paralelos. Para determinar a reta dada pela interseção deles, basta encontrar dois pontos distintos sobre ela. As coordenadas dos pontos de r são precisamente as soluções do sistema linear

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

No Capítulo 1, vimos um método que permite encontrar *todas* as soluções desse sistema. Aqui, bastam duas. Então, vejamos, subtraindo a segunda equação da primeira, obtemos $z = \frac{1}{2}$. Substituindo esse valor na segunda equação fornece $x + y = \frac{1}{2}$. Tomando $x = 0$, obtemos a solução $(0, \frac{1}{2}, \frac{1}{2})$; tomando $y = 0$, obtemos a solução $(\frac{1}{2}, 0, \frac{1}{2})$. Logo, r é a reta que passa pelos pontos $A = (0, \frac{1}{2}, \frac{1}{2})$ e $B = (\frac{1}{2}, 0, \frac{1}{2})$. Como o vetor $\overrightarrow{AB} = (\frac{1}{2}, -\frac{1}{2}, 0)$ é paralelo a r, o vetor $2\overrightarrow{AB} = (1, -1, 0)$ é diretor de r. Logo,

$$\begin{cases} x = \lambda \\ y = \frac{1}{2} - \lambda \\ z = \frac{1}{2} \end{cases} \qquad (\lambda \in \mathbb{R})$$

são equações paramétricas para r. $\diamond$

Vetor normal a um plano

Veremos que, no caso de um sistema ortogonal de coordenadas, os coeficientes que aparecem na equação normal de um plano têm uma interpretação geométrica precisa.

Definição Seja π um plano em $\mathbb{E}^3$. Um vetor $\vec{n} \in \mathbb{V}^3$ é dito *normal* ao plano π se $\vec{n} \neq \vec{0}$ e $\vec{n}$ é ortogonal a todo vetor paralelo a π.

Suponha fixado um sistema ortogonal de coordenadas $\Sigma = (O, \mathcal{B})$ em $\mathbb{E}^3$, e seja π um plano em $\mathbb{E}^3$. Seja $A = (x_0, y_0, z_0)_\Sigma$ um ponto de π e seja $\vec{n} = (a, b, c)_\mathcal{B}$ um vetor normal a π. Dado um ponto $P = (x_1, y_1, z_1)_\mathcal{B}$ de $\mathbb{E}^3$, então $P \in \pi$ se, e somente se $\vec{n}$ e $\overrightarrow{AP}$ são ortogonais, isto é, se, e somente se, $\vec{n} \cdot \overrightarrow{AP} = 0$. Como a base $\mathcal{B}$ de $\mathbb{V}^3$ é ortonormal, essa última igualdade é equivalente a $a(x_1 - x_0) + b(y_1 - y_0) + c(z_1 - z_0) = 0$, ou ainda,

$$ax_1 + by_1 + cz_1 + d = 0$$

em que $d = -(ax_0 + by + 0 + cz_0)$. Ou seja,

$$ax + by + cz + d = 0$$

é uma equação geral de π.

Reciprocamente, mostremos que se $ax + by + cz + d = 0$ é uma equação geral de π, então $\vec{n} = (a, b, c)_\mathcal{B}$ é um vetor normal a π. De fato, seja $\vec{v}$ um vetor paralelo a π e sejam $A = (x_1, y_1, z_1)_\Sigma$ e $B = (x_2, y_2, z_2)_\Sigma$ pontos de π tais que $\vec{v} = \overrightarrow{AB}$. Então, $\vec{v} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)_\mathcal{B}$ e, como $\mathcal{B}$ é ortonormal, temos

$$\vec{n} \cdot \vec{v} = a(x_2 - x_1) + b(y_2 - y_1) + c(z_2 - z_1) = (ax_2 + by_2 + cz_2) - (ax_1 + by_1 + cz_1) = d - d = 0,$$

uma vez que, como tanto A como B estão em π, suas coordenadas satisfazem a equação geral $ax + by + cz + d = 0$ de π.

Nos exemplos que se seguem, está fixado um sistema ortogonal de coordenadas em $\mathbb{E}^3$.

Exemplo 3.3.7 Obtenha uma equação geral do plano π que passa pelo ponto $A = (1, 0, 2)$ e tem vetor normal $\vec{n} = (1, -1, 4)$.

Solução: Sabemos que π tem uma equação geral da forma $x - y + 4z + d = 0$. Falta encontrar d. Como $A \in \pi$, suas coordenadas satisfazem a equação geral. Assim, $1 - 0 + 8 + d = 0$, e, portanto, $d = -9$. Logo, $x - y + 4z - 9 = 0$ é uma equação geral de π. $\diamond$

Exemplo 3.3.8 Obtenha uma equação geral do plano que passa pelo ponto $A = (0, 1, 2)$ e tem vetores diretores $\vec{u} = (4, 1, 2)$, $\vec{v} = (2, 1, -2)$.

Solução: Comecemos por observar que, de fato, $\{\vec{u}, \vec{v}\}$ é LI. Um vetor normal a π é qualquer vetor que seja ortogonal a $\vec{u}$ e a $\vec{v}$. Se $\mathbb{V}^3$ estivesse orientado, poderíamos tomar o produto vetorial $\vec{u} \wedge \vec{v}$ como vetor normal a π. Fixemos, então, uma orientação em $\mathbb{V}^3$. Se a base ortonormal $\mathcal{B} = \{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ de $\mathbb{V}^3$ que compõe o sistema de coordenadas for positiva, saberemos encontrar as coordenadas de $\vec{u} \wedge \vec{v}$ em termos das coordenadas de $\vec{u}$ e de $\vec{v}$ (usando a Proposição 2.7.1). Se $\mathcal{B}$ for uma base negativa, basta inverter o sinal das coordenadas do lado direito de (2.11) para obter as coordenadas do produto vetorial $\vec{u} \wedge \vec{v}$. Em qualquer caso.

$$\det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 4 & 1 & 2 \\ 2 & 1 & -2 \end{bmatrix} = (-4, 12, 2)_\mathcal{B}$$

é um vetor normal a π (pois ele é ortogonal a $\vec{u}$ e a $\vec{v}$, uma vez que ou ele é o igual a $\vec{u} \wedge \vec{v}$ ou igual a $-(\vec{u} \wedge \vec{v})$). Portanto, $\vec{n} = -\frac{1}{2}(-4, 12, 2) = (2, -6, -1)$ também é um vetor normal a π. Logo, π tem uma equação geral da forma $2x - 6y - z + d = 0$. Substituindo as coordenadas de A nela, obtemos $d = 8$. Logo, uma equação geral de π é $2x - 6y - z + 8 = 0$. $\diamond$

Exemplo 3.3.9 Encontre uma equação vetorial para a reta r dada pela interseção dos planos

$$\pi_1 : \ 2x - y - 3 = 0 \qquad \text{e} \qquad \pi_2 : \ 3x + y + 2z - 1 = 0.$$

Solução: Sabemos que $\vec{n_1} = (2, -1, 0)$ é normal a π_1 e que $\vec{n_2} = (3, 1, 2)$ é normal a π_2. Como $\vec{n_1}$ e $\vec{n_2}$ não são paralelos, os planos π_1 e π_2 não são paralelos, e, como consequência, $\pi_1 \cap \pi_2$ é, de fato, uma reta. Agora, um vetor não nulo $\vec{v}$ será diretor de r se for paralelo a r, isto é, se for paralelo a π_1 e a π_2. Mas isso só ocorre se $\vec{v}$ for ortogonal a $\vec{n_1}$ e a $\vec{n_2}$. Como vimos no exemplo acima, o vetor

$$\det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 2 & -1 & 0 \\ 3 & 1 & 2 \end{bmatrix} = (-2, -4, 5)$$

é ortogonal a $\vec{n_1}$ e a $\vec{n_2}$ (ele é $\pm(\vec{n_1} \wedge \vec{n_2})$, dependendo da orientação de $\mathbb{V}^3$.). Assim, podemos tomá-lo como diretor de r. Precisamos, por fim, de um ponto em r, ou seja, um ponto que esteja simultaneamente em π_1 e π_2. Para tanto, basta tomar um ponto cujas coordenadas são uma solução do sistema linear

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

Subtituindo $y = 2x - 3$, obtida a partir da primeira equação, na segunda, obtemos $5x + 2z = 4$. Fazendo $x = 0$ nessa última equação, obtemos a solução $(0, -3, 2)$. Logo,

$$X = (0, -3, 2) + \lambda(-2, -4, 5) \quad (\lambda \in \mathbb{R})$$

é uma equação vetorial para r. $\diamond$

Exercícios Lista 2 - Álgebra Linear I: Exs. 31–35.

3.4 Posições relativas

Nesta seção, veremos como os intrumentos vetoriais que introduzimos no capítulo anterior auxiliam no estudo de posições relativas de retas e planos.

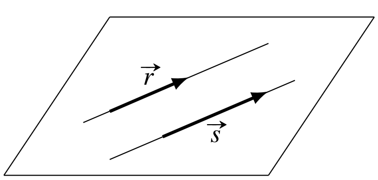
Posição relativa de retas

Lembremos que duas retas em $\mathbb{E}^3$ são ditas *paralelas* se ou são coincidentes ou são coplanares e não têm ponto em comum; são ditas *concorrentes* se têm um único ponto em comum (o que é equivalente a serem coplanares mas não paralelas); e são ditas *reversas* se não são coplanares.

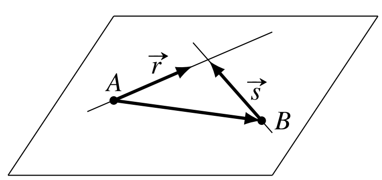
Fica, portanto, claro que dada uma reta r, passando pelo ponto A com vetor diretor $\vec{r}$, e dada uma reta s, passando pelo ponto B com vetor diretor $\vec{s}$, temos:

- r e s são paralelas se, e somente se, $\{\vec{r}, \vec{s}\}$ é LD.
- r e s são reversas se, e somente se, $\{\vec{r}, \vec{s}, \overrightarrow{AB}\}$ é LI.
- r e s são concorrentes se, e somente se, $\{\vec{r}, \vec{s}, \overrightarrow{AB}\}$ é LD e $\{\vec{r}, \vec{s}\}$ é LI.

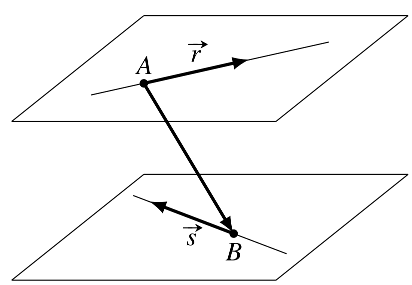
As figuras ilustram as três situações.



retas paralelas



retas concorrentes



retas reversas

Vejamos alguns exemplos, em que se entende fixado um sistema de coordenadas em $\mathbb{E}^3$ (não necessariamente ortogonal).

Exemplo 3.4.1 Estude a posição relativa das retas

$$r: \ X = (1,2,3) + \lambda(0,1,3) \quad (\lambda \in \mathbb{R}) \qquad \text{e} \qquad s: \ X = (0,1,0) + \mu(1,1,1) \quad (\mu \in \mathbb{R}).$$

Solução: O vetor $\vec{r} = (0,1,3)$ é diretor de r, e o vetor $\vec{s} = (1,1,1)$, de s. Como $\{\vec{r}, \vec{s}\}$ é LI, as retas r e s não são paralelas. Para decidir se são concorrentes ou coplanares, considere os pontos $A = (1,2,3) \in r$ e $B = (0,1,0) \in s$ e o vetor $\overrightarrow{AB} = (-1,-1,3)$. Temos $\det \begin{bmatrix} 0 & 1 & 3 \\ 1 & 1 & 1 \\ -1 & -1 & -3 \end{bmatrix} = 2 \neq 0$. Segue que $\{\vec{r}, \vec{s}, \overrightarrow{AB}\}$ é LI. Portanto, r e s são reversas. $\diamond$

Exemplo 3.4.2 Estude a posição relativa das retas

$$r: \ X = (1,2,3) + \lambda(0,1,3) \quad (\lambda \in \mathbb{R}) \qquad \text{e} \qquad s: \ X = (1,3,6) + \mu(0,2,6) \quad (\mu \in \mathbb{R}).$$

Solução: O vetor $\vec{r} = (0,1,3)$ é diretor de r, e o vetor $\vec{s} = (0,2,6)$, de s. Como $\{\vec{r}, \vec{s}\}$ é LD (pois $\vec{s} = 2\vec{r}$), as retas r e s são paralelas. Vejamos se são coincidentes ou não. Considere o ponto $A = (1,2,3) \in r$. Verifiquemos se $A \in s$, ou seja, se existe $\mu \in \mathbb{R}$ tal que $A = (1,3,6) + \mu(0,2,6)$. Essa igualdade é equivalente a

$$\begin{cases} 1 = 1 \\ 2 = 3 + 2\mu \\ 3 = 6 + 6\mu \end{cases}$$

que tem solução dada por $\mu = -\frac{1}{2}$. Logo, $A \in s$, o que faz de r e s reta coincidentes. (Duas retas paralelas são coincidentes precisamente quando têm um ponto em comum.) $\diamond$

Exemplo 3.4.3 Estude a posição relativa das retas

$$r: \ X = (1,2,3) + \lambda(0,1,3) \quad (\lambda \in \mathbb{R}) \qquad \text{e} \qquad s: \begin{cases} x + y + z - 6 = 0 \\ x - y - z + 4 = 0 \end{cases}$$

Solução: O vetor $\vec{r} = (0,1,3)$ é diretor de r. Fazendo $z = 0$ nas equações que definem s, obtemos o ponto $P = (1,5,0) \in s$ e, fazendo $z = 1$, obtemos o ponto $Q = (1,4,1) \in s$. Logo, $\vec{s} = \overrightarrow{PQ} = (0,-1,1)$ é um vetor diretor de s, e $\{\vec{r}, \vec{s}\}$ é LI. Temos os pontos $A = (1,2,3) \in r$ e $P = (1,5,0) \in s$, que definem o vetor $\overrightarrow{AP} = (0,3,-3)$. Como o conjunto $\{\vec{r}, \vec{s}, \overrightarrow{AP}\}$ é LD (pois $\det \begin{bmatrix} 0 & 1 & 3 \\ 0 & -1 & 1 \\ 0 & 3 & -3 \end{bmatrix} = 0$), as retas r e s são concorrentes. Para determinar o ponto de interseção das retas basta, por exemplo, substituir nas equações que definem s as coordenadas de um ponto genérico de r, que são da forma $(1, 2+\lambda, 3+3\lambda)$. Fazendo isso na primeira equação, obtemos $\lambda = 0$ (e, na segunda, também). Logo o ponto de coordenadas $(1,2,3)$ (que é o ponto A) é o ponto de interseção de r e s. $\diamond$

Posição relativa de reta e plano

Em $\mathbb{E}^3$, dada uma reta r e um plano π, há três possibilidades de posições relativas: ou r está contida em π, ou r e π não se interceptam — nesses dois primeiros casos, dizemos que r e π são *paralelos* —, ou a interseção de r e π é apenas um ponto — nesse caso dizemos que r e π são *transversais*.

Se $\vec{r}$ é um vetor diretor da reta r e $\vec{u}, \vec{v}$ são vetores diretores do plano π, fica claro que r e π são paralelos se, e somente se $\{\vec{r}, \vec{u}, \vec{v}\}$ é LD.

O resultado a seguir fornece um instrumento vetorial útil na análise de posições relativas de reta e plano, quando se tem a disposição uma equação geral do plano.

Proposição 3.4.4 *Seja $\Sigma = (O, \mathcal{B})$ um sistema de coordenadas em $\mathbb{E}^3$. Seja $\vec{w} = (m, n, p)_{\mathcal{B}}$ um vetor e seja π um plano em $\mathbb{E}^3$ de equação geral $ax + by + cz + d = 0$. Então, $\vec{w}$ é paralelo a π se, e somente se, $am + bn + cp = 0$.*

Demonstração Tome um ponto $A \in \pi$ e seja $B = A + \vec{w}$. Então, $\vec{w}$ é paralelo a π se, e somente se, $B \in \pi$. Agora, como $A \in \pi$, se $A = (x_0, y_0, z_0)_\Sigma$, então

$$ax_0 + by_0 + cz_0 = -d. \tag{3.4}$$

Como $B = (x_0 + m, y_0 + n, z_0 + p)_\Sigma$, temos que $B \in \pi$ se, e somente se, $a(x_0 + m) + b(y_0 + n) + c(z_0 + p) + d = 0$. Em vista de (3.4), essa igualdade ocorre se, e somente se, $am + bn + cp = 0$. □

Observação Se o sistema de coordenadas na proposição acima fosse ortogonal, o resultado seria imediato, uma vez que, nesse caso $\vec{n} = (a, b, c)_{\mathcal{B}}$ seria um vetor normal a π e, portanto, $\vec{w}$ seria paralelo a π se, e somente se, $\vec{w}$ e $\vec{n}$ fossem ortogonais. O que diz a Proposição 3.4.4 é que o cálculo que realizaríamos para verificar a ortogonalidade entre $\vec{w}$ e $\vec{n}$, no caso de sistema ortogonal, continua valendo como teste de paralelismo entre $\vec{w}$ e π, mesmo se o sistema não for ortogonal.

Como uma reta é paralalela a um plano se, e só se, qualquer vetor diretor da reta for paralelo ao plano, o critério acima pode ser aplicado no estudo da posição relativa entre reta e plano.

Nos exemplos a seguir, sempre estará implícito um sistema de coordenadas não necessariamente ortogonal fixado.

Exemplo 3.4.5 Estude a posição relativa da reta r e plano π, em que

$$r:\ X = (1,1,1) + \alpha(3,2,1) \quad (\alpha \in \mathbb{R}) \qquad \text{e} \qquad \pi:\ X = (1,1,3) + \lambda(1,-1,1) + \mu(0,1,3) \quad (\lambda, \mu \in \mathbb{R}).$$

Solução: Sabemos que uma equação geral de π é dada por

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

Expandindo o determinante, obtemos $\pi : 4x + 3y - z - 4 = 0$. Como $\vec{r} = (3, 2, 1)$ é um vetor diretor de r, o critério da Proposição 3.4.4 nos garante que $\vec{r}$ não é paralelo a π, uma vez que $4 \cdot 3 + 3 \cdot 2 + (-1) \cdot 1 = 17 \neq 0$. Logo, r e π são transversais. Para encontrar o ponto de interseção entre r e π basta substituir as coordenadas de um ponto genérico de r, $(1 + 3\alpha, 1 + 2\alpha, 1 + \alpha)$, na equação geral de π: $4(1 + 3\alpha) + 3(1 + 2\alpha) - (1 + \alpha) - 4 = 0$. Essa equação nos fornece $\alpha = -\frac{2}{17}$. Logo, $r \cap \pi = \left\{\left(\frac{11}{17}, \frac{13}{17}, \frac{15}{17}\right)\right\}$. ◇

Exemplo 3.4.6 Estude a posição relativa da reta r e plano π, em que

$$r:\ X = (2,2,1) + \alpha(3,3,0) \quad (\alpha \in \mathbb{R}) \qquad \text{e} \qquad \pi:\ X = (1,0,1) + \lambda(1,1,1) + \mu(0,0,3) \quad (\lambda, \mu \in \mathbb{R}).$$

Solução: Um vetor diretor de r é $\vec{r} = (3, 3, 0)$ e vetores diretores de π são $\vec{u} = (1, 1, 1)$, $\vec{v} = (0, 0, 3)$. Como $\det \begin{bmatrix} 3 & 3 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix} = 0$, segue que r e π são paralelos (pois o determinante sendo nulo garante que $\{\vec{r}, \vec{u}, \vec{v}\}$ é LD). Para decidir se r está contida em π ou não, tomemos o ponto $A = (2, 2, 1) \in r$. Se A for um ponto de π, como r e π são paralelos, isso implicará que r está contida em π, caso contrário, teremos $r \cap \pi = \emptyset$. Vejamos, então, se existem $\lambda, \mu \in \mathbb{R}$ tais que $A = (1, 0, 1) + \lambda(1, 1, 1) + \mu(0, 0, 3)$, isto é, λ, μ que satisfazem

$$\begin{cases} 2 = 1 + \lambda \\ 2 = \lambda \\ 1 = 1 + \lambda + 3\mu \end{cases}$$

É claro que esse sistema não tem solução. Portanto, $A \notin \pi$, donde se conclui que $r \cap \pi = \emptyset$. $\diamond$

Exemplo 3.4.7 Estude a posição relativa da reta r e plano π, em que

$$r: \begin{cases} x = 1 + \lambda \\ y = 1 - \lambda \\ z = \lambda \end{cases} \quad (\lambda \in \mathbb{R}) \qquad \text{e} \qquad \pi: \ x + y - 2 = 0$$

Solução: Um vetor diretor de r é $\vec{r} = (1, -1, 1)$. Como $1 \cdot 1 + 1 \cdot (-1) + 0 \cdot 1 = 0$, r e π são paralelos. O ponto $A = (1, 1, 0)$ está em r e suas coordenadas satisfazem a equação geral de π. Logo, r está contida em π. $\diamond$

Posição relativa de planos

Dois planos em $\mathbb{E}^3$ são ditos *transversais* se interceptarem-se em uma reta; caso contrário, são ditos *paralelos*. Se são paralelos, podem ser coincidentes ou ter interseção vazia.

Proposição 3.4.8 *Seja $\Sigma = (O, \mathcal{B})$ um sistema de coordenadas em $\mathbb{E}^3$ e sejam π_1 e π_2 planos em $\mathbb{E}^3$, com equações gerais*

$$\pi_1: \ a_1x + b_1y + c_1z + d_1 = 0 \qquad e \qquad \pi_2: \ a_2x + b_2y + c_2z + d_2 = 0.$$

Então, π_1 e π_2 são paralelos se, e somente se, existir $\lambda \in \mathbb{R}$ tal que $\lambda a_1 = a_2$, $\lambda b_1 = b_2$ e $\lambda c_1 = c_2$.

Demonstração Suponha que exista $\lambda \in \mathbb{R}$ tal que $\lambda a_1 = a_2$, $\lambda b_1 = b_2$ e $\lambda c_1 = c_2$. Sejam $\vec{u_1}, \vec{v_1}$ vetores diretores de π_1. Se $\vec{u_1} = (r, s, t)_{\mathcal{B}}$ e $\vec{v_1} = (m, n, p)_{\mathcal{B}}$, então

$$a_2r + b_2s + c_2t = \lambda a_1r + a\lambda b_1s + \lambda c_1t = \lambda(a_1r + b_1s + c_1t) = 0,$$

pois $a_1r + b_1s + c_1t = 0$, pela Proposição 3.4.4, uma vez que, sendo diretor de π_1, $\vec{u_1}$ é paralelo a π_1. De modo similar, mostra-se que $a_2m + b_2n + c_2p = 0$. Portanto, tanto $\vec{u_1}$ como $\vec{v_1}$ são paralelos a π_2. Daí, segue que π_1 e π_2 são paralelos.

Reciprocamente, suponha que π_1 e π_2 sejam paralelos. Sabemos que pelo menos um dos escalares a_1, b_1, c_1 não é nulo. Suponha, por exemplo, que $a_1 \neq 0$. Mostremos que, tomando $\lambda = \frac{a_2}{a_1}$, temos as igualdades desejadas. É claro que $a_2 = \lambda a_1$. Para mostrar que $b_2 = \lambda b_1$, considere o vetor $\vec{w} = (b_2, -a_2, 0)_{\mathcal{B}}$. Pelo teste da Proposição 3.4.4, $\vec{w}$ é paralelo a π_2. Como π_2 e π_1 são paralelos, segue que $\vec{w}$ é paralelo a π_1 Outra aplicação da Proposição 3.4.4 fornece que $a_1b_2 + b_1(-a_2) = 0$, e, portanto, $b_2 = \lambda b_1$. Para mostrar que $c_2 = \lambda c_1$, usa-se um argumento análogo, com o fato de o vetor $\vec{t} = (c_2, 0, -a_2)_{\mathcal{B}}$ ser paralelo a π_2 e, *a fortiori*, a π_1. $\square$

Observação Na Proposição 3.4.8, se o sistema de coordenadas Σ for ortogonal, então $\vec{n_1} = (a_1, b_1, c_1)_{\mathcal{B}}$ é um vetor normal a π_1 e $\vec{n_2} = (a_2, b_2, c_2)_{\mathcal{B}}$ é um vetor normal a π_2. Então, π_1 e π_2 são paralelos se, e somente se, $\vec{n_1}$ e $\vec{n_2}$ são vetores paralelos. Isso ocorre se, e somente se, existe $\lambda \in \mathbb{R}$ tal que $\vec{n_2} = \lambda\vec{n_1}$ (lembre que vetores normais nunca são nulos). O que a Proposição 3.4.8 diz é que mesmo no caso de um sistema que não seja ortogonal, a proporcionalidade entre os três primeiros coeficientes das equações gerais dos planos é equivalente ao paralelismo entre eles.

Nos exemplos, entende-se fixado um sistema de coordenadas não necessariamente ortogonal em $\mathbb{E}^3$.

Exemplo 3.4.9 Estude a posição relativa dos planos $\pi_1 : \ X = (1, 0, 1) + \lambda(1, 1, 1) + \mu(0, 1, 0)$ $(\lambda, \mu \in \mathbb{R})$ e $\pi_2 : \ X = (0, 0, 0) + \alpha(1, 0, 1) + \beta(-1, 0, 3)$ $(\alpha, \beta \in \mathbb{R})$.

Solução: Equações gerais de π_1 e π_3 são obtidas fazendo, respectivamente,

$$\det\begin{bmatrix} x-1 & y & z-1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} = 0 \qquad \text{e} \qquad \det\begin{bmatrix} x & y & z \\ 1 & 0 & 1 \\ -1 & 0 & 3 \end{bmatrix} = 0.$$

Assim, $\pi_1 : \ x - z = 0$ e $\pi_2 : \ y = 0$. Aplicando o critério da Proposição 3.4.8, obtemos que π_1 e π_2 são transversais. A reta r de interseção de π e π_2 é definida por

$$r : \begin{cases} x - z = 0 \\ y = 0 \end{cases} \qquad \diamond$$

Exemplo 3.4.10 Estude a posição relativa dos planos $\pi_1 : 2x - y + z - 1 = 0$ e $\pi_2 : x - \frac{1}{2}y + \frac{1}{2}z - 9 = 0$.

Solução: Tomando $\lambda = \frac{1}{2}$ na Proposição 3.4.8, concluimos que π_1 e π_2 são paralelos. Agora, para que eles fossem coincidentes, seria necessário que o quarto escalar nas equações gerais estivessem na mesma proporção dos demais. Como $-9 \neq \frac{1}{2}(-1)$, segue que $\pi_1 \cap \pi_2 = \emptyset$. $\diamond$

Exercícios Lista 2 - Álgebra Linear I: Exs. 36–41.

3.5 Perpendicularidade e distâncias

Nesta seção, estará fixada uma orientação em $\mathbb{V}^3$ e um sistema ortogonal de coordenadas $\Sigma = (O, \mathcal{B})$ em $\mathbb{E}^3$, em que a base ortonormal $\mathcal{B} = \{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ de $\mathbb{V}^3$ é positiva. Coordenadas de pontos estarão dadas em relação a Σ, e coordenadas de vetores estarão dadas em relação a $\mathcal{B}$.

Comecemos por esclarecer o conceito de perpendicularidade envolvendo retas e planos.

Sejam r e s retas em $\mathbb{E}^3$ com vetores diretores $\vec{r}$ e $\vec{s}$, respectivamente. Dizemos que as retas r e s são *ortogonais* se os vetores $\vec{u}$ e $\vec{v}$ forem ortogonais. Dizemos que as retas r e s são *perpendiculares* se forem ortogonais e concorrentes. (Observe que retas reversas podem ser ortogonais, mas não são perpendiculares.)

Seja r uma reta em $\mathbb{E}^3$ com vetor diretor $\vec{r}$ e seja π um plano em $\mathbb{E}^3$ com vetor normal $\vec{n}$. Dizemos que r e π são *perpendiculares* se os vetores $\vec{r}$ e $\vec{n}$ forem paralelos.

Sejam π_1 e π_2 planos em $\mathbb{E}^3$ com vetores normais $\vec{n_1}$ e $\vec{n_2}$, respectivamente. Dizemos que os planos π_1 e π_2 são *perpendiculares* se os vetores $\vec{n_1}$ e $\vec{n_2}$ forem ortogonais.

Exemplo 3.5.1 Verifique se são ortogonais e perpendiculares as retas $r : X = (1, 1, 1) + \lambda(2, 1, -3)$ $(\lambda \in \mathbb{R})$ e $s : X = (0, 1, 0) + \lambda(-1, 2, 0)$ $(\lambda \in \mathbb{R})$.

Solução: Como vetores diretores de r e s temos $\vec{r} = (2, 1, -3)$ e $\vec{s} = (-1, 2, 0)$, respectivamente. Esses vetores são ortogonais, uma vez que $\vec{r} \cdot \vec{s} = -2 + 2 + 0 = 0$. Logo, r e s são retas ortogonais. Agora, tomando os pontos $A = (1, 1, 1) \in r$ e $B = (0, 1, 0) \in s$, obtemos o vetor $\overrightarrow{AB} = (-1, 0, -1)$. Como $\det \begin{bmatrix} 2 & 1 & -3 \\ -1 & 2 & 0 \\ -1 & 0 & -1 \end{bmatrix} = -11 \neq 0$, segue que $\{\vec{r}, \vec{s}, \overrightarrow{AB}\}$ é LI e, portanto, r e s são reversas. Logo, não são perpendiculares. $\diamond$

Exemplo 3.5.2 Encontre uma equação geral do plano π que contém a origem e é perpendicular à reta $r : X = (1, 1, 0) + \lambda(2, 3, 7)$ $(\lambda \in \mathbb{R})$.

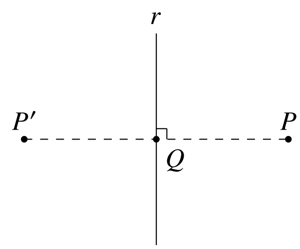
Solução: Para que r e π sejam perpendiculares, é preciso que o vetor diretor $\vec{r} = (2, 3, 4)$ de r seja normal a π. Assim, uma equação geral de π é da forma $2x + 3y + 7z + d = 0$. Como a origem $O = (0, 0, 0)$ é um ponto de π, segue que $d = 0$ e, portanto, $\pi : 2x + 3y + 7z = 0$. $\diamond$

Exemplo 3.5.3 Verifique se os planos $\pi_1 : X = (0, 0, 1) + \lambda(1, 0, 1) + \mu(-1, -1, 1)$ $(\lambda, \mu \in \mathbb{R})$ e $\pi_2 : 2x - 7y + 16z - 40 = 0$ são perpendiculares.

Solução: Como $\vec{u} = (1, 0, 1)$ e $\vec{v} = (-1, -1, 1)$ são vetores diretores de π_1, $\vec{n_1} = \vec{u} \wedge \vec{v} = \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 0 & 1 \\ -1 & -1 & 1 \end{bmatrix} = (1, -2, -1)$ é um vetor normal a π_1. Sabemos que $\vec{n_2} = (2, -4, 16)$ é normal a π_2. Temos $\vec{n_1} \cdot \vec{n_2} = 2 + 14 - 16 = 0$. Logo, $\vec{n_1}$ e $\vec{n_2}$ são ortogonais. Portanto, π_1 e π_2 são perpendiculares. $\diamond$

Exemplo 3.5.4 Considere a reta $r : X = (2,0,-1) + \lambda(3,1,3)$ ($\lambda \in \mathbb{R}$) e o ponto $P = (1,2,1)$. Determine a projeção ortogonal do ponto P sobre a reta r e o ponto P' simétrico de P em relação a r.

Solução: Seja Q a projeção ortogonal de P sobre r. Então, Q é um ponto de r tal que a reta PQ é perpendicular a r, e $P' = P + 2\overrightarrow{PQ}$. Veja a figura.



Como $Q \in r$, temos $Q = (2 + 3\lambda, \lambda, -1 + 3\lambda)$, para algum $\lambda \in \mathbb{R}$. Portanto, $\overrightarrow{PQ} = (1 + 3\lambda, \lambda - 2, -2 + 3\lambda)$. Como queremos que r e a reta PQ sejam perpendiculares, é preciso que os vetores $\vec{r} = (3,1,3)$ e $\overrightarrow{PQ}$ sejam ortogonais. Assim, o número λ que procuramos deve satisfazer $3(1 + 3\lambda) + (\lambda - 2) + 3(-2 + 3\lambda) = \vec{r} \cdot \overrightarrow{PQ} = 0$. Segue que $\lambda = \frac{5}{19}$. Logo, $Q = (\frac{53}{19}, \frac{5}{19}, -\frac{4}{19})$. Também segue que $\overrightarrow{PQ} = \frac{1}{19}(34, -33, -23)$. Assim, como $P' = P + 2\overrightarrow{PQ}$, segue que $P' = \big(1 + \frac{2}{19}34, 2 + \frac{2}{19}(-33), 1 + \frac{2}{19}(-23)\big) = (\frac{87}{19}, -\frac{28}{19}, -\frac{27}{19})$.

Uma outra solução resulta da observação de que $Q = A + \operatorname{proj}_{\vec{r}} \overrightarrow{AP}$, em que A é um ponto arbitrário de r (por exemplo, $A = (2,0,-1)$). Neste caso, temos $\overrightarrow{PQ} = \overrightarrow{AQ} - \overrightarrow{AP} = \operatorname{proj}_{\vec{r}} \overrightarrow{AP} - \overrightarrow{AP}$ e, assim, $P' = P + 2\big(\operatorname{proj}_{\vec{r}} \overrightarrow{AP} - \overrightarrow{AP}\big)$, sem precisarmos determinar Q. $\diamond$

Distância entre ponto e reta

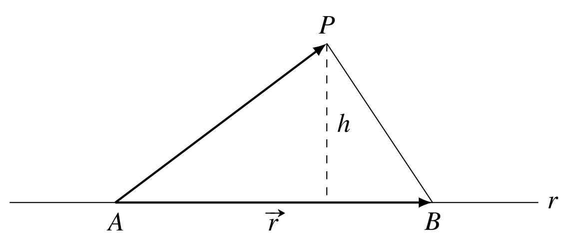
Seja r uma reta em $\mathbb{E}^3$ e seja $P \in \mathbb{E}^3$ um ponto que não está em r. Sabemos, da geometria euclidiana, que a *distância* $d(P,r)$ entre P e r é dada por $d(P,r) = d(P,Q)$, em que Q é o ponto de r de modo que a reta PQ seja perpendicular a r. Se $P \in r$, então $d(P,r) = 0$.

O resultado a seguir fornece uma fórmula para a distância entre um ponto e uma reta que prescinde da determinação do "pé da perpendicular" (isto é, do ponto Q como acima).

Proposição 3.5.5 *Seja r uma reta que passa pelo ponto A e tem vetor diretor $\vec{r}$, e seja P um ponto. Então,*

$$d(P,r) = \frac{\left\|\overrightarrow{AP} \wedge \vec{r}\right\|}{\|\vec{r}\|}. \tag{3.5}$$

Demonstração Se $P \in r$, então $\{\overrightarrow{AP}, \vec{r}\}$ é LD, e a fórmula é válida. Suponha que $P \notin r$ e seja $B = A + \vec{r}$. Então, $d(P,r) = h$, em que h denota a altura do triângulo ABP em relação ao lado AB.



Sabemos que a área do triângulo ABP é dada por $\frac{1}{2}\left\|\overrightarrow{AP} \wedge \overrightarrow{AB}\right\|$. Mas essa área também é dada por $\frac{1}{2}\left\|\overrightarrow{AB}\right\| h$. Igualando as duas expressões para a área, usando que $\overrightarrow{AB} = \vec{r}$ e resolvendo para h, obtemos a fórmula desejada. $\square$

Exemplo 3.5.6 Calcule a distância entre o ponto $P = (1, 1, -1)$ e a reta $r : \begin{cases} x - y - 1 = 0 \\ x + y - z = 0 \end{cases}$.

Solução: Para aplicar a fórmula (3.5), precisamos de um ponto sobre r e um vetor diretor de r. Como r está dada pela interseção de dois planos, sabemos que um vetor diretor de r pode ser obtido pelo produto vetorial entre os vetores normais aos planos. Assim, $\vec{r} = \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & -1 & 0 \\ 1 & 1 & -1 \end{bmatrix} = (1, 1, 2)$ é um vetor diretor de r. Um ponto sobre r é um ponto cujas coordenadas satisfaçam as equações gerais dos dois planos, por exemplo, $A = (0, -1, -1) \in r$. Assim, $\overrightarrow{AP} = (1, 2, 0)$ e, aplicando a fórmula (3.5), obtemos

$$d(P, r) = \frac{\|(1, 2, 0) \wedge (1, 1, 2)\|}{\|(1, 1, 2)\|} = \frac{\|(4, -2, -1)\|}{2} = \frac{\sqrt{14}}{2}.$$

(Uma outra maneira de se obter um vetor diretor de r seria determinar dois pontos A e B em r e tomar $\overrightarrow{AB}$ como diretor de r.) $\diamond$

Distância entre ponto e plano

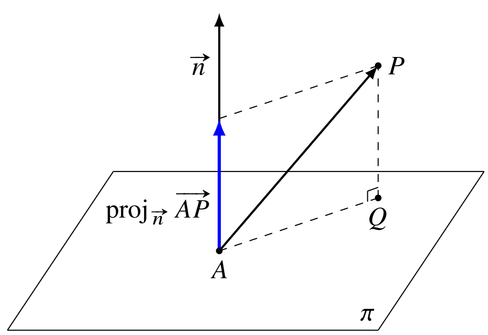
Seja π um plano em $\mathbb{E}^3$ e seja $P \in \mathbb{E}^3$ um ponto que não está em π. A *distância* $d(P, \pi)$ entre P e π é dada por $d(P, \pi) = d(P, Q)$, em que Q é o ponto de π de modo que a reta PQ seja perpendicular a π. Se $P \in \pi$, então $d(P, \pi) = 0$.

Existe uma fórmula para a distância entre um ponto e um plano que evita a determinação do ponto Q mencionado acima. Essa fórmula está dada na próxima proposição e é uma versão tridimensional para uma conhecida fórmula para a distância entre um ponto e uma reta da geometria analítica bidimensional.

Proposição 3.5.7 *Seja π o plano que contém o ponto A e tem vetor normal $\vec{n}$, e seja P um ponto. Então,*

$$d(P, \pi) = \frac{|\overrightarrow{AP} \cdot \vec{n}|}{\|\vec{n}\|}. \tag{3.6}$$

Demonstração Basta notar que a distância $d(P, \pi)$ é dada pelo comprimento da projeção ortogonal de $\overrightarrow{AP}$ na direção ortogonal ao plano π, ou seja, na direção de $\vec{n}$, conforme ilustrado na figura.



Assim,

$$d(P, \pi) = \left\|\mathrm{proj}_{\vec{n}}\,\overrightarrow{AP}\right\| = \left\|\frac{\overrightarrow{AP} \cdot \vec{n}}{\|\vec{n}\|^2}\vec{n}\right\| = \left|\frac{\overrightarrow{AP} \cdot \vec{n}}{\|\vec{n}\|^2}\right| \|\vec{n}\| = \frac{|\overrightarrow{AP} \cdot \vec{n}|}{\|\vec{n}\|^2}\|\vec{n}\| = \frac{|\overrightarrow{AP} \cdot \vec{n}|}{\|\vec{n}\|},$$

que é a fórmula desejada. $\square$

Observação Suponha, na Proposição 3.5.7, que

$$\pi: \ ax + by + cz + d = 0 \qquad \text{e} \qquad P = (x_0, y_0, z_0).$$

Então, um vetor normal a π será $\vec{n} = (a, b, c)$. Agora, dado $A = (x_1, y_1, z_1) \in \pi$, teremos $ax_1 + by_1 + cz_1 + d = 0$. Logo,

$$\overrightarrow{AP} \cdot \vec{n} = (x_0 - x_1, y_0 - y_1, z_0 - z_1) \cdot (a, b, c) = a(x_0 - x_1) + b(y_0 - y_1) + c(z_0 - z_1) = ax_0 + by_0 + cz_0 + d,$$

uma vez que $d = -(ax_1 + by_1 + cz_1)$. Assim, a fórmula (3.6) assume a forma

$$d(P, \pi) = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}. \tag{3.7}$$

Essa é a versão tridimensional para a conhecida fórmula da distância entre um ponto e uma reta da geometria analítica no plano.

Exemplo 3.5.8 (Prova 2, Álgebra Linear I, 2019) Considere o ponto $P = (-1, 1, 1)$ e o plano π dado por:

$$\begin{cases} x = -1 + \lambda + 7\mu \\ y = 1 - 2\lambda + \mu \\ z = \lambda - 5\mu \end{cases} \qquad (\lambda, \mu \in \mathbb{R}) \ .$$

A distância entre o ponto P e plano π é igual a

(A) $\frac{4}{\sqrt{5}}$ (B) $\frac{1}{\sqrt{2}}$ (C) $\frac{2}{\sqrt{3}}$ (D) $\frac{1}{2}$ (E) $\frac{2}{3}$

Solução: Das equações paramétricas no enunciado, vemos que $\vec{u} = (1, -2, 1)$ e $\vec{v} = (7, 1, -5)$ são vetores diretores do plano π e que o ponto $A = (-1, 1, 0)$ pertence ao plano π. Logo, uma equação geral para π é

$$\det \begin{bmatrix} x+1 & y-1 & z \\ 1 & -2 & 1 \\ 7 & 1 & -5 \end{bmatrix} = 0,$$

ou, $9x + 12y + 15z - 3 = 0$. Substituindo na fórmula (3.7), obtemos

$$d(P, \pi) = \frac{|9(-1) + 12(1) + 15(1) - 3|}{\sqrt{9^2 + 12^2 + 15^2}} = \frac{1}{\sqrt{2}} \ .$$

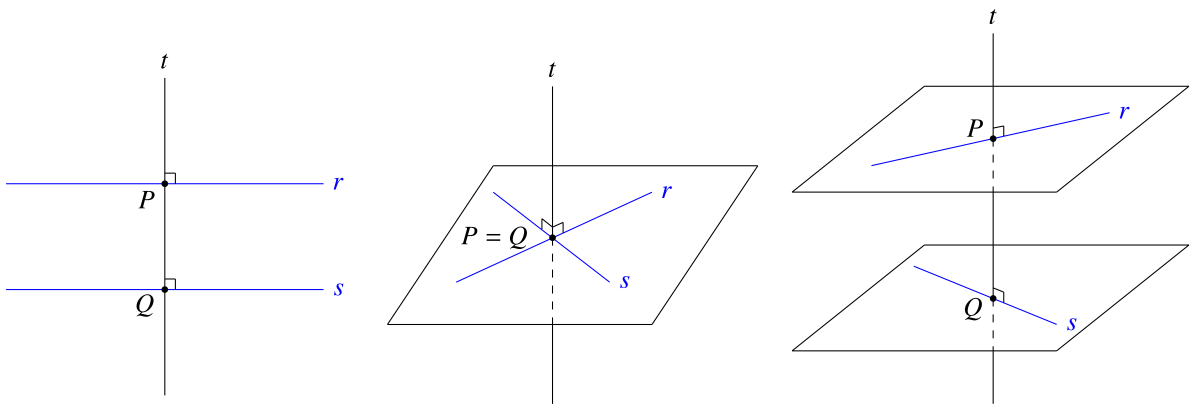
Resposta: (B). $\diamond$

Distância entre retas

Sejam r e s retas não coincidentes em $\mathbb{E}^3$, e seja t uma reta perpendicular a r e a s. A *distância* $d(r, s)$ entre r e s é dada por $d(r, s) = d(P, Q)$, em que P é o ponto de interseção entre r e t e Q é o ponto de interseção entre s e t. Se r e s são coincidentes, então $d(r, s) = 0$. S

Se r e s são paralelas, existem infinitas retas perpendiculares a r e a s. Mas se r e s são concorrentes ou reversas, existe apenas uma perpendicular comum. O leitor deve tentar se convencer deste fato. Segue dessa discussão que se r e s são concorrentes, então $d(r, s) = 0$.

A ilustração abaixo auxilia a visualizar as diferentes configurações.



retas paralelas retas concorrentes retas reversas

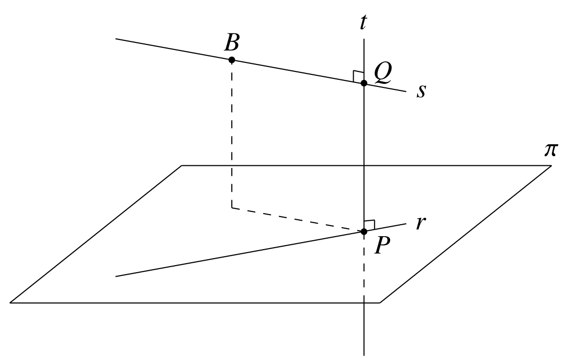
Se r e s são paralelas, então $d(r,s) = d(A,s) = d(B,r)$, quaisquer que sejam $A \in r$ ou $B \in s$. Assim, o cálculo de distâncias entre retas paralelas recai no cálculo de distâncias entre ponto e reta.

Se as retas não são paralelas, o resultado a seguir dá uma fórmula para a distância entre elas que evita passar pela determinação da perpendicular comum e suas interseções com as duas retas.

Proposição 3.5.9 *Sejam r e s retas não paralelas com vetores diretores $\vec{r}$ e $\vec{s}$, respectivamente. Seja $A \in r$ e seja $B \in s$. Então,*

$$d(r,s) = \frac{|[\vec{r}, \vec{s}, \overrightarrow{AB}]|}{\|\vec{r} \wedge \vec{s}\|}. \tag{3.8}$$

Demonstração Como r e s não são paralelas, $\{\vec{r}, \vec{s}\}$ é LI, e, portanto, $\vec{r} \wedge \vec{s} \neq \vec{0}$. Logo, a fórmula (3.8) faz sentido. Se r e s são concorrentes, então $\{\vec{r}, \vec{s}, \overrightarrow{AB}\}$ é LD, o que implica que $[\vec{r}, \vec{s}, \overrightarrow{AB}] = 0$. Logo, a fórmula (3.8) vale neste caso. Vejamos o caso em que r e s são reversas. Seja π o único plano que contém r e é paralelo a s. (Esse é o plano que tem vetores diretores $\vec{r}$, $\vec{s}$ e contém o ponto A.) Então, a distância entre r e s coincide com a distância entre B e π (veja a figura).



Como $\vec{r} \wedge \vec{s}$ é um vetor normal a π, podemos aplicar a Proposição 3.5.7 para obter

$$d(r,s) = d(B,\pi) = \frac{|\overrightarrow{AB} \cdot (\vec{r} \wedge \vec{s})|}{\|\vec{r} \wedge \vec{s}\|} = \frac{|[\vec{r}, \vec{s}, \overrightarrow{AB}]|}{\|\vec{r} \wedge \vec{s}\|},$$

que é o que desejávamos mostrar. □

Uma outra maneira de entender (3.8) é perceber que $d(r,s)$ é precisamente a altura do paralelepípedo definido pelos vetores $\vec{r}$, $\vec{s}$ e $\overrightarrow{AB}$ em relação à face definida por $\vec{r}$ e $\vec{s}$.

Exemplo 3.5.10 Determine a distância entre as retas

$$r:\ X=(1,2,3)+\lambda(0,1,3)\quad(\lambda\in\mathbb{R})\qquad \text{e}\qquad s:\ X=(0,1,0)+\mu(1,1,1)\quad(\mu\in\mathbb{R}).$$

Solução: Já havíamos visto, no Exemplo 3.4.1, que r e s são reversas. De todo modo, elas não são paralelas, uma vez que têm vetores diretores $\vec{r}=(0,1,3)$ e $\vec{s}=(1,1,1)$, respectivamente, e eles não são paralelos. Temos os pontos $A=(1,2,3)\in r$ e $B=(0,1,0)\in s$. Logo, por (3.8), temos

$$d(r,s)=\frac{|[\vec{r},\vec{s},\overrightarrow{AB}]|}{\|\vec{r}\wedge\vec{s}\|}=\frac{|[(0,1,3),(1,1,1),(-1,-1,-3)]|}{\|(0,1,3)\wedge(1,1,1)\|}=\frac{|2|}{\|(-2,3,-1)\|}=\frac{2}{\sqrt{14}}.\quad\diamond$$

Distância entre reta e plano

Seja r uma reta em $\mathbb{E}^3$ e seja π um plano em $\mathbb{E}^3$. Se $r\cap\pi=\emptyset$, isto é, se r e π forem paralelos, mas r não estiver contida em π, então a *distância* $d(r,\pi)$ entre a reta r e o plano π é dada por $d(r,\pi)=d(P,\pi)$, em que P é qualquer ponto de r. Se r estiver contida em π ou r e π forem transversais, então $d(r,\pi)=0$.

Assim, o cálculo da distância entre uma reta e um plano, quando a distância não é nula, fica reduzido ao cálculo da distância entre um ponto e um plano, que já tratamos anteriormente.

Distância entre planos

Se π_1 e π_2 são planos em $\mathbb{E}^3$ tais que $\pi_1\cap\pi_2=\emptyset$, ou seja, são planos paralelos não coincidentes, então a *distância* $d(\pi_1,\pi_2)$ entre os planos π_1 e π_2 satisfaz $d(\pi_1,\pi_2)=d(P,\pi_2)=d(Q,\pi_1)$, quaisquer que sejam $P\in\pi_1$ e $Q\in\pi_2$. Se π_1 e π_2 forem coincidentes ou transversais, $d(\pi_1,\pi_2)=0$.

Novamente, recaímos em um caso já estudado anteriormente.

Exercícios Lista 2 - Álgebra Linear I: Exs. 42–63.

Capítulo 4
Espaços vetoriais

Neste capítulo, será introduzida a noção de espaço vetorial, um tipo de estrutura algébrica que generaliza o espaço $\mathbb{V}^3$ dos vetores tridimensionais.

As referências para este e os demais capítulos destas notas são [3], [4] e [5].

4.1 Definição, exemplos e propriedades básicas

Lembre que no conjunto $\mathbb{V}^3$, formado pelos vetores tridimensionais, foram definidas duas operações: soma e multiplicação por um escalar. (Também foram definidos produto escalar e produto vetorial, mas essas operações não serão consideradas neste momento.) Essas operações tinham propriedades algébricas, listadas nas Proposições 2.1.1 e 2.1.2, que fazem de $\mathbb{V}^3$ um exemplo de um espaço vetorial, conforme a definição a seguir.

Definição Chamamos de *espaço vetorial* um conjunto V munido de duas operações

$$\begin{aligned} V \times V &\longrightarrow V \\ (u,v) &\longmapsto u+v \end{aligned} \qquad \text{e} \qquad \begin{aligned} \mathbb{R} \times V &\longrightarrow V \\ (\lambda,v) &\longmapsto \lambda v, \end{aligned}$$

denominadas *soma* e *multiplicação por escalar*, respectivamente, que satisfazem as seguintes condições:

(EV-1) $u + v = v + u$, para todos $u, v \in V$;
(EV-2) $u + (v + w) = (u + v) + w$, para todos $u, v, w \in V$;
(EV-3) existe $0_V \in V$ tal que $u + 0_V = u$, qualquer que seja $u \in V$;
(EV-4) para todo $u \in V$, existe $-u \in V$ tal que $u + (-u) = 0_V$;
(EV-5) $(\lambda\mu)u = \lambda(\mu u)$, para todos $\lambda, \mu \in \mathbb{R}$ e $u \in V$;
(EV-6) $(\lambda + \mu)u = \lambda u + \mu u$, para todos $\lambda, \mu \in \mathbb{R}$ e $u \in V$;
(EV-7) $\lambda(u + v) = \lambda u + \lambda v$, para todos $\lambda \in \mathbb{R}$ e $u, v \in V$;
(EV-8) $1u = u$, para todo $u \in V$.

Os elementos de um espaço vetorial V são denominados *vetores* de V; o vetor 0_V é chamado *vetor nulo* de V.

Vejamos alguns exemplos de espaços vetoriais.

Exemplo 4.1.1 Como já mencionado, $\mathbb{V}^3$ com as operações usuais de soma de vetores e de multiplicação de um escalar por um vetor é um espaço vetorial, em que $0_{\mathbb{V}^3} = \vec{0}$. As Proposições 2.1.1 e 2.1.2 garantem que as oito condições na definição de espaço vetorial estão satisfeitas.

Exemplo 4.1.2 Seja n um inteiro positivo, e considere o conjunto $\mathbb{R}^n$ formado por todas as sequências de n números reais, isto é,

$$\mathbb{R}^n = \left\{ (a_1, a_2, \dots, a_n) \mid a_i \in \mathbb{R}, \text{ para todo } i = 1, \dots, n \right\}.$$

O conjunto $\mathbb{R}^n$ tem uma estrutura natural de espaço vetorial em que a soma é dada por

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n),$$

e a multiplicação por escalar, por

$$\lambda(a_1, a_2, \dots, a_n) = (\lambda a_1, \lambda a_2, \dots, \lambda a_n).$$

Para demonstrar essa afirmação é necessário verificar cada uma das condições (EV-1)–(EV-8) na definição de espaço vetorial. Essa é uma tarefa simples (e tediosa), que fica a cargo do leitor. O que vale a pena destacar é que, com essas operações, os vetores cujas existências estão postuladas nas condições (EV-3) e (EV-4) são dados por

$$0_{\mathbb{R}^n} = (0, 0, \dots, 0)$$

e

$$-(a_1, a_2, \dots, a_n) = (-a_1, -a_2, \dots, -a_n).$$

Exemplo 4.1.3 Seja I um intervalo contido na reta real $\mathbb{R}$. Considere o conjunto das funções reais definidas em I:

$$\mathcal{F}(I) = \left\{ f : I \to \mathbb{R} \mid f \text{ é uma função} \right\}.$$

Existe uma estrutura natural de espaço vetorial em $\mathcal{F}(I)$ em que as operações são tais que se $f, g \in \mathcal{F}(I)$ e $\lambda \in \mathbb{R}$, então

$$(f + g)(x) = f(x) + g(x) \qquad \text{e} \qquad (\lambda f)(x) = \lambda f(x),$$

para todo $x \in I$. Neste caso, o vetor nulo $0_{\mathcal{F}}(I)$ é a função nula $\mathbf{n} : I \to \mathbb{R}$, definida por $\mathbf{n}(x) = 0$, para todo $x \in I$. E se $f \in \mathcal{F}(I)$, então $-f \in \mathcal{F}(I)$ é a função definida por $(-f)(x) = -f(x)$, para todo $x \in I$. Mais uma vez, a tarefa de verificação de que as condições (EV-1)–(EV-8) estão satisfeitas fica a cargo do leitor.

Exemplo 4.1.4 Um exemplo de espaço vetorial já visto nesta notas, mas não nomeado dessa maneira, são os espaços vetoriais de matrizes. Sejam m, n inteiros positivos e denote (como vimos fazendo) o conjunto de todas as matrizes $m \times n$ por $M_{m\times n}(\mathbb{R})$. Esse conjunto tem uma estrutura natural de espaço vetorial em que as operações são dadas por

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}) \qquad \text{e} \qquad \lambda(a_{ij}) = (\lambda a_{ij}).$$

Essas são as mesmas operações que introduzimos na Seção 1.2. Aqui, $0_{M_{m\times n}(\mathbb{R})}$ é a matriz $m \times n$ nula, cujas entradas são todas iguais a 0, e $-(a_{ij}) = (-a_{ij})$.

Exemplo 4.1.5 Este é um exemplo que será bastante explorado neste curso. Seja n um inteiro positivo. Denote por $\mathcal{P}_n(\mathbb{R})$ o conjunto formado por todos os polinômios de grau menor ou igual a n e o polinômio nulo. Ou seja,

$$\mathcal{P}_n(\mathbb{R}) = \left\{ a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mid a_0, a_1, \dots, a_n \in \mathbb{R} \right\}.$$

Definimos, em $\mathcal{P}_n(\mathbb{R})$ a seguinte operação de soma: dados $p = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n, q = b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n \in \mathcal{P}_n(\mathbb{R})$, define-se

$$p + q = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_n + b_n)x^n.$$

Uma operação de multiplicação por escalar em $\mathcal{P}_n(\mathbb{R})$ é definida da seguinte maneira: dados $\lambda \in \mathbb{R}$ e $p = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in \mathcal{P}_n(\mathbb{R})$, então

$$\lambda p = \lambda a_0 + \lambda a_1 x + \lambda a_2 x^2 + \cdots + \lambda a_n x^n.$$

Com essas operações, $\mathcal{P}_n(\mathbb{R})$ é um espaço vetorial em que $0_{\mathcal{P}_n(\mathbb{R})}$ é o polinômio nulo, ou seja, aquele em que todos os $n+1$ coeficientes são iguais a 0, e dado $p = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in \mathcal{P}_n(\mathbb{R})$, temos $-p = -a_0 - a_1 x - a_2 x^2 - \cdots - a_n x^n$, como o leitor pode verificar.

Observe que $\mathcal{P}_n(\mathbb{R})$ é um subconjunto de $\mathcal{F}(\mathbb{R})$. Na seção seguinte, neste capítulo, veremos o conceito de subespaço vetorial de um espaço vetorial e mostraremos que, de fato, $\mathcal{P}_n(\mathbb{R})$ é um subespaço de $\mathcal{F}(\mathbb{R})$.

Exemplo 4.1.6 (Lista 1 - Álgebra Linear II, ex. 2) Consideremos um exemplo em que as operações em um espaço vetorial não são as "usuais", por assim dizer.

Seja $V = \{u \in \mathbb{R} \mid u > 0\}$. Serão definidas operações no conjunto V que farão dele um espaço vetorial. A fim de evitar ambiguidades na compreensão das operações a serem definidas, utilizaremos, neste exemplo, o símbolo $\oplus$ para denotar a operação de soma no espaço vetorial V e o símbolo $\odot$ para denotar a operação de multiplicação por escalar no espaço vetorial V.

Vamos às definições: dados $u, v \in V$, defina a soma $u \oplus v$ por

$$u \oplus v = uv,$$

e, dados $\lambda \in \mathbb{R}$ e $u \in V$, defina a multiplicação por escalar $\lambda \odot u$ por

$$\lambda \odot u = u^{\lambda}.$$

Observe que, de fato, $u \oplus v$ é um elemento do conjunto V, uma vez que se $u, v \in V$, então $u > 0$ e $v > 0$, o que implica $uv > 0$. Também é verdade que se $\lambda \odot u \in V$, quaisquer que sejam $\lambda \in \mathbb{R}$ e $u \in V$, pois, como $u > 0$, $u^{\lambda} > 0$. Em resumo, as definições de $\oplus$ e $\odot$ são, de fato, operações no conjunto V. Agora, passemos às condições (EV-1)–(EV-8). Por exemplo, para verificar a condição (EV-1), é preciso mostrar que quaisquer que sejam $u, v \in V$, temos $u \oplus v = v \oplus u$. Mas isso é, de fato, verdadeiro, uma vez que o produto entre números reais é comutativo: $uv = vu$. Deixemos a verificação das condições (EV-2)–(EV-4) a cargo do leitor, com apenas dois comentários. O primeiro diz respeito ao candidato ao vetor nulo nesse espaço vetorial. Fica claro que para provar a validade de (EV-3) é preciso tomar $0_V = 1$, que, cabe observar, é um elemento do conjunto V. Ainda, o "oposto" de um elemento u de V, aqui denotado por $\ominus u$, dever satisfazer $\ominus u = \frac{1}{u}$, para que a condição (EV-4) esteja satisfeita.

Façamos a verificação de mais uma das oito propriedades das operações que definem um espaço vetorial. Por exemplo, vejamos (EV-7): seja $\lambda \in \mathbb{R}$ e sejam $u, v \in V$. Então,

$$\lambda \odot (u \oplus v) = \lambda \odot (uv) = (uv)^{\lambda} = u^{\lambda} v^{\lambda} = u^{\lambda} \oplus v^{\lambda} = (\lambda \odot u) \oplus (\lambda \odot v),$$

como queríamos. O leitor deve terminar a verificação das demais condições.

Exemplo 4.1.7 Vejamos, agora, um exemplo de um conjunto munido de duas operações $\oplus$ e $\odot$ que *não* é um espaço vetorial porque as condições (EV-1)–(EV-8) não estão todas satisfeitas.

Seja $U = \{(a, b) \mid a, b \in \mathbb{R}\}$. (O conjunto U nada mais é, como conjunto, que o conjunto $\mathbb{R}^2$). Defina em U as seguintes operações:

$$(a, b) \oplus (c, d) = (a + d, b + c), \text{ para todos } (a, b), (c, d) \in U,$$

e

$$\lambda \odot (a, b) = (\lambda a, \lambda b), \text{ para todos } \lambda \in \mathbb{R} \text{ e } (a, b) \in U.$$

Para mostrar que U não é um espaço vetorial, basta mostrar que pelo menos uma das condições (EV-1)–(EV-8) não está satisfeita (uma vez que, para ser um espaço vetorial, todas as condições deveriam estar satisfeitas). Mostremos, por exemplo, que a condição (EV-1) não vale. Como essa condição faz referência a uma igualdade valer para *todos* os elementos do conjunto em que estão definidas as operações, para mostrar que ela é violada, é suficiente exibir uma escolha particular de elementos u e v para a qual a condição não vale. No nosso caso específico, se tomarmos os elementos $(1, 2)$ e $(3, 5)$ de U, vemos que, por um lado

$$(1, 2) \oplus (3, 5) = (6, 5),$$

e, por outro lado,

$$(3, 5) \oplus (1, 2) = (5, 6).$$

Segue que $(1, 2) \oplus (3, 5) \neq (3, 5) \oplus (1, 2)$. Assim, a condição (EV-1) não está satisfeita e, portanto, U não é um espaço vetorial com essas particulares definições de operações. (Já havíamos visto, no Exemplo 4.1.2, com $n = 2$, que existem operações em $\mathbb{R}^2$ que fazem dele um espaço vetorial.)

