Notas para um Curso de Álgebra Linear

Heitor R. de Assis

18 de janeiro de 2026

Sumário

Capítulo 1

Sistemas Lineares e Matrizes

1.1 Sistemas Lineares

Sistemas lineares são conjuntos de equações (lineares, portanto o nome) que buscamos resolver simultaneamente. Sua origem é antiga e suas aplicações são as mais variadas, estando presentes em diversos campos da ciência. É a partir do seu estudo que a álgebra linear se formou e, reciprocamente, os sistemas lineares formam uma ferramente extremamente útil para esta teoria.

Os coeficientes presentes nas equações são elementos de um **corpo**. Até o ponto presente, o mais comum é trabalharmos com equações contendo *coeficientes reais*. Se faz importante, porém, equações com coeficientes no corpo dos complexos, $\mathbb{C}$. Pode-se pensar que perguntas sobre os números complexos só tem interesse abstrato, mas a variedade de instâncias onde encontramos números complexos já deve ser suficiente para considerá-los.

Definição 1.1.1. Um *corpo* $\mathbb{F}$ é um conjunto em que duas operações estão definidas: uma chamada de multiplicação $\cdot : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ e outra chamada de adição $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$. Estas operações devem satisfazer:

1. $+$ é comutativa: $x + y = y + x$ para todos $x, y \in \mathbb{F}$.
2. $+$ é associativa: $x + (y + z) = (x + y) + z$ para todos $x, y, z \in \mathbb{F}$.
3. Existe um único elemento neutro $0 \in \mathbb{F}$, chamado de zero, tal que $x + 0 = 0$ para todo $x \in \mathbb{F}$.
4. Para cada $x \in \mathbb{F}$, existe um único elemento $(-x) \in \mathbb{F}$ tal que $x + (-x) = 0$.
5. $\cdot$ é comutativa: $x \cdot y = y \cdot x$ para todos $x, y \in \mathbb{F}$.
6. $\cdot$ é associativa: $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ para todos $x, y, z \in \mathbb{F}$.
7. Existe um único elemento diferente de zero e denotado por 1 que satisfaz $x \cdot 1 = x$ para todo $x \in \mathbb{F}$.
8. Para todo $x \in \mathbb{F}$ diferente de zero, existe um único elemento, denotado por x^{-1}, que satisfaz $x^{-1} \cdot x = 1$.
9. Vale a propriedade distributiva, ou seja, $x \cdot (y + z) = x \cdot y + x \cdot z$ para todo $x, y, z \in \mathbb{F}$.

Observação. Discorramos um pouco mais sobre esta definição, para se ter uma noção de sua significância. Raramente na matemática estuda-se um conjunto de elementos por si só, sendo muito mais comum o estudo de um conjunto de elementos juntamente com as operações que os relacionam. Como exemplo, considere os números reais. Quando falamos da reta real $\mathbb{R}$, não estamos apenas falando do conjunto de números que formam a reta real, $\mathbb{R} = \{1, 2, 3, \pi, -2, \sqrt{3}, \cdots\}$, mas também assumimos que $1+2 = 3$, $4\pi+2\pi = 6\pi$ e assim por diante. Da mesma forma, estamos convencionando que $1 \times x = x$, qualquer valor que $x \in \mathbb{R}$ possa ter, que $-1 \times 4 = -4$ e que daí $(-1 \times 4) + 4 = 0$. Infinitas outras operações entre elementos individuais de $\mathbb{R}$ estão de certa forma implícitas quando dizemos estudamos a reta real.

Acontece que é um objetivo do campo da matemática a abstração de conceitos conhecidos. Ao nos depararmos com um conjunto, munido de certas operações que julgamos interessantes, buscamos abstrair o máximo possível. Nos perguntamos, por exemplo: Qual é o conjunto minimal de propriedades que, por si só, são capazes de gerar todas outras? Quando um conjunto, qualquer que seja, satisfaz tais propriedades, quais afirmações podem ser feitas? Por exemplo, a afirmação de que qualquer número multiplicado por zero da zero não está entre as propriedades da Definição 1.1.1. Apesar disso, qualquer conjunto com estas propriedades satisfaz esta afirmação (consegue prová-la?). Esta é uma forma de diferenciarmos as proposições que são intrínsecas de uma certa estrutura daquelas que são específicas de cada modelo. A Definição 1.1.1, portanto, é o resultado da abstração das propriedades essenciais dos números reais.

Esta investigação é útil, inclusive, para testar se as mesmas operações podem ser assumidas sobre diferentes conjuntos. Como exemplo, considere o plano $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$. Temos uma óbvia e bem estabelecida adição neste conjunto, a saber

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) .$$

Podemos generalizar a operação de multiplicação que está presente em $\mathbb{R}$ para este novo conjunto? A primeira tentativa pode ser replicar o que foi feito para a adição, definindo

$$(x_1, y_1) \times (x_2, y_2) = (x_1 \times x_2, y_1 \times y_2) . \tag{1.1}$$

Existe alguma propriedade que valorizamos quando falamos de produtos, que $\mathbb{R}$ com a multiplicação usual possui, mas $\mathbb{R}^2$ com a multiplicação definida acima não satisfaz? A resposta é **sim**. Em $\mathbb{R}$, não existem dois números x e y *diferentes de zero* tais que $x \times y = 0$. Já em $\mathbb{R}^2$, é fácil ver que como definido acima, a multiplicação de duas duplas pode ser zero sem que qualquer uma delas seja zero. Por exemplo,

$$(x, 0) \times (0, y) = (0, 0) .$$

Acontece que esta é uma propriedade importante por alguns motivos diferentes e, portanto, a definição de produto como em (1.1) não nos é adequada. Podemos nos perguntar se existe alguma definição para o produto em $\mathbb{R}^2$ que funcione da mesma forma que a multiplicação em $\mathbb{R}$.

Os exemplos mais importantes de corpos são os já conhecidos $\mathbb{R}$ e $\mathbb{C}$. Estes não são os únicos, mas serão os únicos considerados aqui. Neste curso, portanto, $\mathbb{F}$ será sempre $\mathbb{R}$ ou $\mathbb{C}$ e sua utilização será principalmente a de um simplificador de notação. Na próxima definição, por exemplo, generalizamos o conceito de produto cartesiano de reais, considerando $\mathbb{R}$ e $\mathbb{C}$ ao mesmo tempo.

Definição 1.1.2. Seja $n \in \mathbb{N}$ um natural qualquer. O conjunto $\mathbb{F}^n$ é aquele formado por todas as n-uplas de elementos do corpo $\mathbb{F}$, ou seja,

$$\mathbb{F}^n = \big\{(\alpha_1, \cdots, \alpha_n) \mid \alpha_1, \cdots, \alpha_n \in \mathbb{F}\big\} .$$

Neste texto, as letras gregas α, β, γ, etc. serão utilizadas para representar elementos do corpo $\mathbb{F}$, seja ele $\mathbb{R}$ ou $\mathbb{C}$.

Definição 1.1.3. Um *sistema linear* é um conjunto de equações da forma

$$(S) \qquad \begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = \beta_1, \\ \vdots \\ \alpha_{m1}x_1 + \cdots + \alpha_{mn}x_n = \beta_m, \end{cases}$$

em que $\alpha_{ij} \in \mathbb{F}$ e $\beta_1, \ldots, \beta_m \in \mathbb{F}$ são constantes e $x_1, \ldots, x_n$ são as incógnitas, cujos valores queremos determinar.

Dizemos que o sistema linear (S) é **homogêneo** se $\beta_1 = \cdots = \beta_m = 0$. O conjunto dos elementos $(x_1, \ldots, x_n) \in \mathbb{F}^n$ que resolvem o sistema é chamado de **conjunto solução** do sistema, ou seja,

$$\left\{ (x_1, \ldots, x_n) \in \mathbb{F}^n \mid \sum_{k=1}^{n} \alpha_{jk}x_k = \beta_j \,,\, \forall\, j = 1, \cdots, m \right\} .$$

Quanto à possibilidade de solução, temos somente três casos:

- Um sistema (S) é **compatível determinado** se existe uma única solução, ou seja, se o conjunto solução possui uma única $n-$upla.
- Um sistema (S) é **compatível indeterminado** se existem infinitas soluções, ou seja, se o conjunto solução possui infinitas $n-$uplas.
- Um sistema (S) é **incompatível** se não possui solução, ou seja, se o conjunto solução é vazio.

Exemplo. Resolvemos o sistema

$$(S) \qquad \begin{cases} x + y = 10, \\ x - y = -2. \end{cases}$$

Utilizamos o método de eliminação de variáveis: Somando as duas equações, obtemos

$$2x = 8 \implies x = 4 \implies y = 6 \,.$$

O sistema (S) é portanto compatível determinado e seu conjunto solução é dado pela dupla $(4, 6)$, apenas.

Exemplo. Cada equação de um sistema linear de m variáveis é uma restrição que tem potencial, portanto, de condicionar uma variável as outras. Consideremos por exemplo o caso de um sistema com apenas uma equação.

$$(S) \qquad \{\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta \ .$$

Neste caso podemos isolar qualquer uma das variáveis, colocando-a em função das demais. O conjunto solução deste sistema pode ser escrito, por exemplo, uma vez que escrevemos

$$\begin{aligned} x_1 &= -\frac{\alpha_2}{\alpha_1}x_1 - \cdots - \frac{\alpha_n}{\alpha_1}x_n \\ &= -\sum_{k=2}^{n} \frac{\alpha_k}{\alpha_1}x_k \,. \end{aligned}$$

Logo, o conjunto de soluções de (S) é

$$\left\{ \left(-\sum_{k=2}^{n} \frac{\alpha_k}{\alpha_1}x_k, x_2, \cdots, x_n \right) \mid x_2, \cdots, x_n \in \mathbb{R} \right\} .$$

Como já dito, podemos isolar qualquer umas das variáveis, ou seja também podemos escrever

$$x_2 = -\frac{\alpha_1}{\alpha_2}x_1 - \sum_{k=3}^{n}\frac{\alpha_k}{\alpha_2}x_k\,,$$

de forma que o conjunto solução de (S) pode ser reescrito como

$$\left\{\left(x_1, -\frac{\alpha_1}{\alpha_2}x_1 - \sum_{k=3}^{n}\frac{\alpha_k}{\alpha_2}x_k, x_3, \cdots, x_n\right) \mid x_1, x_3, \cdots, x_n \in \mathbb{R}\right\}.$$

Ambas as formas representam *o mesmo conjunto* de elementos, sendo apenas parametrizações diferentes do mesmo conjunto.

