Álgebra Linear

Vinicius de Oliveira Rodrigues

14 de novembro de 2025

Sumário

Capítulo 1

Introdução

Neste capítulo, estabeleceremos a notação básica sobre a qual trabalharemos ao longo deste texto.

Também, neste capítulo, trabalharemos um mínimo com as noções de anéis e de corpos. Não é o objetivo deste texto dar uma introdução à Teoria de Corpos ou à Teoria de Anéis. Apenas buscamos introduzir um mínimo sobre a linguagem de corpos que seja suficiente para desenvolver a Teoria da Álgebra Linear.

1.1 Noções de Conjuntos

Utilizaremos, a notação usual de teoria dos conjuntos ao longo deste texto.

Assumiremos, implicitamente, e sem qualquer aprofundamento desnecessário, que estamos trabalhando na *Teoria dos Conjuntos de Zermelo-Fraenkel com o Axioma da Escolha (ZFC)*. O conhecimento sobre o que se trata tal teoria não é, em nenhum aspecto, necessário para a compreensão deste texto. Porém, espera-se o entendimento básico sobre o que é a relação de um conjunto, a relação de pertinência $\in$, a relação de subconjunto $\subseteq$, o entendimento de operações básicas como união, interseção, e diferença de conjuntos, a compreensão de notações como $\{a, b, c\}$, da noção de par ordenado, (a, b).

Também espera-se a compreensão das noções básicas de funções, como o entendimento da notação $f : A \to B$, da noção de domínio e imagem de uma função, injetividade, sobrejetividade, bijetividade, invertibilidade de funções e de contradomínios.

As noções de *operações* em um conjunto aparecerão naturalmente ao longo do texto. Em particular, estabelece-se a seguir o que chamaremos de *operação unária* e de *operação binária*.

Lembramos que $A \times B$ denota o produto cartesiano de dois conjuntos A e B, ou seja, o conjunto de todos os pares ordenados (a, b) com $a \in A$ e $b \in B$. Se $n \geq 1$ e A é um conjunto, escrevemos $A^n = \{(a_1, \ldots, a_n) : a_1, \ldots, a_n \in A\}$.

Definição 1.1.1 (Operações)**.** Uma *operação unária* em um conjunto A é uma função $* : A \to A$.

Ao trabalhar com uma operação unária, frequentemente utilizamos notações especiais. Abrevia-se $*(a)$ como $*a$.

Uma *operação binária* em um conjunto A é uma função $* : A \times A \to A$.

Ao trabalhar com uma operação binária, frequentemente utilizamos notações especiais. Abrevia-se $*(a, b)$ como $a * b$.

Por exemplo: ao lidar-se com números reais, a *soma usual* $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ é uma operação binária, de modo que $+(a, b) = a + b$.

Da mesma forma, a operação de *oposto* $- : \mathbb{R} \to \mathbb{R}$ é uma operação unária, de modo que $-(a) = -a$.

1.2 Anéis e corpos

Anéis e corpos possuem como objetivo generalizar estruturas numéricas conhecidas que possuem uma soma e um produto, como os números inteiros, racionais, reais e complexos.

Se R é um conjunto, uma **operação binária** em R é uma função $* : R \times R \to R$

Definição 1.2.1. Um *anel* é um conjunto R munido de duas operações binárias, geralmente denotadas por $+$ e $\cdot$, e por dois elementos destacados, geralmente denotados por 0 e 1, tais que, para todos $a, b, c \in R$, valem:

(A1) $(a + b) + c = a + (b + c)$ (associatividade da soma),

(A2) $a + b = b + a$ (comutatividade da soma),

(A3) $a + 0 = a$ (elemento neutro),

(A4) existe $x \in R$ tal que $a + x = 0$ (elemento oposto),

(M1) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associatividade do produto),

(M2) $a \cdot 1 = 1 \cdot a = a$ (elemento neutro do produto),

(D1) $a \cdot (b + c) = a \cdot b + a \cdot c$ (distributiva à esquerda),

(D2) $(a + b) \cdot c = a \cdot c + b \cdot c$ (distributiva à direita).

Se, além disso, para todos $a, b \in R$ vale que:

(M3) $a \cdot b = b \cdot a$ (comutatividade do produto),

então dizemos que R é um *anel comutativo*.

No contexto de anéis, quando não há confusão, é usual denotar $a \cdot b$ simplesmnete por ab. Além disso, valem as regras usuais de precedência de parênteses, de modo que $a + b \cdot c$ é interpretado como $a + (b \cdot c)$, e **não** como $(a + b) \cdot c$.

Exemplos de anéis incluem os inteiros $\mathbb{Z}$, os números racionais $\mathbb{Q}$, os números reais $\mathbb{R}$ e os números complexos $\mathbb{C}$.

Outro exemplo de anel é o conjunto unitário $\{a\}$ com as operações $a + a = 0$ e $a \cdot a = a$, onde a é qualquer objeto. Denotando-se $1 = 0 = a$, verifica-se facilmente que $\{a\}$ é um anel comutativo. Os anéis dessa forma são chamados de *anéis triviais*.

Os números naturais $\mathbb{N}$ não formam um anel, pois não possuem elementos opostos para a soma.

Vejamos algumas propriedades básicas.

Proposição 1.2.2 (Unicidade do elemento neutro da soma)**.** Seja R um anel. Então só existe um elemento neutro para a soma. Ou seja, se $x \in R$ é tal que para todos $a \in R$ vale $a + x = a$, então $x = 0$.

Demonstração. Como 0 é elemento neutro da soma, temos que $0 + x = x$. Por outro lado, como x é elemento neutro da soma, temos que $x + 0 = x$. Logo, $x = 0$. □

Do mesmo modo, prova-se a unicidade do elemento neutro do produto. Deixamos a redação da demonstração como exercício de fixação ao leitor.

Proposição 1.2.3 (Unicidade do elemento neutro do produto)**.** Seja R um anel. Então só existe um elemento neutro para o produto. Ou seja, se $x \in R$ é tal que para todos $a \in R$ vale $a \cdot x = x \cdot a = a$, então $x = 1$.

Demonstração. Exercício. □

Da mesma forma, os *opostos aditivos* são únicos.

Proposição 1.2.4 (Unicidade dos opostos aditivos)**.** Seja R um anel. Então, para cada $a \in R$, existe um *único* elemento $x \in R$ tal que $a + x = 0$.

Demonstração. A existência de um tal x é garantida pela propriedade A4. Para a unicidade, suponha que x e y são elementos de R tais que $a + x = 0$ e $a + y = 0$. Segue que:

$$x = x + 0 = x + (a + y) = (x + a) + y = 0 + y = y.$$

□

Assim, podemos dar um nome especial para o oposto aditivo de um elemento $a \in R$:

Definição 1.2.5 (Opostos aditivos)**.** Seja R um anel. Então, para cada $a \in R$ o *oposto aditivo* de a, denotado por $-a$, é o único elemento de R que satisfaz $a + (-a) = 0$.

Em particular, veja que $-0 = 0$, pois $0 + 0 = 0$.

Uma propriedade particular interessante do elemento nulo é que ele anula, via multiplicação, qualquer outro elemento.

Notacionalmente, é útil definir uma notação binária para a *diferença* de dois elementos.

Definição 1.2.6 (Diferença)**.** Seja R um anel e $a, b \in R$. A *diferença* de a e b, denotada por $a - b$, é definida como $a + (-b)$.

Algumas propriedades dos elementos opostos são:

Proposição 1.2.7. Seja R um anel e $a, b \in R$. Então:

(i) $-(-a) = a$.

(ii) $-(a + b) = (-a) + (-b)$.

(iii) $-0 = 0$.

(iv) $a \cdot 0 = 0 = 0 \cdot a$.

(v) $(-a)b = a(-b) = -(ab)$.

(vi) $(-a)(-b) = ab$.

(vii) $-a = (-1)a$.

Demonstração. (i) Como $(-a) + a = 0$, temos que $-(-a) = a$.

(ii) Como $(a+b)+((-a)+(-b)) = a+(-a)+b+(-b) = 0+0 = 0$, temos que $-(a+b) = (-a)+(-b)$.

(iii) Note que $0 + 0 = 0$, temos que $-0 = 0$.

(iv) Veja que $0 \cdot a = (0+0) \cdot a = 0 \cdot a + 0 \cdot a$. Assim, $0 = (0 \cdot a) + (-(0 \cdot a)) = [(0 \cdot a) + (0 \cdot a)] + (-(0 \cdot a)) = (0 \cdot a) + [(0 \cdot a) + (-(0 \cdot a))] = 0 \cdot a$.

Analogamente, $a \cdot 0 = 0$.

(v) Note que $ab + (-a)b = [a + (-a)]b = 0b = 0$. Assim, $(-a)b = -(ab)$. Da mesma forma, $a(-b) = -(ab)$.

(vi) Pelo item anterior, temos que $(-a)(-b) = -[a(-b)] = -(-(ab)) = ab$.

(vii) Finalmente, $(-1)a = -(1 \cdot a) = -a$.

$\square$

Assim, as chamadas "regras de sinais" decorrem das definições iniciais sobre anéis, e as operações numéricas de subtração usuais podem ser vistas como casos particulares de somas, fazendo-se uso da noção de opostos.

Nesse contexto, como podemos encarar a divisão? Para isso, precisamos da noção de *inverso multiplicativo*.

Definição 1.2.8. Seja R um anel e $a \in R$. Dizemos que a é *invertível* se existe $b \in R$ tal que $a \cdot b = b \cdot a = 1$.

É de conhecimento do leitor que não existe um elemento $b \in \mathbb{Z}$ tal que $2 \cdot b = 1$. Assim, 2 não é invertível em $\mathbb{Z}$. Por outro lado, 2 é invertível em $\mathbb{Q}$, pois $2 \cdot \frac{1}{2} = \frac{1}{2} \cdot 2 = 1$.

Similarmente ao que ocorre com opostos aditivos, os inversos multiplicativos, caso existam, são únicos.

Proposição 1.2.9. Seja R um anel e $a \in R$ um elemento invertível. Então existe um *único* elemento $b \in R$ tal que $a \cdot b = b \cdot a = 1$.

Demonstração. A existência de b é garantida pela definição de elemento invertível. Para a unicidade, suponha que b e c são elementos de R tais que $a \cdot b = b \cdot a = 1$ e $a \cdot c = c \cdot a = 1$.

Segue que $b = b.1 = b(ac) = (ba)c = 1 \cdot c = c$. $\square$

Definição 1.2.10. Seja R um anel e $a \in R$ um elemento invertível. O *inverso multiplicativo* de a, denotado por a^{-1}, é o único elemento de R tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

Algumas propriedades dos inversos multiplicativos são:

Proposição 1.2.11. Seja R um anel e $a, b \in R$ elementos invertíveis. Então:

(i) ab é invertível e $(ab)^{-1} = b^{-1}a^{-1}$.

(ii) a^{-1} é invertível e $(a^{-1})^{-1} = a$.

(iii) 1 é invertível e $1^{-1} = 1$.

Demonstração. (i) Notemos que $(ab)(b^{-1}a^{-1}) = abb^{-1}a^{-1} = aa^{-1} = 1$. Similarmente, $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}b = 1$. Assim, ab é invertível e $(ab)^{-1} = b^{-1}a^{-1}$.

(ii) Note que $a^{-1}a = aa^{-1} = 1$, logo a^{-1} é invertível e $(a^{-1})^{-1} = a$.

(iii) Como $1 \cdot 1 = 1$, temos que $1^{-1} = 1$.

□

Finalmente, chegamos a noção de corpo. Um corpo é um anel comutativo no qual todo elemento é invertível, a excessão do 0, que deve necessariamente ser diferente de 1.

Definição 1.2.12. Um *corpo* é um anel comutativo $\mathbb{K}$ no qual valem, adicionalmente, as seguintes propriedades:

(NT) $0 \neq 1$ (não trivialidade).

(M4) Para todo $a \in \mathbb{K}$, se $a \neq 0$ então a é invertível (existência de inversos multiplicativos).

Assim, em um corpo, todo elemento não nulo é invertível. De fato, esses são todos, como mostra a proposição a seguir.

Proposição 1.2.13. Seja $\mathbb{K}$ um corpo. Então $0 \in K$ não é invertível.

Demonstração. Suponha por absurdo que 0 é invertível. Então existe $x \in \mathbb{K}$ tal que $0x = 1$. Porém, conforme visto, $0x = 0$, logo, $0 = 1$, violando a não trivialidade do corpo. □

Neste texto, nossos principais exemplos de corpos serão o conjunto dos números reais e o conjunto dos números complexos, e, por vezes, o conjunto dos números racionais, com os quais supomos que o leitor tenha alguma familiaridade, ainda que informal. Porém, como exemplo de um corpo diferente destes, temos:

Exemplo 1.2.14. Seja $\mathbb{F}_2 = \{0, 1\}$, com as operações de soma e produto definidas por:

- $0 + 0 = 0$, $0 + 1 = 1$, $1 + 0 = 1$, $1 + 1 = 0$ (soma módulo 2),
- $0 \cdot 0 = 0$, $0 \cdot 1 = 0$, $1 \cdot 0 = 0$, $1 \cdot 1 = 1$ (produto usual).

Verifica-se que $\mathbb{F}_2$ é um corpo (ver Exercício 1.2).

Exercícios

Exercício 1.1. Prove a Proposição 1.2.3.

Exercício 1.2. Mostre que $\mathbb{F}_2$ é um corpo (ver Exemplo 1.2.14).

Exercício 1.3. Mostre que todo anel em que 0 é invertível é trivial.

Capítulo 2

Sistemas Lineares e Matrizes

Nesse capítulo estudaremos sistemas lineares do ponto de vista teórico e prático. Introduziremos matrizes como uma ferramenta natural para realizar tal estudo, e estudaremos suas propriedades básicas.

2.1 A representação matricial de um sistema linear

No Brasil, é usual um estudo introdutório sobre sistemas lineares de números reais.

Um sistema linear de números reais é, intuitivamente, uma coleção finita de equações da forma:

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

Em que cada a_{ij} e cada b_i é um número real, para $1 \leq i \leq m$ e $1 \leq j \leq n$. Os números a_{ij} são chamados de *coeficientes* do sistema, e os números b_i são chamados de *termos independentes* do sistema.

Caso todos os b_i sejam iguais a zero, dizemos que o sistema é dito *homogêneo.*

Os símbolos $x_1, x_2, \ldots, x_n$ são chamados de *variáveis* do sistema. Resolver um sistema linear consiste em determinar (todos) os valores reais para as variáveis $x_1, x_2, \ldots, x_n$ que satisfaçam todas as equações do sistema.

Exemplo 2.1.1. Considere o seguinte sistema linear:

$$\begin{cases} 2x_1 + 3x_2 = 5 \\ x_1 - x_2 = 1 \end{cases} \tag{2.2}$$

Nesse caso, temos $m = 2$, $n = 2$, $a_{11} = 2$, $a_{12} = 3$, $a_{21} = 1$, $a_{22} = -1$, $b_1 = 5$ e $b_2 = 1$. As variáveis do sistema são x_1 e x_2. O leitor pode verificar, por meios elementares, que a única solução do sistema é dada por $(x_1, x_2) = (2, 1)$.

Enfatizamos que não é um requisito para a leitura deste texto o conhecimento prévio de técnicas para resolver sistemas lineares, uma vez que um dos objetivos deste capítulo é apresentar técnicas gerais que resolvam qualquer sistema linear.

Aproveitando o exemplo anterior, considere o seguinte sistema linear:

$$\begin{cases} 2y_1 + 3y_2 = 5 \\ y_1 - y_2 = 1 \end{cases} \tag{2.3}$$

O sistema linear da Equação (2.3) é essencialmente o mesmo que o sistema linear da Equação (2.2), apenas com as variáveis renomeadas. Por óbvio, as soluções de ambos os sistemas devem ser as mesmas. Por isso, ao estudar soluções de sistemas lineares, nos preocupamos apenas com os coeficientes dos mesmos.

Assim, para estudar sistemas lineares, podemos definir formalmente objetos matemáticos que agem como "tabelas" de números. Tais objetos são as *matrizes.*

Definição 2.1.2. Seja A um anel e m, n inteiros positivos. Uma *matriz de ordem* $m \times n$ (de m linhas e n colunas) com entradas em A é uma família $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ de elementos de A.

Podemos ainda a denotar por:

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

Ou ainda, por:

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

E, finalmente, quando m, n são subentendidos, podemos escrever $A = (a_{ij})_{i,j}$.

O conjunto de todas as matrizes de ordem $m \times n$ com entradas em A é denotado por $M_{m \times n}(A)$. Quando $m = n$, dizemos que a matriz é *quadrada* de ordem n e denotamos por $M_n(A)$.

Os elementos a_{ij} são chamados de *coeficientes* de A.

Neste texto, daremos particular atenção para o corpo $A = \mathbb{R}$, ou, com maior generalidade, para corpos.

Exemplo 2.1.3. Seja $(a_{ij})_{i,j} \in M_2(\mathbb{R})$ a matriz dada por:

$$A = \begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix}$$

Então seus coeficientes são $a_{11} = 2$, $a_{12} = 3$, $a_{21} = 1$ e $a_{22} = -1$.

Neste texto, estudaremos sistemas lineares sobre corpos, com ênfase especial no corpo dos números reais $\mathbb{R}$. O estudo de sistemas lineares sobre corpos, conforme veremos, tem relação direta com o estudo do que é conhecido por *espaço vetorial*, e este é objeto da *Álgebra Linear*.

O estudo de sistemas lineares com coeficientes em anéis é mais geral, e, apesar de possuir muitas semelhanças com o que será apresentado, possui também muitas características próprias estudadas na Teoria de Módulos. Tal estudo não é objeto deste texto.

Definição 2.1.4. Seja $\mathbb{K}$ um corpo e $A = (a_{ij})_{i,j} \in M_{m \times n}(\mathbb{K})$ uma matriz. O *sistema linear homogêneo associado a* A é o sistema linear dado por:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \tag{2.4}$$

Tal sistema linear é denotado por $AX = 0$. O *conjunto solução do sistema linear homogêneo de* A é o conjunto de todas as n-uplas $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ que satisfazem o sistema linear da Equação (2.4).

Se, adicionalmente, $b = (b_1, b_2, \ldots, b_m) \in \mathbb{K}^m$, o *sistema linear associado a* A *e* b é o sistema linear dado por:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.5}$$

Tal sistema linear é denotado por $AX = b$.

O *conjunto solução do sistema linear de* A *com coeficientes independentes* b é o conjunto de todas as n-uplas $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ que satisfazem o sistema linear da Equação (2.5).

É imediato que todo sistema linear pode ser expresso na forma matricial, como acima. Assim, o estudo de matrizes pode ser uma ferramenta poderosa no estudo de sistemas lineares.

Exemplo 2.1.5. Seja $(a_{ij})_{i,j} \in M_2(\mathbb{R})$ a matriz dada por:

$$A = \begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix}$$

O sistema linear homogêneo associado a A é o sistema linear dado por:

$$\begin{cases} 2x_1 + 3x_2 = 0 \\ x_1 - x_2 = 0 \end{cases}$$

Conforme vimos, o produto matricial é uma operação que nasce naturalmente do estudo de sistemas lineares.

Outra operação pertinente é a soma de matrizes, motivada pelo seguinte fato:

Se $A, B \in M_{m\times n}(\mathbb{K})$ são matrizes e $x \in \mathbb{K}^n$ é solução dos sistemas lineares homogêneos $AX = 0$ e $BX = 0$, então, somando as equações linha-a-linha, x será solução do sistema linear resultante. Os coeficientes desse sistema são dados pela soma dos coeficientes de A e B.

Definição 2.1.6. Sejam $A = (a_{ij})_{i,j} \in M_{m\times n}(\mathbb{K})$ e $B = (b_{ij})_{i,j} \in M_{m\times n}(\mathbb{K})$ matrizes. A *soma de matrizes* $A + B$ é a matriz dada como a seguir.

$$A + B = (a_{ij} + b_{ij})_{i,j} \in M_{m\times n}(\mathbb{K})$$

Ou seja:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

2.2 Sistemas Equivalentes

Resgatemos conhecimentos elementares sobre resolução de equações simples: considerando nosso universo de discurso o conjunto dos números reais, como se obtém uma solução para a equação $3x + 2 = 5$? Ora, transformamos tal expressão na equação $3x = 3$, e, após, na equação $x = 1$.

Por que tal processo é válido? Ora, se x é um número real tal que $3x + 2 = 5$, então $3x = 3$, e, se x é um número real tal que $3x = 3$, então $x = 1$. Ou seja, se x é uma solução da equação $3x + 2 = 5$, então necessariamente x é 1. Por outro lado, tais operações são *reversíveis*: se $x = 1$, então $3x = 3$, e, se $3x = 3$, então $3x + 2 = 5$. Assim, podemos dizer, ao menos intuitivamente, que as equações $3x + 2 = 5$, $3x = 3$ e $x = 1$ são *equivalentes*, ou seja, possuem as mesmas soluções.

Perceba que, de modo geral, nosso método de resolução de equações é baseado em transformar uma equação em outra, sucessivamente, preservando todas as suas soluções, de modo a torná-la cada vez mais simples, até que se obtenha uma equação que seja trivialmente resolvida. Como fazer isso de modo generalizado para sistemas lineares? Nesta seção, estudaremos este tema.

O primeiro passo nesta direção é, dentro do nosso formalismo, definir a equivalência entre dois sistemas lineares.

Definição 2.2.1. Seja R um corpo, $A, B \in M_{m\times n}(\mathbb{K})$ matrizes e $b, c \in \mathbb{K}^m$. Dizemos que os sistemas lineares $AX = b$ e $BX = c$ são *equivalentes* se o conjunto solução da equação $AX = b$ é igual ao conjunto solução da equação $BX = c$.

Continuando nosso discurso, consideremos um sistema linear homogêneo qualquer, por exemplo, em $\mathbb{R}$:

$$\begin{cases} 2x_1 + 3x_2 + 5x_3 = 0 \\ 4x_1 + 6x_2 - x_3 = 0 \end{cases} \tag{2.6}$$

Quaisquer $(x_1, x_2, x_3) \in \mathbb{R}^3$ que satisfaça tal sistema satisfará também as equações $2x_1 + 3x_2 + 5x_3 = 0$ e $4x_1 + 6x_2 - x_3 = 0$, bem como a soma destas, $6x_1 + 9x_2 + 4x_3 = 0$, bem como o dobro da primeira, $4x_1 + 6x_2 + 10x_3 = 0$, bem como o dobro da primeira menos o triplo da segunda, $-2x_1 - 3x_2 + 8x_3 = 0$.

Vamos pensar em como sistematizar tal raciocínio. Olhemos para a matriz dos coeficientes do sistema da Equação (2.6):

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

Se montarmos um novo sistema obtido pela soma de múltiplos de linhas de A, obtemos um novo sistema linear homogêneo tal que toda solução do sistema original é também solução do novo sistema.

Segundo nosso exemplo, qualquer solução do sistema da Equação (2.6) é também solução do sistema homogêneo associado à matriz B dada por:

$$B = \begin{pmatrix} 6 & 9 & 4 \\ 4 & 6 & 10 \\ -2 & -3 & 8 \end{pmatrix}$$

Para obter a primeira linha de B, somamos a primeira linha de A com a segunda, e para obter a segunda linha de B, multiplicamos a primeira linha de A por 2. Para obter a terceira linha de B, subtraímos o dobro da primeira linha de A com o triplo da segunda linha de A.

Tais operações consistem em obter o que chamamos de *combinações lineares* das linhas de A.

Para melhor expressar tais operações, definamos o seguinte:

Definição 2.2.2. Seja $\mathbb{K}$ um corpo e n um inteiro positivo. Sejam $v = (a_1, \ldots, a_n) \in \mathbb{K}^n$ e $u = (c_1, \ldots, c_n) \in \mathbb{K}^n$.

Define-se a *soma* de v com u como $v + u = (a_1 + c_1, \ldots, a_n + c_n)$.

Define-se o *produto por escalar* de v por um elemento $b \in \mathbb{K}$ como $b \cdot v = (ba_1, \ldots, ba_n)$.

O elemento nulo de $\mathbb{K}^n$ é o elemento $0 = (0, 0, \ldots, 0)$.

Se $v = (a_1, \ldots, a_n) \in \mathbb{K}^n$, o *oposto* de v é o elemento $-v = (-a_1, -a_2, \ldots, -a_n)$.

Se $v, w \in \mathbb{K}^n$, a *diferença* de v com w é o elemento $v - w = v + (-w)$.

Elementos de $\mathbb{K}^n$ são chamados de *vetores*.

Proposição 2.2.3. Sejam, $u, v, w \in \mathbb{K}^n$ e $\alpha, \beta \in \mathbb{K}$. Então:

(i) $u + (v + w) = (u + v) + w$;

(ii) $u + v = v + u$;

(iii) $u + 0 = u$;

(iv) $u + (-u) = 0$;

(v) $\alpha \cdot (\beta \cdot u) = (\alpha\beta) \cdot u$;

(vi) $1 \cdot u = u$;

(vii) $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$;

(viii) $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$.

