



Universidade Regional do Cariri - URCA
Departamento de Matemática
Programa de Mestrado Profissional
em Matemática em Rede Nacional



Matrizes e algumas aplicações

Francisco Genilson dos Santos Silva

Juazeiro do Norte - CE
2021

Matrizes e algumas aplicações

Francisco Genilson dos Santos Silva

Dissertação de Mestrado apresentada à
Comissão Acadêmica Institucional do
PROFMAT-URCA como requisito parcial
para obtenção do título de Mestre em
Matemática.

Orientador

Prof. Dr. José Tiago Nogueira Cruz

Juazeiro do Norte - CE
2021

Ficha Catalográfica elaborada pela Biblioteca Central da Universidade Regional do Cariri – URCA
Bibliotecária: Ana Paula Saraiva de Sousa CRB: 3/1000

Silva, Francisco Genilson dos Santos.
S586m Matrizes e algumas aplicações/ Francisco Genilson dos Santos
Silva. – Juazeiro do Norte-CE, 2021
177p.

Dissertação apresentada ao Programa de Mestrado Profissional
em Matemática em Redes - PROFMAT
Orientador: Prof. Dr. José Tiago Nogueira Cruz.

1. Matrizes, 2. Ensino Médio, 3. Aplicações, 4. Rotação de
cônicas, 5. Teorema espectral; I. Título.

CDD: 512.943

Matrizes e algumas aplicações

Francisco Genilson dos Santos Silva

Dissertação de Mestrado apresentada à
Comissão Acadêmica Institucional do
PROFMAT-URCA como requisito par-
cial para obtenção do título de Mestre
em Matemática.

BANCA EXAMINADORA

Prof. Dr José Tiago Nogueira Cruz (Orientador)
Universidade Regional do Cariri (URCA)

Prof. Dr. Júnio Moreira de Alencar
Instituto Federal de Educação e Tecnologia do Ceará (IFCE)

Prof. Dr. Flávio França Cruz
Universidade Regional do Cariri (URCA)

*Dedicado à minha esposa Simone, a qual
acreditou em mim e me incentivou nos mo-
mentos em que mais precisei.*

Agradecimentos

À minha esposa Simone, pois esta me apoiou nos momentos mais difíceis e me incentivou a não desistir quando desanimei.

Aos meus pais, que sempre me apoiaram e me ensinaram que a educação é o melhor caminho.

Ao meu orientador, Professor Doutor José Tiago Nogueira Cruz, por acreditar em meu potencial e se mostrar sempre disponível quando as dúvidas apareciam.

Ao Professor Doutor Júnio Moreira de Alencar que, durante toda a graduação me incentivou bastante e por inúmeras vezes, reservou horários para discutirmos os mais variados assuntos de Álgebra Linear.

A todos os professores que tive, desde o início da minha carreira escolar até hoje, pois cada um destes contribuiu, à sua maneira, para a formação da pessoa que sou hoje.

À Banca Examinadora deste trabalho, composta pelos professores Júnio Moreira de Alencar e Fávio França Cruz, pode ter aceito o convite para analisá-la.

A todos que, de alguma maneira contribuíram para a realização deste trabalho.

“— Qual é o melhor caminho para eu subir a montanha?

— Apenas suba, e não pense nisso!”

(Friederich Nietzsche in, ‘A Gaia Ciência’.)

“Para os precavidos não existe o azar; nada de apuros para os preparados”.

(Baltasar Gracián in, ‘A Arte da Prudência’)