Exemplo. Considere o sistema linear

$$(S) \quad \begin{cases} x + 2y - z = 2\,, \\ -3x + 3y - 4z = 0\,, \\ 4x - y + 3z = 3\,. \end{cases}$$

Ao tentarmos o processo de eliminação de variáveis para resolver (S), encontramos o seguinte

$$(l.1) - (l.2) - (l.3) \implies 0 = -1\,.$$

O que isso significa? Que se existir uma solução para este sistema, de modo que as três equações de (S) sejam verdadeiras, então obtemos que $0 = 1$, o que é um absurdo. Só nos resta afirmar que (S) não pode ser assumido verdadeiro, ou seja, não existe instância onde as três equações são simultaneamente satisfeitas.

Posto desta forma, talvez não fique claro o que há de errado com o sistema (S). Vejamos o seguinte: subtraindo a segunda equação da primeira, temos

$$4x - y + 3z = 2\,.$$

Assim, uma solução de (S) é automaticamente uma solução de

$$(S) \quad \begin{cases} x + 2y - z = 2\,, \\ 4x - y + 3z = 2\,, \\ 4x - y + 3z = 3\,. \end{cases}$$

Desta forma, fica claro que não existe tripla que satisfaça as três equações simultaneamente, pois as duas últimas formam resultados incompatíveis.

O que pretendemos fazer neste capítulo obter um método para a resolução de qualquer sistema linear que tenhamos em mãos. Com isso, queremos dizer: (i) determinar se o sistema possui ou não soluções; (ii) Obter fórmulas que identifiquem todas as soluções existentes. Para a primeira tarefa, a relação entre sistemas lineares e matrizes é de fundamental importância, como veremos a frente.

1.2 Resolução de Sistemas Lineares

O estudo dos sistemas lineares já está completamente classificado, de forma que para qualquer sistema linear (S) podemos determinar se ele é compatível ou incompatível e, no primeiro caso, obter todas as soluções possíveis. O método mais difundido é o **método de escalonamento**, que em linha gerais equivale (S) a um sistema linear (S') mais simples, de onde podemos extrair as soluções do primeiro. A noção de equivalência entre os dois sistemas lineares, definida a frente, permitirá afirmar que as suas soluções são as mesmas. Este processo é análogo, por exemplo, ao que fazemos para equações quadráticas, quando utilizamos o método de completamento de quadrados. Para obtermos as soluções de uma equação genérica de segundo grau,

$$ax^2 + bx + c = 0\,, \tag{1.2}$$

argumentamos que qualquer solução de (1.2) é também uma solução de

$$\left(x + \frac{b}{2a}\right)^2 = \left(\frac{b}{2a}\right)^2 - \frac{c}{a}\,. \tag{1.3}$$

A existência de solução (compatibilidade) da equação (1.3) é mais diretamente verificada do que o caso anterior, com a equação 1.2. Se o lado direito for positivo, (1.3) tem como solução/soluções

$$x = \pm\sqrt{\left(\frac{b}{2a}\right)^2 - \frac{c}{a}} - \frac{b}{2a}\,.$$

De modo geral, buscamos equiparar (1.2) a uma equação da forma

$$(x - \alpha)^2 = \beta\,, \tag{1.4}$$

que é evidente quanto a sua solubilidade e quanto a suas soluções. O escalonamento de um sistema linear (S) segue o mesmo princípio, sendo que devemos portanto achar quais são nossas *operações elementares*, ou seja que preservam as soluções, e quais são os sistemas mais simples possíveis. Estas serão as duas próximas definições.

Definição 1.2.1. Podemos definir três *operações elementares* sobre os sistemas lineares. Estas são:

1. Permutar duas equações do sistema.
2. Multiplicar uma das equações do sistema por um $\lambda \in \mathbb{F}$, $\lambda \neq 0$.
3. Substituir uma equação pela soma desta com outra equação.

Quando um sistema (S') é obtido através de (S) por operações elementares, dizemos que os dois sistemas são *equivalentes* e escrevemos $S \sim S'$.

Observação. Podemos substituir as últimas operações por uma alternativa[1]:

1' Substituir uma equação pela combinação linear desta com outra com outra equação.

A justificativa para considerarmos estas operações acima de quaisquer outras é que estas não alteram as soluções do sistema.

Proposição 1.2.2. *Se dois sistemas (S_1) e (S_2) são equivalentes, então eles possuem as mesmas soluções.*

[1] Ao utilizarmos esta definição, porém, um cuidado extra deve ser tomado de não multiplicarmos a equação substituída pelo zero ao somarmos com a múltipla de outra equação.

Demonstração. Precisamos apenas provar que, ao realizar qualquer umas das três operações expostas na Def. 1.2.1 não alteramos as soluções. É evidente que permutar duas linhas e multiplicar uma linha por um escalar (não nulo) não irá trocar as soluções. Para a soma de duas equações, basta notar que qualquer solução de um par de equações

$$(S) \qquad \begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = \beta_1, \\ \alpha_{21}x_1 + \cdots + \alpha_{2n}x_n = \beta_2, \end{cases}$$

vai satisfazer também a equação

$$(\alpha_{11} + \alpha_{12})x_1 + \cdots + (\alpha_{1n} + \alpha_{2n})x_n = \beta_1 + \beta_2\,.$$

Ou seja, satisfaz por exemplo

$$(S') \qquad \begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = \beta_1, \\ (\alpha_{11} + \alpha_{12})x_1 + \cdots + (\alpha_{1n} + \alpha_{2n})x_n = \beta_1 + \beta_2. \end{cases}$$

Reciprocamente, se uma n-upla satisfaz (S'), então da mesma forma esta vai satisfazer também (S). Isso mostra que os dois sistemas compartilham o mesmo conjunto de soluções. □

A seguir, temos resultados que justificam a nomenclatura 'sistemas equivalentes'.

Proposição 1.2.3. *A relação $\sim$ define uma relação de equivalência entre os sistemas, ou seja,*

1. *$S \sim S$.*
2. *Se $S_1 \sim S_2$, então $S_2 \sim S_1$.*
3. *Se $S_1 \sim S_2$ e $S_2 \sim S_3$, então $S_1 \sim S_3$.*

Agora passamos para a caracterização do que seriam nossos sistemas mais simples possíveis.

Definição 1.2.4. Um sistema linear de m equações e n incógnitas é *escalonado* se ele for da forma

$$\begin{cases} x_1 + & \cdots + & & + \cdots + & & + \cdots & + \alpha_{1n}x_n = \beta_1, \\ & & x_{r_1} & + \cdots + & & + \cdots & + \alpha_{2n}x_n = \beta_2, \\ & & & & & & \vdots \\ & & & & x_{r_k} & + \cdots & + \alpha_{kn}x_n = \beta_k, \\ & & & & & & \vdots \\ & & & & & & + 0 \cdot x_n = \beta_m, \end{cases}$$

com $1 \leq r_1 < r_2 < \cdots < r_k \leq n$.

Vejamos a mesma definição posta em termos diferentes:

Definição 1.2.5. Um sistema linear (S) é escalonado se:

1. a primeira variável presente em uma linha estiver a direita da primeira variável da linha superior;
2. a primeira variável presente na equação de uma linha não estiver presente nas linhas acima;
3. a primeira variável de cada equação é seguida do coeficiente 1;

4. qualquer linha nula estiver abaixo de qualquer linha não nula.

Tendo em mãos os sistemas mais simples possíveis, devemos ser capazes de resolvê-los sem mais problemas, tal como argumentamos para a equação quadrática na forma (1.4).

Proposição 1.2.6. *Se (S) é um sistema escalonado, então:*

1. *(S) é incompatível se possui uma linha da forma*

$$0 = \beta\,, \quad \beta \neq 0\,.$$

Caso, contrário, (S) é compatível.

2. *se (S) é compatível e possui mais variáveis do que equações não nulas, então (S) é compatível indeterminado. Caso contrário, (S) é compatível determinado.*

Por fim, a prova de que qualquer sistema linear pode ser resolvido se conclui ao mostrarmos que qualquer sistema (S) é equivalente a um sistema escalonado.

Teorema 1.2.7. *Todo sistema é equivalente a um (único) sistema escalonado. A resolução do sistema é obtida então através da solubilidade do sistema escalonado equivalente.*

Demonstração. A prova é feita expondo o algoritmo para transformar um sistema qualquer, através de operações elementares, em um sistema escalonado. A ideia é sempre que possível deixar apenas uma linha para cada variável, sendo necessárias as operações elementares para zerarem a mesma variável nas demais equações.

Primeiramente, trocamos linhas se necessário para que a primeira equação tenha x_1. Uma vez feito isso (com *qualquer* equação com x_1 ocupando o lugar de primeira equação), zeramos os coeficientes de x_1 nas demais linhas. Para um par de linhas qualquer,

$$\left\{\begin{array}{llll} \alpha_{11}x_1 + & \cdots + & \alpha_{1n}x_n = \beta_1, \\ \alpha_{k1}x_1 + & \cdots + & \alpha_{kn}x_n = \beta_k, \end{array}\right.$$

fazemos $\alpha_{k1}(l.1) - \alpha_{11}(l.k) \to (l.k)$ e obtemos

$$\left\{\begin{array}{lll} \alpha_{11}x_1 + & \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n = & \beta_1, \\ & \gamma_{k2}x_2 + \cdots + \gamma_{kn}x_n = & \alpha_{k1}\beta_1 - \alpha_{11}\beta_k, \end{array}\right.$$

onde $\gamma_{kj} = \alpha_{k1}\alpha_{1j} - \alpha_{11}\alpha_{kj}$. Ao final deste processo, o sistema estará da seguinte forma

$$\left\{\begin{array}{llll} \alpha_{11}x_1 + & \alpha_{12}x_2 + \cdots & +\alpha_{1n}x_n = \beta_1, \\ & \gamma_{22}x_2 + \cdots & +\gamma_{2n}x_n = \delta_2 \\ & \gamma_{32}x_2 + \cdots & +\gamma_{3n}x_n = \delta_2 \\ & & \vdots \\ & \gamma_{n2}x_2 + \cdots & +\gamma_{nn}x_n = \delta_2 \end{array}\right.$$

Em seguida, comparamos os coeficientes da próxima coluna (referente a variável x_2). Procuramos entre as linhas, excetuando a primeira, qual possui a variável x_2 e teremos dois resultados possíveis:

Caso 1: Existe um $\gamma_{k2} \neq 0$. Neste caso, colocamos a k-ésima equação na posição da segunda equação através de uma troca de linhas. Deste modo, podemos supor aqui que $\gamma_{22} \neq 0$. Zeramos o coeficiente da primeira equação através da operação elementar $\gamma_{22}(l.1) - \alpha_{12}(l.2) \to (l.1)$, enquanto que para as demais linhas zeramos fazendo $\gamma_{22}(l.k) - \gamma_{k2}(l.2) \to (l.k)$, da mesma forma como feito antes.