Observação Note que, na definição de espaço vetorial, não há menção ao fato de haver um *único* vetor nulo em um espaço vetorial. Também não é dito que dado um vetor u em um espaço vetorial existe um *único* vetor $-u$ tal que $u + (-u) = 0_V$. Apesar disso, ambas as unicidades podem ser garantidas. É o que faremos a seguir.

Seja V um espaço vetorial.

- Suponha que $0_V, 0'_V \in V$ sejam vetores que satisfazem a condição (EV-3) na definição. Mostremos que $0_V = 0'_V$. Com efeito,

$$
\begin{aligned}
0'_V &= 0'_V + 0_V && \text{pois } 0_V \text{ satisfaz (EV-3)}\\
&= 0_V + 0'_V && \text{por (EV-1)}\\
&= 0_V && \text{pois } 0'_V \text{ satisfaz (EV-3)}
\end{aligned}
$$

- Seja $u \in V$. Suponha que $-u_1$ e $-u_2$ sejam vetores que satisfazem a condição (EV-4) na definição. Mostremos que $-u_1 = -u_2$. Com efeito,

$$
\begin{aligned}
-u_1 &= -u_1 + 0_V && \text{por (EV-3)}\\
&= -u_1 + \big(u + (-u_2)\big) && \text{pois } -u_2 \text{ satisfaz (EV-4)}\\
&= (-u_1 + u) + (-u_2) && \text{por (EV-2)}\\
&= \big(u + (-u_1)\big) + (-u_2) && \text{por (EV-1)}\\
&= 0_V + (-u_2) && \text{pois } -u_1 \text{ satisfaz (EV-4)}\\
&= -u_2 + 0_V && \text{por (EV-1)}\\
&= -u_2 && \text{por (EV-3)}
\end{aligned}
$$

Algumas propriedades que todos os espaços vetoriais satisfazem são decorrências imediadas da definição, como as que constam da proposição a seguir.

Proposição 4.1.8 *Seja V um espaço vetorial. Então, valem:*

(i) $\lambda 0_V = 0_V$, *para todo* $\lambda \in \mathbb{R}$;
(ii) $0u = 0_V$, *para todo* $u \in V$;
(iii) se $\lambda \in \mathbb{R}$ *e* $u \in V$ *são tais que* $\lambda u = 0_V$, *então* $\lambda = 0$ *ou* $u = 0_V$;
(iv) $(-\lambda)u = -(\lambda u) = \lambda(-u)$, *para todos* $\lambda \in \mathbb{R}$ *e* $u \in V$;
(v) $-(-u) = u$, *para todo* $u \in V$;
(vi) se $u, v, w \in V$ *são tais que* $u + v = u + w$, *então* $v = w$.

Observe que, fazendo $\lambda = 1$ em (iv), acima, obtém-se $(-1)u = -u$, para todo $u \in V$.

Demonstração Faremos a demonstração apenas de (i), ficando as demais propriedades a cargo do leitor. Começemos por observar que

$$
\lambda 0_V = \lambda(0_V + 0_V) = \lambda 0_V + \lambda 0_V, \tag{4.1}
$$

em que utilizamos (EV-3) na primeira igualdade e (EV-7), na segunda. Por (EV-4), existe um vetor $-(\lambda 0_V)$ que satisfaz $\lambda 0_V + \big(-(\lambda 0_V)\big) = 0_V$. Assim,

$$
\begin{aligned}
0_V &= \lambda 0_V + \big(-(\lambda 0_V)\big) = (\lambda 0_V + \lambda 0_V) + \big(-(\lambda 0_V)\big) && \text{por (4.1)}\\
&= \lambda 0_V + \Big(\lambda 0_V + \big(-(\lambda 0_V)\big)\Big) && \text{por (EV-2)}\\
&= \lambda 0_V + 0_V && \text{por (EV-4)}\\
&= \lambda 0_V, && \text{por (EV-3)}
\end{aligned}
$$

que é a igualdade desejada. $\square$

Deste ponto em diante, faremos uso das propriedades relacionadas na proposição acima, bem como das oito condições na definição de espaço vetorial, sem fazer menção explícita a elas.

Uma questão notacional: se u e v são vetores em um espaço vetorial, o vetor $u + (-v)$ será, doravante, simplesmente denotado por $u - v$.

Exercícios Lista 1 - Álgebra Linear II: Exs. 1–2.

4.2 Subespaços vetoriais

No Exemplo 4.1.5, observamos que o espaço vetorial $\mathcal{P}_n(\mathbb{R})$ estava, como conjunto, contido no espaço vetorial $\mathcal{F}(\mathbb{R})$. Veremos que, de acordo com a próxima definição, $\mathcal{P}_n(\mathbb{R})$ é, de fato, um subespaço vetorial de $\mathcal{F}(\mathbb{R})$.

Definição Seja V um espaço vetorial. Dizemos que um subconjunto W de V é um *subespaço vetorial* de V se

(i) $0_V \in W$;
(ii) para todos $u, v \in W$, temos $u + v \in W$;
(iii) para todos $\lambda \in \mathbb{R}$ e $u \in W$, temos $\lambda u \in W$.

Em outras palavras, um subespaço vetorial (ou, simplesmente, subespaço) de um espaço vetorial é um subconjunto do espaço vetorial que contém o vetor nulo e é "fechado" para as operações de soma e multiplicação por escalar, no sentido de que quando se soma dois vetores desse subconjunto obtém-se um vetor que ainda está no subconjunto. Analogamente, não se "sai" desse subconjunto tomando-se múltiplos escalares de vetores nele.

Observação Se W é um subespaço do espaço vetorial V, então W é, ele mesmo, um espaço vetorial com operações dadas pela restrições das operações de V a elementos de W. As condições (EV-1)–(EV-8) estão satisfeitas em W porque valiam para todos os vetores de V, em particular, para todos os vetores de W. (É claro que $0_W = 0_V$ e que a condição (EV-4) está satisfeita em W, pois se $u \in W$, então $-u = (-1)u \in W$.)

Vejamos alguns exemplos.

Exemplo 4.2.1 Considere o espaço vetorial $\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$ com as operações usuais (aquelas introduzidas no Exemplo 4.1.2), e o seguinte subconjunto de $\mathbb{R}^3$:

$$W = \left\{ (x, y, z) \in \mathbb{R}^3 \mid x - 2y = 0 \right\}.$$

Isto é, os elementos de W são apenas aquelas ternas ordenadas de números reais (x, y, z) em que $x - 2y = 0$. Por exemplo, $(-2, -1, 9) \in W$, porque $(-2) - 2(-1) = 0$, mas $(2, -1, 0) \neq W$, uma vez que $2 - 2(-1) \neq 0$. Mostremos que W é um subespaço de $\mathbb{R}^3$.

(i) O vetor nulo de $\mathbb{R}^3$ é o elemento $0_{\mathbb{R}^3} = (0, 0, 0)$, como vimos no Exemplo 4.1.2. Como $0 - 2(0) = 0$, segue que, de fato, $0_{\mathbb{R}^3} \in W$.

(ii) Vejamos por que W é fechado para a soma. Sejam w_1 e w_2 dois vetores de W. Então, cada um deles é um elemento de $\mathbb{R}^3$ satisfazendo a condição para estar em W, isto é, $w_1 = (x_1, y_1, z_1)$ com $x_1 - 2y_1 = 0$, e $w_2 = (x_2, y_2, z_2)$ com $x_2 - 2y_2 = 0$. Verifiquemos que $w_1 + w_2 \in W$. Para tanto, é preciso lembrar da definição da operação de soma no espaço vetorial do qual W é um subconjunto, no caso, $\mathbb{R}^3$. Conforme mencionamos, a operação de soma é aquela introduzida no Exemplo 4.1.2. Assim,

$$w_1 + w_2 = (x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2).$$

Para que esse vetor soma esteja em W, é preciso que ele satisfaça a condição que define os vetores de W, o que, de fato, é o caso, uma vez que $(x_1 + x_2) - 2(y_1 + y_2) = (x_1 - 2y_1) + (x_2 - 2y_2) = 0 + 0 = 0$. Logo, W é fechado para a soma.

(iii) Finalmente, verifiquemos que W é fechado para multiplicação por escalar. Se $\lambda \in \mathbb{R}$ e $w = (x, y, z) \in W$, então $\lambda w = (\lambda x, \lambda y, \lambda z)$, já que a multiplicação por escalar em $\mathbb{R}^3$ é aquela definida no Exemplo 4.1.2. Agora,

$\lambda x - 2(\lambda y) = \lambda(x - 2y) = \lambda 0 = 0$, uma vez que $x - 2y = 0$, pois w é um elemento de W. Logo, W é também fechado para multiplicação por escalar.

Como as três condições na definição de subespaço estão satisfeitas, W é um subespaço de $\mathbb{R}^3$.

Tome, agora, o seguinte subconjunto de $\mathbb{R}^3$:

$$L = \left\{ (x, y, z) \in \mathbb{R}^3 \mid x + y = 1 \right\}.$$

Esse subconjunto *não* é um subespaço de $\mathbb{R}^3$, pois, por exemplo, $0_{\mathbb{R}^3} \notin L$. (Na verdade, L não satisfaz nenhuma das três condições para ser um subespaço, mas basta que uma delas não esteja satisfeita para que L não seja um subespaço).

Exemplo 4.2.2 Todo espaço vetorial V tem pelo menos dois subespaços, chamados subespaços triviais de V: o subespaço $\{0_V\}$ formado apenas pelo vetor nulo, e o subespaço total V. Espaços vetoriais, em geral, contém muitos outros subespaços além dos triviais.

Exemplo 4.2.3 Este é um dos exemplos mais importantes de subespaço vetorial. Seja $A \in M_{m\times n}(\mathbb{R})$ e considere o sistema linear homogêneo $AX = 0$ de m equações e n incógnitas. As soluções de $AX = 0$ são elementos de $\mathbb{R}^n$. Mostremos que o conjunto S formado por todas as soluções de $AX = 0$ é um subespaço do espaço vetorial $\mathbb{R}^n$. (Aqui, consideramos as operações usuais de $\mathbb{R}^n$, aquelas do Exemplo 4.1.2).

(i) Temos que $0_{\mathbb{R}^n} = (0, 0, \dots, 0) \in S$, uma vez que um sistema homogêneo sempre tem, pelo menos, a solução trivial.

(ii) Sejam $u = (a_1, a_2, \dots, a_n)$ e $v = (b_1, b_2, \dots, b_n)$ elementos de S, isto é, u e v são soluções do sistema $AX = 0$. Mostremos que $u + v \in S$. Para tanto, é preciso mostrar que $u + v = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$ é solução de $AX = 0$. Mas isso é verdade, uma vez que vale a igualdade matricial

$$A\begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix} = A\left(\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}\right) = A\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + A\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

(iii) Seja $u = (a_1, a_2, \dots, a_n) \in S$ e seja $\lambda \in \mathbb{R}$. Então $\lambda u \in S$, pois

$$A\left(\lambda\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}\right) = \lambda A\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \lambda\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Conclui-se que, de fato, S é um subespaço de $\mathbb{R}^n$.

Observe que o subespaço W de $\mathbb{R}^3$ do Exemplo 4.2.1 anterior é um caso particular deste, em que $m = 1$, $n = 3$ e $A = \begin{bmatrix} 1 & -2 & 0 \end{bmatrix}$.

O conjunto formado por todas as soluções de um sistema linear *não homogêneo* de m equações e n incógnitas ainda é um subconjunto de $\mathbb{R}^n$, mas não é um subespaço de $\mathbb{R}^n$. (Você consegue ver por quê?)

Exemplo 4.2.4 Lembre, do Exemplo 4.1.5, que o conjunto $\mathcal{P}_n(\mathbb{R})$ de todos os polinômios de grau $\leq n$, mais o polinômio nulo, tem uma estrutura natural de espaço vetorial. Convença-se de que se $m < n$, então $\mathcal{P}_m(\mathbb{R})$ é um subespaço de $\mathcal{P}_n(\mathbb{R})$.

Exemplo 4.2.5 Dado um intervalo aberto I na reta real, denote por $C(I)$ o conjunto de todas as funções $f : I \to \mathbb{R}$ contínuas. No curso de Cálculo, vimos que a função nula $\mathbf{n} : I \to \mathbb{R}$ é contínua, que soma de funções contínuas é contínua e que a multiplicação de uma constante por uma função contínua é contínua. Esses três fatos podem ser resumidos na afirmação de que $C(I)$ é um subespaço do espaço vetorial $\mathcal{F}(I)$ de todas as funções reais com domínio I. Se $\mathcal{D}(I)$ denota o conjunto de todas as funções $f : I \to \mathbb{R}$ deriváveis, então, segue, também de resultados vistos no curso de Cálculo, que $\mathcal{D}(I)$ é um subespaço de $C(I)$.

Mais geralmente, temos uma cadeia infinita de subespaços vetoriais:

$$\mathcal{P}_0(\mathbb{R}) \subseteq \mathcal{P}_1(\mathbb{R}) \subseteq \mathcal{P}_2(\mathbb{R}) \subseteq \cdots \subseteq \mathcal{P}(\mathbb{R}) \subseteq \mathcal{D}(\mathbb{R}) \subseteq C(\mathbb{R}) \subseteq \mathcal{F}(\mathbb{R}),$$

em que $\mathcal{P}(\mathbb{R})$ denota o espaço vetorial formado por todos os polinômios, sem limitação no grau.

Exemplo 4.2.6 Agora, um exemplo em um espaço de matrizes. Afirmamos que

$$W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R}) \;\middle|\; b = c \right\}$$

é um subespaço do espaço vetorial $M_2(\mathbb{R})$ (com as operações usuais definidas no Exemplo 4.1.4.) Com efeito, $0_{M_2(\mathbb{R})} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \in W$, claramente; se $A, B \in W$, então $A + B \in W$ (pois se $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ e $B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$, então $A + B = \begin{bmatrix} a + a' & b + b' \\ c + c' & d + d' \end{bmatrix}$ e $b + b' = c + c'$, uma vez que $b = c$ e $b' = c'$, já que A e B são elementos de W); se $\lambda \in \mathbb{R}$ e $A \in W$, então $\lambda A \in W$ (pois se $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, então $\lambda A = \begin{bmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{bmatrix}$ e $\lambda b = \lambda c$, já que $b = c$).

Exemplo 4.2.7 O subconjunto

$$S = \left\{ A \in M_2(\mathbb{R}) \mid \det(A) = 0 \right\}$$

do espaço vetorial $M_2(\mathbb{R})$ não é um subespaço, apesar de satisfazer as condições (i) e (iii) na definição de subespaço, com o leitor pode facilmente verificar. Mas S não satisfaz a condição (ii); por exemplo, $M = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ e $N = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ são matrizes de determinante nulo, portanto, $M, N \in S$. Mas $M + N = I_2$ e $\det(I_2) = 1 \neq 0$. Logo, $M + N \notin S$. O subconjunto S não é fechado para soma. Portanto, não é um subespaço de $M_2(\mathbb{R})$.

O que fizemos nesta seção e na anterior — e faremos nas seguintes — foi explorar consequências das condições que definem um espaço vetorial. Pode-se demonstrar que muitas das propriedades do espaço vetorial $\mathbb{V}^3$, por exemplo, são consequências apenas do fato de, nele, estar definida uma estrutura de espaço vetorial, isto é, de estarem definidas em $\mathbb{V}^3$ operações que satisfazem as condições (EV-1)–(EV-8), e não necessariamente de o conjunto $\mathbb{V}^3$ ser formado por vetores, definidos em termos de segmentos orientados em E^3. Como vimos, há diversos exemplos diferentes de espaços vetoriais. Assim, tudo que foi — e será — demonstrado para um espaço vetorial abstrato, isto é, um conjunto munido de duas operações satisfazendo as condições (EV-1)–(EV-8), será válido em cada um desses exemplos. Este é o poder do método axiomático: por meio da obtenção de consequências lógicas de um determinado conjunto de axiomas (por exemplo, as condições que definem espaço vetorial), alcançamos resultados que são válidos em qualquer instância particular em que os axiomas estão satisfeitos. Isso ficará ainda mais claro nas sessões seguintes, em que os conceitos de combinação linear, dependência linear, base e coordenadas serão introduzidos no contexto mais abstrato dos espaços vetoriais.

Exercícios Lista 1 - Álgebra Linear II: Exs. 3–5.

4.3 Combinações lineares e dependência linear

Seja V um espaço vetorial e sejam $u_1, u_2, \ldots, u_n \in V$. Dizemos que um vetor $v \in V$ é *combinação linear* de $u_1, u_2, \ldots, u_n$ se existirem $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tais que

$$v = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n.$$

(Essa definição dever ser comparada com a definição de combinação linear definida entre vetores de $\mathbb{V}^3$ na Seção 2.2.)

Exemplo 4.3.1 Em $\mathbb{R}^2$, $(-3, 8)$ é combinação linear de $(-1, 0)$ e $(1, 2)$, uma vez que $(-3, 8) = 7(-1, 0) + 4(1, 2)$.

Exemplo 4.3.2 Em $\mathcal{P}_4(\mathbb{R})$, $f = 8x^4 - 3x^2 + 5x + 52$ é combinação linear de $g = x^4 - 2x + 8$ e $h = x^2 - 7x + 4$, pois $f = 8g - 3h$.

Exemplo 4.3.3 Qualquer vetor de $\mathcal{P}_n(\mathbb{R})$ é combinação linear dos seguintes $n+1$ vetores: $1, x, x^2, \dots, x^n$.

Exemplo 4.3.4 Em $\mathbb{R}^2$, $(2,1)$ não é combinação linear de $(1,0)$ e $(-1,0)$, uma vez que qualquer combinação linear desses dois vetores será da forma $(\alpha, 0)$, para algum $\alpha \in \mathbb{R}$.

Definição Seja V um espaço vetorial e seja $S = \{u_1, u_2, \dots, u_n\}$ um subconjunto finito de V. Definimos o *subespaço vetorial de V gerado por S* como sendo o subconjunto de V dado por

$$\big[S\big] = \big\{\, \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n \mid \alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R} \,\big\}.$$

Isto é, $[S]$ é o conjunto formado por *todas* as combinações lineares dos elementos de S.

Por exemplo, vimos no Exemplos 4.3.1 e 4.3.4, que, em $\mathbb{R}^2$, $(-3,8) \in [(-1,0),(1,2)]$, mas $(2,1) \notin [(1,0),(-1,0)]$. Já no Exemplo 4.3.3, foi observado que, em $\mathcal{P}_n(\mathbb{R})$, temos $\big[1, x, x^2, \dots, x^n\big] = \mathcal{P}_n(\mathbb{R})$.

Por convenção, definimos $[\varnothing] = \{0_V\}$.

A proposição a seguir justifica o nome dado a $[S]$.

Proposição 4.3.5 *Seja V um espaço vetorial e seja S um subconjunto finito de S. Então, $[S]$ é um subespaço de V.*

Demonstração Se $S = \varnothing$, a convenção acima dá conta do resultado. Suponha que $S = \{u_1, u_2, \dots, u_n\}$. Mostremos que $[S]$, conforme definido acima, satisfaz as três condições na definição de subespaço de V.

(i) $0_V \in [S]$, pois $0_V = 0u_1 + 0u_2 + \dots + 0u_n$.

(ii) Sejam $v, w \in [S]$. Então, $v = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n$ e $w = \beta_1 u_1 + \beta_2 u_2 + \dots + \beta_n u_n$, com $\alpha_i, \beta_i \in \mathbb{R}$. Logo,

$$\begin{aligned} v + w &= (\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n) + (\beta_1 u_1 + \beta_2 u_2 + \dots + \beta_n u_n) \\ &= (\alpha_1 + \beta_1)u_1 + (\alpha_2 + \beta_2)u_2 + \dots + (\alpha_n + \beta_n)u_n, \end{aligned}$$

e, portanto, $v + w \in [S]$.

(iii) Sejam $\lambda \in \mathbb{R}$ e $v \in [S]$. Então, $v = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n$, com $\alpha_i \in \mathbb{R}$, e, portanto,

$$\begin{aligned} \lambda v &= \lambda(\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n) \\ &= (\lambda\alpha_1)u_1 + (\lambda\alpha_2)u_2 + \dots + (\lambda\alpha_n)u_n. \end{aligned}$$

O que prova que $\lambda v \in [S]$. □

Se $S = \{u_1, u_2, \dots, u_n\}$, é comum denotar $[S]$ por $[u_1, u_2, \dots, u_n]$, simplesmente. (Um alerta: a notação, com colchetes, para o subespaço gerado por um conjunto com 3 vetores pode se confundir, em $\mathbb{V}^3$, com a notação para o produto misto. Geralmente, o contexto será suficiente para evitar ambiguidades.)

Exemplo 4.3.6 Descreva o subespaço W de $\mathbb{R}^3$ definido por $W = [u, v]$, em que $u = (1,-1,1)$ e $v = (0,1,-2)$.

Solução: Dado $w = (x,y,z) \in \mathbb{R}^3$, temos

$$\begin{aligned} w \in [u,v] &\iff \text{existem } \alpha, \beta \in \mathbb{R} \text{ tais que } w = \alpha u + \beta v \\ &\iff \text{existem } \alpha, \beta \in \mathbb{R} \text{ tais que } (x,y,z) = \alpha(1,-1,1) + \beta(0,1,-2) \\ &\iff \text{existem } \alpha, \beta \in \mathbb{R} \text{ tais que } \begin{cases} \alpha = x \\ -\alpha + \beta = y \\ \alpha - 2\beta = z \end{cases} \\ &\iff \text{o sistema linear } \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \text{ nas variáveis } \alpha, \beta, \text{ tem solução.} \end{aligned}$$

Para encontrar, em função de x, y, z, condições para a existência de soluções do sistema $\begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, iremos, como de costume, escalonar sua matriz aumentada:

$$\left[\begin{array}{cc|c} 1 & 0 & x \\ -1 & 1 & y \\ 1 & -2 & z \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & x+y \\ 0 & -2 & -x+z \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & x+y \\ 0 & 0 & x+2y+z \end{array}\right]$$

Assim, o sistema tem solução se, e somente se, $x + 2y + z = 0$. Portanto, dado $w = (x, y, z) \in \mathbb{R}^3$, temos que $w \in W$ se, e somente se, $x + 2y + z = 0$. Em outras palavras,

$$W = \big[(1, -1, 1), (0, 1, -2)\big] = \big\{ (x, y, z) \in \mathbb{R}^3 \mid x + 2y + z = 0 \big\}. \quad \diamond$$

Exemplo 4.3.7 Mostre que, em $\mathcal{F}(\mathbb{R})$, temos $\big[\operatorname{sen}^2(x), \cos^2(x)\big] = \big[1, \cos(2x)\big]$.

Solução: Aqui, pede-se para demonstrar que, no espaço vetorial $\mathcal{F}(\mathbb{R})$, de todas as funções reais com domínio a reta toda, os subespaços $W_1 = [f, g]$ e $W_2 = [p, q]$ coincidem, em que as funções $f, g, p, q \colon \mathbb{R} \to \mathbb{R}$ são definidas por

$$f(x) = \operatorname{sen}^2(x), \qquad g(x) = \cos^2(x), \qquad p(x) = 1, \qquad q(x) = \cos(2x),$$

para todo $x \in \mathbb{R}$.

Com efeito, seja $h \in W_1$. Então, existem $\alpha, \beta \in \mathbb{R}$ tais que $h = \alpha f + \beta g$, isto é, $h(x) = \alpha f(x) + \beta g(x) = \alpha \operatorname{sen}^2(x) + \beta \cos^2(x)$, para todo $x \in \mathbb{R}$. Assim, lembrando que $\cos^2(x) = \frac{1}{2}\big(1 + \cos(2x)\big)$, para todo $x \in \mathbb{R}$, temos

$$\begin{aligned}
h(x) &= \alpha \operatorname{sen}^2(x) + \beta \cos^2(x) \\
&= \alpha\big(1 - \cos^2(x)\big) + \beta \cos^2(x) \\
&= \alpha + (-\alpha + \beta) \cos^2(x) \\
&= \alpha + (-\alpha + \beta) \left(\frac{1}{2}\big(1 + \cos(2x)\big)\right) \\
&= \frac{\alpha + \beta}{2} + \frac{-\alpha + \beta}{2} \cos(2x) \\
&= \frac{\alpha + \beta}{2} p(x) + \frac{-\alpha + \beta}{2} q(x),
\end{aligned}$$

para todo $x \in \mathbb{R}$. Portanto, $h = \frac{\alpha+\beta}{2} p + \frac{-\alpha+\beta}{2} q$, o que prova que $h \in W_2$. Logo, $W_1 \subseteq W_2$. Para demonstrar a inclusão reversa, tome $t \in W_2$. Então, existem $\gamma, \delta \in \mathbb{R}$ tais que $t = \gamma p + \delta q$, ou seja, $t(x) = \gamma p(x) + \delta q(x) = \gamma + \delta \cos(2x)$, para todo $x \in \mathbb{R}$. Logo,

$$\begin{aligned}
t(x) &= \gamma + \delta \cos(2x) \\
&= \gamma + \delta\big(-1 + 2\cos^2(x)\big) \\
&= (\gamma - \delta) + 2\delta \cos^2(x) \\
&= (\gamma - \delta)\big(\operatorname{sen}^2(x) + \cos^2(x)\big) + 2\delta \cos^2(x) \\
&= (\gamma - \delta) \operatorname{sen}^2(x) + (\gamma + \delta) \cos^2(x),
\end{aligned}$$

para todo $x \in \mathbb{R}$, donde se conclui que $t = (\gamma - \delta) f + (\gamma + \delta) g \in W_1$. Assim, também temos $W_2 \subseteq W_1$. Comparando as duas inclusões que foram demonstradas, concluímos que $W_1 = W_2$. $\diamond$

Observação Seja V um espaço vetorial, sejam $u_1, u_2, \dots, u_n \in V$ e seja W um subespaço de V. Para mostrar que $[u_1, u_2, \dots, u_n] \subseteq W$, basta mostrar que $u_1, u_2, \dots, u_n \in W$, uma vez que se esses vetores estão em W, então qualquer combinação linear deles (isto é, qualquer elemento do subespaço gerado por eles) também estará em W, já que, sendo um subespaço de V, W é fechado para somas e multiplicações por escalares.

Poderíamos ter utilizado essa observação no exemplo anterior e mostrado apenas que $f \in W_2$ e $g \in W_2$ a fim de garantir $W_1 = [f, g] \subseteq W_2$. Como $W_2 = [p, q]$, mostrar que $f \in W_2$ é mostrar que a função $\operatorname{sen}^2(x)$ se escreve como combinação linear da função constante igual a 1 e da função $\cos(2x)$. Mas isso é, de fato, verdade, uma vez que, para todo $x \in \mathbb{R}$, temos

$$\operatorname{sen}^2(x) = \frac{1}{2}1 + \left(-\frac{1}{2}\right)\cos(2x).$$

Similarmente, mostrar que $g \in W_2$ é mostrar que a função $\cos^2(x)$ é combinação linear da função constante igual a 1 e da função $\cos(2x)$, e isso é verdade, como já vimos:

$$\cos^2(x) = \frac{1}{2}1 + \frac{1}{2}\cos(2x).$$

A outra inclusão, $W_2 \subseteq W_1$, poderia ser mostrada de modo similar, utilizando as identidades

$$1 = 1\operatorname{sen}^2(x) + 1\cos^2(x) \qquad \text{e} \qquad \cos(2x) = (-1)\operatorname{sen}^2(x) + 1\cos^2(x).$$

Dependência linear

Definição Seja V um espaço vetorial e sejam $u_1, u_2, \dots, u_n \in V$. Dizemos que o conjunto $\{u_1, u_2, \dots, u_n\}$ é *linearmente dependente* (LD) se existirem $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$, não todos nulos, tais que

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n = 0_V.$$

Caso contrário, dizemos que $\{u_1, u_2, \dots, u_n\}$ é *linearmente independente* (LI).

Por convenção, o conjunto vazio $\varnothing$ é LI.

Observação Dado um conjunto $\{u_1, u_2, \dots, u_n\}$ de vetores em um espaço vetorial V, é sempre possível escrever o vetor nulo 0_V como combinação linear de $u_1, u_2, \dots, u_n$, basta tomar escalares todos nulos: $0u_1 + 0u_2 + \cdots + 0u_n = 0_V$. Chamemos essa combinação linear de trivial. O que diz a definição acima é que se houver *outra* combinação linear de $u_1, u_2, \dots, u_n$ que resulta no vetor nulo, além da trivial, então o conjunto $\{u_1, u_2, \dots, u_n\}$ é dito LD.

Assim, para mostrar que um conjunto de vetores $\{u_1, u_2, \dots, u_n\}$ em um espaço vetorial V é LI é preciso mostrar que se $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ são tais que

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n = 0_V,$$

então, necessariamente, $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Exemplo 4.3.8 No espaço vetorial $\mathbb{R}^4$ decida se $\{u, v\}$ é LI ou LD, em que $u = (1, 2, 3, 4)$ e $v = (1, 1, 2, 3)$.

Solução: É preciso decidir se os únicos escalares α, β tais que $\alpha u + \beta v = 0_{\mathbb{R}^4}$ são $\alpha = \beta = 0$ (caso em que $\{u, v\}$ é LI) ou não (e $\{u, v\}$ seria LD). Vejamos,

$$\begin{aligned} \alpha u + \beta v = 0_{\mathbb{R}^4} \quad &\iff \quad \alpha(1,2,3,4) + \beta(1,1,2,3) = (0,0,0,0) \\ &\iff \quad \begin{cases} \alpha + \beta = 0 \\ 2\alpha + \beta = 0 \\ 3\alpha + 2\beta = 0 \\ 4\alpha + 3\beta = 0 \end{cases} \end{aligned}$$

Como, evidentemente, $\alpha = \beta = 0$ é a única solução desse sistema linear, $\{u, v\}$ é LI. $\diamond$

Exemplo 4.3.9 No espaço vetorial $\mathbb{R}^3$, decida se $\{(1, 2, 3), (2, 3, 4), (-1, 0, 1)\}$ é LI ou LD.

Solução: Se $\alpha, \beta, \gamma \in \mathbb{R}$ são tais que

$$\alpha(1,2,3) + \beta(2,3,4) + \gamma(-1,0,1) = (0,0,0), \tag{4.2}$$

então $\begin{cases} \alpha + 2\beta - \gamma = 0 \\ 2\alpha + 3\beta = 0 \\ 3\alpha + 4\beta + \gamma = 0 \end{cases}$. Esse sistema linear é indeterminado, uma vez que $\det \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & 0 \\ 3 & 4 & 1 \end{bmatrix} = 0$ (ver Teorema 1.3.6). Portanto, existe pelo menos uma solução não trivial (são infinitas, na verdade) para a equação (4.2), o que implica que $\{(1,2,3),(2,3,4),(-1,0,1)\}$ é LD. Para determinar uma combinação linear não trivial explícita, basta exibir uma solução não trivial. As soluções do sistema linear são da forma $(-3t, 2t, t)$, $t \in \mathbb{R}$. Tomando, por exemplo, $t = 1$, obtemos a seguinte combinação linear não trivial dos três vetores resultando no vetor nulo:

$$-3(1,2,3) + 2(2,3,4) + 1(-1,0,1) = (0,0,0). \quad \diamond$$

Exemplo 4.3.10 No espaço vetorial $\mathcal{P}(\mathbb{R})$, de todos os polinômios, decida se $\{f, g, h\}$ é LI ou LD, em que $f = t^2 + 1, g = t^2 - 1, h = t + 2$. (Neste exemplo, a variável está sendo denotada por t, ao invés de x, como vínhamos fazendo.)

Solução: Sejam $\alpha, \beta, \gamma \in \mathbb{R}$ tais que $\alpha f + \beta g + \gamma h = 0_{\mathcal{P}(\mathbb{R})}$. Então, $\alpha(t^2+1) + \beta(t^2-1) + \gamma(t+2) = 0$, o que é equivalente a $(\alpha+\beta)t^2 + \gamma t + (\alpha - \beta + 2\gamma) = 0$. Em $\mathcal{P}(\mathbb{R})$ isso vale se, e só se, $\begin{cases} \alpha + \beta = 0 \\ \gamma = 0 \\ \alpha - \beta + 2\gamma = 0 \end{cases}$. A única solução desse sistema é $\alpha = \beta = \gamma = 0$. Logo, $\{f, g, h\}$ é LI. $\diamond$

Exemplo 4.3.11 No espaço vetorial $\mathcal{F}(\mathbb{R})$, o conjunto $\{\operatorname{sen}^2(x), \cos^2(x), \cos(2x)\}$ é LD, uma vez que

$$(-1)\operatorname{sen}^2(x) + 1\cos^2(x) + (-1)\cos(2x) = 0,$$

para todo $x \in \mathbb{R}$.

Observação Seja V um espaço vetorial, são consequências imediatas da definição:

- Dado $u \in V$, temos que $\{u\}$ é LD se, e somente se, $u = 0_V$ (uma vez que se $u \neq 0_V$, então $\lambda u = 0_V$ só ocorre quando $\lambda = 0$).
- Dados $u_1, u_2, \dots, u_n \in V$, temos que $\{u_1, u_2, \dots, u_n\}$ é LD se, e somente se, algum dos u_i é combinação linear dos demais. (O mesmo argumento utilizado na demonstração da Proposição 2.2.3 se aplica aqui.)

Vejamos algumas propriedades hereditárias, por assim dizer, da dependência linear.

Proposição 4.3.12 *Seja V um espaço vetorial e sejam $A, B \subseteq V$ dois conjuntos finitos de vetores de V. Valem as seguintes propriedades:*

(i) Se $0_V \in A$, então A é LD.
(ii) Se $A \subseteq B$ e A é LD, então B é LD.
(iii) Se $A \subseteq B$ e B é LI, então A é LI.

Demonstração Para (i), suponha que $A = \{0_V = u_1, u_2, \dots, u_n\}$, então $1u_1 + 0u_2 + \dots 0u_n = 0_V$ e os escalares utilizados nessa combinação linear não são todos nulos; logo, A é LD. É claro que (ii) e (iii) são equivalentes. Mostremos (ii). Suponha que $A = \{u_1, u_2, \dots, u_n\}$ e que $B = \{u_1, u_2, \dots, u_n, u_{n+1}, \dots, u_r\}$. Se A é LD, então existem $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$, não todos nulos, tais que $\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n = 0_V$. Assim,

$$\beta_1 u_1 + \beta_2 u_2 + \dots + \beta_n u_n + \beta_{n+1} u_{n+1} + \dots + \beta_r u_r = 0_V,$$

em que $\beta_i = \alpha_i$, para $i = 1, \dots, n$, e $\beta_{n+1} = \dots = \beta_r = 0$. Como os escalares utilizados nessa combinação linear não são todos nulos (algum β_i com $1 \leq i \leq n$ é não nulo), B é LD. $\square$

Relações de dependência linear em espaços de funções, como a do Exemplo 4.3.11, são incomuns. No Exercício 13 da Lista 1 de Álgebra Linear II, um método para lidar com a independência linear entre funções é esboçado. No exemplo abaixo, apresentamos mais uma técnica que pode ser útil na demonstração da independência linear em espaços de funções muitas vezes diferenciáveis.

Exemplo 4.3.13 Mostre que $\{e^t, e^{2t}, e^{3t}\}$ é LI no espaço vetorial $\mathcal{F}(\mathbb{R})$.

Solução: Para deixar o argumento mais claro, as funções envolvidas serão denotadas por

$$f\colon \mathbb{R} \longrightarrow \mathbb{R}, \qquad g\colon \mathbb{R} \longrightarrow \mathbb{R}, \qquad h\colon \mathbb{R} \longrightarrow \mathbb{R}.$$
$$t \longmapsto e^t \qquad\qquad t \longmapsto e^{2t} \qquad\qquad t \longmapsto e^{3t}$$

Nossa tarefa consiste em mostrar que se $\alpha, \beta, \gamma \in \mathbb{R}$ são tais que $\alpha f + \beta g + \gamma h = 0_{\mathcal{F}(\mathbb{R})}$, então $\alpha = \beta = \gamma = 0$. Como o vetor nulo $0_{\mathcal{F}(\mathbb{R})}$ do espaço vetorial $\mathcal{F}(\mathbb{R})$ é a função identicamente nula, precisamos mostrar que os únicos valores de α, β, γ que satisfazem

$$\alpha f(t) + \beta g(t) + \gamma h(t) = 0, \quad \text{para todo } t \in \mathbb{R},$$

são $\alpha = \beta = \gamma = 0$. Ou seja, que

$$\alpha e^t + \beta e^{2t} + \gamma e^{3t} = 0, \quad \text{para todo } t \in \mathbb{R}, \tag{4.3}$$

implica $\alpha = \beta = \gamma = 0$. Mas, (4.3) implica, em particular, fazendo $t = 0$, que

$$\alpha + \beta + \gamma = 0. \tag{4.4}$$

Agora, (4.3) é uma igualdade entre funções, a função $\alpha f + \beta g + \gamma h$, do lado esquerdo, e a função nula, do lado direito. Como ambas são diferenciáveis, suas derivadas devem também coincidir. Assim, derivando (4.3), obtemos

$$\alpha e^t + 2\beta e^{2t} + 3\gamma e^{3t} = 0, \quad \text{para todo } t \in \mathbb{R}. \tag{4.5}$$

Em particular, se $t = 0$,

$$\alpha + 2\beta + 3\gamma = 0. \tag{4.6}$$

Derivando, agora, (4.5), obtemos

$$\alpha e^t + 4\beta e^{2t} + 9\gamma e^{3t} = 0, \quad \text{para todo } t \in \mathbb{R}, \tag{4.7}$$

que, avaliada em $t = 0$, fonece

$$\alpha + 4\beta + 9\gamma = 0. \tag{4.8}$$

Dessa discussão, segue que α, β, γ são números reais que satisfazem (juntando (4.4), (4.6) e (4.8))

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{4.9}$$

Como $\det \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix} = 2 \neq 0$, a única solução de (4.9) é a trivial: $\alpha = \beta = \gamma = 0$. Segue que $\{e^t, e^{2t}, e^{3t}\}$ é LI. $\diamond$

Observação A ideia explorada no exemplo acima pode ser generatizada: se $\lambda_1, \lambda_2, \dots, \lambda_n$ são n números reais distintos, então o conjunto $\{e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t}\}$ de n vetores de $\mathcal{F}(\mathbb{R})$ é LI. Basta argumentar avaliando uma combinação linear desses vetores e das $n-1$ primeira derivadas dela em $t = 0$. O argumento funciona porque

$$\det \begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_n^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{bmatrix}$$

é igual ao produto de todos os termos da forma $(\lambda_j - \lambda_i)$, com $1 \leq i < j \leq n$, que, sendo os λ_i todos distintos, é diferente de zero. Esse determinante chama-se *determinante de Vandermonde*. Uma prova da igualdade mencionada pode ser facilmente encontrada na internet, por exemplo, em `https://en.wikipedia.org/wiki/Vandermonde_matrix`.

Exercícios Lista 1 - Álgebra Linear II: Exs. 6–16.

4.4 Bases e dimensão

Na seção anterior, vimos os conceitos de subespaço gerado por um subconjunto e independência (e dependência) linear. Essas ideias serão conjugadas para obtermos o conceito de base de um espaço vetorial.

Definição Seja V um espaço vetorial e seja $\mathcal{B}$ um conjunto finito formado por vetores de V. Dizemos que $\mathcal{B}$ é uma *base* de V se as duas seguintes condições estão satisfeitas:

(i) $\mathcal{B}$ gera V, isto é, $[\mathcal{B}] = V$, e
(ii) $\mathcal{B}$ é LI.

Como vimos na seção anterior, dado um subconjunto finito $\mathcal{B}$ de um espaço vetorial, sempre podemos considerar o subespaço gerado por dele, que denotamos por $[\mathcal{B}]$ e que é formado por todas as combinações lineares dos elementos de $\mathcal{B}$. O que a condição (i) na definição de base exige é que esse subespaço coincida com todo o espaço vetorial V. Assim, uma base de um espaço vetorial V nada mais é do que um conjunto finito que é LI e que esgota, tomando-se combinações lineares de seus elementos, todos os elementos de V.

Comecemos com alguns exemplos para explorar essa definição.

Exemplo 4.4.1 O conjunto $\{(1,0,0),(0,1,0),(0,0,1)\}$ é uma base de $\mathbb{R}^3$. Com efeito, é claro que esse conjunto é LI. Além disso, esse conjunto gera $\mathbb{R}^3$, pois todo vetor de $\mathbb{R}^3$ pode ser escrito como combinação linear dos três vetores no conjunto:

$$(a,b,c) = a(1,0,0) + b(0,1,0) + c(0,0,1).$$

De maneira análoga, é fácil ver que o conjunto de n vetores $\{e_1, e_2, \dots, e_n\}$, em que, para cada $i = 1, \dots, n$,

$$e_i = (0,0,\dots,0,\underset{i}{1},0,0,\dots,0)$$

é uma base de $\mathbb{R}^n$. Essa base é chamada *base canônica* de $\mathbb{R}^n$.

Exemplo 4.4.2 O conjunto de $n+1$ vetores $\{1, t, t^2, \dots, t^n\}$ é uma base de $\mathcal{P}_n(\mathbb{R})$, pois ele é claramente gerador e LI.

Exemplo 4.4.3 O conjunto de 4 vetores

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

é uma base de $M_2(\mathbb{R})$.

As bases dos dois exemplos anteriores também são, às vezes, chamadas de bases canônicas de seus respectivos espaços vetoriais.

Exemplo 4.4.4 O conjunto $\mathcal{B} = \{(1,2,3),(0,2,2),(1,0,-1)\}$ é uma base de $\mathbb{R}^3$. De fato, mostremos que $\mathcal{B}$ gera $\mathbb{R}^3$ e que $\mathcal{B}$ é LI.

(i) $[\mathcal{B}] = \mathbb{R}^3$: aqui, é preciso mostrar que todo vetor de $\mathbb{R}^3$ é combinação linear de $u_1 = (1,2,3), u_2 = (0,2,2), u_3 = (1,0,-1)$. Para tanto, é preciso mostrar que, dado $v = (x,y,z) \in \mathbb{R}^3$, existem $\alpha, \beta, \gamma \in \mathbb{R}^3$ tais que $v = \alpha u_1 + \beta u_2 + \gamma u_3$, ou seja, $(x,y,z) = \alpha(1,2,3) + \beta(0,2,2) + \gamma(1,0,-1) = (\alpha+\gamma, 2\alpha+2\beta, 3\alpha+2\beta-\gamma)$. Assim, mostrar que $\mathcal{B}$ gera $\mathbb{R}^3$ é mostrar que, dados $x,y,z \in \mathbb{R}$, o sistema linear

$$\begin{cases} \alpha + \gamma = x \\ 2\alpha + 2\beta = y \\ 3\alpha + 2\beta - \gamma = z \end{cases} \tag{4.10}$$

nas variáveis α, β, γ, tem solução. Como $\det \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 0 \\ 3 & 2 & -1 \end{bmatrix} = -4 \neq 0$, segue do Teorema 1.3.6 que o sistema (4.10) é possível determinado; em particular, tem solução. Portanto, de fato, $[\mathcal{B}] = \mathbb{R}^3$. (Na realidade, o Teorema 1.3.6 nos garante que (4.10) tem uma única solução. Pelo método do escalonamento, por exemplo, pode-se determinar que a única solução de (4.10) é dada por $\alpha = \frac{x-y+z}{2}, \beta = \frac{-x+2y-z}{2}, \gamma = \frac{x+y-z}{2}$. Em outras palavras, se $v = (x, y, z)$, então $v = (\frac{x-y+z}{2})u_1 + (\frac{-x+2y-z}{2})u_2 + (\frac{x+y-z}{2})u_3$.)

(ii) $\mathcal{B}$ é LI: aqui, o objetivo é mostrar que se $\alpha, \beta, \gamma \in \mathbb{R}$ são tais que $\alpha u_1 + \beta u_2 + \gamma u_3 = (0, 0, 0)$, então necessariamente, $\alpha = \beta = \gamma = 0$. Isso é equivalente a mostrar que o sistema (4.10) com $x = y = z = 0$ é determinado. Mas uma vez, esse fato segue do Teorema 1.3.6.

No exemplo acima, vimos que essencialmente a mesma razão garantiu que o conjunto $\mathcal{B}$ era gerador e LI. Veremos que essa correlação sempre estará presente quando o conjunto em questão tiver a quantidade certa, num sentido a ser precisado, de elementos (cf. Coroláro 4.4.15).

Exemplo 4.4.5 O conjunto $\{(1, 0, 0), (0, 1, 0)\}$ não é uma base de $\mathbb{R}^3$, porque, apesar de ser LI (como é fácil ver), não gera $\mathbb{R}^3$; por exemplo, $(0, 0, 1)$ não é combinação linear de $(1, 0, 0)$ e $(0, 1, 0)$.

Exemplo 4.4.6 O conjunto $\{(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 2, 3)\}$ não é uma base de $\mathbb{R}^3$, porque, apesar de gerar $\mathbb{R}^3$ (todo vetor de $\mathbb{R}^3$ é combinação linear dos três primeiros, e, portanto, dos quatro vetores), não é LI: $(1, 2, 3) = 1(1, 0, 0) + 2(0, 1, 0) + 3(0, 0, 1)$.

Todo espaço vetorial que pode ser gerado por um conjunto finito de vetores tem, pelo menos, uma base. Isso é consequência do Teorema 4.4.8, que veremos a seguir.

O lema abaixo, necessário na demonstração do teorema, é uma consequência quase imediata da definição de subespaço gerado.

Lema 4.4.7 *Seja V um espaço vetorial e sejam $u_1, u_2, \dots, u_n \in V$. Para todo $v \in V$, temos:*

$$v \in [u_1, \dots, u_n] \quad \textit{se, e somente se,} \quad [v, u_1, \dots, u_n] = [u_1, \dots, u_n]\,.$$

Em outras palavras, o lema afirma que um vetor que é combinação linear de outros vetores de um conjunto gerador de um subespaço pode ser omitido desse conjunto gerador sem afetar o subespaço gerado.

Demonstração Como sempre vale que $v \in [v, u_1, \dots, u_n]$, se $[v, u_1, \dots, u_n] = [u_1, \dots, u_n]$, então $v \in [u_1, \dots, u_n]$.

Por outro, lado, a inclusão $[u_1, \dots, u_n] \subseteq [v, u_1, \dots, u_n]$ é sempre válida. Agora, se $v \in [u_1, \dots, u_n]$, como $u_1, \dots, u_n \in [u_1, \dots, u_n]$, segue que $[v, u_1, \dots, u_n] \subseteq [u_1, \dots, u_n]$. □

Teorema 4.4.8 *Seja V um espaço vetorial e sejam $u_1, \dots, u_n \in V$. Se $V = [u_1, \dots, u_n]$, então existe uma base $\mathcal{B}$ de V tal que $\mathcal{B} \subseteq \{u_1, \dots, u_n\}$.*

Demonstração Se o conjunto $\mathcal{A} = \{u_1, \dots, u_n\}$ é LI, então ele é uma base e não há o que demonstrar. Suponha, então, que $\mathcal{A}$ é LD. Assim, algum u_j é combinação linear dos demais, isto é, $u_j \in \left[u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_n\right]$. Daí segue, do Lema 4.4.7, que

$$\left[u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_n\right] = [\mathcal{A}] = V.$$

Se $\mathcal{A}_1 = \{u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_n\}$ for LI, esta será uma base de V contida em $\mathcal{A}$. Se não, repetimos o argumento com $\mathcal{A}_1$ no lugar de $\mathcal{A}$. Desse modo, obteremos uma cadeia estritamente descendente

$$\mathcal{A} \supsetneq \mathcal{A}_1 \supsetneq \mathcal{A}_2 \supsetneq \dots$$

de subconjuntos de $\mathcal{A}$, todos gerando V. Como esses conjuntos são todos finitos e diminuem de tamanho ao longo da cadeia, em algum ponto, a cadeia estacionará em um conjunto que é também LI (já que o conjunto vazio é LI). Nesse ponto, teremos uma base de V contida em $\mathcal{A}$. □

A demonstração do teorema fornece um método para obter uma base contida em um conjunto gerado. Vejamos um exemplo.

Exemplo 4.4.9 Mostre que o conjunto $\mathcal{A} = \{(1,0),(1,1),(1,2)\}$ gera o espaço vetorial $\mathbb{R}^2$ e exiba uma base de $\mathbb{R}^2$ contida nele.

Solução: Vejamos, primeiramente, que $\mathcal{A}$ gera $\mathbb{R}^2$. Para tanto é preciso mostrar que todo vetor de $\mathbb{R}^2$ é combinação linear dos elementos de $\mathcal{A}$. Dado $(a,b) \in \mathbb{R}^2$, temos que $(a,b) \in [\mathcal{A}]$ se, e somente se, existirem $x, y, z \in \mathbb{R}$ tais que $(a,b) = x(1,0) + y(1,1) + x(1,2)$, ou, equivalentemente, x, y, z satisfazerem

$$\begin{cases} x + y + z = a \\ y + 2z = b \end{cases} \tag{4.11}$$

Como a matriz aumentada $\left[\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 0 & 1 & 2 & b \end{array}\right]$ desse sistema já é escalonada, vemos que, para quaisquer $a, b \in \mathbb{R}$, o sistema (4.11) tem solução. Logo, $\mathcal{A}$, de fato, gera $\mathbb{R}^2$. (As soluções de (4.11) são dadas por $(a - b + t, b - 2t, t)$, $t \in \mathbb{R}$; portanto, uma decomposição de (a,b) como combinação linear dos elementos de $\mathcal{A}$ é, por exemplo,

$$(a,b) = (a - b + 1)(1,0) + (b - 2)(1,1) + 1(1,2),$$

em que tomamos $t = 1$. Existem infinitas decomposições como essa, uma para cada escolha do parâmetro t.)

O conjunto $\mathcal{A}$ é LD, pois olhando para (4.11) com $a = b = 0$, vemos que existem combinações lineares não triviais dos elementos de $\mathcal{A}$ que resultam no vetor nulo de $\mathbb{R}^2$, por exemplo, tomando a solução com $t = 2$,

$$2(1,0) + (-4)(1,1) + 2(1,2) = (0,0). \tag{4.12}$$

Assim, $\mathcal{A}$ não é uma base de $\mathbb{R}^2$, mas, pelo Teorema 4.4.8, $\mathcal{A}$ contém uma. A relação de dependência linear (4.12) permite que se escreva, por exemplo, $(1,0)$ como combinação linear de $(1,1)$ e $(1,2)$:

$$(1,0) = \frac{1}{2}\big(4(1,1) + (-2)(1,2)\big) = 2(1,1) - 1(1,2).$$

Assim, pelo Lema 4.4.7, $[(1,1),(1,2)] = [\mathcal{A}] = \mathbb{R}^2$. Ou seja, é possível reduzir o conjunto gerador $\mathcal{A}$ de $\mathbb{R}^2$ para um conjunto que continua gerando $\mathbb{R}^2$, está contido em $\mathcal{A}$, mas tem um elemento a menos. Agora, esse novo conjunto $\mathcal{A}_1 = \{(1,1),(1,2)\}$ é LI, uma vez que $\alpha(1,1) + \beta(1,2) = (0,0)$ implica $\alpha = \beta = 0$, como o leitor pode facilmente verificar. Portanto $\mathcal{A}_1$ é uma base de $\mathbb{R}^2$, que está contida em $\mathcal{A}$. (No Apêndice A, apresenta-se um método mais eficiente para se extrair uma base de um conjunto gerador.) ◇

Dimensão

Nosso objetivo, agora, será mostrar que duas bases de um mesmo espaço vetorial têm sempre a mesma quantidade de elementos. Esse fato seguirá como consequência do resultado a seguir.

Teorema 4.4.10 *Seja V um espaço vetorial, sejam $\mathcal{A} = \{u_1, \dots, u_n\}$ e $\mathcal{B} = \{v_1, \dots, v_m\}$ subconjuntos de V. Se $V = [\mathcal{A}]$ e $\mathcal{B}$ é LI, então $m \leq n$.*

Demonstração A demonstração será feita por contradição. Suponha que $m > n$. Como $\mathcal{A}$ é um conjunto gerador de V, todo vetor de V se escreve como combinação linear dos elementos de $\mathcal{A}$, em particular, isso se aplica aos elementos de $\mathcal{B}$: para cada $j = 1, \dots, m$, existem $\alpha_{ij} \in \mathbb{R}$ $(1 \leq i \leq n)$ tais que

$$v_j = \sum_{i=1}^{n} \alpha_{ij} u_i.$$

Considere o sistema linear homogêneo de n equações e m incógnitas cujos coeficientes são os escalares α_{ij}:

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1m}x_m = 0 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2m}x_m = 0 \\ \quad \vdots \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \cdots + \alpha_{nm}x_m = 0 \end{cases} \tag{4.13}$$

Como $m > n$, pelo Teorema 1.1.8, esse sistema tem pelo menos uma solução não trivial $(\beta_1, \beta_2, \ldots, \beta_m)$. Assim,

$$\begin{aligned} \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_m v_m &= \beta_1 \left(\sum_{i=1}^{n} \alpha_{i1} u_i \right) + \beta_2 \left(\sum_{i=1}^{n} \alpha_{i2} u_i \right) + \cdots + \beta_m \left(\sum_{i=1}^{n} \alpha_{im} u_i \right) \\ &= \left(\sum_{j=1}^{m} \alpha_{1j}\beta_j \right) u_1 + \left(\sum_{j=1}^{m} \alpha_{2j}\beta_j \right) u_2 + \cdots + \left(\sum_{j=1}^{m} \alpha_{nj}\beta_j \right) u_n \end{aligned}$$

Como $(\beta_1, \beta_2, \ldots, \beta_m)$ é uma solução de (4.13), cada um dos coeficientes dos u_i na combinação linear acima é zero. Segue que

$$\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_m v_m = 0_V.$$

Mas os escalares β_j não são todos nulos, pois formam uma solução não trivial de (4.13). Portanto, $\mathcal{B}$ é LD, o que é uma contradição. Logo, $m \leq n$. □

Corolário 4.4.11 *Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ duas bases de um espaço vetorial V. Então, $\mathcal{B}_1$ e $\mathcal{B}_2$ têm a mesma quantidade de elementos.*

Demonstração Suponha que $\mathcal{B}_1$ tenha n elementos e que $\mathcal{B}_2$ tenha m elementos. Como $V = [\mathcal{B}_1]$ e $\mathcal{B}_2$ é LI, segue, do Teorema 4.4.10, que $m \leq n$. Por outro lado, como $V = [\mathcal{B}_2]$ e $\mathcal{B}_1$ é LI, segue, do mesmo resultado, que $n \leq m$. Logo, $n = m$. □

Esse resultado justifica a próxima definição.

Definição Se V é um espaço vetorial que tem uma base com n elementos, diremos que V tem *dimensão n* e escreveremos $\dim(V) = n$. Por convenção, o espaço vetorial nulo tem dimensão 0.

Exemplo 4.4.12 Segue, pelo que vimos nos exemplos no início desta seção, que

- $\dim(\mathbb{R}^n) = n$,
- $\dim\big(\mathcal{P}_n(\mathbb{R})\big) = n + 1$ e
- $\dim\big(M_2(\mathbb{R})\big) = 4$.

Nesse último caso, mais geralmente, $\dim\big(M_{m\times n}(\mathbb{R})\big) = mn$. Uma base de $M_{m\times n}(\mathbb{R})$ é $\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$, em que E_{ij} denota a matriz $m \times n$ com entrada (i, j) igual a 1 e as demais, todas, nulas.

Como vimos no Teorema 4.4.8, todo espaço vetorial que pode ser gerado por um conjunto finito de vetores tem uma base. Esses espaços serão chamado de *espaços vetoriais de dimensão finita*.

Terminamos esta seção com um resultado que é, em certo sentido, dual do Teorema 4.4.8. Aqui, também, precisaremos de um lema preparatório.

Lema 4.4.13 *Seja V um espaço vetorial e seja $\{v_1, \ldots, v_m\}$ um conjunto LI de vetores de V. Se $w \in V$ é tal que $\{w, v_1, \ldots, v_m\}$ é LD, então $w \in [v_1, \ldots, v_m]$.*

Demonstração Como $\{w, v_1, \ldots, v_m\}$ é LD, existem $\beta, \alpha_1, \ldots, \alpha_m \in \mathbb{R}$, não todos nulos, tais que $\beta w + \alpha_1 v_1 + \cdots + \alpha_m v_m = 0_V$. Se $\beta = 0$, teríamos algum $\alpha_i \neq 0$ e $\alpha_1 v_1 + \cdots + \alpha_m v_m = 0_V$, o que é impossível, uma vez que $\{v_1, \ldots, v_m\}$ é LI. Logo, $\beta \neq 0$ e, portanto, $w = -\frac{\alpha_1}{\beta} v_1 - \cdots - \frac{\alpha_m}{\beta} v_m \in [v_1, \ldots, v_m]$. □

Teorema 4.4.14 *Seja V um espaço vetorial de dimensão finita e seja $\mathcal{B}$ um subconjunto LI de V. Então, V tem uma base contendo $\mathcal{B}$.*

Esse resultado é também conhecido como *Teorema do completamento* , uma vez que garante que qualquer conjunto LI em um espaço vetorial de dimensão finita pode ser completado a uma base.

Demonstração Suponha que $\dim(V) = n$ e seja $\mathcal{A} = \{v_1, \ldots, v_m\}$ um conjunto LI formado por vetores de V. Como V tem uma base com n elementos, segue do Teorema 4.4.10 que $m \leq n$. Se $[\mathcal{A}] = V$, então $\mathcal{A}$ é uma base de V e não há o que demonstrar. Suponha que $[\mathcal{A}] \neq V$. Então, existe um vetor $w_1 \in V$ tal que $w_1 \notin [\mathcal{A}]$. Pelo Lema 4.4.13, o conjunto $\mathcal{A}_1 = \{w_1, v_1, \ldots, v_m\}$ é LI. (Se fosse LD, teríamos $w_1 \in [\mathcal{A}]$.) Consideremos, agora, o conjunto $\mathcal{A}_1$. Ele é LI. Se $[\mathcal{A}_1] = V$, então $\mathcal{A}_1$ será uma base de V que contém $\mathcal{A}$. Se não, repita o argumento com $\mathcal{A}_1$ no lugar de $\mathcal{A}$. Esse processo termina em uma base (que contém $\mathcal{A}$), porque em V não existem conjuntos LI com mais do que n vetores. □

Corolário 4.4.15 *Seja V um espaço vetorial de dimensão finita igual a n.*

(i) Se $\mathcal{B}$ é um conjunto gerador de V com n elementos, então $\mathcal{B}$ é uma base de V.
(ii) Se $\mathcal{B}$ é um subconjunto LI de V com n elementos, então $\mathcal{B}$ é uma base de V.

Demonstração (i) Pelo Teorema 4.4.8, existe uma base C de V tal que $C \subseteq \mathcal{B}$. Como C tem n elementos (pois $\dim(V) = n$), segue que $C = \mathcal{B}$. (ii) Pelo Teorema 4.4.14, exise uma base C de V tal que $\mathcal{B} \subseteq C$. Como C tem n elementos, $\mathcal{B} = C$. □

O corolário que acabamos de ver é útil na tarefa de se exibir bases, como ilustra o próximo exemplo.

Exemplo 4.4.16 Mostre que $\mathcal{B} = \{v_1, v_2, v_3, v_4\}$, em que

$$v_1 = (1, 1, 0, -4), \quad v_2 = (3, -1, 2, 1), \quad v_3 = (-2, -1, 1, 1), \quad v_4 = (0, 1, -2, 2),$$

é uma base de $\mathbb{R}^4$.

Solução: Vejamos se $\mathcal{B}$ é LI. Sejam $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$. Então, $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 = 0_{\mathbb{R}^4}$ se, e somente se,

$$\begin{bmatrix} 1 & 3 & -2 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 2 & 1 & -2 \\ -4 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Como $\det \begin{bmatrix} 1 & 3 & -2 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 2 & 1 & -2 \\ -4 & 1 & 1 & 2 \end{bmatrix} = -9 \neq 0$, segue do Teorema 1.3.6 que, necessariamente, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$. Logo, $\mathcal{B}$ é um conjunto LI em $\mathbb{R}^4$. Como $\dim(\mathbb{R}^4) = 4$, e $\mathcal{B}$ é um conjunto LI formado por 4 vetores de $\mathbb{R}^4$, segue, do Corolário 4.4.15, que $\mathcal{B}$ é uma base de $\mathbb{R}^4$. ◇

Exercícios Lista 1 - Álgebra Linear II: Exs. 17–24.