Resumo

O presente trabalho, o qual foi desenvolvido por meio de pesquisa científica do tipo qualitativa, visa mostrar algumas aplicações que as matrizes possuem, tanto dentro da própria Matemática, como também em outras áreas, tendo a preocupação de esboçar os detalhes de forma acessível, evitando portanto a utilização de muitas definições e teoremas, mas sem perder o rigor matemático necessário. A pesquisa é direcionada tanto ao professor do ensino médio, para que este possa vir a aprofundar o assunto em sala de aula, bem como para o aluno de ensino médio que tenha interesse em entender a funcionalidade das matrizes na prática. Parte-se então de uma simples organização de uma lista de compras com seus respectivos preços em uma tabela, passando por problemas que envolvem rotas de voo, criptografia, crescimento populacional, dentre outros. Tais temáticas foram inseridas como exemplos e/ou comentários ao longo do desenvolvimento do texto. Uma delas recebe maior enfoque, a saber, a rotação de cônicas, a qual é discutida tanto com o uso de diagonalização de matrizes, bem como sem a utilização da mesma. Outro ponto abordado é a exponencial de matrizes, tópico este que possui inúmeras aplicações em diversos ramos do conhecimento. Por fim, é realizada a explanação sobre a relação entre números complexos e matrizes, tais como involuções, matrizes hermitianas, quatérnios até chegar ao Teorema Espectral.

Palavras-chave: Matrizes. Ensino Médio. Aplicações. Rotação de cônicas. Teorema Espectral.

Abstract

The present work aims to show the most varied applications that the matrices have, both within mathematics itself, as well as in other areas, with the concern to outline the details in a simple way, so that a high school student can read it completely, but without losing the necessary mathematical rigor. It then departs from a simple organization of a shopping list with their respective prices in a table, going through problems involving flight routes, cryptography, population growth, among others. These themes were inserted as examples and/or comments throughout the development of the text. One of them receives a greater focus, that is, the rotation of conics, which is discussed both with the use of diagonalization of matrices, as well as without the use of it. Another point addressed is the exponential of matrices, a topic that has numerous applications in various branches of knowledge. Finally, an explanation is made about the relationship between complex numbers and matrices, such as involutions, hermitian matrices, quaternions until then reaching the Spectral Theorem.

Key-words: Matrices. High School. Applications. Rotation of conics. Spectral theorem.

Sumário

Agradecimentos	5
Resumo	7
Abstract	8
Introdução	11
1 Matrizes	13
1.1 Matrizes	13
1.1.1 Adição e subtração de matrizes	16
1.1.2 Multiplicação de matrizes	18
1.1.3 Matriz Transposta	23
1.1.4 Matriz Inversa	30
1.2 Matrizes de blocos	34
1.3 Matrizes elementares	40
1.4 Algumas aplicações matriciais	52
1.4.1 Matrizes de Leslie	52
1.4.2 Criptografia	59
2 Determinantes e Sistemas Lineares	66
2.1 Determinantes	66
2.1.1 Propriedades dos determinantes	70
2.1.2 Matriz Adjunta e a Regra de Cramer	75
2.2 Sistemas Lineares	78
3 Transformações Lineares e Matrizes	87
3.1 O plano \mathbb{R}^2 e o espaço tridimensional \mathbb{R}^3	87
3.2 Transformações Lineares	104
3.2.1 Conceito e exemplos	104
3.2.2 Translação, rotação e reflexão de cônicas	114
3.2.3 Identificando formas quadráticas	123
4 Diagonalização de matrizes	130
4.1 O conceito	130
4.2 A diagonalização de matrizes na rotação de cônicas	139
4.3 Exponencial de Matriz	149
5 Sobre números complexos e quatérnios	156
5.1 Números complexos	156
5.2 Quatérnios	158

Considerações Finais	163
Referências	164
Apêndice	169

Introdução

A Álgebra Linear é, de acordo com BOURBAKI [10], um dos mais velhos ramos da Matemática e, ao mesmo tempo, um dos mais novos. Enquanto que, por um lado, conforme MEYER [46], a análise mais antiga de equações simultâneas é encontrada no livro chinês Chiu-chang Suan-shu (Nove Capítulos sobre Aritmética) por volta de 200 a.C., é afirmado por RICARDO [57] que, devido à disponibilidade de computadores e calculadoras de grande manuseio, a Álgebra Linear tem emergido como pré-requisito para muitas áreas de aplicações. Como se pode observar em [3], tais aplicações se estendem desde tópicos como Programação linear geométrica, até Tomografia Computadorizada. Dessa forma, é possível perceber o quão importante e atual é o estudo deste ramo da Matemática, sendo necessário então que seu ensino esteja sempre em aperfeiçoamento.