Demonstração. Escreva $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$ e $w = (w_1, \ldots, w_n)$. Então:

(i) $u+(v+w) = (u_i)_{i=1}^n+(v_i+w_i)_{i=1}^n = (u_i+(v_i+w_i))_{i=1}^n = ((u_i+v_i)+w_i)_{i=1}^n = (u_i+v_i)_{i=1}^n+(w_i)_{i=1}^n = (u + v) + w$;

(ii) $u + v = (u_i)_{i=1}^n + (v_i)_{i=1}^n = (u_i + v_i)_{i=1}^n = (v_i + u_i)_{i=1}^n = (v_i)_{i=1}^n + (u_i)_{i=1}^n = v + u$;

(iii) $u + 0 = (u_i)_{i=1}^n + (0, \ldots, 0) = (u_i + 0)_{i=1}^n = (u_i)_{i=1}^n = u$;

(iv) $u + (-u) = (u_i)_{i=1}^n + (-u_i)_{i=1}^n = (u_i + (-u_i))_{i=1}^n = (0, \ldots, 0) = 0$;

(v) $\alpha \cdot (\beta \cdot u) = \alpha \cdot (\beta u_i)_{i=1}^n = (\alpha(\beta u_i))_{i=1}^n = ((\alpha\beta)u_i)_{i=1}^n = (\alpha\beta) \cdot (u_i)_{i=1}^n = (\alpha\beta) \cdot u$;

(vi) $1 \cdot u = (1u_i)_{i=1}^n = (u_i)_{i=1}^n = u$;

(vii) $\alpha \cdot (u + v) = \alpha \cdot (u_i + v_i)_{i=1}^n = (\alpha(u_i + v_i))_{i=1}^n = (\alpha u_i + \alpha v_i)_{i=1}^n = (\alpha u_i)_{i=1}^n + (\alpha v_i)_{i=1}^n = \alpha \cdot u + \alpha \cdot v$;

(viii) $(\alpha + \beta) \cdot u = ((\alpha + \beta)u_i)_{i=1}^n = (\alpha u_i + \beta u_i)_{i=1}^n = (\alpha u_i)_{i=1}^n + (\beta u_i)_{i=1}^n = \alpha \cdot u + \beta \cdot u$.

□

Definição 2.2.4. Se $v_1, \ldots, v_n \in \mathbb{K}^n$, uma *combinação linear de* $v_1, \ldots, v_n$ é uma n-upla da forma $\sum_{i=1}^n c_i v_i$, com $c_i \in \mathbb{K}$.

Com essa notação, qual a forma geral para se obter, a partir de uma matriz A uma nova matriz C formada de combinações lineares de linhas de A? Tal matriz C, como A, deve possuir n colunas. Porém, o número de linhas de C não possui limite. Imaginemos que C tem p linhas.

Ora, sendo $A = (a_{ij})_{i,j} \in M_{m\times n}(A)$, cada linha de A pode ser encarada como uma n-upla $a_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \in A^n$.

Se $1 \leq k \leq p$, a k-ésima linha de C é uma combinação linear de $a_1, \ldots, a_n$, ou seja, é da forma $\sum_{i=1}^m b_{ki} a_i$, com $b_{ki} \in \mathbb{K}$. A i-ésima linha de C, é, então, $(\sum_{i=1}^m b_{ki} a_{i1}, \ldots, \sum_{i=1}^m b_{ki} a_{in})$. Explicitamente:

$$C = \begin{pmatrix} \sum_{i=1}^m b_{1i} a_{i1} & \cdots & \sum_{i=1}^m b_{1i} a_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^m b_{pi} a_{i1} & \cdots & \sum_{i=1}^m b_{pi} a_{in} \end{pmatrix}$$

Ou seja, escolhendo-se uma matriz $B = (b_{ki})_{k,i} \in M_{p\times m}(A)$ e operando os elementos b_{ik} como acima, obtemos uma matriz C de ordem $p \times n$ cujas linhas são combinações lineares das linhas de A, e todas tais matrizes C são equivalentes a A.

Seria útil formalizar tal operação. Tal operação é muito conhecida e chamada de *produto matricial*.

Definição 2.2.5. Sejam m, n, p inteiros positivos, R um anel, e $A = (a_{ij})_{i,j} \in M_{m\times n}(R)$ e $B = (b_{ki})_{k,i} \in M_{p\times m}(R)$ matrizes. O *produto matricial* de B por A, denotado por BA, é a matriz $C \in M_{p\times n}(R)$ dada por:

$$C = BA = \begin{pmatrix} \sum_{i=1}^m b_{1i} a_{i1} & \cdots & \sum_{i=1}^m b_{1i} a_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^m b_{pi} a_{i1} & \cdots & \sum_{i=1}^m b_{pi} a_{in} \end{pmatrix}$$

Ou seja, $C = (\sum_{i=1}^m b_{ki} a_{ij})_{k,j} \in M_{p\times n}(R)$.

Conforme esperado, o seguinte resultado é, então, verdadeiro:

Proposição 2.2.6. Seja $\mathbb{K}$ um corpo, $A \in M_{m\times n}(\mathbb{K})$ e $B \in M_{p\times m}(\mathbb{K})$ matrizes. Então toda solução do sistema linear homogêneo $AX = 0$ é também solução do sistema linear homogêneo $BAX = 0$.

Demonstração. Escreva $A = (a_{ij})_{i,j} \in M_{m\times n}(\mathbb{K})$ e $B = (b_{ki})_{k,i} \in M_{p\times m}(\mathbb{K})$. Seja $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ uma solução do sistema linear homogêneo $AX = 0$.

Devemos ver que para todo k tal que $1 \leq k \leq p$, vale $\sum_{j=1}^n \left(\sum_{i=1}^m b_{ki} a_{ij}\right) x_j = 0$.

Ora, dado i tal que $1 \leq i \leq m$, temos que $\sum_{j=1}^n a_{ij} x_j = 0$. Daí, temos:

$$\sum_{j=1}^n \left(\sum_{i=1}^m b_{ki} a_{ij}\right) x_j = \sum_{j=1}^n \left(\sum_{i=1}^m b_{ki} a_{ij} x_j\right) = \sum_{i=1}^m \left(\sum_{j=1}^n b_{ki} a_{ij} x_j\right) = \sum_{i=1}^m b_{ki} \left(\sum_{j=1}^n a_{ij} x_j\right) = \sum_{i=1}^m b_{ki} 0 = 0.$$

□

Note, ainda, que a notação $AX = b$ não é por acaso.

Proposição 2.2.7. Seja $\mathbb{K}$ um corpo, $A \in M_{m\times n}(\mathbb{K})$ e $B \in M_{p\times m}(\mathbb{K})$ matrizes.

Seja $b = (b_1, b_2, \ldots, b_m) \in \mathbb{K}^m$ e $x = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$. Considere X a matriz coluna dada por $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ e B a matriz coluna dada por $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$.

Então x é solução do sistema linear $AX = b$ se, e somente se, $AX = B$, ou seja, se, e somente se, sendo $A = (a_{ij})_{i,j}$:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Demonstração. Note que:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}$$

Logo, $AX = B$ se, e somente se para cada i tal que $1 \leq i \leq m$, temos $\sum_{j=1}^n a_{ij}x_j = b_i$, ou seja, se, e somente se x é solução do sistema linear $AX = b$. $\square$

A operação de produto de matrizes é, de fato, uma operação de recombinação de linhas, conforme ilustrado a seguir:

Proposição 2.2.8. Sejam $A = (a_{ij})_{i,j} \in M_{m\times n}(\mathbb{K})$ e $B = (b_{ki})_{k,i} \in M_{p\times m}(\mathbb{K})$. Seja $C = BA = (c_{kj})_{k,j} \in M_{p\times n}(\mathbb{K})$ o produto matricial de B por A. Então cada linha de C é uma combinação linear das linhas de A.

Mais especificamente, sendo $L_1, \ldots, L_m$ as linhas de A, e $K_1, \ldots, K_p$ as linhas de B, temos que cada linha C_k de C pode ser escrita como:

$$C_k = \sum_{i=1}^m b_{ki}L_i.$$

Reciprocamente, se $C \in M_{p\times m}(\mathbb{K})$ é uma matriz tal que cada linha de C é uma combinação linear das linhas de A, então existe uma matriz $B \in M_{p\times m}(\mathbb{K})$ tal que $C = BA$.

Demonstração. Utilizando a notação do enunciado, dados k, j tais que $1 \leq k \leq p$ e $1 \leq j \leq n$, temos que o elemento c_{kj} de C é dado por:

$$c_{kj} = \sum_{i=1}^m b_{ki}a_{ij}.$$

Tal elemento é o j-ésimo elemento de C_k e de $\sum_{i=1}^m b_{ki}L_i$, logo, $C_k = \sum_{i=1}^m b_{ki}L_i$.

Reciprocamente, suponha que C é uma matriz tal que cada linha de C é uma combinação linear das linhas de A. Então, para cada k tal que $1 \leq k \leq p$, existem $b_{k1}, b_{k2}, \ldots, b_{km} \in \mathbb{K}$ tais que $C_k = \sum_{i=1}^m b_{ki}L_i$.

Sendo $B = (b_{ki})_{k,i} \in M_{p\times m}(\mathbb{K})$, temos que $C = BA$. $\square$

Exemplo 2.2.9. Seja $\mathbb{R}$ o corpo dos números reais, $A = \begin{pmatrix} 2 & 3 & 5 \\ 4 & 6 & -1 \end{pmatrix} \in M_{2\times 3}(\mathbb{R})$ e $B = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ -1 & -3 \end{pmatrix} \in M_{3\times 2}(\mathbb{R})$.

Então:

$$BA = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ -1 & -3 \end{pmatrix} \begin{pmatrix} 2 & 3 & 5 \\ 4 & 6 & -1 \end{pmatrix} = \begin{pmatrix} 6 & 9 & 4 \\ 4 & 6 & 10 \\ -10 & -21 & -2 \end{pmatrix}$$

Sendo $L_1 = (2, 3, 5)$ e $L_2 = (4, 6, -1)$ as linhas de A, temos que as linhas de BA são dadas por $1L_1 + 1L_2$, $2L_1 + 0L_2$ e $-1L_1 - 3L_2$.

2.3 Operações com matrizes

Nesta seção, faremos um interlúdio sobre operações gerais de matrizes.

Proposição 2.3.1. Seja R um anel e m, n inteiros positivos. Então $M_{m\times n}(R)$ é, com a soma de matrizes, um *grupo abeliano*, ou seja, satisfaz as seguintes propriedades. Para todos $A, B, C \in M_{m\times n}(R)$:

(i) A soma é associativa, ou seja, $(A + B) + C = A + (B + C)$.

(ii) A soma é comutativa, ou seja, $A + B = B + A$.

(iii) Existe um elemento neutro para a soma, a matriz nula $0_{m\times n} = (0)_{i,j}$.

(iv) Todo elemento tem oposto aditivo, a matriz $-A$ dada por $(-a_{ij})_{i,j}$.

Demonstração. Sejam $A = (a_{ij})_{i,j}, B = (b_{ij})_{i,j}, C = (c_{ij})_{i,j} \in M_{m\times n}(R)$.

(i) Temos que

$$(A+B)+C = (a_{ij}+b_{ij})_{i,j} + (c_{ij})_{i,j} = ((a_{ij}+b_{ij})+c_{ij})_{i,j} = (a_{ij}+(b_{ij}+c_{ij}))_{i,j} = A+(B+C).$$

(ii) Temos que

$$A+B = (a_{ij}+b_{ij})_{i,j} = (b_{ij}+a_{ij})_{i,j} = B+A.$$

(iii) Veja que

$$A+0_{m\times n} = (a_{ij}+0)_{i,j} = (a_{ij})_{i,j} = A.$$

(iv) Note que

$$A+(-A) = (a_{ij}+(-a_{ij}))_{i,j} = (0)_{i,j} = 0_{m\times n}.$$

□

A operação de produto, por sua vez, possui propriedades típicas das de multiplicação em um anel, porém, ela não está definida para quaisquer duas matrizes: os tamanhos devem ser compatíveis.

Proposição 2.3.2. Seja R um anel e m, n, p, q inteiros positivos. Seja $A, A' \in M_{m\times n}(R)$, $B, B' \in M_{p\times m}(R)$ e $C \in M_{q\times p}(R)$ matrizes. Segue que:

(i) O produto é associativo, ou seja, $(CB)A = C(BA)$.

(ii) O produto é distributivo em relação à soma, ou seja, $B(A+A') = BA+BA'$ e $(B+B')A = BA+BA'$.

Demonstração. Sejam $A = (a_{ij})_{i,j} \in M_{m\times n}(R)$, $B = (b_{ki})_{k,i} \in M_{p\times m}(R)$, $C = (c_{lk})_{l,k} \in M_{q\times p}(R)$. Temos que:

$$(CB)A = (\sum_{k=1}^{p} c_{lk}b_{ki})_{l,i}(\sum_{j=1}^{n} a_{ij})_{i,j} = (\sum_{j=1}^{n}\sum_{k=1}^{p} c_{lk}b_{ki}a_{ij})_{l,j}.$$

Por outro lado, temos que:

$$C(BA) = (c_{lk})_{l,k}(\sum_{i=1}^{m} b_{ki}a_{ij})_{k,j} = (\sum_{k=1}^{p}\sum_{j=1}^{n} c_{lk}b_{ki}a_{ij})_{l,j}.$$

Assim, segue a (i).

Para a (ii), temos que, sendo $A' = (a'_{ij})_{i,j} \in M_{m\times n}(R)$:

$$B(A+A') = (b_{ki})_{k,i}(a_{ij}+a'_{ij})_{i,j} = (\sum_{i=1}^{m} b_{ki}(a_{ij}+a'_{ij}))_{k,j} = (\sum_{i=1}^{m} b_{ki}a_{ij} + \sum_{i=1}^{m} b_{ki}a'_{ij})_{k,j} = BA+BA'$$

Analogamente, $(B+B')A = BA+B'A$. □

Em particular, matrizes quadradas formam um anel.

Proposição 2.3.3 (Anel de matrizes)**.** Seja R um anel e n um inteiro positivo. Então $M_n(R)$, com a soma e produto de matrizes, é um anel, cuja identidade multiplicativa é a matriz identidade $I_n = (\delta_{ij})_{i,j}$, onde δ_{ij} é o *símbolo de Kronecker*, ou seja, $\delta_{ij} = 1$ se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$. Mais explicitamente:

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

Ainda mais: se m, n são inteiros positivos e $A \in M_{m\times n}(R)$, então $I_mA = A$ e $AI_n = A$.

Demonstração. Note que, para $A = (a_{ij})_{i,j} \in M_{m\times n}(R)$, temos que $I_m A$ é uma matriz de dimensões $m \times n$. Se i, j são tais que $1 \le i \le m$ e $1 \le j \le n$ o elemento ij de $I_m A$ é dado por $\sum_{k=1}^{n} \delta_{ik} a_{kj} = a_{ij}$.

Analogamente, AI_n é uma matriz de dimensões $m\times n$ e o elemento ij de AI_n é dado por $\sum_{k=1}^{m} a_{ik}\delta_{kj} = a_{ij}$. □

O produto de matrizes não é comutativo.

Exemplo 2.3.4. Seja $A = (a_{ij})_{i,j} \in M_2(\mathbb{R})$ e $B = (b_{ij})_{i,j} \in M_2(\mathbb{R})$ matrizes dadas por:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Então, temos que:

$$AB = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Note que $AB \neq BA$.

Além disso, há matrizes que não são invertíveis.

Exemplo 2.3.5. Considere a matriz $A = (a_{ij})_{i,j} \in M_2(\mathbb{R})$ dada por:

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

Afirmamos que A não é invertível. De fato, suponha por absurdo que exista uma matriz B tal que $AB = BA = I_2$. Escreva $B = (b_{ij})_{i,j} \in M_2(\mathbb{R})$. Então, temos que:

$$I_n = AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ 0 & 0 \end{pmatrix}.$$

Assim:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = AB = \begin{pmatrix} b_{11} & b_{12} \\ 0 & 0 \end{pmatrix}$$

Em particular, $0 = 1$, o que é absurdo.

2.4 Resolução de Sistemas Homogêneos

Com já o básico operacional de matrizes desenvolvido, ganhamos algumas ferramentas para estudar sistemas lineares.

Proposição 2.4.1. Seja $\mathbb{K}$ um corpo e $A \in M_{m\times n}(\mathbb{K})$ uma matriz. Se $B \in M_m(\mathbb{K})$ é uma matriz invertível, então o sistema linear homogêneo $AX = 0$ é equivalente ao sistema linear homogêneo $BAX = 0$.

Demonstração. Já vimos que toda solução do sistema linear homogêneo $AX = 0$ é também solução do sistema linear homogêneo $BAX = 0$.

Sendo B^{-1} a inversa de B, segue que toda solução do sistema linear homogêneo $BAX = 0$ é também solução do sistema linear homogêneo $B^{-1}BAX = 0$. Como $B^{-1}BA = A$, temos que $B^{-1}BAX = 0$ é equivalente a $AX = 0$. □

Agora iniciaremos o estudo direto da resolução de sistemas lineares homogêneos. Mais tarde, veremos como resolver sistemas lineares não homogêneos a partir das técnicas que desenvolveremos aqui.

Voltemos a nossa discussão intuitiva. Considere o seguinte sistema linear homogêneo:

$$\begin{cases} 2x_1 + 3x_2 = 0 \\ x_1 - x_2 = 0 \end{cases}$$

Sistemas homogêneos sempre tem a chamada *solução trivial*: a atribuição do valor 0 a cada variável. Será que o sistema acima possui outras soluções?

Vamos considerar três operações que podemos realizar nesse sistema:

A primeira: multiplicar alguma das linhas por um número não nulo. É imediato que, por exemplo, ao multiplicar a primeira linha por $\frac{1}{2}$, obtemos um sistema equivalente.

$$\begin{cases} x_1 + \frac{3}{2}x_2 = 0 \\ x_1 - x_2 = 0 \end{cases}$$

A segunda: trocar linhas de posição. É imediato que o sistema acima é equivalente ao seguinte:

$$\begin{cases} x_1 - x_2 = 0 \\ x_1 + \frac{3}{2}x_2 = 0 \end{cases}$$

E a terceira: trocar uma linha por ela própria, mais um múltiplo de outra. Por exemplo, o sistema acima é equivalente ao seguinte:

$$\begin{cases} 3x_1 + 2x_2 = 0 \\ x_1 + \frac{3}{2}x_2 = 0 \end{cases}$$

Para se obter esse sistema, somamos, na primeira linha, o dobro da segunda. Para obter novamente a primeira linha original a partir do novo sistema, basta subtrair, da nova primeira linha, o dobro da segunda. Logo, os sistemas são equivalentes.

Nossa afirmação é que, utilizando sistematicamente essas três operações, que chamaremos de *operações elementares nas linhas de um sistema linear*, podemos resolver completamente qualquer sistema linear. No nosso exemplo, olhando novamente para o sistema original, e somando na primeira linha o oposto da segunda (o produto da segunda por -1), obtemos o seguinte sistema:

$$\begin{cases} 0x_1 + \frac{5}{2}x_2 = 0 \\ x_1 - x_2 = 0 \end{cases}$$

Multiplicando a primeira linha por $\frac{2}{5}$ e trocando as linhas de posição, obtemos o seguinte sistema:

$$\begin{cases} x_1 - x_2 = 0 \\ 0x_1 + x_2 = 0 \end{cases}$$

Finalmente, somando na primeira linha, a segunda, obtemos o seguinte sistema:

$$\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

O que mostra que a única solução do sistema original é a solução trivial.

Vamos, nesta seção, sistematizar e formalizar este raciocínio a partir da linguagem de matrizes.

Definição 2.4.2. Sejam m, n inteiros positivos. Chamamos de *operações elementares em linhas* em $M_{m\times n}(\mathbb{K})$ as seguintes operações.

1. Para algum $l \in \{1, \ldots, m\}$ e $\lambda \in \mathbb{K}$ não nulo, $e^1_{l\lambda}((a_{ij})_{i,j}) = (b_{ij})_{i,j}$ onde:

$$b_{ij} = \begin{cases} \lambda a_{ij} & \text{se } i = l \\ a_{ij} & \text{caso contrário.} \end{cases}$$

2. Para $l, l' \in \{1, \ldots, m\}$, $e^2_{ll'}((a_{ij})_{i,j}) = (b_{ij})_{i,j}$ onde:

$$b_{ij} = \begin{cases} a_{l'j} & \text{se } i = l \\ a_{lj} & \text{se } i = l' \\ a_{ij} & \text{caso contrário.} \end{cases}$$

3. Para $l, l' \in \{1, \ldots, m\}$ e $\lambda \in \mathbb{K}$, $e^3_{ll'\lambda}((a_{ij})_{i,j}) = (b_{ij})_{i,j}$ onde:

4.

$$b_{ij} = \begin{cases} a_{lj} + \lambda a_{l'j} & \text{se } i = l \\ a_{ij} & \text{caso contrário.} \end{cases}$$

Tais operações elementares correspondem a produtos por certas matrizes.

Definição 2.4.3. Seja m um inteiro positivo. Chamamos de *matrizes elementares* de $M_m(\mathbb{K})$ as seguintes matrizes.

1. Para algum $l \in \{1,\dots,m\}$ e $\lambda \in \mathbb{K}$ não nulo, $E^1_{l\lambda} = (b_{ij})_{i,j}$, onde:

$$b_{ij} = \begin{cases} \lambda & \text{se } i = j = l \\ 1 & \text{se } i = j \neq l \\ 0 & \text{caso contrário.} \end{cases}$$

2. Para $l, l' \in \{1,\dots,m\}$, $E^2_{ll'} = (b_{ij})_{i,j}$, onde:

$$b_{ij} = \begin{cases} 1 & \text{se } i = j \notin \{l, l'\} \\ 1 & \text{se } i = l, j = l' \\ 1 & \text{se } i = l', j = l \\ 0 & \text{caso contrário.} \end{cases}$$

3. Para $l, l' \in \{1,\dots,m\}$ e $\lambda \in \mathbb{K}$, $E^3_{ll'\lambda} = (b_{ij})_{i,j}$, onde:

$$b_{ij} = \begin{cases} 1 & \text{se } i = j \\ \lambda & \text{se } i = l, j = l' \\ 0 & \text{caso contrário.} \end{cases}$$

Lema 2.4.4. Sejam m, n inteiros positivos e $A \in M_{m\times n}(\mathbb{K})$. Então:

1. Dado $l \in \{1,\dots,m\}$ e $\lambda \in \mathbb{K}$ não nulo, $E^1_{l\lambda}A = e^1_{l\lambda}(A)$.
2. Dado $l, l' \in \{1,\dots,m\}$, $E^2_{ll'}A = e^2_{ll'}(A)$.
3. Dado $l, l' \in \{1,\dots,m\}$ e $\lambda \in \mathbb{K}$, $E^3_{ll'\lambda}A = e^3_{ll'\lambda}(A)$.

Demonstração. Sejam i, j tais que $1 \leq i \leq m$ e $1 \leq j \leq n$. Escreva $A = (a_{ij})_{i,j}$ e, em cada caso, seja $(b_{ij})_{i,j}$ a matriz elementar em questão.

1. Se $i \neq l$, o elemento ij de $E^1_{l\lambda}A$ é dado por $\sum_{k=1}^n b_{ik}a_{kj} = a_{ij}$. Já se $i = l$, o elemento ij de $E^1_{l\lambda}A$ é dado por $\sum_{k=1}^n b_{ik}a_{kj} = \lambda\lambda a_{lj}$. Assim, $E^1_{l\lambda}A = e^1_{l\lambda}(A)$.
2. Se $i \notin \{l, l'\}$, o elemento ij de $E^2_{ll'}A$ é dado por $\sum_{k=1}^n b_{ik}a_{kj} = a_{ij}$. Se $i = l$, o elemento ij de $E^2_{ll'}A$ é dado por $\sum_{k=1}^n b_{lk}a_{kj} = a_{l'j}$. Finalmente, se $i = l'$, o elemento ij de $E^2_{ll'}A$ é dado por $\sum_{k=1}^n b_{l'k}a_{kj} = a_{lj}$. Assim, $E^2_{ll'}A = e^2_{ll'}(A)$.
3. Se $i \neq l$, o elemento ij de $E^3_{ll'\lambda}A$ é dado por $\sum_{k=1}^n b_{ik}a_{kj} = a_{ij}$. Se $i = l$, o elemento ij de $E^3_{ll'\lambda}A$ é dado por $\sum_{k=1}^n b_{lk}a_{kj} = a_{ij} + \lambda a_{l'j}$. Assim, $E^3_{ll'\lambda}A = e^3_{ll'\lambda}(A)$.

$\square$

Lema 2.4.5. Seja m um inteiro positivo e $\mathbb{K}$ um corpo. As matrizes elementares de $M_m(\mathbb{K})$ são invertíveis.

Demonstração. Consideremos $E^1_{l\lambda}$, onde $l \in \{1,\dots,m\}$ e $\lambda \in \mathbb{K}$ não nulo. É imediato que $I_n = e^1_{l\lambda^{-1}}(E^1_{l\lambda})$, de modo que $I_n = E^1_{l\lambda^{-1}}E^1_{l\lambda}$. Da mesma forma, $I_n = E^1_{l\lambda}E^1_{l\lambda^{-1}}$

Agora consideremos $E^2_{ll'}$, onde $l, l' \in \{1,\dots,m\}$. É imediato que $I_n = e^2_{ll'}(E^2_{ll'})$, de modo que $I_n = E^2_{ll'}E^2_{ll'}$.

Finalmente, consideremos $E^3_{ll'\lambda}$, onde $l, l' \in \{1,\dots,m\}$ e $\lambda \in \mathbb{K}$. É imediato que $I_n = e^3_{ll'(-\lambda)}(E^3_{ll'\lambda})$, de modo que $I_n = E^3_{ll'(-\lambda)}E^3_{ll'\lambda}$. Da mesma forma, $I_n = E^3_{ll'\lambda}E^3_{ll'(-\lambda)}$. $\square$

Corolário 2.4.6. Sejam $A, B \in M_{m\times n}(\mathbb{K})$ matrizes tais que A é obtida a partir de B aplicando-se sucessivas operações elementares em linhas. Então os sistemas lineares homogêneos $AX = 0$ e $BX = 0$ são equivalentes.