4.5 Coordenadas

Se $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ é uma base do espaço vetorial V e $v \in V$, então, como $\mathcal{B}$ gera V, existem $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tais que

$$v = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n. \tag{4.14}$$

Como $\mathcal{B}$ é LI, a sequência $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ dos escalares utilizados na expressão (4.14) está univocamente determinada pelos elementos de $\mathcal{B}$ e pela ordem em que eles estão dados. Com efeito, se $\beta_1, \beta_2, \ldots, \beta_n \in \mathbb{R}$ são tais que

$$v = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n,$$

então, igualando ambas as expressões de v como combinação linear dos elementos de $\mathcal{B}$, obtém-se

$$(\alpha_1 - \beta_1)u_1 + (\alpha_2 - \beta_2)u_2 + \cdots + (\alpha_n - \beta_n)u_n = 0_V.$$

Como $\mathcal{B}$ é LI, segue que $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$.

Uma vez fixada uma ordem nos elementos de uma base $\mathcal{B}$ de um espaço vetorial V, dizemos que $\mathcal{B}$ é uma *base ordenada* de V.

O que acabamos de ver permite que se definam coordenadas em relação a uma base ordenada.

Definição Seja V um espaço vetorial de dimensão finita e seja $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ uma base ordenada de V. Dado $v \in V$, sejam $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ tais que

$$v = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n.$$

Dizemos que $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ são as *coordenadas* de v em relação à base ordenada $\mathcal{B}$, e escrevemos

$$v = (\alpha_1, \alpha_2, \ldots, \alpha_n)_{\mathcal{B}}.$$

Observação Cabe um comentário sobre as semelhanças e diferenças entre a nomenclatura e notação para bases e coordenadas utilizadas no Capítulo 2 em comparação com as utilizadas no contexto de espaços vetoriais abstratos. Para o espaço vetorial $\mathbb{V}^3$ dos vetores tridimensionais, definimos uma base como sendo um conjunto ordenado LI composto por três vetores. Essas bases, no contexto abstrato, são o que estamos, a partir de agora, chamando de bases ordenadas. Repare, também, que o conceito de coordenadas visto acima coincide com aquele do Capítulo 2.

Tem-se uma versão para espaços vetoriais gerais da Proposição 2.3.2, que nos permitia "fazer contas em coordenadas":

Proposição 4.5.1 *Seja V um espaço vetorial de dimensão finita igual a n e seja $\mathcal{B}$ uma base ordenada de V. Sejam $v, w \in V$ com $v = (\alpha_1, \alpha_2, \ldots, \alpha_n)_{\mathcal{B}}$ e $w = (\beta_1, \beta_2, \ldots, \beta_n)_{\mathcal{B}}$, e seja $\lambda \in \mathbb{R}$. Então,*

(i) $v + w = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)_{\mathcal{B}}$;
(ii) $\lambda v = (\lambda\alpha_1, \lambda\alpha_2, \ldots, \lambda\alpha_n)_{\mathcal{B}}$.

Demonstração Suponha que $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$. Como $v = \alpha_1 u_1 + \cdots + \alpha_n u_n$ e $w = \beta_1 u_1 + \cdots + \beta_n u_n$, segue que $v + w = (\alpha_1 + \beta_1)u_1 + \cdots + (\alpha_n + \beta_n)u_n$ e $\lambda v = (\lambda\alpha_1)u_1 + \ldots (\lambda\alpha_n)u_n$. □

Exemplo 4.5.2 (Prova 3, Álgebra Linear I, 2015) Considere a base ordenada $\mathcal{B}$ de $M_2(\mathbb{R})$ dada por

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\}.$$

Se (a, b, c, d) denotam as coordenadas da matriz $\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$ em relação à base $\mathcal{B}$, então $8a + 4b + 2c + d$ é igual a

(A) 3 (B) 5 (C) 8 (D) 11 (E) 15

Solução: Fica a cargo do leitor convencer-se de que $\mathcal{B}$ é, de fato, uma base de $M_2(\mathbb{R})$. Temos que $\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} = (a, b, c, d)_{\mathcal{B}}$ se, e somente se,

$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} = a \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + b \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} + c \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a-b+c+d & a+b+c+d \\ a+b+c & a-b+d \end{bmatrix},$$

ou seja, se e somente se, $\begin{cases} a-b+c+d=1 \\ a+b+c+d=2 \\ a+b+c=3 \\ a-b+d=0 \end{cases}$. A solução desse sistema é dada por $a = \frac{3}{2}, b = \frac{1}{2}, c = 1, d = -1$.

Resposta: (E). $\diamond$

Exercícios Lista 1 - Álgebra Linear II: Exs. 25–27.

4.6 Base e dimensão de subespaços

Como vimos na Seção 4.2, um subespaço de um espaço vetorial é, ele também, um espaço vetorial. Nesta seção, aplicaremos as ideias de base e dimensão ao estudo de subespaços de espaços vetoriais.

Começamos com as seguinte consequência dos resultados da Seção 4.4.

Proposição 4.6.1 *Seja V um espaço vetorial de dimensão finita e seja W um subespaço de V. Então, W tem dimensão finita e* $\dim(W) \leq \dim(V)$. *Além disso,* $\dim(W) = \dim(V)$ *se, e somente se,* $W = V$.

Demonstração Se $W = \{0_V\}$, então $\dim(W) = 0 \leq \dim(V)$. Se $W \neq \{0_V\}$, tome um vetor $w_1 \in W$, $w_1 \neq 0_V$. Então, o conjunto $\{w_1\}$ é LI, o que implica, pelo Teorema 4.4.10, que $\dim(V) \geq 1$. Como W é um subespaço de V, temos $[w_1] \subseteq W$. Se $[w_1] = W$, então $\{w_1\}$ é uma base de W, e, portanto, $\dim(W) = 1 \leq \dim(V)$. Se $[w_1] \neq W$, então existe $w_2 \in W$ tal que $w_2 \neq [w_1]$. Do Lema 4.4.13, segue que $\{w_1, w_2\}$ é um subconjunto LI de W e, *a fortiori*, de V. Assim, usando novamente o Teorema 4.4.10, concluímos que $\dim(V) \geq 2$. Sabemos que $[w_1, w_2] \subseteq W$ (porque W é um subespaço de V). Se $[w_1, w_2] = W$, o conjunto $\{w_1, w_2\}$ é uma base de W e $\dim(W) = 2 \leq \dim(V)$. Se $[w_1, w_2] \neq W$, repetimos o argumento. Esse processo não pode ser repetido indefinidamente, pois, a cada passo, construímos em W, e, portanto, em V, um conjunto LI maior, mas os subconjuntos LI de V têm tamanho limitado por $\dim(V)$.

É claro que se $W = V$, então $\dim(W) = \dim(V)$. Reciprocamente, suponha que $n = \dim(W) = \dim(V)$ e seja $\mathcal{B}$ uma base de W. Então $\mathcal{B}$ é um subconjunto LI de V contendo $n = \dim(V)$ elementos. Pelo Corolário 4.4.15, $\mathcal{B}$ é uma base de V. Logo, $V = [\mathcal{B}] = W$. $\square$

Dedicaremos o resto desta seção ao estudo de exemplos em que bases e dimensões de subespaços são determinados.

Nos exemplos que seguem, faremos uso da seguinte observação útil na determinação de bases de subespaços por meio do processo de escalonamento.

Observação Seja V um espaço vetorial de dimensão finita, digamos, $\dim(V) = n$, e seja $\mathcal{B}$ uma base ordenada de V. Dados vetores $u_1, \dots, u_k$ de V, podemos considerar o subespaço $W = [u_1, \dots, u_k]$. Para encontrar uma base para W (e, portanto, determinar sua dimensão), podemos proceder da seguinte maneira: determine as coordenadas dos vetores que geram W em relação à base ordenada $\mathcal{B}$,

$$u_i = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in})_{\mathcal{B}},$$

e considere a matriz $A = (\alpha_{ij}) \in M_{k\times n}(\mathbb{R})$, cujas linhas são as coordenadas, em relação a $\mathcal{B}$, dos geradores de W. Agora, seja $R = (\beta_{ij}) \in M_{k\times n}(\mathbb{R})$ uma matriz que foi obtida a partir de A por meio de uma sequência de operações elementares sobre linhas. Note que as linhas de R são combinações lineares das linhas de A. Portanto, segue da Proposição 4.5.1, que se definirmos os vetores

$$w_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{in})_{\mathcal{B}},$$

cujas coordenadas, em relação à base $\mathcal{B}$ são dadas pelas linhas de R, então $w_i \in [u_1, \dots, u_k] = W$. Logo, $[w_1, \dots, w_k] \subseteq W$. Como A também pode ser obtida a partir de R por meio de uma sequência de operações elementares sobre linhas, o mesmo argumento prova que $W = [u_1, \dots, u_k] \subseteq [w_1, \dots, w_k]$. Portanto, $W = [w_1, \dots, w_k]$.

Agora, suponha que R seja uma matriz escalonada (por exemplo, que tenha sido obtida a partir de A pelo processo de escalonamento). Neste caso, W será gerado pelos vetores cujas coordenadas, em relação à base $\mathcal{B}$, são dadas pelas linhas *não nulas* de R (pois as linhas nulas correspondem ao vetor nulo, que sempre pode ser eliminado de um conjunto gerador). Mais ainda, segue novamente da Proposição 4.5.1, que os vetores cujas coordenadas correspondem às linhas não nulas de R formam um conjunto LI de vetores (uma vez que os pivôs ocorrem em posições distintas); logo formam uma base de W. Assim, $\dim(W)$ coincide com o número de linhas não nulas de R, que é o número de pivôs de R.

Exemplo 4.6.2 Em $\mathcal{P}_2(\mathbb{R})$, determine a dimensão do subespaço $W = \left[(x-1)^2, (x-2)^2, (x-3)^2\right]$.

Solução: Sabemos, da Proposição 4.6.1, que $\dim(W) \leq \dim\big(\mathcal{P}_2(\mathbb{R})\big) = 3$. Vamos utilizar a observação acima para encontrar uma base, e consequentemente, a dimensão, de W. Comecemos por escolher uma base ordenada para $\mathcal{P}_2(\mathbb{R})$, por exemplo, $\mathcal{B} = \{1, x, x^2\}$, e determinemos as coordenadas, em relação a essa base, dos geradores de W:

$$\begin{aligned}(x-1)^2 &= 1 - 2x + x^2 = (1, -2, 1)_{\mathcal{B}}\\ (x-2)^2 &= 4 - 4x + x^2 = (4, -4, 1)_{\mathcal{B}}\\ (x-3)^2 &= 9 - 5x + x^2 = (9, -6, 1)_{\mathcal{B}},\end{aligned}$$

ou seja, $W = [(1, -2, 1)_{\mathcal{B}}, (4, -4, 1)_{\mathcal{B}}, (9, -6, 1)_{\mathcal{B}}]$. Agora, considere a matriz cujas linhas são as coordenadas dos geradores de W e submeta-a ao processo de escalonamento:

$$\begin{bmatrix} 1 & -2 & 1\\ 4 & -4 & 1\\ 9 & -6 & 1\end{bmatrix} \to \begin{bmatrix} 1 & -2 & 1\\ 0 & 4 & -3\\ 0 & 12 & -8\end{bmatrix} \to \begin{bmatrix} 1 & -2 & 1\\ 0 & 4 & -3\\ 0 & 0 & 1\end{bmatrix} = R$$

Pelo comentário feito antes do exemplo, sabemos que $W = [(1, -2, 1)_{\mathcal{B}}, (0, 4, -3)_{\mathcal{B}}, (0, 0, 1)_{\mathcal{B}}]$. Como esses três vetores são, evidentemente, LI, eles formam uma base de W, donde segue que $\dim(W) = 3$, que é precisamente o número pivôs da matriz escalonada R.

Um comentário final: encontramos uma base $\{f_1, f_2, f_3\}$ para o subespaço W, em que

$$\begin{aligned} f_1 &= (1, -2, 1)_{\mathcal{B}} = 1 - 2x + x^2,\\ f_2 &= (0, 4, -3)_{\mathcal{B}} = 4x - x^2,\\ f_3 &= (0, 0, 1)_{\mathcal{B}} = x^2.\end{aligned}$$

Como W também é gerado por $\{(x-1)^2, (x-2)^2, (x-3)^2\}$ e, agora, sabemos que $\dim(W) = 3$, segue que esse conjunto também é uma base de W. ◇

Exemplo 4.6.3 Em $M_2(\mathbb{R})$, determine a dimensão do subespaço $W = [v_1, v_2, v_3]$, em que

$$v_1 = \begin{bmatrix} 2 & -1\\ 2 & 0\end{bmatrix}, \qquad v_2 = \begin{bmatrix} 1 & 2\\ 3 & 4\end{bmatrix}, \qquad v_3 = \begin{bmatrix} 0 & -5\\ -4 & -8\end{bmatrix}.$$

Solução: Como $\dim\big(M_2(\mathbb{R})\big) = 4$, segue que $\dim(W) \leq 4$. Mais do que isso, como Wé gerado por 3 vetores, sabemos, pelo Teorema 4.4.10, que $\dim(W) \leq 3$. Seja $\mathcal{B} = \left\{\begin{bmatrix} 1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix} 0 & 1\\ 0 & 0\end{bmatrix}, \begin{bmatrix} 0 & 0\\ 1 & 0\end{bmatrix}, \begin{bmatrix} 0 & 0\\ 0 & 1\end{bmatrix}\right\}$ a base canônica de $M_2(\mathbb{R})$. Vamos trabalhar com coordenadas em relação a $\mathcal{B}$. Temos, assim,

$$W = \left[(2, -1, 2, 0)_{\mathcal{B}},\ (1, 2, 3, 4)_{\mathcal{B}},\ (0, -5, -4, -8)_{\mathcal{B}}\right].$$

Agora,

$$\begin{bmatrix} 2 & -1 & 2 & 0\\ 1 & 2 & 3 & 4\\ 0 & -5 & -4 & -8\end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 4\\ 2 & -1 & 2 & 0\\ 0 & -5 & -4 & -8\end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 4\\ 0 & -5 & -4 & -8\\ 0 & -5 & -4 & -8\end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 4\\ 0 & -5 & -4 & -8\\ 0 & 0 & 0 & 0\end{bmatrix} = R$$

Logo, $\{w_1, w_2\}$, em que

$$w_1 = (1,2,3,4)_{\mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = v_2 \quad \text{e} \quad w_2 = (0,-5,-4,-8)_{\mathcal{B}} = \begin{bmatrix} 0 & -5 \\ -4 & -8 \end{bmatrix} = v_3,$$

é uma base de W e, portanto, $\dim(W) = 2$, que é o número de pivôs de R. $\diamond$

Exemplo 4.6.4 Encontre uma base e a dimensão do subespaço $S = [u_1, u_2, u_3, u_4]$ de $\mathbb{R}^5$, em que

$$u_1 = (1,0,-1,2,0),\ u_2 = (2,1,3,0,0),\ u_3 = (0,1,-5,4,0),\ u_4 = (1,0,-11,10,0).$$

Solução: Sabemos que $\dim(S) \leq 4$, uma vez que S é gerado por 4 vetores. Como vimos fazendo (e trabalhando com coordenadas em relação à base canônica de $\mathbb{R}^5$), vamos submeter a matriz cujas linhas são formadas pelas coordenadas dos geradores de S ao processo de escalonamento:

$$\begin{bmatrix} 1 & 0 & -1 & 2 & 0 \\ 2 & 1 & 3 & 0 & 0 \\ 0 & 1 & -5 & 4 & 0 \\ 1 & 0 & -11 & 10 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 2 & 0 \\ 0 & 1 & 5 & -4 & 0 \\ 0 & 1 & -5 & 4 & 0 \\ 0 & 0 & -10 & 8 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 2 & 0 \\ 0 & 1 & 5 & -4 & 0 \\ 0 & 0 & -10 & 8 & 0 \\ 0 & 0 & -10 & 8 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 2 & 0 \\ 0 & 1 & 5 & -4 & 0 \\ 0 & 0 & -10 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Assim, $\dim(S) = 3$ (que é o número de pivôs da matriz escalonada ao final do processo) e uma base de S é $\{u_1', u_2', u_3'\}$, em que

$$u_1' = (1,0,-1,2,0) = u_1, \quad u_2' = (0,1,5,-4,0) \quad \text{e} \quad u_3' = (0,0,-10,8,0). \quad \diamond$$

Observação No exemplo acima, encontramos uma base $\{u_1', u_2', u_3'\}$ de S que não está contida no conjunto gerador $\{u_1, u_2, u_3, u_4\}$. O Teorema 4.4.8 garante que é possível extrair de $\{u_1, u_2, u_3, u_4\}$ uma base de S. Para um método que permite extrair uma base de um conjunto gerador de um subespaço, veja o Apêndice A.

Exemplo 4.6.5 Determine a dimensão do subespaço $V = \left\{ p \in \mathcal{P}_3(\mathbb{R}) \mid p(1) = 0 \right\}$ de $\mathcal{P}_3(\mathbb{R})$.

Solução: Começamos por observar que V é, de fato, um subespaço de $\mathcal{P}_3(\mathbb{R})$, como o leitor pode verificar. Assim, $\dim(V) \leq \dim\big(\mathcal{P}_3(\mathbb{R})\big) = 4$. Vamos trabalhar em coordenadas e procurar um conjunto gerador de V. Seja $\mathcal{B} = \{1, t, t^2, t^3\}$ a base canônica de $\mathcal{P}_3(\mathbb{R})$. Dado um vetor genérico $p = (a,b,c,d)_{\mathcal{B}}$ de $\mathcal{P}_3(\mathbb{R})$, determinemos condições necessárias e suficientes sobre suas coordenadas para que p seja um elemento de V. Sabemos que $p \in V$ se, e somente se $p(1) = 0$. Como $p = (a,b,c,d)_{\mathcal{B}}$, segue que $p = a + bt + ct^2 + dt^3$ e, portanto, $p \in V$ se, e somente se, $0 = p(1) = a + b + c + d$. Ou seja, os vetores de V são aqueles elementos de $\mathcal{P}_3(\mathbb{R})$ cujas coordenadas, em relação à base $\mathcal{B}$, satisfazem a equação

$$x + y + z + w = 0,$$

ou, em outras palavras, os elementos de V são precisamente os vetores de $\mathcal{P}_3(\mathbb{R})$ cujas coordenadas, em relação à base $\mathcal{B}$, são soluções do sistema linear homogêneo (em notação matricial)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}.$$

A matriz de coeficientes desse sistema já é escalonada e tem apenas um pivô. As variáveis y, z, w são livres, e, portanto, as soluções do sistema são da forma $(-y - z - w, y, z, w)$, com $y, z, w \in \mathbb{R}$. Logo, os vetores de V são da forma

$$(-y - z - w, y, z, w)_{\mathcal{B}} = y(-1,1,0,0)_{\mathcal{B}} + z(-1,0,1,0)_{\mathcal{B}} + w(-1,0,0,1)_{\mathcal{B}}.$$

(O que fizemos aqui foi "colocar em evidência" as variáveis livres y, z, w a fim de obter a solução geral como combinação linear de um conjunto finito de vetores.) Conclui-se daí que $V = [g_1, g_2, g_3]$, em que

$$g_1 = (-1,1,0,0)_{\mathcal{B}}, \quad g_2 = (-1,0,1,0)_{\mathcal{B}} \quad \text{e} \quad g_3 = (-1,0,0,1)_{\mathcal{B}}.$$

Como $\{g_1, g_2, g_3\}$ é claramente LI, esses três vetores formam uma base de V, donde segue que $\dim(V) = 3$. (Explicitamente, como elementos de $\mathcal{P}_3(\mathbb{R})$, os elementos dessa base de V são $g_1 = -1 + t$, $g_2 = -1 + t^2$ e $g_3 = -1 + t^3$.)

Veja que, neste exemplo, a dimensão do subespaço V é dada pelo número de variáveis livres do sistema linear homogêneo que define as coordenadas dos elementos de V, que por sua vez, é igual ao número total de variáveis (no caso, a dimensão do espaço ambiente $\mathcal{P}_3(\mathbb{R})$) menos o número de pivôs: $4 - 1 = 3$. $\diamond$

No exemplo acima, vimos uma instância em que foi possível determinar a dimensão de um subespaço em termos do número de variáveis livres de um sistema linear homogêneo. Diante de um problema em que se trabalha com coordenadas e o subespaço está descrito em termos de condições que devem ser estar satisfeitas pelas coordenadas de seus elementos, sempre teremos o método aplicado no exemplo anterior à disposição. Isso se deve ao seguinte fato.

Proposição 4.6.6 *Seja* $A \in M_{m\times n}(\mathbb{R})$ *e seja* $S = \{v \in \mathbb{R}^n \mid v$ *é solução do sistema linear homogêneo* $AX = 0\}$. *Então,* S *é um subespaço de* $\mathbb{R}^n$ *e* $\dim(S) = n - t$, *em que* t *é o número de pivôs de uma matriz escalonada obtida a partir de* A *por operações elementares sobre linhas.*

Demonstração Já vimos, no Exemplo 4.2.3 que S é, de fato, um subespaço de $\mathbb{R}^n$. Vejamos qual é sua dimensão. Seja R uma matriz escalonada obtida a partir de A por operações elementares sobre linhas (por exemplo, aplicando-se o processo de escalonamento a A). Então, os sistemas lineares $AX = 0$ e $RX = 0$ têm as mesmas soluções. As soluções de $RX = 0$ são obtidas atribuindo-se parâmetros independentes às variáveis livres e escrevendo, por retrossubstituição, as variáveis pivô em termos das variáveis livres. Digamos, então, que R tenha t pivôs, e consequentemente, o sistema $RX = 0$ tenha $n - t$ variáveis livres. Uma solução genérica de $RX = 0$ se escreve, então, como uma combinação linear de $n - t$ vetores $v_1, v_2, \dots, v_{n-t}$ de $\mathbb{R}^n$, em que os escalares são justamente os $n - t$ parâmetros atribuídos às variáveis livres. Ou, dito de outra maneira, S é gerado por $v_1, v_2, \dots, v_{n-t}$. Como na posição relativa a cada uma das variáveis livres apenas um desses vetores tem coordenada igual a 1, enquanto os demais têm essa coordenada igual a 0, segue que o conjunto $\{v_1, v_2, \dots, v_{n-t}\}$ é, além de gerador de S, também LI — uma base de S. Logo, $\dim(S) = n - t$. $\square$

Exemplo 4.6.7 Encontre uma base e a dimensão do subespaço

$$W = \left\{ p \in \mathcal{P}_2(\mathbb{R}) \;\middle|\; p'(2) = p(1) = \int_0^2 p(x)\,dx \right\}$$

de $\mathcal{P}_2(\mathbb{R})$.

Solução: Fica a cargo do leitor verificar que, de fato, W é um subespaço de $\mathcal{P}_2(\mathbb{R})$. Trabalhemos com coordendas em relação à base canônica $\mathcal{B} = \{1, x, x^2\}$ de $\mathcal{P}_2(\mathbb{R})$. Para encontrar um conjunto gerador para W, tomemos um elemento genérico $p = (a, b, c)_{\mathcal{B}}$ de $\mathcal{P}_2(\mathbb{R})$ e vamos impor as condições para que ele pertença a W. Temos que $p = a + bx + cx^2$. Assim, $p' = b + 2cx$ e, portanto, $p'(2) = b + 4c$. Ainda, $p(1) = a + b + c$, e $\int_0^2 p(x)\,dx = 2a + 2b + \frac{8}{3}c$ (uma vez que uma primitiva de p é $ax + \frac{b}{2}x^2 + \frac{c}{3}x^3$). Logo, $p \in W$ se, e somente se, $\begin{cases} b + 4c = a + b + c \\ a + b + c = 2a + 2b + \frac{8}{3}c \end{cases}$, que é equivalente a $\begin{cases} a - 3c = 0 \\ a + b + \frac{5}{3}c = 0 \end{cases}$. Ou seja, $p \in W$ se, e somente se, suas coordenadas, em relação à base $\mathcal{B}$, satisfazem o sistema linear homogêneo

$$\begin{bmatrix} 1 & 0 & -3 \\ 1 & 1 & \frac{5}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Escalonando a matriz de coeficientes, obtemos

$$\begin{bmatrix} 1 & 0 & -3 \\ 1 & 1 & \frac{5}{3} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & \frac{14}{3} \end{bmatrix} = R$$

Como R tem 2 pivôs, $\dim(W) = 3 - 2 = 1$. Para obter uma base de W, basta encontrar a solução geral do sistema, que, no caso, é $(3z, -\frac{14}{3}z, z)$, com $z \in \mathbb{R}$. Isto é, os elementos de W são da forma $(3z, -\frac{14}{3}z, z)_{\mathcal{B}} = z(3, -\frac{14}{3}, 1)_{\mathcal{B}}$, com

$z \in \mathbb{R}$; em outras palavras, $W = \left[(3, -\frac{14}{3}, 1)_{\mathcal{B}}\right] = \left[3 - \frac{14}{3}x + x^2\right]$. E esse conjunto gerador (de um único vetor) é uma base para W. $\diamond$

Exemplo 4.6.8 Encontre uma base para o subespaço $S = \{(x, y, z, w) \in \mathbb{R}^4 \mid 2x-y+z = 3x+y-z+w = 5y-5z+2w = 0\}$ de $\mathbb{R}^4$.

Solução: S é o subespaço de $\mathbb{R}^4$ formado precisamente pelas soluções do sistema linear homogêneo cuja matriz de coeficientes é $A = \begin{bmatrix} 2 & -1 & 1 & 0 \\ 3 & 1 & -1 & 1 \\ 0 & 5 & -5 & 2 \end{bmatrix}$. Escalonando A, obtemos

$$\begin{bmatrix} 2 & -1 & 1 & 0 \\ 3 & 1 & -1 & 1 \\ 0 & 5 & -5 & 2 \end{bmatrix} \to \begin{bmatrix} 2 & -1 & 1 & 0 \\ 0 & \frac{5}{2} & -\frac{5}{2} & 1 \\ 0 & 5 & -5 & 2 \end{bmatrix} \to \begin{bmatrix} 2 & -1 & 1 & 0 \\ 0 & 5 & -5 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R,$$

que tem 2 pivôs. Logo, $\dim(S) = 4 - 2 = 2$. Para achar uma base de S, olhando para R, obtemos que as soluções do sistema são os vetores (x, y, z, w), em que

$$w = t_1, \quad z = t_2, \quad 5y - 5z + 2w = 0, \quad 2x - y + z = 0.$$

Em termos dos parâmetros atribuídos às variáveis livres z e w, temos:

$$w = t_1, \quad z = t_2, \quad y = t_2 - \frac{2}{5}t_1, \quad x = -\frac{1}{5}t_1.$$

Ou seja, as soluções de $AX = 0$ são da forma

$$\left(-\frac{1}{5}t_1, t_2 - \frac{2}{5}t_1, t_2, t_1\right) = t_1\left(-\frac{1}{5}, -\frac{2}{5}, 0, 1\right) + t_2(0, 1, 1, 0), \quad (t_1, t_2 \in \mathbb{R}).$$

Logo, $S = \left[\, (-\frac{1}{5}, -\frac{2}{5}, 0, 1), (0, 1, 1, 0) \,\right]$ e $\left\{(-\frac{1}{5}, -\frac{2}{5}, 0, 1), (0, 1, 1, 0)\right\}$ é uma base de S. $\diamond$

O processo ilustrado no exemplo anterior pode ser revertido, no sentido que se um subespaço de $\mathbb{R}^n$ é apresentado por um conjunto de geradores, é possível descrevê-lo, também, como subespaço formado pelas soluções de um sistema linear homogêneo adequado. No próximo exemplo faremos exatamente isso. (Note que já havíamos tratado de um caso especial no Exemplo 4.3.6.)

Exemplo 4.6.9 (Lista 1 - Álgebra Linear II, ex. 9(iii)) Descreva o subespaço

$$S = \left[(2, -1, 2, 0), (1, 2, 3, 4), (0, -5, -4, -8)\right]$$

de $\mathbb{R}^4$ como o subespaço das soluções de um sistema linear homogêneo.

Solução: Dado $v = (a, b, c, d) \in \mathbb{R}^4$, temos que $v \in S$ se, e somente se, existirem $\alpha, \beta, \gamma \in \mathbb{R}$ tais que

$$v = \alpha(2, -1, 2, 0) + \beta(1, 2, 3, 4) + \gamma(0, -5, -4, -8),$$

o que é equivalente a existirem $\alpha, \beta, \gamma \in \mathbb{R}$ tais que

$$\begin{cases} 2\alpha + \beta = a \\ -\alpha + 2\beta - 5\gamma = b \\ 2\alpha + 3\beta - 4\gamma = c \\ 4\beta - 8\gamma = d \end{cases}$$

Isso, por sua vez, é dizer que (α, β, γ) é solução do sistema linear

$$\begin{cases} 2x + y = a \\ -x + 2y - 5z = b \\ 2x + 3y - 4z = c \\ 4y - 8z = d \end{cases} \tag{4.15}$$

nas incógnitas x, y, z. Assim, $v \in S$ se, e somente se, (4.15) tiver solução. Para encontrar as condições que garantem a existência de soluções de (4.15), escalonamos sua matriz aumentada:

$$\left[\begin{array}{ccc|c} 2 & 1 & 0 & a \\ -1 & 2 & -5 & b \\ 2 & 3 & -4 & c \\ 0 & 4 & -8 & d \end{array}\right] \to \left[\begin{array}{ccc|c} -1 & 2 & -5 & b \\ 2 & 1 & 0 & a \\ 2 & 3 & -4 & c \\ 0 & 4 & -8 & d \end{array}\right] \to \left[\begin{array}{ccc|c} -1 & 2 & -5 & b \\ 0 & 5 & -10 & a+2b \\ 0 & 7 & -14 & 2b+c \\ 0 & 4 & -8 & d \end{array}\right] \to \left[\begin{array}{ccc|c} -1 & 2 & -5 & b \\ 0 & 1 & -2 & \frac{d}{4} \\ 0 & 5 & -10 & a+2b \\ 0 & 7 & -14 & 2b+c \end{array}\right] \to \left[\begin{array}{ccc|c} -1 & 2 & -5 & b \\ 0 & 1 & -2 & \frac{d}{4} \\ 0 & 0 & 0 & a+2b-\frac{5}{4}d \\ 0 & 0 & 0 & 2b+c-\frac{7}{4}d \end{array}\right]$$

Segue que (4.15) tem solução se, e somente se,

$$\begin{cases} a + 2b - \frac{5}{4}d = 0 \\ 2b + c - \frac{7}{4}d = 0 \end{cases}$$

Assim, S é o subespaço de $\mathbb{R}^4$ formado pelas soluções do sistema linear homogêneo

$$\begin{cases} x_1 + 2x_2 - \frac{5}{4}x_4 = 0 \\ 2x_2 + x_3 - \frac{7}{4}x_4 = 0 \end{cases}$$

nas variáveis x_1, x_2, x_3, x_4. $\diamond$

Observação Vimos, na Proposição 4.6.1, que a dimensão de um subespaço está limitada superiormente pela dimensão do espaço vetorial em que ele está. Nos exemplos desta seção, calculamos a dimensão de diversos subespaços vetoriais. Cabe registrar, agora, que, dado um espaço vetorial V de dimensão finita igual a n, V sempre terá subsepaços de dimensão k, para todo $k \in \{0, 1, \dots, n\}$. Mais do que isso, se $1 \leq k < n$, então V tem infinitos subespaços de dimensão k. Vejamos por quê. O único subespaço de V de dimensão 0 é $\{0_V\}$, conforme a convenção que adotamos, e o único subespaço de V de dimensão n é V (pela Proposição 4.6.1). Assim, se $\dim(V) = 1$, V tem apenas dois subespaços: $\{0_V\}$, de dimensão 0, e o próprio V, de dimensão 1.

Suponha, agora, que $\dim(V) > 1$. Para cada vetor não nulo $v \in V$, o subespaço $[v]$ de V tem dimensão 1 (pois $\{v\}$ é uma base dele). Como $V \neq [v]$ (pois $\dim(V) > 1$), existe $u \in V$ tal que $u \notin [v]$. Para cada $\lambda \in \mathbb{R}$, defina o vetor $v_\lambda = v + \lambda u$. Obtém-se, assim, uma família infinita $[v_\lambda]$, $\lambda \in \mathbb{R}$, de subespaços de V, todos de dimensão 1, diferentes entre si.

Essa ideia pode ser generalizada: para cada $k \in \{1, 2, \dots, n-1\}$, tome k vetores $v_1, v_2, \dots, v_k \in V$ tais que $\{v_1, v_2, \dots, v_k\}$ seja LI (o que sempre é possível, tomando, por exemplo, k vetores de uma base de V) e escolha $u \in V$ tal que $u \notin [v_1, v_2, \dots, v_k]$ (o que também é possível, pois $[v_1, v_2, \dots, v_k] \neq V$). Definindo, para cada $\lambda \in \mathbb{R}$, $V_\lambda = [v_1, v_2, \dots, v_{k-1}, v_k + \lambda u]$, obtemos uma família infinita V_λ ($\lambda \in \mathbb{R}$) de subespaços de V, todos de dimensão k, distintos entre si.

Exercícios Lista 1 - Álgebra Linear II: Exs. 28–38.

4.7 Soma e interseção de subespaços

Se S_1 e S_2 são subespaços de um espaço vetorial, a *interseção* deles é denotada e definida por

$$S_1 \cap S_2 = \left\{ v \in V \mid v \in S_1 \text{ e } v \in S_2 \right\}.$$

Veremos que $S_1 \cap S_2$ é, sempre, um subespaço de V, ao passo que a união

$$S_1 \cup S_2 = \left\{ v \in V \mid v \in S_1 \text{ ou } v \in S_2 \right\}$$

é um subconjunto de V que, em geral, não é um subespaço.

Proposição 4.7.1 *Seja V um espaço vetorial e sejam S_1 e S_2 subespaços de V. Então, $S_1 \cap S_2$ é um subespaço de V.*

Demonstração O vetor nulo 0_V é um elemento de $S_1 \cap S_2$, porque é um elemento tanto de S_1 como de S_2, já que são, ambos, subespaços de V. Sejam, agora, $u, v \in S_1 \cap S_2$ e seja $\lambda \in \mathbb{R}$. Então, como $u \in S_1$, $v \in S_1$ e S_1 é um subespaço de V, segue que $u + v \in S_1$. Analogamente, $u + v \in S_2$, uma vez que ambos, u e v, são vetores de S_2, que também é um subespaço. Assim, $u + v \in S_1 \cap S_2$. Finalmente, como S_1 é um subespaço e $u \in S_1$, segue $\lambda u \in S_1$. Analogamente, $\lambda u \in S_2$, donde conclui-se que $\lambda u \in S_1 \cap S_2$. Ou seja, o subconjunto $S_1 \cap S_2$ de V contém o vetor nulo, é fechado para somas e é fechado para multiplicação por escalares. Logo, $S_1 \cap S_2$ é um subespaço de V. □

Como acabamos de ver, $S_1 \cap S_2$ é um subespaço de V que está contido, por definição, tanto em S_1 quando em S_2. Assim, se V for um espaço vetorial de dimensão finita, teremos

$$\dim(S_1 \cap S_2) \leq \dim(S_1) \quad \text{e} \quad \dim(S_1 \cap S_2) \leq \dim(S_2).$$

No próximo exemplo, encontramos a dimensão da interseção de dois subespaços.

Exemplo 4.7.2 Considere os seguintes subespaços de $\mathbb{R}^3$:

$$S_1 = \big[(0, -1, 2), (2, -1, -1)\big], \quad S_2 = \big[(1, -1, 1), (1, 1, 0)\big].$$

Descreva $S_1 \cap S_2$.

Solução: Quando se pede uma descrição de um subespaço, é frequentemente suficiente exibir um conjunto gerador. Procuremos, assim, um conjunto gerador para $S_1 \cap S_2$. Para tanto, procederemos da seguinte maneira: tomando um elemento arbitrário de S_1, vejamos que condições ele deve satisfazer para pertencer, também, a S_2. Dado $v \in S_1$, sabemos que existem $a, b \in \mathbb{R}$ tais que $v = a(0, -1, 2) + b(2, -1, -1)$, uma vez que S_1 é gerado por $\{(0, -1, 2), (2, -1, -1)\}$. Para que v seja, também um elemento de S_2, é preciso que existam $x, y \in \mathbb{R}$ tais que $v = x(1, -1, 1) + y(1, 1, 0)$. Resumindo, dados $a, b \in \mathbb{R}$, $a(0, -1, 2) + b(2, -1, -1) \in S_2$ se, e somente se, existem $x, y \in \mathbb{R}$ tais que

$$a(0, -1, 2) + b(2, -1, -1) = x(1, -1, 1) + y(1, 1, 0),$$

ou, equivalentemente, comparando coordenadas,

$$\begin{cases} x + y = 2b \\ -x + y = -a - b \\ x = 2a - b \end{cases} . \tag{4.16}$$

Assim, dados $a, b \in \mathbb{R}$, o vetor $a(0, -1, 2) + b(2, -1, -1)$ de S_1 é também um vetor de S_2 se, e somente se, o sistema linear 4.16, nas incógnitas x, y, tiver solução. Vejamos o resultado do processo de escalonamento aplicado à matriz aumentada de (4.16):

$$\left[\begin{array}{rr|c} 1 & 1 & 2b \\ -1 & 1 & -a-b \\ 1 & 0 & 2a-b \end{array}\right] \to \left[\begin{array}{rr|c} 1 & 1 & 2b \\ 0 & 2 & -a+b \\ 0 & -1 & 2a-3b \end{array}\right] \to \left[\begin{array}{rr|c} 1 & 1 & 2b \\ 0 & -1 & 2a-3b \\ 0 & 2 & -a+b \end{array}\right] \to \left[\begin{array}{rr|c} 1 & 1 & 2b \\ 0 & -1 & 2a-3b \\ 0 & 0 & 3a-5b \end{array}\right]$$

Daí, segue que (4.16) tem solução se, e somente se, $3a - 5b = 0$, ou seja, se, e somente se, $b = \frac{3a}{5}$. Logo, os elementos de S_1 que estão também em S_2 são as combinações lineares dos geradores de S_1 da forma

$$a(0, -1, 2) + \frac{3a}{5}(2, -1, -1) = \left(\frac{6a}{5}, -\frac{8a}{5}, \frac{7a}{5}\right) = a\left(\frac{6}{5}, -\frac{8}{5}, \frac{7}{5}\right), \quad \text{com } a \in \mathbb{R}.$$

Portanto, $S_1 \cap S_2 = \left[\left(\frac{6}{5}, -\frac{8}{5}, \frac{7}{5}\right)\right]$, ou, se o leitor preferir, $S_1 \cap S_2 = [(6, -8, 7)]$. Segue dessa descrição da interseção que $\dim(S_1 \cap S_2) = 1$, enquanto $\dim(S_1) = \dim(S_2) = 2$.

(Cabe comentar que poderíamos, alternativamente, ter descrito condições sobre um elemento arbitário de S_2 para que ele pertencesse a S_1, isto é, olhando (4.16) como um sistema nas incóginitas a, b. Esse outro caminho conduziria, evidentemente, à mesma conclusão.) $\diamond$

Enquanto a interseção de dois subespaços é sempre um subespaço, a união raramente o é. O próximo exemplo ilustra esse fenômeno.

Exemplo 4.7.3 Considere os subespaços

$$S_1 = \{ (x, 0) \mid x \in \mathbb{R} \} \quad \text{e} \quad S_2 = \{ (0, y) \mid y \in \mathbb{R} \}$$

do espaço vetorial $\mathbb{R}^2$. (Você deve se convencer que esses subconjuntos são, de fato, subespaços de $\mathbb{R}^2$.) Mostremos que a união $S_1 \cup S_2$ *não é* um subespaço de $\mathbb{R}^2$. De fato, tomemos $u = (1, 0) \in S_1 \subseteq S_1 \cup S_2$ e $v = (0, 1) \in S_2 \subseteq S_1 \cup S_2$. Se $S_1 \cup S_2$ fosse um subespaço de $\mathbb{R}^2$, deveríamos ter $u + v = (1, 1) \in S_1 \cup S_2$. Mas isso não ocorre pois os elementos da união têm primeira ou segunda coordenada nula.

O que falta à união de subespaços para que ela resulte em um subespaço é precisamente a propriedade de ser fechada com respeito a somas. Esse requisito será alcançado com a próxima definição.

Definição Seja V um espaço vetorial e sejam S_1 e S_2 subespaços de V. Definimos a *soma* de S_1 e S_2 como sendo o seguinte conjunto de V:

$$S_1 + S_2 = \{ w \in V \mid w = u_1 + u_2, \text{ para algum } u_1 \in S_1 \text{ e algum } u_2 \in S_2 \}.$$

A soma de subespaços é um subespaço:

Proposição 4.7.4 *Sejam S_1 e S_2 subespaços de um espaço vetorial V. Então, a soma $S_1 + S_2$ é um subespaço de V.*

Demonstração O vetor nulo de V é um elemento da soma, uma vez que $0_V = 0_V + 0_V$, em que o primeiro termo é um elemento de S_1 (pois S_1 é um subespaço) e o segundo termo é um elemento de S_2 (pois S_2 é um subespaço).

Agora, tomemos $w, w' \in S_1 + S_2$. Então, existem $u_1, u_1' \in S_1$ e $u_2, u_2' \in S_2$ tais que

$$w = u_1 + u_2 \quad \text{e} \quad w' = u_1' + u_2'.$$

Logo, $w + w' = (u_1 + u_2) + (u_1' + u_2') = (u_1 + u_1') + (u_2 + u_2')$. Como S_1 e S_2 são subespaços de V, segue que $u_1 + u_1' \in S_1$ e $u_2 + u_2' \in S_2$. Portanto, conseguimos escrever $w + w'$ como soma de um elemento de S_1 com um elemento de S_2, isto é, $w + w' \in S_1 + S_2$.

Finalmente, se $w = u_1 + u_2 \in S_1 + S_2$, com $u_1 \in S_1$ e $u_2 \in S_2$, e $\lambda \in \mathbb{R}$, segue que $\lambda w = \lambda(u_1 + u_2) = (\lambda u_1) + (\lambda u_2)$, que é um elemento da soma $S_1 + S_2$, uma vez que $\lambda u_1 \in S_1$ e $\lambda u_2 \in S_2$. $\square$

Veja que, no Exemplo 4.7.3, temos $S_1 + S_2 = \mathbb{R}^2$.

Dados subespaços S_1 e S_2 de um espaço vetorial V, temos associados a eles dois novos subespaços: a interseção $S_1 \cap S_2$ e a soma $S_1 + S_2$. Sabemos que tanto S_1 quando S_2 contêm $S_1 \cap S_2$. Observe, agora, que ambos estão contidos em $S_1 + S_2$, já que qualquer elemento u_1 de S_1 pode se escrever na forma $u_1 = u_1 + 0_V \in S_1 + S_2$, e qualquer elemento u_2 de S_2 se escreve como $u_2 = 0_V + u_2 \in S_1 + S_2$. Assim, temos as seguintes cadeias de subespaços de V:

$$S_1 \cap S_2 \subseteq S_1 \subseteq S_1 + S_2 \qquad \text{e} \qquad S_1 \cap S_2 \subseteq S_2 \subseteq S_1 + S_2.$$

Daí, extrai-se que se V é um espaço vetorial de dimensão finita, então

$$\dim(S_1 \cap S_2) \leq \min\{ \dim(S_1), \dim(S_2) \} \leq \max\{ \dim(S_1), \dim(S_2) \} \leq \dim(S_1 + S_2).$$

Como a soma $S_1 + S_2$ contém tanto S_1 como S_2, temos que $S_1 \cup S_2 \subseteq S_1 + S_2$. Além disso, $S_1 + S_2$ é o menor subespaço de V que contém a união $S_1 \cup S_2$, no sentido de que se W é um subespaço de V tal que $S_1 \cup S_2 \subseteq W$, então

$S_1 + S_2 \subseteq W$. (Você consegue escrever uma demonstração para esse fato?) Assim, podemos dizer que, em um certo sentido, a soma de subespaços "corrige" o fato de a união não ser um subespaço.

Para descrever a soma de subespaços em termos de conjuntos geradores deles, faremos uso do próximo resultado.

Proposição 4.7.5 *Seja V um espaço vetorial e sejam $u_1, \dots, u_m, v_1, \dots, v_n \in V$. Considere os subespaços $S_1 = [u_1, \dots, u_m]$ e $S_2 = [v_1, \dots, v_n]$ de V. Então,*

$$S_1 + S_2 = \big[u_1, \dots, u_m, v_1, \dots, v_n\big].$$

Demonstração Comecemos por mostrar que $S_1 + S_2 \subseteq [u_1, \dots, u_m, v_1, \dots, v_n]$. Dado $w \in S_1 + S_2$, sabemos que existem $u \in S_1$ e $v \in S_2$ tais que $w = u + v$. Como $S_1 = [u_1, \dots, u_m]$ e $S_2 = [v_1, \dots, v_n]$, existem $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n \in \mathbb{R}$ tais que

$$u = \alpha_1 u_1 + \cdots + \alpha_m u_m \quad \text{e} \quad v = \beta_1 v_1 + \cdots + \beta_n v_n.$$

Então,

$$w = u + v = \alpha_1 u_1 + \cdots + \alpha_m u_m + \beta_1 v_1 + \cdots + \beta_n v_n \in \big[u_1, \dots, u_m, v_1, \dots, v_n\big].$$

Isso prova que $S_1 + S_2 \subseteq [u_1, \dots, u_m, v_1, \dots, v_n]$.

Mostremos que vale a inclusão reversa. Seja $w \in [u_1, \dots, u_m, v_1, \dots, v_n]$. Então, existem $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n \in \mathbb{R}$ tais que

$$\begin{aligned} w &= \alpha_1 u_1 + \cdots + \alpha_m u_m + \beta_1 v_1 + \cdots + \beta_n v_n = (\alpha_1 u_1 + \cdots + \alpha_m u_m) + (\beta_1 v_1 + \cdots + \beta_n v_n) \\ &\in \big[u_1, \dots, u_m\big] + \big[v_1, \dots, v_n\big] = S_1 + S_2. \end{aligned}$$

Assim, $[u_1, \dots, u_m, v_1, \dots, v_n] \subseteq S_1 + S_2$.

Das duas inclusões provadas segue a igualdade desejada. □

Um alerta: da proposição acima segue que a união de bases de dois subespaços é um conjunto gerador da soma deles. Mas, em geral, esse conjunto não será uma base da soma, pois não será LI, como se pode ver no próximo exemplo.

Exemplo 4.7.6 Considere os subespaços

$$S_1 = \big[(1, 0, -1, 0), (1, 2, 0, 1), (1, 0, 1, 1)\big] \quad \text{e} \quad S_2 = \big[(1, 1, 1, 0), (1, 0, 0, -1)\big]$$

de $\mathbb{R}^4$. Descreva $S_1 \cap S_2$ e $S_1 + S_2$, e calcule suas dimensões.

Solução: Vimos, na Proposição 4.7.5, que

$$S_1 + S_2 = \big[(1, 0, -1, 0), (1, 2, 0, 1), (1, 0, 1, 1), (1, 1, 1, 0), (1, 0, 0, -1)\big].$$

Escrevendo as coordenadas, em relação à base canônica de $\mathbb{R}^4$, dos vetores geradores de $S_1 + S_2$ nas linhas de uma matriz e escalonando (como fizemos no Exemplo 4.6.3), obtemos

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 2 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Portanto, $S_1 + S_2 = [(1, 0, -1, 0), (0, 1, 2, 0), (0, 0, 1, -1), (0, 0, 0, -2)]$. Como os vetores nesse conjunto gerador são claramente LI (pois são linhas não nulas de uma matriz escalonada), esse conjunto é uma base de $S_1 + S_2$ e $\dim(S_1 + S_2) = 4$. Em particular, conclui-se que $S_1 + S_2 = \mathbb{R}^4$.

(Note que os conjuntos geradores de S_1 e S_2, por serem, ambos, LI, são, de fato, bases de S_1 e S_2. Segue da discussão que fizemos acima, que a união desses dois conjuntos, apesar de gerar $S_1 + S_2$, não é uma base da soma, já que contém 5 vetores, mas $\dim(S_1 + S_2) = 4$. Aliás, não existem conjuntos LI em $\mathbb{R}^4$ contendo 5 vetores.)

Tratemos, agora, da interseção, como fizemos no Exemplo 4.7.2. Dado $v = a(1,0,-1,0)+b(1,2,0,1)+c(1,0,1,1) \in S_1$, com $a,b,c \in \mathbb{R}$, temos que $v \in S_2$ se, e somente se, existem $x,y \in \mathbb{R}$ tais que $v = x(1,1,1,0) + y(1,0,0,-1)$, ou seja, e somente se, o sistema linear

$$\begin{cases} x+y = a+b+c \\ x = 2b \\ x = -a+c \\ -y = b+c \end{cases} \tag{4.17}$$

nas variáveis x,y tiver solução. Escalonando a matriz aumentada do sistema, obtemos

$$\left[\begin{array}{cc|c} 1 & 1 & a+b+c \\ 1 & 0 & 2b \\ 1 & 0 & -a+c \\ 0 & -1 & b+c \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & a+b+c \\ 0 & -1 & -a+b-c \\ 0 & -1 & -2a-b \\ 0 & -1 & b+c \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & a+b+c \\ 0 & -1 & -a+b-c \\ 0 & 0 & -a-2b+c \\ 0 & 0 & a+2c \end{array}\right]$$

Assim, (4.17) tem solução se, e somente se, $\begin{cases} -a-2b+c = 0 \\ a+2c = 0 \end{cases}$, e estamos, novamente, diante de um sistema linear, desta vez, homogêneo. O escalonamento de sua matriz de coeficientes resulta em

$$\begin{bmatrix} -1 & -2 & 1 \\ 1 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} -1 & -2 & 1 \\ 0 & -2 & 3 \end{bmatrix}.$$

Portanto, c é uma variável livre e, resolvendo por retrossubstituição, obtemos $a = -2c$ e $b = \frac{3c}{2}$. A conclusão a que chegamos é que para (4.17) ter solução (o que, por sua vez, é a condição precisa para o vetor v ser, também, um elemento de S_2) é necessário e suficiente que atribuindo-se valores arbitrários para c, tenhamos $a = -2c$ e $b = \frac{3c}{2}$. Isto quer dizer que os vetores em $S_1 \cap S_2$ são da forma

$$(-2c)(1,0,-1,0) + \frac{3c}{2}(1,2,0,1) + c(1,0,1,1) = \left(\frac{c}{2}, 3c, 3c, \frac{5c}{2}\right) = c\left(\frac{1}{2}, 3, 3, \frac{5}{2}\right),$$

com $c \in \mathbb{R}$. Portanto, $S_1 \cap S_2 = \left[\left(\frac{1}{2}, 3, 3, \frac{5}{2}\right)\right] = [(1,6,6,5)]$. Logo, uma base de $S_1 \cap S_2$ é $\{(1,6,6,5)\}$ e $\dim(S_1 \cap S_2) = 1$. $\diamond$

Como veremos a seguir, o que impede, no exemplo acima, que a dimensão da soma $S_1 + S_2$ seja igual à soma das dimensões de S_1 e de S_2 é justamente a interseção ter dimensão positiva.

Teorema 4.7.7 *Seja V um espaço vetorial de dimensão finita, e sejam S_1 e S_2 subespaços de V. Então,*

$$\dim(S_1 + S_2) + \dim(S_1 \cap S_2) = \dim(S_1) + \dim(S_2).$$

Demonstração Seja $\mathcal{B} = \{u_1, \dots, u_k\}$ uma base de $S_1 \cap S_2$. (Estamos admitindo, assim, que $\dim(S_1 \cap S_2) = k$.) Como $\mathcal{B}$ é um conjunto LI que está contido no subespaço S_1, pelo Teorema do completamento (Teorema 4.4.14), existe uma base de S_1 que contém $\mathcal{B}$, isto é, existem vetores $v_1, \dots, v_m \in S_1$ tais que $C_1 = \{u_1, \dots, u_k, v_1, \dots, v_m\}$ é uma base de S_1. Usando um raciocínio análogo, uma vez que $\mathcal{B}$ está contido também em S_2, garante-se que existem vetores $w_1, \dots, w_n \in S_2$ tais que $C_2 = \{u_1, \dots, u_k, w_1, \dots, w_n\}$ é uma base de S_2. (Em particular, estamos admitindo que $\dim(S_1) = k + m$ e que $\dim(S_2) = k + n$.)

Considere, agora, o conjunto

$$\mathcal{D} = \left\{ u_1, \dots, u_k, v_1, \dots, v_m, w_1, \dots, w_n \right\}.$$

Mostremos que $\mathcal{D}$ é uma base de $S_1 + S_2$.

$\mathcal{D}$ gera $S_1 + S_2$: Vimos, na Proposição 4.7.5, que a união de conjuntos geradores de S_1 e S_2 geram $S_1 + S_2$. Em particular, $S_1 + S_2 = [C_1 \cup C_2] = [u_1, \dots, u_k, v_1, \dots, v_m, w_1, \dots, w_n]$.

$\mathcal{D}$ é LI: Sejam $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_m, \gamma_1, \dots, \gamma_n \in \mathbb{R}$ tais que

$$\alpha_1 u_1 + \dots \alpha_k u_k + \beta_1 v_1 + \dots + \beta_m v_m + \gamma_1 w_1 + \dots + \gamma_n w_n = 0_V. \tag{4.18}$$

Nosso objetivo é mostrar que, necessariamente, todos esses escalares são nulos. Seja $z = \alpha_1 u_1 + \dots \alpha_k u_k + \beta_1 v_1 + \dots + \beta_m v_m$. Como $\mathcal{C}_1$ gera S_1, segue que $z \in S_1$. Por outro lado, segue de (4.18) que $z = -\gamma_1 w_1 - \dots - \gamma_n w_n$. Assim, porque $\mathcal{C}_2$ gera S_2, temos que $z \in S_2$. Segue, portanto, que $z \in S_1 \cap S_2$. Logo, existem $\delta_1, \dots, \delta_k \in \mathbb{R}$ tais que $z = \delta_1 u_1 + \dots + \delta_k u_k$, já que $S_1 \cap S_2$ é gerado por $\mathcal{B}$. Porém, isso implica que

$$-\gamma_1 w_1 - \dots - \gamma_n w_n = \delta_1 u_1 + \dots + \delta_k u_k,$$

o que é equivalente a

$$\delta_1 u_1 + \dots + \delta_k u_k + \gamma_1 w_1 + \dots + \gamma_n w_n = 0_V.$$

Como $\mathcal{C}_2$ é LI, segue que $\delta_1 = \dots = \delta_k = \gamma_1 = \dots = \gamma_n = 0$. Ou seja, já temos que todos os γ_i são nulos. Substituindo essa informação em (4.18), obtemos

$$\alpha_1 u_1 + \dots \alpha_k u_k + \beta_1 v_1 + \dots + \beta_m v_m = 0_V.$$

Como $\mathcal{C}_1$ é LI, $\alpha_1 = \dots = \alpha_k = \beta_1 = \dots = \beta_m = 0$. Assim, todos os $\alpha_i, \beta_i, \gamma_i$ são nulos, o que implica que $\mathcal{D}$ é, de fato, LI.

Vimos que $\mathcal{D}$ é um conjunto LI que gera $S_1 + S_2$ e contém $k + m + n$ vetores. Portanto,

$$\dim(S_1 + S_2) = k + m + n = (k + m) + (k + n) - k = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2),$$

que é o que desejávamos mostrar. □

Retornando ao Exemplo 4.7.6, tínhamos $\dim(S_1) = 3$, $\dim(S_2) = 2$, e calculamos que $\dim(S_1 + S_2) = 4$. Usando a fórmula do Teorema 4.7.7, já poderíamos ter previsto que $\dim(S_1 \cap S_2) = 3 + 2 - 4 = 1$, o que, de fato, se confirmou com o cálculo que efetuamos para encontrar uma base de $S_1 \cap S_2$.

Exemplo 4.7.8 (Prova 3, Álgebra Linear I, 2017) Assinale a alternativa que contém uma afirmação correta sobre os subespaços S_1 e S_2 de $\mathcal{P}_4(\mathbb{R})$ definidos abaixo:

$$S_1 = \left[t^2 - t,\ t - 1\right] \quad \text{e} \quad S_2 = \left\{ p \in \mathcal{P}_4(\mathbb{R}) : p(1) = p(-1) = 0 \right\}.$$

(A) Se $p \in S_1 \cap S_2$ e p não é o polinômio nulo, então p tem grau igual a 2.
(B) S_1 coincide com $\mathcal{P}_2(\mathbb{R})$.
(C) A interseção $S_1 \cap S_2$ contém $\mathcal{P}_1(\mathbb{R})$ como subespaço.
(D) S_1 está contido em S_2.
(E) A soma $S_1 + S_2$ coincide com $\mathcal{P}_4(\mathbb{R})$.

Solução: Vamos começar por determinar as dimensões de S_1 e S_2. Para tanto, trabalhemos em coordenadas em relação à base ordenada $\mathcal{B} = \{1, t, t^2, t^3, t^4\}$ de $\mathcal{P}_4(\mathbb{R})$. Sobre os geradores de S_1, sabemos:

$$t^2 - t = (0, -1, 1, 0, 0)_{\mathcal{B}} \quad \text{e} \quad t - 1 = (-1, 1, 0, 0, 0)_{\mathcal{B}}.$$

Como esses vetores claramente formam um conjunto LI, segue que $\dim(S_1) = 2$. Procuremos geradores de S_2. Dado $p = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 \in \mathcal{P}_4(\mathbb{R})$, temos que $p \in S_2$ se, e somente se,

$$0 = p(1) = a_0 + a_1 + a_2 + a_3 + a_4, \text{ e}$$
$$0 = p(-1) = a_0 - a_1 + a_2 - a_3 + a_4$$

Assim, para que p seja um elemento de S_2 é necessário e suficiente que seus coeficientes sejam solução do sistema linear homogêneo de matriz de coeficientes $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \end{bmatrix}$, que, submetida ao processo de escalonamento, fornece:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -2 & 0 & -2 & 0 \end{bmatrix}.$$

Daí, segue que a_2, a_3, a_4 são variáveis livres e $a_1 = -a_3$ e $a_0 = -a_2 - a_4$. Assim, os elementos de S_2 são da forma

$$(-a_2 - a_4) + (-a_3)t + a_2t^2 + a_3t^3 + a_4t^4 = a_2(-1 + t^2) + a_3(-t + t^3) + a_4(-1 + t^4), \text{ com } a_2, a_3, a_4 \in \mathbb{R}.$$

Logo, $S_2 = \left[-1 + t^2, -t + t^3, -1 + t^4\right]$, ou, em coordenadas,

$$S_2 = \left[\, (-1, 0, 1, 0, 0)_{\mathcal{B}}, (0, -1, 0, 1, 0)_{\mathcal{B}}, (-1, 0, 0, 0, 1)_{\mathcal{B}} \,\right].$$

Sendo, evidentemente, LI, esse conjunto de três vetores é uma base de S_2 e, assim, $\dim(S_2) = 3$. Agora, sabemos que $S_1 + S_2$ é gerado pelo conjunto de cinco vetores obtido unindo-se a base de S_1 com a base de S_2 que encontramos, ou seja,

$$S_1 + S_2 = \left[\, (0, -1, 1, 0, 0)_{\mathcal{B}}, (-1, 1, 0, 0, 0)_{\mathcal{B}}, (-1, 0, 1, 0, 0)_{\mathcal{B}}, (0, -1, 0, 1, 0)_{\mathcal{B}}, (-1, 0, 0, 0, 1)_{\mathcal{B}} \,\right].$$

Para encontrar uma base (e a dimensão de $S_1 + S_2$), escalonamos a matriz cujas linhas são formadas pelas coordenadas dos geradores de $S_1 + S_2$ (ver Observação na página 81):

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

Como a matriz escalonada obtida ao final desse processo tem 4 pivôs, segue que $\dim(S_1 + S_2) = 4$ (e que as linhas não nulas dessa matriz são as coordenadas de vetores em uma base de $S_1 + S_2$.) Do Teorema 4.7.7, extraímos $\dim(S_1 \cap S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 + S_2) = 2 + 3 - 4 = 1$.

Com essas informações, já é possível ver que as alternativas (B), (C), (D) e (E) são todas falsas: $\dim(S_1) = 2$ e $\dim\big(\mathcal{P}_2(\mathbb{R})\big) = 3$, portanto $S_1 \neq \mathcal{P}_2(\mathbb{R})$; $\dim(S_1 \cap S_2) = 1$ e $\dim\big(\mathcal{P}_1(\mathbb{R})\big) = 2$, logo $\mathcal{P}_1(\mathbb{R}) \not\subseteq S_1 \cap S_2$; se $S_1 \subseteq S_2$, teríamos $S_1 + S_2 = S_2$, mas esses espaços têm dimensões diferentes; $\dim(S_1 + S_2) = 4$ e $\dim\big(\mathcal{P}_4(\mathbb{R})\big) = 5$, portanto, $S_1 + S_2 \neq \mathcal{P}_4(\mathbb{R})$.