O processo de ensino e aprendizagem de Matemática ainda possui muitos obstáculos a serem superados. Um deles, como aponta FIORENTINI [23], é o fato do aluno não entendê-la como a escola lhe ensina e, na maioria das vezes, sentem dificuldades em utilizar o conhecimento “adquirido”, até mesmo quando aprovados na disciplina. Com base nisso, buscou-se apresentar um dos inúmeros tópicos da matemática, a saber, as matrizes, de uma maneira interdisciplinar, apontando aplicações que estas possuem em diferentes campos do conhecimento, respeitando assim a importância da interdisciplinaridade, onde esta, de acordo com TERRADAS [60], é uma metodologia de ensino inovadora e importante, não somente para a Educação Matemática, mas também para os diversos âmbitos do saber.

Um dos principais objetivos do presente trabalho é fornecer uma abordagem de tópicos matriciais, de modo que estes possam ser entendidos por alunos de Ensino Médio. Nessa perspectiva, todos os resultados apresentados foram obtidos utilizando propriedades de matrizes, as quais são apresentada no Capítulo 1, juntamente com as operações relacionadas às mesmas. Os exemplos de aplicações das matrizes aqui utilizados são intercalados à medida que é desenvolvida a teoria, a fim de alinhar a interdisciplinaridade e, conseqüentemente despertar a curiosidade do leitor quanto ao assunto, além de estabelecer uma conexão entre o conteúdo e a realidade, o que está de acordo com o pensamento de DENCKER [17], segundo o qual a interdisciplinaridade surge nos anos 70 como uma resposta às necessidades de um tratamento mais integrador da realidade.

A abordagem descrita acima também levanta a possibilidade do professor de Ensino Médio discutir tópicos relacionados às matrizes de uma forma mais dinâmica, não se atendo apenas à resolução de exercícios, mas também a situações práticas em que tais temas surgem. O mesmo se aplica ao ensino superior, pois como aponta FAVARÃO e ARAÚJO [22], os professores acabam dando atenção ao conteúdo em si e não à sua interligação com a situação da qual emerge, gerando assim a clássica dissociação entre teoria e prática. Vale ressaltar que não está sendo colocada de lado a importância da teoria, pelo contrário, as aplicações não só realçam o embasamento teórico, bem como lhes acrescenta beleza. Outrossim, principalmente para os alunos que pretendem ingressar no campo acadêmico voltado à áreas exatas do conhecimento, é de

fundamental importância que a teoria seja consolidada. Todavia, como ressalta SOARES e ALMEIDA [58], os alunos que pretendem ingressar no Ensino Superior acabam por criar expectativas que não condizem com a realidade acadêmica, gerando assim conflitos que acabam por aumentar os níveis de desistência. Por este motivo, é defendido que,

Num mundo extremamente competitivo, a universidade precisa se preocupar com o estudante universitário, promovendo condições para o seu desenvolvimento integral, tentando desenvolver suas potencialidades ao máximo para que possa atingir seu nível de excelência pessoal e estar preparado para um papel atuante na sociedade (SANTOS apud CUNHA e CARRILHO [15], p. 216).

Dessa maneira, buscou-se escrever a teoria de maneira simples e clara, porém sem fugir do rigor matemático, com exceção de raros casos de demonstrações que exigem indução matemática, as quais, quando possível, foram provadas por uma ideia intuitiva de indução, isto é, mostrando casos iniciais de modo a encontrar um padrão e, em seguida, generalizando a ideia. O motivo de não ter sido abordada a indução matemática deve-se à busca de evitar possíveis digressões, mantendo o foco sobre as matrizes e suas propriedades. Porém, a discussão de tal assunto direcionada a alunos de Ensino Médio pode ser encontrada em [16] e [51].

O Capítulo 1 é direcionado ao desenvolvimento da teoria acerca das matrizes, apontando suas diferentes propriedades e classificações. Entre elas, destacam-se duas que não são comuns ao Ensino Médio: Matrizes de blocos e Matrizes elementares. Além disso, a última seção de tal capítulo trata exclusivamente do crescimento populacional por meio de matrizes de Leslie. Já o Capítulo 2 versa a respeito dos determinantes das matrizes e das propriedades dos mesmos, onde a justificativa de algumas destas ficarão mais claras devido à matrizes elementares.