Caso 2: Não existe γ_{k2} diferente de zero. Neste caso, α_{21} deve ser não nulo e podemos simplesmente passar para a próxima variável.

É fácil ver que é possível prosseguir desta forma até obtermos um sistema que satisfaça as condições 1, 2 e 4 de um sistema escalonado. Para concluir, basta multiplicar cada linha pelo inverso do coeficiente que segue cada primeira variável ($\alpha_{11}^{-1}, \gamma_{k2}^{-1}$, etc.). $\square$

Observação. Não provamos aqui a parte da unicidade, ou seja, que se (S) é ao mesmo tempo equivalente a dois sistemas escalonados (S_1) e (S_2), então estes na verdade são exatamente o mesmo sistema. O leitor pode ver a prova deste fato em livros de álgebra linear mais avançados.

A obtenção da unicidade é mais além a consequência de todas as quatro condições assumidas na Def. 1.2.5. Alguns materiais, por exemplo, definem sistemas escalonados aqueles que satisfazem 1, 2 e 4, apenas. Desta forma, porém, dois sistemas como

$$(S_1) \quad \begin{cases} 2x+ & & -4z = 2\,, \\ & -\quad 3y & +3z = 4\,. \end{cases} \qquad (S_2) \quad \begin{cases} x+ & & -2z = 1\,, \\ & -\quad y & +z = 4/3\,. \end{cases}$$

seriam ambos escalonados e equivalentes.

1.3 Matrizes sobre um corpo

Definição 1.3.1. Uma matriz $m \times n$ com coeficientes em $\mathbb{F}$ é uma função

$$\begin{array}{rccc} A\colon & \{1,\dots,m\} \times \{1,\dots,n\} & \to & \mathbb{F} \\ & (i,j) & \to & A(i,j)\,. \end{array}$$

Denotaremos os elementos da imagem dessa função pela letra minúscula a_{ij}, ou seja, $A(i,j) = a_{ij}$. Com isso, por vezes escrevemos simplesmente $A = (a_{ij})$. Por fim, dada uma matriz cujos termos são a_{ij} podemos representá-la da seguinte forma:

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

Dizemos que a matriz tem m linhas e n colunas e que a_{ij} é o elemento da i-ésima linha e $j-$ésima coluna de A. O conjunto das matrizes $m \times n$ é denotado por $M_{m\times n}(\mathbb{F})$ ou $M(m \times n, \mathbb{F})$. Se $m = n$, denotamos $M_{m\times n}(\mathbb{F})$ apenas por $M_n(\mathbb{F})$ ou $M(n, \mathbb{F})$.

Definição 1.3.2. Dados quaisquer $m, n \in \mathbb{N}$, podemos definir duas operações sobre $M_{m\times n}(\mathbb{F})$:

1. Adição $(+)$: $M_{m\times n}(\mathbb{F}) \times M_{m\times n}(\mathbb{F}) \to M_{m\times n}(\mathbb{F})$ é dada por $(a+b)_{ij} = a_{ij} + b_{ij}$.
2. Multiplicação por um escalar $(\cdot)$: $\mathbb{F} \times M_{m\times n}(\mathbb{F}) \to M_{m\times n}(\mathbb{F})$ é dada por $(\lambda \cdot a)_{ij} = \lambda a_{ij}$.

Estas operações já cobrem o básico para que possamos falar de resolução de sistemas lineares utilizando matrizes. Este conjunto, no entanto, possui outras possibilidades de operações, que adicionam estruturas interessantes.

Definição 1.3.3. Podemos definir também novas operações sobre o conjunto de **todas** as matrizes, independente do tamanho:

1. Multiplicação de matrizes $(\cdot)$: $M_{m\times n}(\mathbb{F})\times M_{n\times p}(\mathbb{F}) \to M_{m\times p}(\mathbb{F})$ é dada por $(a\times b)_{ij} = (c_{ij})$, em que $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$.

2. Transposta de matrizes $t : M_{m\times n}(\mathbb{F}) \to M_{n\times m}(\mathbb{F})$ é dada por $(a^t)_{ij} = (a_{ji})$.

A próxima operação é característica de matrizes sobre o corpo dos complexos.

1. Adjunta de matrizes $* : M_{m\times n}(\mathbb{C}) \to M_{n\times m}(\mathbb{C})$ é dada por $(a^*)_{ij} = (\overline{a_{ji}})$.

Note que todas as operações até agora definidas tem como subconjuntos invariantes as coleções de matrizes quadradas de um tamanho $n \in \mathbb{N}$ fixado. Estas matrizes, portanto, formam um subconjunto dotado de rica estrutura.

Observação. Se $a_1, \dots, a_n \in \mathbb{F}$, definimos a somatória por $\sum_{i=1}^{n} a_i := a_1 + \cdots + a_n$. Algumas propriedades da somatória:

1. $\alpha(\sum_{i=1}^{n} a_i) = \sum_{i=1}^{n} \alpha a_i$, para todo $\alpha \in \mathbb{F}$.
2. $\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i = \sum_{i=1}^{n}(a_i + b_i)$.
3. $\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} = \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij}$.
4. $(\sum_{i=1}^{n} a_i)(\sum_{j=1}^{m} b_j) = \sum_{j=1}^{m} \sum_{i=1}^{n} a_i b_j$.

Definição 1.3.4. A matriz identidade de $M_n(\mathbb{F})$ é a matriz $n \times n$ dada por $\mathrm{I}_n = (\delta_{ij})$, em que o delta de Kronecker δ_{ij} é o símbolo que significa

$$\delta_{ij} = \begin{cases} 1, & \text{se } i = j, \\ 0, & \text{se } i \neq j\,. \end{cases}$$

Esta matriz portanto é escrita como

$$\mathrm{I}_n = \underbrace{\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}}_{n \text{ colunas}} \left.\vphantom{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}\right\} n \text{ linhas}$$

A nomenclatura vem do fato de que I_n se comporta como a identidade de $\mathbb{R}$, ou seja o número 1. De fato, para qualquer $A \in M(n \times m, \mathbb{F})$, temos

$$B = \mathrm{I}_n A \implies B \in M(n \times m, \mathbb{F})\,, \quad b_{ij} = \sum_{k=1}^{n} \delta_{ik} a_{kj} = a_{ij}\,,$$

ou seja

$$\mathrm{I}_n A = A\,.$$

Não podemos, porém, fazer o produto $A\mathrm{I}_n$ se $m \neq n$. O caso em que $m \neq n$ é portanto especial, como veremos na Seção 1.4.

Proposição 1.3.5. *As operações de matrizes satisfazem as seguintes propriedades para $A \in M_{m\times n}(\mathbb{F})$, $A' \in M_{m\times n}(\mathbb{F})$, $B \in M_{n\times p}(\mathbb{F})$, $B' \in M_{n\times p}(\mathbb{F})$ e $C \in M_{p\times \ell}(\mathbb{F})$ (Para as propriedades envolvendo a operação adjunta, considere $\mathbb{F} = \mathbb{C}$).*

1. *$A(BC) = (AB)C$.*

2. $(A + A')B = AB + A'B$ *e* $A(B + B') = AB + AB'$.

3. $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

4. $(A + B)^t = A^t + B^t$ *e* $(A + B)^* = A^* + B^*$.

5. $(AB)^t = B^t A^t$ *e* $(AB)^* = B^* A^*$.

6. $(\alpha A)^t = \alpha A^t$ *e* $(\alpha A)^* = \overline{\alpha} A^*$.

7. $(A^t)^t = A$ *e* $((A^*)^*) = A$.

Podemos definir três operações elementares sobre as matrizes $M_{m\times n}(\mathbb{F})$, de forma a espelhar as operações sobre sistemas lineares que nos levam aos sistemas escalonados.

Definição 1.3.6. Chamamos de operações elementares sobre $M_{m\times n}(\mathbb{F})$ as seguintes operações:

1. Permutar duas linhas da matriz.
2. Multiplicar uma das linhas da matriz por um $\lambda \in \mathbb{F}$, $\lambda \neq 0$.
3. Somar uma das linhas da matriz por um múltiplo de outra linha.

Se uma matriz B é obtida de uma matriz A através de um conjunto de operações elementares dizemos que B é equivalente a A e denotamos $B \sim A$.

O mesmo resultado obtido para sistemas lineares, que justifica a ideia de igualdade entre sistemas relacionados por operações equivalentes, pode ser citado no caso matricial também.

Proposição 1.3.7. *A relação* $\sim$ *define uma relação de equivalência em* $M_{m\times n}(\mathbb{F})$*, ou seja:*

1. $A \sim A$.

2. Se $A \sim B$*, então* $B \sim A$.

3. Se $A \sim B$ *e* $B \sim C$*, então* $A \sim C$.

Demonstração. Veja o Exercício **??**. $\square$

1.4 Matrizes quadradas

Dentre todas as matrizes possíveis, existem subconjuntos que são mais comportados - ou mais estruturados - que os demais. Estes são os conjuntos das matrizes quadradas, ou seja com o mesmo número de colunas e linhas. A unicidade destes subconjuntos está no fato de que podemos definir um **elemento neutro** para a multiplicação sem ambiguidade. De fato, para qualquer matriz $A \in M(n \times m, \mathbb{F})$, podemos achar matrizes I e I' tais que

$$IA = A = AI' .$$

Porém, pela definição da multiplicação de matrizes, I e I' devem ser diferentes (seus tamanhos não coincidem). Por exemplo, para

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

temos

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

Para $A \in M(n \times m, \mathbb{F})$, portanto, o elemento neutro à direita é I_m, enquanto que o elemento neutro à esquerda é I_n. Em $M(n, \mathbb{F})$, portanto, ambos I e I' são o mesmo, o que caracteriz I_n como o elemento neutro da multiplicação em $M(n, \mathbb{F})$.