Estudemos agora o seguinte sistema:

$$\begin{cases} 2x_1 + 6x_2 + 4x_3 + x_4 = 0 \\ 3x_1 - 3x_2 - 4x_3 + x_4 = 0 \end{cases} \tag{2.7}$$

Vamos resolvê-lo utilizando operações elementares em linhas. A matriz associada a esse sistema é:

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

Multiplicando a primeira linha por $\frac{1}{2}$, obtemos:

$$A_1 = \begin{pmatrix} 1 & 3 & 2 & \frac{1}{2} \\ 3 & -3 & -4 & 1 \end{pmatrix}$$

Agora, subtraindo da segunda linha o triplo da primeira, obtemos:

$$A_2 = \begin{pmatrix} 1 & 3 & 2 & \frac{1}{2} \\ 0 & -15 & -12 & -\frac{3}{2} \end{pmatrix}$$

Multiplicando a segunda linha por $-\frac{1}{15}$, obtemos:

$$A_3 = \begin{pmatrix} 1 & 3 & 2 & \frac{1}{2} \\ 0 & 1 & \frac{4}{5} & \frac{1}{10} \end{pmatrix}$$

Finalmente, subtraindo da primeira linha o triplo da segunda, obtemos:

$$A_4 = \begin{pmatrix} 1 & 0 & -\frac{2}{5} & \frac{1}{5} \\ 0 & 1 & \frac{4}{5} & \frac{1}{10} \end{pmatrix}$$

Assim, o sistema da Equação (2.7) é equivalente ao seguinte sistema:

$$\begin{cases} x_1 - \frac{2}{5}x_3 + \frac{1}{5}x_4 = 0 \\ x_2 + \frac{4}{5}x_3 + \frac{1}{10}x_4 = 0 \end{cases} \tag{2.8}$$

Ou, de forma mais simples:

$$\begin{cases} x_1 = \frac{2}{5}x_3 - \frac{1}{5}x_4 \\ x_2 = -\frac{4}{5}x_3 - \frac{1}{10}x_4 \end{cases}$$

Escolhidos quaisquer valores reais de x_3 e x_4, e utilizando-se as expressões acima para x_1, x_2, obtemos uma solução do sistema original, e todas as soluções do sistema original são obtidas dessa forma. Em outras palavras, o conjunto solução do sistema da Equação (2.7) é dado por:

$$\left\{ \left(\frac{2}{5}x_3 - \frac{1}{5}x_4, -\frac{4}{5}x_3 - \frac{1}{10}x_4, x_3, x_4 \right) \mid x_3, x_4 \in \mathbb{R} \right\}$$

Nosso objetivo é mostrar que esse método é geral. Notemos que o sistema da Equação (2.8) é notável: basta observá-lo para, de imediato, escrever seu conjunto solução.

Definiremos, abaixo, uma classe de matrizes semelhantes a essa, e, a seguir, mostraremos que para qualquer matriz, existe uma sequência finita de operações elementares em linhas que a transforma em uma matriz dessa classe. A demonstração desse teorema será construtiva, de modo que ela nos dará o algoritmo necessário para resolver qualquer sistema linear homogêneo.

Definição 2.4.7. Seja m, n inteiros positivos e $\mathbb{K}$ um corpo.

Seja $A \in M_{m\times n}(\mathbb{K})$ e i entre 1 e m.

Dizemos que a i-ésima linha de A possui pivô se ela é não nula e $j \in \{1, \ldots, n\}$ tal que $a_{ij} = 1$ e $a_{i'j} = 0$ para todo $i' \neq i$, e, além disso, a_{ij} é o primeiro elemento não nulo da linha. Nesse caso, o elemento a_{ij} é chamado de *pivô* da i-ésima linha de A, e j é a *posição do pivô* da i-ésima linha de A.

Dizemos que uma matriz $A \in M_{m\times n}(\mathbb{K})$ está na *forma escalonada* se, para todo $i, i' \in \{1, \ldots, m\}$ com $i < i'$, valem as seguintes condições.

a) Se a i-ésima linha de A é nula, então a i'-ésima linha de A também é nula.

b) Se a i-ésima linha de A não é nula, então ela possui um pivô.

c) Caso as linhas i, i' de A não são nulas, e $j_i, j_{i'}$ são as posições dos pivôs dessas linhas, então $j_i < j_{i'}$.

Para $i \in \{1, \ldots, m\}$ com $i \leq S$, a posição (i, j_i) é dita ser um *pivô* de A.

Observe que a matriz identidade I_n e a matriz nula estão ambas na forma escalonada.
Como ocorre na discussão, acima, temos:

Teorema 2.4.8. Seja m, n inteiros positivos e $\mathbb{K}$ um corpo. Então, para toda matriz $A \in M_{m\times n}(\mathbb{K})$ na forma escalonada as soluções do sistema $AX = 0$ são dadas, segundo a notação acima, pelas n-uplas $(x_j : 1 \leq j \leq n)$, onde x_j é qualquer se para nenhum i, (i, j) é um pivô de A. Já se (i, j) é um pivô de A, então $x_j = \sum (-a_{ij'}x_{j'} : j < j'$, j' não é pivô de $A)$.

Teorema 2.4.9. Seja m, n inteiros positivos e $\mathbb{K}$ um corpo. Então, para toda matriz $A \in M_{m\times n}(\mathbb{K})$, existe uma sequência finita de operações elementares em linhas que transforma A em uma matriz na forma escalonada.

Demonstração. Provaremos por indução em n.

Se $n = 1$, então A é uma matriz coluna. Se A é nula, então A já está na forma escalonada.

Caso contrário, trocando duas linhas de posição, obtemos uma matriz coluna cujo primeiro elemento é não nulo. Multiplicando a primeira linha pelo inverso de seu novo elemento, obtemos uma matriz coluna cujo primeiro elemento é 1. Finalmente, somando em cada outra linha i não nula o múltiplo da primeira linha pelo oposto do elemento na posição $i1$, obtemos a matriz escalonada abaixo:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Para a hipótese de indução, suponha que toda matriz $A \in M_{m\times n}(\mathbb{K})$ pode ser escalonada por uma sequência finita de operações elementares em linhas.

Se $A \in M_{m\times(n+1)}(\mathbb{K})$ e $B \in M_{m\times n}(\mathbb{K})$ é a matriz obtida a partir de A eliminando-se a última coluna, então, por hipótese de indução, existe uma sequência finita de operações elementares em linhas que transforma B em uma matriz na forma escalonada. Perceba que se e é uma operação elementar em linhas e π é a operação que elimina de uma matriz sua última coluna, então $e \circ \pi(A) = \pi \circ e(A)$. Logo, sendo $e_1, \ldots e_l$ operações elementares em linhas tais que $e_1 \circ \cdots \circ e_l(B)$ é a matriz na forma escalonada, temos que $\pi \circ e_1 \circ \cdots \circ e_l(A) = e_1 \circ \cdots \circ e_l \circ \pi(A) = e_1 \circ \cdots \circ e_l(B)$ está na forma escalonada.

Ou seja, aplicando-se em A a sequência de operações elementares em linhas $e_1, \ldots e_l$ que escalona B, obtemos uma matriz C que, a menos de sua última coluna, está na forma escalonada, S.

Se A' está já na forma escalonada, não há nada para fazer. Caso contrário:

- Caso 1: existem linhas $i < i'$ de C tais que $C_{i'}$ não é nula e C_i é nula. Como S está na forma escalonada, o único elemento não nulo de $C_{i'}$ está na última coluna. Podemos multiplicá-lo pelo seu inverso e, como no caso base, utilizar esta linha para zerar as outras entradas da última coluna da matriz, e, em seguida, reordená-la para transformá-la em uma matriz na forma escalonada.

- Caso 2: alguma linha não nula não possui pivô. Notemos que cada pivô de S é um pivô de C. Seja c_{ij} o primeiro elemento não nulo de C que não é pivô.

  Se, por absurdo $j \leq n$, c_{ij} não é pivô de S, logo, existe $j' < j$ tal que $c_{ij'}$ é pivô de S. Assim, a i-ésima linha de S possui pivô, e, portanto, a i-ésima linha de C também possui pivô, o que é um absurdo. Logo, o único elemento não nulo de C_i está na última coluna e agimos como no caso anterior.

- Caso 3: existem linhas $i < i'$ de C tais que ambas são não nulas e, sendo $a_{ij}, a_{i'j'}$ os pivôs dessas linhas, $j > j'$.

  Ora, como S está na forma escalonada, isso só é possível se $j = n + 1$. Logo, a i-ésima linha de C é nula, o que implica que a i'-ésima linha de S é nula, mas $a_{i'j'}$ é pivô de S, o que é um absurdo.

$\square$

Proposição 2.4.10. Seja $A \in M_{m\times n}(\mathbb{K})$ uma matriz e $B \in M_{m\times n}(\mathbb{K})$ uma matriz na forma escalonada equivalente a B. Seja S o número de pivôs de B.

- m representa o número de equações do sistema linear homogêneo $AX = 0$.
- n representa o número de incógnitas do sistema linear homogêneo $AX = 0$.
- O número $S \leq m$ representa o número de variáveis dependentes do sistema linear homogêneo $AX = 0$ – variáveis que são determinadas a partir das demais.
- O número $n - S$ representa o número de variáveis independentes do sistema linear homogêneo $AX = 0$ – variáveis que podem ser escolhidas livremente. Ele é chamado de *grau de liberdade* do sistema linear homogêneo $AX = 0$.
- O sistema linear homogêneo $AX = 0$ possui soluções não triviais se, e somente se $n - S > 0$. Assim, se $m < n$, então o sistema linear homogêneo $AX = 0$ possui soluções não triviais.

2.5 Invertendo matrizes

O processo de escalonamento é uma ferramenta que é útil, também, para inverter matrizes.
Primeiro, definiremos invertibilidades laterais.

Definição 2.5.1. Seja R um anel e $a \in M_n(R)$. Dizemos que a é **invertível à esquerda** se existe $b \in M_n(R)$ tal que $ba = I_n$. Nesse caso, b é chamado de um **inverso à esquerda** de a.
Dizemos que a é **invertível à direita** se existe $c \in M_n(R)$ tal que $ac = I_n$. Nesse caso, c é chamado de um **inverso à direita** de a.

Quando inversos laterais de ambos os lados existem, eles coincidem, e, portanto, o elemento é invertível.

Proposição 2.5.2. Seja R um anel e $a \in M_n(R)$. Se a é invertível à esquerda e à direita, então a é invertível e os inversos laterais coincidem.

Demonstração. Suponha que b é um inverso à esquerda de a e que c é um inverso à direita de a. Então, temos

$$b = b1 = b(ac) = (ba)c = 1c = c.$$

Portanto, $b = c$, e assim a é invertível com (único) inverso $b = c$. □

Veremos mais adiante neste texto que há anéis para os quais existem elementos invertíveis apenas de um lado. Porém, conforme veremos a seguir, isso não ocorre para matrizes quadradas sobre corpos.

Lema 2.5.3. Seja $A \in M_n(\mathbb{K})$. São equivalentes:

(i) A é invertível.

(ii) A é invertível à esquerda.

(iii) O sistema linear homogêneo $AX = 0$ possui apenas a solução trivial.

(iv) Existe uma sequência finita de operações elementares nas linhas de $M_n(\mathbb{K})$ que transforma A na matriz identidade I_n.

(v) A pode ser escrita como um produto de matrizes elementares.

Demonstração. Façamos a demonstração das implicações.

(i)⇒(ii) Imediato.

(ii)⇒(iii) Suponha que A é invertível à esquerda. Então, existe $B \in M_n(\mathbb{K})$ tal que $BA = I_n$.

Seja $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ uma solução do sistema linear homogêneo $AX = 0$. Segue que:

$$A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = 0.$$

Multiplicando ambos os lados da equação por B à esquerda, obtemos

$$BA\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix} = B0 = 0.$$

Ou seja, $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Portanto, o sistema linear homogêneo $AX = 0$ possui apenas a solução trivial.

(iii)⇒(iv) Suponha que o sistema linear homogêneo $AX = 0$ possui apenas a solução trivial. Existem operações elementares nas linhas que transformam A em uma matriz escalonada S.

Como o sistema linear homogêneo $AX = 0$ possui apenas a solução trivial, todas as colunas de S possuem pivôs.

Como o número de linhas de S é igual ao número de colunas de S, segue que toda linha i de S possui pivô.

Sejam $j_1, j_2, \ldots, j_n$ as posições dos pivôs de S das linhas $1, 2, \ldots, n$, respectivamente. Temos que $1 \leq j_1 < j_2 < \cdots < j_n \leq n$. Portanto, $j_k = k$ para todo $k = 1, 2, \ldots, n$. Assim, S é uma matriz cuja diagonal é composta de pivôs, e, os elementos fora da diagonal, por estarem em uma coluna com pivô, são nulos. Ou seja, S é a matriz identidade.

(iv)⇒(v) Por hipótese, existem operações elementares $e_1, \ldots, e_k$ tais que $e_k \circ \ldots e_1(A) = I_n$.

Sendo $E_1, \ldots, E_k$ as matrizes elementares correspondentes às operações elementares $e_1, \ldots, e_k$, respectivamente, temos que:

$$E_k \cdots E_1 A = I_n.$$

Como as matrizes elementares são invertíveis, segue que $A = E_1^{-1} \cdots E_k^{-1}$.

Como o inverso de uma matriz elementar é uma matriz elementar, segue que A pode ser escrita como um produto de matrizes elementares.

(v)⇒(i) Suponha que A pode ser escrita como um produto de matrizes elementares. Como as matrizes elementares são invertíveis, segue que A é um produto de matrizes invertíveis, e, portanto, A é invertível.

□

Notemos um fato interessante: na demonstração anterior, a implicação (iv)⇒(v) nos dá um algoritmo sobre como escrever uma matriz invertível como um produto de matrizes elementares.

Vejamos um exemplo de seu funcionamento.

Exemplo 2.5.4. Seja

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

Queremos escrever A como um produto de matrizes elementares.

Para isso, aplicamos operações elementares nas linhas de A até que ela se transforme na matriz identidade:

$$A \xrightarrow{L_1 \leftrightarrow L_3} \begin{pmatrix}1 & 0 & 0\\ 1 & 3 & 2\\ 2 & 1 & 1\end{pmatrix} \xrightarrow{L_2 \leftarrow L_2 - L_1} \begin{pmatrix}1 & 0 & 0\\ 0 & 3 & 2\\ 2 & 1 & 1\end{pmatrix} \xrightarrow{L_3 \leftarrow L_3 - 2L_1} \begin{pmatrix}1 & 0 & 0\\ 0 & 3 & 2\\ 0 & 1 & 1\end{pmatrix}$$
$$\xrightarrow{L_2 \leftarrow \frac{1}{3}L_2} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & \frac{2}{3}\\ 0 & 1 & 1\end{pmatrix} \xrightarrow{L_3 \leftarrow L_3 - L_2} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & \frac{2}{3}\\ 0 & 0 & \frac{1}{3}\end{pmatrix} \xrightarrow{L_3 \leftarrow 3L_3} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & \frac{2}{3}\\ 0 & 0 & 1\end{pmatrix}$$
$$\xrightarrow{L_2 \leftarrow L_2 - \frac{2}{3}L_3} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} = I_3.$$

Portanto, A pode ser escrita como um produto de matrizes elementares.

Temos que:

$$A = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}E_5^{-1}E_6^{-1}E_7^{-1}.$$

Na expressão acima, E_i^{-1} é a inversa da matriz elementar correspondente à i-ésima operação e_i elementar aplicada.

Se e_i for uma troca de linhas, $E_i^{-1} = E_i$ é a matriz identidade com estas mesmas linhas trocadas.

Se e_i for uma multiplicação de uma linha por um escalar não nulo, E_i^{-1} é a matriz identidade com a mesma linha multiplicada pelo inverso do escalar.

Se e_i for a adição de um múltiplo de uma linha a outra, E_i^{-1} é a matriz identidade com o mesmo múltiplo da mesma linha subtraída da outra. Portanto,

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

A demonstração do teorema anterior também nos dá um algoritmo para inverter matrizes: basta aplicar operações elementares nas linhas de A até que ela se transforme em I_n. A inversa será, então, a matriz $E_k \dots E_1$, onde E_i é a matriz elementar correspondente à i-ésima operação elementar aplicada, e_i. Porém, $E_k \dots E_1$ é a matriz $e_k \circ \dots \circ e_1(I_n)$, ou seja, é a matriz que resulta de aplicar as mesmas operações elementares em I_n.

Já se ao aplicar operações elementares em A chegarmos em uma matriz escalonada diferente da identidade, esta não terá pivôs em todas as colunas (ou seria a identidade), e, portanto, o sistema linear homogêneo $AX = 0$ possuirá soluções não triviais, o que, pela proposição, implica em A não ser invertível.

Exemplo 2.5.5. Consideremos a seguinte matriz:

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

Para invertê-la, aplicaremos, simultaneamente, as mesmas operações elementares nela e em I_4, até que A se transforme em I_4. A matriz a qual I_4 se transformar será A^{-1}.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{L_2 \leftarrow L_2 - 2L_1} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -2 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{L_3 \leftarrow L_3 - L_1} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -2 & 1 & 0 \\ 0 & -1 & 1 & -1 & 0 & 1 \end{array}\right) \xrightarrow{L_3 \leftarrow L_3 + L_2} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -2 & 1 & 0 \\ 0 & 0 & 4 & -3 & 1 & 1 \end{array}\right)$$

$$\xrightarrow{L_3 \leftarrow \frac{1}{4}L_3} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -2 & 1 & 0 \\ 0 & 0 & 1 & -\frac{3}{4} & \frac{1}{4} & \frac{1}{4} \end{array}\right) \xrightarrow{L_2 \leftarrow L_2 - 3L_3} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \\ 0 & 0 & 1 & -\frac{3}{4} & \frac{1}{4} & \frac{1}{4} \end{array}\right)$$

$$\xrightarrow{L_1 \leftarrow L_1 + L_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \\ 0 & 0 & 1 & -\frac{3}{4} & \frac{1}{4} & \frac{1}{4} \end{array}\right)$$

Portanto,

$$A^{-1} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \\ -\frac{3}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Exemplo 2.5.6. Consideremos a seguinte matriz:

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

Vamos tentar invertê-la.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 3 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{L_2 \leftarrow L_2 - 2L_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right)$$

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

Observe que a última linha da matriz à esquerda ficou toda nula, enquanto a parte aumentada à direita não é nula. Isso indica que a forma escalonada de A não é a identidade, e, portanto, A não é invertível.

Agora vamos tirar mais alguns corolários teóricos a partir do já exposto.

Corolário 2.5.7. Seja $A \in M_n(\mathbb{K})$. As seguintes afirmações são equivalentes:

(i) A é invertível.

(ii) A é invertível à esquerda.

(iii) A é invertível à direita.

Demonstração. A equivalência (i)$\Leftrightarrow$(ii) já foi estabelecida, e é claro que (i)$\Rightarrow$(iii).

Para ver que (iii)$\Rightarrow$(i), suponha que A é invertível à direita e seja $B \in M_n(\mathbb{K})$ tal que $AB = I_n$.

Segue que B é invertível à esquerda, e, assim, B é invertível e sua inversa é A. Logo, $A = B^{-1}$ é invertível (e sua inversa é B). □

Corolário 2.5.8. Sejam $A_1, A_2, \ldots, A_k \in M_n(\mathbb{K})$ matrizes invertíveis.

Se o produto $A_1 A_2 \cdots A_k$ é invertível, então cada matriz A_i é invertível ($i = 1, 2, \ldots, k$).

Demonstração. Provaremos o corolário por indução em k.

Para $k = 1$, o resultado é trivial. Para $k = 2$, suponha que $A_1 A_2$ é invertível. Então, existe $B \in M_n(\mathbb{K})$ tal que $B(A_1 A_2) = I_n$. Assim, BA_1 é um inverso à esquerda de A_2, e, portanto, A_2 é invertível. Além disso, B é invertível, e, assim, $A_1 = B^{-1} A_2^{-1}$ é invertível.

Suponha que o resultado seja válido para k. Seja $A_1, A_2, \ldots, A_{k+1} \in M_n(\mathbb{K})$ tais que $A_1 A_2 \cdots A_{k+1}$ é invertível. Pelo caso $k = 2$, A_{k+1} é invertível e $A_1 A_2 \cdots A_k$ é invertível, e então, por hipótese de indução, $A_1, A_2, \ldots, A_k$ são invertíveis. □

2.6 Sistemas lineares não homogêneos

Já possuímos todas as ferramentas teóricas para tratar de sistemas não homogêneos. O último ingrediente se traduz na seguinte afirmação:

Lema 2.6.1. Seja $A \in M_{m \times (n+1)}(\mathbb{K})$. Escreva $A = (a_{ij})_{i,j}$. Os seguintes sistemas lineares são equivalentes:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = a_{1(n+1)} \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = a_{2(n+1)} \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = a_{m(n+1)}. \end{cases}$$

E:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + a_{1(n+1)}x_{n+1} = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + a_{2(n+1)}x_{n+1} = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + a_{m(n+1)}x_{n+1} = 0 \\ x_{n+1} = -1. \end{cases}$$

Isso nos dá um método prático para resolver sistemas lineares não homogêneos.

- Dado um sistema linear não homogêneo $AX = b$, podemos formar a matriz aumentada $(A|b)$, em que é adicionada à A uma coluna extra dada pelos elementos de b.
- O sistema $AX = b$ é, então, equivalente ao sistema homogêneo associado à matriz aumentada $(A|b)X = 0$ juntado com a equação extra $x_{n+1} = -1$.
- Escalonando-se $(A|b)$ encontramos uma matriz da forma $(S|c)$, onde S está na forma escalonada.
- O sistema original é, então, equivalente ao sistema $(S|c)X = c$ junto da a equação extra $x_{n+1} = -1$, que é equivalente ao sistema $SX = c$.
- Como acontece com os sistemas homogêneos, as soluções de $SX = c$ são notáveis. Porém, há um caso adicional: se temos uma linha correspondente a uma equação do tipo $0 = 1$, o sistema não possui solução. Nesse caso, dizemos que o sistema é *impossível*, ou *inconsistente*.
- Caso contrário, dizemos que o sistema é *possível*, ou *consistente*, e, como acontece no caso homogêneo, ou teremos infinitas soluções (no caso de alguma coluna de S não possuir pivô, caso no qual dizemos que o sistema é *indeterminado*), ou uma solução única (no caso de todas as colunas de S possuírem pivô, e nesse caso, dizemos que o sistema é *determinado*).

Capítulo 3

Introdução

Neste capítulo introduziremos a noção de Espaço Vetorial.

3.1 Espaços Vetoriais

Um espaço vetorial é uma estrutura algébrica definida sobre um corpo fixado que generaliza as propriedades dos vetores em um espaço euclidiano. Ele é composto por um conjunto de elementos chamados vetores, juntamente com duas operações: a adição de vetores e a multiplicação por escalares.

Definição 3.1.1. Seja $\mathbb{K}$ um corpo. Um *espaço vetorial* sobre $\mathbb{K}$, também chamado de $\mathbb{K}$-espaço vetorial, é um conjunto V munido de operações $+ : V \times V \to V$ e $\cdot : \mathbb{K} \times V \to V$, além de um elemento distinguido 0, tais que, para todos $u, v, w \in V$ e $\alpha, \beta \in \mathbb{K}$, as seguintes propriedades são satisfeitas:

A1. (Associatividade da adição) $(u + v) + w = u + (v + w)$.

A2. (Comutatividade da adição) $u + v = v + u$.

A3. (Elemento neutro da adição) $u + 0 = u$.

A4. (Elemento oposto da adição) Existe $x \in V$ tal que $u + x = 0$.

M1. (Compatibilidade do produto) $\alpha(\beta u) = (\alpha\beta)u$.

M2. (Elemento neutro do produto) $1u = u$, onde 1 é o elemento neutro de $\mathbb{K}$.

D1. (Distributividade de vetores) $\alpha(u + v) = \alpha u + \alpha v$.

D2. (Distributividade de escalares) $(\alpha + \beta)u = \alpha u + \beta u$.

Notemos que as quatro primeiras propriedades nos dizem que $(V, +, 0)$ é um grupo abeliano. Deste modo, conforme visto na introdução, o elemento x da propriedade A4 é único e é denotado por $-u$.

Vejamos alguns exemplos.

Exemplo 3.1.2 (Espaço Vetorial $\mathbb{K}^n$)**.** Dado um corpo $\mathbb{K}$, o conjunto $\mathbb{K}^n$ (o espaço das n-tuplas de elementos de $\mathbb{K}$) é um espaço vetorial sobre $\mathbb{K}$. Tal fato foi provado implicitamente em 2.2.3.

Exemplo 3.1.3 (O espaço de polinômios)**.** Seja $\mathbb{K}$ um corpo. O conjunto $\mathbb{K}[x]$ dos polinômios com coeficientes em $\mathbb{K}$, denotado por $\mathbb{K}[x]$, é o espaço das expressões formais da forma

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

onde $n \in \mathbb{N}$ e $a_0, a_1, \ldots, a_n \in \mathbb{K}$. Vale que se $a_0, \ldots, a_n, b_0, \ldots, b_m \in \mathbb{K}$, então:

$$a_0 + a_1x + \cdots + a_nx^n = b_0 + b_1x + \cdots + b_nx^n \Leftrightarrow \forall i \leq n\, a_i = b_i.$$

O conjunto $\mathbb{K}[x]$ pode ser construído formalmente como sendo o conjunto de todas as sequências $(a_0, a_1, a_2, \ldots)$ de elementos de $\mathbb{K}$ tais que existe $n \in \mathbb{N}$ com $a_i = 0$ para todo $i > n$. O polinômio nulo 0 é dado pela sequência nula, ou seja, $0 = (0, 0, 0, \ldots)$.