Resta provar que (A) é, de fato, verdadeira. Sabemos que $\dim(S_1 \cap S_2) = 1$. Procuremos uma base para esse subespaço. Dado $p \in S_1$, sabemos que p se escreve na forma $p = a(t^2 - t) + b(t - 1) = -b + (-a + b)t + at^2$, com $a, b \in \mathbb{R}$. Vamos impor as condições para que p seja, também, um elemento de S_2:

$$\begin{aligned} 0 &= p(1) = -b + (-a + b) + a \\ 0 &= p(-1) = -b - (a - b) + a = 2a - 2b \end{aligned}$$

A primeira equação não fornece informação nova, mas a segunda, sim: $b = a$, ou seja, os elementos de $S_1 \cap S_2$ são da forma

$$a(t^2 - t) + a(t - 1) = a(-1 + t^2), \text{ com } a \in \mathbb{R}.$$

Logo, $S_1 \cap S_2 = \left[-1 + t^2\right]$. E, com efeito, todo polinômio não nulo em $S_1 \cap S_2$ tem grau 2. *Resposta:* (A) ◇

Exercícios Lista 1 - Álgebra Linear II: Exs. 39–49.

4.8 Soma direta de subespaços

A dimensão da soma de subespaços coincide com a soma das dimensões deles precisamente quando a interseção for nula. Esse fato segue diretamente do Teorema 4.7.7 e merece uma nomenclatura própria.

Definição Sejam S_1 e S_2 subespaços de um espaço vetorial V. Se $S_1 \cap S_2 = \{0_V\}$, dizemos que a soma $S_1 + S_2$ é *direta*. A notação, neste caso, para a soma dos subespaços S_1 e S_2 é $S_1 \oplus S_2$.

Exemplo 4.8.1 A soma dos subespaços

$$W_1 = \left\{ (x, 0, 0) \mid x \in \mathbb{R} \right\} \quad \text{e} \quad W_2 = \left\{ (0, y, z) \mid y, z \in \mathbb{R} \right\}$$

de $\mathbb{R}^3$ é direta, pois, claramente, $W_1 \cap W_2 = \{(0, 0, 0)\}$. Em particular, $\dim(W_1 \oplus W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2) = 1 + 2 - 0 = 3$. Portanto, $W_1 \oplus W_2 = \mathbb{R}^3$.

Proposição 4.8.2 *Sejam S_1 e S_2 subespaços de um espaço vetorial de dimensão finita V. Então, a soma de S_1 e S_2 é direta se, e somente se,* $\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2)$.

Demonstração A soma $S_1 + S_2$ é direta, por definição, se, e somente se, $S_1 \cap S_2 = \{0_V\}$, o que, por sua vez, é equivalente a $\dim(S_1 \cap S_2) = 0$. Mas, pelo Teorema 4.7.7, isso ocorre precisamente quando $\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2)$. □

Exemplo 4.8.3 Retomando o Exemplo 4.7.2, em que consideramos os subespaços

$$S_1 = \left[(0, -1, 2), (2, -1, -1) \right] \quad \text{e} \quad S_2 = \left[(1, -1, 1), (1, 1, 0) \right]$$

de $\mathbb{R}^3$, podemos afirmar que a soma $S_1 + S_2$ não é direta, uma vez que $S_1 \cap S_2 = [(6, -8, 7)] \neq \{(0, 0, 0)\}$. Utilizando o critério da Proposição 4.8.2, poderíamos ter argumentado simplesmente que a soma não é direta porque $\dim(S_1) + \dim(S_2) = 2 + 2 = 4 \neq \dim(S_1 + S_2)$, uma vez que todos os subespaços de $\mathbb{R}^3$ têm dimensão limitada superiormente por $\dim(\mathbb{R}^3) = 3$, em particular, $S_1 + S_2$. Agora, como $\dim(S_1 \cap S_2) = 1$, segue que $\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2) = 2 + 2 - 1 = 3 = \dim(\mathbb{R}^3)$. Assim, $S_1 + S_2 = \mathbb{R}^3$. (Veja a Proposição 4.6.1.) Isso quer dizer que todo vetor de $\mathbb{R}^3$ se escreve como uma soma de um vetor de S_1 com um vetor de S_2. Por exemplo, tome o vetor $(-3, 2, 7) \in \mathbb{R}^3$; esse vetor pode ser decomposto como uma soma

$$(-3, 2, 7) = (-4, -1, 8) + (1, 3, -1),$$

com $(-4, -1, 8) = 3(0, -1, 2) + (-2)(2, -1, -1) \in S_1$ e $(1, 3, -1) = (-1)(1, -1, 1) + 2(1, 1, 0) \in S_2$. Porém, essa decomposição não é única, por exemplo, também é possível escrever

$$(-3, 2, 7) = (-16, 15, -6) + (13, -13, 13),$$

em que $(-16, 15, -6) = (-7)(0, -1, 2) + (-8)(2, -1, -1) \in S_1$ e $(13, -13, 13) = 13(1, -1, 1) + 0(1, 1, 0) \in S_2$. A não unicidade da decomposição é consequência do fato de a soma $S_1 + S_2$ não ser direta, como veremos a seguir.[1]

Proposição 4.8.4 *Sejam S_1 e S_2 subespaços de um espaço vetorial V. Então a soma $S_1 + S_2$ é direta se, e somente se, todo vetor de $S_1 + S_2$ se escrever de maneira única na forma $u_1 + u_2$, com $u_1 \in S_1$ e $u_2 \in S_2$.*

Demonstração Suponha que a soma $S_1 + S_2$ seja direta e tome $v \in S_1 + S_2$. Se

$$v = u_1 + u_2 \quad \text{e} \quad v = u_1' + u_2',$$

com $u_1, u_1' \in S_1$ e $u_2, u_2' \in S_2$, então teremos $u_1 + u_2 = u_1' + u_2'$, o que implica que $u_1 - u_1' = u_2' - u_2$, que é um vetor de $S_1 \cap S_2$ (uma vez que o lado esquerdo está em S_1, o direito em S_2 e eles são iguais). Mas, como a soma é direta, $S_1 \cap S_2 = \{0_V\}$. Portanto, $u_1 - u_1' = 0_V$ e $u_2' - u_2 = 0_V$. Portanto $u_1 = u_1'$ e $u_2 = u_2'$.

Reciprocamente, suponha que os vetores de $S_1 + S_2$ se decompõem de maneira única como soma de um vetor de S_1 com um vetor de S_2. Então, se $u \in S_1 \cap S_2$, temos a decomposição $u = u + 0_V$, com $u \in S_1$ e $0_V \in S_2$ e a decomposição $u = 0_V + u$, com $0_V \in S_1$ e $u \in S_2$. Da unicidade, segue que $u = 0_V$. Ou seja, o único vetor em $S_1 \cap S_2$ é o vetor nulo de V. Logo, a soma é direta. □

[1] A título de curiosidade, todo vetor de $\mathbb{R}^3$ tem infinitas decomposições distintas. Vale em geral que

$$(a, b, c) = \big((-2a - 3b - c + 5t)(0, -1, 2) + (-a - 2b - c + 3t)(2, -1, -1)\big) + \big((3a + 4b + 2c - 7t)(1, -1, 1) + t(1, 1, 0)\big),$$

para qualquer $t \in \mathbb{R}$. Você consegue imaginar como essa expressão foi obtida?

A proposição que acabamos de demonstrar será especialmente útil quando estivermos tratando de espaços vetoriais com produto interno, no próximo capítulo.

Exemplo 4.8.5 Mostre que $\mathbb{R}^3 = S \oplus T$, em que S e T são os subespaços de $\mathbb{R}^3$ definidos por

$$S = \left\{ (x, y, z) \in \mathbb{R}^3 \mid x - 2y - z = 0 \right\} \quad \text{e} \quad T = \left\{ (x, y, z) \in \mathbb{R}^3 \mid 2x - y = 0 \text{ e } -x + y + 3z = 0 \right\}.$$

Além disso, dado $v \in \mathbb{R}^3$, encontre $v_1 \in S$ e $v_2 \in T$ tais que $v = v_1 + v_2$.

Solução: Comecemos por mostrar que a soma $S + T$ é direta. Dado $v = (x, y, z) \in \mathbb{R}^3$, temos que $v \in S \cap T$ se, e somente se,

$$\begin{cases} x - 2y - z = 0 \\ 2x - y = 0 \\ -x + y + 3z = 0 \end{cases}. \tag{4.19}$$

Como $\det \begin{bmatrix} 1 & -2 & -1 \\ 2 & -1 & 0 \\ -1 & 1 & 3 \end{bmatrix} = 8 \neq 0$, segue que a única solução de (4.19) é $(0, 0, 0)$. Logo, $S \cap T = \{(0, 0, 0)\}$ e, portanto, a soma $S + T$ é direta. Utilizando as ideias da Seção 4.6, podemos encontrar

$$S = \left[(2, 1, 0), (1, 0, 1) \right] \quad \text{e} \quad T = \left[(-3, -6, 1) \right].$$

Segue que $\dim(S) = 2$ e $\dim(T) = 1$. Assim, $\dim(S \oplus T) = 2 + 1 = 3 = \dim(\mathbb{R}^3)$. Logo, $\mathbb{R}^3 = S \oplus T$.

Agora, dado $v = (x, y, z) \in \mathbb{R}^3$, queremos encontrar $v_1 \in S$ e $v_2 \in T$ tais que $v = v_1 + v_2$. Como $v_1 \in S_1$, precisamos determinar $\alpha, \beta \in \mathbb{R}$ tais que $v_1 = \alpha(2, 1, 0) + \beta(1, 0, 1)$. Analogamente, como $v_2 \in S_2$, precisamos determinar $\gamma \in \mathbb{R}$ tal que $v_2 = \gamma(-3, -6, 1)$. Resumindo, é preciso encontrar $\alpha, \beta, \gamma \in \mathbb{R}$ tais que

$$(x, y, z) = \alpha(2, 1, 0) + \beta(1, 0, 1) + \gamma(-3, -6, 1),$$

ou, equivalentemente, tais que

$$\begin{cases} 2\alpha + \beta - 3\gamma = x \\ \alpha - 6\gamma = y \\ \beta + \gamma = z \end{cases}$$

resolvendo esse sistema (possível determinado), obtemos

$$\alpha = \frac{3x - 2y - 3z}{4}, \quad \beta = \frac{-x + 2y + 9z}{8}, \quad \gamma = \frac{x - 2y - z}{8}.$$

Assim,

$$v_1 = \frac{3x - 2y - 3z}{4}(2, 1, 0) + \frac{-x + 2y + 9z}{8}(1, 0, 1) = \left(\frac{11x - 6y - 3z}{8}, \frac{3x - 2y - 3z}{4}, \frac{-x + 2y + 9z}{8} \right)$$

e

$$v_2 = \frac{x - 2y - z}{8}(-3, -6, 1) = \left(\frac{-3x + 6y + 3z}{8}, \frac{-3x + 6y + 3z}{4}, \frac{x - 2y - z}{8} \right). \quad \diamond$$

Exercícios Lista 1 - Álgebra Linear II: Exs. 50–51.

Capítulo 5
Espaços vetoriais com produto interno

Na Seção 2.4, vimos que o espaço vetorial $\mathbb{V}^3$, dos vetores tridimensionais, está dotado de uma operação, à época designada produto escalar, que permitia dar um tratamento mais geométrico, por meio do conceito de ângulo, a algumas relações entre vetores.

Nesta seção, apresentaremos, de modo axiomático, o conceito de produto interno em um espaço vetorial, que generaliza o de produto escalar de $\mathbb{V}^3$.

5.1 Produto interno

A Proposição 2.4.2 apresentava propriedades algébricas do produto escalar em $\mathbb{V}^3$. Muitos dos resultados obtidos a respeito da geometria de $\mathbb{V}^3$ eram consequência dessa proposição.

Aqui faremos uma alteração de ponto de vista e definiremos, em um espaço vetorial abstrato, uma operação, a ser chamada de produto interno, por intermédio das propriedades que se deseja que ela tenha, a fim de que argumentos de natureza geométrica possam ser feitos nesse contexto mais geral.

Definição Um *produto interno* em um espaço vetorial V é uma função

$$\begin{aligned} V \times V &\longrightarrow \mathbb{R} \\ (u, v) &\longmapsto \langle u, v \rangle \end{aligned}$$

que satisfaz

(PI-1) $\langle u_1 + u_2, v \rangle = \langle u_1, v \rangle + \langle u_2, v \rangle$, para todos $u_1, u_2, v \in V$;
(PI-2) $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$, para todos $u, v \in V$ e $\lambda \in \mathbb{R}$;
(PI-3) $\langle u, v \rangle = \langle v, u \rangle$, para todos $u, v \in V$;
(PI-4) para todo $u \in V$, $u \neq 0_V$, vale $\langle u, u \rangle > 0$.

Como já mencionamos, o produto escalar em $\mathbb{V}^3$ é um produto interno. Vejamos outros exemplos.

Exemplo 5.1.1 A função definida por

$$\Big\langle (a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n) \Big\rangle = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

é um produto interno no espaço vetorial $\mathbb{R}^n$. Que as condições (PI-1)–(PI-3) na definição de produto interno estão satisfeitas segue imediatamente da definição das operações de soma e multiplicação por escalar em $\mathbb{R}^n$. A condição (PI-4) é consequência do fato de uma soma de quadrados de números reais ser sempre um número positivo se pelo menos um de seu termos for não nulo. Essa operação será doravante designada o *produto interno usual* de $\mathbb{R}^n$.

Exemplo 5.1.2 Além do produto interno usual, existem outros produtos internos em $\mathbb{R}^2$. Por exemplo,

$$\Big\langle (a,b),(c,d)\Big\rangle = 2ac + bd$$

também define um produto interno em $\mathbb{R}^2$, como o leitor pode facilmente verificar. (Veja o Exercício 53 da Lista 1 - Álgebra Linear II para ainda mais um exemplo de produto interno em $\mathbb{R}^2$.)

Exemplo 5.1.3 Considere o espaço vetorial $\mathcal{P}_n(\mathbb{R})$, dos polinômios de grau menor ou igual a n, e a função definida por

$$\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt,$$

para todos $f,g \in \mathcal{P}_n(\mathbb{R})$. Mostremos que essa função é um produto interno no espaço vetorial $\mathcal{P}_n(\mathbb{R})$. Vale iniciar notando que, de fato, essa função está bem definida, uma vez que para todos $f,g \in \mathcal{P}_n(\mathbb{R})$, o produto fg é uma função contínua em $[0,1]$ e, portanto, integrável nesse intervalo. As condições (PI-1) e (PI-2) são propriedades da integração: se $f,g,h \in \mathcal{P}_n(\mathbb{R})$, então

$$\begin{aligned}\langle f+g,h\rangle &= \int_0^1 \big(f(t)+g(t)\big)h(t)\,dt = \int_0^1 \big(f(t)h(t)+g(t)h(t)\big)\,dt\\ &= \int_0^1 f(t)h(t)\,dt + \int_0^1 g(t)h(t)\,dt = \langle f,h\rangle + \langle g,h\rangle,\end{aligned}$$

e se $\lambda \in \mathbb{R}$, segue

$$\langle \lambda f,g\rangle = \int_0^1 \lambda f(t)g(t)\,dt = \lambda\int_0^1 f(t)g(t)\,dt = \lambda\langle f,g\rangle.$$

A condição (PI-3) é óbvia. Resta (PI-4); vejamos. Comecemos por lembrar que o vetor nulo de $\mathcal{P}_n(\mathbb{R})$ é o polinômio nulo. Assim, se $f \in \mathcal{P}_n(\mathbb{R})$ e $f \neq 0_{\mathcal{P}_n(\mathbb{R})}$, existe $t_0 \in [0,1]$ tal que $f(t_0) \neq 0$, o que implica que $f(t_0)^2 > 0$. Como f é uma função contínua no intervalo $[0,1]$, existem $c,d \in \mathbb{R}$ com $0 \leq c < d \leq 1$ tais que $f(t)^2 > 0$, para todo $t \in (c,d)$, enquanto $f(t)^2 \geq 0$ para todo $t \in [0,c] \cup [d,1]$. Portanto,

$$\langle f,f\rangle = \int_0^1 f(t)^2\,dt = \int_0^c f(t)^2\,dt + \int_c^d f(t)^2\,dt + \int_d^1 f(t)^2\,dt > 0,$$

já que $\int_c^d f(t)^2\,dt > 0$ e os outros dois termos na soma acima não são negativos.

O argumento utilizado também prova que a função definida acima é um produto interno no espaço $\mathcal{P}(\mathbb{R})$ de todos os polinômios, sem limitação no grau. Ainda, não há nada de especial na escolha dos extremos de integração: para quaisquer $a,b \in \mathbb{R}$ com $a < b$, $\langle f,g\rangle = \int_a^b f(t)g(t)\,dt$ define um produto interno em $\mathcal{P}_n(\mathbb{R})$ e em $\mathcal{P}(\mathbb{R})$.

Mais geralmente, $\langle f,g\rangle = \int_a^b f(t)g(t)\,dt$ é um produto interno no espaço vetorial $C([a,b])$ de todas as funções reais contínuas definidas no intervalo $[a,b]$. (Mas não é um produto interno no espaço vetorial $C(\mathbb{R})$. Por quê?)

Exemplo 5.1.4 Há outro produto interno comumente utilizado em $\mathcal{P}_n(\mathbb{R})$, que é uma versão discreta do produto interno definido no exemplo anterior. Fixe m números reais distintos $a_1,a_2,\dots,a_m$, com $m > n$. Dados $f,g \in \mathcal{P}_n(\mathbb{R})$, defina

$$\langle f,g\rangle = f(a_1)g(a_1) + f(a_2)g(a_2) + \cdots + f(a_m)g(a_m).$$

Fica a cargo do leitor verificar que as condições (PI-1)–(PI-3) estão satisfeitas (para essas, a hipótese que $m > n$ não é necessária). Para a condição (PI-4), suponha que $f \in \mathcal{P}_n(\mathbb{R})$ seja tal que $0 = \langle f,f\rangle = f(a_1)^2 + f(a_2)^2 + \cdots + f(a_m)^2$. Então, $f(a_1) = f(a_2) = \cdots = f(a_m) = 0$. Como f tem grau menor ou igual a n, se não for nulo, f terá no máximo n raízes distintas. Como os a_i são distintos entre si e $m > n$, segue que f tem que ser o polinômio nulo. (A título de ilustração do argumento, considere o polinômio $p(t) = t^2 - 3t + 2 \in \mathcal{P}_2(\mathbb{R})$. Se tomarmos apenas dois pontos, por exemplo, $a_1 = 1$ e $a_2 = 2$, teremos $\langle p,p\rangle = p(1)^2 + p(2)^2 = 0$, ao passo que p não é o polinômio nulo.)

As propriedades listadas a seguir são consequências imediatas da definição.

Proposição 5.1.5 *Seja V um espaço vetorial com produto interno, e sejam $u, v, w \in V$ e $\lambda \in \mathbb{R}$. Valem:*

(i) $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$;
(ii) $\langle u, \lambda v \rangle = \lambda \langle u, v \rangle$;
(iii) $\langle u, 0_V \rangle = 0$;
(iv) $\langle u, u \rangle = 0$ *se, e somente se* $u = 0_V$.

Demonstração (i) segue de (PI-1) e (PI-3); (ii) segue de (PI-2) e (PI-3). Provemos (iii): temos $\langle u, 0_V \rangle = \langle u, 0_V + 0_V \rangle = \langle u, 0_V \rangle + \langle u, 0_V \rangle$, por (i). Assim $\langle u, 0_V \rangle$ é um número real que é igual ao seu dobro. Só existe um número real com essa propriedade, 0. Finalmente, para (iv), já sabemos, por (PI-4), que se $u \neq 0_V$, então $\langle u, u \rangle \neq 0$. Reciprocamente, se $u = 0_V$, então $\langle u, u \rangle = \langle 0_V, 0_V \rangle = 0$, por (iii). □

Norma e distância

Continuando com a analogia que estabelecemos entre o produto escalar e o produto interno, definiremos norma e distância, em um espaço vetorial abstrato, em termos de um produto interno.

Definição Seja V um espaço vetorial com produto interno $\langle\ ,\ \rangle$. Dado $u \in V$, definimos a *norma* de u por

$$\|u\| = \sqrt{\langle u, u \rangle}.$$

Dados $u, v \in V$, definimos a *distância* entre u e v por

$$\operatorname{dist}(u, v) = \|u - v\|.$$

(Observe que, em vista da condição (PI-4), a norma de um vetor está sempre definida. Além disso, a função dist satisfaz as condições esperadas de uma função distância[1].)

Exemplo 5.1.6 No espaço vetorial $\mathbb{R}^n$ com o produto interno usual (cf. Exemplo 5.1.1), dados $u = (a_1, a_2, \dots, a_n)$, $v = (b_1, b_2, \dots, b_n) \in \mathbb{R}^n$, temos

$$\|u\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

e

$$\operatorname{dist}(u, v) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}.$$

A norma e distância assim definidas em $\mathbb{R}^n$ costumam ser designadas *euclidianas*.

Observação Não é demais enfatizar que norma e distância são conceitos relativos ao produto interno. Assim, por exemplo, se tomarmos em $\mathbb{R}^2$ os seguintes dois produtos internos

$$\big\langle (a, b), (c, d) \big\rangle_1 = ac + bd \qquad \text{e} \qquad \big\langle (a, b), (c, d) \big\rangle_2 = 2ac + bd,$$

teremos definidas em $\mathbb{R}^2$ duas normas (e, portanto, duas distâncias). Por exemplo,

$$\|(1,1)\|_1 = \sqrt{\big\langle (1,1), (1,1) \big\rangle_1} = \sqrt{2} \qquad \text{e} \qquad \|(1,1)\|_2 = \sqrt{\big\langle (1,1), (1,1) \big\rangle_2} = \sqrt{3}.$$

Propriedades imediatas da norma são enunciadas na próxima proposição.

---

[1] que são

(i) $\operatorname{dist}(u, v) = 0$ se, e somente se, $u = v$;
(ii) $\operatorname{dist}(u, v) = \operatorname{dist}(v, u)$;
(iii) $\operatorname{dist}(u, v) \leq \operatorname{dist}(u, w) + \operatorname{dist}(w, v)$,

quaisquer que sejam $u, v, w \in V$ (A última condição é consequência do Corolário 5.1.9, que veremos adiante.)

Proposição 5.1.7 *Seja V um espaço vetorial com produto interno e sejam $u \in V$, $\lambda \in \mathbb{R}$. Valem:*

(i) se $u \neq 0_V$, então $\|u\| > 0$;
(ii) $\|u\| = 0$ se, e somente se, $u = 0_V$;
(iii) $\|\lambda u\| = |\lambda| \, \|u\|$.

Demonstração De (PI-4), se $u \neq 0_V$, então $\langle u, u\rangle > 0$ e, portanto, $\|u\| = \sqrt{\langle u, u\rangle} > 0$, provando (i). A equivalência em (ii) segue da Proposição 5.1.5 (iv). Para (iii), temos que $\|\lambda u\| = \sqrt{\langle \lambda u, \lambda u\rangle} = \sqrt{\lambda^2 \langle u, u\rangle} = \sqrt{\lambda^2}\sqrt{\langle u, u\rangle} = |\lambda| \, \|u\|$, em que a segunda igualdade segue de (PI-2) e Proposição 5.1.5 (ii). □

Os próximos dois resultados são propriedades não tão imediatas da norma.

Teorema 5.1.8 *Seja V um espaço vetorial com produto interno. Então, para todos $u, v \in V$, vale*

$$|\langle u, v\rangle| \leq \|u\| \, \|v\| \,. \tag{5.1}$$

Além disso, em (5.1)*, vale a igualdade se, e somente se, $\{u, v\}$ é LD.*

A desigualdade (5.1) é conhecida como *desigualdade de Cauchy-Schwarz* .

Demonstração Se $v = 0_V$, então $\langle u, v\rangle = 0$ e $\|v\| = 0$. Portanto, neste caso, a desigualdade vale trivialmente. Suponha que $v \neq 0_V$. Então, para todo $\lambda \in \mathbb{R}$, temos

$$0 \leq \|u + \lambda v\|^2 = \langle u + \lambda v, u + \lambda v\rangle = \langle u, u\rangle + \lambda \langle u, v\rangle + \lambda \langle v, u\rangle + \lambda^2 \langle v, v\rangle = \|v\|^2 \lambda^2 + 2 \langle u, v\rangle \lambda + \|u\|^2 \,.$$

Assim, a função $f(x) = \|v\|^2 x^2 + 2\langle u, v\rangle x + \|u\|^2$ é uma função polinomial de grau dois (pois $\|v\|^2 \neq 0$, já que $v \neq 0_V$) que assume, sempre, valores não negativos. Logo, seu discriminante não é positivo: $\big(2\langle u, v\rangle\big)^2 - 4\|v\|^2\|u\|^2 \leq 0$, ou, equivalentemente $|\langle u, v\rangle| \leq \|u\| \, \|v\|$.

É fácil ver que se $\{u, v\}$ é LD, então $|\langle u, v\rangle| = \|u\| \, \|v\|$. Reciprocamente, suponha que $|\langle u, v\rangle| = \|u\| \, \|v\|$. Mostremos que $\{u, v\}$ é LD. Considere o sistema linear homogêneo

$$\begin{bmatrix} \|u\|^2 & \langle u, v\rangle \\ \langle u, v\rangle & \|v\|^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{5.2}$$

Como $\det \begin{bmatrix} \|u\|^2 & \langle u, v\rangle \\ \langle u, v\rangle & \|v\|^2 \end{bmatrix} = \|u\|^2 \|v\|^2 - \langle u, v\rangle^2 = 0$, segue que (5.2) tem uma solução não trivial (α, β). Logo,

$$\begin{aligned} \|\alpha u + \beta v\|^2 &= \langle \alpha u + \beta v, \alpha u + \beta v\rangle \\ &= \alpha\big(\|u\|^2 \alpha + \langle u, v\rangle \beta\big) + \beta\big(\langle u, v\rangle \alpha + \|v\|^2 \beta\big) \\ &= \alpha 0 + \beta 0 \qquad \text{pois } (\alpha, \beta) \text{ é solução de (5.2)} \\ &= 0 \end{aligned}$$

Portanto, $\alpha u + \beta v = 0_V$, donde segue que $\{u, v\}$ é LD. □

Uma consequência da desigualdade de Cauchy-Schwarz é o seguinte resultado.

Corolário 5.1.9 *Seja V um espaço vetorial com produto interno. Então,*

$$\|u + v\| \leq \|u\| + \|v\| \,, \tag{5.3}$$

para todos $u, v \in V$.

Chama-se (5.3) de *desigualdade triangular* (pois ela evoca o fato de que a soma dos comprimentos de dois lados de um triângulo sempre excede o comprimento do terceiro).

Demonstração Dados $u, v \in V$, temos

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \|u\|^2 + 2\langle u, v\rangle + \|v\|^2 \\
&\leq \|u\|^2 + 2\|u\|\,\|v\| + \|v\|^2 \qquad \text{por Cauchy-Schwarz} \\
&= \big(\|u\| + \|v\|\big)^2,
\end{aligned}$$

o que implica a desigualdade desejada. □

Exemplo 5.1.10 Considere o produto interno definido por

$$\langle f, g\rangle = \int_0^{\pi} f(t)g(t)\,dt$$

no espaço vetorial $C([0,\pi])$ e os vetores $u = \cos(t)$, $v = \operatorname{sen}(t)$. Calcule $\|u\|$ e $\langle u, v\rangle$. Como fica a desigualdade de Cauchy-Schwarz nesse espaço?

Solução: Por definição,

$$\|u\|^2 = \langle u, u\rangle = \int_0^{\pi} \cos^2(t)\,dt = \frac{1}{2}\int_0^{\pi} \big(1+\cos(2t)\big)\,dt = \frac{1}{2}\left(t + \frac{\operatorname{sen}(2t)}{2}\right)\Bigg|_0^{\pi} = \frac{\pi}{2}.$$

Portanto, $\|u\| = \sqrt{\frac{\pi}{2}}$. Similarmente,

$$\langle u, v\rangle = \int_0^{\pi} \cos(t)\operatorname{sen}(t)\,dt = \frac{1}{2}\int_0^{\pi} \operatorname{sen}(2t)\,dt = \frac{1}{2}\left(-\frac{\cos(2t)}{2}\right)\Bigg|_0^{\pi} = 0.$$

Finalmente, neste contexto, a desigualdade de Cauchy-Scharwz expressa-se por

$$\left|\int_0^{\pi} f(t)g(t)\,dt\right| \leq \left(\int_0^{\pi} f(t)^2\,dt\right)^{\frac{1}{2}} \left(\int_0^{\pi} g(t)^2\,dt\right)^{\frac{1}{2}},$$

quaiquer que sejam $f, g \in C([0,\pi])$. ◇

Exercícios Lista 1 - Álgebra Linear II: Exs. 52–56.

5.2 Ortogonalidade

Diremos que dois vetores u e v de um espaço vetorial com produto interno $\langle\ ,\ \rangle$ são *ortogonais* se $\langle u, v\rangle = 0$, e, neste caso, usaremos a notação $u \perp v$. (Mais uma vez, estamos emprestando nomenclatura e notação utilizadas no estudo de vetores de $\mathbb{V}^3$.)

Vimos, no Exemplo 5.1.10, que, em $C([0,\pi])$, com o produto interno definido naquele contexto, $\cos(t) \perp \operatorname{sen}(t)$.

Observação Assim como norma e distância, ortogonalidade também é um conceito relativo ao produto interno definido no espaço vetorial. Dois vetores de um espaço vetorial podem ser ortogonais com relação a um produto interno e deixarem de ser com relação a outro. Por exemplo, considere os seguintes dois produtos internos definidos em $\mathcal{P}(\mathbb{R})$:

$$\langle p, q\rangle_1 = \int_0^1 p(t)q(t)\,dt \qquad \text{e} \qquad \langle p, q\rangle_2 = \int_{-1}^1 p(t)q(t)\,dt.$$

Se tomarmos os vetores $f(t) = t$ e $g(t) = t^2$ de $\mathcal{P}(\mathbb{R})$, então f e g não são ortogonais com relação a $\langle\ ,\ \rangle_1$, pois

$$\langle f, g\rangle_1 = \int_0^1 t^3\, dt = \left.\frac{t^4}{4}\right|_0^1 = \frac{1}{4} \neq 0,$$

mas são ortogonais com relação a $\langle\ ,\ \rangle_2$, uma vez que

$$\langle f, g\rangle_2 = \int_{-1}^1 t^3\, dt = \left.\frac{t^4}{4}\right|_{-1}^1 = 0;$$

Assim como ocorria em $\mathbb{V}^3$, ortogonalidade garante independência linear:

Proposição 5.2.1 *Seja V um espaço vetorial com produto interno. Se $u_1, u_2, \dots, u_n \in V$ são vetores não nulos dois a dois ortogonais, então $\{u_1, u_2, \dots, u_n\}$ é LI.*

Demonstração Sejam $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ tais que $\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n = 0_V$. Então, para cada $i = 1, \dots, n$, temos

$$0 = \langle 0_V, u_i\rangle = \langle \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n, u_i\rangle = \lambda_1 \langle u_1, u_i\rangle + \lambda_2 \langle u_2, u_i\rangle + \cdots + \lambda_n \langle u_n, u_i\rangle.$$

Como, por hipótese, $\left\langle u_j, u_i\right\rangle = 0$, para todo $j \neq i$, o lado direito da expressão acima é igual a $\lambda_i \langle u_i, u_i\rangle = \lambda_i \|u_i\|^2$. Assim, $\lambda_i \|u_i\|^2 = 0$, mas o fato de u_i ser não nulo implica $\|u_i\| \neq 0$. Daí, obtemos $\lambda_i = 0$, para todo $i = 1, \dots, n$. Logo, o conjunto $\{u_1, u_2, \dots, u_n\}$ é LI. □

Definição Seja V um espaço vetorial de dimensão finita com produto interno e seja $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ uma base de V. Dizemos que $\mathcal{B}$ é *ortogonal* se $u_i \perp u_j$, para todos $i \neq j$. Dizemos que $\mathcal{B}$ é *ortonormal* se for ortogonal e, adicionalmente, $\|u_i\| = 1$, para todos $i = 1, \dots, n$.

Exemplo 5.2.2 A base canônica de $\mathbb{R}^n$ é ortonormal com relação ao produto interno usual.

Exemplo 5.2.3 A função $\langle A, B\rangle = \operatorname{tr}(B^{\intercal} A)$ define um produto interno no espaço vetorial $M_2(\mathbb{R})$, em que $\operatorname{tr}(X)$ denota o *traço* da matriz X, isto é, a soma dos elementos em sua diagonal principal, e $X^{\intercal}$ denota a matriz transposta da matriz X. (Veja o exercício 73 da Lista 1 - Álgebra Linear II.) Explicitamente,

$$\left\langle \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \right\rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}.$$

A base canônica $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ de $M_2(\mathbb{R})$ é ortonormal com relação a esse produto interno, como é rotineiro verificar. Este exemplo se generaliza de maneira óbvia para espaços de matrizes de tamanho superior.

Exemplo 5.2.4 A base canônica $\{1, t, t^2\}$ de $\mathcal{P}_2(\mathbb{R})$ não é ortogonal (e, portanto, também não é ortonormal) com relação ao produto interno $\langle p, q\rangle = \int_0^1 p(t)q(t)\, dt$, pois, conforme vimos, $t \not\perp t^2$.

Coordenadas de vetores em relação a bases ortonormais são particularmente convenientes, um fenômeno já encontrado no estudo do espaço vetorial $\mathbb{V}^3$. (Veja a Proposição 2.4.1.)

Proposição 5.2.5 *Seja V um espaço vetorial de dimensão finita com produto interno $\langle\ ,\ \rangle$ e seja $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ uma base ortonormal ordenada de V. Então,*

(i) para todo $u \in V$,

$$u = \big(\langle u, e_1\rangle, \langle u, e_2\rangle, \dots, \langle u, e_n\rangle\big)_{\mathcal{B}};$$

(ii) dados $v = (\alpha_1, \alpha_2, \dots, \alpha_n)_{\mathcal{B}}, w = (\beta_1, \beta_2, \dots, \beta_n)_{\mathcal{B}}$, valem

$$\langle v, w\rangle = \alpha_1\beta_1 + \alpha_2\beta_2 + \cdots + \alpha_n\beta_n \qquad e \qquad \|v\| = \sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2}.$$

Demonstração Sejam $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tais que $u = (\lambda_1, \lambda_2, \dots, \lambda_n)_{\mathcal{B}}$. Mostremos que $\lambda_i = \langle u, e_i \rangle$, para todo $i = 1, \dots, n$. Com efeito,

$$\langle u, e_i \rangle = \langle \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n, e_i \rangle = \lambda_1 \langle e_1, e_i \rangle + \lambda_2 \langle e_2, e_i \rangle + \cdots + \lambda_n \langle e_n, e_i \rangle = \lambda_i \|e_i\|^2,$$

pois $\mathcal{B}$ é ortogonal. Como, adicionalmente, $\|e_i\| = 1$, segue que $\lambda_i = \langle u, e_i \rangle$, provando (i).

Vejamos (ii):

$$\begin{aligned}
\langle v, w \rangle &= \langle \alpha_1 e_1 + \cdots + \alpha_n e_n, \beta_1 e_1 + \cdots + \beta_n e_n \rangle \\
&= \alpha_1 \beta_1 \|e_1\|^2 + \alpha_2 \beta_2 \|e_2\|^2 + \cdots + \alpha_n \beta_n \|e_n\|^2 && \text{pois } \langle e_i, e_j \rangle = 0 \text{ se } i \neq j \\
&= \alpha_1 \beta_1 + \alpha_2 \beta_2 + \cdots + \alpha_n \beta_n && \text{pois } \|e_i\| = 1.
\end{aligned}$$

A afirmação sobre a norma de v é um caso particular da identidade demonstrada acima. □

Observação O trabalho para encontrar uma base ortonormal de um espaço vetorial V com produto interno concentra-se na determinação de uma base ortogonal $\{v_1, v_2, \dots, v_n\}$ de V, uma vez que, de posse de tal base, o conjunto $\left\{\frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \dots, \frac{v_n}{\|v_n\|}\right\}$ é, claramente, uma base ortonormal de V. A primeira (e mais trabalhosa) etapa desse processo será discutida a seguir.

Exemplo 5.2.6 O conjunto $\{(0,0,1), (1,-1,0), (1,1,0)\}$ é uma base ortogonal de $\mathbb{R}^3$ com relação ao produto interno usual. Segue, da observação acima, que

$$\mathcal{B} = \left\{(0,0,1), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)\right\}$$

é uma base ortonormal de $\mathbb{R}^3$. Determine as coordenadas de $v = (-7, 2, 3)$ em relação à base $\mathcal{B}$.

Solução: Como $\mathcal{B}$ é ortonormal, podemos aplicar o item (i) da Proposição 5.2.5:

$$\begin{aligned}
\big\langle (-7,2,3), (0,0,1) \big\rangle &= 3 \\
\left\langle (-7,2,3), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right) \right\rangle &= -\frac{9}{\sqrt{2}} \\
\left\langle (-7,2,3), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) \right\rangle &= -\frac{5}{\sqrt{2}}
\end{aligned}$$

Assim,

$$v = \left(3, -\frac{9}{\sqrt{2}}, -\frac{5}{\sqrt{2}}\right)_{\mathcal{B}}. \quad \diamond$$

Processo de ortogonalização de Gram-Schmidt

Todo espaço vetorial de dimensão finita com produto interno tem bases ortonormais. Isso seguirá do próximo resultado. Em sua demonstração, será apresentado um procedimento que, a partir de um conjunto LI de vetores em um espaço vetorial com produto interno, produz um conjunto ortogonal de vetores que geram o mesmo subespaço. Esse procedimento é conhecido como *processo de ortogonalização de Gram-Schmidt*.

Teorema 5.2.7 *Seja V um espaço vetorial com produto interno e seja $\{u_1, u_2, \dots, u_n\}$ um conjunto LI de vetores de V. Então, existem $v_1, v_2, \dots, v_n \in V$ vetores não nulos dois a dois ortogonais entre si tais que, para todo $i = 1, \dots, n$, $[v_1, \dots, v_i] = [u_1, \dots, u_i]$.*

Demonstração Comecemos por definir $v_1 = u_1$. Então, é claro que $v_1 \neq 0_V$ e que $[v_1] = [u_1]$.

Agora, defina

$$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\|v_1\|^2} v_1.$$

Então, $v_2 \neq 0_V$, pois se v_2 fosse o vetor nulo, teríamos $u_2 \in [v_1] = [u_1]$, o que contradiz o fato de $\{u_1, u_2\}$ ser LI. Além disso, v_2 e v_1 são ortogonais, uma vez que

$$\langle v_2, v_1\rangle = \left\langle u_2 - \frac{\langle u_2, v_1\rangle}{\|v_1\|^2} v_1, v_1 \right\rangle = \langle u_2, v_1\rangle - \frac{\langle u_2, v_1\rangle}{\|v_1\|^2} \langle v_1, v_1\rangle = \langle u_2, v_1\rangle - \langle u_2, v_1\rangle = 0.$$

A penúltima igualdade segue do fato de que $\langle v_1, v_1\rangle = \|v_1\|^2$. Finalmente, como $\{v_1, v_2\}$ é um conjunto formado por dois vetores não nulos que são ortogonais, $\{v_1, v_2\}$ é LI (pela Proposição 5.2.1). Mas, por sua própria definição, $v_2 \in [v_1, u_2]$; portanto, $[v_1, v_2] \subseteq [v_1, u_2] = [u_1, u_2]$ (já que $[v_1] = [u_1]$). Assim, $[v_1, v_2]$ é um subespaço de dimensão 2 do espaço $[u_1, u_2]$, que também tem dimensão 2. Segue, da Proposição 4.6.1, que $[v_1, v_2] = [u_1, u_2]$.

Repetimos o processo na definição de v_3, utilizando os vetores v_1 e v_2 já definidos:

$$v_3 = u_3 - \frac{\langle u_3, v_1\rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2\rangle}{\|v_2\|^2} v_2.$$

Aqui, também $v_3 \neq 0_V$, uma vez que $u_3 \neq [u_1, u_2] = [v_1, v_2]$, já que $\{u_1, u_2, u_3\}$ é LI; e v_3 é ortogonal a v_1 e a v_2, pois

$$\begin{aligned}
\langle v_3, v_1\rangle &= \left\langle u_3 - \frac{\langle u_3, v_1\rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2\rangle}{\|v_2\|^2} v_2, v_1 \right\rangle \\
&= \langle u_3, v_1\rangle - \frac{\langle u_3, v_1\rangle}{\|v_1\|^2} \langle v_1, v_1\rangle - \frac{\langle u_3, v_2\rangle}{\|v_2\|^2} \langle v_2, v_1\rangle \\
&= \langle u_3, v_1\rangle - \langle u_3, v_1\rangle \qquad \text{uma vez que } \langle v_1, v_1\rangle = \|v_1\|^2 \text{ e } \langle v_2, v_1\rangle = 0 \\
&= 0.
\end{aligned}$$

A demonstração que $\langle v_3, v_2\rangle = 0$ é análoga. Temos, assim, um conjunto $\{v_1, v_2, v_3\}$ de três vetores não nulos, dois a dois ortogonais entre si, e, portanto, LI. Logo, $[v_1, v_2, v_3]$ é um subespaço vetorial do espaço vetorial $[v_1, v_2, u_3] = [u_1, u_2, u_3]$, que também tem dimensão 3. Logo, $[v_1, v_2, v_3] = [u_1, u_2, u_3]$.

Prosseguimos dessa maneira: uma vez definidos $v_1, \dots, v_i$, define-se

$$v_{i+1} = u_{i+1} - \frac{\langle u_{i+1}, v_1\rangle}{\|v_1\|^2} v_1 - \frac{\langle u_{i+1}, v_2\rangle}{\|v_2\|^2} v_2 - \cdots - \frac{\langle u_{i+1}, v_i\rangle}{\|v_i\|^2} v_i.$$

A verificação que $v_{i+1} \neq 0_V$, que v_{i+1} é ortogonal a $v_1, v_2, \dots, v_i$ e que $[v_1, v_2, \dots, v_{i+1}] = [u_1, u_2, \dots, u_{i+1}]$ é análoga à que fizemos nos casos $i = 1$ e 2. □

Corolário 5.2.8 *Seja V um espaço vetorial de dimensão finita com produto interno. Então, V tem uma base ortonormal.*

Demonstração Seja $\{u_1, \dots, u_n\}$ uma base de V. Submetendo-a ao processo de ortonormalização de Gram-Schmidt, obtemos um conjunto $\{v_1, \dots, v_n\}$ de vetores não nulos, dois a dois ortogonais — e, portanto, LI —, cujo subespaço gerado coincide com $[u_1, \dots, u_n] = V$. Logo, $\{v_1, \dots, v_n\}$ é uma base ortogonal de V e, portanto, $\left\{\frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \dots, \frac{v_n}{\|v_n\|}\right\}$ é uma base ortonormal de V. □

Vejamos dois exemplos.

Exemplo 5.2.9 Em $\mathbb{R}^3$, com o produto interno usual, encontre uma base ortonormal para o subespaço $S = \big[(0,2,1), (-1,1,2)\big]$.

Solução: O conjunto $\{u_1, u_2\}$, em que $u_1 = (0,2,1)$ e $u_2 = (-1,1,2)$ é LI e, portanto, uma base de S. Podemos submetê-lo ao processo de ortogonalização de Gram-Schmidt. Faça $v_1 = u_1 = (0,2,1)$ e

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1.$$

Temos

$$\langle u_2, v_1 \rangle = \Big\langle (-1, 1, 2), (0, 2, 1) \Big\rangle = 0 + 2 + 2 = 4$$

e

$$\|v_1\|^2 = \|(0, 2, 1)\|^2 = 0 + 4 + 1 = 5.$$

Logo,

$$v_2 = (-1, 1, 2) - \frac{4}{5}(0, 2, 1) = \left(-1, -\frac{3}{5}, \frac{6}{5}\right).$$

Assim, $\{v_1, v_2\}$ é uma base ortogonal de S. Agora, como já calculamos, $\|v_1\|^2 = 5$ e

$$\|v_2\|^2 = \left\|\left(-1, -\frac{3}{5}, \frac{6}{5}\right)\right\|^2 = 1 + \frac{9}{25} + \frac{36}{25} = \frac{70}{25}.$$

Portanto,

$$\|v_1\| = \sqrt{5} \qquad \text{e} \qquad \|v_2\| = \frac{\sqrt{70}}{5}.$$

Assim,

$$\left\{\left(0, \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right), \left(-\frac{5}{\sqrt{70}}, -\frac{3}{\sqrt{70}}, \frac{6}{\sqrt{70}}\right)\right\}$$

é uma base ortonormal de S. $\diamond$

Exemplo 5.2.10 (Lista 1 - Álgebra Linear II, ex. 66) Encontre uma base ortonormal para o subespaço $W = \left[1, t, t^2\right]$ do espaço vetorial $\mathcal{P}(\mathbb{R})$, com relação ao produto interno definido por

$$\langle p, q \rangle = \int_0^1 f(t)g(t)\, dt.$$

Solução: Sejam $u_1 = 1$, $u_2 = t$ e $u_3 = t^2$. Por ser LI, $\{u_1, u_2, u_3\}$ é uma base de W. Apliquemos a ela o processo de ortogonalização de Gram-Schmidt. Seja $v_1 = 1$ e seja

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1.$$

Como

$$\langle u_2, v_1 \rangle = \int_0^1 t\, dt = \left.\frac{t^2}{2}\right|_0^1 = \frac{1}{2} \qquad \text{e} \qquad \|v_1\|^2 = \int_0^1 1\, dt = t\Big|_0^1 = 1,$$

segue que

$$v_2 = t - \frac{1/2}{1} 1 = t - \frac{1}{2}.$$

Agora,

$$v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2.$$

Já sabemos que $\|v_1\|^2 = 1$. Efetuando os cálculos dos demais escalares envolvidos nessa expressão, obtemos

$$\langle u_3, v_1 \rangle = \int_0^1 t^2\, dt = \left.\frac{t^3}{3}\right|_0^1 = \frac{1}{3},$$

$$\langle u_3, v_2\rangle = \int_0^1 t^2\left(t-\frac{1}{2}\right)dt = \int_0^1 \left(t^3-\frac{t^2}{2}\right)dt = \left(\frac{t^4}{4}-\frac{t^3}{6}\right)\Bigg|_0^1 = \frac{1}{4}-\frac{1}{6} = \frac{1}{12}.$$

e

$$\|v_2\|^2 = \int_0^1 \left(t-\frac{1}{2}\right)^2 dt = \left.\frac{(t-\frac{1}{2})^3}{3}\right|_0^1 = \frac{1}{3}\left(\frac{1}{8}+\frac{1}{8}\right) = \frac{1}{12}.$$

Logo, obtemos que

$$v_3 = t^2 - t + \frac{1}{6}.$$

Assim,

$$\left\{1,\ t-\frac{1}{2},\ t^2-t+\frac{1}{6}\right\}$$

é uma base ortogonal de W. Resta normalizá-la, isto é, dividir cada vetor pela sua norma. Já sabemos que $\|v_1\| = 1$ e $\|v_2\| = \frac{1}{2\sqrt{3}}$. Resta calcular $\|v_3\|$. Vejamos,

$$\|v_3\|^2 = \int_0^1 \left(t^2-t+\frac{1}{6}\right)^2 dt = \int_0^1 \left(t^4-2t^3+\frac{4t^2}{3}-\frac{t}{3}+\frac{1}{36}\right)dt = \left(\frac{t^5}{5}-\frac{t^4}{2}+\frac{4t^3}{9}-\frac{t^2}{6}+\frac{t}{36}\right)\Bigg|_0^1 = \frac{1}{180}.$$

Logo, $\|v_3\| = \frac{1}{6\sqrt{5}}$ e, portanto,

$$\left\{1,\ \sqrt{3}(2t-1),\ \sqrt{5}(6t^2-6t+1)\right\}$$

é uma base ortonormal para W. $\diamond$

Projeção ortogonal

Vimos que, em um espaço vetorial com produto interno, temos uma noção de distância entre vetores. Então, faz sentido considerar o seguinte

Problema Seja V um espaço vetorial com produto interno. Dados um subespaço W de V e um vetor $u \in V$, é possível encontrar $w_0 \in W$ tal que $\operatorname{dist}(u, w_0) < \operatorname{dist}(u, w)$, para todo $w \in W$, $w \neq w_0$?

Em outras palavras, pode-se encontrar um elemento do subespaço W cuja distância a u seja a menor possível? Esse problema também é conhecido como *problema da melhor aproximação*. Quando ele tem solução, dizemos que o elemento $w_0 \in W$ encontrado é a melhor aproximação de u por elementos de W.

Veremos que quando W é um subespaço de dimensão finita, o problema da melhor aproximação sempre terá solução. Porém comecemos com a seguinte

Observação Se o problema da melhor aproximação tiver solução, ela é única. (Com efeito, dado um subespaço W de um espaço vetorial V com produto interno e um vetor $u \in V$, se $w_0 \in W$ é tal que $\operatorname{dist}(u, w_0) < \operatorname{dist}(u, w)$, para todo $w \in W$, $w \neq w_0$, então qualquer outro vetor de W que não seja w_0 distará de u mais do que w_0 dista e, portanto, não será uma solução do problema da melhor aproximação; w_0 é a única.)

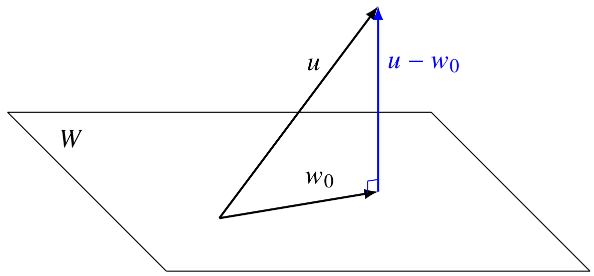
Sabemos que, na geometria euclidiana, a menor distância entre os pontos de uma reta ou de um plano e um determinado ponto fixo ocorre precisamente entre esse ponto e sua projeção ortogonal sobre a reta ou o plano. Em espaços vetoriais abstratos, temos uma versão análoga desse fato. Para enunciá-la, será útil introduzir uma notação.

Definição Seja V um espaço vetorial com produto interno, seja W um subespaço de V e seja $u \in V$. Dizemos que u é *ortogonal* a W, e escrevemos $u \perp W$ se $u \perp w$, para todo $w \in W$.

Podemos, agora, definir projeção ortogonal de um vetor sobre um subespaço.

Definição Seja V um espaço vetorial com produto interno, seja W um subespaço de V e seja $u \in V$. Se existir $w_0 \in W$ tal que $(u - w_0) \perp W$, dizemos que w_0 é uma *projeção ortogonal* de u sobre W.

Uma ilustração inspirada na geometria de $\mathbb{V}^3$ pode ajudar a compreensão desse conceito:



O próximo resultado relaciona a projeção ortogonal com o problema da melhor aproximação.

Proposição 5.2.11 *Seja V um espaço vetorial com produto interno, seja W um subespaço de V e seja $u \in V$. Se $w_0 \in W$ é uma projeção ortogonal de u sobre W, então w_0 é uma solução do problema da melhor aproximação de u por elementos de W.*

Demonstração Precisamos mostrar que $\operatorname{dist}(u, w) > \operatorname{dist}(u, w_0)$, para todo $w \in W$, $w \neq w_0$. Com efeito, dado $w \in W$, $w \neq w_0$, temos

$$\begin{aligned}\|u - w\|^2 &= \|(u - w_0) + (w_0 - w)\|^2 \\ &= \big\langle (u - w_0) + (w_0 - w), (u - w_0) + (w_0 - w) \big\rangle \\ &= \|u - w_0\|^2 + 2\langle u - w_0, w_0 - w\rangle + \|w_0 - w\|^2\end{aligned}$$

Como w_0 é uma projeção ortogonal de u sobre W, temos que $u - w_0$ é ortogonal a todo vetor de W, em particular, $u - w_0$ é ortogonal a $w_0 - w$. Assim, o termo central da soma acima é nulo: $\langle u - w_0, w_0 - w\rangle = 0$. Logo,

$$\|u - w\|^2 = \|u - w_0\|^2 + \|w_0 - w\|^2 > \|u - w_0\|^2,$$

uma vez que, $\|w_0 - w\| \neq 0$, pois w e w_0 são distintos. A conclusão desejada segue, lembrando que $\operatorname{dist}(u, w) = \|u - w\|$ e $\operatorname{dist}(u, w_0) = \|u - w_0\|$, por definição de distância em V. □

Segue, do resultado acima, que, se existir, a projeção ortogonal de um vetor sobre um subespaço é única, uma vez que ela é uma solução do problema da melhor aproximação, que, como vimos, é única.

Assim, se V é um espaço vetorial com produto interno, W é um subespaço de V e $u \in V$, se existir, a projeção ortogonal de u sobre W será denotada por $\operatorname{proj}_W u$.

Resta-nos encontrar condições que garantem a existência da projeção ortogonal. Uma resposta é apresentada a seguir.

Teorema 5.2.12 *Seja V um espaço vetorial com produto interno e seja W um subespaço de V de dimensão finita. Então, para todo $u \in V$, existe a projeção ortogonal de u sobre W e ela é dada por*

$$\operatorname{proj}_W u = \frac{\langle u, w_1\rangle}{\|w_1\|^2} w_1 + \frac{\langle u, w_2\rangle}{\|w_2\|^2} w_2 + \cdots + \frac{\langle u, w_n\rangle}{\|w_n\|^2} w_n, \tag{5.4}$$

em que $\{w_1, w_2, \ldots, w_n\}$ é uma base ortogonal de W.

Duas observações, neste momento, cabem. A primeira é que, como vimos no Corolário 5.2.8, sempre haverá uma base ortogonal para W; na realidade, basta tomarmos uma base qualquer de W e ortogonalizá-la pelo processo de Gram-Schmidt. A segunda dá conta precisamente desse processo; o modo como são definidos os vetores no processo de ortogonalização de Gram-Schmidt pode ser descrito, em termos da projeção ortogonal, da seguinte maneira:

$$v_{i+1} = u_{i+1} - \operatorname{proj}_{[u_1,\dots,u_i]} u_{i+1}.$$

Vamos à demonstração do teorema.

Demonstração Seja $u \in V$ e seja $\{w_1, w_2, \dots, w_n\}$ uma base ortogonal de W. Mostremos que o vetor

$$w = \frac{\langle u, w_1\rangle}{\|w_1\|^2}w_1 + \frac{\langle u, w_2\rangle}{\|w_2\|^2}w_2 + \cdots + \frac{\langle u, w_n\rangle}{\|w_n\|^2}w_n$$

satisfaz as condições para ser a projeção ortogonal de u sobre W. Primeiramente, está claro que $w \in W$. Resta mostrar que $u - w \perp W$, isto é, que $u - w$ é ortogonal a todo vetor de W. Para tanto, basta mostrar que $u - w$ é ortogonal a w_i, para todo $i = 1, \dots, n$, já que essa condição implicará que $u - w$ é ortogonal a toda combinação linear de $w_1, \dots, w_n$. Assim, vejamos: dado $i = 1, \dots, n$, temos

$$\begin{aligned}\langle u - w, w_i\rangle &= \left\langle u - \left(\frac{\langle u, w_1\rangle}{\|w_1\|^2}w_1 + \frac{\langle u, w_2\rangle}{\|w_2\|^2}w_2 + \cdots + \frac{\langle u, w_n\rangle}{\|w_n\|^2}w_n\right), w_i\right\rangle \\ &= \langle u, w_i\rangle - \frac{\langle u, w_1\rangle}{\|w_1\|^2}\langle w_1, w_i\rangle - \frac{\langle u, w_2\rangle}{\|w_2\|^2}\langle w_2, w_i\rangle - \cdots - \frac{\langle u, w_n\rangle}{\|w_n\|^2}\langle w_n, w_i\rangle .\end{aligned}$$

Como $\left\langle w_j, w_i\right\rangle = 0$ se $j \neq i$ e $\langle w_i, w_i\rangle = \|w_i\|^2$, a igualdade acima resume-se a

$$\langle u - w, w_i\rangle = \langle u, w_i\rangle - \frac{\langle u, w_i\rangle}{\|w_i\|^2}\langle w_i, w_i\rangle = 0.$$

Logo, $u - w \perp W$ e, portanto, $w = \operatorname{proj}_W u$. □

Em particular, se V é um espaço vetorial de dimensão finita com produto interno, W é um subespaço de V e $u \in V$, então a projeção ortogonal de u sobre W sempre existirá.

Exemplo 5.2.13 Em $\mathbb{R}^3$ com o produto interno usual, encontre $\operatorname{proj}_W v$, em que $v = (1, -1, 0)$ e

$$W = \big[(1, 2, 1), (2, 1, 2)\big].$$

Solução: Para utilizarmos a fórmula (5.4), necessitamos de uma base ortogonal de W. Chamemos $u_1 = (1, 2, 1)$ e $u_2 = (2, 1, 2)$, de modo que $W = [u_1, u_2]$. O conjunto $\{u_1, u_2\}$ é LI e, portanto, uma base de W; mas essa base não é ortogonal, uma vez que $\langle u_1, u_2\rangle = 6 \neq 0$. Vamos ortogonalizá-la por meio do processo de Gram-Schmidt. Começamos fazendo $v_1 = u_1 = (1, 2, 1)$ e definimos

$$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\|v_1\|^2}v_1.$$

Como

$$\langle u_2, v_1\rangle = \big\langle (2, 1, 2), (1, 2, 1)\big\rangle = 6 \qquad \text{e} \qquad \|v_1\|^2 = \langle v_1, v_1\rangle = \big\langle (1, 2, 1), (1, 2, 1)\big\rangle = 6,$$

segue que

$$v_2 = (2, 1, 2) - \frac{6}{6}(1, 2, 1) = (1, -1, 1).$$

Como v_1 e v_2 são ortogonais e geram o mesmo subespaço que u_1 e u_2, segue que $\{v_1, v_2\}$ é uma base ortogonal de W. Agora, podemos usar (5.4):

$$\begin{aligned}
\operatorname{proj}_W v &= \frac{\langle v, v_1\rangle}{\|v_1\|^2}v_1 + \frac{\langle v, v_2\rangle}{\|v_2\|^2}v_2\\
&= \frac{\big\langle (1,-1,0),(1,2,1)\big\rangle}{\|(1,2,1)\|^2}(1,2,1) + \frac{\big\langle (1,-1,0),(1,-1,1)\big\rangle}{\|(1,-1,1)\|^2}(1,-1,1)\\
&= \frac{-1}{6}(1,2,1) + \frac{2}{3}(1,-1,1)\\
&= \left(\frac{1}{2}, -1, \frac{1}{2}\right). \quad \diamond
\end{aligned}$$

Exemplo 5.2.14 Em $\mathcal{P}_2(\mathbb{R})$, com o produto interno

$$\langle f, g\rangle = f(0)g(0) + f(1)g(1) + f(2)g(2),$$

encontre o vetor de $\mathcal{P}_1(\mathbb{R})$ mais próximo de $q = \frac{1}{2}(t^2 + 3t)$.

Solução: Como $\mathcal{P}_1(\mathbb{R})$ é um subespaço de $\mathcal{P}_2(\mathbb{R})$, trata-se de encontrar a melhor aproximação de q por vetores de $\mathcal{P}_1(\mathbb{R})$. Sabemos, pela Proposição 5.2.11, que a resposta é dada pela projeção ortogonal de q sobre $\mathcal{P}_1(\mathbb{R})$. Para utilizar (5.4), precisamos de uma base ortogonal de $\mathcal{P}_1(\mathbb{R})$. Sabemos que $\{1, t\}$ é uma base de $\mathcal{P}_1(\mathbb{R})$, mas essa base não é ortogonal, uma vez que $\langle 1, t\rangle = (1)(0) + (1)(1) + (1)(2) = 3 \neq 0$. Apliquemos o processo de ortogonalização de Gram-Schmidt: $v_1 = 1$ e

$$v_2 = t - \frac{\langle t, 1\rangle}{\|1\|^2}1.$$

Temos

$$\langle t, 1\rangle = (0)(1) + (1)(1) + (2)(1) = 3$$

e

$$\|1\|^2 = 1^2 + 1^2 + 1^2 = 3.$$

Assim, $v_2 = t - 1$. Como $\{v_1, v_2\}$ é uma base ortogonal de $\mathcal{P}_1(\mathbb{R})$, segue que

$$\operatorname{proj}_{\mathcal{P}_1(\mathbb{R})} q = \frac{\langle q, v_1\rangle}{\|v_1\|^2}v_1 + \frac{\langle q, v_2\rangle}{\|v_2\|^2}v_2.$$

Agora,

$$\begin{aligned}
\langle q, v_1\rangle &= \left\langle \frac{1}{2}(t^2+3t), 1\right\rangle = (0)(1) + (2)(1) + (5)(1) = 7\\
\|v_1\|^2 &= \|1\|^2 = 3 \qquad \text{como já vimos}\\
\langle q, v_2\rangle &= \left\langle \frac{1}{2}(t^2+3t), t-1\right\rangle = (0)(-1) + (2)(0) + (5)(1) = 5\\
\|v_2\|^2 &= \|t-1\|^2 = (-1)^2 + 0^2 + 1^2 = 2.
\end{aligned}$$

Logo,

$$\operatorname{proj}_{\mathcal{P}_1(\mathbb{R})} = \frac{7}{3}1 + \frac{5}{2}(t-1) = \frac{5}{2}t - \frac{1}{6},$$

que é, de todos os vetores de $\mathcal{P}_1(\mathbb{R})$, o mais próximo de q.