O plano e o espaço tridimensional são discutidos no Capítulo 3, porém seu principal foco são as transformações lineares, pois estas serão utilizadas para exemplificar situações em que as cônicas aparecem em posições incomuns, como por exemplo rotacionadas ou reflexionadas em torno de uma reta, além da identificação de formas quadráticas. Cabe lembrar que será admitido o conhecimento prévio das cônicas por parte do leitor, omitindo-se assim os detalhes relativos às mesmas. Mais exemplos destes podem ser encontrados em [12].

A Diagonalização de matrizes é também um tema incomum ao Ensino Médio, sendo portanto tratado no Capítulo 4, tanto em sua definição, como em sua aplicação à rotação de cônicas, fornecendo assim uma segunda maneira de realizar o processo descrito no Capítulo 3. Ademais, é realizada uma breve exposição sobre a exponencial de uma matriz, tema este também incomum ao Ensino Médio, mas que possui vasta aplicabilidade.

O Capítulo 5 é destinado à relação estabelecida entre números complexos e matrizes, mais precisamente a representação de um número complexo por uma matriz e os quatérnios, onde estes são uma extensão dos números complexos. O complemento deste Capítulo é deixado como Apêndice, onde são tratadas as matrizes hermitianas e suas propriedades, além de apresentar um teorema de bastante valor na Álgebra Linear, o Teorema Espectral.

1 Matrizes

Iniciaremos este capítulo com a definição de matrizes, conceito este que será de extrema importância para o desenvolvimento de todos os capítulos posteriores a este.

1.1 Matrizes

Definição 1.1 Dados $m, n \in \mathbb{N}$, chama-se **matriz** \mathbf{A} , toda tabela retangular \mathbf{A} , composta por números reais a_{ij} , distribuídos em m linhas e n colunas, onde $i, j \in \mathbb{N}$ são os **índices** que permitem identificar um dado elemento na matriz que está na i -ésima linha e j -ésima coluna. A representação de uma tal matriz é dada por

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

Exemplo 1.1 A fim de melhor compreender a definição acima, vejamos alguns exemplos de matrizes a seguir:

$$\mathbf{A} = \begin{bmatrix} -7 & 9 \\ 16 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \frac{1}{2} & -1 \\ -\frac{1}{2} & 1 \\ 0 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$
$$\mathbf{D} = \begin{bmatrix} 10 & 12 & -3 & 77 & 32 \\ -4 & 86 & 16 & 99 & -74 \\ 13 & 38 & -9 & 15 & -35 \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

As matrizes \mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D} e \mathbf{E} acima são do tipo 2×2 , 3×2 , 3×3 , 3×5 e 4×4 , respectivamente. Na matriz \mathbf{A} , a **entrada** a_{12} , isto é, o número que está na primeira linha e na segunda coluna, é 9, enquanto que $b_{12} = -1$ na matriz \mathbf{B} . Já na matriz \mathbf{D} , podemos observar que $d_{31} = 13$ e $d_{25} = -74$.

Vale lembrar que a igualdade entre duas matrizes $\mathbf{A} = [a_{ij}]_{m \times n}$ e $\mathbf{B} = [b_{ij}]_{m \times n}$ ocorre se, e somente se, $a_{ij} = b_{ij}$, para todo $i = 1, 2, \dots, m$ e todo $j = 1, 2, \dots, n$. Por exemplo, as matrizes

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix}$$

são diferentes, tendo em vista que $a_{11} = 1 \neq 0 = b_{11}$.