As operações de adição e multiplicação por escalar são definidas da maneira usual. Ou seja, dados $p(x) = a_0 + a_1x + \cdots + a_nx^n$ e $q(x) = b_0 + b_1x + \cdots + b_mx^m$ em $\mathbb{K}[x]$, define-se

$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_k + b_k)x^k, \text{ onde } k = \max\{n, m\},$$
$$\alpha p(x) = (\alpha a_0) + (\alpha a_1)x + \cdots + (\alpha a_n)x^n, \text{ para todo } \alpha \in \mathbb{K}.$$

Com tais operações, $\mathbb{K}[x]$ é um espaço vetorial sobre $\mathbb{K}$. Com efeito, dados $p(x), q(x), r(x) \in \mathbb{K}[x]$ e $\alpha, \beta \in \mathbb{K}$, e escrevendo $p(x) = \sum_{i=0}^n a_ix^i$, $q(x) = \sum_{i=0}^n b_ix^i$ e $r(x) = \sum_{i=0}^n c_ix^i$ (é possível utilizarmos o mesmo n, pois podemos completar com coeficientes nulos), temos:

A1. $(p(x)+q(x))+r(x) = \sum_{i=0}^n(a_i+b_i)+\sum_{i=0}^n c_ix^i = \sum_{i=0}^n a_i+\sum_{i=0}^n(b_i+c_i)x^i = p(x)+(q(x)+r(x))$.

A2. $p(x) + q(x) = \sum_{i=0}^n(a_i + b_i)x^i = \sum_{i=0}^n(b_i + a_i)x^i = q(x) + p(x)$.

A3. $p(x) + 0 = \sum_{i=0}^n a_ix^i + \sum_{i=0}^n 0x^i = \sum_{i=0}^n a_ix^i = p(x)$.

A4. $p(x) + (-p(x)) = \sum_{i=0}^n a_ix^i + \sum_{i=0}^n(-a_i)x^i = \sum_{i=0}^n 0x^i = 0$.

M1. $\alpha(\beta p(x)) = \alpha \sum_{i=0}^n(\beta a_i)x^i = \sum_{i=0}^n(\alpha\beta a_i)x^i = (\alpha\beta)p(x)$.

M2. $1p(x) = \sum_{i=0}^n(1a_i)x^i = \sum_{i=0}^n a_ix^i = p(x)$.

D1. $\alpha(p(x) + q(x)) = \alpha \sum_{i=0}^n(a_i + b_i)x^i = \sum_{i=0}^n(\alpha a_i + \alpha b_i)x^i = \alpha p(x) + \alpha q(x)$.

D2. $(\alpha + \beta)p(x) = \sum_{i=0}^n((\alpha + \beta)a_i)x^i = \sum_{i=0}^n(\alpha a_i + \beta a_i)x^i = \alpha p(x) + \beta p(x)$.

Existe também o produto de polinômios. Porém, abordaremos tal operação apenas quando necessário, uma vez que espaços vetoriais, no geral, não são munidos de produtos, e introduzi-los agora pode causar confusão.

Exemplo 3.1.4 (Matrizes)**.** Fixados inteiros $m, n \geq 1$ e um corpo $\mathbb{K}$, o conjunto $\mathbb{M}_{m\times n}(\mathbb{K})$ das matrizes $m \times n$ com entradas em $\mathbb{K}$ é um $\mathbb{K}$-espaço vetorial, conforme já visto.

Exemplo 3.1.5 (Espaços de funções)**.** Seja $\mathbb{K}$ um corpo e I um conjunto qualquer. O conjunto de todas as funções $f : I \to \mathbb{K}$, denotado por $\mathbb{K}^I$, é um espaço vetorial sobre $\mathbb{K}$ com as operações definidas por

$$(f + g)(x) = f(x) + g(x), \text{ para todo } x \in I,$$
$$(\alpha f)(x) = \alpha(f(x)), \text{ para todo } x \in I \text{ e } \alpha \in \mathbb{K}.0(x) \qquad = 0, \text{ para todo } x \in I.$$

O oposto $-f$ de uma função $f \in \mathbb{K}^I$ é a função definida por $(-f)(x) = -(f(x))$, para todo $x \in I$. De fato, dados $f, g, h \in \mathbb{K}^I$ e $\alpha, \beta \in \mathbb{K}$, temos:

A1. $(f + g) + h = f + (g + h)$. De fato, dado $x \in I$, temos $((f + g) + h)(x) = (f + g)(x) + h(x) = (f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) = f(x) + (g + h)(x) = (f + (g + h))(x)$.

A2. $f + g = g + f$. De fato, dado $x \in I$, temos $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$.

A3. $f + 0 = f$. De fato, dado $x \in I$, temos $(f + 0)(x) = f(x) + 0(x) = f(x) + 0 = f(x)$.

A4. $f + (-f) = 0$. De fato, dado $x \in I$, temos $(f + (-f))(x) = f(x) + (-f)(x) = f(x) - f(x) = 0$.

M1. $\alpha(\beta f) = (\alpha\beta)f$. De fato, dado $x \in I$, temos $(\alpha(\beta f))(x) = \alpha((\beta f)(x)) = \alpha(\beta(f(x))) = (\alpha\beta)(f(x)) = ((\alpha\beta)f)(x)$.

M2. $1f = f$. De fato, dado $x \in I$, temos $(1f)(x) = 1(f(x)) = f(x) = (f)(x)$.

D1. $\alpha(f + g) = \alpha f + \alpha g$. De fato, dado $x \in I$, temos $(\alpha(f + g))(x) = \alpha((f + g)(x)) = \alpha(f(x) + g(x)) = \alpha(f(x)) + \alpha(g(x)) = (\alpha f)(x) + (\alpha g)(x) = (\alpha f + \alpha g)(x)$.

D2. $(\alpha+\beta)f = \alpha f+\beta f$. De fato, dado $x \in I$, temos $((\alpha+\beta)f)(x) = (\alpha+\beta)(f(x)) = \alpha(f(x))+\beta(f(x)) = (\alpha f)(x) + (\beta f)(x) = (\alpha f + \beta f)(x)$.

Vejamos algumas propriedades elementares de espaços vetoriais.

Proposição 3.1.6. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Sejam $u, v \in V$ e $\alpha, \beta \in \mathbb{K}$. Então:

(i) $0u = 0$.

(ii) $\alpha 0 = 0$.

(iii) Se $\alpha u = 0$, então $\alpha = 0$ ou $u = 0$.

(iv) $(-1)u = -u$.

Demonstração. (i) Temos que $0u = (0 + 0)u = 0u + 0u$. Somando o oposto de $0u$ em ambos os lados, obtemos $0 = 0u$.

(ii) Temos que $\alpha 0 = \alpha(0 + 0) = \alpha 0 + \alpha 0$. Somando o oposto de $\alpha 0$ em ambos os lados, obtemos $0 = \alpha 0$.

(iii) Se $\alpha = 0$, a afirmação é verdadeira. Suponhamos então que $\alpha \neq 0$. Neste caso, existe $\alpha^{-1} \in \mathbb{K}$ tal que $\alpha\alpha^{-1} = 1$. Multiplicando ambos os lados da igualdade $\alpha u = 0$ por α^{-1}, obtemos $1u = \alpha^{-1}0$, ou seja, $u = 0$.

(iv) Temos que $(-1)u + u = (-1 + 1)u = 0u = 0$. Portanto, $(-1)u$ é o oposto de u, ou seja, $(-1)u = -u$. □

3.2 Subespaços

Notemos que nos exemplos anteriores, a verificação de que cada um dos exemplos é um espaço vetorial é simples, porém longa, uma vez que todas as oito propriedades devem ser verificadas. Neste processo, foi necessário utilizar propriedades algébricas já conhecidas de outros conjuntos, como os do próprio corpo $\mathbb{K}$.

Porém, em muitos casos, é possível verificar que um subconjunto de um espaço vetorial é, ele próprio, um espaço vetorial, sem a necessidade de verificar todas as oito propriedades.

Definição 3.2.1. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Um subconjunto $W \subseteq V$ é um *subespaço* de V se, com as operações herdadas de V, W é um espaço vetorial sobre $\mathbb{K}$. Ou seja, W é um subespaço de V se, para todo $u, v \in W$ e $\alpha \in \mathbb{K}$, $u + v \in W$, $\alpha u \in W$ e $0 \in W$, onde 0 é o elemento neutro de V, e, com essas operações, W satisfaz as oito propriedades de espaço vetorial.

A verdade é que verificar as oito propriedades de espaço vetorial é desnecessário na definição de subespaço, conforme mostra a proposição a seguir.

Proposição 3.2.2. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Um subconjunto $W \subseteq V$ é um subespaço de V se, e somente se:

(i) $0 \in W$, onde 0 é o elemento neutro de V;

(ii) para todo $u, v \in W$, e $\alpha \in \mathbb{K}$, $\alpha u + v \in W$.

Demonstração. A implicação direta é simples: dados $u, v \in W$ e $\alpha \in \mathbb{K}$, como W é um subespaço de V, temos que $\alpha u \in W$, e, portanto, $\alpha u + v \in W$. Além disso, $0 \in W$.

Reciprocamente, suponha que W satisfaça (i) e (ii).

Primeiro dados $u, v \in W$, temos que $1u + v = u + v \in W$, ou seja, W é fechado para a adição. Além disso, dados $u \in W$ e $\alpha \in \mathbb{K}$, temos que $\alpha u + 0 = \alpha u \in W$, ou seja, W é fechado para a multiplicação escalar. Note ainda que se $u \in W$, então $(-1)u = -u \in W$.

Agora notemos que as propriedades de espaço vetorial são automáticas. De fato, dados $u, v, w \in W$ e $\alpha, \beta \in \mathbb{K}$, temos:

A1. $(u + v) + w = u + (v + w)$, pois isso vale em V.

A2. $u + v = v + u$, pois isso vale em V.

A3. $u + 0 = u$, pois $0 \in W$ e isso vale em V.

A4. Existe $x \in W$ tal que $u + x = 0$, pois $-u \in W$ e isso vale em V.

M1. $\alpha(\beta u) = (\alpha\beta)u$, pois isso vale em V.

M2. $1u = u$, pois isso vale em V.

D1. $\alpha(u + v) = \alpha u + \alpha v$, pois isso vale em V.

D2. $(\alpha+\beta)u = \alpha u + \beta u$, pois isso vale em V.

$\square$

Alguns exemplos:

Exemplo 3.2.3 (O subespaço trivial)**.** Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Então $W = \{0\}$ é um subespaço de V. Com efeito, $\alpha 0 + 0 = 0 \in W$ para todo $\alpha \in \mathbb{K}$.

Tal espaço é denotado por (0) ou $\{0\}$, e é chamado de *subespaço trivial* de V.

Exemplo 3.2.4 (O próprio espaço vetorial)**.** Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Então $W = V$ é um subespaço de V.

Exemplo 3.2.5. $\{(x,0) : x \in \mathbb{R}\}$ é um subespaço de $\mathbb{R}^2$. Com efeito, $(0,0) \in W$ e, dados $(x,0),(y,0) \in W$ e $\alpha \in \mathbb{R}$, temos que $\alpha(x,0)+(y,0) = (\alpha x + y, 0) \in W$.

Analogamente, $\{(0,y) : y \in \mathbb{R}\}$ é um subespaço de $\mathbb{R}^2$.

Exemplo 3.2.6 (O espaço solução)**.** Seja $\mathbb{K}$ um corpo e $A \in M_{m\times n}(\mathbb{K})$. O conjunto solução de $AX = 0$, ou seja, o conjunto $\{(x_1,\ldots,x_n) \in \mathbb{K}^n : A(x_1,\ldots,x_n)^T = 0\}$ é um subespaço de $\mathbb{K}^n$. Com efeito, $(0,\ldots,0) \in W$ e, dados $u = (u_1,\ldots,u_m), v = (v_1,\ldots,v_m) \in W$ e $\alpha \in \mathbb{K}$, temos que

$$A\begin{pmatrix} \alpha u_1 + v_1 \\ \vdots \\ \alpha u_m + v_m \end{pmatrix} = \alpha A \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} + A \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = \alpha 0 + 0 = 0, \tag{3.1}$$

ou seja, $\alpha u + v \in W$.

Exemplo 3.2.7 (Polinômios de grau $\leq n$)**.** Seja $\mathbb{K}$ um corpo. O conjunto $\mathbb{P}_n(\mathbb{K})$ dos polinômios com coeficientes em $\mathbb{K}$ e grau menor ou igual a n (ou 0) é um subespaço de $\mathbb{K}[x]$.

Com efeito, o polinômio nulo está em $\mathbb{P}_n(\mathbb{K})$ e, dados $p(x), q(x) \in \mathbb{P}_n(\mathbb{K})$ e $\alpha \in \mathbb{K}$, temos que $\alpha p(x) + q(x) \in \mathbb{P}_n(\mathbb{K})$.

Exemplo 3.2.8. O conjunto $\{(1,2,3)\}$ não é um subespaço de $\mathbb{R}^3$, pois não contém o vetor nulo.

Exemplo 3.2.9. O conjunto $\mathbb{Z}^2 \subseteq \mathbb{R}^2$ não é um subespaço de $\mathbb{R}^2$, pois não é fechado para a multiplicação escalar.

Exemplo 3.2.10. $\{(x,0) : x \geq 0\} \cup \{(0,y) : y \geq 0\}$ não é um subespaço de $\mathbb{R}^2$, pois não é fechado para a adição.

Assim, a união de dois subespaços não é um subespaço. Apesar disso, qualquer interseção de subespaços é um subespaço.

Proposição 3.2.11. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $\mathcal{S}$ uma coleção qualquer de subespaços de V. Então $\bigcap \mathcal{S} = \bigcap_{W\in\mathcal{S}} W$ é um subespaço de V. Além disso, este é o maior subespaço de V contido em cada subespaço de $\mathcal{S}$ no sentido da inclusão $\subseteq$.

Demonstração. Seja $U = \bigcap \mathcal{S}$. Note que $0 \in U$, pois 0 pertence a todo subespaço de V. Além disso, dados $u, v \in U$ e $\alpha \in \mathbb{K}$, temos que $u, v \in W$ para todo $W \in \mathcal{S}$, e, portanto, $\alpha u + v \in W$ para todo $W \in \mathcal{S}$, ou seja, $\alpha u + v \in U$. Assim, U é um subespaço de V.

Para a última afirmação, seja U um subespaço de V tal que $U \subseteq W'$ para todo $W' \in \mathcal{S}$. Então $U \subseteq \bigcap_{W\in\mathcal{S}} W = U$. $\square$

3.3 Subespaços gerados e combinações lineares

Definição 3.3.1. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $A \subseteq V$. O *subespaço gerado* por A, denotado por $\langle A \rangle$, é o menor subespaço de V que contém A. Mais precisamente, $\langle A \rangle$ é o único subespaço W de V que satisfaz:

(i) $A \subseteq W$;

(ii) se U é um subespaço de V tal que $A \subseteq U$, então $W \subseteq U$.

É claro que devemos ver que o subespaço gerado por A está bem definido, ou seja, que tal subespaço existe e é único.

Lema 3.3.2. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $A \subseteq V$. Então existe um único subespaço de V que satisfaz (i) e (ii) da definição anterior.

Demonstração. Começaremos provando a existência. Seja $\mathcal{S}$ a coleção de todos os subespaços de V que contêm A, ou seja:

$$\mathcal{S} = \{W \subseteq V : W \text{ é um subespaço de } V \text{ e } A \subseteq W\}.$$

Note que $\mathcal{S}$ é não vazia, pois $V \in \mathcal{S}$. Seja $W = \bigcap_{U \in \mathcal{S}} U$. Pela Proposição 3.2.11, W é um subespaço de V. Além disso, $A \subseteq W = \bigcap_{U \in \mathcal{S}} U$, pois $A \subseteq U$ para todo $U \in \mathcal{S}$. Assim, (i) é satisfeita.

Para (ii), seja U' um subespaço de V tal que $A \subseteq U'$. Então $U' \in \mathcal{S}$ e, portanto, $W = \bigcap_{U \in \mathcal{S}} U \subseteq U'$.

Agora provaremos a unicidade. Suponha que W_1 e W_2 são subespaços de V que satisfazem (i) e (ii).

Por (i) aplicado à W_1, temos que $A \subseteq W_1$. Assim, por (ii) aplicado à W_2, com $U = W_1$, temos que $W_2 \subseteq W_1$. Analogamente, por (i) aplicado à W_2, temos que $A \subseteq W_2$. Assim, por (ii) aplicado à W_1, com $U = W_2$, temos que $W_1 \subseteq W_2$.

Portanto, $W_1 = W_2$. □

Convém darmos uma descrição mais construtiva e visualizável do subespaço gerado por um conjunto $A \subseteq V$. Para isso, introduziremos o conceito de combinação linear.

Definição 3.3.3. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $A \subseteq V$. Uma *combinação linear* de elementos de A é um vetor da forma

$$\sum_{i<n} \alpha_i v_i$$

onde $(\alpha_n)_{i<n}$ é uma sequência finita de escalares em $\mathbb{K}$ e $(v_i)_{i<n}$ é uma sequência finita de vetores em A dois-a-dois distintos.

Na notação acima, adotamos a convenção de que, se $n = 0$, a soma vazia é igual ao vetor nulo, 0. Dessa forma, 0 é uma combinação linear de elementos de A qualquer que seja A, mesmo que A seja o conjunto vazio.

Proposição 3.3.4. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $A \subseteq V$. Então $\langle A \rangle$ é o conjunto de todas as combinações lineares de elementos de A. Ou seja:

$$\langle A \rangle = \left\{ \sum_{i<n} \alpha_i v_i : n \in \mathbb{N}, (\alpha_i)_{i<n} \text{ é uma sequência finita em } \mathbb{K}, (v_i)_{i<n} \text{ é uma sequência finita em } A \right\}.$$

Demonstração. Seja $W = \{\sum_{i<n} \alpha_i v_i : n \in \mathbb{N}, (\alpha_i)_{i<n}$ é uma sequência finita em $\mathbb{K}$, $(v_i)_{i<n}$ é uma sequência finita injetora em $A\}$.

Vejamos primeiro que $W \subseteq \langle A \rangle$. Procedemos por indução em n para mostrar que dada uma sequência finita injetora $(v_i)_{i<n}$ em A e uma sequência finita $(\alpha_i)_{i<n}$ em $\mathbb{K}$, $\sum_{i<n} \alpha_i v_i \in \langle A \rangle$.

- Se $n = 0$ temos que $\sum_{i<n} \alpha_i v_i = 0 \in \langle A \rangle$.
- Para o passo indutivo, temos que $\sum_{i<n+1} \alpha_i v_i = \alpha_n v_n + \sum_{i<n} \alpha_i v_i$. Como $v_n \in \langle A \rangle$, $\alpha_n \in \mathbb{K}$ e $\sum_{i<n} \alpha_i v_i \in \langle A \rangle$ pela hipótese indutiva, segue que $\sum_{i<n+1} \alpha_i v_i \in \langle A \rangle$.

Reciprocamente, para ver que $\langle A \rangle \subseteq W$, veremos que W é um subespaço de V que contém A. De fato, $0 \in W$. Dados $u, v \in W$, completando a sequência de escalares com 0 quando necessário, podemos escrever $u = \sum_{i<n} \alpha_i v_i$ e $v = \sum_{i<n} \beta_i v_i$, onde $(v_i)_{i<n}$ é uma sequência finita injetora em A e $(\alpha_i)_{i<n}$, $(\beta_i)_{i<n}$ são sequências finitas em $\mathbb{K}$. Assim, para $\gamma \in \mathbb{K}$, temos que

$$\gamma u + v = \sum_{i<n} \gamma \alpha_i v_i + \sum_{i<n} \beta_i v_i = \sum_{i<n} (\gamma \alpha_i + \beta_i) v_i \in W.$$

Para ver que $A \subseteq W$, note que, para $v \in A$, temos que $v = 1 \cdot v \in W$. □

Um caso particular importante é o seguinte:

Corolário 3.3.5. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $A = \{v_1, \dots, v_n\}$ um subconjunto finito de V, em que $v_i \neq v_j$ para $i \neq j$. Então

$$\langle A \rangle = \left\{ \sum_{i=1}^{n} \alpha_i v_i : (\alpha_i)_{i=1}^{n} \text{ é uma sequência finita em } \mathbb{K} \right\}.$$

Proposição 3.3.6. Seja V um espaço vetorial e $A, B \subseteq V$.
Então $A \subseteq \langle B \rangle$ se, e somente se $\langle A \rangle \subseteq \langle B \rangle$.
Em particular, se $A \subseteq B$, então $\langle A \rangle \subseteq \langle B \rangle$.

Demonstração. Suponha que $A \subseteq \langle B \rangle$. Como $\langle A \rangle$ é o menor subespaço de V que contém A, e $\langle B \rangle$ é um subespaço de V que contém A, segue que $\langle A \rangle \subseteq \langle B \rangle$.

Reciprocamente, se $\langle A \rangle \subseteq \langle B \rangle$, como $A \subseteq \langle A \rangle \subseteq \langle B \rangle$, segue que $A \subseteq \langle B \rangle$.

Para a última afirmação, note que se $A \subseteq B$, como $B \subseteq \langle B \rangle$, segue que $A \subseteq \langle B \rangle$. □

3.4 Dependência linear e bases

Definição 3.4.1 (Dependência linear)**.** Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $A \subseteq V$. Dizemos que A é *linearmente independente* se, para toda sequência finita injetora $(v_i)_{i<n}$ em A e toda sequência finita $(\alpha_i)_{i<n}$ em $\mathbb{K}$, a seguinte implicação é verdadeira:

$$\sum_{i<n} \alpha_i v_i = 0 \implies \alpha_i = 0 \text{ para todo } i < n.$$

Caso contrário, dizemos que A é *linearmente dependente*.

Definição 3.4.2 (Base)**.** Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Uma *base* de V é um conjunto $B \subseteq V$ que é linearmente independente e gera V, ou seja, $\langle B \rangle = V$.

Lema 3.4.3. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $A \subseteq V$. Se $v \in V$ é tal que $v \in \langle A \rangle \setminus A$, então $A \cup \{v\}$ é linearmente dependente.

Demonstração. Seja $v \in \langle A \rangle \setminus A$. Então existem $n \in \mathbb{N}$, uma sequência finita $(\alpha_i)_{i<n}$ em $\mathbb{K}$ e uma sequência finita injetora $(v_i)_{i<n}$ em A tais que $v = \sum_{i<n} \alpha_i v_i$. Assim, $0 = v - v = \sum_{i<n} \alpha_i v_i + (-1)v$. Note que $-1 \neq 0$ e $v \notin A$, logo $A \cup \{v\}$ é linearmente dependente. □

Lema 3.4.4. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $A \subseteq V$. Se A é linearmente independente e $v \in V \setminus A$, então $A \cup \{v\}$ é linearmente independente se, e somente se, $v \notin \langle A \rangle$.

Demonstração. Seja $v \in V \setminus A$.

Pelo lema anterior, se $v \in \langle A \rangle$, então $A \cup \{v\}$ é linearmente dependente.

Reciprocamente, se $A \cup \{v\}$ é linearmente dependente, existem $n \in \mathbb{N}$, uma sequência finita $(\alpha_i)_{i<n}$ em $\mathbb{K}$, não toda nula, e uma sequência finita injetora $(u_i)_{i<n}$ em $A \cup \{v\}$ tais que $\sum_{i<n} \alpha_i u_i = 0$.

Como A é linearmente independente, v deve aparecer em $(u_i)_{i<n}$. Assim, em realidade, existem $m \in \mathbb{N}$, uma sequência finita $(\beta_i)_{i<m}$ em $\mathbb{K}$, não toda nula, e uma sequência finita injetora $(w_i)_{i<m}$ em A tais que $\sum_{i<m} \beta_i w_i + \gamma v = 0$, com $\gamma \neq 0$.

Novamente, como A é linearmente independente, $\gamma \neq 0$. Assim, $v = -\frac{1}{\gamma} \sum_{i<m} \beta_i w_i \in \langle A \rangle$. □

Definição 3.4.5. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Dizemos que V é *finitamente gerado* se existe um conjunto finito $A \subseteq V$ tal que $\langle A \rangle = V$.

Proposição 3.4.6. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e G um gerador finito de V. Então existe $B \subseteq G$ que é uma base de V.

Demonstração. Provaremos por indução no tamanho de G.

Se $|G| = 0$, então $G = \emptyset$. Tome $B = \emptyset$. Então B é uma base de V.

Suponha que o resultado seja verdadeiro para todo gerador finito de tamanho n, independentemente do espaço vetorial.

Seja G um gerador finito de tamanho $n + 1$. Seja $v \in G$ qualquer e seja $G' = G \setminus \{v\}$ e $W = \langle G' \rangle$. Por hipótese de indução, existe $B' \subseteq G'$ que é uma base de W.

Caso 1: se $B' \cup \{v\}$ é linearmente independente, então $G = G' \cup \{v\} \subseteq \langle B' \cup \{v\} \rangle$, o que implica que $V = \langle G \rangle \subseteq \langle B' \cup \{v\} \rangle$. Assim, $B = B' \cup \{v\}$ é um conjunto LI gerador de V, ou seja, uma base de V.

Caso 2: se $B' \cup \{v\}$ é linearmente dependente. Nesse caso, como B' é LI, temos que $v \in \langle B' \rangle \subseteq W = \langle G' \rangle$. Assim, $G \subseteq \langle B' \rangle$, e, portanto, $V = \langle G \rangle \subseteq \langle B' \rangle$. Logo, $B = B'$ é uma base de V. □

Note que a demonstração acima nos dá um algoritmo para encontrar uma base a partir de um gerador finito. Façamos um exemplo prático disso.