Podemos interpretar o resultado da seguinte maneira. Pretende-se determinar dentre todas as retas (que são os gráficos de elementos de $\mathcal{P}_1(\mathbb{R})$) aquela que "melhor se ajusta" à parábola $y = \frac{1}{2}(x^2 + 3x)$ nos pontos $x_1 = 0$, $x_2 = 1$ e $x_3 = 2$. O que vimos é que a reta procurada é a de equação $y = \frac{5}{2}x - \frac{1}{6}$. $\diamond$

Observação Se W é um subespaço de dimensão finita de um espaço vetorial com produto interno e $u \in V$, é possível encontrar a projeção de u sobre W a partir de uma base arbitrária, não necessariamente ortogonal de W. Indicamos ao leitor interessado nesse método a leitura das pp. 130–131 de [3].

Exercícios Lista 1 - Álgebra Linear II: Exs. 57–74.

5.3 O complemento ortogonal

Dado um espaço vetorial V com produto interno e um subconjunto X de V, definimos o *subespaço ortogonal* a X como sendo o seguinte subconjunto de V:

$$X^\perp = \left\{ v \in V \mid v \perp x, \text{ para todo } x \in X \right\}.$$

Por convenção, definimos $\varnothing^\perp = V$.

Observe que na definição não se exige que X seja um subespaço de V; X pode ser qualquer subconjunto de V. Apesar disso, $X^\perp$ é sempre um subespaço, o que justifica a nomenclatura introduzida. Esse é o conteúdo do nosso próximo resultado.

Proposição 5.3.1 *Seja V um espaço vetorial com produto interno e seja X um subconjunto de V. Então, $X^\perp$ é um subespaço de V.*

Demonstração Basta considerar o caso $X \neq \varnothing$. Primeiramente, $0_V \in X^\perp$, porque $0_V \perp v$, para todo $v \in V$, em particular, $0_V \perp x$, para todo $x \in X$. Sejam, agora, $u, v \in X^\perp$. É preciso mostrar que $u + v \in X^\perp$. Com efeito, para todo $x \in X$, temos $\langle u + v, x\rangle = \langle u, x\rangle + \langle v, x\rangle = 0 + 0 = 0$. Logo, de fato, $u + v \in X^\perp$. Finalmente, mostremos que $X^\perp$ é fechado para multiplicação por escalares: sejam $\lambda \in \mathbb{R}$ e $u \in X^\perp$; então, para todo $x \in X$, $\langle \lambda u, x\rangle = \lambda \langle u, x\rangle = \lambda 0 = 0$, o que mostra que $\lambda u \in X^\perp$. Assim, conclui-se que $X^\perp$ é um subespaço de V. □

Sendo um subespaço, podemos aplicar ferramentas e conceitos da Álgebra Linear a $X^\perp$. Antes de tratar de um exemplo, notemos a seguinte

Observação Seja V um espaço vetorial com produto interno e seja X um subconjunto finito de V. Então, $X^\perp = [X]^\perp$. (De fato, como $X \subseteq [X]$, então $[X]^\perp \subseteq X^\perp$. Para a outra inclusão, suponha que $X = \{u_1, \dots, u_n\}$, seja $v \in X^\perp$ e seja $w \in [X]$. Então, existem $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tais que $w = \lambda_1 u_1 + \dots + \lambda_n u_n$. Assim, $\langle v, w\rangle = \langle v, \lambda_1 u_1 + \dots + \lambda_n u_n\rangle = \lambda_1 \langle v, u_1\rangle + \dots + \lambda_n \langle v, u_n\rangle = \lambda_1 0 + \dots + \lambda_n 0 = 0$. Logo, $v \in [X]^\perp$, o que prova que vale, também $X^\perp \subseteq [X]^\perp$.)

Exemplo 5.3.2 Considere, em $\mathbb{R}^4$ o produto interno definido por

$$\Big\langle (a_1, a_2, a_3, a_4), (b_1, b_2, b_3, b_4) \Big\rangle = 2a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4.$$

Encontre uma base para $S^\perp$, em que $S = \big[(1, 2, 0, -1)\big]$.

Solução: Primeiramente, convença-se que a função definida é, de fato, um produto interno em $\mathbb{R}^4$. Em segundo lugar, como vimos na observação que precede este exemplo, $S^\perp = \{(1, 2, 0, -1)\}^\perp$. Agora, dado $v = (x, y, z, w) \in \mathbb{R}^4$, temos $v \in \{(1, 2, 0, -1)\}^\perp$ se, e somente se, $v \perp (1, 2, 0, -1)$, isto é, se e somente se, $\Big\langle (x, y, z, w), (1, 2, 0, -1) \Big\rangle = 0$. Portanto,

$$v = (x, y, z, w) \in S^\perp \quad \text{se, e somente se,} \quad 2x + 2y - w = 0.$$

Assim, $S^\perp = \{(x, y, z, w) \in \mathbb{R}^4 \mid 2x + 2y - w = 0\}$. Agora, basta argumentarmos como vimos fazendo desde a Seção 4.6. Os vetores de $S^\perp$ são os vetores da forma

$$\left(-\alpha + \frac{\gamma}{2}, \alpha, \beta, \gamma\right) = \alpha(-1, 1, 0, 0) + \beta(0, 0, 1, 0) + \gamma\left(\frac{1}{2}, 0, 0, 1\right), \text{ com } \alpha, \beta, \gamma \in \mathbb{R},$$

já que essas são as soluções do sistema linear $2x + 2y - w = 0$ (com uma equação, nas incóginitas x, y, z, w). Assim,

$$S^{\perp} = \left[(-1,1,0,0), (0,0,1,0), \left(\frac{1}{2}, 0, 0, 1\right) \right].$$

O conjunto gerador de $S^{\perp}$ apresentado acima é LI e, portanto, uma base dele. Em particular, $\dim(S^{\perp}) = 3$. $\diamond$

Note que, no exemplo, a dimensão de $S^{\perp}$ é complementar (em relação à dimensão do espaço ambiente $\mathbb{R}^4$) à dimensão de S: $\dim(S) + \dim(S^{\perp}) = 1 + 3 = 4 = \dim(\mathbb{R}^4)$. Isso é um fato geral, como mostra o próximo resultado.

Teorema 5.3.3 *Seja V um espaço vetorial com produto interno e seja W um subespaço de dimensão finita de V. Então, $V = W \oplus W^{\perp}$.*

Em particular, se V tem dimensão finita, $\dim(V) = \dim(W) + \dim(W^{\perp})$*, para todo subespaço W de V.*

Demonstração A última afirmação é uma propriedade de somas diretas (ver Proposição4.8.2).

Provemos a decomposição $V = W \oplus W^{\perp}$. É preciso mostrar duas coisas: que $V = W + W^{\perp}$ e que $W \cap W^{\perp} = \{0_V\}$. Comecemos pela primeira. Dado $u \in V$, como W tem dimensão finita, o Teorema 5.2.12 garante que existe a projeção ortogonal de u sobre W. Assim, temos

$$u = (\operatorname{proj}_W u) + (u - \operatorname{proj}_W u).$$

Como $\operatorname{proj}_W u \in W$ e $u - \operatorname{proj}_W u \in W^{\perp}$, segue $u \in W + W^{\perp}$. Assim, todo vetor de V se decompõe como soma de um vetor em W e um vetor em $W^{\perp}$. Logo, $V = W + W^{\perp}$. Para mostrar que a soma é direta, tome $v \in W \cap W^{\perp}$. Então, $\langle v, v \rangle = 0$ (pela própria definição de $W^{\perp}$). Logo, $v = 0_V$, o que implica que $W \cap W^{\perp} = \{0_V\}$, como queríamos. $\square$

Como a decomposição de um vetor em uma soma de termos de uma soma direta é única (pela Proposição 4.8.4), segue que a decomposição que encontramos, na demonstração, para um vetor de V como soma de um elemento de W e um elemento de $W^{\perp}$ é única.

Em vista desse teorema, $W^{\perp}$ é, às vezes, chamado de *complemento ortogonal* de W.

Exemplo 5.3.4 Em $\mathbb{R}^3$ com o produto interno usual, decomponha $v = (1,2,2)$ como uma soma $v = x + y$, em que $x \in W = \big[(1,1,0), (0,1,1)\big]$ e $y \in W^{\perp}$.

Solução: Vimos, na demonstração do Teorema 5.3.3, que devemos tomar $x = \operatorname{proj}_W v$ e $y = v - x$. O conjunto $\{(1,1,0), (0,1,1)\}$ é uma base de W, mas não é ortogonal. Aplicando a ele o processo de ortogonalização de Gram-Schmidt, obtemos

$$\begin{aligned}
v_1 &= (1,1,0), \\
v_2 &= (0,1,1) - \frac{\big\langle (0,1,1), (1,1,0) \big\rangle}{\|(1,1,0)\|^2}(1,1,0) = (0,1,1) - \frac{1}{2}(1,1,0) = \left(-\frac{1}{2}, \frac{1}{2}, 1\right).
\end{aligned}$$

Assim $\{v_1, v_2\}$ é uma base ortogonal de W. Portanto,

$$\begin{aligned}
\operatorname{proj}_W v &= \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2} v_2 \\
&= \frac{\big\langle (1,2,2), (1,1,0) \big\rangle}{\|(1,1,0)\|^2}(1,1,0) + \frac{\Big\langle (1,2,2), \left(-\frac{1}{2}, \frac{1}{2}, 1\right) \Big\rangle}{\Big\|\left(-\frac{1}{2}, \frac{1}{2}, 1\right)\Big\|^2} \left(-\frac{1}{2}, \frac{1}{2}, 1\right) \\
&= \frac{3}{2}(1,1,0) + \frac{5/2}{3/2}\left(-\frac{1}{2}, \frac{1}{2}, 1\right) \\
&= \left(\frac{2}{3}, \frac{7}{3}, \frac{5}{3}\right).
\end{aligned}$$

Logo, a decomposição procurada é $v = x + y$, com

$$x = \text{proj}_W\, v = \left(\frac{2}{3}, \frac{7}{3}, \frac{5}{3}\right) \in W,$$

$$y = v - x = (1,2,2) - \left(\frac{2}{3}, \frac{7}{3}, \frac{5}{3}\right) = \left(\frac{1}{3}, -\frac{1}{3}, \frac{1}{3}\right) \in W^{\perp}. \quad \diamond$$

Exemplo 5.3.5 (Prova 1, Álgebra Linear II, 2011) Seja V um espaço vetorial com produto interno e sejam S e W subespaços de V. Assinale a afirmação **falsa**.

(A) Se $u \in S$, $v \in S^{\perp}$ e $u + v = 0_V$, então $u = v = 0_V$.
(B) Se $S \subseteq W$, então $W^{\perp} \subseteq S^{\perp}$.
(C) Se $\{u_1, \dots, u_n\}$ é uma base ortogonal de V, então $v = \sum_{j=1}^{n} \text{proj}_{[u_j]}\, v$, para todo $v \in V$.
(D) Se $S \subseteq W$, então $S^{\perp} \subseteq W^{\perp}$.
(E) Se $A = \{u_1, \dots, u_p\} \subseteq S$, $B = \{v_1, \dots, v_q\} \subseteq S^{\perp}$ e A e B são linearmente independentes, então $A \cup B$ é linearmente independente.

Solução: A afirmação falsa ocorre na alternativa (D); por exemplo, em qualquer espaço vetorial V não nulo, temos $\{0_V\}^{\perp} = V$ (pois todo vetor de V é ortogonal ao vetor nulo) e $V^{\perp} = \{0_V\}$ (pois se $v \in V$ é tal que $v \in V^{\perp}$, então $\|v\|^2 = \langle v, v\rangle = 0$). Assim, tomando $S = \{0_V\}$ e $W = V$, obtemos um contraexemplo de (D).

Mostremos, agora, que as demais afirmações são verdadeiras.

(A) Se $u + v = 0_V$, então, $u = -v$. Como $v \in S^{\perp}$ e $S^{\perp}$ é um subespaço de V segue que $-v \in S^{\perp}$. Logo temos um vetor $u = -v$ que pertence, simultaneamente, a S e a $S^{\perp}$. Portanto, ele é ortogonal a si mesmo: $0_V = \langle u, u\rangle = \|u\|^2$. Segue que $u = 0_V$ e, também, $v = -u = -0_V = 0_V$.

(B) Para mostrar que $W^{\perp} \subseteq S^{\perp}$, é preciso mostrar que todo vetor de $W^{\perp}$ é ortogonal a todo vetor de S. Mas isso é válido, uma vez que todo vetor de $W^{\perp}$ é ortogonal a todo vetor de W e W contém S.

(C) Dado $v \in V$, sejam $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tais que $v = \sum_{j=1}^{n} \lambda_j u_j$. Para cada $i = 1, \dots, n$, temos $\langle v, u_i\rangle = \left\langle \sum_{j=1}^{n} \lambda_j u_j, u_i \right\rangle = \sum_{j=1}^{n} \lambda_j \langle u_j, u_i\rangle = \lambda_i \|u_i\|^2$, uma vez que os vetores de $\{u_1, \dots, u_n\}$ são dois a dois ortogonais entre si. Segue que $\lambda_i = \frac{\langle v, u_i\rangle}{\|u_i\|^2}$. Portanto, $\lambda_i u_i = \frac{\langle v, u_i\rangle}{\|u_i\|^2} u_i = \text{proj}_{[u_i]}\, v$.

(E) Sejam $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q \in \mathbb{R}$ tais que $\sum_{i=1}^{p} \alpha_i u_i + \sum_{j=1}^{q} \beta_j v_j = 0_V$. A tarefa é mostrar que todos os α_i e todos os β_j são iguais a zero. Denote $u = \sum_{i=1}^{p} \alpha_i u_i$ e $v = \sum_{j=1}^{q} \beta_j v_j$. Então, $u \in S$ e $v \in S^{\perp}$, já que $A \subseteq S$ e $B \subseteq S^{\perp}$. Pelo que vimos na alternativa (a) isso implica que $0_V = u = \sum_{i=1}^{p} \alpha_i u_i$ e que $0_V = v = \sum_{j=1}^{q} \beta_j v_j$. Como A e B são LI, segue que $\alpha_i = 0$, para todo $i = 1, \dots, p$, e $\beta_j = 0$, para todo $j = 1, \dots, q$.

Resposta: (D) $\diamond$

Exercícios Lista 1 - Álgebra Linear II: Exs. 75–82.

Capítulo 6
Transformações lineares

Neste capítulo, estudaremos relações entre espaços vetoriais possivelmente distintos. A ferramenta para isso serão funções entre os espaços vetoriais que respeitam as operações.

6.1 Definição e exemplos

Definição Sejam U e V espaços vetoriais. Dizemos que uma função $T\colon U \to V$ é uma *transformação linear* se satisfizer as seguintes condições:

(TL-1) $T(u_1 + u_2) = T(u_1) + T(u_2)$, para todos $u_1, u_2 \in U$;
(TL-2) $T(\lambda u) = \lambda T(u)$, para todos $u \in U$, $\lambda \in \mathbb{R}$.

Embora os espaços U e V sejam possivelmente diferentes, e, portanto, possuam operações de soma e multiplicação por escalar diferentes, observe que são utilizadas as mesmas notações (+ e concatenação) para elas nos dois espaços. Em geral, o próprio contexto impedirá interpretações ambíguas.

A seguir, veremos uma série de exemplos de transformações lineares (e, também, de algumas funções que não são transformações lineares.)

Exemplo 6.1.1 A função $T\colon \mathbb{R}^3 \to \mathcal{P}_2(\mathbb{R})$, definida por $T(a,b,c) = a + bt + ct^2$, para todo $(a,b,c) \in \mathbb{R}^3$, é uma transformação linear.[1] Mostremos que, de fato, T satisfaz as condições da definição. Dados $(a_1,b_1,c_1), (a_2,b_2,c_2) \in \mathbb{R}^3$, temos

$$\begin{aligned} T\big((a_1,b_1,c_1) + (a_2,b_2,c_2)\big) &= T(a_1+a_2, b_1+b_2, c_1+c_2) \\ &= (a_1+a_2) + (b_1+b_2)t + (c_1+c_2)t^2 \\ &= (a_1 + b_1 t + c_1 t^2) + (a_2 + b_2 t + c_2 t^2) \\ &= T(a_1,b_1,c_1) + T(a_2,b_2,c_2). \end{aligned}$$

Logo, T satisfaz (TL-1). Para a segunda condição, tome $(a,b,c) \in \mathbb{R}^3$ e $\lambda \in \mathbb{R}$. Então,

$$\begin{aligned} T\big(\lambda(a,b,c)\big) &= T(\lambda a, \lambda b, \lambda c) \\ &= (\lambda a) + (\lambda b)t + (\lambda c)t^2 \\ &= \lambda(a + bt + ct^2) \\ &= \lambda T(a,b,c). \end{aligned}$$

[1] Fizemos, aqui, um pequeno abuso da notação usual de função: deveríamos, a rigor, escrever $T\big((a,b,c)\big)$ para a imagem pela função T do elemento (a,b,c) de seu domínio. Essa simplificação notacional sempre será adotada quando não implicar ambiguidades.

Portanto, T satisfaz, também (TL-2). Assim, T é uma transformação linear.

(É comum denotar uma função, em termos de seu domínio, contradomínio e da definição da imagem de um elemento de seu domínio, por um diagrama, como o que foi utilizado, por exemplo, na definição de espaço vetorial no início da Seção 4.1. Assim, aqui, poderíamos ter denotado T e sua definição por

$$\begin{aligned} T\colon \mathbb{R}^3 &\longrightarrow \mathcal{P}_2(\mathbb{R}) \\ (a,b,c) &\longmapsto a+bt+ct^2 \end{aligned}$$

Daí, poderíamos concluir, por exemplo, que $T(1,-1,0) = 1-t$ e que $T(0,-2,3) = -2t+3t^2$.)

Exemplo 6.1.2 A função

$$\begin{aligned} L\colon \mathbb{R}^2 &\longmapsto \mathbb{R}^3 \\ (x,y) &\longmapsto (2x, x+y, 3y) \end{aligned}$$

é uma transformação linear. Com efeito, L satisfaz (TL-1), pois

$$\begin{aligned} L\big((x_1,y_1)+(x_2,y_2)\big) &= L(x_1+x_2, y_1+y_2) \\ &= \big(2(x_1+x_2), (x_1+x_2)+(y_1+y_2), 3(y_1+y_2)\big) \\ &= (2x_1+2x_2, x_1+y_1+x_2+y_2, 3y_1+3y_2) \\ &= (2x_1, x_1+y_1, 3y_1) + (2x_2, x_2+y_2, 3y_2) \\ &= L(x_1,y_1)+L(x_2,y_2), \end{aligned}$$

quaisquer que sejam $(x_1,y_1),(x_2,y_2) \in \mathbb{R}^2$. E L satisfaz (TL-2), uma vez que

$$\begin{aligned} L\big(\lambda(x,y)\big) &= L(\lambda x, \lambda y) \\ &= (2\lambda x, \lambda x+\lambda y, 3\lambda y) \\ &= \lambda(2x, x+y, 3y) \\ &= \lambda L(x,y), \end{aligned}$$

para todos $(x,y) \in \mathbb{R}^2$ e $\lambda \in \mathbb{R}$.

Algumas transformações geométricas familiares são lineares, como é o caso do próximo exemplo.

Exemplo 6.1.3 No plano cartesiano, a reflexão em relação ao eixo x é uma transformação linear. De fato, essa transformação é dada por

$$\begin{aligned} S\colon \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x,y) &\longmapsto (x,-y) \end{aligned}$$

e é linear porque $S\big((x,y)+(x',y')\big) = S(x+x', y+y') = \big(x+x', -(y+y')\big) = (x+x', -y-y') = (x,-y)+(x',-y') = S(x,y)+S(x',y')$ e $S\big(\lambda(x,y)\big) = S(\lambda x, \lambda y) = (\lambda x, -\lambda y) = \lambda(x,-y) = \lambda S(x,y)$, para todos $(x,y),(x',y') \in \mathbb{R}^2$ e $\lambda \in \mathbb{R}$.

Se V é um espaço vetorial, uma transformação linear $T\colon V \to V$ é, também, designada um *operador linear* de V. Esse é o caso do exemplo anterior.

Algumas funções familiares do Cálculo também são lineares.

Exemplo 6.1.4 O operador de derivação no espaço dos polinômios

$$\begin{aligned} D\colon \mathcal{P}(\mathbb{R}) &\longrightarrow \mathcal{P}(\mathbb{R}) \\ p &\longmapsto p' \end{aligned}$$

é linear, pois, como visto no curso de Cálculo, $D(p+q) = (p+q)' = p'+q' = D(p)+D(q)$ e $D(\lambda p) = (\lambda p)' = \lambda p'$, para todos $p,q \in \mathcal{P}(\mathbb{R})$ e $\lambda \in \mathbb{R}$.

Exemplo 6.1.5 Se V é um espaço vetorial qualquer, então a função identidade $\mathrm{I}_V : V \to V$, definida por $\mathrm{I}_V(v) = v$, para todo $v \in V$, é uma transformação linear, chamada de *operador identidade* de V, como é fácil ver.

Se U e V são espaços vetoriais, também é fácil de ver que a função nula $N : U \to V$, dada por $N(u) = 0_V$, para todo $u \in U$, é uma transformação linear.

Exemplo 6.1.6 A função

$$\begin{aligned} F\colon \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x^2, 0) \end{aligned}$$

não é uma transformação linear, pois, por exemplo, não satisfaz a condição (TL-1): $F\big((1,0)+(1,1)\big) = F(2,1) = (4,0)$, ao passo que $F(1,0)+F(1,1) = (1,0)+(1,0) = (2,0)$. (Observe que a condição (TL-1) exigia que a igualdade expressa nela fosse satisfeita para todos os pares de elementos do domínio. Aqui, mostramos que ela não é satisfeita para uma determinada escolha de pares, o que é suficiente para garantir que F não é uma transformação linear. Note que F também não satisfaz (TL-2).)

Exemplo 6.1.7 A função

$$\begin{aligned} G\colon \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longmapsto (x+1, y+z) \end{aligned}$$

não é uma transformação linear, pois, por exemplo, não satisfaz a condição (TL-2): $G\big(2(1,1,3)\big) = G(2,2,6) = (3,8)$, mas $2G(1,1,3) = 2(2,4) = (4,8)$. (Verifique que G também não satisfaz (TL-1).)

Enunciamos, a seguir, duas propriedades satisfeitas por transformações lineares.

Proposição 6.1.8 *Sejam U e V espaços vetoriais e seja $T : U \to V$ uma transformação linear. Então,*

(i) $T(0_U) = 0_V$;
(ii) $T(-u) = -T(u)$, para todo $u \in U$.

Demonstração Temos $T(0_U) = T(0_U + 0_U) = T(0_U) + T(0_U)$, pois T é linear. Somando $-T(0_U)$ a ambos os lados da igualdade, obtemos (i). Para (ii), basta lembrar que $-u = (-1)u$, para todo $u \in U$ (ver Proposição 4.1.8). □

Esse resultado pode ser utilizado para mostrar que uma determinada função entre espaços vetorias não é uma transformação linear. Por exemplo, no Exemplo 6.1.7, temos $G(0,0,0) = (1,0) \neq (0,0)$. Logo, G não pode ser linear. Entretanto, esse critério deve ser aplicado com cuidado, pois há funções entre espaços vetoriais que preservam o vetor nulo, mas não são lineares. Por exemplo, a função $\det\colon M_2(\mathbb{R}) \to \mathbb{R}$ satisfaz $\det \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$, mas não é linear. (Você consegue mostrar isso?)

O próximo resultado diz que uma transformação linear está completamente determinada uma vez determinadas as imagens dos elementos em uma base do domínio.

Proposição 6.1.9 *Seja U um espaço vetorial de dimensão finita e seja V um espaço vetorial qualquer. Se $\{e_1, \dots, e_n\}$ é uma base de V e $v_1, \dots, v_n \in V$, então existe uma única transformação linear $T : U \to V$ tal que $T(e_i) = v_i$, para todo $i = 1, \dots, n$.*

Na proposição acima, observe que os vetores $v_1, \dots, v_n$ de V não precisam ser distintos.

Demonstração Há duas coisas a serem demonstradas, que (i) existe uma transformação linear T com a propriedade desejada, e que (ii) T é a única transformação linear com essa propriedade.

Comecemos por (i). Para definir uma função $T : U \to V$, dado $u \in U$, sejam $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ tais que $u = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n$, defina

$$T(u) = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n.$$

(Isto é, $T(u)$ é aquela combinação linear de $v_1, \dots, v_n$ cujos escalares são as coordenadas de u em relação à base $\{e_1, \dots, e_n\}$.) Então, $T(e_i) = v_i$, para todo $i = 1, \dots, n$, por definição. Mostremos que T é uma transformação linear. Sejam $u, u' \in U$ e sejam $\lambda_1, \dots, \lambda_n, \lambda'_1, \dots, \lambda'_n \in \mathbb{R}$ tais que

$$u = \lambda_1 e_1 + \cdots + \lambda_n e_n \quad \text{e} \quad u' = \lambda_1' e_1 + \cdots + \lambda_n' e_n.$$

Então, $u + u' = (\lambda_1 + \lambda_1')e_1 + \cdots + (\lambda_n + \lambda_n')e_n$ e, portanto,

$$\begin{aligned} T(u + u') &= (\lambda_1 + \lambda_1')v_1 + \cdots + (\lambda_n + \lambda_n')v_n \\ &= (\lambda_1 v_1 + \cdots + \lambda_n v_n) + (\lambda_1' v_1 + \cdots + \lambda_n' v_n) \\ &= T(u) + T(u'), \end{aligned}$$

e, se $\alpha \in \mathbb{R}$, então $\alpha u = (\alpha\lambda_1)e_1 + \cdots + (\alpha\lambda_n)e_n$, o que implica, pela definição de T, que

$$T(\alpha u) = (\alpha\lambda_1)v_1 + \cdots + (\alpha\lambda_n)v_n = \alpha(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \alpha T(u).$$

Logo, T é, de fato, linear e satisfaz, como vimos, a propriedade desejada.

(ii) Seja, agora, $S\colon U \to V$ uma transformação linear que satisfaz $S(e_i) = v_i$, para todo $i = 1, \dots, n$. Dado $u \in U$, se $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ são tais que $u = \lambda_1 e_1 + \cdots + \lambda_n e_n$, então,

$$\begin{aligned} S(u) &= S(\lambda_1 e_1 + \cdots + \lambda_n e_n) \\ &= \lambda_1 S(e_1) + \cdots + \lambda_n S(e_n) && \text{pois } S \text{ é linear} \\ &= \lambda_1 v_1 + \cdots + \lambda_n v_n \\ &= T(u). \end{aligned}$$

Logo, S e T são funções iguais (isto é, elas têm mesmo domínio, mesmo contradomínio e as imagens em cada elemento do domínio coincidem). Portanto, T é a única transformação linear de U em V que manda e_i em v_i, para todo $i = 1, \dots, n$. □

Esse resultado é especialmente útil para construirmos exemplos de transformações lineares.

Exemplo 6.1.10 Encontre uma transformação linear $T\colon \mathbb{R}^4 \to \mathcal{P}_2(\mathbb{R})$ que satisfaça $T(1, 1, 1, 1) = t^2$.

Solução: Vimos, na Proposição 6.1.9, que basta fixarmos as imagens dos elementos de uma base de $\mathbb{R}^4$ para definir uma transformação linear com domínio $\mathbb{R}^4$. Como queremos ter controle sobre a imagem de $(1, 1, 1, 1)$, tomemos uma base de $\mathbb{R}^4$ que contém esse elemento, por exemplo, $\mathcal{B} = \{(1, 1, 1, 1), (1, 1, 1, 0), (1, 1, 0, 0), (1, 0, 0, 0)\}$. (Convença-se que, de fato, $\mathcal{B}$ é uma base de $\mathbb{R}^4$.) Agora, basta escolher as imagens, em $\mathcal{P}_2(\mathbb{R})$, desses quatro vetores, de modo a respeitar a condição exigida. Uma possibilidade é definir

$$T(1, 1, 1, 1) = t^2; \quad T(1, 1, 1, 0) = t + t^2; \quad T(1, 1, 0, 0) = 1; \quad T(1, 0, 0, 0) = 2 - 3t + 4t^2. \tag{6.1}$$

A Proposição 6.1.9 garante que existe uma transformação linear T com essa propriedade (e que ela é a única satisfazendo (6.1)). (Essa escolha para as imagens dos elementos de $\mathcal{B}$ é totalmente arbitrária, desde que $(1, 1, 1, 1)$ seja enviado em t^2, como requerido. Cada escolha dará origem a uma transformação linear diferente, todas satisfazendo a condição desejada.)

Para darmos uma expressão explícita para a imagem por T de um elemento arbitrário de $\mathbb{R}^4$, basta conhecer suas coordenadas em relação à base $\mathcal{B}$. Assim, dado $(a, b, c, d) \in \mathbb{R}^4$, temos

$$(a, b, c, d) = d(1, 1, 1, 1) + (c - d)(1, 1, 1, 0) + (b - c)(1, 1, 0, 0) + (a - b)(1, 0, 0, 0).$$

Logo,

$$\begin{aligned} T(a, b, c, d) &= T\big(d(1, 1, 1, 1) + (c - d)(1, 1, 1, 0) + (b - c)(1, 1, 0, 0) + (a - b)(1, 0, 0, 0)\big) \\ &= dT(1, 1, 1, 1) + (c - d)T(1, 1, 1, 0) + (b - c)T(1, 1, 0, 0) + (a - b)T(1, 0, 0, 0) && \text{pois } T \text{ é linear} \\ &= dt^2 + (c - d)(t + t^2) + (b - c)1 + (a - b)(2 - 3t + 4t^2) && \text{por (6.1)} \\ &= (2a - b - c) + (-3a + 3b + c - d)t + (4a - 4b + c)t^2. \end{aligned}$$

Veja que, como queríamos, $T(1,1,1,1) = t^2$. $\diamond$

Exercícios Lista 2 - Álgebra Linear II: Exs. 1–4.

6.2 Núcleo e imagem

Estão associados a uma transformação linear dois subespaços vetoriais, definidos como segue.

Definição Sejam U e V espaços vetoriais e seja $T\colon U \to V$ uma transformação linear. Definimos o *núcleo* de T como sendo o seguinte subconjunto de U:

$$\operatorname{Ker}(T) = \left\{ u \in U \mid T(u) = 0_V \right\}.$$

Definimos a *imagem* de T como sendo o seguinte subconjunto de V:

$$\operatorname{Im}(T) = \left\{ v \in V \mid \text{existe algum } u \in U \text{ tal que } v = T(u) \right\}.$$

Proposição 6.2.1 *Sejam U e V espaços vetoriais e seja $T\colon U \to V$ uma transformação linear. Então,* $\operatorname{Ker}(T)$ *é um subespaço de U, e* $\operatorname{Im}(T)$ *é um subespaço de V.*

Demonstração Comecemos mostrando que o núcleo de T é um subespaço de U. Em primeiro lugar, $0_U \in \operatorname{Ker}(T)$, pois, como vimos na Proposição 6.1.8, $T(0_U) = 0_V$. Agora sejam $u, u' \in \operatorname{Ker}(T)$ e $\lambda \in \mathbb{R}$. Mostremos que $u+u', \lambda u \in \operatorname{Ker}(T)$. Pela definição de núcleo, temos $T(u) = 0_V$ e $T(u') = 0_V$. Então, como T é linear, $T(u+u') = T(u)+T(u') = 0_V + 0_V = 0_V$, o que prova que $u + u' \in \operatorname{Ker}(T)$. Finalmente, $T(\lambda u) = \lambda T(u) = \lambda 0_V = 0_V$. Portanto, $\lambda u \in \operatorname{Ker}(T)$. Logo, $\operatorname{Ker}(T)$ é um subespaço de U.

Tratemos, agora, da imagem de T. Como $T(0_U) = 0_V$, segue que $0_V \in \operatorname{Im}(T)$. Agora, sejam $v, v' \in \operatorname{Im}(T)$ e seja $\lambda \in \mathbb{R}$. Então, pela definição de imagem, existem $u, u' \in U$ tais que $v = T(u)$ e $v' = T(u')$. Assim, $v + v' = T(u) + T(u') = T(u + u') \in \operatorname{Im}(T)$, e $\lambda v = \lambda T(u) = T(\lambda u) \in \operatorname{Im}(T)$. Segue que $\operatorname{Im}(T)$ é um subespaço de V. $\square$

Exemplo 6.2.2 Determinemos, nos Exemplos 6.1.1–6.1.5, o núcleo e a imagem de cada uma das transformações lineares lá definidas.

Para a transformação linear $T\colon \mathbb{R}^3 \to \mathcal{P}_2(\mathbb{R})$, dada por $T(a,b,c) = a + bt + ct^2$, temos que $(a,b,c) \in \operatorname{Ker}(T)$ se, e somente se, $a + bt + ct^2 = T(a,b,c) = 0$, o que ocorre se, e somente se $a = b = c = 0$. Portanto, $\operatorname{Ker}(T) = \{(0,0,0)\}$. É fácil ver que $\operatorname{Im}(T) = \mathcal{P}_2(\mathbb{R})$, pois qualquer que seja $q \in \mathcal{P}_2(\mathbb{R})$, digamos, $q = a_0 + a_1 t + a_2 t^2$, temos $q = T(a_0, a_1, a_2) \in \operatorname{Im}(T)$.

Vejamos, agora, $L\colon \mathbb{R}^2 \to \mathbb{R}^3$, definida por $L(x,y) = (2x, x+y, 3y)$. Temos que $(x,y) \in \operatorname{Ker}(L)$ se, e somente se, $(2x, x+y, 3y) = T(x,y) = (0,0,0)$, o que implica $x = y = 0$. Portanto, $\operatorname{Ker}(L) = \{(0,0)\}$. Para a determinar imagem de L, tome $v = (a,b,c) \in \mathbb{R}^3$ e vejamos quais as condições necessárias e suficientes para que $v \in \operatorname{Im}(L)$. Temos que $v \in \operatorname{Im}(L)$ se, e somente se, existe $u = (x,y) \in \mathbb{R}^2$ tal que $v = L(u)$, ou seja, se e somente se, existem $x, y \in \mathbb{R}$ tais que $(a,b,c) = L(x,y) = (2x, x+y, 3y)$. Portanto, $v \in \operatorname{Im}(L)$ se, e somente se, existir solução para o sistema linear

$$\begin{cases} 2x = a \\ x + y = b \\ 3y = c \end{cases} \tag{6.2}$$

nas variáveis x, y, z. Escalonando a matriz aumentada do sistema, obtemos

$$\left[\begin{array}{cc|c} 2 & 0 & a \\ 1 & 1 & b \\ 0 & 3 & c \end{array}\right] \to \left[\begin{array}{cc|c} 2 & 0 & a \\ 0 & 1 & -\frac{a}{2} + b \\ 0 & 3 & c \end{array}\right] \to \left[\begin{array}{cc|c} 2 & 0 & a \\ 0 & 1 & -\frac{a}{2} + b \\ 0 & 0 & \frac{3a}{2} - 3b + c \end{array}\right]$$

Assim, (6.2) tem solução se, e só se, $\frac{3a}{2} - 3b + c = 0$. Portanto, $\operatorname{Im}(L) = \left\{ (a,b,c) \in \mathbb{R}^3 \mid \frac{3a}{2} - 3b + c = 0 \right\}$, que é um subespaço de dimensão 2 de $\mathbb{R}^3$.

Para a reflexão $S\colon \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (x, -y)$, é fácil de ver que $\mathrm{Ker}(S) = \{(0, 0)\}$ e que $\mathrm{Im}(S) = \mathbb{R}^2$ (o que podia ser antecipado por argumentos puramente geométricos: o único vetor que é enviado na origem é a própria origem, e todo ponto do plano é reflexão de sua própria reflexão).

Consideremos o operador de derivação $D\colon \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$, $p \mapsto p'$. Se $p = a_0 + a_1 t + \cdots + a_n t^n \in \mathcal{P}(\mathbb{R})$ é tal que $p' = T(p) = 0$, então $a_1 + 2a_2 t + \cdots + na_n t^{n-1} = 0$, o que implica $a_1 = a_2 = \cdots = a_n = 0$. Assim, $\mathrm{Ker}(D) = \{a_0 \mid a_0 \in \mathbb{R}\} = \mathcal{P}_0(\mathbb{R})$. (Isso já era conhecido do Cálculo: os únicos polinômios de derivada nula são os constantes.) Já $\mathrm{Im}(D) = \mathcal{P}(\mathbb{R})$, uma vez que qualquer polinômio é derivada de um polinômio, uma primitiva dele: $a_0 + a_1 t + \cdots + a_n t^n = D\big(a_0 t + \frac{a_1}{2} t^2 + \cdots + \frac{a_n}{n+1} t^{n+1}\big)$.

Finalmente, $\mathrm{Ker}(\mathrm{I}_V) = \{0_V\}$ e $\mathrm{Im}(\mathrm{I}_V) = V$, como é fácil ver. E, se $N\colon U \to V$ é a transformação nula, então $\mathrm{Ker}(N) = U$ (todo vetor de U é mandado, via N, para 0_V) e $\mathrm{Im}(N) = \{0_V\}$ (o único vetor de V que é atingido via N é o vetor nulo).

Como núcleo e imagem são subespaços, podemos utilizar as ferramentas desenvolvidas para o estudo de espaço vetoriais para analisá-los. Por exemplo, podemos tratar de bases e dimensões para eles.

Exemplo 6.2.3 Encontre uma base para o núcleo da transformação linear $T\colon \mathbb{R}^3 \to \mathbb{R}^2$, definida por $T(x, y, z) = (x + y + z, 3x - 2y)$, para todo $(x, y, z) \in \mathbb{R}^3$.

Solução: Deixamos a cargo do leitor a verificação de que T, assim definida, é, de fato, linear. Seja $u = (x, y, z) \in \mathbb{R}^3$. Então,

$$
\begin{aligned}
u \in \mathrm{Ker}(T) &\iff T(u) = 0_{\mathbb{R}^2} \\
&\iff (x + y + z, 3x - 2y) = T(x, y, z) = (0, 0) \\
&\iff \begin{cases} x + y + z = 0 \\ 3x - 2y = 0 \end{cases}
\end{aligned}
$$

Como as soluções desse sistema são dadas por $(-\frac{2z}{5}, -\frac{3z}{5}, z)$, com $z \in \mathbb{R}$, como o leitor pode verificar, segue que $u \in \mathrm{Ker}(T)$ se, e somente se, existir $z \in \mathbb{R}$ tal que $u = (-\frac{2z}{5}, -\frac{3z}{5}, z) = z(-\frac{2}{5}, -\frac{3}{5}, 1)$. Assim, $\mathrm{Ker}(T) = \big[(-\frac{2}{5}, -\frac{3}{5}, 1)\big] = \big[(-2, -3, 5)\big]$. Como esse conjunto gerador unitário é LI, segue que $\big\{ (-2, -3, 5) \big\}$ é uma base de $\mathrm{Ker}(T)$. Em particular, $\dim\big(\mathrm{Ker}(T)\big) = 1$. $\diamond$

Exemplo 6.2.4 Encontre uma base para a imagem da transformação linear

$$
\begin{aligned}
T\colon \mathcal{P}_2(\mathbb{R}) &\longrightarrow \mathcal{P}(\mathbb{R}) \\
p(t) &\longmapsto tp'(t) + t^2 p(t)
\end{aligned}
$$

Solução: A função T é uma transformação linear porque se $p(t), q(t) \in \mathcal{P}_2(\mathbb{R})$ e $\lambda \in \mathbb{R}$, então

$$
\begin{aligned}
T\big(p(t) + q(t)\big) &= t\big(p(t) + q(t)\big)' + t^2\big(p(t) + q(t)\big) \\
&= tp'(t) + tq'(t) + t^2 p(t) + t^2 q(t) \\
&= \big(tp'(t) + t^2 p(t)\big) + \big(tq'(t) + t^2 q(t)\big) \\
&= T\big(p(t)\big) + T\big(q(t)\big)
\end{aligned}
$$

e

$$
\begin{aligned}
T\big(\lambda p(t)\big) &= t\big(\lambda p(t)\big)' + t^2(\lambda p(t)) \\
&= \lambda t p'(t) + \lambda t^2 p(t) \\
&= \lambda\big(tp'(t) + t^2 p(t)\big) \\
&= \lambda T\big(p(t)\big).
\end{aligned}
$$

Vamos à determinação de sua imagem. Dado $f(t) \in \mathcal{P}(\mathbb{R})$, temos que

$$\begin{aligned}
f(t) \in \operatorname{Im}(T) \quad &\iff \quad \text{existe } p(t) = a_0 + a_1 t + a_2 t^2 \in \mathcal{P}_2(\mathbb{R}) \text{ tal que } f(t) = T\big(p(t)\big) = tp'(t) + t^2 p(t)\\
&\iff \quad \text{existem } a_0, a_1, a_2 \in \mathbb{R} \text{ tais que } f(t) = t(a_1 + 2a_2 t) + t^2(a_0 + a_1 t + a_2 t^2)\\
&\iff \quad \text{existem } a_0, a_1, a_2 \in \mathbb{R} \text{ tais que } f(t) = a_0 t^2 + a_1(t + t^3) + a_2(2t^2 + t^4)\\
&\iff \quad f(t) \in \big[t^2,\ t + t^3,\ 2t^2 + t^4\big].
\end{aligned}$$

Assim, $\operatorname{Im}(T) = \big[t^2,\ t+t^3,\ 2t^2+t^4\big]$. Como esse conjunto de três vetores é LI (são polinômios de graus distintos), segue que $\big\{\, t^2,\ t+t^3,\ 2t^2+t^4 \,\big\}$ é uma base de $\operatorname{Im}(T)$. Em particular $\dim\big(\operatorname{Im}(T)\big) = 3$. $\diamond$

Conhecer o núcleo de uma transformação linear permite que se afirme sobre sua eventual injetividade. (Referimos o leitor ao Apêndice B para os conceitos de injetividade, sobrejetividade e bijetividade no contexto de funções.)

Proposição 6.2.5 *Sejam U e V espaços vetoriais e seja $T\colon U \to V$ uma transformação linear. Então, T é injetora se, e somente se,* $\operatorname{Ker}(T) = \{0_U\}$.

Demonstração Suponha que T é injetora. Mostremos que $\operatorname{Ker}(T) = \{0_U\}$. Já sabemos que $\{0_U\} \subseteq \operatorname{Ker}(T)$ (afinal de contas, $\operatorname{Ker}(T)$ é um subespaço de U). Verifiquemos que vale a inclusão reversa. Seja $u \in U$ tal que $u \in \operatorname{Ker}(T)$. Então, $T(u) = 0_V = T(0_U)$, como vimos na Proposição 6.1.8. Como T é injetora, segue que $u = 0_U$. Portanto, também temos $\operatorname{Ker}(T) \subseteq \{0_U\}$, provando o que queríamos.

Reciprocamente, suponha que $\operatorname{Ker}(T) = \{0_U\}$. Mostremos que isso implica que T é injetora. Sejam $u, u' \in U$ tais que $T(u) = T(u')$. Então, $T(u - u') = T(u) - T(u') = 0_V$ (a primera igualdade segue do fato de T ser linear). Ou seja, $u - u' \in \operatorname{Ker}(T)$, que, por hipótese, contém apenas o vetor nulo de U. Portanto, $u - u' = 0_U$ e, assim, $u = u'$, provando que T é injetora. $\square$

O critério de injetividade contido na Proposição 6.2.5 é extremamente útil e será usado quase que exclusivamente quando tivermos que decidir se uma transformação linear é ou não injetora.

Exemplo 6.2.6 O operador linear

$$\begin{aligned} T\colon \mathcal{P}_n(\mathbb{R}) &\longrightarrow \mathcal{P}_n(\mathbb{R})\\ p &\longmapsto p' + p'' \end{aligned}$$

é injetor?

Solução: Note que a função T está bem definida, uma vez que se $p \in \mathcal{P}_n(\mathbb{R})$, então, certamente, $p' + p'' \in \mathcal{P}_n(\mathbb{R})$. Além disso, T é linear (como o leitor pode verificar). Para responder à pergunta, determinemos o núcleo de T. Dado $p = \sum_{i=0}^n a_i t^i \in \mathcal{P}_n(\mathbb{R})$, temos que $p \in \operatorname{Ker}(T)$ se, e somente se, $p' + p'' = T(p) = 0$, ou seja, se e somente se,

$$\sum_{i=1}^{n-1} \big(ia_i + (i+1)ia_{i+1}\big)t^{i-1} + na_n t^{n-1} = 0. \tag{6.3}$$

(Verifique essa expressão.) Comparando graus, do maior para o menor, obtemos que (6.3) só pode ocorrer se $a_1 = \cdots = a_n = 0$. Assim, $\operatorname{Ker}(T) = \{a_0 \mid a_0 \in \mathbb{R}\}$, que não é o subespaço nulo. Logo, T não é injetor. $\diamond$

Na sequência dos Exemplos 6.1.1–6.1.2, temos que são injetoras apenas T, L, S e I_V. (A transformação nula $N\colon U \to V$ é injetora se, e somente se, $\dim(U) = 0$, um caso, por assim dizer, "patológico".)

Veremos, agora, um exemplo extremamente importante de transformações lineares, por serem, em um sentido que ficará claro mais adiante, os exemplos paradigmáticos de transformações lineares entre espaços vetoriais de dimensão finita.

Exemplo 6.2.7 Seja $A \in M_{m\times n}(\mathbb{R})$ e considere a função $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$ dada por $T_A(v) = Av$, para todo $v \in \mathbb{R}^n$. (Aqui, estamos denotando os elementos de $\mathbb{R}^n$ e $\mathbb{R}^m$ por vetores-coluna, isto é, estamos fazendo a identificação natural entre $\mathbb{R}^n$ e $M_{n\times 1}(\mathbb{R})$ e entre $\mathbb{R}^m$ e $M_{m\times 1}(\mathbb{R})$ que já vínhamos fazendo desde o estudo de sistemas lineares no Capítulo 1.)

(i) Mostre que T_A é uma transformação linear.
(ii) Descreva $\operatorname{Ker}(T_A)$ e $\operatorname{Im}(T_A)$.

Solução: (i) Dados $v, v' \in \mathbb{R}^n$, temos $T_A(v + v') = A(v + v') = Av + Av' = T_A(v) + T_A(v')$; e, se $v \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, então $T_A(\lambda v) = A(\lambda v) = \lambda(Av) = \lambda T_A(v)$, o que prova que T_A é uma transformação linear.

(ii) Dado $v \in \mathbb{R}^n$, temos que $v \in \mathrm{Ker}(T_A)$ se, e somente se, $0_{\mathbb{R}^m} = T_A(v) = Av$, ou seja, se e somente se, v for solução do sistema linear homogêneo cuja matriz de coeficiente é A. Assim, $\mathrm{Ker}(T_A)$ coincide com o subespaço de $\mathbb{R}^n$ formado pelas soluções de $AX = 0$. Em particular, segue da Proposição 4.6.6 que $\dim\big(\mathrm{Ker}(T_A)\big) = n - t$, em que t é o número de pivôs de uma matriz escalonada obtida a partir de A por operações elementares sobre linhas.

Finalmente, dado $w \in \mathbb{R}^m$, temos que $w \in \mathrm{Im}(T_A)$ se, e somente se, existe $v \in \mathbb{R}^n$ tal que $w = T_A(v) = Av$, o que, por sua vez, é equivalente a w ser uma combinação linear das colunas de A (com coeficientes dados pelas entradas de v). Ou seja, $\mathrm{Im}(T_A)$ coincide com o subespaço de $\mathbb{R}^m$ gerado pelas colunas de A. Assim, $\dim\big(\mathrm{Im}(T_A)\big)$ é igual ao posto-coluna de A, que, por sua vez, é igual a t (ver Apêndice A para os detalhes.) $\diamond$

Veremos, na próxima seção, que essa relação entre as dimensões de $\mathrm{Ker}(T_A)$ e $\mathrm{Im}(T_A)$ é uma propriedade de todas as transformações lineares entre espaços vetoriais de dimensão finita.

Exemplo 6.2.8 Encontre uma base para o núcleo e uma base para a imagem de $T_A \colon \mathbb{R}^3 \to \mathbb{R}^2$, em que

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

Solução: Como vimos no exemplo anterior, $\mathrm{Ker}(T_A)$ é o conjunto formado pelas soluções do sistema linear homogêneo

$$\begin{bmatrix} 2 & 0 & -4 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Escalonando A, obtemos

$$\begin{bmatrix} 2 & 0 & -4 \\ -1 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} 2 & 0 & -4 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, x_2 e x_3 são variáveis livres, enquanto x_1 é variável pivô. As soluções do sistema são

$$(2x_3, x_2, x_3) = x_2(0, 1, 0) + x_3(2, 0, 1), \quad \text{com } x_2, x_3 \in \mathbb{R}.$$

Logo, $\mathrm{Ker}(T_A) = \big[(0, 1, 0), (2, 0, 1)\big]$. Como esse conjunto gerador com dois elementos é LI, ele é uma base para $\mathrm{Ker}(T_A)$. Vimos, também, que $\mathrm{Im}(T_A)$ é o subespaço de $\mathbb{R}^2$ gerado pelas colunas de A, ou seja,

$$\mathrm{Im}(T_A) = \big[(2, -1), (0, 0), (-4, 2)\big] = \big[(2, -1)\big].$$

Assim $\big\{ (2, -1) \big\}$ é uma base de $\mathrm{Im}(T_A)$. $\diamond$

Para transformações lineares arbitrárias, o seguinte resultado é bastante útil na tarefa de descrever a imagem.

Proposição 6.2.9 *Sejam U e V espaços vetoriais e seja $T \colon U \to V$ uma transformação linear. Se $u_1, u_2, \dots, u_n \in U$ são tais que $U = [u_1, u_2, \dots, u_n]$, então* $\mathrm{Im}(T) = \big[T(u_1), T(u_2), \dots, T(u_n)\big]$.

Demonstração Como $T(u_i) \in \mathrm{Im}(T)$, para todo $i = 1, \dots, n$, segue que $\big[T(u_1), T(u_2), \dots, T(u_n)\big] \subseteq \mathrm{Im}(T)$. Para a outra inclusão, tome $v \in V$ tal que $v \in \mathrm{Im}(T)$. Então, existe $u \in U$ tal que $v = T(u)$. Como $U = [u_1, u_2, \dots, u_n]$, existem $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ tais que $u = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n$. Assim,

$$\begin{aligned} v = T(u) &= T(\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n) \\ &= \lambda_1 T(u_1) + \lambda T(u_2) + \cdots + \lambda_n T(u_n) \qquad \text{pois } T \text{ é linear.} \end{aligned}$$

Logo, $v \in \big[T(u_1), T(u_2), \dots, T(u_n)\big]$, o que prova que $\mathrm{Im}(T) \subseteq \big[T(u_1), T(u_2), \dots, T(u_n)\big]$, como queríamos. $\square$

A Proposição 6.2.9 deve ser utilizada com cuidado. Se $\{u_1, \dots, u_n\}$ é uma base de U, então ela diz que $\{T(u_1), \dots, T(u_n)\}$ gera $\operatorname{Im}(T)$, mas, em geral, esse conjunto não será uma base de $\operatorname{Im}(T)$, pois não será LI.

Exemplo 6.2.10 Encontre uma base para a imagem da transformação linear

$$\begin{aligned} T\colon \mathbb{R}^3 &\longrightarrow \mathcal{P}_3(\mathbb{R}) \\ (a, b, c) &\longmapsto (a+b) + (a+c)t + (-b+c)t^2 + (a+c)t^3 \end{aligned}$$

Solução: Convença-se que T é, de fato, uma transformação linear. A base canônica $\{(1,0,0), (0,1,0), (0,0,1)\}$ de $\mathbb{R}^3$ é, em particula, um conjunto gerador. Pela Proposição 6.2.9, $\operatorname{Im}(T) = \big[T(1,0,0), T(0,1,0), T(0,0,1)\big]$. Por definição da transformação linear T,

$$T(1,0,0) = 1 + t + t^3, \qquad T(0,1,0) = 1 - t^2, \qquad T(0,0,1) = t + t^2 + t^3.$$

Logo, $\{1+t+t^3, 1-t^2, t+t^2+t^3\}$ gera $\operatorname{Im}(T)$, mas não é uma base, pois não é LI: $t+t^2+t^3 = (1+t+t^3) - (1-t^2)$. Por causa dessa relação de dependência linear existente entre os geradores de $\operatorname{Im}(T)$, sabemos que $\operatorname{Im}(T)$ pode ser gerado por $\{1+t+t^3, 1-t^2\}$. Como, agora sim, esse último conjunto é LI, ele é uma base de $\operatorname{Im}(T)$. $\diamond$

Exercícios Lista 2 - Álgebra Linear II: Exs. 5–16.

6.3 Teorema do núcleo e da imagem

O fenômeno ocorrido no Exemplo 6.2.7, em que as dimensões do núcleo e da imagem de uma transformação linear estão relacionadas, é geral, como mostra o próximo importante resultado.

Teorema 6.3.1 *Seja U um espaço vetorial de dimensão finita, seja V um espaço vetorial arbitrário e seja $T\colon U \to V$ uma transformação linear. Então,* $\operatorname{Ker}(T)$ *e* $\operatorname{Im}(T)$ *têm dimensão finita e vale*

$$\dim\big(\operatorname{Ker}(T)\big) + \dim\big(\operatorname{Im}(T)\big) = \dim(U). \tag{6.4}$$

Demonstração Como $\operatorname{Ker}(T)$ é um subespaço de U e U tem dimensão finita, segue que $\operatorname{Ker}(T)$ tem dimensão finita.

Vejamos, agora, o que se pode dizer de $\operatorname{Im}(T)$. Seja $\{u_1, \dots, u_k\}$ uma base de $\operatorname{Ker}(T)$ (em particular, estamos supondo que $k = \dim\big(\operatorname{Ker}(T)\big)$). Como esse subconjunto de U é LI, pelo Teorema do completamento (Teorema 4.4.14), ele pode ser completado a uma base de U, isto é, existem $u_{k+1}, \dots, u_n \in U$ tais que $\{u_1, \dots, u_k, u_{k+1}, \dots, u_n\}$ é uma base de U (e, assim, $n = \dim(U)$). Mostremos que $\mathcal{D} = \{T(u_{k+1}), \dots, T(u_n)\}$ é uma base de $\operatorname{Im}(T)$. Com efeito, vimos, na Proposição 6.2.9, que

$$\operatorname{Im}(T) = \big[T(u_1), \dots, T(u_k), T(u_{k+1}), \dots, T(u_n)\big].$$

Mas, pela construção dessa base de U, $u_1, \dots, u_k \in \operatorname{Ker}(T)$, isto é, $T(u_1) = \cdots = T(u_k) = 0_V$. Assim, esses vetores podem ser removidos do conjunto gerador de $\operatorname{Im}(T)$, resultando em $\operatorname{Im}(T) = \big[\mathcal{D}\big]$. Para concluir, resta mostrar que $\mathcal{D}$ é LI. Sejam $\lambda_{k+1}, \dots, \lambda_n \in \mathbb{R}$ tais que

$$\lambda_{k+1} T(u_{k+1}) + \cdots + \lambda_n T(u_n) = 0_V. \tag{6.5}$$

Porque T é uma transformação linear, (6.5) é equivalente a

$$T(\lambda_{k+1} u_{k+1} + \cdots + \lambda_n u_n) = 0_V,$$

o que implica que $\lambda_{k+1} u_{k+1} + \cdots + \lambda_n u_n \in \operatorname{Ker}(T)$. Assim, existem $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ tais que

$$\lambda_{k+1} u_{k+1} + \cdots + \lambda_n u_n = \lambda_1 u_1 + \cdots + \lambda_k u_k, \tag{6.6}$$

uma vez que $\{u_1, \ldots, u_k\}$ é uma base de $\operatorname{Ker}(T)$. Agora, de (6.6), segue que

$$\lambda_1 u_1 + \cdots + \lambda_k u_k + (-\lambda_{k+1})u_{k+1} + \cdots + (-\lambda_n)u_n = 0_U. \tag{6.7}$$

Como $\{u_1, \ldots, u_k, u_{k+1}, \ldots, u_n\}$ é LI (pois é uma base de U), obtém-se que todos os coeficientes na combinação linear do lado esquerdo de (6.7) são nulos. Em particular, $\lambda_{k+1} = \cdots = \lambda_n = 0$, provando que $\mathcal{D}$ é LI.

Assim $\mathcal{D}$ é uma base de $\operatorname{Im}(T)$ e, portanto, $\operatorname{Im}(T)$ tem dimensão finita. Além disso, a dimensão de $\operatorname{Im}(T)$ é igual ao número de vetores de $\mathcal{D}$; portanto, $\dim\big(\operatorname{Im}(T)\big) = n - k = \dim(U) - \dim\big(\operatorname{Ker}(T)\big)$, que pode se reescrever na forma (6.4). □

Esse resultado é conhecido como *Teorema do núcleo e da imagem* ou *Teorema da dimensão*.

O Teorema do núcleo e da imagem tem inúmeras aplicações. Por exemplo, é possível utilizá-lo para argumentar sobre a injetividade de uma transformação linear conhecendo a dimensão de sua imagem:

Exemplo 6.3.2 Mostre que a transformação linear

$$\begin{aligned} T\colon \mathcal{P}_2(\mathbb{R}) &\longrightarrow \mathcal{P}(\mathbb{R}) \\ p(t) &\longmapsto tp'(t) + t^2 p(t) \end{aligned}$$

é injetora.

Solução: Vimos, no Exemplo 6.2.4, que $\dim\big(\operatorname{Im}(T)\big) = 3$. Usando o Teorema do núcleo e da imagem, conclui-se que

$$\dim\big(\operatorname{Ker}(T)) = \dim\big(\mathcal{P}_2(\mathbb{R})\big) - \dim\big(\operatorname{Im}(T)\big) = 3 - 3 = 0.$$

Logo, $\operatorname{Ker}(T) = \{0_{\mathcal{P}_2(\mathbb{R})}\}$ e, portanto, segue da Proposição 6.2.5 que T é injetora. ◇

Uma função é dita sobrejetora quando todos os pontos de seu contradomínio são atingidos por imagens de pontos no domínio (ver Apêndice B). Assim, se $T\colon U \to V$ é uma transformação linear, T será sobrejetora precisamente quando $\operatorname{Im}(T) = V$. Visto isso, uma consequência útil do Teorema do núcleo e da imagem é

Corolário 6.3.3 *Sejam U e V espaços vetoriais de dimensão finita, com* $\dim(U) = \dim(V)$*, e seja $T\colon U \to V$ uma transformação linear. São equivalentes:*

(a) T é bijetora;
(b) T é injetora;
(c) T é sobrejetora.

Demonstração Juntando a hipótese $\dim(U) = \dim(V)$ com (6.4), obtemos

$$\dim\big(\operatorname{Im}(T)\big) = \dim(V) - \dim\big(\operatorname{Ker}(T)\big).$$

Assim, $\dim\big(\operatorname{Im}(T)\big) = \dim(V)$ se, e somente se, $\dim\big(\operatorname{Ker}(T)\big) = 0$. Em vista das Proposições 4.6.1 e 6.2.5, segue que T é sobrejetora se, e somente se, T for injetora. Portanto, as duas condições são equivalentes a T ser bijetora. □

Não é demais ressalvar que a hipótese de U e V serem espaços vetoriais de *mesma dimensão finita* é essencial: no Exemplo 6.3.2, T é injetora, mas obviamente não é sobrejetora; no Exemplo 6.2.3, T não é injetora (pois $\dim\big(\operatorname{Ker}(T)\big) = 1 \neq 0$), mas é sobrejetora (uma vez que $\dim\big(\operatorname{Im}(T)) = \dim(\mathbb{R}^3) - \dim\big(\operatorname{Ker}(T)\big) = 3 - 1 = 2 = \dim(\mathbb{R}^2)$).

Exemplo 6.3.4 Neste ponto, retomaremos o Exemplo 6.2.7 para ver uma segunda demonstração (diferente da que foi apresentada no Apêndice A) de que o posto-linha e o posto-coluna de uma matriz coincidem.

Seja $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$. Podemos pensar nas linhas de A como vetores de $\mathbb{R}^n$: para cada $i = 1, \ldots, m$, seja

$$u_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \in \mathbb{R}^n,$$

e nas colunas de A como elementos de $\mathbb{R}^m$: para cada $j = 1, \ldots, n$, seja

$$v_j = (a_{1j}, a_{2j}, \dots, a_{mj}) \in \mathbb{R}^m.$$

Denotamos por $L(A) = [u_1, u_2, \dots, u_m]$ o subespaço de $\mathbb{R}^n$ gerado pelas m linhas de A e por $C(A) = [v_1, v_2, \dots, v_n]$ e por $C(A)$ o subespaço de $\mathbb{R}^m$ gerado pelas n colunas de A. Nosso objetivo será mostrar que $\dim\big(L(A)\big) = \dim\big(R(A)\big)$.