A seguir, veremos algumas matrizes que recebem nomes especiais devido às diferentes propriedades que elas podem apresentar:

a) *Matriz linha*: é toda matriz que possui n colunas e uma única linha, isto é, uma matriz da forma

$$\mathbf{A} = [a_{11} \quad a_{12} \quad \cdots \quad a_{1n}]$$

b) *Matriz coluna*: é toda matriz que possui m linhas e uma única coluna, tendo então a seguinte forma:

$$\mathbf{B} = \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{bmatrix}$$

c) *Matriz quadrada*: é toda matriz $\mathbf{A} = [a_{ij}]_{m \times n}$, tal que $m = n$, ou seja, o número de linhas é igual ao número de colunas. Dizemos neste caso que a matriz \mathbf{A} possui **ordem** n . As matrizes \mathbf{A} , \mathbf{C} e \mathbf{E} do Exemplo 1.1 são quadradas de ordem 2, 3 e 4, respectivamente. Vale lembrar que, nas matrizes quadradas $\mathbf{A} = [a_{ij}]_{n \times n}$, as entradas a_{ij} , tais que $i = j$ constituem um conjunto chamado de **diagonal principal**, ao passo que as entradas com $i + j = n + 1$ formam um conjunto denominado **diagonal secundária**.

Vejam, por exemplo, quais são as diagonais principal e secundária da matriz \mathbf{C} do Exemplo 1.1, observando que

- $a_{11} = 1$; $a_{22} = 5$; $a_{33} = 9 \Rightarrow$ Diagonal principal é dada por $\{1, 5, 9\}$.
- Como \mathbf{C} possui ordem 3, devemos ter $i + j = 4$. Assim, as entradas que atendem a essa exigência são: $a_{13} = 3$; $a_{22} = 5$; $a_{31} = 7$. Portanto, temos que a diagonal secundária é dada por $\{3, 5, 7\}$.

Agora podemos visualizar abaixo as duas diagonais citadas anteriormente:

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Diagonal principal de \mathbf{C}

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Diagonal secundária de \mathbf{C}

Vale lembrar que a soma das entradas de uma matriz quadrada \mathbf{A} é chamado de **traço** da matriz \mathbf{A} e é representado por $\text{tr}(\mathbf{A})$. Mais formalmente, dada uma matriz $\mathbf{A} = [a_{ij}]_{n \times n}$, tem-se

$$\text{tr}(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn}.$$

d) *Matriz nula*: é chamada assim, qualquer matriz $\mathbf{A} = [a_{ij}]_{m \times n}$, tal que $a_{ij} = 0$ para todo $i = 1, 2, \dots, m$ e todo $j = 1, 2, \dots, n$, o que quer dizer que todas as suas entradas são iguais a 0. A matriz \mathbf{E} do Exemplo 1.1 é um exemplo de matriz desse tipo. Representaremos por \mathbf{O} uma matriz nula de ordem $n \times n$.

e) *Matriz Identidade*: São assim denominadas todas as matrizes quadradas $\mathbf{A} = [a_{ij}]_{n \times n}$,

tais que todos os elementos da diagonal principal são iguais a 1. Observemos logo abaixo alguns exemplos de matrizes identidade:

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

Por simplicidade, nos referiremos a uma matriz identidade de ordem n por meio da expressão $\mathbf{I}_{n \times n}$, ou simplesmente por \mathbf{I} , nos casos em que não houver dúvida sobre qual é a ordem dessa matriz.

f) *Matriz diagonal*: Essa é a generalização de uma matriz identidade, de modo que ao invés de possuir todos os elementos de sua diagonal principal iguais a 1, estes são constantes reais não necessariamente todas iguais. Quanto às demais entradas (não pertencentes à diagonal principal) de uma tal matriz, são todas iguais a zero. Vejamos logo abaixo alguns exemplos de matrizes diagonais:

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \quad \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix} \quad \begin{bmatrix} a & 0 & 0 & 0 & 0 & 0 \\ 0 & b & 0 & 0 & 0 & 0 \\ 0 & 0 & c & 0 & 0 & 0 \\ 0 & 0 & 0 & d & 0 & 0 \\ 0 & 0 & 0 & 0 & e & 0 \\ 0 & 0 & 0 & 0 & 0 & f \end{bmatrix},$$

onde a, b, c, d, e e f são constantes reais quaisquer.

g) *Matriz triangular*: As matrizes triangulares podem ser classificadas como *superiores*, quando todas as entradas abaixo da diagonal principal são iguais a zero, ou *inferiores*, quando todas as entradas acima da diagonal principal são iguais a zero. Vejamos abaixo um exemplo de cada caso:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 10 \end{bmatrix}$$