Exemplo 3.4.7. Seja $V = \mathbb{R}[x]$ o espaço vetorial dos polinômios com coeficientes reais. Considere $W = \langle 0, 1, x^2+1, x^2-1, x^3+3x+2\rangle$. Vamos encontrar uma base de W.

Para tanto, vamos, recursivamente, construir bases para $\langle\emptyset\rangle, \langle 0\rangle, \langle 0,1\rangle, \langle 0,1,x^2+1\rangle, \langle 0,1,x^2+1,x^2-1\rangle$ e $\langle 0,1,x^2+1,x^2-1,x^3+3x+2\rangle$.

- Temos que $B_0 = \emptyset$ é uma base de $\langle\emptyset\rangle$.
- Para obter uma base de $\langle 0\rangle$, basta ver se $B_0 \cup \{0\}$ é ou não LI. Como $\{0\}$ é LD, temos que $B_1 = B_0 = \emptyset$ é uma base de $\langle 0\rangle$.
- Para obter uma base de $\langle 0,1\rangle$, vejamos se $B_1 \cup \{1\}$ é ou não LI. Como $\{1\}$ é LI, temos que $B_2 = B_1 \cup \{1\} = \{1\}$ é uma base de $\langle 0,1\rangle$.
- Para obter uma base de $\langle 0,1,x^2+1\rangle$, vejamos se $B_2 \cup \{x^2+1\}$ é ou não LI. Como $x^2+1 \notin \langle B_2\rangle = \{\alpha : \alpha \in \mathbb{R}\}$, temos que $B_3 = B_2 \cup \{x^2+1\} = \{1, x^2+1\}$ é uma base de $\langle 0,1,x^2+1\rangle$.
- Para obter uma base de $\langle 0,1,x^2+1,x^2-1\rangle$, vejamos se $B_3 \cup \{x^2-1\}$ é ou não LI. Como $x^2-1 = (x^2+1) - 2\cdot 1 \in \langle B_3\rangle$, temos que $B_4 = B_3 = \{1, x^2+1\}$ é uma base de $\langle 0,1,x^2+1,x^2-1\rangle$.
- Finalmente, para obter uma base de $\langle 0,1,x^2+1,x^2-1,x^3+3x+2\rangle$, vejamos se $B_4 \cup \{x^3+3x+2\}$ é ou não LI. Como $x^3+3x+2 \notin \langle B_4\rangle = \{\alpha + \beta(x^2+1) : \alpha, \beta \in \mathbb{R}\}$, temos que $B_5 = B_4 \cup \{x^3+3x+2\} = \{1, x^2+1, x^3+3x+2\}$ é uma base de $\langle 0,1,x^2+1,x^2-1,x^3+3x+2\rangle$.

Corolário 3.4.8. Todo espaço finitamente gerado possui uma base finita.

Em geral, vale que todo espaço vetorial tem base. Porém, a prova deste fato não é elementar, e depende de um uso não trivial do Axioma da Escolha, sendo, portanto, um resultado de existência não construtivo: há espaços vetoriais não finitamente gerados cujas bases não são descritíveis por um algoritmo concreto, como o que vimos em ação acima.

O nosso próximo objetivo é mostrar que, em um espaço vetorial finitamente gerado, todas as bases possuem o mesmo tamanho.

Proposição 3.4.9. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Se G é um gerador finito de V e L é um conjunto linearmente independente de V, então $|L| \leq |G|$.

Demonstração. Suponha que $|L| > |G|$. Veremos que L é linearmente dependente. Sem perda de generalidade, $|L| = n+1 > n = |G|$.

Enumeremos $G = \{u_i : 1 \leq i \leq n\}$ e $L = \{v_{j+1} : 1 \leq j \leq n+1\}$.

Como G gera V, para cada $j \in \{1, \ldots, n+1\}$ existem $a_{1j}, \ldots, a_{nj} \in \mathbb{K}$ tais que $u_i = \sum_{j=1}^{n} a_{ij} v_j$.

Vejamos que L é linearmente dependente. Basta ver que existem $\beta_1, \ldots, \beta_{n+1} \in \mathbb{K}$, não todos nulos tais que $\sum_{j=1}^{n+1} \beta_j v_j = 0$, ou seja, tais que:

$$0 = \sum_{j=1}^{n+1} \beta_j v_j = \sum_{j=1}^{n+1} \beta_j \sum_{i=1}^{n} a_{ij} u_i = \sum_{i=1}^{n} \sum_{j=1}^{n+1} a_{ij} \beta_j u_i$$

É suficiente ver que o seguinte sistema tenha solução não trivial:

$$\begin{cases} a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1(n+1)}\beta_{n+1} = 0 \\ a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2(n+1)}\beta_{n+1} = 0 \\ \vdots \\ a_{n1}\beta_1 + a_{n2}\beta_2 + \cdots + a_{n(n+1)}\beta_{n+1} = 0. \end{cases}$$

Sendo este um sistema de n equações e $n+1$ incógnitas, como ao escaloná-lo obtemos no máximo n pivôs, existe ao menos uma coluna sem pivô. Assim, existe uma solução não trivial. □

Corolário 3.4.10. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Se V é finitamente gerado, então todas as bases de V possuem o mesmo tamanho.

Demonstração. Como V é finitamente gerado, existe uma base finita B de V. Seja $n = |B|$. Veremos que toda base de V possui tamanho n.

Seja C uma base qualquer de V. Em particular, C é um gerador de V. Como B é um gerador finito de V, pelo resultado anterior, temos que $|C| \leq |B|$.

Assim, C é um gerador finito de V. Como B é linearmente independente, pelo resultado anterior, temos que $|B| \leq |C|$.

Assim, $|C| = |B| = n$. □

Novamente, é um resultado verdadeiro que todas as bases de um espaço vetorial possuem o mesmo tamanho, mesmo que o espaço não seja finitamente gerado. Porém, há de se definir o que significa dizer que dois conjuntos infinitos possuem o mesmo tamanho. Em geral, dizemos que dois conjuntos possuem a mesma cardinalidade (tamanho, ou número de elementos) se existe uma bijeção entre eles. Porém, o *número de elementos* de um conjunto infinito é um tema da combinatória infinita e da teoria dos conjuntos que foge do escopo deste texto, e não será tratado aqui.

Definição 3.4.11. Seja V um espaço vetorial. A dimensão de V, denotada por $\dim V$, é a cardinalidade de uma (toda) base de V. Se a dimensão de V é finita, dizemos que V é de *dimensão finita*. Caso contrário, dizemos que V é de *dimensão infinita*.

Corolário 3.4.12. Um espaço vetorial possui dimensão finita se, e somente se, é finitamente gerado.

Exemplo 3.4.13 (A base canônica de $\mathbb{K}^n$)**.** O conjunto $B = \{e_1, e_2, \ldots, e_n\} \subseteq \mathbb{K}^n$, onde e_i é o vetor cujas coordenadas são todas nulas, exceto a i-ésima, que é igual a 1, é uma base de $\mathbb{K}^n$.

Assim, $\dim \mathbb{K}^n = n$.

Exemplo 3.4.14. O conjunto $B = \{1, x, x^2, x^3, \ldots\} \subseteq \mathbb{K}[x]$ é uma base de $\mathbb{K}[x]$.

Assim, $\mathbb{K}[x]$ é de dimensão infinita.

Por outro lado, $P_n(\mathbb{K})$, o espaço dos polinômios de grau no máximo n, é gerado pelo conjunto LI $B_n = \{1, x, x^2, \ldots, x^n\}$. Assim, $\dim P_n(\mathbb{K}) = n + 1$.

Proposição 3.4.15. Seja V um espaço vetorial de dimensão finita, n. Então todo conjunto LI de tamanho n é uma base de V, e todo conjunto gerador de tamanho n é uma base de V.

Demonstração. Seja $L \subseteq V$ um conjunto LI de tamanho n. Se L não gera V, temos que existe $v \in V$ tal que $v \notin \langle L \rangle$. Assim, $L \cup \{v\}$ é LI, o que é um absurdo, pois $|L \cup \{v\}| = n + 1 > n = \dim V$.

Agora seja $G \subseteq V$ um conjunto gerador de tamanho n. Se G não é LI, existem $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$, não todos nulos, e $v_1, \ldots, v_n \in G$ tais que $\sum_{i=1}^{n} \alpha_i v_i = 0$. Sem perda de generalidade, suponha que $\alpha_1 \neq 0$. Assim, $v_1 = -\frac{1}{\alpha_1} \sum_{i=2}^{n} \alpha_i v_i \in \langle G \setminus \{v_1\} \rangle$. Logo, $G \subseteq \langle v_2, \ldots, v_n \rangle$, o que implica que $V = \langle G \rangle \subseteq \langle v_2, \ldots, v_n \rangle$. Assim, $v_2, \ldots, v_n$ é um conjunto gerador de V com $n - 1 < n = \dim V$ elementos, o que é um absurdo. □

A hipótese de dimensão finita é essencial no resultado acima.

Exemplo 3.4.16. Considere o espaço vetorial $\mathbb{K}[x]$. O conjunto $L = \{1, x, x^2, \ldots\}$ é uma base de $\mathbb{K}[x]$.

O conjunto $\{x, x^2, x^3, \ldots\}$ é LI, mas não é uma base de $\mathbb{K}[x]$, pois não gera 1.

Finalmente, note que $\{0, 1, x, x^2, \ldots\}$ é um gerador de $\mathbb{K}[x]$, mas não é uma base de $\mathbb{K}[x]$, pois é linearmente dependente.

Note que esses três conjuntos bijetam entre si (são enumeráveis)

Proposição 3.4.17. Seja V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e $W \subseteq V$ um subespaço vetorial. Então $\dim W \leq \dim V$ e $\dim W = \dim V$ se, e somente se $W = V$.

Demonstração. Seja $n = \dim V$ e $m = \dim W$. Toda base de $\mathcal{W}$ é um conjunto LI de V, logo, $m \leq n$.

Se $\dim W = \dim V$, toda base de W é uma base de V, logo, $W = V$.

Reciprocamente, se $W = V$, então $\dim W = \dim V$. □

Podemos ainda estender conjuntos L.I. a bases, o que sempre é possível em geral. Para dimensão finita, também podemos fazer isso de forma concreta.

Proposição 3.4.18. Seja V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e $L \subseteq V$ linearmente independente. Seja G um conjunto gerador de V.

Então existe $G' \subseteq G$ tal que $G' \cup L$ é uma base de V.

Demonstração. Seja $n = \dim V$ e $m = |L|$. Se $m = n$, então L é uma base de V e podemos tomar $G' = \emptyset$.

Suponha que $m < n$ e seja $k = n - m$. Recursivamente, construiremos $G' = \{v_1, \ldots, v_k\} \subseteq G$ tal que $L_i = L \cup \{v_1, \ldots, v_i\}$ é LI para todo $i \in \{1, \ldots, k\}$.

Como G é um gerador de V e L não gera V, existe $v_1 \in G$ tal que $v_1 \notin \langle L \rangle$. Assim, $L_1 = L \cup \{v_1\}$ é LI. Suponha que $v_1, \ldots, v_i$ já foram construídos, com $i < k$, tais que $L_i = L \cup \{v_1, \ldots, v_i\}$ é LI. Como G é um gerador de V e L_i não gera V, existe $v_{i+1} \in G$ tal que $v_{i+1} \notin \langle L_i \rangle$. Assim, $L_{i+1} = L \cup \{v_1, \ldots, v_{i+1}\}$ é LI.

Ao final, $L \cup G'$ é um conjunto LI com $m + k = n$ elementos, logo, é uma base de V. □

Façamos outro exemplo prático.

Exemplo 3.4.19. Considere o espaço vetorial $\mathbb{R}^4$. Seja $L = \{(1,0,0,0), (1,1,0,0)\}$. Note que L é LI, mas não é uma base de $\mathbb{R}^4$, pois não gera $\mathbb{R}^4$.

Considere $G = \{(1,1,0,0), (0,1,0,0), (1,3,2,4), (1,2,2,4)\}$ e seja $W = \langle G \rangle$.

Temos que $L \subseteq W$, pois $(1,0,0,0) = (1,1,0,0) - (0,1,0,0)$. Vamos encontrar $G' \subseteq G$ tal que $L \cup G'$ é uma base de W.

Note que $|L| = 2$ e $\dim W \leq \dim \mathbb{R}^4 = 4$. Assim, estenderemos L a uma base de W com no máximo 2 vetores de G.

Será que $G \subseteq \langle L \rangle$? Sim: $(1,1,0,0) \in \langle L \rangle$ e $(0,1,0,0) \in \langle L \rangle$, pois $(0,1,0,0) = (1,1,0,0) - (1,0,0,0)$.

Porém, $(1,3,2,4) \notin \langle L \rangle$, logo, seja $\mathcal{B}' = L \cup \{(1,3,2,4)\}$.

Será que $\mathcal{B}'$ é uma base de W? Isso ocorre se, e somente se $(0,2,2,4) \in \langle \mathcal{B}' \rangle$. A resposta é positiva: resolvendo-se um sistema, vemos que $(0,2,2,4) = (1,3,2,4) - (1,1,0,0) + 1(1,0,0,0)$.

Exemplo 3.4.20. Seja $V = M_{m \times n}(\mathbb{K})$ o espaço vetorial das matrizes $m \times n$ com entradas em um corpo $\mathbb{K}$. Considere o conjunto $E_{ij} \in V$, $1 \leq i \leq m$, $1 \leq j \leq n$, onde E_{ij} é a matriz cujas entradas são todas nulas, exceto a entrada na i-ésima linha e j-ésima coluna, que é igual a 1.

Temos que $B = \{E_{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$ é uma base de V. De fato, note que para todo $A \in V$, $A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij}$, onde a_{ij} é a entrada de A na i-ésima linha e j-ésima coluna. Assim, B gera V. Além disso, B é LI, pois se $A = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} E_{ij} = 0$, então todas as entradas da matriz nula são nulas, o que implica que $\alpha_{ij} = 0$ para todo i, j. Logo, $\dim V = |B| = mn$.

3.5 Somas de espaços

Vimos que se U e W são subespaços de um espaço vetorial V, então $U \cap W$ também é um subespaço de V, e este é o maior subespaço de V contido em ambos U e W. Será que existe o menor subespaço de V que contém ambos U e W? A resposta é positiva, e este subespaço é chamado de soma de U e W.

Definição 3.5.1. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e U, W subespaços de V. A *soma* de U e W, denotada por $U + W$, é o conjunto

$$U + W = \{u + w : u \in U, w \in W\}.$$

Proposição 3.5.2. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e U, W subespaços de V. Então $U + W$ é um subespaço de V. Além disso, se U, W possuem dimensão finita, então $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$.

Demonstração. Vejamos que $U + W$ é um subespaço de V.

- $0_V \in U + W$, pois $0 = 0 + 0 \in U + W$.
- Se $x, y \in U + W$ e $\alpha \in \mathbb{K}$, então existem $u_1, u_2 \in U$ e $w_1, w_2 \in W$ tais que $x = u_1 + w_1$ e $y = u_2 + w_2$. Então:

$$x + \alpha y = (u_1 + w_1) + \alpha(u_2 + w_2) = (u_1 + \alpha u_2) + (w_1 + \alpha w_2) \in U + W,$$

Logo, $U + W$ é um subespaço de V. É imediato que U, W estão contidos em V. Se X é um subespaço de V que contém ambos U e W, então, para quaisquer $u \in U$ e $w \in W$, temos $u + w \in X$, logo $U + W \subseteq X$. Assim, $U + W$ é o menor subespaço de V que contém ambos U e W.

Agora suponha que U e W possuem dimensão finita. Então $U + W$ é finitamente gerado: sejam $\mathcal{B}_U$ e $\mathcal{B}_W$ bases de U e W, respectivamente. Então $U, W \subseteq \langle \mathcal{B}_U \cup \mathcal{B}_W \rangle$, logo, $U + W \subseteq \langle \mathcal{B}_U \cup \mathcal{B}_W \rangle$. Assim, $U + W$ é finitamente gerado.

Como $U \cap W$ é um subespaço de U, existe $\mathcal{D}$ base de $U \cap W$.

Temos que $\mathcal{D}$ é um subconjunto LI de U, logo, existe $\mathcal{D}'_U \subseteq U$ disjunto de $\mathcal{D}$ tal que $\mathcal{D}_U = \mathcal{D} \cup \mathcal{D}'_U$ é uma base de U.

Analogamente, existe $\mathcal{D}'_W \subseteq W$ disjunto de $\mathcal{D}$ tal que $\mathcal{D}_W = \mathcal{D} \cup \mathcal{D}'_W$ é uma base de W.

Afirmamos que $\mathcal{D}'_W \cap \mathcal{D}_U = \emptyset$, e que $\mathcal{C} = \mathcal{D}_U \cup \mathcal{D}'_W = \mathcal{D} \cup \mathcal{D}'_U \cup \mathcal{D}'_W$ é uma base de $U + W$.

- $\mathcal{D}'_W \cap \mathcal{D}_U = \emptyset$: de fato, se $x \in \mathcal{D}'_W \cap \mathcal{D}_U$, então $x \in W$ e $x \in U$, logo, $x \in U \cap W$. Assim, $x \in \langle \mathcal{D} \rangle$ e $x \notin \mathcal{D}$ (pois $x \in \mathcal{D}'_W$, que é disjunto de $\mathcal{D}$), logo, $\mathcal{D} \cup \{x\} \subseteq \mathcal{D}_W$ é linearmente dependente, o que é um absurdo.

- $\mathcal{C}$ é LI: seja $(\alpha_v : v \in \mathcal{C})$ uma família de escalares tal que $\sum_{v \in \mathcal{C}} \alpha_v v = 0$. Então:

$$\begin{aligned} 0 &= \sum_{v \in \mathcal{C}} \alpha_v v \\ &= \sum_{v \in \mathcal{D}} \alpha_v v + \sum_{v \in \mathcal{D}'_U} \alpha_v v + \sum_{v \in \mathcal{D}'_W} \alpha_v v \end{aligned} \tag{3.2}$$

  Assim:

$$\sum_{v \in \mathcal{D}} \alpha_v v + \sum_{v \in \mathcal{D}'_U} \alpha_v v = - \sum_{v \in \mathcal{D}'_W} \alpha_v v \in U \cap W.$$

  Logo, existem $(\beta_v : v \in \mathcal{D})$ tais que

$$\sum_{v \in \mathcal{D}'_W} -\alpha_v = \sum_{v \in \mathcal{D}} \beta_v v.$$

  Assim:

$$0 = \sum_{v \in \mathcal{D}'_W} \alpha_v + \sum_{v \in \mathcal{D}} \beta_v v. \tag{3.3}$$

  Como $\mathcal{D}$ é LD, temos que $\beta_v = 0$ para todo $v \in \mathcal{D}$ e $\alpha_v = 0$ para todo $v \in \mathcal{D}'_W$. Substituindo na Equação 3.2, temos que $\alpha_v = 0$ para todo $v \in \mathcal{D}'_U \cup \mathcal{D} = \mathcal{D}_U$, pois este último é LI.

- $\mathcal{C}$ é gerador de $U + W$: seja $x \in U + W$. Então existem $u \in U$ e $w \in W$ tais que $x = u + w$. Como $\mathcal{D}_U$ é base de U, existem $(\alpha_v : v \in \mathcal{D}_U)$ tais que $u = \sum_{v \in \mathcal{D}_U} \alpha_v v$. Analogamente, existem $(\beta_v : v \in \mathcal{D}_W)$ tais que $w = \sum_{v \in \mathcal{D}_W} \beta_v v$. Assim:

$$x = u + w = \sum_{v \in \mathcal{D}_U} \alpha_v v + \sum_{v \in \mathcal{D}_W} \beta_v v = \sum_{v \in \mathcal{D}_U} \alpha_v v + \sum_{v \in \mathcal{D}'_W} \beta_v v + \sum_{v \in \mathcal{D}} (\alpha_v + \beta_v) v.$$

Agora verifiquemos que vale a fórmula para a dimensão. Temos:

$$\begin{aligned} \dim(U + W) &= |\mathcal{C}| = |\mathcal{D}_U| + |\mathcal{D}'_W| \\ &= |\mathcal{D}_U| + (|\mathcal{D}_W| - |\mathcal{D}|) \\ &= \dim(U) + \dim(W) - \dim(U \cap W). \end{aligned}$$

□

3.6 Coordenadas

Nessa seção, introduziremos coordenadas relativas a uma base de um espaço vetorial.

Definição 3.6.1. Seja V um espaço vetorial de dimensão finita $n > 0$. Uma **base ordenada** de V é uma sequência $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ tal que $\{v_1, v_2, \ldots, v_n\}$ é uma base de V.

Note que, em particular, se $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ é uma base ordenada de V, então $v_i \neq v_j$ se $i \neq j$, caso contrário a imagem de $\mathcal{B}$, que é uma base de V, teria menos do que $n = \dim V$ elementos.

Proposição 3.6.2. Seja V um espaço vetorial de dimensão finita $n > 0$ e $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ uma sequência de elementos de V. Então, $\mathcal{B}$ é uma base ordenada de V se, e somente se, para todo $v \in V$, existem um único $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ tal que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$

Demonstração. Suponha que $\mathcal{B}$ é uma base ordenada de V. Como a imagem de $\mathcal{B}$ gera V, existem escalares $\alpha_1, \ldots, \alpha_n$ tais que $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. Suponha que existem $(\beta_1, \ldots, \beta_n)$ tais que $v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$. Então:

$$0 = v - v = (\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_n - \beta_n)v_n.$$

Como a imagem de $\mathcal{B}$ é LI, temos que $\alpha_i - \beta_i = 0$ para todo i, ou seja, $\alpha_i = \beta_i$ para todo i. Assim, $(\alpha_1, \ldots, \alpha_n) = (\beta_1, \ldots, \beta_n)$.

Reciprocamente, suponha que, para todo $v \in V$, existe um único $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ tal que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. É imediato que a imagem de $\mathcal{B}$ gera V. Temos que $v_i \neq v_j$ se $i \neq j$, caso contrário, para $v = v_i$, teríamos duas representações distintas de v. Finalmente, seja $(\alpha_1, \ldots, \alpha_n)$ tal que $\sum_{i=1}^n \alpha_i v_i = 0$. Então, como $0 \in V$, pela hipótese, $(\alpha_1, \ldots, \alpha_n) = (0, \ldots, 0)$. Assim, a imagem de $\mathcal{B}$ é LI. $\square$

Definição 3.6.3. Seja V um espaço vetorial de dimensão finita $n > 0$ e $\mathcal{B} = (v_1, \ldots, v_n)$ uma base ordenada de V.

Define-se, para cada $v \in V$, o **vetor de coordenadas** de v em relação a $\mathcal{B}$, denotado por $(v)_{\mathcal{B}}$, como o vetor $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ tal que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$.

Proposição 3.6.4. Seja V um espaço vetorial de dimensão finita $n > 0$ e $\mathcal{B} = (v_1, \ldots, v_n)$ uma base ordenada de V. Então para todos $u, v \in V$ e $\alpha \in \mathbb{K}$, temos:

(i) $(u + v)_{\mathcal{B}} = (u)_{\mathcal{B}} + (v)_{\mathcal{B}}$.

(ii) $(\alpha v)_{\mathcal{B}} = \alpha (v)_{\mathcal{B}}$.

Demonstração. Seja $(u)_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)$ e $(v)_{\mathcal{B}} = (\beta_1, \ldots, \beta_n)$. Então:

$$u + v = \sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^n \beta_i v_i = \sum_{i=1}^n (\alpha_i + \beta_i) v_i,$$

logo, $(u + v)_{\mathcal{B}} = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) = (u)_{\mathcal{B}} + (v)_{\mathcal{B}}$.

Além disso:

$$\alpha v = \alpha \sum_{i=1}^n \beta_i v_i = \sum_{i=1}^n (\alpha \beta_i) v_i,$$

logo, $(\alpha v)_{\mathcal{B}} = (\alpha \beta_1, \ldots, \alpha \beta_n) = \alpha (v)_{\mathcal{B}}$. $\square$

Capítulo 4

Transformações Lineares

Nesse capítulo estudaremos transformações lineares

4.1 Transformações lineares

Definição 4.1.1. Seja V e W espaços vetoriais sobre um mesmo corpo $\mathbb{K}$. Uma função $T : V \to W$ é uma **transformação linear** se, para quaisquer $u, v \in V$ e $\alpha \in \mathbb{K}$, tivermos:

- $T(u + v) = T(u) + T(v)$
- $T(\alpha u) = \alpha T(u)$

Exemplo 4.1.2. A função $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (2x, 3y)$ é uma transformação linear.

Proposição 4.1.3. Se $T : V \to W$ é uma transformação linear, então $T(0_V) = 0_W$ e $T(-v) = -T(v)$ para todo $v \in V$.

Demonstração. Temos que $T(0_V) = T(0_V + 0_V) = T(0_V) + T(0_V)$, logo $T(0_V) = 0_W$. Além disso, $T(-v) = T(-1 \cdot v) = -1 \cdot T(v) = -T(v)$. □

Definição 4.1.4. O *núcleo* de uma transformação linear $T : V \to W$ é o conjunto $\ker(T) = \{v \in V : T(v) = 0_W\}$. A *imagem* de T é o conjunto $\mathrm{Im}(T) = \{T(v) : v \in V\}$.

4.2 O núcleo e a imagem de uma transformação linear

Definição 4.2.1. Seja $T : V \to W$ uma transformação linear. O *núcleo* de T é o conjunto $\ker(T) = \{v \in V : T(v) = 0_W\}$. A *imagem* de T é o conjunto $\mathrm{Im}(T) = \{T(v) : v \in V\}$.

Proposição 4.2.2. Seja $T : V \to W$ uma transformação linear. Então, $\ker(T)$ é um subespaço de V e $\mathrm{Im}(T)$ é um subespaço de W.