Conforme vimos no Exemplo 6.2.7, $C(A) = \operatorname{Im}(T_A)$, em que $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$ é a transformação linear definida por $T_A(v) = Av$, para todo $v \in \mathbb{R}^n$. Pelo Teorema do núcleo e da imagem, segue que $\dim\big(C(A)\big) = n - \dim\big(\operatorname{Ker}(T_A)\big)$. Agora, também vimos naquele exemplo que se t denota o número de pivôs de uma matriz escalonada R obtida a partir de A por operações elementares sobre linhas, então $\dim\big(\operatorname{Ker}(T_A)\big) = n-t$. Segue que $\dim\big(C(A)\big) = n-(n-t) = t$. Mas t é a dimensão do espaço gerado pelas linhas de R (porque R é escalonada), que coincide com $L(A)$ (pois operações elementares sobre linhas não altera o espaço gerado pelas linhas). Portanto, $\dim\big(L(A)\big) = \dim\big(R(A)\big)$. Esse número é comumente chamado *posto* da matriz A.

Transformações bijetoras

Toda função bijetora é inversível. Em particular, se U e V são espaços vetoriais e $T\colon U \to V$ é uma transformação linear bijetora, então existe uma função $T^{-1}\colon V \to U$ que é a inversa de T. Essa função é tal que para todos $v \in V$ e $u \in U$,

$$T^{-1}(v) = u \quad \text{se, e somente se,} \quad T(u) = v.$$

(Consulte o Apêndice B para detalhes sobre invertibilidade de funções.) Não é imediato da definição de função inversa que T^{-1} seja, também, linear. Esse é o conteúdo do próximo resultado.

Proposição 6.3.5 *Sejam U e V espaços vetoriais e seja $T\colon U \to V$ uma transformação linear. Se T é bijetora, então a função inversa $T^{-1}\colon V \to U$ é uma transformação linear.*

Demonstração A tarefa é mostrar que T^{-1} satisfaz as condições (TL-1) e (TL-2). Sejam $v, v' \in V$. Então, $v = T\big(T^{-1}(v)\big)$ e $v' = T\big(T^{-1}(v')\big)$, e, portanto,

$$v + v' = T\big(T^{-1}(v)\big) + T\big(T^{-1}(v')\big) = T\big(T^{-1}(v) + T^{-1}(v')\big),$$

pois T é linear. Segue que $T^{-1}(v+v') = T^{-1}(v) + T^{-1}(v')$. Se $\lambda \in \mathbb{R}$, então

$$\lambda v = \lambda T\big(T^{-1}(v)\big) = T\big(\lambda T^{-1}(v)\big),$$

uma vez que T é linear. Daí, obtemos $T^{-1}\big(\lambda v\big) = \lambda T^{-1}(v)$. Logo, T^{-1} é, de fato, uma transformação linear. □

Exemplo 6.3.6 Mostre que a transformação linear

$$\begin{aligned} T\colon \mathbb{R}^3 &\longrightarrow \mathcal{P}_2(\mathbb{R}) \\ (a,b,c) &\longmapsto (a+c) + (a-c)t + bt^2 \end{aligned}$$

é bijetora e determine sua inversa.

Solução: Porque $\dim(\mathbb{R}^3) = 3 = \dim\big(\mathcal{P}_2(\mathbb{R})\big)$, segue do Corolário 6.3.3, que para mostrar que T é bijetora, é suficiente mostrar, por exemplo, que T é injetora. Vejamos, dado $(a,b,c) \in \mathbb{R}^3$, temos $(a,b,c) \in \operatorname{Ker}(T)$ se, e somente se, $(a+c)+(a-c)t+bt^2 = T(a,b,c) = 0$. Isso ocorre apenas quando $a = b = c = 0$. Assim, $\operatorname{Ker}(T) = \{(0,0,0)\}$, o que, pela Proposição 6.2.5, garante que T é injetora; e, como vimos, bijetora.

Agora, procuremos uma expressão para inversa $T^{-1}\colon \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}^3$ de T. Dado $p = \alpha_0 + \alpha_1 t + \alpha_2 t^2 \in \mathcal{P}_2(\mathbb{R})$, temos que $T^{-1}(p) = (a,b,c)$ se, e somente se, $T(a,b,c) = p$, o que, por sua vez, é equivalente a

$$(a+c) + (a-c)t + bt^2 = \alpha_0 + \alpha_1 t + \alpha_2 t^2.$$

Em $\mathcal{P}_2(\mathbb{R})$, a igualdade acima ocorre se, e só se,

$$\begin{cases} a + c = \alpha_0 \\ a - c = \alpha_1 \\ b = \alpha_2 \end{cases}$$

Resolvendo o sistema para a, b, c, em termos de $\alpha_0, \alpha_1, \alpha_2$, obtemos

$$a = \frac{\alpha_0 + \alpha_1}{2}, \qquad b = \alpha_2, \qquad c = \frac{\alpha_0 - \alpha_1}{2},$$

de modo que

$$T^{-1}(\alpha_0 + \alpha_1 t + \alpha_2 t^2) = \left(\frac{\alpha_0 + \alpha_1}{2}, \alpha_2, \frac{\alpha_0 - \alpha_1}{2}\right). \quad \diamond$$

Há outras consequências do Teorema do núcleo e da imagem:

Corolário 6.3.7 *Sejam U e V espaços vetoriais de dimensão finita e seja $T\colon U \to V$ uma transformação linear. Valem:*

(i) Se T é injetora, então $\dim(U) \leq \dim(V)$.
(ii) Se T é sobrejetora, então $\dim(U) \geq \dim(V)$.
(iii) Se T é bijetora, então $\dim(U) = \dim(V)$

Demonstração Para mostrar (i), suponha que T é injetora. Então, $\dim\big(\mathrm{Ker}(T)\big) = 0$; logo, segue do Teorema do núcleo e da imagem que

$$\dim(U) = \dim\big(\mathrm{Ker}(T)\big) + \dim\big((\mathrm{Im}(T)\big) = \dim\big((\mathrm{Im}(T)\big) \leq \dim(V),$$

pois $\mathrm{Im}(T)$ é um subespaço de V. Para ver que (ii) vale, suponha que T é sobrejetora, ou seja, que $\mathrm{Im}(T) = V$. Pelo Teorema do núcleo e da imagem, segue que

$$\dim(U) = \dim\big(\mathrm{Ker}(T)\big) + \dim\big((\mathrm{Im}(T)\big) = \dim\big(\mathrm{Ker}(T)\big) + \dim(V) \geq \dim(V).$$

Finalmente, (iii) segue de (i) e (ii). $\square$

Esse corolário pode ser lido na contrapositiva: sejam U e V são espaços vetoriais de dimensão finita. Então,

- se $\dim(U) > \dim(V)$, não existe transformação linear $U \to V$ injetora;
- se $\dim(U) < \dim(V)$, não existe transformação linear $U \to V$ sobrejetora.

Exemplo 6.3.8 (Prova 1, Álgebra Linear II, 2016) Sejam V e W espaços vetoriais de dimensão finita e $T\colon V \to W$ uma transformação linear. Considere as seguintes afirmações:

I. Se T é injetora, então $\dim(V) \leq \dim(W)$.
II. Se $\dim(V) \leq \dim(W)$, então T é injetora.
III. Se T é sobrejetora, então $\dim(V) \geq \dim(W)$.

Está correto o que se afirma em

(A) I e II, apenas. (B) II, apenas. (C) II e III, apenas. (D) I e III, apenas. (E) I, II e III.

Solução: Vimos, no Corolário 6.3.7, que I e III são verdadeiras. A afirmação II é a recíproca da afirmação I. Ela é falsa. Há muitas transformações lineares $T\colon V \to W$ não injetoras quando $\dim(V) \leq \dim(W)$. Por exemplo, se $V = \mathbb{R}^2$ e $W = \mathbb{R}^3$, a transformação nula (isto é, a transformação linear $T\colon \mathbb{R}^2 \to \mathbb{R}^3$, dada por $T(x, y) = (0, 0, 0)$, para todo $(x, y) \in \mathbb{R}^2$) não é injetora, apesar de $\dim(\mathbb{R}^2) = 2 < 3 = \dim(\mathbb{R}^3)$. *Resposta:* (D) $\diamond$

Exercícios Lista 2 - Álgebra Linear II: Exs. 17–27.

6.4 Operações com transformações lineares

Sejam U, V, W espaços vetoriais e sejam $T: U \to V$ e $S: V \to W$ transformações lineares. Então, está definida a função composta (ver Apêndice B) $S \circ T: U \to W$, que satisfaz

$$(S \circ T)(u) = S\big(T(u)\big), \quad \text{para todo } u \in U.$$

Se $T_1: U \to V$ é outra transformação linear também está definida a função $T + T_1: U \to V$, por

$$(T + T_1)(u) = T(u) + T_1(u), \quad \text{para todo } u \in U.$$

E, finalmente, se $\alpha \in \mathbb{R}$ é um escalar fixado, podemos considerar a função $\alpha T: U \to V$, definida por

$$(\alpha T)(u) = \alpha T(u), \quad \text{para todo } u \in U.$$

Todas essas construções resultam em transformações lineares:

Proposição 6.4.1 *Sejam U, V, W espaços vetoriais, sejam $T, T_1: U \to V$, $S: V \to W$ transformações lineares e seja $\alpha \in \mathbb{R}$. Então $S \circ T$, $T + T_1$ e αT são transformações lineares.*

Demonstração Sejam $u, u' \in U$ e $\lambda \in U$. Então,

$$\begin{aligned}
(S \circ T)(u + u') &= S\big(T(u + u')\big) \\
&= S\big(T(u) + T(u')\big) && \text{pois } T \text{ é linear} \\
&= S\big(T(u)\big) + S\big(T(u')\big) && \text{pois } S \text{ é linear} \\
&= (S \circ T)(u) + (S \circ T)(u').
\end{aligned}$$

Similarmente,

$$(S \circ T)(\lambda u) = S\big(T(\lambda u)\big) = S\big(\lambda T(u)\big) = \lambda\big(S(T(u)\big) = \lambda(S \circ T)(u).$$

Fica a cargo do leitor mostrar que $T + T_1$ e αT satisfazem as condições para serem transformações lineares. □

Dado um espaço vetorial V, se $T: V \to V$ é um operador linear de V, podemos considerar o operador $T \circ T: V \to V$. Esse operador linear será denotado por T^2. De maneira análoga, se n é um inteiro maior do que 2, denotaremos

$$T^n = T \circ T \circ \cdots \circ T,$$

em que há n ocorrências de T.

Essas construções serão especialmente relevantes no estudo de autovalores de um operador linear, assunto a ser tratato no Capítulo 7.

Exercícios Lista 2 - Álgebra Linear II: Exs. 28–30.

6.5 Matriz de uma transformação linear

Neste seção, será introduzida a mais importante ferramenta de análise de uma transformação linear entre espaços vetoriais de dimensão finita, a matriz da transformação linear.

Sejam U e V espaços vetoriais de dimensão finita, digamos $\dim(U) = n$ e $\dim(V) = m$, e seja $T: U \to V$ uma transformação linear. Fixe uma base ordenada $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ de U e uma base ordenada $C = \{v_1, v_2, \dots, v_m\}$ de V.

Para cada $j = 1, \dots, n$, o vetor $T(u_j)$ pertence ao espaço vetorial V; logo, se escreve como combinação linear dos elementos da base C, ou seja, existem $a_{1j}, a_{2j}, \dots, a_{mj} \in \mathbb{R}$ tais que

$$T(u_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{mj}v_m. \tag{6.8}$$

Considere, agora, a matriz $[T]_{\mathcal{B}C}$, de tamanho $m \times n$, cuja j-ésima coluna é formada pelas coordenadas $(a_{1j}, a_{2j}, \ldots, a_{mj})$ de $T(u_j)$ em relação à base C:

$$[T]_{\mathcal{B}C} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1j} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2j} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mj} & \ldots & a_{mn} \end{bmatrix}$$

Essa matriz é denominada *matriz da transformação linear T em relação às bases $\mathcal{B}$ e C*.

Veremos que esse elemento de $M_{m\times n}(\mathbb{R})$ codifica a informação completa sobre a transformação linear T. Mas, primeiramente, a fim de fixar a definição, vejamos exemplos de construção da matriz de uma transformaçã linear.

Exemplo 6.5.1 Considere a transformação linear

$$\begin{aligned} T\colon \mathbb{R}^3 &\longrightarrow \mathcal{P}_1(\mathbb{R}) \\ (a, b, c) &\longmapsto (2a + c) + (b - 2c)t \end{aligned}$$

e as bases canônicas $C = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ e $\mathcal{D} = \{1, t\}$ de $\mathbb{R}^3$ e de $\mathcal{P}_1(\mathbb{R})$, respectivamente. Encontre $[T]_{\mathcal{B}C}$.

Solução: Comecemos por deteminar o tamanho da matriz $[T]_{\mathcal{B}C}$. Conforme vimos, o número de colunas de $[T]_{\mathcal{B}C}$ é igual à dimensão do domínio de T, no caso, $3 = \dim(\mathbb{R}^3)$; o número de linhas de $[T]_{\mathcal{B}C}$ é igual é igual à dimensão do contradomínio de T, no caso, $2 = \dim\big(\mathcal{P}_1(\mathbb{R})\big)$. Assim, já sabemos que $[T]_{\mathcal{B}C} \in M_{2\times 3}(\mathbb{R})$. Para determinar suas entradas, precisamos encontrar as coordenadas, em relação a C, de cada um dos três vetores de $\mathcal{B}$. Vejamos,

$$T(1, 0, 0) = 2 = 2 \cdot 1 + 0 \cdot t \tag{6.9}$$
$$T(0, 1, 0) = t = 0 \cdot 1 + 1 \cdot t \tag{6.10}$$
$$T(0, 0, 1) = 1 - 2t = 1 \cdot 1 + (-2) \cdot t \tag{6.11}$$

A primeira coluna de $[T]_{\mathcal{B}C}$ contém as coordenadas, em relação a C, da imagem por T do primeiro vetor da base ordenada $\mathcal{B}$ (que é $(1, 0, 0)$), assim, conforme (6.9), as duas entradas na primeira coluna de $[T]_{\mathcal{B}C}$ são 2 e 0:

$$[T]_{\mathcal{B}C} = \begin{bmatrix} 2 & & \\ 0 & & \end{bmatrix}.$$

Falta preencher as outras duas colunas de $[T]_{\mathcal{B}C}$, temporariamente em branco. Para a segunda coluna, precisamos das coordenadas, em relação a C, de $T(0, 1, 0)$, uma vez que $(0, 1, 0)$ é o segundo vetor de $\mathcal{B}$; logo, de (6.10), obtemos

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

Finalmente, sendo $(0, 0, 1)$ o terceiro vetor da base ordenada $\mathcal{B}$, obtemos de (6.11) que as entradas na terceira coluna de $[T]_{\mathcal{B}C}$ são as coordenadas de $T(0, 0, 1)$ em relação à base C:

$$[T]_{\mathcal{B}C} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix}. \quad \diamond$$

O tamanho da matriz de uma transformação linear só depende das *dimensões* de seu domínio e contradomínio. Porém, suas entradas são sensíveis às *bases ordenadas* específicas que foram escolhidas.

Exemplo 6.5.2 Para a mesma transformação linear do exemplo anterior, determine $[T]_{\mathcal{B}_1C}$, $[T]_{\mathcal{B}C_1}$ e $[T]_{\mathcal{B}_1C_1}$, em que $\mathcal{B}$ e C são as bases canônicas de $\mathbb{R}^3$ e $\mathcal{P}_1(\mathbb{R})$, respectivamente, a base ordenada $\mathcal{B}_1$ de $\mathbb{R}^3$ é dada por

$$\mathcal{B}_1 = \{ (1,1,1), (1,1,0), (1,0,0) \},$$

e a base ordenada C_1 de $\mathcal{P}_1(\mathbb{R})$ é dada por

$$C_1 = \{ 1+t, 2+t \}.$$

Solução: Note que $\mathcal{B}_1$ e C_1 são, de fato, bases de seus respectivos espaços. As três matrizes têm tamanho 2×3. Comecemos por $[T]_{\mathcal{B}_1 C}$. Precisamos das coordenadas, em relação à base C, das imagens, por T, dos vetores de $\mathcal{B}_1$. Como

$$\begin{aligned} T(1,1,1) &= 3 - t = 3 \cdot 1 + (-1) \cdot t \\ T(1,1,0) &= 2 + t = 2 \cdot 1 + 1 \cdot t \\ T(1,0,0) &= 2 = 2 \cdot 1 + 0 \cdot t, \end{aligned}$$

segue que

$$[T]_{\mathcal{B}_1 C} = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \end{bmatrix}.$$

Para $[T]_{\mathcal{B} C_1}$, é preciso decompor as imagens, por T, de cada um dos elementos de $\mathcal{B}$ como combinação linear dos elementos de C_1. Efetuando os cálculos necessários[2], obtemos

$$\begin{aligned} T(1,0,0) &= 2 = (-2)(1+t) + 2(2+t) \\ T(0,1,0) &= t = 2(1+t) + (-1)(2+t) \\ T(0,0,1) &= 1 - 2t = (-5)(1+t) + 3(2+t). \end{aligned}$$

Portanto,

$$[T]_{\mathcal{B}_1 C} = \begin{bmatrix} -2 & 2 & -5 \\ 2 & -1 & 3 \end{bmatrix}.$$

Finalmente, de

$$\begin{aligned} T(1,1,1) &= 3 - t = (-5)(1+t) + 4(2+t) \\ T(1,1,0) &= 2 + t = 0(1+t) + 1(2+t) \\ T(1,0,0) &= 2 = (-2)(1+t) + 2(2+t), \end{aligned}$$

segue que

$$[T]_{\mathcal{B}_1 C_1} = \begin{bmatrix} -5 & 0 & -2 \\ 4 & 1 & 2 \end{bmatrix}. \quad \diamond$$

Nos dois exemplos anteriores, as quatro matrizes $[T]_{\mathcal{B}C}$, $[T]_{\mathcal{B}_1 C}$, $[T]_{\mathcal{B}C_1}$ e $[T]_{\mathcal{B}_1 C_1}$ são diferentes, mas são, todas, matrizes da *mesma* transformação linear. Qual é a relação que existe entre elas? Essa pergunta será abordada na Seção 6.7.

Quando se trata de um operador linear, podemos considerar a matriz do operador em relação à mesma base na "saída" e "chegada".

Exemplo 6.5.3 Seja V um espaço vetorial de dimensão finita igual a n, seja I_V o operador identidade de V e seja $\mathcal{B}$ uma base de V. Descreva $[\mathrm{I}_V]_{\mathcal{B}\mathcal{B}}$.

Solução: O operador identidade satisfaz $\mathrm{I}_V(v) = v$, para todo $v \in V$, em particular, fixa também os elementos de $\mathcal{B}$. Mais detalhadamente, se $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$, então, $\mathrm{I}_V(v_j) = v_j$, para todo $i = 1, \dots, n$. Isso que dizer que a j-ésima coluna de $[\mathrm{I}_V]_{\mathcal{B}\mathcal{B}}$ contém zeros em todas as suas entradas, exceto na posição j, onde ocorre o número 1. Essa matriz é a matriz identidade: $[\mathrm{I}_V]_{\mathcal{B}\mathcal{B}} = I_n$. $\diamond$

[2] Por exemplo, para determinar as coordenadas de $T(0,0,1) = 1 - 2t$ em relação à base C_1, escrevemos $1 - 2t = \lambda(1+t) + \mu(2+t) = (\lambda + 2\mu) + (\lambda + \mu)t$, donde segue que $\begin{cases} \lambda + 2\mu = 1 \\ \lambda + \mu = -2 \end{cases}$. A solução desse sistema é $\lambda = -5$ e $\mu = 3$.

Tomando bases diferentes, a matriz do operador identidade deixa de ser a matriz identidade, como mostra o próximo exemplo.

Exemplo 6.5.4 Calcule $[\mathrm{I}]_{\mathcal{B}\mathcal{C}}$, em que I denota o operador identidade de $\mathbb{R}^3$ e as bases ordenadas $\mathcal{B}$ e $\mathcal{C}$ de $\mathbb{R}^3$ são dadas por

$$\mathcal{B} = \{ (1,1,1), (1,1,0), (1,0,0) \} \qquad \text{e} \qquad \mathcal{C} = \{ (-1,0,-1), (-1,1,0), (0,1,0) \}.$$

Solução: Sabemos, do exemplo anterior, que $[\mathrm{I}]_{\mathcal{B}\mathcal{B}} = [\mathrm{I}]_{\mathcal{C}\mathcal{C}} = I_3$. Mas,

$$[\mathrm{I}]_{\mathcal{B}\mathcal{C}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & -1 \\ 1 & 2 & 1 \end{bmatrix},$$

uma vez que, como o leitor pode, ele mesmo, calcular,

$$\begin{aligned} \mathrm{I}(1,1,1) &= (1,1,1) = (-1)(-1,0,-1) + 0(-1,1,0) + 1(0,1,0) \\ \mathrm{I}(1,1,0) &= (1,1,0) = 0(-1,0,-1) + (-1)(-1,1,0) + 2(0,1,0) \\ \mathrm{I}(1,0,0) &= (1,0,0) = 0(-1,0,-1) + (-1)(-1,1,0) + 1(0,1,0) \quad \diamond \end{aligned}$$

Exemplo 6.5.5 Voltemos à família de transformações lineares estudadas no Exemplo 6.2.7. Dada uma matriz $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$, seja $T_A \colon \mathbb{R}^n \to \mathbb{R}^m$ a transformação linear definida por $T_A(v) = Av$, para todo $v \in \mathbb{R}^n$. Sejam $\mathcal{B}$ e $\mathcal{C}$ as bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente. Então, $[T_A]_{\mathcal{B}\mathcal{C}} = A$. Com efeito, se $\mathcal{B} = \{e_1, \dots, e_n\}$ e $\mathcal{D} = \{f_1, \dots, f_m\}$, então

$$T(e_j) = (a_{1j}, a_{2j}, \dots, a_{mj}) = a_{1j}f_1 + a_{2j}f_2 + \cdots + a_{mj}f_m,$$

para todo $j = 1, \dots, n$.

Como usar a matriz de uma transformação linear

Como já dito, o principal uso da matriz de uma transformação linear é no sentido de codificá-la (a transformação) em um objeto compacto, uma matriz. É disso que trataremos agora.

Antes de começar, introduziremos um notação. Vimos, na Seção 4.5, que, dado um espaço vetorial U de dimensão finita igual a n e uma base ordenada $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ de U, associamos a cada elemento $u \in U$ um elemento de $\mathbb{R}^n$, suas coordendas: se $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ são tais que $u = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n$, escrevemos

$$u = (\lambda_1, \lambda_2, \dots, \lambda_n)_{\mathcal{B}}.$$

Também nos será útil, por meio da identificação usual dos elementos de $\mathbb{R}^n$ com matrizes de $M_{n\times 1}(\mathbb{R})$, utilizar a seguinte notação para nos referirmos às coordenadas de u em relação à base $\mathcal{B}$:

$$[u]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

É comum designar $[u]_{\mathcal{B}}$ de *vetor de coordenadas* de u em relação à base ordenada $\mathcal{B}$.

Aqui, é importante notar que, como demonstra a Proposição 4.5.1, as operações de soma e multiplicação por escalar em U, por meio da associação de cada elemento de U com seu vetor de coordenadas, refletem-se precisamente nas operações usuais que conferem a $M_{n\times 1}(\mathbb{R})$ uma estrutura de espaço vetorial (conforme o Exemplo 4.1.4). Em resumo, se $u, v \in U$ e $\lambda \in \mathbb{R}$, então

$$[u+v]_{\mathcal{B}} = [u]_{\mathcal{B}} + [v]_{\mathcal{B}} \qquad \text{e} \qquad [\lambda u]_{\mathcal{B}} = \lambda [u]_{\mathcal{B}}.$$

Voltemos, agora, à matriz de uma transformação linear. Sejam U e V espaços vetoriais de dimensão finita, digamos $\dim(U) = n$ e $\dim(V) = m$, e seja $T : U \to V$ uma transformação linear. Fixe uma base ordenada $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ de U e uma base ordenada $C = \{v_1, v_2, \dots, v_m\}$ de V e considere a matriz $[T]_{\mathcal{B}C} = (a_{ij}) \in M_{m\times n}(\mathbb{R})$. Tomemos, agora, um elemento $u \in U$ e vejamos como utilizar a matriz $[T]_{\mathcal{B}C}$ para descrever $T(u)$. Suponha que $u = (\lambda_1, \lambda_2, \dots, \lambda_n)_{\mathcal{B}}$. Então,

$$\begin{aligned}
T(u) &= T(\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n) \\
&= \lambda_1 T(u_1) + \lambda_2 T(u_2) + \cdots + \lambda_n T(u_n) \\
&= \lambda_1(a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_1) + \lambda_2(a_{12}v_1 + a_{22}v_2 + \cdots + a_{m2}v_1) + \dots \\
&\qquad + \lambda_n(a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{m1}v_n) \qquad \text{por (6.8)} \\
&= (a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1n}\lambda_n)v_1 + (a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2n}\lambda_n)v_2 + \dots \\
&\quad (a_{m1}\lambda_1 + a_{m2}\lambda_2 + \cdots + a_{mn}\lambda_n)v_m.
\end{aligned}$$

A expressão que obtivemos acima permite descrever as coordenadas de $T(u)$ em relação a C em termos das coordenadas de u e das entradas de $[T]_{\mathcal{B}C}$: mostramos que a i-ésima coordenada de $T(u)$ em relação à base C coincide com o produto da i-ésima linha da matriz $[T]_{\mathcal{B}C}$ pelo vetor de coordenadas de u em relação à base $\mathcal{B}$, ou em termos de um produto entre matrizes,

$$[T(u)]_C = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_n \end{bmatrix} = [T]_{\mathcal{B}C}[u]_{\mathcal{B}}.$$

Essa igualdade é tão importante, que vamos enunciá-la na forma de um teorema (que acabamos de demonstrar).

Teorema 6.5.6 *Sejam U e V espaços vetoriais de dimensão finita, seja $T : U \to V$ uma transformação linear e sejam $\mathcal{B}$ e C bases ordenadas de U e de V, respectivamente. Então,*

$$[T(u)]_C = [T]_{\mathcal{B}C}[u]_{\mathcal{B}}, \tag{6.12}$$

para todo $u \in U$.

Vejamos um exemplo em que usamos (6.12) para encontrar a imagem de um elemento por uma transformação linear.

Exemplo 6.5.7 Retomando os Exemplos 6.5.1 e 6.1.2, vimos que para a transformação linear

$$\begin{aligned} T : \mathbb{R}^3 &\longrightarrow \mathcal{P}_1(\mathbb{R}) \\ (a, b, c) &\longmapsto (2a + c) + (b - 2c)t \end{aligned}$$

temos que

$$[T]_{\mathcal{B}_1 C_1} = \begin{bmatrix} -5 & 0 & -2 \\ 4 & 1 & 2 \end{bmatrix}, \tag{6.13}$$

em que as bases envolvidas são

$$\mathcal{B}_1 = \{\, (1, 1, 1), (1, 1, 0), (1, 0, 0) \,\} \qquad \text{e} \qquad C_1 = \{\, 1 + t, 2 + t \,\}.$$

Considere o vetor $u = (2, 1, -1)$ de $\mathbb{R}^3$. Encontre o vetor $T(u)$ de $\mathcal{P}_1(\mathbb{R})$ usando a matriz $[T]_{\mathcal{B}_1 C_1}$.

Solução: Sabemos, do Teorema 6.5.6, que

$$[T(u)]_{C_1} = [T]_{\mathcal{B}_1 C_1}[u]_{\mathcal{B}_1}. \tag{6.14}$$

A matriz $[T]_{\mathcal{B}_1 C_1}$ é conhecida. Vejamos o segundo fator nesse produto, o vetor de coordenadas de u em relação à base $\mathcal{B}_1$: dados $\alpha, \beta, \gamma \in \mathbb{R}$, temos que $u = (\alpha, \beta, \gamma)_{\mathcal{B}_1}$ se, e somente se,

$$(2, 1, -1) = \alpha(1, 1, 1) + \beta(1, 1, 0) + \gamma(1, 0, 0).$$

Resolvendo o sistema linear que decorre dessa igualdade, obtemos

$$\alpha = -1, \qquad \beta = 2, \qquad \gamma = 1.$$

Assim,

$$[u]_{\mathcal{B}_1} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}. \tag{6.15}$$

Logo, substituindo (6.13) e (6.15) em (6.14), obtemos

$$[T(u)]_{C_1} = \begin{bmatrix} -5 & 0 & -2 \\ 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}.$$

Isso quer dizer que $T(u) = 3(1 + t) + 0(2 + t) = 3 + 3t$, já que os vetores de $\mathcal{P}_1(\mathbb{R})$ que compõem a base C_1 são $1 + t$ e $2 + t$, nessa ordem. $\diamond$

O exemplo anterior é um tanto artificial, pois já conhecíamos a expressão de $T(a, b, c)$ para qualquer $(a, b, c) \in \mathbb{R}^3$. O fato é que podemos fornecer apenas a matriz de uma transformação linear a fim de descrevê-la completamente.

Exemplo 6.5.8 Considere a transformação linear $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ tal que

$$[T]_{\mathcal{D}\mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ -3 & 1 \end{bmatrix},$$

em que as bases ordenadas de $\mathcal{D}$ de $\mathbb{R}^2$ e $\mathcal{E}$ de $\mathbb{R}^3$, respectivamente, são dadas por

$$\mathcal{D} = \big\{ (0, 1), (1, 1) \big\}, \qquad \mathcal{E} = \big\{ (0, -1, 0), (2, 1, 0), (0, 1, 2) \big\}.$$

Encontre a expressão de $T(x, y)$ e determine $\mathrm{Im}(T)$.

Solução: Antes de começar, cabem dois comentários. Note que $\mathcal{D}$ e $\mathcal{E}$, de fato, são LI, e, portanto, bases de seus respectivos espaços. Observe, também, que $[T]_{\mathcal{D}\mathcal{E}} \in M_{3\times 2}(\mathbb{R})$, o que está de acordo com as dimensões de $\mathbb{R}^2$ e $\mathbb{R}^3$. Sabemos, pelo Teorema 6.5.6, que, para todo $(x, y) \in \mathbb{R}^2$, vale:

$$[T(x, y)]_{\mathcal{E}} = [T]_{\mathcal{D}\mathcal{E}}[(x, y)]_{\mathcal{D}}.$$

Fixado um vetor $(x, y) \in \mathbb{R}^2$, suponha que $\alpha, \beta \in \mathbb{R}$ sejam tais que

$$[(x, y)]_{\mathcal{D}} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

Isso quer dizer que $(x, y) = \alpha(0, 1) + \beta(1, 1) = (\beta, \alpha + \beta)$. Assim,

$$\begin{cases} \beta = x \\ \alpha + \beta = y \end{cases},$$

donde segue que $\alpha = -x + y$ e $\beta = x$. Portanto,

$$[T(x,y)]_{\mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} -x+y \\ x \end{bmatrix} = \begin{bmatrix} -x+y \\ -3x+2y \\ 4x-3y \end{bmatrix}.$$

Logo, $T(x,y) = (-x+y)(0,-1,0) + (-3x+2y)(2,1,0) + (4x-3y)(0,1,2) = (-6x+4y, 2x-2y, 8x-6y)$.

Agora, para determinar a imagem de T, sabemos, da Proposição 6.2.9, que, como $\mathcal{D}$ gera $\mathbb{R}^2$, então, $\operatorname{Im}(T) = \big[T(0,1), T(1,1)\big] = \big[(4,-2,-6), (-2,0,2)\big]$. Poderíamos ter encontrado $\operatorname{Im}(T)$ sem ter tido que encontrar, antes, uma expressão da imagem por T de elemento genérico do domínio. O argumento seria o seguinte. Como

$$[(0,1)]_{\mathcal{D}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{e} \qquad [(1,1)]_{\mathcal{D}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

segue que

$$[T(0,1)]_{\mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} \qquad \text{e} \qquad [T(1,1)]_{\mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

Logo, $\operatorname{Im}(T) = \big[T(0,1), T(1,1)\big] = \big[(1,2,-3)_{\mathcal{E}}, (0,-1,1)_{\mathcal{E}}\big]$, que resulta no mesmo conjunto gerador para a imagem que encontramos acima. Observe que, para essa segunda solução, a informação a respeito de quais são os elementos de $\mathcal{B}$ é dispensável.

Observação O argumento que fizemos no final do exemplo acima generaliza-se. Sejam U e V espaços vetoriais de dimensão finita, digamos $\dim(U) = n$ e $\dim(V) = m$, sejam $\mathcal{B}$ e $\mathcal{C}$ bases ordenadas de U e de V, respectivamente, e seja $A \in M_{m\times n}(\mathbb{R})$. Considere a transformação linear $T\colon U \to V$ tal que $[T]_{\mathcal{B}\mathcal{C}} = A$. Então, $\operatorname{Im}(T)$ é gerada pelas colunas de A lidas como vetores de coordenadas em relação à base $\mathcal{C}$.

Exemplo 6.5.9 Considere a transformação linear $T\colon M_2(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ que satisfaz

$$[T]_{\mathcal{B}\mathcal{C}} = \begin{bmatrix} 3 & 6 & 1 & -2 \\ 0 & 0 & -2 & 1 \\ 1 & 2 & -1 & 0 \end{bmatrix},$$

em que a base ordenada $\mathcal{B}$ de $M_2(\mathbb{R})$ é dada por

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right\}$$

e $\mathcal{C}$ denota uma base ordenada fixa de $\mathcal{P}_2(\mathbb{R})$. Descreva o núcleo de T.

Solução: Os dados são compatíveis: o tamanho de $[T]_{\mathcal{B}\mathcal{C}}$ está de acordo com as dimensões de $M_2(\mathbb{R})$ e de $\mathcal{P}_2(\mathbb{R})$, e $\mathcal{B}$, de fato, é uma base de $M_2(\mathbb{R})$. Mas parece haver falta de informação no enunciado, uma vez que não é dito quem são os elementos de $\mathcal{C}$. Veremos que, apesar de não termos informação completa sobre T, ainda assim, com a informação de que dispomos, é possível descrever completamente seu núcleo.

Dado $u \in M_2(\mathbb{R})$, sabemos que $u \in \operatorname{Ker}(T)$ se, e somente se, $T(u) = 0_{\mathcal{P}_2(\mathbb{R})}$. Dois vetores são iguais precisamente quando têm as mesmas coordenadas em relação a uma base; além disso, o vetor nulo é o único vetor cujas coordenadas são todas nulas. Assim, usando (6.12), podemos escrever

$$u \in \operatorname{Ker}(T) \quad \Longleftrightarrow \quad \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = [0_{\mathcal{P}_2(\mathbb{R})}]_{\mathcal{C}} = [T(u)]_{\mathcal{C}} = [T]_{\mathcal{B}\mathcal{C}}[u]_{\mathcal{B}}. \tag{6.16}$$

Suponha que $u = (\alpha, \beta, \gamma, \delta)_{\mathcal{B}}$. Então, a condição que define os vetores de $\operatorname{Ker}(T)$, em (6.16), pode ser reescrita em termos de as coordenadas de u em relação à base $\mathcal{B}$ serem soluções de um sistema linear homogêneo:

$$u = (\alpha, \beta, \gamma, \delta)_{\mathcal{B}} \in \operatorname{Ker}(T) \iff \begin{bmatrix} 3 & 6 & 1 & -2 \\ 0 & 0 & -2 & 1 \\ 1 & 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Para resolver esse sistema, recorremos ao processo de escalonamento:

$$\begin{bmatrix} 3 & 6 & 1 & -2 \\ 0 & 0 & -2 & 1 \\ 1 & 2 & -1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & -2 & 1 \\ 3 & 6 & 1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 4 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

que é escalonada. Assim, β e δ são variáveis livres, e, escrevendo as variáveis pivô em termos das livres, obtemos que os vetores de $\operatorname{Ker}(T)$ são exatamente os vetores da forma

$$\left(-2\beta + \frac{\delta}{2}, \beta, \frac{\delta}{2}, \delta\right)_{\mathcal{B}} = \beta(-2, 1, 0, 0)_{\mathcal{B}} + \delta\left(\frac{1}{2}, 0, \frac{1}{2}, 1\right)_{\mathcal{B}},$$

com $\beta, \delta \in \mathbb{R}$. Assim,

$$\operatorname{Ker}(T) = \left[(-2, 1, 0, 0)_{\mathcal{B}}, \left(\frac{1}{2}, 0, \frac{1}{2}, 1\right)_{\mathcal{B}}\right] = \left[(-2, 1, 0, 0)_{\mathcal{B}},\ (1, 0, 1, 2)_{\mathcal{B}}\right].$$

Como

$$(-2, 1, 0, 0)_{\mathcal{B}} = (-2)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 1\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 0\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + 0\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

e

$$(1, 0, 1, 2)_{\mathcal{B}} = 1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 0\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 1\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + 2\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 2 & 0 \end{bmatrix},$$

concluímos que

$$\operatorname{Ker}(T) = \left[\begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}, \begin{bmatrix} 2 & -2 \\ 2 & 0 \end{bmatrix}\right]. \quad \diamond$$

Observação Aqui, também, o método pode ser generalizado. Sejam U e V espaços vetoriais de dimensão finita, digamos $\dim(U) = n$ e $\dim(V) = m$, sejam $\mathcal{B}$ e $\mathcal{C}$ bases ordenadas de U e de V, respectivamente, e seja $A \in M_{m\times n}(\mathbb{R})$. Considere a transformação linear $T\colon U \to V$ tal que $[T]_{\mathcal{BC}} = A$. Então, $\operatorname{Ker}(T)$ é formado pelas soluções do sistema linear homogêneo cuja matriz de coeficientes é A lidas como coordenadas em relação à base $\mathcal{B}$. Em particular, $\dim\bigl(\operatorname{Ker}(T)\bigr) = n - \operatorname{posto}(A)$ e, usando o Teorema do núcleo e da imagem (Teorema 6.3.1), $\dim\bigl(\operatorname{Im}(T)\bigr) = \operatorname{posto}(A)$.

Exercícios Lista 2 - Álgebra Linear II: Exs. 31–39.

6.6 Matriz da transformação composta

Nesta seção, veremos o efeito nas matrizes das operações de soma, multiplicação por escalar e composição envolvendo transformações lineares de que tratamos na Seção 6.4.

Começamos com a soma e a multiplicação por escalar. A composição será tratada separadamente, adiante.

Proposição 6.6.1 *Sejam U e V espaços vetoriais de dimensão finita, sejam $T_1, T_2\colon U \to V$ transformações lineares e sejam $\mathcal{B}$ e $\mathcal{C}$ bases ordenadas de U e de V, respectivamente. Então,*

(i) $[T_1 + T_2]_{\mathcal{BC}} = [T_1]_{\mathcal{BC}} + [T_2]_{\mathcal{BC}}$, *e*
(ii) $[\lambda T_1]_{\mathcal{BC}} = \lambda [T_1]_{\mathcal{BC}}$, *para todo* $\lambda \in \mathbb{R}$.

Demonstração Suponha que $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ e que $\mathcal{C} = \{v_1, v_2, \dots, v_m\}$. Ainda, escreva $[T_1]_{\mathcal{BC}} = (a_{ij})$ e $[T_2]_{\mathcal{BC}} = (b_{ij})$. Então, para cada $j = 1, \dots, n$, temos

$$\begin{aligned}(T_1 + T_2)(u_j) &= T_1(u_j) + T_2(u_j)\\ &= (a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{mj}v_m) + (b_{1j}v_1 + b_{2j}v_2 + \cdots + b_{mj}v_m) && \text{por definição de } [T_1]_{\mathcal{BC}} \text{ e } [T_2]_{\mathcal{BC}}\\ &= (a_{1j} + b_{1j})v_1 + (a_{2j} + b_{2j})v_2 + \cdots + (a_{mj} + b_{mj})v_m.\end{aligned}$$

Daí segue que $[T_1 + T_2]_{\mathcal{BC}} = (c_{ij})$, em que $c_{ij} = a_{ij} + b_{ij}$, para todos $i = 1, \dots, m$ e $j = 1, \dots, n$. Isso demonstra (i). Agora, tome $\lambda \in \mathbb{R}$, Então, para todo $j = 1, \dots, n$, vale

$$\begin{aligned}T_1(\lambda u_j) &= \lambda T_1(u_j)\\ &= \lambda(a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{mj}v_m)\\ &= (\lambda a_{1j})v_1 + (\lambda a_{2j})v_2 + \cdots + (\lambda a_{mj})v_m,\end{aligned}$$

donde segue (ii). □

Costumamos dizer que "a matriz da soma é a soma das matrizes" e que "a matriz de uma multiplicação por escalar é a multiplicação do escalar pela matriz". Para a composição, uma outra operação familiar entre matrizes surge: o produto.

Proposição 6.6.2 *Sejam U, V, W espaços vetoriais de dimensão finita e sejam $T\colon U \to V$ e $S\colon V \to W$ transformações lineares. Sejam $\mathcal{B}$, $\mathcal{C}$ e $\mathcal{D}$ bases ordenadas de U, de V e de W, respectivamente. Então, a transformação linear composta $S \circ T\colon U \to W$ está definida e*

$$[S \circ T]_{\mathcal{BD}} = [S]_{\mathcal{CD}}[T]_{\mathcal{BC}},$$

em que a concatenação de matrizes do lado direito da igualdade expressa o produto entre elas.

Demonstração Como o contradomínio de T coincide com o domínio de S, a composta $S \circ T\colon U \to W$ está definida e é uma transformação linear, como vimos na Proposição 6.4.1. Suponha que $\dim(U) = n$, $\dim(V) = m$ e $\dim(W) = r$. Então $[T]_{\mathcal{BC}} \in M_{m\times n}(\mathbb{R})$ e $[S]_{\mathcal{CD}} \in M_{r\times m}(\mathbb{R})$. Logo, o produto de matrizes $[S]_{\mathcal{CD}}[T]_{\mathcal{BC}}$ está definido e tem tamanho $r \times n$, que coincide com o tamanho de $[S \circ T]_{\mathcal{BD}}$. Resta mostrar que as matrizes têm as mesmas entradas.

Suponha que $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$. Lembre, da definição de matriz de uma transformação linear, que a j-ésima coluna de $[S \circ T]_{\mathcal{BD}}$ é o vetor de coordenadas $[(S \circ T)(u_j)]_{\mathcal{D}}$. Como $(S \circ T)(u_j) = S\big(T(u_j)\big)$, temos

$$\begin{aligned}[(S \circ T)(u_j)]_{\mathcal{D}} &= \big[S\big(T(u_j)\big)\big]_{\mathcal{D}}\\ &= [S]_{\mathcal{CD}}[T(u_j)]_{\mathcal{C}} && \text{pelo Teorema 6.5.6 aplicado a } S\\ &= [S]_{\mathcal{CD}}[T]_{\mathcal{BC}}[u_j]_{\mathcal{B}} && \text{pelo Teorema 6.5.6 aplicado a } T\end{aligned}$$

Mas, $[u_j]_{\mathcal{B}}$ é uma matriz-coluna de n linhas cujas entradas são todas nulas, exceto a j-ésima, que é igual a 1 (uma vez que u_j é o j-ésimo vetor da base ordenada $\mathcal{B}$). Assim, o efeito de multiplicar uma matriz por $[u_j]_{\mathcal{B}}$ à direita é precisamente selecionar a j-ésima coluna dessa matriz. Concluímos dessa discussão que, para cada $j = 1, \dots, n$, a j-ésima coluna da matriz $[S \circ T]_{\mathcal{BD}}$ coincide com a j-ésima coluna da matriz $[S]_{\mathcal{CD}}[T]_{\mathcal{BC}}$. Como essas matrizes têm o mesmo tamanho, elas são iguais. □

Será útil introduzir uma simplificação na notação de matriz de um operador linear, a ser adotada deste ponto em diante. Se U é um espaço vetorial de dimensão finita, $T\colon U \to U$ é um operador linear e $\mathcal{B}$ é uma base de U, então a matriz (quadrada) $[T]_{\mathcal{BB}}$ será denotada simplesmente por $[T]_{\mathcal{B}}$.

Exemplo 6.6.3 Considere os seguintes operadores lineares de $\mathbb{R}^2$:

$$\begin{aligned}F\colon \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 & \qquad G\colon \mathbb{R}^2 &\longrightarrow \mathbb{R}^2\\ (x, y) &\longmapsto (x, x - y) & (x, y) &\longmapsto (x + y, 2x)\end{aligned}$$

Determine as matrizes dos operadores lineares $F + G$, $3F$, $F \circ G$, $G \circ F$ e F^2 em relação à base canônica de $\mathbb{R}^2$.

Solução: A base canônica de $\mathbb{R}^2$ será denotada por can, ou seja, $\mathsf{can} = \{(1, 0), (0, 1)\}$. Neste exemplo, usaremos as Proposições 6.6.1 e 6.6.2. Para encontrar $[F]_{\mathsf{can}}$, calculamos

$$F(1, 0) = (1, 1) = 1(1, 0) + 1(0, 1) \qquad \text{e} \qquad F(0, 1) = (0, -1) = 0(1, 0) + (-1)(0, 1).$$

Assim,

$$[F]_{\mathsf{can}} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}.$$

Analogamente,

$$G(1, 0) = (1, 2) = 1(1, 0) + 2(0, 1) \qquad \text{e} \qquad G(0, 1) = (1, 0) = 1(1, 0) + 0(0, 1)$$

implicam que

$$[G]_{\mathsf{can}} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}.$$

Logo,

$$\begin{aligned}
[F + G]_{\mathsf{can}} &= [F]_{\mathsf{can}} + [G]_{\mathsf{can}} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}, \\
[3F]_{\mathsf{can}} &= 3[F]_{\mathsf{can}} = 3 \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 3 & -3 \end{bmatrix}, \\
[F \circ G]_{\mathsf{can}} &= [F]_{\mathsf{can}}[G]_{\mathsf{can}} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \\
[G \circ F]_{\mathsf{can}} &= [G]_{\mathsf{can}}[F]_{\mathsf{can}} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix}.
\end{aligned}$$

E, como $F^2 = F \circ F$,

$$[F^2]_{\mathsf{can}} = [F]^2_{\mathsf{can}} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Em particular, note que os operadores lineares $F \circ G$ e $G \circ F$ são diferentes. Além disso, o fato de $[F^2]_{\mathsf{can}} = I_2$ permite que se conclua que $F^2 = \mathrm{I}_{\mathbb{R}^2}$ (você consegue demonstrar isso?). $\diamond$

Exercícios Lista 2 - Álgebra Linear II: Exs. 40–43.

6.7 Mudança de base

Esta seção será dedicada a responder à pergunta surgida ao final do Exemplo 6.5.2: qual é a relação entre matrizes de uma mesma transformação linear em relação a diferentes escolhas de pares de bases?

Começamos com um resultado fundamental na construção de uma resposta adequada.

Proposição 6.7.1 *Sejam U e V espaços vetoriais de dimensão finita, seja $T: U \to V$ uma transformação linear e sejam $\mathcal{B}$ e $\mathcal{C}$ bases ordenadas de U e de V, respectivamente. Então, T é bijetora se, e somente se, a matriz $[T]_{\mathcal{BC}}$ for inversível. Além disso, neste caso, $[T]^{-1}_{\mathcal{BC}} = [T^{-1}]_{\mathcal{CB}}$.*

Demonstração Denote $n = \dim(U)$.

Suponha, primeiramente, que T é bijetora. Pelo Corolário 6.3.7, $\dim(V) = n$, e, portanto, $[T]_{\mathcal{BC}}$ é uma matriz quadrada de tamanho n. Ainda, a bijetividade de T implica, pela Proposição 6.3.5, que existe a transformação inversa

$T^{-1}: V \to U$ e ela satisfaz $T \circ T^{-1} = \mathrm{I}_V$ e $T^{-1} \circ T = \mathrm{I}_U$. Logo, usando a fórmula para a matriz da composta, obtida na Proposição 6.6.2,

$$[T]_{\mathcal{B}\mathcal{C}}[T^{-1}]_{\mathcal{C}\mathcal{B}} = [T \circ T^{-1}]_{\mathcal{C}} = [\mathrm{I}_V]_{\mathcal{C}} = I_n,$$

como vimos no Exemplo 6.5.3. Analogamente,

$$[T^{-1}]_{\mathcal{C}\mathcal{B}}[T]_{\mathcal{B}\mathcal{C}} = [T^{-1} \circ T]_{\mathcal{B}} = [\mathrm{I}_U]_{\mathcal{B}} = I_n.$$

Logo, $[T]_{\mathcal{B}\mathcal{C}}$ é inversível, e $[T]_{\mathcal{B}\mathcal{C}}^{-1} = [T^{-1}]_{\mathcal{C}\mathcal{B}}$.

Reciprocamente, suponha que $[T]_{\mathcal{B}\mathcal{C}}$ é inversível. Em particular, $[T]_{\mathcal{B}\mathcal{C}}$ é quadrada de tamanho n e $\dim(U) = n = \dim(V)$. Seja $A = [T]_{\mathcal{B}\mathcal{C}}^{-1}$. Considere a transformação linear $S: V \to U$ tal que $[S]_{\mathcal{C}\mathcal{B}} = A$. Mostremos que $S \circ T = \mathrm{I}_U$ e que $T \circ S = \mathrm{I}_V$. Isso implicará que T é bijetora (e que $S = T^{-1}$). Por um lado,

$$[S \circ T]_{\mathcal{B}} = [S]_{\mathcal{C}\mathcal{B}}[T]_{\mathcal{B}\mathcal{C}} = A[T]_{\mathcal{B}\mathcal{C}} = I_n = [\mathrm{I}_U]_{\mathcal{B}}.$$

Isso quer dizer que as transformações lineares $S \circ T$ e I_U, que já têm domínio e contradomínio comuns, coincidem em todos os elementos da base $\mathcal{B}$. Logo, $S \circ T = \mathrm{I}_U$. De modo análogo,

$$[T \circ S]_{\mathcal{C}} = [T]_{\mathcal{B}\mathcal{C}}[S]_{\mathcal{C}\mathcal{B}} = [T]_{\mathcal{B}\mathcal{C}}A = I_n = [\mathrm{I}_V]_{\mathcal{C}},$$

e isso implica $T \circ S = \mathrm{I}_V$, como queríamos. □

Exemplo 6.7.2 Mostre que a transformação linear

$$\begin{aligned} L: \mathbb{R}^2 &\longrightarrow \mathcal{P}_1(\mathbb{R}) \\ (x, y) &\longmapsto (3x - 4y) + (-x + 2y)t \end{aligned}$$

é bijetora e encontre uma expressão para sua inversa.

Solução: Considere as bases canônicas $\mathcal{B} = \{(1,0), (0,1)\}$ e $\mathcal{C} = \{1, t\}$ de $\mathbb{R}^2$ e $\mathcal{P}_1(\mathbb{R})$, respectivamente. Então, como $L(1,0) = 3 - t$ e $L(0,1) = -4 + 2t$, segue que

$$[L]_{\mathcal{B}\mathcal{C}} = \begin{bmatrix} 3 & -4 \\ -1 & 2 \end{bmatrix}.$$

Como $\det \begin{bmatrix} 3 & -4 \\ -1 & 2 \end{bmatrix} = 2 \neq 0$, essa matriz é inversível. Pelo que vimos acima, L é bijetora e a transformação linear inversa $L^{-1}: \mathcal{P}_1(\mathbb{R}) \to \mathbb{R}^2$ é tal que

$$[L^{-1}]_{\mathcal{C}\mathcal{B}} = [L]_{\mathcal{B}\mathcal{C}}^{-1} = \begin{bmatrix} 3 & -4 \\ -1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 2 \\ \frac{1}{2} & \frac{3}{2} \end{bmatrix}.$$

Isso quer dizer que $L^{-1}(1) = 1(1,0) + \frac{1}{2}(0,1) = \left(1, \frac{1}{2}\right)$ e que $L^{-1}(t) = 2(1,0) + \frac{3}{2}(0,1) = \left(2, \frac{3}{2}\right)$. Mas, conhecendo as imagens por L^{-1} de elementos de uma base de $\mathcal{P}_1(\mathbb{R})$, conhecemos a imagem por L^{-1} de qualquer elemento de $\mathcal{P}_1(\mathbb{R})$:

$$L^{-1}(a + bt) = aL^{-1}(1) + bL^{-1}(t) = a\left(1, \frac{1}{2}\right) + b\left(2, \frac{3}{2}\right) = \left(a + 2b, \frac{a + 3b}{2}\right). \quad \diamond$$

Já temos condição de responder à questão colocada no início da seção.

Teorema 6.7.3 *Sejam U e V espaços vetoriais de dimensão finita, seja $T: U \to V$ uma transformação linear, sejam $\mathcal{B}$ e $\mathcal{B}'$ bases ordenadas de U e sejam $\mathcal{C}$ e $\mathcal{C}'$ bases ordenadas de V. Então,*

$$[T]_{\mathcal{B}'\mathcal{C}'} = P[T]_{\mathcal{B}\mathcal{C}}Q^{-1}, \tag{6.17}$$

em que $P = [\mathrm{I}_V]_{\mathcal{C}\mathcal{C}'}$ e $Q = [\mathrm{I}_U]_{\mathcal{B}\mathcal{B}'}$.

Demonstração Para efeitos de verificação de compatibilidade de tamanhos das matrizes envolvidas, suponha que $\dim(U) = n$ e que $\dim(V) = m$. Então, $[T]_{\mathcal{B}\mathcal{C}}, [T]_{\mathcal{B}'\mathcal{C}'} \in M_{m\times n}(\mathbb{R})$. Além disso, $P = [\mathrm{I}_V]_{\mathcal{C}\mathcal{C}'} \in M_m(\mathbb{R})$ e $Q = [\mathrm{I}_U]_{\mathcal{B}\mathcal{B}'} \in M_n(\mathbb{R})$. Como a transformação identidade I_U é obviamente bijetora, segue da Proposição 6.7.1 que Q é inversível (e sua inversa Q^{-1} tem mesmo tamanho que Q). Portanto, o produto de três matrizes no lado direito de (6.17) está definido e tem tamanho igual ao da matriz no lado esquerdo. Resta-nos, assim, demonstrar a igualdade.

Considere a transformação composta $T \circ \mathrm{I}_U : U \to V$. Como I_U é a transformação identidade de U, é claro que $T \circ \mathrm{I}_U = T$. De modo análgo, $\mathrm{I}_V \circ T = T$. Da Proposição 6.6.2, obtemos

$$[T]_{\mathcal{B}'\mathcal{C}'}[\mathrm{I}_U]_{\mathcal{B}\mathcal{B}'} = [T \circ \mathrm{I}_U]_{\mathcal{B}\mathcal{C}'} = [T]_{\mathcal{B}\mathcal{C}'} = [\mathrm{I}_V \circ T]_{\mathcal{B}\mathcal{C}'} = [\mathrm{I}_V]_{\mathcal{C}\mathcal{C}'}[T]_{\mathcal{B}\mathcal{C}}.$$

Para obter, finalmente, (6.17), basta, agora, multiplicar ambos os termos dessa igualdade por Q^{-1} à esquerda. □

Aqui cabem alguns comentários. O primeiro é que para lembrar da expressão (6.17), o seguinte diagrama pode ser útil:

$$\begin{tikzcd}
U_{\mathcal{B}} \arrow[r, "T"] \arrow[d, "\mathrm{I}_U"'] & V_{\mathcal{C}} \arrow[d, "\mathrm{I}_V"] \\
U_{\mathcal{B}'} \arrow[r, "T"] & V_{\mathcal{C}'}
\end{tikzcd}$$

Veja que, no diagrama, indicamos, além dos espaços vetoriais e das transformações entre eles, as bases envolvidas na determinação das matrizes de transformações lineares.

Em segundo lugar, observe que, como I_V também é uma transformação linear bijetora, a matriz P também é inversível. Assim, (6.17) poderia ter sido apresentada na seguinte forma alternativa:

$$[T]_{\mathcal{B}\mathcal{C}} = P^{-1}[T]_{\mathcal{B}'\mathcal{C}'}Q. \tag{6.18}$$

(Enquanto (6.17) fornece $[T]_{\mathcal{B}'\mathcal{C}'}$ a partir de $[T]_{\mathcal{B}\mathcal{C}}$, (6.18) faz o oposto, dá uma expressão para $[T]_{\mathcal{B}\mathcal{C}}$ em função de $[T]_{\mathcal{B}'\mathcal{C}'}$. Claro que uma das fórmula pode ser obtida a partir da outra, e ambas a partir do diagrama acima.)

O Teorema 6.7.3 será mais frequentemente utilizado para o estudo de operadores lineares. O corolário a seguir é uma consequência imediada do teorema (na forma (6.18), com $U = V$, $\mathcal{B} = \mathcal{C}$ e $\mathcal{B}' = \mathcal{C}'$).

Corolário 6.7.4 *Seja V um espaço vetorial de dimensão finita, seja $T : V \to V$ um operador linear e sejam $\mathcal{B}$ e $\mathcal{C}$ bases ordenadas de V. Então,*

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P, \tag{6.19}$$

em que $P = [\mathrm{I}_V]_{\mathcal{C}\mathcal{B}}$.

Aqui, o diagrama também ajuda:

$$\begin{tikzcd}
V_{\mathcal{C}} \arrow[r, "T"] \arrow[d, "\mathrm{I}_V"'] & V_{\mathcal{C}} \arrow[d, "\mathrm{I}_V"] \\
V_{\mathcal{B}} \arrow[r, "T"] & V_{\mathcal{B}}
\end{tikzcd}$$

As matrizes dos operadores identidade que aparecem nas fórmula vistas acima são chamadas *matrizes de mudança de base*. Elas, além de relacionarem $[T]_{\mathcal{B}\mathcal{C}}$ e $[T]_{\mathcal{B}'\mathcal{C}'}$, também servem para "mudar coordenadas", como mostra o próximo resultado.

Proposição 6.7.5 *Seja V um espaço vetorial de dimensão finita e sejam $\mathcal{B}$ e $\mathcal{C}$ bases ordenadas de V. Então, para todo $u \in V$, temos*

$$[u]_{\mathcal{B}} = P[u]_{\mathcal{C}},$$

em que $P = [\mathrm{I}_V]_{\mathcal{CB}}$.

Demonstração Sabemos, do Teorema 6.5.6 (aplicado ao operador linear $\mathrm{I}_V : V \longrightarrow V$, usando as bases $\mathcal{C}$ no domínio e $\mathcal{B}$ no contradomínio), que

$$[u]_{\mathcal{B}} = [\mathrm{I}_V(u)]_{\mathcal{B}} = [\mathrm{I}_V]_{\mathcal{CB}}[u]_{\mathcal{C}},$$

que é o que desejávamos. □

Note que, como $\mathrm{I}_V(u) = u$, para todo $u \in V$, as colunas de $[\mathrm{I}_V]_{\mathcal{CB}}$ são precisamente as coordenadas, em relação à base $\mathcal{B}$ dos vetores que compõem a base $\mathcal{C}$, na ordem em que eles estão listados em $\mathcal{C}$.

Mais um comentário a respeito de matrizes de mudança de base merece registro. Sabemos que uma matriz de mudança de base é sempre inversível (já que é a matriz do operador identidade, que é bijetor), mas vale mais. Se V é um espaço vetorial de dimensão finita e $\mathcal{B}$ e $\mathcal{C}$ são bases ordenadas de V, então

$$[\mathrm{I}_V]_{\mathcal{CB}} = [\mathrm{I}_V]_{\mathcal{BC}}^{-1}. \tag{6.20}$$

Isso segue da Proposição 6.7.1 e do fato óbvio que $\mathrm{I}_V^{-1} = \mathrm{I}_V$.

Observação O conceito de matriz de mudança de base já havia surgido no contexto de vetores de $\mathbb{V}^3$, mais especificamente, na ocasião em que vimos o Teorema 2.6.1. Veja que a notação que foi utilizada então se expressa, agora, da seguinte maneira: $M_{\mathcal{CB}} = [\mathrm{I}_{\mathbb{V}^3}]_{\mathcal{CB}}$, de sorte que há uniformidade na nomenclatura.

Exemplo 6.7.6 Considere o operador linear de $\mathbb{R}^2$ definido por

$$\begin{aligned} T\colon \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x - y, x + 2y). \end{aligned}$$

Encontre as matrizes $[T]_{\mathcal{B}}$ e $[T]_{\mathcal{C}}$, em que $\mathcal{B}$ denota a base canônica de $\mathbb{R}^2$ e $\mathcal{C} = \{(1, 3), (-1, 0)\}$.

Solução: Temos $T(1, 0) = (1, 1)$ e $T(0, 1) = (-1, 2)$, o que acarreta $[T]_{\mathcal{B}} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$. Agora, usaremos (6.19) para determinar $[T]_{\mathcal{C}}$. Para tanto, é preciso encontrar $P = [\mathrm{I}_{\mathbb{R}^2}]_{\mathcal{CB}}$. Como

$$\mathrm{I}_{\mathbb{R}^2}(1, 3) = (1, 3) = 1(1, 0) + 3(0, 1) \qquad \text{e} \qquad \mathrm{I}_{\mathbb{R}^2}(-1, 0) = (-1, 0) = (-1)(1, 0) + 0(0, 1),$$

segue que $P = \begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix}$. Logo[3],

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P = \begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 0 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 7 & -1 \\ 13 & 2 \end{bmatrix}.$$

Obviamente, uma outra solução seria calcular diretamente:

$$T(1, 3) = (-2, 7) = \frac{7}{3}(1, 3) + \frac{13}{3}(-1, 0) \qquad \text{e} \qquad T(-1, 0) = (-1, -1) = \left(-\frac{1}{3}\right)(1, 3) + \frac{2}{3}(-1, 0). \quad \diamond$$

---

[3] Aqui usaremos a utilíssima fórmula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

que vale sempre que $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \neq 0$.