Matriz triangular superior

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 10 \end{bmatrix}$$

Matriz triangular inferior

h) *Matriz de Vandermonde*¹: São matrizes $m \times n$ da forma

$$\mathbf{V} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} \end{bmatrix},$$

¹Este nome é em homenagem a Alexandre Théophile Vandermonde (1735-1796), um matemático francês que fez uma variedade de contribuições à matemática, mas que talvez seja melhor conhecido como o primeiro europeu a fornecer uma exposição lógica completa da teoria dos determinantes [46].

onde $x_i \neq x_j$, para todo $i \neq j$.

A seguir definiremos as operações de adição/subtração e multiplicação de matrizes. Quanto à divisão de matrizes, veremos posteriormente que esta não é definida, devido ao fato de que nem toda matriz possui uma *matriz inversa*.

1.1.1 Adição e subtração de matrizes

Definição 1.2. Dadas duas matrizes $\mathbf{A} = [a_{ij}]_{m \times n}$ e $\mathbf{B} = [b_{ij}]_{m \times n}$ a soma $\mathbf{A} + \mathbf{B}$ é igual a uma matriz $\mathbf{C} = [c_{ij}]_{m \times n}$ tal que

$$[a_{ij}]_{m \times n} + [b_{ij}]_{m \times n} = [c_{ij}]_{m \times n}.$$

Podemos concluir, a partir da definição acima, que a soma de duas (ou mais) matrizes é ainda uma matriz, a qual está definida apenas para o caso em que \mathbf{A} e \mathbf{B} possuem o mesmo número de linhas e o mesmo número de colunas.

Exemplo 1.2. Ana, Bruna e Carla fizeram compras de arroz, feijão e carne durante o mês de janeiro nos estabelecimentos A, B e C, respectivamente. A tabela abaixo mostra as quantidades totais de cada um desses alimentos comprados por elas:

	Arroz	Feijão	Carne
Ana	20 kg	10 kg	10 kg
Bruna	25 kg	12 kg	13 kg
Carla	30 kg	18 kg	15 kg

No mês seguinte, elas realizaram compras dos mesmos alimentos nos mesmos estabelecimentos, porém em quantidades diferentes, conforme mostra o próximo quadro:

	Arroz	Feijão	Carne
Ana	23 kg	12 kg	8 kg
Bruna	20 kg	13 kg	11 kg
Carla	27 kg	15 kg	17 kg

Quanto aos preços, em reais, de cada quilo de alimento nos respectivos estabelecimentos, estes estão organizados no seguinte quadro:

	A	B	C
Arroz	2,50	2,30	3,20
Feijão	4,30	4,00	4,80
Carne	25,30	28,90	36,60

Para saber o total de quilos de cada alimento que cada uma delas comprou durante esses dois meses, podemos representar as compras realizadas em janeiro e fevereiro pelas respectivas matrizes \mathbf{J} e \mathbf{F} abaixo:

$$\mathbf{J} = \begin{bmatrix} 20 & 10 & 10 \\ 25 & 12 & 13 \\ 30 & 18 & 15 \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} 23 & 12 & 8 \\ 20 & 13 & 11 \\ 27 & 15 & 17 \end{bmatrix}$$

Agora, realizando a soma das duas últimas matrizes, decorre

$$\begin{aligned}
 \mathbf{J} + \mathbf{F} &= \begin{bmatrix} 20 & 10 & 10 \\ 25 & 12 & 13 \\ 30 & 18 & 15 \end{bmatrix} + \begin{bmatrix} 23 & 12 & 8 \\ 20 & 13 & 11 \\ 27 & 15 & 17 \end{bmatrix} \\
 &= \begin{bmatrix} 20 + 23 & 10 + 12 & 10 + 8 \\ 25 + 20 & 12 + 13 & 13 + 11 \\ 30 + 27 & 18 + 15 & 15 + 17 \end{bmatrix} \\
 &= \begin{bmatrix} 43 & 22 & 18 \\ 45 & 25 & 24 \\ 57 & 33 & 32 \end{bmatrix}
 \end{aligned}$$

Logo, podemos concluir que, durante esses dois meses

- Ana comprou um total de 43 kg de arroz, 22 kg de feijão e 18 kg de carne.
- Bruna comprou um total de 45 kg de arroz, 25 kg de feijão e 24 kg de carne.
- Carla comprou um total de 57 kg de arroz, 33 kg de feijão e 32 kg de carne.