Proposição 1.4.1. *A matriz identidade* $\mathrm{I}_n \in M_n(\mathbb{F})$ *é a única matriz tal que, para todo* $A \in M_n(\mathbb{F})$*,*

$$A\mathrm{I}_n = \mathrm{I}_n A = A.$$

Demonstração. Suponha que $B \in M_n(\mathbb{F})$ é uma matriz tal que

$$AB = BA = A\,, \quad \forall\, A \in M_n(\mathbb{F})\,.$$

Em particular, segue daí que

$$\mathrm{I}_n = \mathrm{I}_n B = B\,.$$

Logo, toda matriz B com esta característica deve ser igual à I_n, o que prova que a identidade é única. □

A partir do momento em que temos uma multiplicação e um elemento neutro, podemos nos perguntar quais elementos podemos inverter, ou seja, para quais matrizes (quadradas) A, existe uma matriz B tal que $AB = BA = \mathrm{I}_n$.

Definição 1.4.2. Uma matriz $A \in M_n(\mathbb{F})$ é inversível se existe $B \in M_n(\mathbb{F})$ tal que

$$AB = BA = \mathrm{I}_n.$$

Neste caso B é denotado por A^{-1} e é chamado de inversa de A.

De uma maneira análoga à Proposição 1.4.1, é possível provar que a inversa de qualquer matriz é única.

Observação. Há um nome especial para um conjunto munido das três operações vistas até aqui - soma, multiplicação por escalar e produto - para matrizes quadradas. Estes conjuntos são **algebras**. Vimos acima que qualquer outro $M(n \times m, \mathbb{F})$, com $m \neq n$, não é uma álgebra.

Estudemos agora a definição do determinante de uma matriz. Esta quantidade é definida apenas para matrizes quadradas e a sua definição deriva do seguinte desenvolvimento: Considere um sistema genérico com duas variáveis e duas equações,

$$(S_1) \qquad \begin{cases}\alpha_{11}x_1 + \alpha_{12}x_2 = \beta_1,\\ \alpha_{21}x_1 + \alpha_{22}x_2 = \beta_2.\end{cases}$$

Podemos resolve-lo de maneira simples, substituindo uma das equações na restante. Por exemplo, a segunda equação nos diz que[2]

$$x_1 = \frac{1}{\alpha_{21}}(\beta_2 - \alpha_{22}x_2)\,.$$

[2]Note que se $\alpha_{21} = \alpha_{22} = 0$, então ou o sistema só possui uma equação (se $\beta = 0$) ou o sistema é incompatível ($\beta \neq 0$). Estamos assumindo portanto, que α_{21} é diferente de zero.

Substituindo esta igualdade na primeira equação, temos

$$\frac{\alpha_{11}}{\alpha_{21}}(\beta_2 - \alpha_{22}x_2) + \alpha_{12}x_2 = \beta_1\,,$$

ou seja,

$$(\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})x_2 = \alpha_{21}\beta_1 - \alpha_{11}\beta_1\,. \tag{1.5}$$

Portanto, se $\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} \neq 0$, obtemos a solução do sistema,

$$x_2 = \frac{\alpha_{21}\beta_1 - \alpha_{11}\beta_1}{\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}}\,, \quad x_1 = \frac{\beta_1 - \alpha_{12}x_2}{\alpha_{11}}\,.$$

Por outro lado, se $\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} = 0$, então ou o sistema (S_1) é incompatível (se o lado direito de (1.5) for diferente de zero) ou ele possui infinitas soluções (se o lado direito de (1.5) for igual a zero). Esta combinação, por ter este caráter definidor da solubilidade de (S_1), merece especial atenção e é ela que recebe o nome de determinante da matriz que define o sistema (a relação entre sistemas lineares e matrizes será vista na próxima seção). Podemos fazer o mesmo raciocínio para sistemas 3×3 e a quantidade que obtemos é o que definimos ser o determinante de uma matriz 3×3. O cálculo é um tanto trabalhoso, por isso não faremos ele aqui. Apenas apresentaremos a definição dos determinantes de matrizes quadradas de qualquer tamanho.

Definição 1.4.3. Seja $A \in M_n(\mathbb{F})$ uma matriz quadrada de tamanho n. Então definimos o **determinante** de A por

$$\det A = \sum_{i=1}^{n} (-1)^{i+1} a_{1i} \det A(1|i)\,,$$

onde $A(j|i)$ é a matriz de tamanho $n-1$ obtida de A retirando a $j-$ésima coluna e a $i-$ésima linha, ou seja

$$(A(j|i))_{lk} = \begin{cases} a_{lk}\,, & \text{se } l < j\,, k < i\,, \\ a_{(l+1)k}\,, & \text{se } l \geq j\,, k < i\,, \\ a_{l(k+1)}\,, & \text{se } l < j\,, k \geq i\,, \\ a_{(l+1)(k+1)}\,, & \text{se } l \geq j\,, k \geq i\,. \end{cases}$$

Observação. Para os tamanhos dois e três, podemos expandir a definição acima explicitamente. O primeiro caso retorna a quantidade que já encontramos,

$$A = (a_{ij}) \in M_2(\mathbb{F}) \Rightarrow \det A = a_{11}a_{22} - a_{21}a_{12}\,,$$

enquanto que para o caso de tamanho três, podemos calcular com a Regra de Sarrus,

$$\left|\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right.$$

$$\det A = \underbrace{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}}_{\text{somas das diagonais principais}} - \underbrace{\left(a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33}\right)}_{\text{somas das diagonais secundárias}}.$$

Se estudamos matrizes por conta própria, sem a motivação dos sistemas lineares, ainda assim o determinante é uma quantidade importante, pois nos dá a verificação de invertibilidade.

Teorema 1.4.4. *Seja $A \in M_n(\mathbb{F})$ uma matriz quadrada de tamanho qualquer. Então, A é inversível se, e somente se,* $\det A \neq 0$.

Mais do que isso, podemos escrever quem será a matriz inversa, em termos do determinante. Para isso é necessário também a definição da matriz adjunta de uma matriz quadrada.

Definição 1.4.5. Dada uma matriz $A \in M_n(\mathbb{F})$, definimos a matriz adjunta de A por

$$\operatorname{Adj}(A) = \Delta^T \,,$$

onde $\Delta = (\Delta_{ij})$ é a matriz dos cofatores de A, que foram apresentados na Def. 1.4.3,

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

Teorema 1.4.6. *Seja $A \in M_n(\mathbb{F})$ uma matriz quadrada de tamanho n qualquer. Temos A é inversível se, e somente se,* $\det A \neq 0$. *No caso afirmativo, temos*

$$A^{-1} = \frac{1}{\det A} \operatorname{Adj}(A)\,.$$

1.5 Matrizes e Sistemas Lineares

Agora que já vimos as definições e propriedades importantes do conjunto das matrizes, mostremos como elas se relacionam com a resolução de sistemas lineares.

Considere o sistema linear

$$(S) \qquad \begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = \beta_1, \\ \vdots \\ \alpha_{m1}x_1 + \cdots + \alpha_{mn}x_n = \beta_m, \end{cases} .$$

O que podemos perceber é que cada uma das m equações de (S) pode ser vista como uma multiplicação de uma linha matricial com uma coluna matricial. O lado esquerdo da primeira, por exemplo, é simplesmente $\sum_j \alpha_{1j}x_j$ e ela pode ser vista como

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \beta_1\,.$$

Como isso vale para cada uma das equações de (S) e a matriz coluna é sempre a mesma, com elementos sendo as variáveis x_i, podemos reescrever todas as equações de uma forma concisa através de uma multiplicação de matrizes

$$\underbrace{\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}}_{B}\,.$$

Podemos resolver um sistema linear, portanto, analisando as matrizes que o compõe, em particular a matrizes de coeficientes $A = (\alpha_{ij})$ e $B = (\beta_i)$. A matriz de variáveis, X, é secundária e tem um papel tal qual o argumento x de uma função $f : D \to CD$, de representar qual é a ação de f sobre o seu domínio D. De fato, note o seguinte: Se a matriz A de coeficientes é inversível, então podemos tomar $X = A^{-1}B$, de modo que

$$AX = A(A^{-1}B) = (AA^{-1})B = IB = B\,.$$

Desta forma, solucionamos (S) apenas olhando para a matriz A e nem mesmo B foi necessária. Chegamos portanto a uma primeira forma, por vezes falha pois nem toda matriz pe inversível, de se revolver um sistema (S) qualquer.

Proposição 1.5.1. *Considere um sistema (S) colocado em forma matricial $AX = B$, com A uma matriz quadrada de tamanho n. Se A é inversível, então (S) possui solução $X = A^{-1}B$, ou seja,*

$$x_i = \sum_{j=1}^{n} \gamma_{ij}\beta_j\,, \quad i = 1, \cdots, n\,,$$

onde $A^{-1} = (\gamma_{ij})$.

Note porém que a questão da solubilidade de um sistema (S) geral não depende de modo algum da inversa de uma matriz, inclusive pois na maioria dos sistemas, a matriz A de coeficientes não será quadrada e portanto não temos o que é a noção de uma inversa. Os sistemas quadrados são apenas um caso particular onde as propriedades de matrizes no ajudam a simplificar a resolução. O determinante de uma matriz é um quantificador importante para isso, pois o Teorema 1.4.6 nos diz que A possui inversa se e somente se seu determinante não é zero. O resultado que temos nos diz mais ainda.

Teorema 1.5.2. *Considere um sistema qualquer*

$$(S) \qquad \begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = \beta_1, \\ \vdots \\ \alpha_{m1}x_1 + \cdots + \alpha_{mn}x_n = \beta_m, \end{cases}$$

e seja $A = (\alpha_{ij})$ a matriz dos coeficientes de (S). Então temos o seguinte:

1. *Se $\det A \neq 0$ então (S) possui uma única solução. Se $B = 0$, esta é a solução trivial $x_1 = x_2 = \cdots = x_n = 0$.*
2. *Se $\det A = 0$, então (S) é incompatível ou possui infinitas soluções. Se $B = 0$, então (S) necessariamente possui infinitas soluções.*

Podemos expressar este resultado no sentido contrário, olhando para uma matriz quadrada qualquer e tirando conclusões dos sistemas que ela pode formar.

Teorema 1.5.3. *Seja $A = (\alpha_{ij})$ uma matriz quadrada de tamanho n qualquer. Então temos o seguinte:*

1. *Se $\det A \neq 0$ então qualquer sistema (S) da forma $AX = B$ possui uma única solução. Para $B = 0$, a solução é trivial, $x_1 = x_2 = \cdots = x_n = 0$.*
2. *Se $\det A = 0$, então existe uma $B \in M_{n\times 1}(\mathbb{F})$ tal que o sistema (S) da forma $AX = B$ é incompatível. Para $B = 0$, (S) possui infinitas soluções.*

Observação. Os resultados que obtivemos até aqui para sistemas quadrados através da analise dos determinantes são apenas de natureza classificatória, ou seja, não nos dão explicitamente as soluções procuradas a não ser em casos particulares como sistemas homogêneos. Como os determinantes também nos permitem calcular as inversas das matrizes (com a matriz adjunta entrando também na fórmula), eles podem também ser usado para a resolução completa dos sistemas lineares no caso em que existe uma inversa. Este método de cálculo da inversa, porém, costuma ser mais trabalhoso que o método de escalonamento e portanto não costuma valer a pena utilizá-lo.