Exemplo 6.7.7 (Prova 2, Álgebra Linear II, 2015) Sabendo que a matriz da transformação linear $T\colon \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}^2$ em relação às bases ordenadas $\mathcal{B} = \{1, x + x^2, x^2\}$ e $C = \{(1, -1), (1, 1)\}$ de $\mathcal{P}_2(\mathbb{R})$ e de $\mathbb{R}^2$, respectivamente, é $\begin{bmatrix} 3 & 0 & 1 \\ 1 & 2 & -1 \end{bmatrix}$, o vetor $T(x^2 - x + 1)$ é

(A) $(5, -3)$ (B) $(2, -8)$ (C) $(5, 3)$ (D) $(-2, -8)$ (E) $(-5, -5)$

Solução: Seja $v = x^2 - x + 1$. Sabemos, pelo Teorema 6.5.6, que $[T(v)]_C = [T]_{\mathcal{B}C}[v]_{\mathcal{B}}$. Agora, pela Proposição 6.7.5, $[v]_{\mathcal{B}} = P[v]_{\mathcal{D}}$, em que $\mathcal{D}$ denota a base canônica de $\mathcal{P}_2(\mathbb{R})$ (isto é, $\mathcal{D} = \{1, x, x^2\}$), e $P = [\mathrm{I}]_{\mathcal{D}\mathcal{B}}$, em que I denota o operador identidade de $\mathcal{P}_2(\mathbb{R})$. É claro que $[v]_{\mathcal{D}} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$. Agora, sabemos que $P = [\mathrm{I}]_{\mathcal{D}\mathcal{B}} = [\mathrm{I}]^{-1}_{\mathcal{B}\mathcal{D}}$, e essa segunda matriz é fácil de descrever :

$$[\mathrm{I}]_{\mathcal{B}\mathcal{D}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

(O que foi feito aqui foi escrever as coordenadas dos vetores de $\mathcal{B}$ em relação à base $\mathcal{D}$, o que é fácil, pois $\mathcal{D}$ é a base canônica.) Procedemos, agora, à inversão dessa matriz:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right]$$

Assim,

$$P = [\mathrm{I}]_{\mathcal{D}\mathcal{B}} = [\mathrm{I}]^{-1}_{\mathcal{B}\mathcal{D}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

Concluímos, portanto, que

$$[v]_{\mathcal{B}} = P[v]_{\mathcal{D}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}.$$

Logo,

$$[T(v)]_C = [T]_{\mathcal{B}C}[v]_{\mathcal{B}} = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}.$$

Assim, $T(v) = 5(1, -1) + (-3)(1, 1) = (2, -8)$. *Resposta:* (B)

Poderíamos ir além e encontrar a expressão para $T(a + bx + cx^2)$. Basta fazer

$$[a + bx + cx^2]_{\mathcal{B}} = P[a + bx + cx^2]_{\mathcal{D}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ b \\ -b + c \end{bmatrix},$$

donde

$$[T(a + bx + cx^2)]_C = [T]_{\mathcal{B}C}[a + bx + cx^2]_{\mathcal{B}} = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ -b + c \end{bmatrix} = \begin{bmatrix} 3a - b + c \\ a + 3b - c \end{bmatrix}.$$

Portanto, $T(a + bc + cx^2) = (3a - b + c)(1, -1) + (a + 3b - c)(1, 1) = (4a + 2b, -2a + 4b - 2c)$. $\diamond$

Observação Neste capítulo, vimos que se U e V são espaços vetoriais de dimensão finita, com $\dim(U) = n$ e $\dim(V) = m$, e se $\mathcal{B}$ e C são bases ordenadas de U e de V, respectivamente, a cada transformação linear $T\colon U \to V$ associa-se uma matriz $[T]_{\mathcal{B}C} \in M_{m\times n}(\mathbb{R})$. Dois comentários a respeito dessa associação são devidos:

(i) Essa associação é injetora, no sentido de que se T_1 e T_2 são transformações lineares de U em V tais que $[T_1]_{\mathcal{B}C} = [T_2]_{\mathcal{B}C}$, então $T_1 = T_2$. (Isso decorre do fato de, além de terem o mesmo domínio e contradomínio, as tranformações lineares T_1 e T_2 têm a mesma imagem em cada elemento do domínio. Com efeito, dado $u \in U$, temos

$$\begin{aligned}[T_1(u) - T_2(u)]_C &= [T_1(u)]_C - [T_2(u)]_C = [T_1]_{\mathcal{B}C}[u]_{\mathcal{B}} - [T_2]_{\mathcal{B}C}[u]_{\mathcal{B}} \\ &= \big([T_1]_{\mathcal{B}C} - [T_2]_{\mathcal{B}C}\big)[u]_{\mathcal{B}} = 0[u]_{\mathcal{B}} = 0.\end{aligned}$$

Logo, $T_1(u) = T_2(u)$, uma vez que o único vetor de V que têm todas as coordenadas nulas é 0_V.)

(ii) Essa associação é também sobrejetora; em outras palavras, dada $A \in M_{m\times n}(\mathbb{R})$ existe uma (única, como acabamos de ver) transformação linear $T\colon U \to V$ tal que $[T]_{\mathcal{B}C} = A$. (Com efeito, suponha que $\mathcal{B} = \{u_1, \dots, u_n\}$ e $C = \{v_1, \dots, v_m\}$, então, se $A = (a_{ij})$, basta tomar T como sendo a transformação linear definida por $T(u_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{mj}v_m$, para todo $j = 1, \dots, n$.)

Note que essas observações foram tacitamente assumidas nos Exemplos 6.5.8, 6.5.9 e 6.7.7, nos quais, dada uma matriz, assumimos que estava definida uma única transformação linear cuja matriz coincidia com a matriz dada.

Exercícios Lista 2 - Álgebra Linear II: Exs. 44–48.

Capítulo 7
Diagonalização de operadores

Neste capítulo, procedemos a uma análise mais profunda de operadores lineares em espaço de dimensão finita a fim de descrevê-los da maneira mais simples possível, em um sentido que ficará claro à medida que progredimos.

Para introduzir o conceito principal deste capítulo, o de autovetor de um operador linear, comecemos por um exemplo simples, aparentemente não diretamente relacionado com o estudo de operadores lineares.

Seja n um inteiro positivo. Dizemos que uma matriz $D = (d_{ij}) \in M_n(\mathbb{R})$ é *diagonal* se $d_{ij} = 0$ para todos $i, j = 1, \dots, n$ tais que $i \neq j$. Em outras palavras, D é diagonal se apenas as entradas em sua diagonal principal são eventualmente não nulas, isto é, se D o seguinte formato:

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{n-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

(Destacamos as entradas na diagonal principal de D com a cor vermelha. Todas as demais entradas são nulas.) Em algumas ocasiões, quando for conveniente, denotaremos a matriz diagonal $D \in M_n(\mathbb{R})$ cujas entradas na diagonal principal são $\lambda_1, \lambda_2, \dots, \lambda_n$, nessa ordem, por $D = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

De acordo com a definição do produto entre matrizes, fica claro que as potências da matriz diagonal D são também matrizes diagonais, dadas por

$$D^r = \begin{bmatrix} \lambda_1^r & 0 & \dots & 0 & 0 \\ 0 & \lambda_2^r & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{n-1}^r & 0 \\ 0 & 0 & \dots & 0 & \lambda_n^r \end{bmatrix},$$

para todo inteiro positivo r. Usando a notação introduzida acima, se $D = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, então $D^r = \operatorname{diag}(\lambda_1^r, \lambda_2^r, \dots, \lambda_n^r)$.

Em geral, calcular a potência de uma matriz quadrada é muito custoso, pois são muitas as operações envolvendo suas entradas a serem efetuadas. Porém, se D é diagonal, como vimos, suas potências são calculadas de modo imediado. Há um caso intermediário, em que, apesar de não se tratar de uma matriz diagonal, é bastante rápido o cálculo de suas potências. Este é o caso destacado na definição a seguir.

Definição Seja n um inteiro positivo. Dizemos que uma matriz $A \in M_n(\mathbb{R})$ é *diagonalizável* se existir uma matriz inversível $P \in M_n(\mathbb{R})$ tal que $P^{-1}AP$ seja uma matriz diagonal.

Voltando à nossa discussão sobre potências, se $A \in M_n(\mathbb{R})$ é diagonalizável, digamos, $P^{-1}AP = D$, com $P, D \in M_n(\mathbb{R})$, P inversível e D diagonal, então, para calcular a r-ésima potência de A, procedemos da seguinte maneira: de $P^{-1}AP = D$, segue que $A = PDP^{-1}$; assim,

$$A^r = (PDP^{-1})^r = (PDP^{-1})(PDP^{-1})\dots(PDP^{-1}),$$

em que o produto do lado direito tem r fatores da forma PDP^{-1}. Nessa expressão, toda ocorrência da matriz P, exceto pela mais à esquerda, vem acompanhada de P^{-1} multiplicada por ela à esquerda:

$$(PDP^{-1})(PDP^{-1})\dots(PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)D(P^{-1}P)D\dots D(P^{-1}P)DP^{-1}.$$

Cada um dos produtos $P^{-1}P$ é igual à matriz identidade I_r. Portanto, todas as ocorrências de P e P^{-1} entre D's se cancelam, de modo que temos, ao final,

$$A^r = (PDP^{-1})^r = PD^rP^{-1},$$

que resulta em uma expressão para a potência de A envolvendo um número muito menor de cálculos a serem efetuados entre as entradas de A.

Em resumo, se A for diagonalizável, suas potências são fáceis de serem calculadas. Nem toda matriz é diagonalizável, entretanto. Será objeto deste capítulo determinar condições necessárias e suficientes para tanto. E, em caso afirmativo, veremos como encontrar as matrizes P e D.

7.1 Autovalores e autovetores

Introduzimos o conceito principal deste capítulo. (E, posteriormente, veremos como ele se relaciona ao conceito de matriz diagonalizável.)

Definição Seja V um espaço vetorial e seja $T\colon V \longrightarrow V$ um operador linear em V. Dizemos que um vetor $v \in V$ é um *autovetor* de T se $v \neq 0_V$ e existir $\lambda \in \mathbb{R}$ tal que $T(v) = \lambda v$. Neste caso, λ é chamado *autovalor* de T, e dizemos que v é um autovetor de T associado ao autovalor λ.

Autovetores também são conhecidos, na literatura, como *vetores próprios* ou *vetores característicos*. Similarmente, autovalores são também chamados de *valores próprios* ou *valores característicos*.

Exemplo 7.1.1 Considere o operador linear T de $\mathbb{R}^2$ definido por $T(x, y) = (x, -y)$, para todo $(x, y) \in \mathbb{R}^2$. Como $T(1, 0) = (1, 0) = 1(1, 0)$, de acordo com a definição acima, 1 é um autovalor de T e $(1, 0)$ é um autovetor de T associado ao autovalor 1. Ainda, $T(0, 1) = (0, -1) = (-1)(0, 1)$, donde segue que -1 é um autovalor de T e $(0, 1)$ é um autovetor de T associado ao autovalor -1. Agora, $(2, 3)$ **não é** um autovetor de T, pois $T(2, 3) = (2, -3)$ e não existe $\lambda \in \mathbb{R}$ que satisfaça $(2, -3) = T(2, 3) = \lambda(2, 3)$.

Um autovetor só pode estar associado a um único autovalor, pois se $v \in V$ é um vetor não nulo tal que $T(v) = \lambda v$ e $T(v) = \mu v$, como $\lambda, \mu \in \mathbb{R}$, então $\lambda v = \mu v$, o que implica $(\lambda - \mu)v = 0_V$. Como v não é o vetor nulo, segue (do item (iii) da Proposição 4.1.8) que $\lambda - \mu = 0$, ou, ainda, que $\lambda = \mu$.

Mas, se $\lambda \in \mathbb{R}$ é um autovalor de um operador linear T, então há infinitos autovetores de T associados a λ. Isso é o que veremos a seguir.

Dado um espaço vetorial V e um operador linear $T\colon V \longrightarrow V$, se $\lambda \in \mathbb{R}$, utilizaremos a seguinte notação:

$$V(\lambda) = \{v \in V \mid T(v) = \lambda v\}$$

Se $\lambda \in \mathbb{R}$ é um autovalor de T, então $V(\lambda)$ é o conjunto formado por todos os autovetores de T associados a λ, mais o vetor nulo (uma vez que $T(0_V) = 0_V = \lambda 0_V$). A proposição abaixo mostra que se λ é um autovalor de T, então $V(\lambda)$ não só é um conjunto infinito, como é um subespaço de V.

Proposição 7.1.2 *Seja V um espaço vetorial, seja $T\colon V \longrightarrow V$ um operador linear de V e seja $\lambda \in \mathbb{R}$. Então, $V(\lambda) = \operatorname{Ker}(T - \lambda \mathrm{I}_V)$. Em particular, $V(\lambda)$ é um subespaço de V.*

Demonstração Tome $v \in V$. Então,

$$\begin{aligned} v \in V(\lambda) &\iff T(v) = \lambda v \\ &\iff T(v) - \lambda v = 0_V \\ &\iff \big(T - \lambda \operatorname{I}_V\big)(v) = 0_V \\ &\iff v \in \operatorname{Ker}(T - \lambda \operatorname{I}_V), \end{aligned}$$

que é o que desejávamos mostrar. (Observe que ser um elemento de $V(\lambda)$ é ser um elemento do núcleo de um operador — não do operador T, mas do operador $T - \lambda \operatorname{I}_V$.) □

Definição Se T é um operador linear no espaço vetorial V e $\lambda \in \mathbb{R}$ é um autovalor de T, o subespaço $V(\lambda)$ é chamado *autoespaço* de V associado ao autovalor λ.

Dados um operador linear T em um espaço vetorial V e um escalar λ, como vimos na Proposição 7.1.2, $V(\lambda)$ é sempre um subespaço de V e, portanto, sempre contém, pelo menos, o vetor nulo de V. Para λ ser um autovalor de T é preciso que haja, em $V(\lambda)$, pelo menos um vetor não nulo (isto é, que T tenha um autovetor associado λ). Assim, temos o seguinte critério: λ é um autovalor de T se, e somente se, $V(\lambda) \neq \{0_V\}$.

Exemplo 7.1.3 Determine todos os autovalores e todos os autovetores do operador linear de $\mathbb{R}^2$ definido por $T(x, y) = (x, -y)$, para todo $(x, y) \in \mathbb{R}^2$.

Solução: Dado $\lambda \in \mathbb{R}$, vejamos para que valores de λ existe $(x, y) \neq (0, 0)$ tal que $T(x, y) = \lambda(x, y)$. Então, $T(x, y) = \lambda(x, y)$ se, e somente se, $(x, -y) = (\lambda x, \lambda y)$, o que, por sua vez, ocorre se, e somente se

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

Se $x \neq 0$, a primeira equação implica $\lambda = 1$ e, assim, da segunda, obtemos $y = 0$. Se $x = 0$, como estamos procurando soluções com $(x, y) \neq (0, 0)$, então $y \neq 0$ e, neste caso, $\lambda = -1$ e $x = 0$.

Conclusão: os únicos autovalores de T são 1 e -1. Além disso, $V(1) = \{(x, 0) \mid x \in \mathbb{R}\}$ e $V(-1) = \{(0, y) \mid y \in \mathbb{R}\}$. Em outras palavras, os autovetores de T associados ao autovalor 1 são os vetores da forma $(x, 0)$, com $x \in \mathbb{R}, x \neq 0$; os autovetores de T associados ao autovalor -1 são os vetores da forma $(0, y)$, com $y \in \mathbb{R}$, $y \neq 0$. (Para qualquer $\lambda \in \mathbb{R}$, $\lambda \neq 1, -1$, temos $V(\lambda) = \{(0, 0)\}$.) ◇

Exemplo 7.1.4 Considere o operador linear T em $\mathcal{P}_2(\mathbb{R})$ tal que

$$[T]_{\mathcal{B}} = \begin{bmatrix} -10 & 4 & 6 \\ -8 & 2 & 6 \\ -8 & 4 & 4 \end{bmatrix},$$

em que $\mathcal{B} = \{1 + t, 1 + t^2, t + t^2\}$. Mostre que -2 é um autovalor de T e determine $\dim\big(V(-2)\big)$.

Solução: Sabemos que -2 será um autovalor de T se, e somente se, $V(-2) \neq \{0_{\mathcal{P}_2(\mathbb{R})}\}$, o que, em vista de $V(-2)$ ser um subespaço de $\mathcal{P}_2(\mathbb{R})$, é equivalente a $\dim\big(V(-2)\big) \neq 0$. Assim, se determinarmos $\dim\big(V(-2)\big)$, respondemos às duas questões simultaneamente. Vimos que $V(-2) = \operatorname{Ker}\big(T - (-2) \operatorname{I}_{\mathcal{P}_2(\mathbb{R})}\big)$. Vamos, assim, estudar o núcleo do operador $T - (-2) \operatorname{I}_{\mathcal{P}_2(\mathbb{R})} = T + 2 \operatorname{I}_{\mathcal{P}_2(\mathbb{R})}$. Dado $p \in \mathcal{P}_2(\mathbb{R})$, sabemos que

$$
\begin{aligned}
p \in \operatorname{Ker}(T + 2\,\mathrm{I}_{\mathcal{P}_2(\mathbb{R})}) \quad &\Longleftrightarrow \quad (T + 2\,\mathrm{I}_{\mathcal{P}_2(\mathbb{R})})(p) = 0_{\mathcal{P}_2(\mathbb{R})} \\
&\Longleftrightarrow \quad \big[(T + 2\,\mathrm{I}_{\mathcal{P}_2(\mathbb{R})})(p)\big]_{\mathcal{B}} = [0_{\mathcal{P}_2(\mathbb{R})}]_{\mathcal{B}} = \begin{bmatrix}0\\0\\0\end{bmatrix} \\
&\Longleftrightarrow \quad [T + 2\,\mathrm{I}_{\mathcal{P}_2(\mathbb{R})}]_{\mathcal{B}}[p]_{\mathcal{B}} = \begin{bmatrix}0\\0\\0\end{bmatrix} \qquad \text{pelo Teorema 6.5.6} \\
&\Longleftrightarrow \quad \big([T]_{\mathcal{B}} + 2[\mathrm{I}_{\mathcal{P}_2(\mathbb{R})}]_{\mathcal{B}}\big)[p]_{\mathcal{B}} = \begin{bmatrix}0\\0\\0\end{bmatrix} \qquad \text{pela Proposição 6.6.1} \\
&\Longleftrightarrow \quad \big([T]_{\mathcal{B}} + 2I_3\big)[p]_{\mathcal{B}} = \begin{bmatrix}0\\0\\0\end{bmatrix} \qquad \text{pois } [\mathrm{I}_{\mathcal{P}_2(\mathbb{R})}]_{\mathcal{B}} = I_3 \text{ (Exemplo 6.5.3)} \\
&\Longleftrightarrow \quad \left(\begin{bmatrix}-10 & 4 & 6\\-8 & 2 & 6\\-8 & 4 & 4\end{bmatrix} + 2\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}\right)[p]_{\mathcal{B}} = \begin{bmatrix}0\\0\\0\end{bmatrix} \\
&\Longleftrightarrow \quad \begin{bmatrix}-8 & 4 & 6\\-8 & 4 & 6\\-8 & 4 & 6\end{bmatrix}[p]_{\mathcal{B}} = \begin{bmatrix}0\\0\\0\end{bmatrix}
\end{aligned}
$$

Assim, $p \in \operatorname{Ker}(T + 2\,\mathrm{I}_{\mathcal{P}_2(\mathbb{R})})$ se, e somente se, suas coordenadas, em relação à base $\mathcal{B}$, são solução do sistema linear homogêneo (em notação matricial)

$$
\begin{bmatrix}-8 & 4 & 6\\-8 & 4 & 6\\-8 & 4 & 6\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}. \tag{7.1}
$$

Sabemos, pela Proposição 4.6.6, que o subespaço de $\mathbb{R}^3$ formado pelas soluções de (7.1) tem dimensão igual ao número de variáveis livres do sistema, que, no caso, fica claro ser igual a 2. Portanto, as soluções de (7.1) são combinações lineares de dois vetores, digamos u_1 e u_2, em $\mathbb{R}^3$ com $\{u_1, u_2\}$ LI. Segue da Proposição 4.5.1, que os elementos de $V(-2) = \operatorname{Ker}(T + 2\,\mathrm{I}_{\mathcal{P}_2(\mathbb{R})})$ são as combinações lineares dos vetores q_1 e q_2 de $\mathcal{P}_2(\mathbb{R})$ cujas coordenadas, em relação à base $\mathcal{B}$, são u_1 e u_2. Logo, $\{q_1, q_2\}$ é uma base de $V(-2)$, e, portanto, $\dim\big(V(-2)\big) = 2$.

Para determinar explicitamente q_1 e q_2, é preciso encontrar as soluções de (7.1). Resolvendo esse sistema, obtemos que suas soluções são da forma

$$
\left(\frac{y}{2} + \frac{3z}{4}, y, z\right) = y\left(\frac{1}{2}, 1, 0\right) + z\left(\frac{3}{4}, 0, 1\right), \quad y, z \in \mathbb{R}.
$$

Segue que o subespaço de $\mathbb{R}^3$ formado pelas soluções de (7.1) é $\big[\big(\frac{1}{2}, 1, 0\big), \big(\frac{3}{4}, 0, 1\big)\big] = \big[(1, 2, 0), (3, 0, 4)\big]$. Assim, $V(-2) = \big[(1, 2, 0)_{\mathcal{B}}, (3, 0, 4)_{\mathcal{B}}\big] = \big[3 + t + 2t^2, 3 + 7t + 4t^2\big]$, uma vez que $(1, 2, 0)_{\mathcal{B}} = 3 + t + 2t^2$ e $(3, 0, 4)_{\mathcal{B}} = 3 + 7t + 4t^2$. $\diamond$

Exemplo 7.1.5 Seja V um espaço vetorial. Determine todos os autovalores e autovetores do operador identidade I_V e do operador nulo N de V.

Solução: Para todo $v \in V$, temos $\mathrm{I}_V(v) = v = 1v$. Assim, 1 é o único autovalor de I_V e $V(1) = V$. Logo, todo vetor não nulo de V é um autovetor de I_V associado a 1. (Note que $\mathrm{I}_V - 1\,\mathrm{I}_V = N$, e, portanto, $V(1) = \operatorname{Ker}(\mathrm{I}_V - 1\,\mathrm{I}_V) = \operatorname{Ker}(N) = V$.) O operador nulo N de V satisfaz $N(v) = 0_V = 0v$, para todo $v \in V$. Assim, 0 é o único autovalor de N e $V(0) = V$. (Aqui, $N - 0\,\mathrm{I}_V = V$, e, assim, $V(0) = \operatorname{Ker}(N - 0\,\mathrm{I}_V) = \operatorname{Ker}(N) = V$).

Mais geralmente, fixe $\lambda \in \mathbb{R}$ e considere o operador T_λ de V definido por $T_\lambda(v) = \lambda v$, para todo $v \in V$. (Em outras palavras, $T_\lambda = \lambda\,\mathrm{I}_V$.) Então, λ é o único autovalor de T_λ e $V(\lambda) = \operatorname{Ker}(T_\lambda - \lambda\,\mathrm{I}_V) = \operatorname{Ker}(N) = V$. Os dois operadores de que tratamos acima são casos especiais: $\mathrm{I}_V = T_1$ e $N = T_0$. $\diamond$

Observação Há uma relação entre injetividade e autovalores de um operador linear: se T é um operador linear em um espaço vetorial V, então são equivalentes:

(a) T é injetor;
(b) 0 é autovalor de T;
(c) $\dim\big(V(0)\big) \neq 0$.

Com efeito, já vimos que 0 é autovalor se, e somente se $V(0) \neq \{0_V\}$, o que, por sua vez é equivalente a $\dim\big(V(0)\big) \neq 0$. Portanto, (b) e (c) são equivalentes. A equivalência entre (a) e (b) segue da observação que $V(0) = \mathrm{Ker}(T - 0\,\mathrm{I}_V) = \mathrm{Ker}(T)$ e da Proposição 6.2.5.

Exercícios Lista 2 - Álgebra Linear II: Exs. 49–55.

7.2 O polinômio característico

Vimos, na seção anterior, que se λ é um autovalor de um operador linear T no espaço vetorial V, então o autoespaço de T associado a λ é dado por $V(\lambda) = \mathrm{Ker}(T - \lambda\,\mathrm{I}_V)$, e, portanto, os autovetores de T associados a λ são todos os vetores de $V(\lambda)$ com exceção do vetor nulo. Resta dispormos de um instrumento para determinar o conjunto completo de autovalores de T. Isso será provido pelo polinômio característico de T, como veremos nesta seção.

Definição Seja $A = (a_{ij}) \in M_n(\mathbb{R})$. O *polinômio característico* de A é definido por

$$p_A(t) = \det(A - tI_n) = \det \begin{bmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{bmatrix}.$$

Observe que, de fato, $p_A(t)$ é um polinômio na variável t e que $p_A(t)$ tem grau n, isso segue do fato de que ao se expandir o determinante que define o polinômio característico de A, obteremos uma soma de $n!$ termos, cada um deles sendo um produto de n fatores que são entradas da matriz $A - tI_n$. O único desses termos que resultará em um polinômio de grau n é dado pelo produto dos elementos na diagonal principal de $A - tI_n$; os demais termos no determinante serão polinômios de grau menor do que n. O coeficiente do termo líder desse termo de grau n é dado por $(-1)^n$. Além disso, sabemos que o termo de grau zero de $p_A(t)$ é dado por $p_A(0) = \det(A - 0I_n) = \det(A)$. Assim,

$$p_A(t) = \alpha_n t^n + \alpha_{n-1} t^{n-1} + \cdots + \alpha_1 t + \alpha_0,$$

em que $\alpha_n = (-1)^n$ e $\alpha_0 = \det(A)$. (Os demais coeficientes do polinômio característico de A também podem ser expressos em termos das entradas de A, em especial, não é difícil ver que $\alpha_{n-1} = (-1)^{n-1}\,\mathrm{tr}(A)$, mas, para os objetivos destas notas, não será necessário entrar nesses detalhes.)

Duas matrizes diferentes podem ter o mesmo polinômio característico. Isso ocorre, especialmente, em um caso que já encontramos antes.

Definição Dadas $A, B \in M_n(\mathbb{R})$, dizemos que A e B são *semelhantes* se existir uma matriz $P \in M_n(\mathbb{R})$ inversível tal que $A = P^{-1}BP$.

É claro que se existe $P \in M_n(\mathbb{R})$ inversível tal que $A = P^{-1}BP$, então também existe $Q \in M_n(\mathbb{R})$ inversível tal que $B = Q^{-1}AQ$, basta tomar $Q = P^{-1}$. Ou seja, a semalhança entre A e B independe da ordem em que essas matrizes são listadas.

Proposição 7.2.1 *Matrizes semelhantes têm o mesmo polinômio característico.*

Demonstração Sejam $A, B \in M_n(\mathbb{R})$ e seja $P \in M_n(\mathbb{R})$ inversível tal que $A = P^{-1}BP$. Então,

$$\begin{aligned} p_A(t) &= \det(A - tI_n) = \det(P^{-1}BP - tI_n) \\ &= \det(P^{-1}BP - tP^{-1}I_nP) = \det\big(P^{-1}(B - tI_n)P\big) \\ &= \det(P^{-1})\det(B - tI_n)\det(P) = \det(B - tI_n)\det(P^{-1})\det(P) \\ &= \det(B - tI_n)\det(P)^{-1}\det(P) = \det(B - tI_n) \\ &= p_B(t), \end{aligned}$$

como desejávamos. □

Observação Mesmo matrizes não semelhantes podem ter o mesmo polinômio característico. Por exemplo, isso ocorre com $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ e $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. (Você consegue ver por que essas matrizes não são semelhantes?)

Havíamos encontrado matrizes semelhantes no Corolário 6.7.4: se $\mathcal{B}$ e $\mathcal{C}$ são bases ordenadas de um espaço vetorial V de dimensão finita e T é um operador linear de V, então $[T]_{\mathcal{B}}$ e $[T]_{\mathcal{C}}$ são matrizes semelhantes. Assim, podemos falar em polinômio característico de um operador linear, conforme a definição a seguir.

Definição Seja V um espaço vetorial de dimensão finita e seja $T\colon V \to V$ um operador linear. O *polinômio característico* de T é definido por

$$p_T(t) = p_A(t),$$

em que $A = [T]_{\mathcal{B}}$ e $\mathcal{B}$ é uma base ordenada qualquer de V.

O polinômio característico de T não depende da escolha da base ordenada $\mathcal{B}$ de V: se $\mathcal{C}$ é outra base ordenada de V, as matrizes $A = [T]_{\mathcal{B}}$ e $B = [T]_{\mathcal{C}}$ são semelhantes, e, pela Proposição 7.2.1, $p_A(t) = p_B(t)$.

Vale registrar que $p_T(t)$ é um polinômio de grau n, em que $n = \dim(V)$.

A importância do polinômio característico reside no fato de ele conter, nas suas raízes, os autovalores do operador linear, como veremos na próxima proposição, cuja demonstração movimentará diversos resultados que vimos acumulando ao longo dessas notas.

Proposição 7.2.2 *Seja V um espaço vetorial de dimensão finita, seja $T\colon V \to V$ um operador linear de V e seja $\lambda \in \mathbb{R}$. Então, λ é um autovalor de T se, e somente se, $p_T(\lambda) = 0$.*

Demonstração Suponha que $\dim(V) = n$ e fixe uma base ordenada $\mathcal{B}$ de V. Sabemos que λ é um autovalor de T se, e somente se, $V(\lambda) \neq \{0_V\}$. Mas, vimos na Proposição 7.1.2 que $V(\lambda) = \operatorname{Ker}(T - \lambda\,\mathrm{I}_V)$. Logo,

$$\begin{aligned} \lambda \text{ é um autovalor de } T \quad &\Longleftrightarrow \quad \operatorname{Ker}(T - \lambda\,\mathrm{I}_V) \neq \{0_V\} \\ &\Longleftrightarrow \quad T - \lambda\,\mathrm{I}_V \text{ não é injetor} \quad \text{(pela Proposição 6.2.5)} \\ &\Longleftrightarrow \quad T - \lambda\,\mathrm{I}_V \text{ não é bijetor} \quad \text{(pelo Corolário 6.3.3)} \\ &\Longleftrightarrow \quad [T]_{\mathcal{B}} - \lambda I_n = [T - \lambda\,\mathrm{I}_V]_{\mathcal{B}} \text{ não é inversível} \quad \text{(pela Proposição 6.7.1)} \\ &\Longleftrightarrow \quad p_T(\lambda) = \det\big([T]_{\mathcal{B}} - \lambda I_n\big) = 0 \quad \text{(pelo Teorema 1.3.6)}, \end{aligned}$$

demonstrando, assim, a equivalência no enunciado. □

Como um polinômio de grau n tem no máximo n raízes distintas, essa proposição estabelece um limite superior para o número de autovalores que um operador linear em um espaço vetorial de dimensão finita:

Corolário 7.2.3 *Seja T um operador linear em um espaço vetorial de dimensão finita igual a n. Então, T possui no máximo n autovalores distintos.*

Exemplo 7.2.4 Determine os autovalores e autovetores do operador linear T de $\mathbb{R}^2$ que satisfaz

$$[T]_{\mathsf{can}} = \begin{bmatrix} 8 & -10 \\ 3 & -3 \end{bmatrix},$$

em que can denota a base canônica de $\mathbb{R}^2$.

Solução: Vimos, na Proposição 7.2.2, que os autovalores de T são as raízes de

$$\begin{aligned} p_T(t) &= \det\big([T]_{\mathsf{can}} - tI_2\big) = \det\left(\begin{bmatrix} 8 & -10 \\ 3 & -3 \end{bmatrix} - t\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\begin{bmatrix} 8-t & -10 \\ 3 & -3-t \end{bmatrix} = (8-t)(-3-t)+30 \\ &= t^2 - 5t + 6 = (t-3)(t-2) \end{aligned}$$

Assim, T tem dois autovalores: 2 e 3.

Passamos, agora, à determinação dos autoespaços $V(2)$ e $V(3)$.

Sabemos que $V(2) = \operatorname{Ker}(T - 2\,\mathrm{I}_{\mathbb{R}^2})$. Logo, dado $v \in \mathbb{R}^2$, temos que $v \in V(2)$ se, e somente se, $(T - 2\,\mathrm{I}_{\mathbb{R}^2})(v) = (0,0)$. Tomando coordenadas em relação à base canônica, se $v = (x,y)$, então $v = (x,y)_{\mathsf{can}}$, e teremos que $v \in V(2)$ se, e somente se, $\big([T]_{\mathsf{can}} - 2I_2\big)[v]_{\mathsf{can}} = [(0,0)]_{\mathsf{can}}$, isto é, se, e somente se,

$$\begin{bmatrix} 6 & -10 \\ 3 & -5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{7.2}$$

já que $[T]_{\mathsf{can}} - 2I_2 = \begin{bmatrix} 8 & -10 \\ 3 & -3 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & -10 \\ 3 & -5 \end{bmatrix}$. As soluções de (7.2) são $\left(\frac{5y}{3}, y\right)$, $y \in \mathbb{R}$. Logo, $V(2) = \left[\left(\frac{5}{3}, 1\right)\right] = [(5,3)]$. Em particular, descobrimos que $\dim\big(V(2)\big) = 1$, pois $\{(5,3)\}$ é uma base de $V(2)$.

Por outro lado, $V(3) = \operatorname{Ker}(T - 3\,\mathrm{I}_{\mathbb{R}^2})$. Como $[T]_{\mathsf{can}} - 3I_2 = \begin{bmatrix} 8 & -10 \\ 3 & -3 \end{bmatrix} - 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & -10 \\ 3 & -6 \end{bmatrix}$, argumentando de maneira análoga, concluí-se que os vetores de $V(3)$ são as soluções de

$$\begin{bmatrix} 5 & -10 \\ 3 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{7.3}$$

que são $(2y, y)$, $y \in \mathbb{R}$. Logo, $V(3) = [(2,1)]$, e $\dim\big(V(3)\big) = 1$, já que $\{(2,1)\}$ é uma base de $V(3)$.

Cabe registrar que o conjunto $\mathcal{B} = \{(5,3), (2,1)\}$, formado pela união das bases de $V(2)$ e $V(3)$ que encontramos, é LI. Como $\mathcal{B}$ contém 2 vetores, segue que $\mathcal{B}$ é uma base de $\mathbb{R}^2$. E é uma base muito especial, ela é formada por autovetores de T: $(5,3)$ é um autovetor de T associado ao autovalor 2 e $(2,1)$ é um autovetor de T associado ao autovalor 3. Isso quer dizer que a matriz $[T]_{\mathcal{B}}$ é diagonal e contém os autovalores de T em sua diagonal principal:

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

Esse fenômeno — $\mathbb{R}^2$ ter uma base formada por autovetores de T — será mais detalhadamente investigado na próxima seção. $\diamond$

Exemplo 7.2.5 Determine os autovalores e autovetores do operador linear T de $\mathbb{R}^2$ que satisfaz $[T]_{\mathsf{can}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Solução: O polinômio característico de T é dado por

$$p_T(t) = \det\big([T]_{\mathsf{can}} - tI_2\big) = \det\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} - t\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\begin{bmatrix} -t & -1 \\ 1 & -t \end{bmatrix} = t^2 + 1$$

Como $p_T(t)$ não tem raízes reais, T não tem autovalores (e, portanto, não tem autovetores). $\diamond$

Exemplo 7.2.6 Determine os autovalores e autovetores do operador linear T de $\mathbb{R}^3$ que satisfaz $[T]_{\mathsf{can}} = \begin{bmatrix} -4 & 0 & -3 \\ 0 & 2 & 0 \\ 6 & 0 & 5 \end{bmatrix}$.

Solução: O polinômio característico de T é dado por

$$p_T(t) = \det\big([T]_{\mathsf{can}} - tI_3\big) = \det\left(\begin{bmatrix} -4 & 0 & -3 \\ 0 & 2 & 0 \\ 6 & 0 & 5 \end{bmatrix} - t\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = \det\begin{bmatrix} -4-t & 0 & -3 \\ 0 & 2-t & 0 \\ 6 & 0 & 5-t \end{bmatrix}.$$

Desenvolvendo esse determinante por expansão em cofatores ao longo da segunda linha, obtemos

$$p_T(t) = (2-t)\det\begin{bmatrix} -4-t & -3 \\ 6 & 5-t \end{bmatrix} = (2-t)\big((-4-t)(5-t)+18\big) = (2-t)(t^2-t-2) = -(t-2)^2(t+1).$$

Portanto, os autovalores de T são 2 e -1, as raízes de $p_T(t)$.

Sabemos que $V(2) = \operatorname{Ker}(T - 2\,\mathrm{I}_{\mathbb{R}3})$. Como $[T]_{\mathsf{can}} - 2I_3 = \begin{bmatrix} -6 & 0 & -3 \\ 0 & 0 & 0 \\ 6 & 0 & 3 \end{bmatrix}$, segue que $V(2)$ é formado pelas soluções de

$$\begin{bmatrix} -6 & 0 & -3 \\ 0 & 0 & 0 \\ 6 & 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

que são da forma

$$\left(-\frac{z}{2}, y, z\right) = y(0,1,0) + z\left(-\frac{1}{2}, 0, 1\right), \quad y, z \in \mathbb{R}.$$

Portanto, $V(2) = \big[(0,1,0), \big(-\frac{1}{2}, 0, 1\big)\big] = [(0,1,0), (-1,0,2)]$.

Já $V(-1) = \operatorname{Ker}\big(T - (-1)\,\mathrm{I}_{\mathbb{R}^3}\big) = \operatorname{Ker}(T + \mathrm{I}_{\mathbb{R}^3})$. Porque $[T]_{\mathsf{can}} + I_3 = \begin{bmatrix} -3 & 0 & -3 \\ 0 & 3 & 0 \\ 6 & 0 & 6 \end{bmatrix}$, o autoespaço $V(-1)$ é formado pelas soluções de

$$\begin{bmatrix} -3 & 0 & -3 \\ 0 & 3 & 0 \\ 6 & 0 & 6 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

e essas são da forma

$$(-z, 0, z) = z(-1, 0, 1), \quad z \in \mathbb{R}.$$

Logo, $V(-1) = [(-1,0,1)]$.

Aqui, também, se unirmos as bases $\{(0,1,0), (-1,0,2)\}$ de $V(2)$ e $\{(-1,0,1)\}$ de $V(-1)$, obtemos o conjunto $\mathcal{B} = \{(0,1,0), (-1,0,2), (-1,0,1)\}$, que é LI e contém 3 vetores. Logo, $\mathcal{B}$ é uma base de $\mathbb{R}^3$ e, como seus elementos são autovetores de T, obtemos

$$[T]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \quad \diamond$$

Exercícios Lista 2 - Álgebra Linear II: Exs. 56–58.

7.3 Diagonalização

Nesta seção, será estabelecida uma relação entre matrizes diagonalizáveis e operadores em espaços que possuem bases formadas por autovetores.

Definição Seja V um espaço vetorial de dimensão finita e seja $T\colon V \to V$ um operador linear de V. Dizemos que T é *diagonalizável* se existir uma base ordenada $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ seja uma matriz diagonal.

A seguinte caracterização de operadores diagonalizáveis é extremamente útil.

Proposição 7.3.1 *Seja V um espaço vetorial de dimensão finita e seja $T\colon V \to V$ um operador linear de V. Então, T é diagonalizável se, e somente se, V possuir uma base formada por autovetores de T.*

Demonstração Suponha que T seja diagonalizável. Tome uma base ordenada $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ de V tal que $[T]_{\mathcal{B}} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Então, pela própria definição de matriz de um operador, segue que, para todo $i = 1, \dots, n$, temos $T(v_i) = \lambda_i v_i$. Como os vetores de $\mathcal{B}$ são todos não nulos, eles são, todos, autovetores de T. (Precisamente, v_i é um autovetor de T associado ao autovalor λ_i).

Reciprocamente, suponha que V tenha uma base ordenada $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$, em que, para todo $i = 1, \dots, n$, v_i é um autovetor de T. Então, existem $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ (não necessariamente distintos) tais que $T(v_i) = \lambda_i v_i$, para todo $i = 1, \dots, n$. Logo, $[T]_{\mathcal{B}} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. $\square$

Os Exemplos 7.2.4 e 7.2.6, que vimos na seção anterior, são exemplos de operadores diagonalizáveis.

Exemplo 7.3.2 O operador linear S de $\mathbb{R}^2$ definido por $S(x, y) = (x + y, y)$, para todo $(x, y) \in \mathbb{R}^2$, não é diagonalizável. Vejamos por quê. Os autovalores de S são as raízes de seu polinômio característico; para encontrar o polinômio característico de S, é necessária a matriz de S em relação a alguma base de $\mathbb{R}^2$, por exemplo, can, a canônica. Como $S(1, 0) = (1, 0)$ e $S(0, 1) = (1, 1)$, segue que $[S]_{\mathsf{can}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Portanto, $p_S(t) = \det \begin{bmatrix} 1-t & 1 \\ 0 & 1-t \end{bmatrix} = (1-t)^2$. Assim, o único autovalor de S é 1. Os autovetores de S são todos associados ao autovalor 1 e coincidem com os vetores não nulos de $V(1) = \operatorname{Ker}(S - 1\,\mathrm{I}_{\mathbb{R}^2}) = \operatorname{Ker}(S - \mathrm{I}_{\mathbb{R}^2})$. Porque $[S]_{\mathsf{can}} - I_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, sabemos que $V(1)$ é formado pelas soluções de

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Portanto, $V(1) = [(1, 0)]$. Não existe base de $\mathbb{R}^2$ formada por autovetores de S, pois qualquer autovetor de S é múltiplo de $(1, 0)$, o que implica que um conjunto formado por dois autovetores de S será sempre LD. Logo, S não é diagonalizável. (Ou, colocando de modo equivalente, não existe base de $\mathbb{R}^2$ com a propriedade de a matriz de S em relação a ela ser diagonal.)

Operadores diagonalizáveis são especialmente bem comportados. Por exemplo, é simples encontrar suas iterações.

Exemplo 7.3.3 Considere o operador linear $T\colon \mathbb{R}^2 \to \mathbb{R}^2$, definido por $T(x, y) = (8x - 10y, 3x - 3y)$, para todo $(x, y) \in \mathbb{R}^2$. Então, $[T]_{\mathsf{can}} = \begin{bmatrix} 8 & -10 \\ 3 & -3 \end{bmatrix}$, em que can denota a base canônica de $\mathbb{R}^2$. Assim, T é exatamente o operador linear de $\mathbb{R}^2$ de que tratamos no Exemplo 7.2.4. Naquela ocasião, vimos que T é diagonalizável e que tomando a base $\mathcal{B} = \{(5, 3), (2, 1)\}$ de $\mathbb{R}^2$, obtemos

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

Vamos procurar uma expressão para T^n, a n-ésima potência de T, para um inteiro positivo n. (Como vimos na página 123, T^n é o operador de $\mathbb{R}^2$ que resulta da composição de T consigo mesmo $(n-1)$ vezes.)

Pelo Corolário 6.7.4,

$$P^{-1} [T]_{\mathsf{can}} P = [T]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix},$$

em que $P = [\mathrm{I}_{\mathbb{R}^2}]_{\mathcal{B},\mathsf{can}}$. Assim, como fizemos no início deste capítulo,

$$P^{-1}[T]_{\mathsf{can}}^n P = \left(P^{-1}[T]_{\mathsf{can}}P\right)^n = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}^n = \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix}. \tag{7.4}$$

A Proposição 6.6.2 nos diz que $[T]_{\mathsf{can}}^n = [T^n]_{\mathsf{can}}$. Utilizando esse fato com (7.4), obtemos

$$P^{-1}[T^n]_{\mathsf{can}}P = \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix},$$

que, multiplicada por P à esquerda e por P^{-1} à direita, resulta em

$$[T^n]_{\mathsf{can}} = P \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix} P^{-1}. \tag{7.5}$$

O cálculo de P é imediato (por can se tratar da base canônica de $\mathbb{R}^2$):

$$P = [\mathrm{I}_{\mathbb{R}^2}]_{\mathcal{B},\mathsf{can}} = \begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix}.$$

Portanto, obtemos de (7.5), que

$$[T^n]_{\mathsf{can}} = \begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 3 & -5 \end{bmatrix} = \begin{bmatrix} 2 \cdot 3^{n+1} - 5 \cdot 2^n & 10(2^n - 3^n) \\ 3^{n+1} - 3 \cdot 2^n & 3 \cdot 2^{n+1} - 5 \cdot 3^n \end{bmatrix}.$$

Dessa expressão, extrai-se que

$$T^n(1,0) = \left(2 \cdot 3^{n+1} - 5 \cdot 2^n,\ 3^{n+1} - 3 \cdot 2^n\right) \qquad \text{e} \qquad T^n(0,1) = \left(10(2^n - 3^n),\ 3 \cdot 2^{n+1} - 5 \cdot 3^n\right).$$

Portanto, para qualquer $(x, y) \in \mathbb{R}^2$, tem-se

$$\begin{aligned} T^n(x,y) &= T^n\big(x(1,0) + y(0,1)\big) = xT^n(1,0) + yT^n(0,1) \\ &= x\big(2 \cdot 3^{n+1} - 5 \cdot 2^n,\ 3^{n+1} - 3 \cdot 2^n\big) + y\big(10(2^n - 3^n),\ 3 \cdot 2^{n+1} - 5 \cdot 3^n\big) \\ &= \big((2 \cdot 3^{n+1} - 5 \cdot 2^n)x + 10(2^n - 3^n)y,\ (3^{n+1} - 3 \cdot 2^n)x + (3 \cdot 2^{n+1} - 5 \cdot 3^n)y\big). \end{aligned}$$

A expressão obtida para $T^n(x, y)$ é o que menos importa. Mais importante é perceber que se um operador é diagonalizável, podemos encontrar a imagem de um elemento de seu domínio por uma potência do operador, por maior que ela seja, sem precisar realizar o esforço de compor o operador com ele mesmo o número de vezes necessário. ◇

Observação Cabe, neste momento, a fim de relacionar o conceito de matriz diagonalizável, visto no início deste capítulo, com o de operador linear diagonalizável, retomar o Exemplo 6.2.7 no caso em que $m = n$. Vimos que cada matriz $A \in M_n(\mathbb{R})$ define um operador linear $T_A \colon \mathbb{R}^n \to \mathbb{R}^n$ dado por $T_A(v) = Av$, para todo $v \in \mathbb{R}^n$. Sabemos, também, que, se can denota a base canônica de $\mathbb{R}^n$, então $[T_A]_{\mathsf{can}} = A$ (ver Exemplo 6.5.5). O resultado é o que se espera: A é uma matriz diagonalizável se, e somente se, T_A é um operador diagonalizável. Vejamos por quê.

Suponha que A seja diagonalizável. De acordo com a definição na p. 139, existe uma matriz inversível $P \in M_n(\mathbb{R})$ e uma matriz diagonal $D \in M_n(\mathbb{R})$ tais que $P^{-1}AP = D$. Vamos mostrar que $P = (c_{ij})$ é uma matriz de mudança de base apropriada. Para tanto, para cada $j = 1, \dots, n$ seja $v_j \in \mathbb{R}^n$ tal que $[v_j]_{\mathsf{can}} = \begin{bmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{bmatrix}$ (isto é, as coordenadas de v_j como elemento de $\mathbb{R}^n$ são as entradas na j-ésima coluna de P). Mostremos que $\mathcal{B} = \{v_1, \dots, v_n\}$ é uma base ordenada de $\mathbb{R}^n$. Como esse conjunto ordenado contém n vetores, será suficiente verificar que ele é LI. Sejam $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tais que $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_{\mathbb{R}^n}$. Em coordenadas (em relação à base canônica) isso é equivalente a

$$P\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix} = \begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix}.$$

Como P é inversível, segue que $\alpha_1 = \cdots = \alpha_n = 0$. Assim $\mathcal{B}$ é, de fato, uma base ordenada de $\mathbb{R}^n$. Além disso, pela escolha que fizemos dos elementos de $\mathcal{B}$, temos $[\mathrm{I}_{\mathbb{R}^n}]_{\mathcal{B},\mathsf{can}} = P$. Daí segue, do Corolário 6.7.4, que

$$[T_A]_{\mathcal{B}} = P^{-1}[T_A]_{\mathsf{can}}P = P^{-1}AP = D,$$

que é diagonal. Portanto, T_A é diagonalizável (os elementos de $\mathcal{B}$ — lembre, as colunas de P — são autovetores de T_A e as entradas na diagonal principal de D são os autovalores correspondentes).

Reciprocamente, suponha, agora, que T_A seja diagonalizável e seja $\mathcal{B}$ uma base ordenada de $\mathbb{R}^n$ formada por autovetores de T_A. Então, denotando $P = [\mathrm{I}_{\mathbb{R}^n}]_{\mathcal{B},\mathsf{can}}$, segue, do Corolário 6.7.4, que

$$P^{-1}AP = P^{-1}[T_A]_{\mathsf{can}}P = [T_A]_{\mathcal{B}},$$

que é uma matriz diagonal (uma vez que os elementos de $\mathcal{B}$ são autovetores de T). Logo, A é diagonalizável.

Por causa dessa observação, dada uma matriz $A \in M_n(\mathbb{R})$, dizemos que um vetor não nulo $v \in \mathbb{R}^n$ é um *autovetor* de A se v for um autovetor do operador linear T_A, e dizemos que um escalar $\lambda \in \mathbb{R}$ é um *autovalor* de A se λ for um autovalor de T_A.

Podemos utilizar essa relação entre uma matriz $A \in M_n(\mathbb{R})$ e o operador linear T_A de $\mathbb{R}^n$ para calcular potências de A.

Exemplo 7.3.4 Dada $A = \begin{bmatrix}-5 & -1 & -2\\ 12 & 2 & 6\\ 8 & 2 & 3\end{bmatrix}$, encontre A^{500}.

Solução: Considere o operador linear $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ tal que $[T]_{\mathsf{can}} = A$, em que can denota a base canônica de $\mathbb{R}^3$ (ou seja, $T = T_A$). Vejamos se T é diagonalizável. Temos

$$p_T(t) = \det\begin{bmatrix}-5-t & -1 & -2\\ 12 & 2-t & 6\\ 8 & 2 & 3-t\end{bmatrix} = -t^3 + 3t + 2 = -(t+1)^2(t-2).$$

Assim, os autovalores de T são -1 e 2. Vejamos se é possível encontrar uma base de $\mathbb{R}^3$ formada por autovetores de T.

Os autovetores de T associados ao autovalor -1 são os vetores não nulos em $V(-1) = \operatorname{Ker}\big(T - (-1)\,\mathrm{I}_{\mathbb{R}^3}\big) = \operatorname{Ker}(T + \mathrm{I}_{\mathbb{R}^3})$. Dado $v = (x, y, z) \in \mathbb{R}^3$, temos que $v \in V(-1)$ se, e somente se,

$$\begin{bmatrix}-4 & -1 & -2\\ 12 & 3 & 6\\ 8 & 2 & 4\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix},$$

ou seja, se e somente se, para quaisquer escolhas de $y, z \in \mathbb{R}$, tivermos $x = -\frac{1}{4}y - \frac{1}{2}z$. Logo, os vetores em $V(-1)$ são da forma

$$\left(-\frac{1}{4}y - \frac{1}{2}z, y, z\right) = y\left(-\frac{1}{4}, 1, 0\right) + z\left(-\frac{1}{2}, 0, 1\right),$$

com $y, z \in \mathbb{R}$. Portanto, $\{(-1, 4, 0), (-1, 0, 2)\}$ é uma base para $V(-1)$.

Os elementos do autoespaço $V(2) = \operatorname{Ker}(T - 2\,\mathrm{I}_{\mathbb{R}^3})$ são os vetores de $\mathbb{R}^3$ cujas coordenadas são soluções do sistema linear homogêneo

$$\begin{bmatrix}-7 & -1 & -2\\ 12 & 0 & 6\\ 8 & 2 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix}.$$

Resolvendo o sistema, obtemos $V(2) = \left\{\left(-\frac{1}{2}z, \frac{3}{2}, z\right) \mid z \in \mathbb{R}\right\}$, donde segue que $V(2) = [(-1, 3, 2)]$.

Observe, agora, que $\mathcal{B} = \{(-1, 4, 0), (-1, 0, 2), (-1, 3, 2)\}$ é LI (aqui é preciso fazer as contas para se convencer desse fato), e, portanto, uma base de $\mathbb{R}^3$, que é formada por autovetores de T, os dois primeiros associados a -1 e o terceiro, a 2. Assim,

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

Denotando $P = [\mathrm{I}_{\mathbb{R}^3}]_{\mathcal{B},\mathsf{can}}$, segue, do Corolário 6.7.4, que

$$A = [T]_{\mathsf{can}} = P[T]_{\mathcal{B}}P^{-1},$$

donde obtemos

$$A^{500} = \left(P[T]_{\mathcal{B}}P^{-1}\right)^{500} = P[T]_{\mathcal{B}}^{500}P^{-1}. \tag{7.6}$$

É fácil ver que $P = [\mathrm{I}_{\mathbb{R}^3}]_{\mathcal{B},\mathsf{can}} = \begin{bmatrix} -1 & -1 & -1 \\ 4 & 0 & 3 \\ 0 & 2 & 2 \end{bmatrix}$. Invertendo-a, obtemos

$$P^{-1} = \begin{bmatrix} -1 & -1 & -1 \\ 4 & 0 & 3 \\ 0 & 2 & 2 \end{bmatrix}^{-1} = \frac{1}{6}\begin{bmatrix} -6 & 0 & -3 \\ -8 & -2 & -1 \\ 8 & 2 & 4 \end{bmatrix}.$$

Substituindo os valores encontrados em (7.6), obtemos

$$\begin{aligned}
A^{500} &= \begin{bmatrix} -1 & -1 & -1 \\ 4 & 0 & 3 \\ 0 & 2 & 2 \end{bmatrix}\left(\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}\right)^{500}\left(\frac{1}{6}\begin{bmatrix} -6 & 0 & -3 \\ -8 & -2 & -1 \\ 8 & 2 & 4 \end{bmatrix}\right) \\
&= \frac{1}{6}\begin{bmatrix} -1 & -1 & -1 \\ 4 & 0 & 3 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} (-1)^{500} & 0 & 0 \\ 0 & (-1)^{500} & 0 \\ 0 & 0 & 2^{500} \end{bmatrix}\begin{bmatrix} -6 & 0 & -3 \\ -8 & -2 & -1 \\ 8 & 2 & 4 \end{bmatrix} \\
&= \frac{1}{6}\begin{bmatrix} -1 & -1 & -1 \\ 4 & 0 & 3 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^{500} \end{bmatrix}\begin{bmatrix} -6 & 0 & -3 \\ -8 & -2 & -1 \\ 8 & 2 & 4 \end{bmatrix} \\
&= \frac{1}{3}\begin{bmatrix} 7 - 2^{502} & 1 - 2^{500} & 2 - 2^{501} \\ -12 + 3 \cdot 2^{502} & 3 \cdot 2^{500} & -6 + 3 \cdot 2^{501} \\ -8 + 2^{503} & -2 + 2^{501} & -1 + 2^{502} \end{bmatrix} \qquad \diamond
\end{aligned}$$

Como último exemplo nesta seção, vejamos como dada uma matriz diagonaliável encontrar informações sobre o operador linear definido por ela.

Exemplo 7.3.5 (Prova 2, Álgebra Linear II, 2019) Seja $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ o operador linear tal que

$$M^{-1}[T]_{\mathsf{can}}M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

em que $M = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ e can denota a base canônica de $\mathbb{R}^3$. Então, $T(5, 3, 4)$ é igual a

(A) $(5, 11, 13)$ (B) $(8, 5, 9)$ (C) $(10, 6, 8)$ (D) $(9, 8, 5)$ (E) $(7, 7, 6)$

Solução: A igualdade no enunciado nos mostra que T é diagonalizável, com autovalores $1, 2, 3$ e que $\mathcal{B} = \{(1,0,1), (0,1,1), (1,1,0)\}$ (o conjunto formado pelas colunas de M) é uma base ordenada de $\mathbb{R}^3$ formada por autovetores de T associados, respectivamente, a $1, 2, 3$. (Isso, pois $M = [\mathrm{I}_{\mathbb{R}^3}]_{\mathcal{B},\mathsf{can}}$ e, assim, usando o Corolário 6.7.4, $[T]_{\mathcal{B}} = M^{-1}[T]_{\mathsf{can}}M$.) Portanto,

$$\begin{aligned} T(1,0,1) &= 1(1,0,1) = (1,0,1) \\ T(0,1,1) &= 2(0,1,1) = (0,2,2) \\ T(1,1,0) &= 3(1,1,0) = (3,3,0). \end{aligned}$$

Temos informação suficiente sobre T (imagens dos vetores em uma base do domínio) para determinar a imagem por T de qualquer vetor de $\mathbb{R}^3$. Em particular, como[1]

$$(5,3,4) = 3(1,0,1) + 1(0,1,1) + 2(1,1,0),$$

segue que

$$T(5,3,4) = 3T(1,0,1) + 1T(0,1,1) + 2T(1,1,0) = 3(1,0,1) + 1(0,2,2) + 2(3,3,0) = (9,8,5).$$

Resposta: (D). $\diamond$

Nosso objetivo, na próxima seção, será sistematizar as ideias exploradas nesta seção e determinar condições necessárias e suficientes para um operador linear em um espaço vetorial de dimensão finita ser diagonalizável.

Exercícios Lista 2 - Álgebra Linear II: Exs. 59–63.

7.4 Operadores diagonalizáveis

Esta seção será dedicada ao Teorema 7.4.3, que caracteriza os operadores diagonalizáveis de um espaço vetorial de dimensão finita e suas consequências.

É preciso começar com dois resultados preliminares.

Lema 7.4.1 *Seja T um operador linear em um espaço vetorial V e sejam $\lambda_1, \lambda_2, \dots, \lambda_k$ autovalores de T dois a dois distintos entre si. Se $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$ são conjuntos LI finitos satisfazendo $\mathcal{B}_i \subseteq V(\lambda_i)$, para todo $i = 1, \dots, k$, então a união $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$ é LI.*

Demonstração Demonstremos o lema no caso em que $k = 2$. Sejam, então, $\lambda, \mu \in \mathbb{R}$ autovalores de T, com $\lambda \neq \mu$ e sejam $\mathcal{B} = \{u_1, u_2, \dots, u_p\}$ um conjunto LI de autovetores de T associados a λ e $\mathcal{C} = \{v_1, v_2, \dots, v_q\}$ um conjunto LI de autovetores de T associados a μ. Para mostrar que $\mathcal{B} \cup \mathcal{C} = \{u_1, u_2, \dots, u_p, v_1, v_2, \dots, v_q\}$, é LI, precisamos tomar uma combinação linear dos elementos desse conjunto que resulta no vetor nulo de V e mostrar que todos os escalares na combinação linear são nulos. Assim, sejam $\alpha_1, \alpha_2, \dots, \alpha_p, \beta_1, \beta_2, \dots, \beta_q \in \mathbb{R}$ tais que

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_p u_p + \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_q v_q = 0_V. \tag{7.7}$$

Aplicando o operador T em ambos os lados de (7.7), obtemos

$$\alpha_1 T(u_1) + \alpha_2 T(u_2) + \cdots + \alpha_p T(u_p) + \beta_1 T(v_1) + \beta_2 T(v_2) + \cdots + \beta_q T(v_q) = T(0_V) = 0_V. \tag{7.8}$$

[1] Você deve ser capaz de encontrar as coordenadas, em relação à base $\mathcal{B}$ de um vetor arbitrário:

$$(x,y,z) = \frac{x-y+z}{2}(1,0,1) + \frac{-x+y+z}{2}(0,1,1) + \frac{x+y-z}{2}(1,1,0).$$

Como os u_i são autovetores de T associados a λ, temos $T(u_i) = \lambda u_i$, para cada $i = 1, \dots, k$. De maneira análoga, $T(v_j) = \mu v_j$, para cada $j = 1, \dots, q$. Assim, (7.8) implica

$$\alpha_1\lambda u_1 + \alpha_2\lambda u_2 + \cdots + \alpha_p\lambda u_p + \beta_1\mu v_1 + \beta_2\mu v_2 + \cdots + \beta_q\mu v_q = 0_V. \tag{7.9}$$

Por outro, lado, multiplicando os dois lados de (7.7) por λ resulta em

$$\alpha_1\lambda u_1 + \alpha_2\lambda u_2 + \cdots + \alpha_p\lambda u_p + \beta_1\lambda v_1 + \beta_2\lambda v_2 + \cdots + \beta_q\lambda v_q = \lambda 0_V = 0_V. \tag{7.10}$$

Subtraindo (7.9) de (7.10), obtemos

$$(\lambda - \mu)\beta_1 v_1 + (\lambda - \mu)\beta_2 v_2 + \cdots + (\lambda - \mu)\beta_q v_q = 0_V.$$

Mas o conjunto $\mathcal{C}$ é, por hipótese, LI. Assim, necessariamente, $(\lambda - \mu)\beta_i = 0$, para todo $i = 1, \dots, q$. Como $\lambda \neq \mu$, segue que $\beta_1 = \beta_2 = \cdots = \beta_q = 0$. Substituindo esses valores em (7.7), obtemos

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_p u_p = 0_V,$$

que, por sua vez, implica $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$, já que $\mathcal{B}$ também é LI. Isso prova que $\mathcal{B} \cup \mathcal{C}$ é LI.