Lema 4.2.3. Seja $T : V \to W$ uma transformação linear. Então, $\ker(T)$ é um subespaço de V e $\mathrm{Im}(T)$ é um subespaço de W.

Demonstração. Vejamos que $\ker(T)$ é um subespaço de V.

- $0_V \in \ker(T)$, pois $T(0_V) = 0_W$.
- Se $u, v \in \ker(T)$ e $\alpha \in \mathbb{K}$, então $T(u) = 0_W$ e $T(v) = 0_W$. Temos o seguinte:

$$T(u + \alpha v) = T(u) + \alpha T(v) = 0_W + \alpha 0_W = 0_W.$$

Portanto, $\ker(T)$ é um subespaço de V. Vejamos que $\mathrm{Im}(T)$ é um subespaço de W.

- $0_W \in \mathrm{Im}(T)$, pois $T(0_V) = 0_W$.
- Se $x, y \in \mathrm{Im}(T)$ e $\alpha \in \mathbb{K}$, então existem $u, v \in V$ tais que $x = T(u)$ e $y = T(v)$. Temos o seguinte:

$$x + \alpha y = T(u) + \alpha T(v) = T(u + \alpha v) \in \mathrm{Im}(T),$$

□

Lema 4.2.4. Seja $T : V \to W$ uma transformação linear. Então T é injetora se, e somente se, para todo conjunto linearmente independente $L \subseteq V$, o conjunto $T[L] = \{T(v) : v \in L\}$ é linearmente independente em W.

Demonstração. Seja T injetora. Seja $L \subseteq V$ um conjunto linearmente independente. Sejam $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ e $v_1, \ldots, v_n \in L$ distintos tais que $\sum_{i=1}^n \alpha_i T(v_i) = 0_W$. Então $T\left(\sum_{i=1}^n \alpha_i v_i\right) = 0_W$. Como T é injetora, temos que $\sum_{i=1}^n \alpha_i v_i = 0_V$. Como L é linearmente independente, temos que $\alpha_i = 0$ para todo i. Assim, $T[L]$ é linearmente independente.

Reciprocamente, se T não é injetora, então existe $v \in V$, $v \neq 0_V$, tal que $T(v) = 0_W$. Então, o conjunto $L = \{v\}$ é linearmente independente, mas $T[L] = \{0_W\}$ não é linearmente independente. □

Lema 4.2.5. Seja $T : V \to W$ uma transformação linear. Então T é sobrejetora se, e somente se, para todo conjunto gerador $G \subseteq W$, o conjunto $T[G] = \{T(v) : v \in G\}$ é um conjunto gerador de W.

Demonstração. Seja T sobrejetora. Seja $G \subseteq V$ um conjunto gerador de V. Seja $w \in W$. Como T é sobrejetora, existe $v \in V$ tal que $T(v) = w$. Como G gera V, existem $u_1, \ldots, u_m \in G$ e $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ tais que $v = \sum_{i=1}^m \alpha_i u_i$. Então:

$$w = T(v) = T\left(\sum_{i=1}^m \alpha_i u_i\right) = \sum_{i=1}^m \alpha_i T(u_i),$$

logo, $T[G]$ gera W.

Reciprocamente, se T não é sobrejetora, V é um conjunto gerador de V tal que $T[V] = \mathrm{Im}(T)$ não gera W. □

Teorema 4.2.6 (Teorema do núcleo e da imagem)**.** Seja $T : V \to W$ uma transformação linear. Então:

$$\dim(V) = \dim(\ker(T)) + \dim(\mathrm{Im}(T)).$$

Demonstração. Veremos a prova apenas para dimensão finita. Seja $\mathcal{B}$ uma base de $\ker(T)$. Seja $\mathcal{C}$ um conjunto disjunto de $\mathcal{B}$ tal que $\mathcal{B} \cup \mathcal{C}$ é uma base de V. Mostraremos que $T[\mathcal{C}]$ é uma base de $\mathrm{Im}(T)$ e que $T|_{\mathcal{C}}$ é injetora.

Sejam $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ e $v_1, \ldots, v_n \in \mathcal{C}$ distintos tais que $\sum_{i=1}^n \alpha_i T(v_i) = 0_W$. Então $T\left(\sum_{i=1}^n \alpha_i v_i\right) = 0_W$. Como $\mathcal{B}$ é base de $\ker(T)$, temos que $\sum_{i=1}^n \alpha_i v_i \in \langle \mathcal{B} \rangle$. Logo, existem $(\beta_j)_{j<m}$ tais que $\sum_{i=1}^n \alpha_i v_i = \sum_{j<m} \beta_j b_j$. Como $\mathcal{B} \cup \mathcal{C}$ é LI, temos que $\alpha_i = 0$ para todo i.

Assim, $T[\mathcal{C}]$ é LI. Além disso, $T|_{\mathcal{C}}$ é injetora. pois se $T(v) = T(u)$ para $u, v \in \mathcal{C}$, então $T(v) - T(u) = T(v - u) = 0_W$, e, pelo argumento anterior, $1 = -1 = 0$, um absurdo.

Resta ver que $T[\mathcal{C}]$ gera $\mathrm{Im}(T)$. Seja $w \in \mathrm{Im}(T)$. Então, existe $v \in V$ tal que $T(v) = w$. Como $\mathcal{B} \cup \mathcal{C}$ é base de V, existem $(\alpha_i)_{i<n}$ e $(\beta_j)_{j<m}$ tais que $v = \sum_{i<n} \alpha_i b_i + \sum_{j<m} \beta_j c_j$, com $b_i \in \mathcal{B}$ e $c_j \in \mathcal{C}$. Assim:

$$w = T(v) = T\left(\sum_{i<n} \alpha_i b_i + \sum_{j<m} \beta_j c_j\right) = \sum_{i<n} \alpha_i T(b_i) + \sum_{j<m} \beta_j T(c_j).$$

Como $T(b_i) = 0_W$ para todo i, temos que $w = \sum_{j<m} \beta_j T(c_j)$, logo, $T[\mathcal{C}]$ gera $\mathrm{Im}(T)$. Portanto, temos:

$$\dim(V) = |\mathcal{B}| + |\mathcal{C}| = \dim(\ker(T)) + \dim(\mathrm{Im}(T)).$$

□

4.3 O espaço das Transformações Linares

Até então, estudamos transformações lineares individualmente. Nesta seção, veremos que elas, por si só, formam um espaço vetorial.

Definição 4.3.1. Sejam V, W espaços vetoriais. O conjunto de todas as transformações lineares de V em W é denotado por $L(V, W)$.

Dadas $S, T \in L(V, W)$ e $\alpha \in \mathbb{K}$, definimos:

- A soma $S + T \in L(V, W)$ por $(S + T)(v) = S(v) + T(v)$, para todo $v \in V$.

- O produto $\alpha T \in L(V, W)$ por $(\alpha T)(v) = \alpha(T(v))$, para todo $v \in V$.
- A transformação linear nula é dada por $0_{L(V,W)}(v) = 0_W$, para todo $v \in V$.

Note que a noção de soma e produto de função por escalar é coerente com a definição que se dá em outras áreas básicas da matemática, como o cálculo diferencial e integral. É imediato verificar que a transformação nula é uma transformação linear. Como de costume, quando claro pelo contexto, ela será denotada apenas por 0.

Exemplo 4.3.2. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (2x, 3y)$ e $S : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $S(x, y) = (x + y, xy)$. Então, a soma $T + S : \mathbb{R}^2 \to \mathbb{R}^2$ é dada por:

$$(T + S)(x, y) = T(x, y) + S(x, y) = (2x, 3y) + (x + y, xy) = (3x + y, 3y + xy).$$

Além disso, o produto $2T : \mathbb{R}^2 \to \mathbb{R}^2$ é dado por:

$$(2T)(x, y) = 2(T(x, y)) = 2(2x, 3y) = (4x, 6y).$$

Proposição 4.3.3. O conjunto $L(V, W)$, munido das operações de soma e produto por escalar definidas acima, é um espaço vetorial sobre o corpo $\mathbb{K}$ com elemento neutro 0. Em particular, a soma de transformações lineares é uma transformação linear, e o produto de uma transformação linear por um escalar é uma transformação linear.

Além disso, dada $T \in L(V, W)$, o oposto de T é a transformação linear $-T \in L(V, W)$ definida por $(-T)(v) = -(T(v))$ para todo $v \in V$.

Demonstração. Vamos começar verificando a última afirmação. Sejam $S, T \in L(V, W)$ e $\alpha \in \mathbb{K}$. Para quaisquer $u, v \in V$ e $\beta \in \mathbb{K}$, temos o seguinte.

- Para a soma, temos:

$$\begin{aligned}(S + T)(u + \beta v) &= S(u + \beta v) + T(u + \beta v) \\ &= S(u) + \beta S(v) + T(u) + \beta T(v) \\ &= (S(u) + T(u)) + \beta(S(v) + T(v)) \\ &= (S + T)(u) + \beta(S + T)(v).\end{aligned}$$

- Para o produto por escalar, temos:

$$\begin{aligned}(\alpha T)(u + \beta v) &= \alpha T(u + \beta v) \\ &= \alpha(T(u) + \beta T(v)) \\ &= \alpha T(u) + \alpha\beta T(v) \\ &= (\alpha T)(u) + \beta(\alpha T)(v).\end{aligned}$$

Agora verificaremos todas as propriedades de espaço vetorial para $L(V, W)$:

- (Associatividade da soma) Para quaisquer $R, S, T \in L(V, W)$ e $v \in V$, temos:

$$\begin{aligned}((R + S) + T)(v) &= (R + S)(v) + T(v) \\ &= (R(v) + S(v)) + T(v) \\ &= R(v) + (S(v) + T(v)) \\ &= R(v) + (S + T)(v) \\ &= (R + (S + T))(v).\end{aligned}$$

  Logo, $(R + S) + T = R + (S + T)$.

- (Comutatividade da soma) Para quaisquer $S, T \in L(V, W)$ e $v \in V$, temos:

$$(S + T)(v) = S(v) + T(v) = T(v) + S(v) = (T + S)(v).$$

  Logo, $S + T = T + S$.

- (Elemento neutro da soma) Para qualquer $T \in L(V, W)$ e $v \in V$, temos:

$$(T + 0)(v) = T(v) + 0_W = T(v).$$

Logo, $T + 0 = T$.

- (Elemento oposto da soma) Para qualquer $T \in L(V, W)$, definimos $-T \in L(V, W)$ por $(-T)(v) = -(T(v))$ para todo $v \in V$. Note que $-T$=$(-1)T$, assim, T é linear. Para qualquer $v \in V$, temos:

$$(T + (-T))(v) = T(v) + (-T)(v) = T(v) - T(v) = 0_W.$$

Logo, $T + (-T) = 0$.

- (Neutralidade do produto por escalar) Para qualquer $T \in L(V, W)$ e $v \in V$, temos:

$$(1T)(v) = 1(T(v)) = T(v).$$

Logo, $1T = T$.

- (Associatividade do produto por escalar) Para quaisquer $\alpha, \beta \in \mathbb{K}$, $T \in L(V, W)$ e $v \in V$, temos:

$$((\alpha\beta)T)(v) = (\alpha\beta)(T(v)) = \alpha(\beta(T(v))) = (\alpha(\beta T))(v).$$

Logo, $(\alpha\beta)T = \alpha(\beta T)$.

- (Distributividade do produto por escalar em relação à soma de vetores) Para qualquer $\alpha \in \mathbb{K}$, $S, T \in L(V, W)$ e $v \in V$, temos:

$$\begin{aligned}(\alpha(S + T))(v) &= \alpha((S + T)(v)) \\ &= \alpha(S(v) + T(v)) \\ &= \alpha S(v) + \alpha T(v) \\ &= (\alpha S)(v) + (\alpha T)(v) \\ &= (\alpha S + \alpha T)(v).\end{aligned}$$

Logo, $\alpha(S + T) = \alpha S + \alpha T$.

- (Distributividade do produto por escalar em relação à soma de escalares) Para quaisquer $\alpha, \beta \in \mathbb{K}$, $T \in L(V, W)$ e $v \in V$, temos:

$$((\alpha + \beta)T)(v) = (\alpha + \beta)(T(v)) = \alpha(T(v)) + \beta(T(v)) = (\alpha T)(v) + (\beta T)(v) = (\alpha T + \beta T)(v).$$

Logo, $(\alpha + \beta)T = \alpha T + \beta T$.

□

Assim, o espaço das transformações lineares entre dois espaços vetoriais é, ele próprio, um espaço vetorial. Note que, ao olhar para ele, cada transformação linear é vista como um vetor desse espaço. Assim, é possível falar de transformações lineares que possuem esse espaço como seu domínio, uma vez que todo espaço vetorial pode ser o domínio de uma transformação linear.

Não vamos lidar com isso nesse momento, e sim estudar outra operação entre funções: a composição.

Proposição 4.3.4. Sejam U, V, W espaços vetoriais e $S \in L(U, V)$ e $T \in L(V, W)$. Então $T \circ S \in L(U, W)$. Ou seja, a composição de transformações lineares é uma transformação linear.

Demonstração. Seja $u, v \in U$ e $\alpha \in \mathbb{K}$. Temos:

$$\begin{aligned}(T \circ S)(u + \alpha v) &= T(S(u + \alpha v)) \\ &= T(S(u) + \alpha S(v)) \\ &= T(S(u)) + \alpha T(S(v)) \\ &= (T \circ S)(u) + \alpha(T \circ S)(v).\end{aligned}$$

Logo, $T \circ S$ é uma transformação linear. □

Proposição 4.3.5. Sejam U, V, W, X espaços vetoriais, $R \in L(W, X)$, $S \in L(V, W)$ e $T \in L(U, V)$. Então:

- (Associatividade da composição) $R \circ (S \circ T) = (R \circ S) \circ T$.
- (Elemento neutro da composição) Se $\mathrm{id}_V : V \to V$ e $\mathrm{id}_W : W \to W$ são as transformações identidades em V e W, então, para qualquer $T \in L(U, V)$ e qualquer $S \in L(V, W)$, temos $\mathrm{id}_V \circ T = T$ e $S \circ \mathrm{id}_W = S$.
- (Distributividade à direita) Para quaisquer $S_1, S_2 \in L(V, W)$ e qualquer $T \in L(U, V)$, temos $(S_1 + S_2) \circ T = S_1 \circ T + S_2 \circ T$.
- (Distributividade à esquerda) Para quaisquer $S \in L(V, W)$ e $T_1, T_2 \in L(U, V)$, temos $S \circ (T_1 + T_2) = S \circ T_1 + S \circ T_2$.
- (Compatibilidade com o produto por escalar) Para qualquer $S \in L(V, W)$, qualquer $T \in L(U, V)$ e qualquer $\alpha \in \mathbb{K}$, temos $(\alpha S) \circ T = \alpha(S \circ T)$.

Demonstração. As demonstrações são diretas, baseando-se na definição de composição de funções e nas propriedades das transformações lineares. Os dois primeiros itens são propriedades gerais de composição de funções, não dependendo da linearidade, e são deixadas como exercício.

Provaremos as outras.

- (Distributividade à direita) Sejam $S_1, S_2 \in L(V, W)$, $T \in L(U, V)$ e $u \in U$. Temos:

$$\begin{aligned} ((S_1 + S_2) \circ T)(u) &= (S_1 + S_2)(T(u)) \\ &= S_1(T(u)) + S_2(T(u)) \\ &= (S_1 \circ T)(u) + (S_2 \circ T)(u) \\ &= (S_1 \circ T + S_2 \circ T)(u). \end{aligned}$$

  Logo, $(S_1 + S_2) \circ T = S_1 \circ T + S_2 \circ T$.

- (Distributividade à esquerda) Sejam $S \in L(V, W)$, $T_1, T_2 \in L(U, V)$ e $u \in U$. Temos:

$$\begin{aligned} (S \circ (T_1 + T_2))(u) &= S((T_1 + T_2)(u)) \\ &= S(T_1(u) + T_2(u)) \\ &= S(T_1(u)) + S(T_2(u)) \\ &= (S \circ T_1)(u) + (S \circ T_2)(u) \\ &= (S \circ T_1 + S \circ T_2)(u). \end{aligned}$$

  Logo, $S \circ (T_1 + T_2) = S \circ T_1 + S \circ T_2$.

- (Compatibilidade com o produto por escalar) Sejam $S \in L(V, W)$, $T \in L(U, V)$, $\alpha \in \mathbb{K}$ e $u \in U$. Temos:

$$\begin{aligned} ((\alpha S) \circ T)(u) &= (\alpha S)(T(u)) \\ &= \alpha S(T(u)) \\ &= \alpha(S \circ T)(u). \end{aligned}$$

  Logo, $(\alpha S) \circ T = \alpha(S \circ T)$.

□

4.4 Isomorfismos

Conforme já mencionado, um isomorfismo é uma transformação linear bijetora. Se há um isomorfismo entre dois espaços vetoriais, dizemos que eles são isomorfos. Nesse caso, no que diz respeito à estrutura de espaço vetorial, eles são essencialmente o mesmo espaço, e a função isomorfismo pode ser vista como um dispositivo de "tradução" entre os dois espaços.

A relação de isomorfismo entre espaços vetoriais é uma relação de equivalência na classe dos espaços vetoriais sobre um mesmo corpo.

Proposição 4.4.1. Sejam U, V e W espaços vetoriais sobre um corpo $\mathbb{K}$. Então:

(a) $\mathrm{id}_U : U \to U$ é um isomorfismo. Portanto, U é isomorfo a si mesmo.

(b) Se $T : U \to V$ é um isomorfismo e $S : V \to W$ é um isomorfismo, então $S \circ T : U \to W$ é um isomorfismo. Portanto, se U é isomorfo a V e V é isomorfo a W, então U é isomorfo a W.

(c) Se $T : U \to V$ é um isomorfismo, então sua inversa $T^{-1} : V \to U$ é um isomorfismo. Portanto, se U é isomorfo a V, então V é isomorfo a U.

Demonstração. (a) É claro que id_U é bijetora é linear. Para (b), é um fato geral que a composição de funções bijetoras é bijetora, e, como já vimos, a composição de funções lineares é linear.

Para (c), se T é um isomorfismo, então sua inversa T^{-1} também é bijetora. Resta ver que T^{-1} é linear.

Sejam $u, v \in V$ e $\alpha \in \mathbb{K}$. Então existem $x, y \in U$ tais que $T(x) = u$ e $T(y) = v$. Assim:

$$\begin{aligned} T^{-1}(u + \alpha v) &= T^{-1}(T(x) + \alpha T(y)) \\ &= T^{-1}(T(x + \alpha y)) \\ &= x + \alpha y \\ &= T^{-1}(u) + \alpha T^{-1}(v). \end{aligned}$$

□

Abaixo, daremos um importante exemplo de isomorfismo.

Definição 4.4.2. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$. Para cada $x \in V$, o vetor coluna das coordenadas de x na base ordenada $\mathcal{B}$, denotado por $[x]_{\mathcal{B}}$, é dado por:

$$[x]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

em que $(\alpha_1, \alpha_2, \ldots, \alpha_n) = (x)_{\mathcal{B}}$, ou seja, em que $x = \sum_{i=1}^{n} \alpha_i v_i$ e $\mathcal{B} = (v_1, v_2, \ldots, v_n)$.

Da mesma forma que a transformação $x \mapsto (x)_{\mathcal{B}}$ é um isomorfismo entre V e $\mathbb{K}^n$, a transformação $x \mapsto [x]_{\mathcal{B}}$ também é um isomorfismo entre V e $M_{n\times 1}(\mathbb{K})$.

Lema 4.4.3. Sejam V um espaço vetorial sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$, e $\mathcal{B}$ uma base ordenada de V. A transformação $[\,\cdot\,]_{\mathcal{B}} : V \to M_{n\times 1}(\mathbb{K})$ é um isomorfismo.

Demonstração. Verificaremos a linearidade. Sejam $u, w \in V$ e $\beta \in \mathbb{K}$, com $(\alpha_1, \ldots, \alpha_n) = (u)_{\mathcal{B}}$ e $(\gamma_1, \ldots, \gamma_n) = (w)_{\mathcal{B}}$.

Sejam as coordenadas de u e w na base $\mathcal{B}$ dadas por $(\alpha_1, \ldots, \alpha_n)$ e $(\gamma_1, \ldots, \gamma_n)$, respectivamente. Então as coordenadas de $u + w$ e de βw na base $\mathcal{B}$ são dadas por $(\alpha_1 + \gamma_1, \ldots, \alpha_n + \gamma_n)$ e $(\beta\gamma_1, \ldots, \beta\gamma_n)$, respectivamente. Logo:

$$\begin{aligned} [u + \beta w]_{\mathcal{B}} &= \begin{pmatrix} \alpha_1 + \beta\gamma_1 \\ \alpha_2 + \beta\gamma_2 \\ \vdots \\ \alpha_n + \beta\gamma_n \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \beta \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix} \\ &= [u]_{\mathcal{B}} + \beta [w]_{\mathcal{B}}. \end{aligned}$$

O que prova que $[\,\cdot\,]_{\mathcal{B}}$ é linear.

Para a injetividade, veremos que o núcleo é trivial. Seja $u \in V$ tal que $[u]_{\mathcal{B}} = 0$. Então as coordenadas de u na base $\mathcal{B}$ são todas nulas, ou seja, $u = 0$.

Para a sobrejetividade, temos, pelo Teorema do Núcleo e Imagem, que $\dim(\mathrm{Im}([\,\cdot\,]_{\mathcal{B}})) = \dim(V) = n = \dim(M_{n\times 1}(\mathbb{K}))$. □

4.5 Matrizes de transformações lineares

Quem são as transformações lineares entre espaços vetoriais finitamente gerados?
Primeiro, estudemos uma propriedade relativa a bases.

Teorema 4.5.1. Sejam V e W espaços vetoriais sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$ e $\mathcal{B}$ uma base de V. Então, para toda $f : \mathcal{B} \to W$ função, existe uma única transformação linear $T : V \to W$ tal que $T(v) = f(v)$ para todo $v \in \mathcal{B}$.

Demonstração. Ordenemos a base $\mathcal{B}$ como $\mathcal{C} = (v_1, \ldots, v_n)$.
Seja $u \in V$ e $(\alpha_1, \ldots, \alpha_n) = (u)_\mathcal{C}$. Definimos:

$$T(u) = \sum_{i=1}^{n} \alpha_i f(v_i).$$

T é linear, pois dados $u, w \in V$ e $\beta \in \mathbb{K}$, com $(\alpha_1, \ldots, \alpha_n) = (u)_\mathcal{C}$ e $(\gamma_1, \ldots, \gamma_n) = (w)_\mathcal{C}$, temos:

$$\begin{aligned} T(u + \beta w) &= T\left(\sum_{i=1}^{n} (\alpha_i + \beta\gamma_i) v_i\right) \\ &= \sum_{i=1}^{n} (\alpha_i + \beta\gamma_i) f(v_i) \\ &= \sum_{i=1}^{n} \alpha_i f(v_i) + \beta \sum_{i=1}^{n} \gamma_i f(v_i) \\ &= T(u) + \beta T(w). \end{aligned}$$

Temos que T estende f, pois as coordenadas de v_j na base $\mathcal{C}$ são $(0, \ldots, 0, 1, 0, \ldots, 0)$ (com 1 na j-ésima posição), logo:

$$T(v_j) = \sum_{i=1}^{n} \delta_{ij} f(v_i) = f(v_j).$$

Finalmente, seja $S : V \to W$ outra transformação linear que estende f. Seja $u \in V$ e $(\alpha_1, \ldots, \alpha_n) = (u)_\mathcal{C}$. Então:

$$S(u) = S\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i S(v_i) = \sum_{i=1}^{n} \alpha_i f(v_i) = T(u).$$

Assim, $S = T$. □

Quem são todas as transformações lineares de $\mathbb{K}^n$ em $\mathbb{K}$?

Exemplo 4.5.2. Seja $T : \mathbb{K}^n \to \mathbb{K}$ uma transformação linear. Temos que $T(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i T(e_i)$, onde e_i é o i-ésimo vetor da base canônica de $\mathbb{K}^n$.

Logo, para qualquer escolha $T(e_i) = a_i \in \mathbb{K}$, existe uma única transformação linear $T : \mathbb{K}^n \to \mathbb{K}$ dada por $T(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i$.

Essas são todas as transformações lineares de $\mathbb{K}^n$ em $\mathbb{K}$.

Quem são todas as transformações lineares de $\mathbb{K}^n$ em $\mathbb{K}^m$?

Exemplo 4.5.3. Seja $T : \mathbb{K}^n \to \mathbb{K}^m$ uma transformação linear. Temos que $T(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i T(e_j)$, onde e_i é o i-ésimo vetor da base canônica de $\mathbb{K}^n$.

Logo, para qualquer escolha $T(e_j) = (a_{1j}, a_{2j}, \ldots, a_{mj}) \in \mathbb{K}^m$, existe uma única transformação linear $T : \mathbb{K}^n \to \mathbb{K}^m$ dada por

$$T(x_1, \ldots, x_n) = \left(\sum_{j=1}^{n} a_{1j} x_j, \sum_{j=1}^{n} a_{2j} x_j, \ldots, \sum_{j=1}^{n} a_{mj} x_j\right).$$

Note que, colocando os vetores em formato de coluna, temos que as coordenadas de $T(x_1, \ldots, x_n)$ são dadas pela seguinte coluna:

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

Abaixo, estenderemos esse raciocínio para transformações lineares entre espaços vetoriais quaisquer de dimensão finita positiva.