Mesmo o método de escalonamento para a resolução de sistemas lineares é facilitado com a utilização da representação matricial. Para utilizarmos este, porém, precisamos primeiro definir o que é uma matriz escalonada, que espelha a definição para sistemas lineares. Antes disso, citamos que, para uma matriz qualquer, o primeiro termo não nulo de uma linha é chamado de **pivô**.

Definição 1.5.4. Uma matriz $A \in M(n \times m, \mathbb{F})$ está na forma escalonada se:

1. o pivô de uma linha estiver a direita do pivô da linha superior;
2. em cada coluna com um pivô, todas as demais entradas forem nulas;
3. o pivô de cada linha (quando houver) é 1;
4. qualquer linha nula estiver abaixo de qualquer linha não nula.

Lembremos, por outro lado, que as informaões de um sistema linear qualquer se encontram nos coeficientes que acompanham as variáveis e também nos coeficientes do lado esquerdo. Isso motiva a seguinte deifinição.

Definição 1.5.5. Seja (S) um sistema linear qualquer

$$(S) \quad \begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = \beta_1, \\ \qquad\qquad \vdots \\ \alpha_{n1}x_1 + \cdots + \alpha_{nn}x_n = \beta_n. \end{cases}$$

Chamamos de matriz aumentada do sistema (S) a matriz dada por

$$(A|B) = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} & \beta_n \end{pmatrix}.$$

Diremos que uma matriz aumentada está na forma escalonada quando a matriz dada pelas primeiras n colunas estiver na forma escalonada. Ao realizarmos as operações elementares sobre a matriz aumentada de (S) e alcançarmos uma matriz escalonada, obteremos uma nova matriz que representa um sistema escalonado e que é equivalente ao sistema (S).

1.6 Algoritmo de inversão de matrizes

Suponha que queiramos achar a inversa de uma matriz quadrada $A \in M(n, \mathbb{F})$. Posto de outra forma, buscamos $B \in M(n, \mathbb{F})$ tal que $AB = \mathrm{I}_n = BA$. Olhamos para a primeira identidade. A igualdade entre duas matrizes $n \times n$ é também a igualdade de n matrizes coluna, de modo que assumir $BA = \mathrm{I}_n$ é equivalente a

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots \quad, \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{1n} \\ \vdots \\ b_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Como os coeficientes a_{ij} já são conhecidos, o que estamos fazendo nada mais é do que resolvendo n sistemas lineares simultaneamente. Cada um dos sistemas nos dará uma das n colunas da matriz inversa B, se ela existir. Pelo Teorema 1.5.2, se $\det A \neq 0$, então todos estes sistemas serão solúveis. Mais do que isso, esta condição nos dá que A será equivalente à matriz identidade e, como o método de escalonamento com matrizes só depende da matriz dos coeficientes, podemos resolver todos os sistemas simultaneamente.

Teorema 1.6.1 (Algoritmo de inversão)**.** *Seja $A \in M(n, \mathbb{F})$ uma matriz quadrada, com $\det A \neq 0$. Formemos a matriz aumentada*

$$(A|I) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Ao transformarmos esta matriz, através de operações elementares, até que tenhamos uma matriz da forma $(I|B)$, então B será a inversa de A.

Ainda por causa da estreita relação entre sistemas lineares e matrizes, o método de escalonamento (através de operações elementares e portanto matrizes equivalentes) nos dá mais um critério para invertibilidade.

Teorema 1.6.2. *Seja $A \in M_n(\mathbb{F})$ uma matriz quadrada de tamanho qualquer. Então, A é inversível se, e somente se, $A \sim \mathrm{I}_n$, ou seja, a identidade pode ser obtida de A através de operações elementares. Neste caso a mesma sequência de operações elementares que leva A à I_n, também leva I_n à A.*

Demonstração. Veja o Exercício **??**. □

Capítulo 2

Espaços Vetoriais

A noção de espaço vetorial é o que decorre da abstração das estruturas que aparecem quando estudamos sistemas lineares ou vetores no plano ($\mathbb{R}^2$) ou no espaço ($\mathbb{R}^3$). Tomamos todas as propriedades que os conjuntos desejados satisfazem e as colocamos como axiomas, para estudarmos conjuntos gerais que satisfazem estas condições. Esta é uma forma de tratarmos todos os casos conhecidos, e aqueles que por ventura ainda não foram considerados, ao mesmo tempo, ou seja, estudamos as propriedades e os resultados que são **comuns** a qualquer estrutura que satisfaça estes axiomas. O preço que pagamos por esta generalidade é a abstração, por vezes um fator que confunde, mas que pode com ainda mais eficácia elucidar.

2.1 Definições

Definição 2.1.1. Um *espaço vetorial* V sobre um corpo $\mathbb{F}$ consiste de:

1. Um conjunto V, cujos elementos são chamados vetores.
2. Um corpo $\mathbb{F}$, cujos elementos são chamados escalares.
3. Duas operações $+ : V \times V \to V$ e $\cdot : F \times V \to V$ que satisfazem:

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

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

(A3) existe em V um elemento, chamemos este de o, que é neutro em relação a adição, ou seja,
$$u + o = u\,, \quad \forall\, u \in V;$$

(A4) para todo $u \in V$ existe um elemento $-u \in V$ tal que $u + (-u) = o$;

(M1) para todos $u, v \in V$, $\alpha \in \mathbb{F}$,
$$\alpha(u + v) = \alpha u + \alpha v\,;$$

(M2) para todos $u \in V$, $\alpha, \beta \in \mathbb{F}$
$$(\alpha + \beta)u = \alpha u + \beta u\,;$$

(M3) para todos $u \in V$, $\alpha, \beta \in \mathbb{F}$

$$(\alpha\beta)u = \alpha(\beta u)\,;$$

(M4) para todo $u \in V$,

$$1 \cdot u = u\,.$$

Chamaremos de *vetores* os elementos de um espaço vetorial.

Começamos a ver as recompensas de buscar a abstração dos conceitos usuais que conhecemos. O próximo resultado enumera um conjunto de propriedades que todo espaço vetorial irá satisfazer.

Proposição 2.1.2. *Seja V um espaço vetorial sobre um corpo $\mathbb{F}$. Então:*

1. *para todo $u \in V$, temos $(-1)u = -u$;*
2. *para todo $u \in V$, temos $0 \cdot u = o$;*
3. *para todo $\alpha \in V$, temos $\alpha \cdot o = o$;*
4. *se $\alpha u = o$, então ou temos $\alpha = 0$ ou $u = o$.*

Estas são propriedades que podem parecer triviais, mas tem a vantagem de serem propriedades universais dos espaços vetoriais. Esta é uma ideia que deve ser apreciada: o preço de um resultado (geralmente quantificado pela dificuldade da prova) acompanha inversamente a sua generalidade.

Definição 2.1.3. Seja V um espaço vetorial sobre F. Um *subespaço vetorial* de V é um subconjunto $W \subset V$ tal que:

1. $o \in W$.
2. Se $u, v \in W$ então $u + v \in W$.
3. Se $\alpha \in F$ e $u \in W$ então $\alpha u \in W$.

Podemos ainda substituir as duas últimas propriedades por

4 Se $\alpha \in F$ e $u, v \in W$ então $u + \alpha v \in W$.

Esta definição de subespaço vetorial é uma forma mais eficaz de dizer que o subconjunto $W \subset V$ será um espaço vetorial **com relação às operações de** V. A definição acima já basta, não tendo provado todas as propriedades $(S1 - M4)$, porque os elementos de W estão por definição em V, o que já confere a maioria das identidades necessárias. Tudo o que precisamos, basicamente, é pedir que W seja fechado perante as operações de soma e multiplicação por escalar.

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

$$W = \{(x, y) \mid ax + by = c\}\,,$$

ou seja, W é uma reta. Por definição, W será um subespaço vetorial de $\mathbb{R}^2$ se

1. $o \in W$;
2. para todos $u, v \in W$ e $\alpha \in \mathbb{R}$, tenhamos

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

É fácil ver que a primeira condição será satisfeita quando, e apenas quando, $d = 0$. De outra forma, já podemos concluir que uma reta que não passa pela origem não pode ser um subespaço vetorial.

Em seguida, consideremos $d = 0$ e olhemos a segunda condição. Sejam

$$u = (x_1, y_1)\,, v = (x_2, y_2) \in W\,, \quad \alpha \in \mathbb{R}\,.$$

Segue daí que

$$u + \alpha v = (x_1 + \alpha x_2, y_1 + \alpha y_2)$$

e podemos calcular que

$$a(x_1 + \alpha x_2) + b(y_1 + \alpha y_2) = \underbrace{ax_1 + by_1}_{0} + \alpha\,\underbrace{(x_2 + b\alpha y_2)}_{0} = 0\,.$$

Concluímos, portanto, que W será um subespaço de $\mathbb{R}^2$ se, e somente se, $d = 0$, ou seja se a reta passa pela origem.

Proposição 2.1.4. *Se U e W são subespaços vetoriais de V, então $U \cap W$ também é um subespaço vetorial de V. O conjunto $U + W = \{u + w \mid u \in U, w \in W\}$ também é um subespaço vetorial de V.*

O caso em que a interseção de dois subespaços é trivial (note que esta nunca será o vazio, pois ao menos o vetor nulo deve estar nos dois subespaços) é interessante. Mais a frente será possível decompor um espaço vetorial V da maneira mais eficaz possível com esta noção.

Definição 2.1.5. Se U e W são subespaços vetoriais de V e $U \cap W = \{0\}$, então $U + W$ é chamado de *soma direta* de U com W, denotada por $U \oplus W$.

A importância da soma direta está no fato de que ela nos permite descrever os elementos da soma como uma combinação única de elementos em cada um dos subespaços iniciais.