O passo seguinte é mostrar que o resultado vale para $k = 3$. Seremos um pouco mais econômicos nos detalhes. Sejam λ, μ, ρ autovalores distintos de T e sejam $\mathcal{B} = \{u_1, \dots, u_p\} \subseteq V(\lambda)$, $\mathcal{C} = \{v_1, \dots, v_p\} \subseteq V(\mu)$ e $\mathcal{D} = \{w_1, \dots, w_t\} \subseteq V(\rho)$ conjuntos LI. Dados $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q, \gamma_1, \dots, \gamma_t \in \mathbb{R}$ tais que

$$\alpha_1\, u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_t w_t = 0_V, \tag{7.11}$$

aplicando T a (7.11), por um lado, e multiplicando (7.11) por λ, por outro, após subtrair uma expressão da outra, obteremos

$$(\lambda - \mu)\beta_1 v_1 + \dots (\lambda - \mu)\beta_q v_q + (\lambda - \rho)\gamma_1 w_1 + \dots (\lambda - \rho)\gamma_t w_t = 0_V,$$

que é uma combinação linear dos elementos de $\mathcal{C} \cup \mathcal{D}$ resultando no vetor nulo de V. Mas, vimos que o resultado é válido para $k = 2$. Assim, necessariamente, essa combinação linear deve ser trivial, isto é, $(\lambda - \mu)\beta_i = 0$, para todo $i = 1, \dots, q$, e $(\lambda - \rho)\gamma_j = 0$, para todo $j = 1, \dots, t$. Como $\lambda \neq \mu$ e $\lambda \neq \rho$, segue que $\beta_1 = \cdots = \beta_q = \gamma_1 = \cdots = \gamma_t = 0$, o que substituído em (7.11) implica que $\alpha_1 = \cdots = \alpha_p = 0$. Logo, $\mathcal{B} \cup \mathcal{C} \cup \mathcal{D}$ é LI.

Como vimos acima, a validade do resultado para $k = 2$ implica sua validade para $k = 3$. De modo análogo, podemos demonstrar que o caso $k = 4$ é consequência do caso $k = 3$, já demonstrado, e assim por diante, obtendo a validade do resultado para qualquer valor inteiro positivo de k.[2] □

O próximo lema, além de ser utilizado na demonstração do Teorema 7.4.3, tem interesse em si próprio e será utilizado diversas vezes nas análises de operadores que veremos nos exercícios e aplicações.

Como vimos na Seção 7.2, os autovalores de um operador linear T em um espaço vetorial de dimensão finita são precisamente as raízes do polinômio característico $p_T(t)$ de T. Assim, se $\lambda \in \mathbb{R}$ é um autovalor de T, então $p_T(t)$ é divisível por $t - \lambda$. Chamamos de *multiplicidade* de λ como raiz de $p_T(t)$ o maior inteiro positivo r tal que $p_T(t)$ seja divisível por $(t - \lambda)^r$. (É frequente chamar r também de *multiplicidade algébrica* de λ.) O resultado a seguir estabelece uma relação entre a multiplicidade de um autovalor como raiz do polinômio característico e a dimensão do autoespaço associado.

Lema 7.4.2 *Seja T um operador linear em um espaço vetorial V de dimensão finita e seja $\lambda \in \mathbb{R}$ um autovalor de T de multiplicidade r como raiz do polinômio característico de T. Então,* $1 \leq \dim\big(V(\lambda)\big) \leq r$.

Demonstração Como V tem dimensão finita, $V(\lambda)$ tem, também, dimensão finita. Que $\dim\big(V(\lambda)\big) \geq 1$ segue do fato de λ ser um autovalor de T. Demonstremos a segunda desigualdade. Tome uma base $\{v_1, \dots, v_s\}$ de $V(\lambda)$ (estamos, assim, denotando $s = \dim\big(V(\lambda)\big)$) e estenda-a a uma base ordenada $\mathcal{B} = \{v_1, \dots, v_s, v_{s+1}, \dots, v_n\}$ de V (o que é sempre possível, pelo Teorema 4.4.14). Então, a matriz $[T]_{\mathcal{B}}$ tem a forma

[2] Esse é um exemplo de argumento em que se utiliza o chamado *princípio de indução finita*.

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda I_s & A \\ 0 & B \end{bmatrix},$$

em que $A \in M_{s\times(n-s)}(\mathbb{R})$, $B \in M_{n-s}(\mathbb{R})$ e 0 denota a matriz nula de tamanho $(n-s)\times s$. Logo,

$$p_T(t) = \det\big([T]_{\mathcal{B}} - tI_n\big) = \det \begin{bmatrix} (\lambda - t)I_n & A \\ 0 & B - tI_{n-s} \end{bmatrix} = (\lambda - t)^s \det(B - tI_{n-s}). \tag{7.12}$$

A última igualdade pode ser obtida, por exemplo, por expansão em cofatores, sucessivamente, ao longo da primeira coluna. Podemos reescrever a igualdade obtida em (7.12) da seguinte maneira: $p_T(t) = (t-\lambda)^r q(t)$, com $q(t) = (-1)^s p_B(t)$. Assim, há, pelo menos, s fatores $t - \lambda$ em $p_T(t)$. Isso quer dizer que $r \geq s = \dim\big(V(\lambda)\big)$. □

Para um autovalor λ de um operador linear T em um espaço vetorial de dimensão finita, costuma-se chamar o número $\dim\big(V(\lambda)\big)$ de *multiplicidade geométrica* de λ. Usando essa linguagem, o Lema 7.4.2 diz que para qualquer autovalor de T, sua multiplicidade geométrica está limitada superiormente por sua multiplicidade algébrica.

Um operador pode não ser diagonalizável por não ter autovalores suficientes (isso é o que ocorreu no Exemplo 7.2.5), ou por não ter autovetores suficientes (como no Exemplo 7.3.2). Esses são, essencialmente, os dois únicos possíveis obstáculos que podem impedir um operador de ser diagonalizável, como veremos no próximo resultado (que já estamos prontos para demonstrar), em que apresentamos um critério necessário e suficiente para um operador linear em um espaço vetorial de dimensão finita ser diagonalizável.

Teorema 7.4.3 *Seja T um operador linear em um espaço vetorial V de dimensão finita. Então, T é diagonalizálvel se, e somente se,*

(i) todas as raízes do polinômio característico $p_T(t)$ de T forem reais, e
(ii) para cada autovalor λ de T, sua multiplicidade como raiz de $p_T(t)$ for igual a $\dim\big(V(\lambda)\big)$.

Demonstração Suponha, inicialmente, que T seja diagonalizável e que $\lambda_1, \lambda_2 \dots, \lambda_k$ seja a lista completa de seus autovalores distintos. Seja $\mathcal{B}$ uma base de V formada por autovetores de T que foi ordenada de modo que os r_1 primeiros elementos de $\mathcal{B}$ sejam autovetores associados a λ_1, que os r_2 seguintes elementos de $\mathcal{B}$ sejam associados a λ_2 e, assim por diante, até os r_k últimos elementos de $\mathcal{B}$, que são autovetores de T associados a λ_k. Assim,

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 I_{r_1} & 0 & \dots & 0 \\ 0 & \lambda_2 I_{r_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k I_{r_k} \end{bmatrix},$$

em que os 0s são blocos nulos de tamanhos apropriados. Como essa matriz é diagonal, segue que

$$p_T(t) = (\lambda_1 - t)^{r_1}(\lambda_2 - t)^{r_2}\dots(\lambda_k - t)^{r_k}.$$

Logo, as raízes de $p_T(t)$ são $\lambda_1, \lambda_2 \dots, \lambda_k$ (de multiplicidades $r_1, r_2, \dots, r_k$, respectivamente), que são, todas, números reais. Resta-nos, portanto, mostrar que para cada $i = 1, \dots, k$, a multiplicidade algébrica de λ_i coincide com sua multiplicidade geométrica. Pelo Lema 7.4.2, já temos uma desigualdade: $\dim\big(V(\lambda_i)\big) \leq r_i$. Mas, o conjunto de r_i autovetores de T associados a λ_i que estão em $\mathcal{B}$ formam um subconjunto LI do espaço vetorial $V(\lambda_i)$. Portanto, $r_i \leq \dim\big(V(\lambda_i)\big)$. (Isso é consequência do Teorema 4.4.10: o tamanho dos subconjuntos LI está limitado superiormente pelo tamanho dos subconjuntos geradores, em particular, está limitado superiormente pela dimensão do espaço vetorial em que se encontram.) Logo, $r_i = \dim\big(V(\lambda_i)\big)$.

Reciprocamente, suponha, agora, que valem (i) e (ii) no enunciado do teorema e denote $n = \dim(V)$. Como todas as raízes de $p_T(t)$ são reais, esse polinômio se fatora completamente:

$$p_T(t) = (\lambda_1 - t)^{r_1}(\lambda_2 - t)^{r_2}\dots(\lambda_k - t)^{r_k}, \tag{7.13}$$

com $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}$ distintos. Assim, os autovalores de T são $\lambda_1, \lambda_2, \dots, \lambda_k$. Para cada $i = 1, \dots, k$, seja $\mathcal{B}_i$ uma base de $V(\lambda_i)$. O Lema 7.4.1 garante que $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_k$ é LI. Além disso, esse conjunto contém $r_1 + r_2 + \dots + r_k$

elementos, pela condição (ii). De (7.13), segue que essa soma coincide com o grau de $p_T(t)$, que, como sabemos, é igual a n. Em resumo, $\mathcal{B}$ é um conjunto LI contendo n elementos. Logo, $\mathcal{B}$ é uma base de V. Como seus elementos são todos autovetores de T, concluímos que T é diagonalizável. □

Veremos uma sequência de exemplos em que as condições do teorema serão exploradas. Porém, uma observação bastante útil a respeito da aplicação do teorema deve ser feita antes.

Observação Uma vez dado um operador linear T em um espaço vetorial de dimensão finita, para decidirmos se ele é ou não diagonalizável, precisamos, primeiramente verificar se todas as raízes de seu polinômio característico $p_T(t)$ são reais. Se $p_T(t)$ tiver uma raiz complexa não real, já temos nossa resposta: T não é diagonalizável.

Suponha, agora, que todas as raízes de $p_T(t)$ sejam números reais. Ainda assim, é possível que T não seja diagonalizável, basta que, para um de seus autovalores, a multiplicidade geométrica seja diferente da algébrica. Assim, para mostrar que T é diagonalizável (caso ele, de fato, o seja), é preciso verificar, para cada um dos autovalores de T se as multiplicidades algébricas e geométricas coincidem. As multiplicidades algébricas dos autovalores se obtêm de uma fatoração completa do polinômio característico como produto de termos lineares. Já as multiplicidades geométricas são obtidas por meio do cálculo das dimensões dos autoespaços associados.

O que se quer destacar aqui é que se λ é um autovalor de T de multiplicidade algébrica igual a 1, não há cálculo necessário a ser feito. Do Lema 7.4.2 segue que, neste caso, $1 \leq \dim\big(V(\lambda)\big) \leq 1$ e, portanto, $\dim\big(V(\lambda)\big) = 1$. Ou seja, a igualdade desejada entre as multiplicidades do autovalor λ está garantida.

Assim, um operador linear cujos autovalores são todos reais só deixa eventualmente de ser diagonalizável se algum autovalor de multiplicidade algébrica *maior* do que 1 tiver multiplicidade geométrica menor do que ela.

Exemplo 7.4.4 Decida se é ou não diagonalizável o operador linear $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ tal que $[T]_{\mathsf{can}} = \begin{bmatrix} 0 & 3 & 3 \\ -1 & -4 & -1 \\ 2 & 2 & -1 \end{bmatrix}$.

Solução: Começamos por determinar os autovalores de T:

$$p_T(t) = \det \begin{bmatrix} -t & 3 & 3 \\ -1 & -4-t & -1 \\ 2 & 2 & -1-t \end{bmatrix} = -t^3 - 5t^2 - 3t + 9 = -(t-1)(t+3)^2.$$

(As raízes de $p_T(t)$ foram encontradas por inspeção: sabemos que se $p_T(t)$ tiver raízes racionais, elas pertencerão ao conjunto $\{\pm 1, \pm 3, \pm 9\}$. Veja o Apêndice C para alguns fatos sobre polinômios e suas raízes.) Assim, os autovalores de T são 1, de multiplicidade 1, e -3, de multiplicidade 2. Como vimos na observação que precede este exemplo, sabemos que $\dim\big(V(1)\big) = 1$. Resta verificar a condição (ii) no Teorema 7.4.3 para o autovalor -3. Temos que $V(-3) = \operatorname{Ker}(T + 3\,\mathrm{I}_{\mathbb{R}^3})$. Para calcular a dimensão desse autoespaço podemos proceder de maneira indireta, por meio do Teorema do núcleo e da imagem (Teorema 6.3.1). Sabemos, pela Observação na p. 129, que a imagem de $T + 3\,\mathrm{I}_{\mathbb{R}^3}$ é gerada pelas colunas de

$$[T + 3\,\mathrm{I}_{\mathbb{R}^3}]_{\mathsf{can}} = [T]_{\mathsf{can}} + 3[\mathrm{I}_{\mathbb{R}^3}]_{\mathsf{can}} = [T]_{\mathsf{can}} + 3I_3 = \begin{bmatrix} 3 & 3 & 3 \\ -1 & -1 & -1 \\ 2 & 2 & 2 \end{bmatrix}.$$

Isto é, $\operatorname{Im}(T + 3\,\mathrm{I}_{\mathbb{R}^3}) = [\,(3,-1,2)\,]$, que, obviamente, tem dimensão igual a 1. Segue que $\dim\big(V(-3)\big) = \dim\big(\operatorname{Ker}(T + 3\,\mathrm{I}_{\mathbb{R}^3})\big) = \dim(\mathbb{R}^3) - \dim\big(\operatorname{Im}(T + 3\,\mathrm{I}_{\mathbb{R}^3})\big) = 3 - 1 = 2$, coincidindo, portanto, com a multiplicidade algébrica de -3. Logo, pelo Teorema 7.4.3, T é um operador diagonalizável.

Conseguimos dar uma resposta para o que se pediu sem precisar exibir uma base de $\mathbb{R}^3$ formada por autovetores de T. Mas sabemos que uma tal base existe. Vejamos, agora, como encontrá-la. Sabemos que se encontramos uma base para $V(1)$ (que será formada por um único vetor, uma vez que $\dim\big(V(1)\big) = 1$) e uma base para $V(-3)$ (que terá dois vetores), a união dessas duas bases será um conjunto LI (pelo Lema 7.4.1) de 3 vetores em $\mathbb{R}^3$, portanto, uma base de $\mathbb{R}^3$, que é formada por autovetores de T. Tratemos de um autoespaço por vez.

Como $V(1) = \operatorname{Ker}(T - \mathrm{I}_{\mathbb{R}^3})$, os elementos de $V(1)$ são os vetores $v \in \mathbb{R}^3$ tais que $[T - \mathrm{I}_{\mathbb{R}^3}]_{\mathsf{can}}[v]_{\mathsf{can}} = [(T - \mathrm{I}_{\mathbb{R}^3})(v)]_{\mathsf{can}} = [0_{\mathbb{R}^3}]_{\mathsf{can}}$. Porque

$$[T - \mathrm{I}_{\mathbb{R}^3}]_{\mathsf{can}} = [T]_{\mathsf{can}} - [\mathrm{I}_{\mathbb{R}^3}]_{\mathsf{can}} = [T]_{\mathsf{can}} - I_3 = \begin{bmatrix} -1 & 3 & 3 \\ -1 & -5 & -1 \\ 2 & 2 & -2 \end{bmatrix},$$

segue que, dado $v = (x, y, z) \in \mathbb{R}^3$, temos $v \in V(1)$ se, e somente se,

$$\begin{bmatrix} -1 & 3 & 3 \\ -1 & -5 & -1 \\ 2 & 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{7.14}$$

Escalonando a matriz de coeficientes desse sistema, obtemos

$$\begin{bmatrix} -1 & 3 & 3 \\ -1 & -5 & -1 \\ 2 & 2 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & 3 & 3 \\ 0 & -8 & -4 \\ 0 & 8 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & 3 & 3 \\ 0 & -8 & -4 \\ 0 & 0 & 0 \end{bmatrix}.$$

(Essa matriz tem dois pivôs, o que implica que o subespaço de $\mathbb{R}^3$ formado pelas soluções do sistema (7.14) tem dimensão $3 - 2 = 1$, o que não se trata de novidade, uma vez que esse subespaço é justamente $V(1)$, cuja dimensão já sabíamos ser igual a 1.) Assim, z é a única variável livre e, por retrossubstituição, obtemos $y = -\frac{z}{2}$ e $x = \frac{3z}{2}$. Logo, os vetores em $V(1)$ são da forma

$$\left(\frac{3z}{2}, -\frac{z}{2}, z\right) = z\left(\frac{3}{2}, -\frac{1}{2}, 1\right),$$

com $z \in \mathbb{R}$. Portanto, $V(1) = \left[\left(\frac{3}{2}, -\frac{1}{2}, 1\right)\right] = [(3, -1, 2)]$.

Procedamos à determinação de uma base de $V(-3)$. Por definição, $V(-3) = \mathrm{Ker}(T - 3\,\mathrm{I}_{\mathbb{R}^3})$. Logo, dado $v = (x, y, z) \in \mathbb{R}^3$, temos que $v \in V(-3)$ se, e somente se,

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

O escalonamento é imediato; obtemos que y e z são variáveis livres e $x = -y - z$. Assim, os elementos de $V(-3)$ são da forma

$$(-y - z, y, z) = y(-1, 1, 0) + z(-1, 0, 1), \quad z \in \mathbb{R}.$$

Portanto, $V(-3) = [(-1, 1, 0), (-1, 0, 1)]$.

O conjunto $\mathcal{B} = \{(3, -1, 2), (-1, 1, 0), (-1, 0, 1)\}$ é uma base de $\mathbb{R}^3$ formada por autovetores de T. Em particular,

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

Finalmente, podemos utilizar o Corolário 6.7.4 para obter $P^{-1}[T]_{\mathsf{can}}P = [T]_{\mathcal{B}}$, em que $P = [\mathrm{I}_{\mathbb{R}^3}]_{\mathcal{B},\mathsf{can}} = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$, ou, explicitamente,

$$\begin{bmatrix} 3 & -1 & -1 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 3 & 3 \\ -1 & -4 & -1 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & -1 & -1 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}. \quad \diamond$$

Exemplo 7.4.5 Decida se é ou não diagonalizável o operador linear $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ tal que $[T]_{\mathsf{can}} = \begin{bmatrix} -8 & -5 & 1 \\ 13 & 8 & -2 \\ -5 & -3 & 1 \end{bmatrix}$.

Solução: O polinômio característico de T é

$$p_T(t) = \det \begin{bmatrix} -8-t & -5 & 1 \\ 13 & 8-t & -2 \\ -5 & -3 & 1-t \end{bmatrix} = -t^3 + t^2 = -t^2(t-1).$$

Os autovalores de T são 0 (de multiplicidade 2) e 1 (de multiplicidade 1). Como sabemos que $\dim\bigl(V(1)\bigr) = 1$, T só poderá deixar de ser diagonalizável caso $\dim\bigl(V(0)\bigr) < 2$. Dado $v = (x, y, z) \in \mathbb{R}^3$, temos que $v \in V(0)$ se, e somente se,

$$\begin{bmatrix} -8 & -5 & 1 \\ 13 & 8 & -2 \\ -5 & -3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Fazendo operações elementares sobre as linhas da matriz de coeficientes desse sistema, obtemos

$$\begin{bmatrix} -8 & -5 & 1 \\ 13 & 8 & -2 \\ -5 & -3 & 1 \end{bmatrix} \to \begin{bmatrix} -8 & -5 & 1 \\ 13 & 8 & -2 \\ -13 & -8 & 2 \end{bmatrix} \to \begin{bmatrix} -8 & -5 & 1 \\ 13 & 8 & -2 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} -8 & -5 & 1 \\ -3 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} -8 & -5 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix},$$

que é escalonada. Segue, assim, que $y = -3z$ e $x = 2z$. Assim, $V(0) = [(2, -3, 1)]$, que tem dimensão 1. Logo, T não é diagonalizável. (Isto é, não existe uma base de $\mathbb{R}^3$ formada por autovetores de T.) $\diamond$

Exemplo 7.4.6 Decida se é ou não diagonalizável o operador linear $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ tal que $[T]_{\mathsf{can}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ -2 & -1 & 3 \end{bmatrix}$.

Solução: O polinômio característico de T é

$$p_T(t) = \det \begin{bmatrix} 1-t & 1 & 1 \\ 0 & 2-t & 0 \\ -2 & -1 & 3-t \end{bmatrix} = -(t-2)(t^2 - 4t + 5).$$

Como as raízes do fator $t^2 - 4t + 5$ de $p_T(t)$ não são reais (uma vez que tem discriminante $(-4)^2 - 4 \cdot 1 \cdot 5 = -4$), as raízes de $p_T(t)$ não são todas reais. Logo, T não é diagonalizável. $\diamond$

Algumas consequências imediatas do Teorema 7.4.3 são as seguintes.

Corolário 7.4.7 *Seja T um operador linear em um espaço vetorial V de dimensão finita igual a n. Se tiver n autovalores distintos, T será diagonaliável.*

Demonstração Se $\lambda_1, \lambda_2, \dots, \lambda_n$ são n autovalores distintos de T, então $p_T(t)$ será divisível por $(t-\lambda_1)(t-\lambda_2)\dots(t-\lambda_n)$. Uma comparação de graus implica que

$$p_T(t) = (-1)^n(t-\lambda_1)(t-\lambda_2)\dots(t-\lambda_n).$$

Ou seja as raízes de $p_T(t)$ são precisamente $\lambda_1, \dots, \lambda_n$, todas reais. Além disso, cada uma delas tem multiplicidade igual a 1 e, portanto, pelo Lema 7.4.2, $\dim\bigl(V(\lambda_i)\bigr) = 1$, para todo $i = 1, \dots, n$. Logo, T é diagonalizável. $\square$

É preciso alertar que a recíproca desse resultado está longe de ser verdadeira: há diversos operadores diagonalizáveis com menos autovalores do que a dimensão do espaço em que estão definidos, como ocorre, a título de ilustração, no Exemplo 7.4.4.

Corolário 7.4.8 *Seja T um operador linear em um espaço vetorial V de dimensão finita cujos autovalores são $\lambda_1, \lambda_2, \dots, \lambda_k$. Então, T é diagonalizálvel se, e somente se,*

$$\dim\big(V(\lambda_1)\big) + \dim\big(V(\lambda_2)\big) + \cdots + \dim\big(V(\lambda_k)\big) = \dim(V).$$

Demonstração Seja $n = \dim(V)$. Se T é diagonalizável, então $p_T(t) = (-1)^n(t-\lambda_1)^{r_1}(t-\lambda_2)^{r_2}\dots(t-\lambda_k)^{r_k}$ e, para cada $i = 1, \dots, k$, $r_i = \dim\big(V(\lambda_i)\big)$. Portanto, como o grau de $p_T(t)$ coincide com a dimensão de V, temos $n = r_1 + r_2 + \cdots + r_k = \dim\big(V(\lambda_1)\big) + \dim\big(V(\lambda_2)\big) + \cdots + \dim\big(V(\lambda_k)\big)$.

Reciprocamente, suponha que $\dim\big(V(\lambda_1)\big) + \dim\big(V(\lambda_2)\big) + \cdots + \dim\big(V(\lambda_k)\big) = n$ e tome, para cada $i = 1, \dots, n$, uma base $\mathcal{B}_i$ de $V(\lambda_i)$. Pelo Lema 7.4.1, $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$ é um conjunto LI, que contém $n = \dim(V)$ elementos. Logo, $\mathcal{B}$ é uma base de V formada por autovetores de T, o quer dizer que T é diagonalizável. $\square$

Exemplo 7.4.9 (Prova 2, Álgebra Linear II, 2016) Sejam V um espaço vetorial de dimensão 5 e $T\colon V \to V$ um operador linear cujo polinômio característico é $p_T(t) = -t^3(t-2)^2$. Então, T é diagonalizável se, e somente se,

(A) $\dim\big(\mathrm{Ker}(T - 2\,\mathrm{I}_V)\big) = 2$
(B) $\dim\big(\mathrm{Im}(T)\big) - \dim\big(\mathrm{Ker}(T - 2\,\mathrm{I}_V)\big) = 0$
(C) $\dim\big(\mathrm{Im}(T - 2\,\mathrm{I}_V)\big) - \dim\big(\mathrm{Im}(T)\big) = 1$
(D) $\dim\big(\mathrm{Im}(T)\big) = 2$
(E) $\dim\big(\mathrm{Im}(T - 2\,\mathrm{I}_V)\big) + \dim\big(\mathrm{Ker}(T)\big) = 5$

Solução: Os autovalores de T são 0 e 2. Sabemos, pelo Corolário 7.4.8, que T é diagonalizável se, e somente se, $\dim\big(V(0)\big) + \dim\big(V(2)\big) = \dim(V) = 5$. Agora, $V(0) = \mathrm{Ker}(T - 0\,\mathrm{I}_V) = \mathrm{Ker}(T)$ e $V(2) = \mathrm{Ker}(T - 2\,\mathrm{I}_V)$. Pelo Teorema do núcleo e da imagem (Teorema 6.3.1), $\dim\big(\mathrm{Ker}(T)\big) = \dim(V) - \dim\big(\mathrm{Im}(T)\big) = 5 - \dim\big(\mathrm{Im}(T)\big)$. Assim, T é diagonalizável se, e somente se, $\big(5 - \dim\big(\mathrm{Im}(T)\big)\big) + \dim\big(\mathrm{Ker}(T - 2\,\mathrm{I}_V)\big) = 5$, o que, por sua vez, é equivalente a $\dim\big(\mathrm{Im}(T)\big) - \dim\big(\mathrm{Ker}(T - 2\,\mathrm{I}_V)\big) = 0$. *Resposta:* (B) (Fica a cargo do leitor verificar que as igualdades nas demais alternativas não são equivalentes à da alternativa (B).). $\diamond$

Terminamos esta seção com uma observação a respeito de dois conceitos que são frequentemente confundidos, apesar de serem completamente independentes, quais sejam, o de um operador linear ser diagonalizável e o de ser bijetor. Não há relação de dependência entre esses dois conceitos, conforme comprovam os exemplos abaixo, em que exibimos operadores lineares $T\colon \mathbb{R}^2 \to \mathbb{R}^2$, com $[T]_{\mathsf{can}} = A \in M_2(\mathbb{R})$:

(i) Se $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, então T é tanto bijetor (pois $T = \mathrm{I}_{\mathbb{R}^2}$) quanto diagonalizável (can é uma base de $\mathbb{R}^2$ formada por autovetores de T).

(ii) Se $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, então T não é bijetor (pois não é injetor, uma vez que $(0,1) \in \mathrm{Ker}(T)$), mas é diagonalizável (can é uma base de $\mathbb{R}^2$ formada por autovetores de T).

(iii) Se $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, então T é bijetor (pois $\mathrm{Im}(T) = [(1,1),(1,0)] = \mathbb{R}^2$), mas não é diagonalizável (já que $\mathrm{Im}(T - \mathrm{I}_{R^2}) = [(0,1),(0,0)] = [(0,1)]$ e, portanto, $\dim(V(1)) = \dim\big(\mathrm{Ker}(T - \mathrm{I}_{\mathbb{R}^2})\big) = 2 - \dim\big(\mathrm{Im}(T - \mathrm{I}_{\mathbb{R}^2})\big) = 1$, ao passo que $p_T(t) = (t-1)^2$.)

(iv) Se $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, então T não é bijetor (pois, por exemplo, $(0,1) \in \mathrm{Ker}(T)$), nem diagonaliável (já que $p_T(t) = t^2$, ao passo que $\dim\big(V(0)\big) = \dim\big(\mathrm{Ker}(T)\big) = 2 - \dim\big(\mathrm{Im}(T)\big) = 1$, uma vez que $\mathrm{Im}(T) = [(0,1),(0,0)] = [(0,1)]$).

Exercícios Lista 2 - Álgebra Linear II: Exs. 64–81.

Apêndice A
O posto de uma matriz

Veremos que a dimensão do espaço gerado pelas linhas de uma matriz (seu posto-linha) sempre coincide com a dimensão do espaço gerado por suas colunas (o posto-coluna). Como consequência da demonstração desse fato, obteremos um método de extração de bases a partir de conjuntos geradores.

A.1 Posto-linha e posto-coluna

Dada uma matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

de $M_{m\times n}(\mathbb{R})$, podemos considerar suas linhas como vetores de $\mathbb{R}^n$ e suas colunas como vetores de $\mathbb{R}^m$, da seguinte maneira: para cada $i = 1, \dots, m$, seja

$$u_i = (a_{i1}, a_{i2}, \dots, a_{in}) \in \mathbb{R}^n,$$

e, para cada $j = 1, \dots, n$, seja

$$v_j = (a_{1j}, a_{2j}, \dots, a_{mj}) \in \mathbb{R}^m.$$

Denotaremos por $L(A)$ o subespaço de $\mathbb{R}^n$ gerado pelas linhas de A, isto é,

$$L(A) = [u_1, u_2, \dots, u_m],$$

e por $C(A)$ o subespaço de $\mathbb{R}^m$ gerado pelas colunas de A:

$$C(A) = [v_1, v_2, \dots, v_n].$$

A dimensão de $L(A)$ será denominada *posto-linha* de A e será denotada por $\mathrm{pl}(A)$. A dimensão de $C(A)$ é chamada *posto-coluna* de A e denotada por $\mathrm{pc}(A)$.

Teorema A.1.1 *Para qualquer* $A \in M_{m\times n}(\mathbb{R})$*, vale* $\mathrm{pl}(A) = \mathrm{pc}(A)$.

Demonstração Seja $R \in M_{m\times n}(\mathbb{R})$ uma matriz escalonada obtida a partir de A por meio de uma sequência de operações elementares sobre linhas (por exemplo, pelo processo de escalonamento). Uma vez que as linhas de R são combinações lineares das linhas de A, segue que $L(R) \subseteq L(A)$. Como operações elementares sobre linhas podem ser revertidas, A também pode ser obtida a partir de R por uma sequência de operações elementares sobre linhas, e, portanto, temos, de fato, $L(A) = L(R)$. Em particular, $\mathrm{pl}(A) = \mathrm{pl}(R)$. Agora, como R é escalonada, suas linhas não nulas (que são aquelas que contêm pivôs) são LI e, portanto, formam uma base para $L(R)$. Logo, se R tem r pivôs,

$$\text{pl}(A) = \text{pl}(R) = r. \tag{A.1}$$

Considere, agora, o sistema linear homogêneo

$$AX = 0. \tag{A.2}$$

Uma solução $(\alpha_1, \alpha_2, \dots, \alpha_n)$ de (A.2) origina uma relação de dependência linear entre as colunas de A:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0_{\mathbb{R}^m}.$$

E, reciprocamente, qualquer relação de dependência linear entre as coluna de A dá origem a uma solução de A. Assim, como (A.2) têm as mesmas soluções que o sistema linear homogêneo

$$RX = 0,$$

as relações de dependência linear entre as colunas de A são as mesmas que entre as colunas de R. Portanto, $\text{pc}(A) = \text{pc}(R)$. Cada coluna de R que não contém um pivô é combinação linear das colunas que contêm pivô e a precedem. Logo, $C(R)$ é gerado pelas colunas de R que contêm pivôs. Como essas colunas são, evidentemente, LI, elas formam uma base para $C(R)$. Sendo r o número de colunas de R com pivô, segue

$$\text{pc}(A) = \text{pc}(R) = r. \tag{A.3}$$

Juntando (A.1) e (A.3), obtemos $\text{pl}(A) = \text{pc}(A)$. □

Definição Dada uma matriz $A \in M_{m\times n}(\mathbb{R})$, o *posto* de A é definido por $\text{posto}(A) = \text{pl}(A)$.

Vimos, no Teorema A.1.1, que o posto de A coincide com o posto-coluna de A e com o número de pivôs de uma matriz escalonada obtida a partir de A pelo processo de escalonamento. Em particular, $\text{posto}(A) \leq \min\{m, n\}$. Além disso, apesar de a matriz escalonada obtida por operações elementares sobre linhas a partir de A não ser única, por depender de escolhas feitas ao longo do processo de escalonamento, o número de linhas não nulas dessa matriz (ou seja, o número de pivôs dela) depende apenas de A, pois coincide com o posto-linha de A. Resumidamente, se R_1 e R_2 são duas matrizes escalonadas obtidas a partir de A por meio de operações elementares sobre linhas, ambas têm o mesmo número de linhas não nulas, mesmo não sendo, necessariamente, matrizes iguais.

A.2 Extração de bases

Voltando à demonstração do Teorema A.1.1, vimos que, dada uma matriz $A \in M_{m\times n}(\mathbb{R})$, se $R \in M_{m\times n}(\mathbb{R})$ é uma matriz escalonada obtida a partir de A por meio de uma sequência de operações elementares sobre linhas, então as relações de dependência linear entre as colunas de A são as mesmas que existem entre as colunas de R. Vimos, também, que as colunas de R que contêm pivôs formam uma base para $C(R)$. Logo, não só é possível concluir que $\text{pc}(A) = \text{pc}(R)$, como também que as colunas de A que ocupam as mesmas posições que as colunas de R que contêm pivôs formam uma base para $C(A)$.

O Teorema 4.4.8 garantia que qualquer conjunto gerador finito de um espaço vetorial contém uma base, porém sua demonstração não fornecia um método eficiente de se obter uma base a partir desse conjunto gerador (ou, como também se costuma dizer, para se extrair do conjunto gerador uma base) — era preciso investigar as relações de dependência linear existentes entre esses geradores. A observação que fizemos no parágrafo anterior permite encontrar um procedimento eficiente, como veremos a seguir.

Método para extração de uma base de um conjunto gerador de um subespaço

Sejam $v_1, v_2, \dots, v_n \in \mathbb{R}^m$ e considere o subespaço

$$S = [v_1, v_2, \ldots, v_n]$$

de $\mathbb{R}^m$. Seja $A \in M_{m\times n}(\mathbb{R})$ a matriz que cujas *colunas* são $v_1, v_2, \ldots, v_n$, e seja $R \in M_{m\times n}(\mathbb{R})$ uma matriz escalonada obtida a partir de A por meio de uma sequência de operações elementares sobre *linhas* (por exemplo, pelo processo de escalonamento). Então, como acabamos de ver, uma base de S é formada pelos vetores v_j tais que a j-ésima coluna de R contém um pivô.

Exemplo A.2.1 Voltemos ao Exemplo 4.6.4, em que se pedia uma base para o subespaço $S = [u_1, u_2, u_3, u_4]$ de $\mathbb{R}^5$, em que

$$u_1 = (1, 0, -1, 2, 0),\ u_2 = (2, 1, 3, 0, 0),\ u_3 = (0, 1, -5, 4, 0),\ u_4 = (1, 0, -11, 10, 0).$$

Aplicando o método que acabamos de descrever, devemos construir uma matriz 5×4 cujas colunas são os vetores u_1, u_2, u_3, u_4 e submetê-la ao processo de escalonamento:

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ -1 & 3 & -5 & -11 \\ 2 & 0 & 4 & 10 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 5 & -5 & -10 \\ 0 & -4 & 4 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -10 & -10 \\ 0 & 0 & 8 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} \color{red}{1} & 2 & 0 & 1 \\ 0 & \color{red}{1} & 1 & 0 \\ 0 & 0 & \color{red}{-10} & -10 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Os pivôs da matriz escalonada obtida ao final do processo estão destacados em vermelho; eles ocorrem nas colunas 1, 2 e 3. Logo, $\{u_1, u_2, u_3\}$ é uma base de S.

Exemplo A.2.2 Encontre uma base para o subespaço $S = [\mathcal{A}]$ de $\mathbb{R}$, em que

$$\mathcal{A} = \big\{ (1, 0, -1, 2, 0),\ (2, 1, 3, 8, 0),\ (3, 1, 2, 10, 0),\ (0, 1, 6, 4, 0) \big\},$$

que esteja contida em $\mathcal{A}$.

Solução: Procedemos conforme o método apresentado, sujeitando a matriz cujas colunas são os elementos de $\mathcal{A}$ ao processo de escalonamento:

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 1 & 1 \\ -1 & 3 & 2 & 6 \\ 2 & 8 & 10 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 5 & 5 & 6 \\ 0 & 4 & 4 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} \color{red}{1} & 2 & 3 & 0 \\ 0 & \color{red}{1} & 1 & 1 \\ 0 & 0 & 0 & \color{red}{1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Como os pivôs da matriz escalonada ocorrem nas colunas 1, 2 e 4, segue que uma base de S é formada pelos primeiro, segundo e quarto elementos do conjunto $\mathcal{A}$: $\{(1, 0, -1, 2, 0), (2, 1, 3, 8, 0), (0, 1, 6, 4, 0)\}$. $\diamond$

Observação O método descrito e ilustrado acima aplica-se a subespaços do espaço vetorial $\mathbb{R}^m$. Ele, na verdade, pode se aplicar a qualquer subespaço de um espaço de dimensão finita, uma vez fixada uma base ordenada e trabalhando em coordenadas.

Mais detalhadamente, seja V um espaço vetorial de dimensão finita igual a m e seja $\mathcal{B}$ uma base ordenada de V. Dados $v_1, v_2, \ldots, v_n \in V$, considere o subespaço $S = [v_1, v_2, \ldots, v_n]$ de V. Tome, agora, para cada $j = 1, \ldots, n$, as coordenadas de v_j em relação à base $\mathcal{B}$:

$$v_j = (a_{1j}, a_{2j}, \ldots, a_{mj})_{\mathcal{B}},$$

e construa a matriz $A \in M_{m\times n}(\mathbb{R})$ que contém, em sua j-ésima coluna, as coordenadas de v_j em relação à base $\mathcal{B}$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

Uma base para S é formada por aqueles vetores v_j tais que a coluna j da matriz escalonada obtida a partir de A pelo processo de escalonamento contém um pivô.

Apêndice B
Um pouco sobre funções

Este apêndice tem por objetivo recordar alguns conceitos, provavelmente conhecidos do leitor, a respeito de funções entre dois conjuntos. Daremos especial atenção aos conceitos de injetividade e sobrejetividade e a condições equivalentes a eles envolvendo a noção de composição de funções.

B.1 Definições

Sejam X e Y conjuntos quaisquer (não necessariamente dotados de estrutura de espaço vetorial ou qualquer outra estrutura). O *produto cartesiano* $X \times Y$ é o conjunto formado por todos os pares ordenados (x, y), com $x \in X$ e $y \in Y$.

Por definição, uma *função* de X em Y é um subconjunto f do produto cartesiano $X \times Y$ tal que para todo $x \in X$, existe um único $y \in Y$ tal que $(x, y) \in f$. Essa é a definição formal de função, mas veja que ela coincide com a noção intuitiva que o leitor deve já deve ter: uma função tem um domínio — o conjunto X —, um contradomínio — o conjunto Y — e uma "regra" que associa a cada elemento do domínio um elemento do contradomínio, chamado sua imagem. Na definição formal, a "regra" nada mais é do que a listagem completa de todo os pontos do domínio acompanhados de suas imagens. Para nos aproximarmos da notação com a que o leitor provavelmente é mais familiar, escreveremos $f(x) = y$ para denotar que $(x, y) \in f$ e indicaremos o fato de f ser uma função com domínio X e contradomínio Y por $f\colon X \to Y$.

Uma função $f\colon X \to Y$ é dita

- *injetora* se para todos $x, x' \in X$, $x \neq x'$ implicar $f(x) \neq f(x')$;
- *sobrejetora* se para todo $y \in Y$ existir (ao menos um) $x \in X$ tal que $y = f(x)$;
- *bijetora* se for injetora e sobrejetora.

As condições de injetividade e sobrejetividade são completamente independentes.

Exemplo B.1.1 A função $f\colon \{1, 2, 3\} \to \{4, 5, 6, 7\}$, definida por

$$f(1) = 4, \qquad f(2) = 5, \qquad f(3) = 6,$$

é injetora, mas não é sobrejetora.

A função $g\colon \{1, 2, 3\} \to \{4, 5\}$, definida por

$$g(1) = 4 \qquad g(2) = 4, \qquad g(3) = 5,$$

é sobrejetora, mas não é injetora.

A função $h\colon \{1, 2, 3\} \to \{4, 5, 6\}$, definida por

$$h(1) = 4, \qquad h(2) = 5, \qquad h(3) = 6,$$

é bijetora.

Dizemos que duas funções $f\colon X \to Y$ e $g\colon Z \to W$ são iguais, e, neste caso, escrevemos $f = g$, se $X = Z$, $Y = W$ e $f(x) = g(x)$, para todo $x \in X$. (Dito de outra forma, f e g são iguais se tiverem mesmo domínio, mesmo contradomínio e as imagens de cada ponto do domínio comum coincidirem.)

B.2 Composição de funções

Se X, Y, Z são conjuntos e $f\colon X \to Y$ e $g\colon Y \to Z$ são funções, pode-se definir a *função composta* $g \circ f\colon X \to Z$, dada por

$$(g \circ f)(x) = g\big(f(x)\big), \quad \text{para todo } x \in X.$$

Observe que esta definição funciona bem, uma vez que para todo $x \in X$, temos $f(x) \in Y$, que é o domínio da função g.

Exemplo B.2.1 Considere a função $f\colon \{1, 2, 3\} \to \{4, 5, 6, 7\}$, definida por

$$f(1) = 7, \qquad f(2) = 4, \qquad f(3) = 6,$$

e a função $g\colon \{4, 5, 6, 7\} \to \{8, 9\}$, definida por

$$g(4) = 8, \qquad g(5) = 8, \qquad g(6) = 9, \qquad g(7) = 8.$$

Como o contradomínio de f (que é o conjunto $\{4, 5, 6, 7\}$) coincide com o domínio de g, a função composta $g \circ f$ está definida, tem domínio coincidente com o domínio de f, contradomínio coincidente com o contradomínio de g (em símbolos, $g \circ f\colon \{1, 2, 3\} \to \{8, 9\}$) e é definida por

$$(g \circ f)(1) = g\big(f(1)\big) = g(7) = 8, \qquad (g \circ f)(2) = g\big(f(2)\big) = g(4) = 8, \qquad (g \circ f)(3) = g\big(f(3)\big) = g(6) = 9.$$

A operação de composição de funções é associativa, no sentido de que se $f\colon X \to Y$, $g\colon Y \to Z$, $h\colon Z \to W$, então $(h \circ g) \circ f = h \circ (g \circ f)$. (É fácil ver que ambas têm domínio X e contradomínio W, e que, as imagens por ambas em um elemento $x \in X$ são iguais a $h\big(g\big(f(x)\big)\big)$.) Mas a operação de composição *não é comutativa*. Isso quer dizer que se $f\colon X \to Y$ e $g\colon Y \to X$, então as composições $g \circ f$ e $f \circ g$ estão ambas definidas, mas, em geral, $g \circ f \neq f \circ g$. Isso é claro quando $X \neq Y$. Mas, mesmo quando $X = Y$, $g \circ f$ e $f \circ g$ não costumam coincidir.

Exemplo B.2.2 Considere a função $f\colon \{1, 2, 3\} \to \{1, 2, 3\}$, dada por

$$f(1) = 2, \qquad f(2) = 3, \qquad f(3) = 1,$$

e a função $g\colon \{1, 2, 3\} \to \{1, 2, 3\}$, dada por

$$g(1) = 1, \qquad g(2) = 3, \qquad g(3) = 2.$$

Então, $g \circ f$ e $f \circ g$ estão, ambas, definidas, têm domínios e contradomínios coincidentes, mas $g \circ f \neq f \circ g$, pois, por exemplo,

$$(g \circ f)(1) = 3 \neq 2 = (f \circ g)(1).$$

Dado um conjunto X, chamamos de *função identidade* do conjunto X a função $\mathrm{id}_X\colon X \to X$ definida por $\mathrm{id}_X(x) = x$, para todo $x \in X$. Claramente, id_X é uma função bijetora.

As funções identidade funcionam como "elementos neutros" para a composição: se X e Y são conjuntos e $f\colon X \to Y$ é uma função, então

$$f \circ \mathrm{id}_X = f \qquad \text{e} \qquad \mathrm{id}_Y \circ f = f.$$

Exercícios Sejam X, Y, Z conjuntos e sejam $f\colon X \to Y$, $g\colon Y \to Z$ funções. Demonstre as implicações abaixo.

1. Se f e g são injetoras, então $g \circ f$ é injetora.
2. Se f e g são sobrejetoras, então $g \circ f$ é sobrejetora.
3. Se $g \circ f$ é injetora, então f é injetora.
4. Se $g \circ f$ é sobrejetora, então g é sobrejetora.

Funções bijetoras são *inversíveis* no sentido de que se $f\colon X \to Y$ é uma função bijetora, então existe uma função $g\colon Y \to X$ que "desfaz o que f faz", mais precisamente

Proposição B.2.3 *Sejam X e Y conjuntos e seja $f\colon X \to Y$ uma função bijetora. Então, existe uma única função $g\colon Y \to X$ tal que $g \circ f = \mathrm{id}_X$ e $f \circ g = \mathrm{id}_Y$.*

Essa função g é chamada de *função inversa* da função f e é comumente denotada por f^{-1}.

Demonstração A fim de definir $g\colon Y \to X$, para cada $y \in Y$ é preciso definir o elemento de X dado por $g(y)$. Como f é sobrejetora, existe um elemento $x \in X$ tal que $f(x) = y$. De fato, só temos uma escolha para tal x, pois se $x' \in X$ também satisfaz $f(x') = y$, então $x = x'$, uma vez que f é, também, injetora. Com este elemento x em mãos, definimos $g(y) = x$.

Tendo a função g já definida, verifiquemos que ela satisfaz as condições desejadas. Para cada $x \in X$, temos $f(x) \in Y$ e, segundo a definição de g, $g\big(f(x)\big) = x'$, em que $x' \in X$ é tal que $f(x') = f(x)$. Mas, como f é injetora, $x' = x$. Portanto, para todo $x \in X$, temos $g\big(f(x)\big) = x$. Logo, $g \circ f = \mathrm{id}_X$. Para a outra composta, tome $y \in Y$. Então $g(y) = x$, em que $x \in X$ é tal que $f(x) = y$. Portanto, $f\big(g(y)\big) = f(x) = y$. Isso prova que $f \circ g = \mathrm{id}_X$.

Finalmente, mostremos a unicidade de g. Seja $h\colon Y \to X$ tal que $h \circ f = \mathrm{id}_X$ e $f \circ h = \mathrm{id}_Y$. Então,

$$g = g \circ \mathrm{id}_Y = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_X \circ h = h,$$

o que prova que g é a única que satisfaz ambas as condições. □

Em resumo, foi mostrado que se $f\colon X \to Y$, então existe a função inversa $f^{-1}\colon Y \to X$ de f, e essa função satisfaz a seguinte condição: para todos $x \in X$ e $y \in Y$, temos

$$f(x) = y \quad \text{se, e somente se,} \quad f^{-1}(y) = x.$$

Exercícios Sejam X, Y, Z conjuntos e sejam $f\colon X \to Y$, $g\colon Y \to Z$ funções. Demonstre as implicações abaixo.

1. Se f é bijetora, então a função inversa $f^{-1}\colon Y \to X$ é bijetora e $(f^{-1})^{-1} = f$.
2. Se f e g são bijetoras, então $g \circ f\colon X \to Z$ é bijetora e $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Terminamos observando que a Proposição B.2.3 tem uma recíproca:

Proposição B.2.4 *Sejam X e Y conjuntos e seja $f\colon X \to Y$ uma função. Se existem funções $g, h\colon Y \to X$ tais que $g \circ f = \mathrm{id}_X$ e $f \circ h = \mathrm{id}_Y$, então f é bijetora e $f^{-1} = g = h$.*

Demonstração Como $g \circ f = \mathrm{id}_X$, segue que f é injetora (pois id_X é injetora). Como $f \circ h = \mathrm{id}_Y$, segue que f é sobrejetora (pois id_Y é sobrejetora). Logo, f é bijetora. Além disso,

$$g = g \circ \mathrm{id}_Y = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_X \circ h = h.$$

Finalmente, pela unicidade da inversa, $f^{-1} = g$. □

Apêndice C
Polinômios e suas raízes

Este apêndice contém alguns resultados a respeito de raízes e fatorações de polinômios comumente cobertos nos cursos de matemática no Ensino Médio.

C.1 Polinômios com coeficientes reais

Todos os polinômios de que trataremos nesta seção terão coeficientes reais, isto é, serão expressões da forma $a_0 + a_1 t + \cdots + a_n t^n$, com $a_0, a_1, \ldots, a_n \in \mathbb{R}$.

Se $f(t)$ é um polinômio de grau n, denotaremos esse fato por $\operatorname{gr}\big(f(t)\big) = n$.

Lembre que o produto entre dois polinômios é, novamente, um polinômio que é obtido por distributividade, utilizando as identidades

$$t^n t^m = t^{m+n} \qquad \text{e} \qquad at = ta, \quad \text{para todo } a \in \mathbb{R}.$$

Assim, se $\operatorname{gr}\big(f(t)\big) = n$ e $\operatorname{gr}\big(g(t)\big) = m$, então $\operatorname{gr}\big(f(t)g(t)\big) = n + m$.

Se $f(t) = a_0 + a_1 t + \cdots + a_n t_n$ é um polinômio e $\beta \in \mathbb{R}$, será denotado por $f(\beta)$ o número real obtido pela substituição da variável t na expressão de $f(t)$ por β:

$$f(\beta) = a_0 + a_1\beta + \cdots + a_n\beta^n.$$

Assim, é fácil ver que se $f(t), g(t), h(t), p(t)$ são polinômios tais que $h(t) = f(t)g(t)$ e $p(t) = f(t) + g(t)$, então $h(\beta) = f(\beta)g(\beta)$ e $p(\beta) = f(\beta) + g(\beta)$.

Dados dois polinômios $f(t), g(t)$, dizemos que $f(t)$ é divisível por $g(t)$ se existir um polinômio $h(t)$ tal que $f(t) = g(t)h(t)$. Um caso particular de divisibilidade entre polinômios é o seguinte:

Lema C.1.1 *Seja $\beta \in \mathbb{R}$. Então, para todo inteiro positivo n, o polinômio $t^n - \beta^n$ é divisível por $t - \beta$.*

Demonstração Segue da identidade

$$t^n - \beta^n = (t - \beta)(t^{n-1} + \beta t^{n-2} + \beta^2 t^{n-3} + \cdots + \beta^{n-2} t + \beta^{n-1}),$$

que pode ser verificada efetuando-se o produto. □

Dado um polinômio $f(t) = a_0 + a_1 t + \cdots + a_n t^n$, com $a_0, a_1, \ldots, a_n \in \mathbb{R}$, dizemos que um número real β é *raiz* de $f(t)$ se $f(\beta) = 0$.

Proposição C.1.2 *O número real β é raiz do polinômio $f(t)$ se, e somente se, $f(t)$ for divisível por $t - \beta$.*

Demonstração Se $f(t)$ é divisível por $t - \beta$, então existe um polinômio $g(t)$ tal que $f(t) = g(t)(t - \beta)$. Assim, $f(\beta) = g(\beta)(\beta - \beta) = g(\beta)0 = 0$. Logo, β é raiz de $f(t)$.

Para a recíproca, observe que $t-\beta$ divide $f(t)-f(\beta)$, uma vez que

$$f(t)-f(\beta)=a_1(t-\beta)+a_2(t^2-\beta^2)+\cdots+a_n(t^n-\beta^n).$$

Pelo Lema C.1.1, cada um dos termos no lado esquerdo da igualdade é divisível por $t-\beta$, o que faz a soma ser divisível por $t-\beta$. Em particular, se $f(\beta)=0$, então $f(t)$ é divisível por $t-\beta$. □

Assim, uma vez encontrada uma raiz β de um polinômio $f(t)$, a busca pelas demais raízes se reduz à busca das raízes de um polinômio $g(t)$ que, por satisfazer $f(t)=(t-\beta)g(t)$, tem grau menor do que o de $f(t)$ (precisamente, $\operatorname{gr}\big(g(t)\big)=\operatorname{gr}\big(f(t)\big)-1$). Iterando esse processo, no caso em que cada um dos quocientes tiver pelo menos uma raiz, obtém-se uma fatoração completa de $f(t)$ como um produto de polinômios de grau 1:

$$f(t)=a(t-\beta_1)^{r_1}(t-\beta_2)^{r_2}\dots(t-\beta_k)^{r_k},$$

em que $a\in\mathbb{R}$ e $\beta_1,\beta_2,\dots,\beta_k$ são as raízes distintas de $f(t)$. Para cada $i=1,\dots,k$, o expoente r_i é chamado multiplicidade de β_i como raiz de $f(t)$. Mais geralmente, mesmo que $f(t)$ não se fatore completamente como produto de fatores de grau 1, se β é uma raiz de $f(t)$, dizemos que β tem multiplicidade r se $(t-\beta)^r$ divide $f(t)$, mas $(t-\beta)^{r+1}$ não divide $f(t)$.

A tarefa de encontrar as raízes (se existirem) de um polinômio é difícil, mas há alguns testes que podem ser realizados em alguns casos. Por exemplo, 0 é raiz de um polinômio $f(t)$ se, e somente se, o coeficiente do termo de grau 0 em $f(t)$ for igual a 0. Para polinômios com coeficientes inteiros, quando isso não ocorre, há ainda um teste útil para as possíveis raízes racionais.

Proposição C.1.3 *Seja $f(t)=a_0+a_1t+a_2t^2+\cdots+a_nt^n$ um polinômio de grau n (isto é, tal que $a_n\neq 0$), cujos coeficientes são todos números inteiros, com $a_0\neq 0$. Se $\beta=\frac{r}{s}$ é uma raiz racional de $f(t)$, com r e s inteiros relativamente primos, então r divide a_0 e s divide a_n.*

Demonstração Como $\beta=\frac{r}{s}$ é raiz de $f(t)$, temos que $a_0+a_1\frac{r}{s}+a_2\frac{r^2}{s^2}+\cdots+a_n\frac{r^n}{s^n}=0$. Multiplicando essa igualdade, em ambos os lados, por s^n, obtemos

$$a_0s^n+a_1rs^{n-1}+a_2r^2s^{n-2}+\cdots+a_{n-1}r^{n-1}s+a_nr^n=0. \tag{C.1}$$

Assim, $a_0s^n=r(-a_1s^{n-1}-a_2rs^{n-2}-\cdots-a_{n-1}r^{n-2}s-a_nr^{n-1})$. Portanto, r divide o inteiro a_0s^n. Como r e s são primos entre si, segue que r divide a_0. De (C.1) também obtemos que $a_nr^n=s(-a_0s^{n-1}-a_1rs^{n-2}-\cdots-a_{n-1}r^{n-1})$, donde, usando novamente que r e s são primos entre si, segue que s divide a_n. □

Essa proposição permite a seguinte estratégia: se $f(t)=a_0+a_1t+\cdots+a_nt^n$ é um polinômio com coeficientes inteiros, com $a_0\neq 0$ e $a_n\neq 0$, e $f(t)$ tiver uma raiz racional β, então $\beta\in\left\{\frac{r}{s}\mid r \text{ é divisor de } a_0 \text{ e } s \text{ é divisor de } a_n\right\}$, um conjunto finito que pode ser inspecionado por substituição de cada um de seus elementos em $f(t)$.

Exemplo C.1.4 Encontre as raízes de $f(t)=3t^3+4t^2-6t-8$.

Solução: Os divisores de -8 são $\pm1,\pm2,\pm4,\pm8$ e os divisores de 3 são $\pm1,\pm3$. Assim, se $f(t)$ tiver raízes racionais, elas estarão no conjunto $\left\{\pm1,\pm2,\pm4,\pm8,\pm\frac{1}{3},\pm\frac{2}{3},\pm\frac{4}{3},\pm\frac{8}{3}\right\}$. Verificando cada um desses dezesseis valores, encontramos $f(-\frac{4}{3})=0$. Logo, $t+\frac{4}{3}$ divide $f(t)$ e obtemos a fatoração $f(t)=3(t+\frac{4}{3})(t^2-2)$. Como as raízes de t^2-2 são $\pm\sqrt{2}$, obtemos uma fatoração completa: $f(t)=3\big(t+\frac{4}{3}\big)\big(t-\sqrt{2}\big)\big(t+\sqrt{2}\big)$. ◇

Referências

1. H. Anton e C. Rorres, *Álgebra Linear com Aplicações*, 10a. ed., Bookmam, Porto Alegre, 2012.
2. P. Boulos e I. Camargo, *Geometria Analítica: um tratamento vetorial*, 2a. ed., Pearson Education do Brasil, São Paulo, 2003.
3. M. Barone Júnior, *Álgebra Linear*, 3a. ed., IME-USP, São Paulo, 1988.
4. C. A. Callioli, H. H. Domingues e R. C. F. Costa, *Álgebra Linear e Aplicações*, 6a. ed., Atual Editora, São Paulo, 1998.
5. W. K. Nicholson, *Álgebra Linear*, 2a. ed., McGraw-Hill, São Paulo, 2006.

Índice Remissivo

ângulo entre vetores, 28

autoespaço, 141
autovalor
  de um operador linear, 140
  de uma matriz, 149
autovetor
  de um operador linear, 140
  de uma matriz, 149

base, 75
  canônica, 75
  ordenada, 80
  ortogonal, 100
  ortonormal, 100
base (de $\mathbb{V}^3$), 23
  negativa, 33
  ortonormal, 26
  positiva, 33
bases
  de mesma orientação, 33
  de orientações opostas, 33

Cauchy-Schwarz, 98
combinação linear, 69
combinação linear (em $\mathbb{V}^3$), 20
complemento ortogonal, 109
coordenadas
  de um ponto, 43
  de um vetor, 80
  de um vetor (em $\mathbb{V}^3$), 24

dependência linear, 72
dependência linear (em $\mathbb{V}^3$), 20
desigualdade
  de Cauchy-Schwarz, 98
  triangular, 98
determinante, 12
  de Vandermonde, 75
dimensão, 78
distância, 97
  entre planos, 62
  entre ponto e plano, 59
  entre ponto e reta, 58
  entre pontos, 45
  entre reta e plano, 62
  entre retas, 60

eliminação gaussiana, 5
equação
  geral do plano, 49
  linear, 1
  vetorial
    da reta, 45
    do plano, 48
equações
  na forma simétrica da reta, 46
  paramétricas
    da reta, 46
    do plano, 49
escalar, 19
escalonamento, 5
espaço vetorial, 63
  de funções, 64
  de matrizes, 64
  de polinômios, 64

Gram-Schmidt, 101

imagem, 115
interseção
  de subespaços, 86

LD, 72
LD (em $\mathbb{V}^3$), 20
LI, 72
LI (em $\mathbb{V}^3$), 20

matriz
  aumentada, 3
  da transformação composta, 131
  de mudança de base, 134
  de mudança de base (em $\mathbb{V}^3$), 32
  de uma transformação linear, 124
  de Vandermonde, 75
  diagonal, 139
  diagonalizável, 139
  elementar, 10
  escalonada, 5
  identidade, 9

inversível, 9
transposta, 12
triangular superior, 12
matrizes
semelhantes, 143
melhor aproximação, 104
multiplicação por escalar (em $\mathbb{V}^3$), 19
multiplicidade, 152
algébrica, 152
geométrica, 153

núcleo, 115
norma, 97
norma (em $\mathbb{V}^3$), 18

operação elementar, 4
operador
identidade, 113
linear, 112

pivô, 5
planos
perpendiculares, 57
polinômio característico
de um operador, 144
de uma matriz, 143
posto
de uma matriz, 121, 160
posto-coluna, 159
posto-linha, 159
problema da melhor aproximação, 104
processo
de escalonamento, 5
de ortogonalização de Gram-Schmidt, 101
produto
escalar, 28
interno, 95
misto, 38
vetorial, 34
projeção ortogonal, 105
projeção ortogonal (em $\mathbb{V}^3$), 30

retas
concorrentes, 53
ortogonais, 57
paralelas, 53
perpendiculares, 57
reversas, 53
retrossubstituição, 7

segmento orientado, 17
sistema de coordenadas, 43
ortogonal, 43
sistema linear, 1
homogêneo, 8
impossível, 2
possível determinado, 2
possível indeterminado, 2
sistemas lineares equivalentes, 4
solução trivial, 8
soma
de subespaços, 88
de vetores (em $\mathbb{V}^3$), 18
direta, 93
subespaço
gerado, 70
ortogonal, 108
vetorial, 67

Teorema
da dimensão, 120
do completamento, 79
do núcleo e da imagem, 120
traço
de uma matriz, 100
transformação linear, 111

valor
característico, 140
próprio, 140
variável
livre, 7
pivô, 7
vetor
característico, 140
de coordenadas, 126
diretor de uma reta, 45
normal a um plano, 51
nulo, 63
nulo (de $\mathbb{V}^3$), 18
oposto, 18
próprio, 140
vetores
coplanares, 21
diretores do plano, 48
ortogonais, 99
ortogonais (em $\mathbb{V}^3$), 26
paralelos, 18