Definição 4.5.4. Sejam V e W espaços vetoriais sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$ e $1 \leq \dim W = m < \infty$. Sejam $\mathcal{B} = (v_1, \ldots, v_n)$ uma base ordenada de V e $\mathcal{C} = (w_1, \ldots, w_m)$ uma base ordenada de W. A *matriz de T em relação às bases $\mathcal{B}$ e $\mathcal{C}$* é a matriz $[T]_{\mathcal{BC}} \in M_{m\times n}(\mathbb{K})$ cuja j-ésima coluna é dada por $[T(v_j)]_{\mathcal{C}}$, ou seja:

$$[T]_{\mathcal{BC}} = \begin{pmatrix} | & | & & | \\ [T(v_1)]_{\mathcal{C}} & [T(v_2)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & | & & | \end{pmatrix}.$$

Formalmente, para cada $j \in \{1, \ldots, n\}$, seja $(\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{mj}) = (T(v_j))_{\mathcal{C}}$. Então:

$$[T]_{\mathcal{BC}} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}.$$

O mesmo fenômeno descrito anteriormente ocorre aqui, como expressado formalmente pela seguinte proposição.

Proposição 4.5.5. Sejam V e W espaços vetoriais sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$ e $1 \leq \dim W = m < \infty$. Sejam $\mathcal{B} = (v_1, \ldots, v_n)$ uma base ordenada de V e $\mathcal{C} = (w_1, \ldots, w_m)$ uma base ordenada de W. Para toda transformação linear $T : V \to W$ e todo $x \in V$, temos:

$$[T]_{\mathcal{BC}}[x]_{\mathcal{B}} = [T(x)]_{\mathcal{C}}.$$

Além disso, $[T]_{\mathcal{BC}}$ é a única matriz em $M_{m\times n}(\mathbb{K})$ que satisfaz essa propriedade. Ou seja, se $A \in M_{m\times n}(\mathbb{K})$ é tal que $A[x]_{\mathcal{B}} = [T(x)]_{\mathcal{C}}$ para todo $x \in V$, então $A = [T]_{\mathcal{BC}}$.

Demonstração. Seja $x \in V$ e $(\alpha_1, \ldots, \alpha_n) = (x)_{\mathcal{B}}$. Temos:

$$\begin{aligned} [T]_{\mathcal{BC}}\,[x]_{\mathcal{B}} &= \begin{pmatrix} | & | & & | \\ [T(v_1)]_{\mathcal{C}} & [T(v_2)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & | & & | \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \\ &= \sum_{i=1}^{n} \alpha_i [T(v_i)]_{\mathcal{C}} \\ &= \left[\sum_{i=1}^{n} \alpha_i T(v_i) \right]_{\mathcal{C}} \\ &= \left[T\left(\sum_{i=1}^{n} \alpha_i v_i \right) \right]_{\mathcal{C}} \\ &= [T(x)]_{\mathcal{C}}. \end{aligned}$$

Agora, seja $A \in M_{m\times n}(\mathbb{K})$ tal que $A[x]_{\mathcal{B}} = [T(x)]_{\mathcal{C}}$ para todo $x \in V$. Seja $j \in \{1, \ldots, n\}$. Então, tomando $x = v_j$, temos que $(v_j)_{\mathcal{B}}$ é o vetor coluna com 1 na j-ésima posição e 0 nas demais, ou seja, $(v_j)_{\mathcal{B}} = (\delta_{ij})_{i=1}^{n}$.

$$A[e_j] = A[v_j]_{\mathcal{B}} = [T(v_j)]_{\mathcal{C}},$$

onde e_j é o j-ésimo vetor da base canônica de $\mathbb{K}^n$. Logo, sendo $A_1, A_2, \ldots, A_n$ as colunas de A, temos que:

$$
\begin{aligned}
A[v_j]_{\mathcal{B}} &= \begin{pmatrix} | & | & & | \\ A_1 & A_2 & \cdots & A_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} \delta_{1j} \\ \delta_{2j} \\ \vdots \\ \delta_{nj} \end{pmatrix} \\
&= \sum_{i=1}^{n} \delta_{ij} A_i \\
&= A_j.
\end{aligned}
$$

Por outro lado, $A[v_j]_{\mathcal{B}} = [T(v_j)]_{\mathcal{C}}$. Assim, a j-ésima coluna de A é igual à j-ésima coluna de $[T]_{\mathcal{BC}}$. Como isso vale para todo $j \in \{1, \ldots, n\}$, concluímos que $A = [T]_{\mathcal{BC}}$. □

Dessa forma, toda transformação linear entre espaços vetoriais finitamente gerados pode ser representada por uma matriz, desde que escolhidas bases adequadas para os espaços envolvidos. Isso mostra que é possível estabelecer uma correspondência entre transformações lineares e matrizes, o que é fundamental para o estudo e aplicação de transformações lineares em diversas áreas da matemática e suas aplicações.

Essa correspondência vai além. Em verdade, ela é um isomorfismo entre o espaço vetorial das transformações lineares de V em W e o espaço vetorial das matrizes $M_{m\times n}(\mathbb{K})$.

Teorema 4.5.6. Sejam V e W espaços vetoriais sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$ e $1 \leq \dim W = m < \infty$. Sejam $\mathcal{B} = (v_1, \ldots, v_n)$ uma base ordenada de V e $\mathcal{C} = (w_1, \ldots, w_m)$ uma base ordenada de W. A aplicação $\Phi : \mathcal{L}(V, W) \to M_{m\times n}(\mathbb{K})$ dada por $\Phi(T) = [T]_{\mathcal{BC}}$ é um isomorfismo de espaços vetoriais.

Demonstração. Verificaremos a linearidade, injetividade e sobrejetividade de Φ.

Sejam $T, S \in \mathcal{L}(V, W)$ e $\beta \in \mathbb{K}$. Veremos que $[T + \beta S]_{\mathcal{BC}} = [T]_{\mathcal{BC}} + \beta[S]_{\mathcal{BC}}$. Para isso, veremos que, sendo $A = [T]_{\mathcal{BC}} + \beta[S]_{\mathcal{BC}}$, temos $A[x]_{\mathcal{B}} = [(T + \beta S)(x)]_{\mathcal{C}}$ para todo $x \in V$.

Com efeito, temos:

$$
\begin{aligned}
A[x]_{\mathcal{B}} &= \left([T]_{\mathcal{BC}} + \beta[S]_{\mathcal{BC}}\right)[x]_{\mathcal{B}} \\
&= [T]_{\mathcal{BC}}[x]_{\mathcal{B}} + \beta[S]_{\mathcal{BC}}[x]_{\mathcal{B}} \\
&= [T(x)]_{\mathcal{C}} + \beta[S(x)]_{\mathcal{C}} \\
&= [T(x) + \beta S(x)]_{\mathcal{C}} \\
&= [(T + \beta S)(x)]_{\mathcal{C}}.
\end{aligned}
$$

Assim, pela unicidade da matriz que satisfaz essa propriedade, temos que $A = [T + \beta S]_{\mathcal{BC}}$. Logo, Φ é linear.

Vejamos que Φ é injetora. Basta ver que $\ker \Phi = \{0\}$, ou seja, que se $[T]_{\mathcal{BC}} = 0$, então $T = 0$. Se $[T]_{\mathcal{BC}} = 0$, então, para todo $v \in V$, temos $[T(v)]_{\mathcal{C}} = [T]_{\mathcal{BC}}[v]_{\mathcal{B}} = 0$. Como o mapa $[\,\cdot\,]_{\mathcal{C}}$ é um isomorfismo, isso implica que $T(v) = 0$ para todo $v \in V$, ou seja, $T = 0$. Portanto, Φ é injetora.

Resta ver que Φ é sobrejetora. Seja $A \in M_{m\times n}(\mathbb{K})$. Definimos $f : \mathcal{B} \to W$ como a função que associa a cada $v_j \in \mathcal{B}$ o vetor cujas coordenadas na base $\mathcal{C}$ são dadas pela j-ésima coluna de A, ou seja, $f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$, onde a_{ij} é o elemento da i-ésima linha e j-ésima coluna de A. Pela caracterização por bases, existe uma única transformação linear $T : V \to W$ que estende f. Pela definição de f, temos que $[T(v_j)]_{\mathcal{C}}$ é a j-ésima coluna de A. Logo, $[T]_{\mathcal{BC}} = A$. □

Corolário 4.5.7. Sejam V e W espaços vetoriais sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$ e $1 \leq \dim W = m < \infty$. Então $\dim L(V, W) = mn$.

Demonstração. Pelo teorema anterior, $\mathcal{L}(V, W)$ é isomorfo a $M_{m\times n}(\mathbb{K})$. Como $\dim M_{m\times n}(\mathbb{K}) = mn$, temos que $\dim \mathcal{L}(V, W) = mn$. □

Tal isomorfismo também respeita composições, transformando-as em multiplicações de matrizes.

Teorema 4.5.8. Sejam U, V e W espaços vetoriais sobre um corpo $\mathbb{K}$, com $1 \leq \dim U = p < \infty$, $1 \leq \dim V = n < \infty$ e $1 \leq \dim W = m < \infty$. Sejam $\mathcal{A} = (u_1, \ldots, u_p)$ uma base ordenada de U, $\mathcal{B} = (v_1, \ldots, v_n)$ uma base ordenada de V e $\mathcal{C} = (w_1, \ldots, w_m)$ uma base ordenada de W. Sejam $T : U \to V$ e $S : V \to W$ transformações lineares. Então:

$$[S \circ T]_{\mathcal{AC}} = [S]_{\mathcal{BC}}[T]_{\mathcal{AB}}.$$

Demonstração. Seja $u \in U$. Temos:

$$\begin{aligned}[S]_{\mathcal{BC}}\,[T]_{\mathcal{AB}}[u]_{\mathcal{A}} &= [S]_{\mathcal{BC}}[T(u)]_{\mathcal{B}} \\ &= [S(T(u))]_{\mathcal{C}}\end{aligned}$$

Pela unicidade da matriz que satisfaz essa propriedade, concluímos que $[S \circ T]_{\mathcal{AC}} = [S]_{\mathcal{BC}}[T]_{\mathcal{AB}}$. □

Finalmente, a invertibilidade também é respeitada.

Proposição 4.5.9. Sejam V e W espaços vetoriais sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$ e $1 \leq \dim W = m < \infty$. Sejam $\mathcal{B} = (v_1, \ldots, v_n)$ uma base ordenada de V e $\mathcal{C} = (w_1, \ldots, w_m)$ uma base ordenada de W. Seja $T : V \to W$ uma transformação linear. Então:

(a) $[\mathrm{id}_V]_{\mathcal{BB}} = I_n$, onde I_n é a matriz identidade de ordem n.

(b) Se T é um isomorfismo se, e somente se $n = m$ e $[T]_{\mathcal{BC}}$ é invertível. Nesse caso, $[T^{-1}]_{\mathcal{CB}} = ([T]_{\mathcal{BC}})^{-1}$.

Demonstração. O primeiro item é imediato, uma vez que $[\mathrm{id}_V(v_j)]_{\mathcal{B}} = [v_j]_{\mathcal{B}}$ é o vetor coluna com 1 na j-ésima posição e 0 nas demais, ou seja, $[\delta_{ij}]_{i=1}^n$.

Para o segundo item, suponha que T é um isomorfismo. Então, necessariamente, $n = m$ e $[T]_{\mathcal{CB}}$ está bem definida. Do item anterior, segue que:

$$I_n = [\mathrm{id}_V]_{\mathcal{BB}} = [T^{-1} \circ T]_{\mathcal{BB}} = [T^{-1}]_{\mathcal{CB}}[T]_{\mathcal{BC}}.$$

E que:

$$I_n = [\mathrm{id}_W]_{\mathcal{CC}} = [T \circ T^{-1}]_{\mathcal{CC}} = [T]_{\mathcal{BC}}[T^{-1}]_{\mathcal{CB}}.$$

Logo, $[T]_{\mathcal{BC}}$ é invertível e $[T^{-1}]_{\mathcal{CB}} = ([T]_{\mathcal{BC}})^{-1}$.

Reciprocamente, suponha que $n = m$ e que $[T]_{\mathcal{BC}}$ é invertível. Seja A a matriz inversa de $[T]_{\mathcal{BC}}$. Seja S uma transformação linear tal que $[S]_{\mathcal{CB}} = A$. Então temos o seguinte:

$$[S \circ T]_{\mathcal{BB}} = [S]_{\mathcal{CB}}[T]_{\mathcal{BC}} = A[T]_{\mathcal{BC}} = I_n = [\mathrm{id}_V]_{\mathcal{BB}}.$$

Logo, $S \circ T = \mathrm{id}_V$. Similarmente, $T \circ S = \mathrm{id}_W$. Portanto, T é um isomorfismo e $S = T^{-1}$. □

Vamos enfatizar alguns casos particulares importantes.

Definição 4.5.10. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$. Sejam $\mathcal{B} = (v_1, \ldots, v_n)$ e $\mathcal{C} = (w_1, \ldots, w_n)$ bases ordenadas de V.

A matriz de mudança de base de $\mathcal{B}$ para $\mathcal{C}$ é a matriz P tal que:

$$P[x]_{\mathcal{B}} = [x]_{\mathcal{C}}$$

para todo vetor $x \in V$.

Ou seja, $P = [\mathrm{id}_V]_{\mathcal{BC}}$.

A matriz de mudança de base de $\mathcal{B}$ para $\mathcal{C}$ é denotada por $I_{\mathcal{BC}}$.

Dessa forma, $I_{\mathcal{BC}}$ é a matriz cuja j-ésima coluna é dada por $[v_j]_{\mathcal{C}}$, e $I_{\mathcal{BC}} = [\mathrm{id}_V]_{\mathcal{BC}}$. Pelo que vimos, é imediato que:

Corolário 4.5.11. Sejam V um espaço vetorial sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$. Sejam $\mathcal{B} = (v_1, \ldots, v_n)$ e $\mathcal{C} = (w_1, \ldots, w_n)$ bases ordenadas de V. Então, $I_{\mathcal{BC}}$ é invertível e $I_{\mathcal{CB}} = (I_{\mathcal{BC}})^{-1}$.

Demonstração. Temos que $I_{\mathcal{BC}} = [\mathrm{id}_V]_{\mathcal{BC}}$. Como id_V é um isomorfismo, pelo teorema anterior, $I_{\mathcal{BC}}$ é invertível e $[(\mathrm{id}_V)^{-1}]_{\mathcal{CB}} = (I_{\mathcal{BC}})^{-1} = [\mathrm{id}_V]_{\mathcal{CB}} = I_{\mathcal{CB}}$. □

Outro caso particular importante é o seguinte: se temos bases ordenadas $\mathcal{B}, \mathcal{B}'$ de V e $\mathcal{C}, \mathcal{C}'$ de W, como relacionar $[T]_{\mathcal{BC}}$ e $[T]_{\mathcal{B}'\mathcal{C}'}$?

Proposição 4.5.12. Sejam V e W espaços vetoriais sobre um corpo $\mathbb{K}$, com $1 \leq \dim V = n < \infty$ e $1 \leq \dim W = m < \infty$. Sejam $\mathcal{B}, \mathcal{B}'$ bases ordenadas de V e $\mathcal{C}, \mathcal{C}'$ bases ordenadas de W. Seja $T : V \to W$ uma transformação linear. Então:

$$[T]_{\mathcal{B}'\mathcal{C}'} = I_{\mathcal{CC}'}[T]_{\mathcal{BC}}I_{\mathcal{B}'\mathcal{B}}.$$

Demonstração. Pelos teoremas anteriores, temos:

$$\begin{aligned} I_{\mathcal{C}\mathcal{C}'}[T]_{\mathcal{B}\mathcal{C}} I_{\mathcal{B}'\mathcal{B}} &= [\mathrm{id}_W]_{\mathcal{C}\mathcal{C}'}[T]_{\mathcal{B}\mathcal{C}}[\mathrm{id}_V]_{\mathcal{B}'\mathcal{B}} \\ &= [\mathrm{id}_W \circ T]_{\mathcal{B}\mathcal{C}'}[\mathrm{id}_V]_{\mathcal{B}'\mathcal{B}} \\ &= [\mathrm{id}_W \circ T \circ \mathrm{id}_V]\mathcal{B}'\mathcal{C}' \\ &= [T]_{\mathcal{B}'\mathcal{C}'}. \end{aligned}$$

□

Capítulo 5

Determinantes

Neste capítulo introduziremos determinantes e provaremos suas propriedades básicas.

5.1 Introdução

Os determinantes são funções que associam a cada matriz quadrada um escalar, de forma que essa associação capture certas propriedades importantes da matriz.

Para defini-los, pensarmos em uma matriz quadrada $n \times n$ como uma coleção de n vetores em um espaço vetorial de dimensão n.

Em nossa discussão inicial, pensaremos no corpo dos números reais. No corpo dos números reais, o determinante mede o "volume"do paralelogramo (ou paralelepípedo) formado por esses vetores.

Listaremos algumas propriedades que esperamos que o determinante satisfaça. Abaixo, todos os vetores são elementos de $\mathbb{R}^n$.

a) $\det(e_1, e_2, \ldots, e_n) = 1$, em que e_i são os vetores da base canônica, pois este é o paralelogramo unitário.

b) Temos que $\det(v_1, v_2, \ldots, v_n) = 0$ caso tenhamos dois vetores repetidos, pois o paralelogramo colapsa em um espaço de dimensão menor, resultando em volume zero.

c) Ao multiplicar um vetor por algum escalar, o determinante é multiplicado por esse escalar. Ou seja, para todos os escalares $\alpha \in \mathbb{R}$ e qualquer coordenada i, temos que $\det(v_1, \ldots, \alpha v_i, \ldots, v_n) = \alpha \det(v_1, v_2, \ldots, v_n)$. Notemos que isso abre margem para a possibilidade de o determinante ser negativo, o que tem a ver com a orientação do paralelogramo.

d) Se $v_1, v_2, \ldots, v_n$ são vetores e v_i' é outro vetor, então $\det(v_1, \ldots, v_{i-1}, v_i + v_i', v_{i+1}, \ldots, v_n) = \det(v_1, v_2, \ldots, v_n) + \det(v_1, \ldots, v_{i-1}, v_i', v_{i+1}, \ldots, v_n)$. Tal propriedade é esperada, pois o paralelogramo determinado pelos vetores do lado esquerdo da igualdade pode ser visto como uma junção do paralelogramo determinado pelos vetores do primeiro determinante do lado direito com o paralelogramo determinado pelos vetores do segundo determinante do lado direito, o que faz as áreas serem somadas. Note que isso também abre margem para volumes negativos, o que será ignorado na discussão inicial, mas que tem a ver, novamente, com a orientação.

As propriedades c) e d) indicam que o determinante é uma função *linear em cada coordenada*.

Olhando para elas isoladamente, é possível definir o seguinte conceito. Abaixo, note que $(R^n)^n = R^n \times R^n \times \cdots \times R^n$ (n vezes).

Definição 5.1.1. Seja R um anel comutativo. Uma função $f : (R^n)^n \to R$ é dita *multilinear* se, para cada $i \in \{1, \ldots, n\}$, a função obtida ao fixar todos os argumentos de f exceto o i-ésimo é linear. Formalmente, se para todos $v_1, \ldots, v_n, v_j' \in R^n$ e todo $\alpha \in R$, temos:

- A função f abre para soma coordenada-a-coordenada, ou seja:

$$f(v_1, \ldots, v_{j-1}, v_j + v_j', v_{j+1}, \ldots, v_n) = f(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) + f(v_1, \ldots, v_{j-1}, v_j', v_{j+1}, \ldots, v_n).$$

- A função f respeita o produto por escalar em cada coordenada, ou seja:

$$f(v_1, \ldots, v_{j-1}, \alpha v_j, v_{j+1}, \ldots, v_n) = \alpha f(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n).$$

Agora olharemos para a propriedade b). Na presença da multilinearidade, ela é equivalente a outra propriedade no corpo dos números reais.

Proposição 5.1.2. Seja $f : (R^n)^n \to R$ uma função multilinear. Considere as seguintes afirmações:

a) (f é antissimétrica) Para quaisquer $v_1, \dots, v_n \in R^n$ e $i, j \in \{1, \dots, n\}$ com $i \neq j$, temos que:

$$f(v_1, \dots, v_n) = -f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n).$$

b) (f é alternada) Para quaisquer $v_1, \dots, v_n \in R^n$ e $i, j \in \{1, \dots, n\}$ com $i \neq j$, se $v_i = v_j$, então:

$$f(v_1, \dots, v_n) = 0.$$

Então a propriedade (b) implica a propriedade (a). Além disso, se, em R, $2(= 1 + 1)$ é tal que para todo $a \in R$, $2a = 0$ implica $a = 0$ então (a) implica (b).

Demonstração. Suponha que (b) seja verdadeira. Temos que:

$$\begin{aligned}
0 &= f(v_1, \dots, v_{i-1}, v_i + v_j, v_{i+1}, \dots, v_{j-1}, v_i + v_j, v_{j+1}, \dots, v_n) \\
&= f(v_1, \dots, v_n) + f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_n) + f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_n) \\
&\quad + f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n) \\
&= f(v_1, \dots, v_n) + f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n).
\end{aligned}$$

Agora suponha que (a) seja verdadeira, além da hipótese adicional sobre R, e sejam dados $v_1, \dots, v_n \in R^n$ tais que existem $i, j \in \{1, \dots, n\}$ com $i \neq j$ e $v_i = v_j$. Então, aplicando (a) e trocando v_i por v_j na expressão de (a), temos:

$$f(v_1, \dots, v_n) = -f(v_1, \dots, v_n).$$

Logo, $2f(v_1, \dots, v_n) = 0$, o que implica que $f(v_1, \dots, v_n) = 0$. □

Algumas propriedades de funções n-lineares são:

Proposição 5.1.3. Seja $f : (R^n)^n \to R$ uma função n-linear. Então, para quaisquer $v_1, \dots, v_n \in R^n$ e $i \in \{1, \dots, n\}$, se $v_i = 0$, então:

$$f(v_1, \dots, v_n) = 0.$$

Demonstração. Seja $v_1, \dots, v_n \in R^n$ tais que $v_i = 0$ para algum i. Então:

$$\begin{aligned}
f(v_1, \dots, v_n) &= f(v_1, \dots, v_{i-1}, 0 + 0, v_{i+1}, \dots, v_n) \\
&= f(v_1, \dots, v_{i-1}, 0, v_{i+1}, \dots, v_n) + f(v_1, \dots, v_{i-1}, 0, v_{i+1}, \dots, v_n) \\
&= 2f(v_1, \dots, v_{i-1}, 0, v_{i+1}, \dots, v_n).
\end{aligned}$$

Pela hipótese sobre R, temos que $f(v_1, \dots, v_n) = 0$. □

Proposição 5.1.4. Seja $f : (R^n)^n \to R$ uma função n-linear e alternada. Então, para quaisquer $v_1, \dots, v_n \in R^n$ e $i, j \in \{1, \dots, n\}$ com $i \neq j$, se $r \in R$, então:

$$f(v_1, \dots, v_{i-1}, v_i + rv_j, v_{i+1}, \dots, v_n) = f(v_1, \dots, v_n).$$

Demonstração. Seja $v_1, \dots, v_n \in R^n$ e $i, j \in \{1, \dots, n\}$ com $i \neq j$. Então:

$$\begin{aligned}
f(v_1, \dots, v_{i-1}, v_i + rv_j, v_{i+1}, \dots, v_n) &= f(v_1, \dots, v_n) + rf(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_n) \\
&= f(v_1, \dots, v_n) + r \cdot 0 \\
&= f(v_1, \dots, v_n).
\end{aligned}$$

□

Na definição de determinante, o que usaremos é a propriedade de f ser alternada. Estamos prontos para enunciar a definição de determinante.

Definição 5.1.5. Seja R um anel comutativo. O *determinante* $n \times n$ de R é a única função $\det : (R^n)^n \to R$ multilinear, alternada, e que vale 1 em $(e_1, e_2, \dots, e_n)$, em que e_i são os vetores da base canônica de R^n.

É claro que, por enquanto, não provamos a existência nem a unicidade de tal função. Isso será feito na seção a seguir.

5.2 Existência e unicidade de determinantes

No fim da seção anterior, chegamos à definição de uma função determinante, mas não provamos que tal função existe ou que é única.

Iniciaremos com a existência.

Lema 5.2.1. Seja R um anel comutativo e $r \in R$. Então existe uma função multilinear e antissimétrica $f : (R^n)^n \to R$ tal que $f(e_1, e_2, \ldots, e_n) = r$.

Lema 5.2.2. Seja R um anel comutativo e $g : (R^n)^n \to R$ uma função multilinear e alternada. Seja $i \in \{1, \ldots, n+1\}$. Então a função $f : (R^{n+1})^{n+1} \to R$ definida por

$$f(v_1, v_2, \ldots, v_{n+1}) = \sum_{j=1}^{n+1} (-1)^{j+i} v_{ij} g(v_1^{(i)}, \ldots, v_{j-1}^{(i)}, v_{j+1}^{(i)}, \ldots, v_{n+1}^{(i)}),$$

em que v_{ij} é a j-ésima coordenada do vetor v_i e $v_k^{(i)}$ é o vetor obtido de v_k ao remover sua i-ésima coordenada, é multilinear e alternada. Além disso, $f(e_1, e_2, \ldots, e_{n+1}) = g(e_1, e_2, \ldots, e_n)$.