Proposição 2.1.6. *Seja V um espaço vetorial e U, W subconjuntos de V. São equivalentes:*

1. *$U + W$ é uma soma direta;*
2. *Cada vetor $v \in U + W$ pode ser escrito de maneira única como $v = u + w$ com $u \in U, w \in W$.*

Até aqui consideramos apenas subespaços gerais, não concretos. Vejamos agora qual é o principal método de obter um subespaço vetorial, de alguma forma mais concreto. A ideia é a seguinte: Considere V um espaço vetorial e $u \in V$ um vetor qualquer. Note que se $U \leq V$ é um subespaço vetorial e $u \in U$, então U deve conter também qualquer vetor αu, $\alpha \in \mathbb{F}$, pela terceira condição na Definição 2.1.3. Se U contém um outro vetor v, então U deve conter além dos múltiplos αu e βv, com α, β escalares quaisquer, as **combinações lineares**

$$w = \alpha u + \beta v\,, \quad \forall\, \alpha, \beta \in \mathbb{F}\,.$$

Como consideramos aqui U um subespaço qualquer, temos a afirmação de que *todo subespaço de V que contem u e v deve também conter todas as combinações lineares entre os dois vetores.* Delas, podemos definir o menor subespaço contendo u e v.

Definição 2.1.7. Sejam $v_1, v_2, \cdots, v_k$ vetores de um espaço vetorial V. Chamamos de combinação linear qualquer elemento de V da forma

$$u = \alpha_1 v_1 + \cdots + \alpha_n v_n\,, \quad \alpha_j \in \mathbb{F}\,.$$

Para a próxima definição, ressaltemos a diferença entre um subconjunto de V e um subespaço de V. Um subconjunto é apenas uma coleção de elementos de V, sem qualquer requerimento portanto sobre as relações entre os mesmos. A definição de subespaço é mais forte, e pede que a coleção seja **fechada quanto as operações de V**. Entre outras coisas, vemos que **um subespaço de V é sempre um subconjunto (infinito) de V**. Um subconjunto de vetores V será definido com chaves $\{\}$, enquanto que para o subespaço contendo os mesmos vetores, como veremos agora, empregaremos os colchetes $[]$.

Definição 2.1.8. Sejam $v_1, v_2, \cdots, v_k$ vetores de um espaço vetorial V. O subespaço de V gerado por $S = \{v_1, \cdots, v_k\}$ é definido da forma

$$[S] = \left\{ \sum_{i=1}^{k} \alpha_i v_i \, ; \, \alpha_i \in \mathbb{F} \, , i = 1, \cdots, k \right\} .$$

Se queremos explicitar quais são os vetores geradores, escrevemos também $[v_1, \cdots, v_k]$.

Quando evocamos a noção de menor ou maior para subconjuntos, queremos dizer que U é maior que V se o último está contido no primeiro, ou seja se todo elemento de V é um elemento de U também. Desta forma, ponhamos a afirmação anterior em termos mais técnicos.

Proposição 2.1.9. *Seja $S = \{v_1, v_2, \cdots, v_k\}$ um subconjunto finito do espaço vetorial V. Então $[S]$ é o menor subespaço vetorial de V que contem os vetores v_i, $i = 1, \cdots, n$, ou seja, qualquer subespaço de V que contenha S irá conter também $[S]$.*

2.2 Bases e dimensão

O nosso objetivo quando falamos de subespaços gerados é descrever os espaços com os quais trabalhamos em função de um certo conjunto (finito) de vetores. Isso é o que já é feito, mesmo que inconscientemente, quando definimos o plano $\mathbb{R}^2$ da forma

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

Segue desta definição e da noção de soma de vetores que qualquer elemento de $\mathbb{R}^2$ pode ser escrito da forma

$$(x, y) = x(1, 0) + y(0, 1) , \quad x, y \in \mathbb{R}$$

ou seja, como uma combinação linear dos vetores $e_1 = (1, 0)$ e $e_2 = (0, 1)$.

Exemplo. Vejamos que podemos fazer a mesma coisa com qualquer subespaço de $\mathbb{R}^n$. Por exemplo, considere em $V \subset \mathbb{R}^3$ o seguinte subespaço

$$V = \{(x, y, z) \mid 2x - y + 5z = 0\} .$$

Uma possível parametrização de V, sabemos, é a seguinte

$$\begin{aligned} V =& \{(x, y, z) \mid y = 2x + 5z\} \\ =& \{(x, 2x + 5z, z) \mid x, z \in \mathbb{R}\} . \end{aligned}$$

Da mesma forma como antes, podemos então escrever cada $v \in V$ como

$$v = x(1, 2, 0) + z(0, 5, 1) , \quad x, z \in \mathbb{R} .$$

Esta equação é valida para todo $v \in V$ e para cada um, mudamos apenas o valor dos parâmetros x e z. Quando colocamos um par de vetores, chamemos $v_1, v_2 \in V$, nesta descrição, é um trabalho mais simples obter a soma $v_1 + v_2$, por exemplo.

$$v_1 = x_1(1,2,0) + z_1(0,5,1)\,, \quad v_2 = x_2(1,2,0) + z_2(0,5,1)$$

$$\Rightarrow v_1 + v_2 = (x_1 + x_2)(1,2,0) + (z_1 + z_2)(0,5,1)\,.$$

Portanto, buscamos ser capaz de fazer o mesmo para qualquer espaço vetorial que tenhamos. Como buscamos uma descrição que seja a mais simples possível, procuramos o menor conjunto de vetores que nos permita fazer a descrição do espaço da maneira como acima, ou seja, que haja como um **conjunto gerador**.

Definição 2.2.1. Seja V um espaço vetorial sobre $\mathbb{F}$. Um *conjunto gerador* de V é um subconjunto $\{v_1, \cdots, v_k\} \subset V$ tal que

$$V = [v_1, \cdots, v_k]\,,$$

ou seja, tal que qualquer elemento de V seja uma combinação linear dos vetores v_i, $i \in \{1, \cdots, k\}$.

Existem espaços vetoriais onde não é possível achar tal conjunto com um número finito de elementos apenas, mas não trataremos destes espaços neste curso, a não ser possivelmente em exercícios complementares. Portanto, para todo espaço vetorial que consideremos, será possível obter um número mínimo de vetores geradores. Estes são chamados de espaços vetoriais *finitamente gerados*.

Exemplo. Considere $V = \mathbb{R}^3$ e $U \leq V$ dado por

$$U = \{(\alpha + \beta + 2\gamma, 2\beta + 2\gamma, -\alpha + \beta)\,, \alpha, \beta, \gamma \in \mathbb{R}\}\,.$$

Como já fizemos anteriormente, cada elemento de U é da forma

$$u = \alpha \underbrace{(1,0,-1)}_{u_1} + \beta \underbrace{(1,2,1)}_{u_2} + \gamma \underbrace{(2,2,0)}_{u_3}\,,$$

de modo que $\{u_1, u_2, u_3\}$ é um conjunto gerador de U.

Note, porém, que $u_3 = u_1 + u_2$. Desta forma, qualquer combinação linear destes três vetores pode ser posta como combinação linear de u_1 e u_2, apenas.

$$\alpha u_1 + \beta u_2 + \gamma u_3 = (\alpha + \gamma)u_1 + (\beta + \gamma)u_2\,.$$

Conseguimos portanto descrever U com apenas dois vetores, ao invés de três.

A condição não satisfeita por u_1, u_2, u_3 no último exemplo, mas que é satisfeita por u_1 e u_2, e que nos dará um conjunto ótimo, é o que denotamos por independência linear.

Definição 2.2.2. Um conjunto $L = \{u_1, \ldots, u_n\} \subset V$ é *linearmente independente* (L.I.) se

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0 \implies \alpha_1 = \cdots = \alpha_n = 0. \tag{2.1}$$

Um conjunto $L \subset$ que não é L.I. é denominado linearmente dependente (L.D.).

Tomemos alguma linhas a mais para entender o que significa a condição (2.1). A negação dela é que existem escalares $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$ tais que não são *todos* simultaneamente nulos e tais que

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0\,.$$

Suponhamos, por exemplo, que $\alpha_1 \neq 0$. Neste caso, podemos isolar o vetor u_1 e escrever

$$u_1 = -\frac{\alpha_2}{\alpha_1} u_2 + \cdots - \frac{\alpha_n}{\alpha_1} u_n \,.$$

Em outras palavras, o vetor u_1 é uma combinação linear dos demais vetores de L. Entre outras coisas, isso implica que qualquer vetor que possa ser escrito como combinação linear dos vetores de L pode, na verdade, ser escrito como combinação linear de $L \backslash \{u_1\}$.

Proposição 2.2.3. *Um subconjunto $S \subset V$ é L.D. se, e somente se, um de seus vetores pode ser escrito como combinação linear dos demais. Em outras palavras, S é L.D. se, e somente se, existe um vetor $v_j \in S$ tal que $v_j \in [S \backslash \{v_j\}]$, ou ainda*

$$v_j = \alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1} + \alpha_{j+1} v_{j+1} + \cdots + \alpha_k v_k \,.$$

O mesmo não é o caso se L for L.I. De fato, se todo $v \in [L]$ é escrito como combinação linear de $L \backslash \{u_1\}$, então

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = v = \beta_2 u_2 + \cdots + \beta_n u_n \,,$$

para toda n-upla de escalares $\alpha_1, \cdots, \alpha_n$. Ora, tome em particular $\alpha_1 = 1$ e todos os outros escalares nulos, de modo que

$$u_1 - \beta_2 u_2 - \cdots - \beta_n u_n = 0 \,.$$

Isso conclui que L deve ser L.D.

Vejamos algumas propriedades de conjuntos L.I. e L.D.

Proposição 2.2.4. *Seja S um subconjunto finito de V. Então:*

1. *Se S é L.I., então qualquer $L \subset S$ é também L.I.*
2. *Se S é L.D., então qualquer $L \supset S$ é também L.D.*
3. *Se $v \in [S]$, então o conjunto gerado por $S \cup \{v\}$ é igual a $[S]$.*
4. *Se $o \in S$, então S é L.D.*
5. *Se $S = \{v\}$ é um conjunto unitário, então S é L.I. se, e somente se, $v \neq 0$.*
6. *Se S é L.I. e $v \in S$ é um elemento qualquer de S, então $S \backslash \{v\}$ gera um subespaço estritamente menor que $[S]$.*

A afirmação 6. é de grande importância, pois implica que um modo de verificarmos que um conjunto é o menor conjunto gerador de um espaço é apurar se este é L.I. ou L.D. Tal conjunto minimal é o que chamamos de **base de** V.

Definição 2.2.5. Seja V um espaço linear. Uma base de V é um conjunto L.I. de vetores que gera V.