Demonstração. Verificaremos a multilinearidade. Sejam $k \in \{1, \ldots, n+1\}$, $v_1, \ldots, v_{n+1}, v_k' \in R^{n+1}$ e $\alpha \in R$.

$$\begin{aligned}
f(v_1, \ldots, v_{k-1}, v_k + \alpha v_k', v_{k+1}, \ldots, v_{n+1}) &= (-1)^{k+i}(v_{ik} + \alpha v_{ik'})\, g(v_1^{(i)}, \ldots, v_{k-1}^{(i)}, v_{k+1}^{(i)}, \ldots, v_{n+1}^{(i)}) \\
&\quad + \sum_{\substack{j=1 \\ j \neq k}}^{n+1} (-1)^{j+i} v_{ij}\, g(v_1^{(i)}, \ldots, (v_k + \alpha v_k')^{(i)}, \ldots, v_{n+1}^{(i)}) \\
&= (-1)^{k+i}(v_{ik} + \alpha v_{ik'})\, g(v_1^{(i)}, \ldots, v_{k-1}^{(i)}, v_{k+1}^{(i)}, \ldots, v_{n+1}^{(i)}) \\
&\quad + \sum_{\substack{j=1 \\ j \neq k}}^{n+1} (-1)^{j+i} v_{ij} \big(g(v_1^{(i)}, \ldots, v_k^{(i)}, \ldots, v_{n+1}^{(i)}) \\
&\qquad + g(v_1^{(i)}, \ldots, \alpha v_k'^{(i)}, \ldots, v_{n+1}^{(i)}) \big) \\
&= f(v_1, \ldots, v_{n+1}) + \alpha f(v_1, \ldots, v_{k-1}, v_k', v_{k+1}, \ldots, v_{n+1}).
\end{aligned}$$

Agora vejamos que f é alternada.

Sejam $v_1, \ldots, v_{n+1} \in R^{n+1}$ tais que existem $k, l \in \{1, \ldots, n+1\}$ com $k \neq l$ e $v_k = v_l$. Veremos que $f(v_1, \ldots, v_{n+1}) = 0$. Perceba que na expressão que define a função, todas as parcelas para as quais $j \notin \{k, l\}$ são nulas. Assim:

$$\begin{aligned}
f(v_1, \ldots, v_{n+1}) &= (-1)^{k+i} v_{ik} g(v_1^{(i)}, \ldots, v_{k-1}^{(i)}, v_{k+1}^{(i)}, \ldots, v_{n+1}^{(i)}) \\
&\quad + (-1)^{l+i} v_{il} g(v_1^{(i)}, \ldots, v_{l-1}^{(i)}, v_{l+1}^{(i)}, \ldots, v_{n+1}^{(i)})
\end{aligned}$$

Perceba que $(v_1^{(i)}, \ldots, v_{k-1}^{(i)}, v_{k+1}^{(i)}, \ldots, v_{n+1}^{(i)})$ é a mesma sequência que $(v_1^{(i)}, \ldots, v_{l-1}^{(i)}, v_{l+1}^{(i)}, \ldots, v_{n+1}^{(i)})$, exceto pela aplicação de $|k - l - 1|$ transposições. Logo:

$$\begin{aligned}
f(v_1, \ldots, v_{n+1}) &= (-1)^{k+i} v_{ik} g(v_1^{(i)}, \ldots, v_{k-1}^{(i)}, v_{k+1}^{(i)}, \ldots, v_{n+1}^{(i)}) \\
&\quad + (-1)^{l+i} v_{il} g(v_1^{(i)}, \ldots, v_{l-1}^{(i)}, v_{l+1}^{(i)}, \ldots, v_{n+1}^{(i)}) \\
&= (-1)^{k+i} v_{ik} g(v_1^{(i)}, \ldots, v_{k-1}^{(i)}, v_{k+1}^{(i)}, \ldots, v_{n+1}^{(i)}) \\
&\quad + (-1)^{l+i} v_{ik} (-1)^{k-l+1} g(v_1^{(i)}, \ldots, v_{k-1}^{(i)}, v_{k+1}^{(i)}, \ldots, v_{n+1}^{(i)}) \\
&= 0.
\end{aligned}$$

Para a última afirmação, note que ao calcular $f(e_1, e_2, \ldots, e_{n+1})$, a única parcela que não é nula é aquela em que $j = i$, o que nos dá exatamente g aplicada na base canônica de R^n, que é r. □

Em particular, existem funções determinantes. Agora podemos provar a unicidade dessas funções.

Lema 5.2.3. Seja R um anel. Para todo $n \geq 1$ e todo $r \in R$, existe no máximo uma função multilinear e alternada $f : (R^n)^n \to R$ tal que $f(e_1, e_2, \ldots, e_n) = r$.

Demonstração. Uma transposição é uma função bijetora $\tau : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$ que troca dois elementos de posição e mantém os demais fixos.

Se $f : (R^n)^n \to R$ é multilinear e alternada, se $v_1, v_2, \dots, v_n \in R^n$ e τ é uma transposição, então $f(v_1, v_2, \dots, v_n) = -f(v_{\tau(1)}, v_{\tau(2)}, \dots, v_{\tau(n)})$.

Para cada função bijetora $\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$, existem $k \geq 0$ transposições $\tau_1, \tau_2, \dots, \tau_k$ tais que:

$$\tau_k \circ \tau_{k-1} \circ \cdots \circ \tau_1 = \sigma.$$

Tal fato pode ser provado por indução em n e é deixado ao leitor. Notemos que:

$$f(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = (-1)^k f(v_{\tau_1 \circ \tau_2 \circ \cdots \circ \tau_k(1)}, \dots, v_{\tau_1 \circ \tau_2 \circ \cdots \circ \tau_k(n)}) = (-1)^k f(v_1, v_2, \dots, v_n).$$

Para cada $\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$ bijetora, fixamos um número k_σ qualquer para o qual existem transposições $\tau_1, \tau_2, \dots, \tau_{k_\sigma}$ tais que $\tau_{k_\sigma} \circ \tau_{k_\sigma - 1} \circ \cdots \circ \tau_1 \circ \sigma$ é a identidade e definimos $\operatorname{sgn}(\sigma) = (-1)^{k_\sigma}$.

Suponha agora que f, g são funções como no enunciado. Temos que, dados $v_1, v_2, \dots, v_n \in R^n$, seja, para cada $j \in \{1, 2, \dots, n\}$, $v_j = \sum_{i=1}^n v_{ij} e_i$, em que $v_{ij} \in R$.

$$\begin{aligned} f(v_1, v_2, \dots, v_n) &= f\left(\sum_{i_1=1}^n v_{i_1 1} e_{i_1}, \sum_{i_2=1}^n v_{i_2 2} e_{i_2}, \dots, \sum_{i_n=1}^n v_{i_n n} e_{i_n}\right) \\ &= \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_n=1}^n v_{i_1 1} v_{i_2 2} \cdots v_{i_n n} f(e_{i_1}, e_{i_2}, \dots, e_{i_n}) \end{aligned}$$

As parcelas para as quais há repetição de índices i_j são nulas, pois f é alternada. Assim, as únicas parcelas são nulas são aquelas para as quais $(i_1, i_2, \dots, i_n)$ é uma sequência injetora em $\{1, 2, \dots, n\}$. Mas isso implica que tal sequência é uma bijeção σ de $\{1, 2, \dots, n\}$ em $\{1, 2, \dots, n\}$.

Seja S_n o conjunto das bijeções de $\{1, 2, \dots, n\}$ em $\{1, 2, \dots, n\}$. A expressão anterior é igual à

$$\begin{aligned} f(v_1, v_2, \dots, v_n) &= \sum_{\sigma \in S_n} v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n} f(e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n} (-1)^{k_\sigma} f(e_1, e_2, \dots, e_n) \\ &= r \sum_{\sigma \in S_n} (-1)^{k_\sigma} v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n}. \end{aligned}$$

De forma completamente análoga, temos que

$$g(v_1, v_2, \dots, v_n) = r \sum_{\sigma \in S_n} (-1)^{k_\sigma} v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n}.$$

Logo, $f(v_1, v_2, \dots, v_n) = g(v_1, v_2, \dots, v_n)$ para todos $v_1, v_2, \dots, v_n \in R^n$. □

Isso termina a prova da existência e unicidade do determinante.

Definição 5.2.4. Seja R um anel comutativo. O *determinante* $n \times n$ em R é a função $\det : (R^n)^n \to R$ multilinear, alternada, e que vale 1 em $(e_1, e_2, \dots, e_n)$, em que e_i são os vetores da base canônica de R^n.

Na demonstração anterior, fixamos um número arbitrário k_σ para cada bijeção σ. No entanto, é possível mostrar que o número de transposições necessárias para transformar σ na identidade tem sempre a mesma paridade, de modo que a expressão $(-1)^{k_\sigma}$, que nos dá um sinal, é bem definida. Uma das formas de provar isso é usando determinantes! Faremos isso como aplicação.

Exemplo 5.2.5. Para $n = 1$, a função determinante em um anel comutativo R é dada por $\det(v_1) = v_{11}$.

Exemplo 5.2.6. Para $n = 2$, a função determinante em um anel comutativo R é dada por $\det(v_1, v_2) = v_{11} v_{22} - v_{12} v_{21}$.

Exemplo 5.2.7. Para $n = 3$, a função determinante em um anel comutativo R é dada por $\det(v_1, v_2, v_3) = v_{11} v_{22} v_{33} + v_{12} v_{23} v_{31} + v_{13} v_{21} v_{32} - v_{13} v_{22} v_{31} - v_{11} v_{23} v_{32} - v_{12} v_{21} v_{33}$.

Exemplo 5.2.8 (Expansão de Laplace em uma linha)**.** Para $n \geq 1$, a função determinante $n+1 \times n+1$ em um anel comutativo R é dada pela expressão recursiva:

$$\det(v_1, v_2, \ldots, v_{n+1}) = \sum_{j=1}^{n+1} (-1)^{j+i} v_{ij} \det(v_1^{(i)}, \ldots, v_{j-1}^{(i)}, v_{j+1}^{(i)}, \ldots, v_{n+1}^{(i)}),$$

em que i está fixo.

Em linguagem matricial, sendo $A \in M_{n+1}(R)$, temos:

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

Em que i é uma linha fixa, e $A(i|j)$ é a matriz obtida de A ao remover a linha i e a coluna j.

Proposição 5.2.9. Seja R um anel comutativo e $n \geq 1$. Para toda bijeção $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, se k e s são inteiros positivos para os quais existem transposições $\tau_1, \tau_2, \ldots, \tau_k$ e $\rho_1, \rho_2, \ldots, \rho_s$ tais que $\tau_k \circ \tau_{k-1} = \sigma$ e $\rho_s \circ \rho_{s-1} \circ \cdots \circ \rho_1 = \sigma$, então k e s têm a mesma paridade.

Demonstração. Seja σ uma bijeção de $\{1, 2, \ldots, n\}$ em $\{1, 2, \ldots, n\}$. Temos que $\det(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}) = (-1)^k \det(e_1, e_2, \ldots, e_n) = (-1)^k$. De forma análoga, $\det(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}) = (-1)^s$. Logo, $(-1)^k = (-1)^s$, o que implica que k e s têm a mesma paridade. $\square$

Definição 5.2.10. Dado um anel comutativo R e $n \geq 1$, para cada bijeção $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, definimos o *sinal* de σ como sendo

$$\operatorname{sgn}(\sigma) = (-1)^{k_\sigma},$$

em que k_σ é um número de transposições necessárias para transformar σ na identidade.

Corolário 5.2.11. Seja R um anel comutativo e $n \geq 1$. Então para todos $v_1, v_2, \ldots, v_n \in R^n$, temos que

$$\det(v_1, \ldots, v_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n},$$

E, se $r \in R$ e f é uma função multilinear e alternada tal que $f(e_1, e_2, \ldots, e_n) = r$, então $f(v_1, v_2, \ldots, v_n) = r \det(v_1, v_2, \ldots, v_n)$ para todos $v_1, v_2, \ldots, v_n \in R^n$.

5.3 Propriedades dos determinantes

Teorema 5.3.1 (Binet)**.** Sejam $A, B \in M_n(R)$. Então $\det(AB) = \det(A)\det(B)$.

Demonstração. Fixe A. Defina $f : (R^n)^n \to R$ por $f(v_1, v_2, \ldots, v_n) = \det(Av_1, Av_2, \ldots, Av_n)$. Note que f é multilinear e alternada e que $f(e_1, e_2, \ldots, e_n) = \det(A)$. Logo, $f(v_1, v_2, \ldots, v_n) = \det(A)\det(v_1, v_2, \ldots, v_n)$.

Por outro lado, $g : (R^n)^n \to R$ definida por $g(v_1, v_2, \ldots, v_n) = \det(A)\det(v_1, v_2, \ldots, v_n)$ também é multilinear e alternada e vale $\det(A)$ em $(e_1, e_2, \ldots, e_n)$. Logo, $f(v_1, v_2, \ldots, v_n) = g(v_1, v_2, \ldots, v_n)$ para todos $v_1, v_2, \ldots, v_n \in R^n$. Em particular, para as colunas de B, temos:

$$\det(AB) = f(\operatorname{col}_1(B), \operatorname{col}_2(B), \ldots, \operatorname{col}_n(B)) = g(\operatorname{col}_1(B), \operatorname{col}_2(B), \ldots, \operatorname{col}_n(B)) = \det(A)\det(B).$$

$\square$

Proposição 5.3.2. Seja R um anel comutativo e $A \in M_n(R)$. Então $\det(A) = \det(A^t)$, onde t é a matriz transporta de A.

Demonstração. Seja $A = (a_{ij})_{i,j}$. Temos que $A^t = (b_{ij})_{i,j} = (a_{ji})_{i,j}$. Logo,

$$\begin{aligned}\det(A^t) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1)1} b_{\sigma(2)2} \cdots b_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.\end{aligned}$$

Em cada produtório, fazendo $j = \sigma(i)$, temos: que $i = \sigma^{-1}(j)$. Logo,

$$\prod_{i=1}^{n} a_{i\sigma(i)} = \prod_{j=1}^{n} a_{\sigma^{-1}(j)j}.$$

Assim:

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma^{-1}(j)j}.$$

No somatório, fazendo $\rho = \sigma^{-1}$, temos que $\sigma = \rho^{-1}$.

Assim:

$$\begin{aligned}\det(A^t) &= \sum_{\rho \in S_n} \operatorname{sgn}(\rho^{-1}) \prod_{j=1}^{n} a_{\rho(j)j} \\ &= \sum_{\rho \in S_n} \operatorname{sgn}(\rho) \prod_{j=1}^{n} a_{\rho(j)j} \\ &= \det(A).\end{aligned}$$

Para essa penúltima igualdade, devemos ver que $\operatorname{sgn}(\rho^{-1}) = \operatorname{sgn}(\rho)$. S $\tau_1, \tau_2, \ldots, \tau_k$ são transposições tais que $\tau_k \circ \cdots \circ \tau_1 = \rho$, então $\tau_1 \circ \cdots \circ \tau_k = \rho^{-1}$. Do fato que a inversa de uma transposição é ela mesma, temos que $\tau_1 \circ \cdots \circ \tau_k = \rho^{-1}$. Logo, um número de transposições necessárias para obter ρ^{-1} é o mesmo que para obter ρ. Portanto, $\operatorname{sgn}(\rho^{-1}) = \operatorname{sgn}(\rho)$. $\square$

Proposição 5.3.3 (Expansão de Laplace em uma coluna)**.** Para $n \geq 1$, $A \in M_{n+1}(R)$, e $j \in \{1, \ldots, n+1\}$, temos que

$$\det(A) = \sum_{j=1}^{n+1} (-1)^{j+i} a_{ji} \det(A(i|j)),$$

em que $A(i|j)$ é a matriz obtida de A ao remover a linha i e a coluna j.

Demonstração. Como visto, $\det(A) = \det(A^t)$. Seja $\bar{i} = j$. Escreva $A^t = (b_{ij})_{i,j}$ e $A = (a_{ij})_{i,j}$. Temos que $b_{ij} = a_{ji}$ para todos i, j. Usando a expansão de Laplace em na linha $\bar{i}$ de A^t, temos:

$$\det(A) = \det(A^t) = \sum_{k=1}^{n+1} (-1)^{\bar{i}+k} b_{\bar{i}k} \det((A^t)(\bar{i}|k)).$$

Note que $(A^t)(\bar{i}|k) = (A(k|\bar{i}))^t$. Logo, $\det((A^t)(\bar{i}|k)) = \det(A(k|\bar{i}))$.

Assim:

$$\det(A) = \sum_{k=1}^{n+1} (-1)^{\bar{i}+k} a_{k\bar{i}} \det(A(k|\bar{i})).$$

Lembrando que $j = \bar{i}$, e trocando a letra muda k por i na expressão acima, temos o resultado desejado. $\square$

Exemplo 5.3.4 (Regra de Sarrus)**.** Seja R um anel comutativo e $A \in M_3(R)$ dada por

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

Então:

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

Demonstração. Fazendo a expansão de Laplace na primeira coluna, temos que:

$$\det(A) = a_{11} \det(A(1|1)) - a_{21} \det(A(2|1)) + a_{31} \det(A(3|1)).$$

Ou seja:

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

□

Agora vejamos como calcular o determinante em blocos.

Proposição 5.3.5. Seja R um anel comutativo e $n, m \geq 1$. Sejam $A \in M_n(R)$, $B \in M_{n\times m}(R)$, $0 \in M_{m\times n}(R)$, e $D \in M_m(R)$.

$$\det\begin{pmatrix} A & B \\ 0 & D\end{pmatrix} = \det(A)\det(D).$$

Demonstração. Primeiro, vamos provar que, para todo $A \in M_n(R)$ e $B \in M_{n\times m}(R)$, temos que

$$\det\begin{pmatrix} A & B \\ 0 & I_m\end{pmatrix} = \det(A).$$

Para tanto, seja $f : M_{n+m}(R) \to R$ dada por $f(X) = \det\begin{pmatrix} X & B \\ 0 & I_m\end{pmatrix}$. É fácil ver que f é multilinear e alternada em relação às primeiras n colunas de X. Além disso, $f(I_n) = \det\begin{pmatrix} I_n & B \\ 0 & I_m\end{pmatrix} = 1$, uma vez que as últimas m linhas podem, via escalonamento, eliminar a matriz B sem alterar o determinante. Logo, $f(X) = \det(X)$ para todo $X \in M_n(R)$.

Agora, fixemos $A \in M_n(R)$ e definimos $g : M_m(R) \to R$ por $g(Y) = \det\begin{pmatrix} A & B \\ 0 & Y\end{pmatrix}$. É fácil ver que g é multilinear e alternada em relação às últimas m colunas de Y. Além disso, $g(I_m) = \det\begin{pmatrix} A & B \\ 0 & I_m\end{pmatrix}$. Logo, $g(Y) = \det(A)\det(Y)$ para todo $Y \in M_m(R)$. Em particular, para $Y = D$, temos o resultado desejado. □

5.4 A matriz adjunta e a fórmula de Cramer

Definição 5.4.1. Seja R um anel comutativo e $A \in M_n(R)$.

Se $i, j \in \{1, \ldots, n\}$, o *cofator* (i, j) de A é o elemento de R dado por

$$c_{ij} = (-1)^{i+j}\det(A(i|j)).$$

A transposta $(c_{ji})_{i,j}$ da matriz dos cofatores de A é chamada de *matriz adjunta* de A e é denotada por $\operatorname{adj}(A)$.

Proposição 5.4.2. Seja R um anel comutativo e $A \in M_n(R)$. Então $A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \det(A)I_n$.

Demonstração. Seja $B = \operatorname{adj}(A)$. Note que o elemento (i, j) de AB é dado por

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n} a_{ik}(-1)^{k+j}\det(A(k|j)).$$

Se $i = j$, pela expansão de Laplace na linha i de A, temos que $(AB)_{ii} = \det(A)$.

Se $i \neq j$, considere a matriz C obtida de A ao substituir a linha j pela linha i. Pela expansão de Laplace na linha j de C, temos que

$$0 = \det(C) = \sum_{k=1}^{n} a_{ik}(-1)^{k+j}\det(A(k|j)) = (AB)_{ij}.$$

Logo, $AB = \det(A)I_n$. De forma análoga, podemos mostrar que $BA = \det(A)I_n$. □

Corolário 5.4.3. Seja R um anel comutativo e $A \in M_n(R)$. Então A é invertível em $M_n(R)$ se, e somente se, $\det(A)$ é invertível em R, e, nesse caso, $A^{-1} = \det(A)^{-1}\operatorname{adj}(A)$.

Em particular, se R é um corpo, então A é invertível em $M_n(R)$ se, e somente se, $\det(A) \neq 0$.

Demonstração. Se A é invertível em $M_n(R)$, então existe $B \in M_n(R)$ tal que $AB = I_n$. Logo, $\det(A)\det(B) = \det(AB) = \det(I_n) = 1$, o que implica que $\det(A)$ é invertível em R.

A reciproca é imediata a partir da proposição anterior. □

Finalmente, provaremos a fórmula de Cramer para sistemas lineares.

Proposição 5.4.4. Seja R um anel comutativo, $A \in M_n(R)$ e $b \in R^n$. Se $\det(A)$ é invertível em R, então o sistema linear $Ax = b$ tem solução única $(\alpha_1, \dots, \alpha_n)$ dada por

$$\alpha_i = \det(A_i)\det(A)^{-1},$$

em que A_i é a matriz obtida de A ao substituir a coluna i por b.

Demonstração. Seja $v = (\alpha_1, \dots, \alpha_n)$ a solução do sistema. Seja $[v]$ a matriz coluna associada a v, ou seja, $[v] = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$. Para cada $i \in \{1, \dots, n\}$, seja A_i a matriz obtida de A ao substituir a coluna i por b.

Note que $A[v] = [b]$. Multiplicando ambos os lados por $\operatorname{adj}(A)$, temos que $\operatorname{adj}(A)Av = \operatorname{adj}(A)[b]$.

Assim,

$$\det(A)v = \operatorname{adj}(A)[b].$$

Logo, para cada $j \in \{1, \dots, n\}$, temos que

$$\det(A)\alpha_j = (\operatorname{adj}(A)[b])_j = \sum_{i=1}^{n} c_{ji}b_i = \sum_{i=1}^{n} (-1)^{i+j}\det(A(i|j)b_i.$$

Note que a soma acima é exatamente a expansão de Laplace na coluna j de A_j. Logo,

$$\det(A)\alpha_j = \det(A_j),$$

o que implica que

$$\alpha_j = \det(A_j)\det(A)^{-1}.$$

□

Proposição 5.4.5. Seja R um anel comutativo $A \in M_n(R)$. Se A é uma matriz triangular superior ou inferior, então

$$\det(A) = a_{11}a_{22}\cdots a_{nn}.$$

Demonstração. Visualmente, temos que:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Lembremos que $\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}$. Note que, se σ não é a identidade, então existe $k \in \{1, \dots, n-1\}$ tal que $\sigma(k) \neq k$. Sendo k o menor com essa propriedade, temos que $\sigma(k) > k$, o que implica que $a_{\sigma(k)k} = 0$. Logo, a única parcela não nula na soma acima é aquela em que σ é a identidade, o que nos dá o resultado desejado. □

Capítulo 6

Diagonalização

Neste capítulo, estudaremos a diagonalização de matrizes e transformações lineares.

No processo de diagonalização, dado um operador linear $T : V \to V$, onde V é um espaço vetorial sobre um corpo $\mathbb{K}$ de dimensão finita positiva, busca-se uma base ordenada $\mathcal{B}$ de V tal que a matriz associada $[T]_{\mathcal{B}}$ seja uma matriz diagonal.

6.1 Autovalores, autovetores e autoespaços

Note que, se o que foi descrito no parágrafo anterior for possível, cada elemento v da base $\mathcal{B}$ deve satisfazer $T(v) = \lambda v$ para algum escalar $\lambda \in \mathbb{K}$ e ser não nulo.

Definição 6.1.1. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e $T : V \to V$ um operador linear. Um vetor $v \in V \setminus \{0\}$ é chamado de *autovetor* de T se existir $\lambda \in \mathbb{K}$ tal que $T(v) = \lambda v$.

O escalar λ é chamado de *autovalor* de T associado ao autovetor v.

O *autoespaço* de T associado a um autovalor λ é o conjunto

$$\mathrm{Aut}_\lambda(T) = \{v \in V : T(v) = \lambda v\}.$$

Assim, $\mathrm{Aut}_\lambda(T)$ é o conjunto de todos os autovetores de T associados a λ, e o vetor nulo.

Lema 6.1.2. Na notação acima, $\mathrm{Aut}_\lambda(T)$ é um subespaço vetorial de V e $\mathrm{Aut}_\lambda(T) = \ker(\lambda \,\mathrm{id} - T)$.

Demonstração. Basta provar a igualdade de conjuntos. Para tanto, note que, para todo $v \in V$:

$$v \in \mathrm{Aut}_\lambda(T) \iff T(v) = \lambda v \iff \lambda v - T(v) = 0 \iff (\lambda \,\mathrm{id} - T)(v) = 0 \iff v \in \ker(\lambda \,\mathrm{id} - T).$$

□

Como encontrar autovalores e autovetores? Uma forma natural é buscar autovetores em coordenadas. Para tanto, temos o seguinte.

Proposição 6.1.3. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ de dimensão finita positiva, $T : V \to V$ um operador linear, e $\mathcal{B}$ uma base ordenada de V.

Então $\lambda \in \mathbb{K}$ é um autovalor de T se, e somente se, $\det(\lambda I_n - [T]_{\mathcal{B}}) = 0$.

Demonstração. Note que λ é um autovalor de T se, e somente se, existe $v \in V \setminus \{0\}$ tal que $T(v) = \lambda v$. Isso é equivalente a dizer que existe $v \in \ker(\lambda \,\mathrm{id} - T) \setminus \{0\}$, ou seja, tal que $(\lambda \,\mathrm{id} - T)(v) = 0$.

Em coordenadas, isso é equivalente a dizer que existem $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$, não todos nulos, tais que

$$(\lambda I_n - [T]_{\mathcal{B}}) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = 0.$$

Ou seja, que o sistema homogêneo associado à matriz $\lambda I_n - [T]_{\mathcal{B}}$ tem solução não trivial.

Pelo visto no capítulo anterior, isso é equivalente a dizer que $\det(\lambda I_n - [T]_{\mathcal{B}}) = 0$.

□

Índice Remissivo