Podemos nos perguntar agora qual o método geral para se obter uma base de um espaço vetorial V qualquer. A ideia é que começamos com um vetor $v \in V$ e acrescentamos vetores ao conjunto $B_1 = \{v\}$ até obtermos um que gera V, sempre verificando se o presente conjunto é L.I. Não temos porém um método eficaz que nos ajude a escolher o próximo vetor a ser adicionado.

Exemplo. Seja $V = \mathbb{R}^4$ e defina o subespaço dado pela solução do sistema homogêneo

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

Escalonar este sistema é uma tarefa trivial e ao final chegaremos ao sistema escalonado

$$\begin{cases} x + 9z + 9w = 0, \\ y - 5z + 6w = 0\ ,. \end{cases}$$

Obtemos desta forma que o subespaço W das soluções deste sistema será dado por

$$W = \{(9(z+w), 5z + 6w, z, w) \mid z, w \in \mathbb{R}\}\,.$$

Em outras palavras, cada elemento de W é da forma

$$v = z(9,5,1,0) + w(9,6,0,1)\,,$$

de modo que chegamos a um conjunto gerador de W

$$W = [(9,5,1,0), (9,6,0,1)]\,.$$

Estes dois vetores formam um conjunto L.I., pois

$$\begin{aligned} \alpha(9,5,1,0) + \beta(9,6,0,1) = 0 &\implies (9(\alpha+\beta), 5\alpha + 6\beta, \alpha, \beta) = (0,0,0,0) \\ &\implies \alpha = \beta = 0\,. \end{aligned}$$

Concluímos que $\{(9,5,1,0), (9,6,0,1)\}$ é uma base para W.

Para o caso geral de subespaços de $\mathbb{R}^n$ dados por sistemas lineares homogêneos, um algoritmo pode ser descrito.

Exemplo. Seja $V \subset \mathbb{R}^n$ o subespaço linear definido como o conjunto solução do sistema linear homogêneo

$$(S) \qquad \begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = 0, \\ \vdots \\ \alpha_{m1}x_1 + \cdots + \alpha_{mn}x_n = 0, \end{cases} \qquad .$$

Seja A a matriz dos coeficientes de (S) e seja

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

a matriz escalonada equivalente à matriz A. A última equação do sistema é da forma

$$x_k + \beta_{m(k+1)}x_{k+1} + \cdots + \beta_{mn}x_n = 0\,,$$

ou seja,

$$x_k = -\beta_{m(k+1)}x_{k+1} - \cdots - \beta_{mn}x_n\,. \tag{2.2}$$

A próxima equação será da forma

$$x_{l-1} + \beta_{(m-1)(l+1)}x_{l+1} + \cdots + \beta_{(m-1)k}x_k + \cdots + \beta_{(m-1)n}x_n = 0\,,$$

com $l < k$, de modo que de (2.2), temos

$$x_{l-1} = \left(-\beta_{(m-1)(l+1)} x_{l+1} - \cdots\right) + \left(\beta_{(m-1)k} \beta_{m(k+1)} x_{k+1} + \cdots + \beta_{(m-1)k} \beta_{mn} x_n\right) + \\ - \left(\beta_{m(k+1)} x_{k+1} + \cdots + \beta_{mn} x_n\right) .$$

Fazemos isto até chegarmos à primeira equação, até todas as variáveis dependentes estarem em termos das variáveis independentes. Com isto, conseguimos saber exatamente quais são os vetores geradores de V.

É evidente que uma base não é algo único de um espaço vetorial. Para o mesmo V, existem infinitas bases. Uma quantidade, o número de elementos, permanece constante quando olhamos para todas elas. A condição de que B seja um conjunto gerador de V nos permite fixar um número suficiente e a condição de que B seja L.I. não permite que esse número seja menor. Esta constância é que nos permite falar sobre **dimensão**.

Teorema 2.2.6. *Seja V um espaço vetorial. Então toda base de V possui o mesmo número de vetores.*

Definição 2.2.7. Seja V um espaço vetorial. A dimensão de V, escrita $\dim V$, é o número de vetores de uma (e portanto toda) base de V.

Vejamos agora alguns resultados que costumam facilitar nossas contas.

Proposição 2.2.8. *Seja V um espaço vetorial de dimensão n. Então:*

1. *todo subconjunto S de tamanho n L.I. é uma base para V.*
2. *todo subconjunto S gerador de V de tamanho n é uma base para V.*
3. *qualquer subconjunto de m elementos, com $m > n$, é L.D.*
4. *nenhum subconjunto de m elementos, com $m < n$, pode gerar V.*

Teorema 2.2.9 (Teorema do Completamento)**.** *Seja V um espaço vetorial de dimensão n, $k < n$ e seja $\{v_1, \cdots, v_k\}$ um subconjunto L.I. de V com k elementos. Então existem $n - k$ vetores, que denotaremos por $v_{k+1}, \cdots, v_n$, tais que*

$$B = \{v_1, \cdots, v_k, v_{k+1}, \cdots, v_n\}$$

é uma base de V.

Vejamos alguns resultados mais sobre dimensão, dizendo respeito a subespaços. O primeiro é um análogo do resultado que temos para tamanho de conjuntos. Sabemos que se U, W são dois conjuntos quaisquer, então o tamanho da união $U \cup W$ é igual a soma dos tamanhos de U e W menos o tamanho da interseção. No contextos de espaços vetoriais, o tamanho é substituído pela dimensão.

Proposição 2.2.10. *Seja V um espaço vetorial e sejam U, W dois subespaços vetoriais de V. Então*

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

2.3 Obtendo bases para subespaços de $\mathbb{R}^n$

O método de obtenção de uma base para qualquer subespaço U de $\mathbb{R}^n$ começa com um conjunto de geradores, do qual extraímos um conjunto L.I. para formar uma base. Consideremos então um conjunto

$$S = \{v_1, \cdots, v_m\}$$

com m vetores em $\mathbb{R}^n$. Dizer que S é L.I. significa dizer que existe uma m-upla de escalares $(\alpha_1, \cdots, \alpha_m) \neq (0, \cdots, 0)$ tal que

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0\,. \tag{2.3}$$

Se explicitamos todas as coordenadas destes m vetores,

$$v_1 = (v_{11}, \cdots, v_{1n})\,, \quad v_2 = (v_{21}, \cdots, v_{2n})\,, \cdots, \quad v_m = (v_{m1}, \cdots, v_{mn})\,, \tag{2.4}$$

então a equação (2.3) é equivalente ao sistema linear

$$\begin{cases} v_{11}\alpha_1 + v_{21}\alpha_2 + \cdots + v_{m1}\alpha_m = 0 \\ v_{12}\alpha_1 + v_{22}\alpha_2 + \cdots + v_{m2}\alpha_m = 0 \\ \vdots \\ v_{1n}\alpha_1 + v_{2n}\alpha_2 + \cdots + v_{mn}\alpha_m = 0 \end{cases}$$

Portanto, S será L.I. se este sistema possui apenas solução única, enquanto será L.D. se este sistema possuir qualquer outra solução (ou seja, infinitas soluções). A solução, sabemos, pode ser obtida através do escalonamento da matriz

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

que nada mais é do que a matriz cujas colunas são os vetores de S. Pelo que vimos durante a discussão de sistemas lineares, o sistema acima terá solução única se ao escalonarmos obtivermos

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

onde o bloco superior for a identidade de tamanho n. Em particular, se temos $m > n$, chegar a matriz da forma acima é impossível. Já encontramos este resultado na Proposição 2.2.8, item 3.

Por outro lado, podemos trabalhar com a matriz cujas linhas são os vetores de S, ao invés da coluna. Vejamos como: as seguintes afirmações são verdades para qualquer $S = \{v_1, \cdots, v_m\} \subset \mathbb{R}^n$.

1. O subespaço gerado $[S]$ é o mesmo que o espaço gerado quando trocamos dois vetores de lugar,
$$[v_1, \cdots, v_k, \cdots, v_j, \cdots, v_m] = [v_1, \cdots, v_j, \cdots, v_k, \cdots, v_m]\,.$$

2. O subespaço gerado $[S]$ é o mesmo que o espaço gerado quando trocamos um vetor pela sua combinação linear com outro vetor,

$$[v_1, \cdots, v_k, \cdots, v_j, \cdots, v_m] = [v_1, \cdots, v_k, \cdots, \alpha v_k + \beta v_j, \cdots, v_m] \,.$$

Levando em conta estas afirmações, vemos que se coletamos os vetores de S em uma matriz na forma (utilizamos aqui a mesma notação para os vetores dada em (2.4))

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

então qualquer matriz equivalente a B terá suas linhas gerando o mesmo subespaço que S. Por fim, afirmamos que qualquer matriz escalonada *sem linhas nulas* reúne vetores, em suas linhas, que formam um conjunto L.I. De fato, suponhamos que E seja uma matriz $m \times n$ escalonada e sejam $u_1, \cdots, u_m$ seus vetores linha. Se olharmos para o pivô da primeira linha, então todos os elementos restantes desta coluna, suponhamos que seja a $k-$ésima coluna, são nulos. Logo, na $k-$ésima entrada da equação

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0$$

nos dará $\alpha_1 = 0$. Prosseguindo desta forma para cada linha, tendo cada uma delas um pivô, obteremos $\alpha_2 = \cdots = \alpha_m = 0$.

2.4 Coordenadas e mudança de base

Como já mencionado, o objetivo ao introduzirmos o conceito de base é descrever qualquer elemento de V uniformemente, da maneira mais simples possível. O que queremos, em outras palavras, é um **sistema de coordenadas** para V. Esta nomenclatura já é usual no exemplo canônico, $\mathbb{R}^3$, onde sabemos que os vetores

$$\vec{i} = (1,0,0)\,, \quad \vec{j} = (0,1,0)\,, \quad \vec{k} = (0,0,1)$$

formam um sistema de coordenadas para $\mathbb{R}^3$. Com estes, qualquer ponto $p \in \mathbb{R}^3$ pode ser escrito como

$$p = (x,y,z) = x\vec{i} + y\vec{j} + z\vec{k}\,.$$

Neste caso, x, y e z são as coordenadas do ponto p. Sabemos que $\{\vec{i}, \vec{j}, \vec{k}\}$ é uma base de $\mathbb{R}^3$. Por outro lado, $\{\vec{k}, \vec{i}, \vec{j}\}$ também é uma base de $\mathbb{R}^3$, mas quando escrevemos (x, y, z) convencionamos que a coordenada x é a que acompanha $\vec{i}$ e não $\vec{k}$. Ao falarmos de coordenadas, portanto, importa a ordem com que olhamos os vetores da base, para que saibamos diferenciar, por exemplo, os pontos $(1, 2, 0)$ e $(0, 1, 2)$. Segue disso o conceito de **base ordenada**.

Definição 2.4.1. Seja V um espaço vetorial de dimensão n. Uma base ordenada B de V é uma bijeção entre $\{1, 2, \cdots, n\}$ e V, tal que a imagem é uma base de V. Frequentemente denotamos uma base ordenada de V por $(v_1, \cdots, v_n)$, com parênteses ao invés de chaves[1].

Com isso, prosseguimos como no exemplo acima, para $\mathbb{R}^3$, e podemos definir as coordenadas de qualquer vetor.

[1]Tratando de vetores em $\mathbb{R}^n$, porém, onde os elementos já possuem parênteses, utilizaremos também as chaves para falar sobre bases ordenadas e esperamos que não haverá ambiguidade.

Definição 2.4.2. Seja V um espaço vetorial e seja $B = (v_1, \cdots, v_n)$ uma base ordenada de V. Dado um vetor
$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n \,,$$
os escalares $\alpha_1, \cdots, \alpha_n$ são denominados as *coordenadas* de v na base B.

Notação : Com o advento das coordenadas, podemos colocar os elementos de qualquer espaço vetorial na mesma condição que os pontos de $\mathbb{R}^3$, ou qualquer $\mathbb{R}^n$. Por exemplo, para uma base ordenada $B = (v_1, \cdots, v_n)$ e um vetor qualquer

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n \,,$$

podemos escrever este como $(\alpha_1, \cdots, \alpha_n)$. Isso, claro, só é uma descrição eficaz se está fixada qual a base em questão. Por este motivo, usamos também a notação

$$(v)_B = (\alpha_1, \cdots, \alpha_n)$$

para simbolizar as coordenadas de v *na base* B. Por fim, podemos coletar também todas as coordenadas de um vetor qualquer em um vetor coluna ao invés de um vetor linha,

$$(v)_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} .$$

Exemplo. Considere V um espaço vetorial qualquer de dimensão v e $B = (v_1, \cdots, v_n)$ uma base ordenada de V. Então, as coordenadas de cada v_k na própria base B são triviais. Temos

$$(v_1)_B = (1, 0, \cdots, 0)\,, \quad (v_2)_B = (0, 1, \cdots, 0)\,, \quad \cdots \quad , \quad (v_n)_B = (0, \cdots, 0, 1)\,.$$

É importante citar aqui que, ao condicionarmos nosso estudo a bases ordenadas, obteremos unicidade para as coordenadas de qualquer vetor.

Proposição 2.4.3. *Seja V um espaço vetorial e $B = (v_1, \cdots, v_n)$ uma base ordenada de V. Então, as coordenadas de um vetor na base B são unicamente determinadas pelo vetor.*

Demonstração. Estamos afirmando aqui que $v \in V$ não pode ser representado por duas matrizes coluna diferentes. De fato, se
$$(v)_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} ,$$
então temos
$$\alpha_1 v_1 + \cdots + \alpha_n v_n = v = \beta_1 v_1 + \cdots + \beta_n v_n \,.$$
Segue disso que
$$(\alpha_1 - \beta_1) v_1 + \cdots + (\beta_1 - \beta_n) v_n = 0 \,.$$
Ora, como B é L.I., obtemos

$$\alpha_1 - \beta_1 = \cdots = \alpha_n - \beta_n = 0 \,,$$

ou seja, $\alpha_k = \beta_k$, para todo $k \in \{1, \cdots, n\}$. □

Podemos nos perguntar agora como são as coordenadas dos vetores de uma certa base B em uma base C diferente. Considere, por exemplo,

$$B = (v_1, \cdots, v_n)\,, \quad C = (u_1, \cdots, u_n)$$

duas bases ordenadas de um espaço vetorial V de dimensão n. Como C gera todo o V, cada u_k pode ser escrito como uma combinação linear dos elementos da base B. Utilizaremos os símbolos α_{lk}, com $l = 1, \cdots, n$, para as coordenadas de u_k em B, ou seja,

$$u_1 = \alpha_{11}v_1 + \alpha_{21}v_2 + \cdots + \alpha_{n1}v_n\,.$$

$$u_2 = \alpha_{12}v_1 + \alpha_{22}v_2 + \cdots + \alpha_{n2}v_n\,.$$

$$\cdots$$

$$u_n = \alpha_{1n}v_1 + \alpha_{2n}v_2 + \cdots + \alpha_{nn}v_n\,.$$

Chamamos a matriz $M = (\alpha_{ij})$ de *matriz mudança da base* B para a base C. Explicitamente,

$$M = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix}.$$

A matriz M é construída portanto descrvendo os vetores de C em termos de B e construindo cada **coluna** de M com as **linhas** das combinações lineares obtidas. A justificativa para fazermos esta troca, ou seja pegarmos a transposta da relação obtida, é a seguinte: suponha que $w \in V$ seja um vetor com coordenadas $(\beta_1, \cdots, \beta_n)$ na base B e com coordenadas $(\gamma_1, \cdots, \gamma_n)$ na base C. Então,

$$w = \gamma_1 u_1 + \cdots + \gamma_n u_n\,,$$

de onde segue que podemos escrever

$$\begin{aligned} w =& \gamma_1(\alpha_{11}v_1 + \cdots + \alpha_{n1}v_n) + \gamma_2(\alpha_{12}v_1 + \cdots + \alpha_{n2}v_n) + \\ &+ \gamma_n(\alpha_{1n}v_1 + \cdots + \alpha_{nn}v_n)\,. \end{aligned}$$

Coletando todos os coeficientes que seguem cada v_k, teremos

$$\begin{aligned} w =& (\gamma_1\alpha_{11} + \cdots + \gamma_n\alpha_{1n})v_1 + (\gamma_1\alpha_{21} + \cdots + \gamma_n\alpha_{2n})v_2 + \\ &+ (\gamma_1\alpha_{n1} + \cdots + \gamma_n\alpha_{nn})v_n\,. \end{aligned}$$

Isso nos dá as coordenadas de w na base C, sendo que a k-ésima coordenada será

$$\beta_k = (\alpha_{k1}\gamma_1 + \cdots + \alpha_{kn}\gamma_n)\,,$$

que pode também ser escrito como $\beta_k = \sum_{j=1}^{n} \alpha_{kj}\gamma_j$. Portanto, podemos ver que

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

Esta fórmula nos diz que: *a matriz que transforma as coordenadas de uma base C nas coordenadas de uma base B é a transposta da matriz que transforma os vetores de C nos vetores de B.* Escrevamos isto na forma de uma proposição.

Proposição 2.4.4. *Considere V um espaço vetorial de dimensão n e sejam $B = (v_k)$, $C = (u_k)$ duas bases ordenadas de V. Se $M = (\alpha_{ij})$ é a matriz que descreve os vetores de C em relação à base B, através das fórmulas*

$$u_k = \sum_{j=1}^{n} \alpha_{jk} v_j \,, \quad k = 1, \cdots, n \,,$$

então M transforma as coordenadas de um vetor qualquer na base C para as coordenadas na base B, através da igualdade matricial

$$(w)_B = M(w)_C \,.$$

Comumente denotamos a matriz mudança de base (de B para C) por $I_{B\to C}$, para explicitar quais são as bases em questão.

Exemplo. Considere $V = \mathbb{R}^3$, B a base canônica de V e

$$C = \{(1, 1, 0), (0, 1, 0), (0, 0, 3)\}$$

uma segunda base ordenada de V. Para cada vetor de V, é fácil obter $(v)_B$, enquanto achar $(v)_C$ é mais trabalhoso. Suponhamso que precisamos das coordenadas dos seguintes quatro vetores

$$w_1 = (2, 3, 12) \,, w_2 = (0, 3, 1) \,, w_3 = (-1, -1, -1) \,, w_4 = (5, -5, 0)$$

na base C. A matriz mudança de base permite que a gente resolva este problema para diversos vetores, simultaneamente. Vamos então obter esta matriz.

Como vimos, para isso precisamos colocar os vetores de B em relação aos vetores de C. Temos

$$(1, 0, 0) = \alpha_1(1, 1, 0) + \alpha_2(0, 1, 0) + \alpha_3(0, 0, 3) \,,$$

$$(0, 1, 0) = \beta_1(1, 1, 0) + \beta_2(0, 1, 0) + \beta_3(0, 0, 3) \,,$$

$$(0, 0, 1) = \gamma_1(1, 1, 0) + \gamma_2(0, 1, 0) + \gamma_3(0, 0, 3) \,.$$

Isso nos dará três sistemas lineares, o primeiro sendo

$$\begin{cases} \alpha_1 = 1 \\ \alpha_1 + \alpha_2 = 0 \\ \alpha_3 = 0 \end{cases} \implies \alpha = 1 \,, \alpha_2 = -1 \,, \alpha_3 = 0 \,.$$

Deixamos a cargo do leitor resolver os demais. Acharemos

$$M = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$$

Portanto, agora podemos achar $(w_k)_C$ utilizando apenas a multiplicação de martizes,

$$(w_1)_C = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 12 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} \qquad (w_2)_C = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 1/3 \end{pmatrix} \,,$$

$$(w_3)_C = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -1/3 \end{pmatrix} \qquad (w_4)_C = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} 5 \\ -5 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -10 \\ 0 \end{pmatrix} \,.$$

Exemplo. Suponhamos que tomamos a base C como sendo a própria base B, ou seja, estamos perguntando qual a matriz que irá levar as coordenadas de um vetor qualquer, $(v)_B$, no mesmo vetor coluna,

$$(v)_B = I_{B\to B}(v)_B \,.$$

Ora, a única matriz cuja multiplicação por qualquer vetor coluna preserva este vetor é a identidade. Teremos portanto $I_{B\to B} = I_n$, isto para qualquer base ordenada B que escolhamos.

De fato, podemos calcular $I_{B\to B}$ como temos feito. Se $B = (v_1, \cdots, v_n)$, podemos ver que

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

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

$$\cdots$$

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

A transposta da matriz que aparece nestas equações nada mais é do que I_n.

A seguir, listemos algumas propriedades desta matriz.

Proposição 2.4.5. *Sejam V um espaço vetorial de dimensão n, e $B = (u_k), C = (v_k)$ e $D = (w_k)$ bases ordenadas de V. Então,*

1. $I_{B\to D} = I_{B\to C} I_{C\to D}$;
2. $I_{B\to C} = I^{-1}_{C\to B}$.