A subtração de matrizes é definida de forma análoga, de modo que, dadas duas matrizes $\mathbf{A} = [a_{ij}]_{m \times n}$ e $\mathbf{B} = [b_{ij}]_{m \times n}$ então $\mathbf{A} - \mathbf{B}$ é igual a uma matriz $\mathbf{D} = [d_{ij}]_{m \times n}$, tal que

$$[a_{ij}]_{m \times n} - [b_{ij}]_{m \times n} = [d_{ij}]_{m \times n}.$$

Com estas definições, passemos às propriedades da soma de matrizes. As propriedades para a subtração são análogas. Consideremos então as matrizes $\mathbf{A} = [a_{ij}]_{m \times n}$, $\mathbf{B} = [b_{ij}]_{m \times n}$ e $\mathbf{C} = [c_{ij}]_{m \times n}$. São válidas as seguintes afirmações:

a) Comutatividade: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.

Sabemos que, pela comutatividade dos números reais, se $a, b \in \mathbb{R}$, então vale $a + b = b + a$. Daí, como as entradas das matrizes adotadas até aqui são números reais, temos que $a_{ij} + b_{ij} = b_{ij} + a_{ij}$, e portanto

$$\mathbf{A} + \mathbf{B} = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = \mathbf{B} + \mathbf{A}$$

b) Associatividade: $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$.

Segue diretamente da associatividade dos números reais, isto é, dados três números reais a, b e c , tem-se que $(a + b) + c = a + (b + c)$. Dessa forma, decorre que

$$\begin{aligned}
 (\mathbf{A} + \mathbf{B}) + \mathbf{C} &= ([a_{ij}] + [b_{ij}]) + [c_{ij}] \\
 &= [a_{ij} + b_{ij}] + [c_{ij}] \\
 &= [(a_{ij} + b_{ij}) + c_{ij}] \\
 &= [a_{ij} + (b_{ij} + c_{ij})] \\
 &= [a_{ij}] + [b_{ij} + c_{ij}] \\
 &= [a_{ij}] + ([b_{ij}] + [c_{ij}]) \\
 &= \mathbf{A} + (\mathbf{B} + \mathbf{C})
 \end{aligned}$$

c) Elemento neutro: Para toda matriz $\mathbf{A} = [a_{ij}]_{m \times n}$, existe uma matriz $\mathbf{O} = [o_{ij}]_{mn}$, tal que $\mathbf{A} + \mathbf{O} = \mathbf{A}$.

Povemos inicialmente a existência. se ocorre $\mathbf{A} + \mathbf{O} = \mathbf{A}$, então segue que

$$[a_{ij} + o_{ij}] = [a_{ij}] \Rightarrow a_{ij} + o_{ij} = a_{ij} \Rightarrow o_{ij} = a_{ij} - a_{ij} = 0.$$

Como i e j são arbitrários, decorre que $o_{ij} = 0$ para todo i e todo j . Logo a matriz \mathbf{O} existe e é uma matriz nula com o mesmo número de linhas e colunas de \mathbf{A} . Quanto à unicidade, suponhamos que exista uma outra matriz $\mathbf{O}' = [o'_{ij}]$, tal que $\mathbf{A} + \mathbf{O}' = \mathbf{A}$. Decorre a partir daí que

$$\mathbf{A} + \mathbf{O} = \mathbf{A} + \mathbf{O}' \Rightarrow a_{ij} + o_{ij} = a_{ij} + o'_{ij} \Rightarrow o_{ij} = o'_{ij} \Rightarrow \mathbf{O} = \mathbf{O}'.$$

Portanto, o elemento neutro de qualquer matriz é único.

d) Elemento oposto: Para toda matriz \mathbf{A} , existe uma única matriz \mathbf{A}' , tal que $\mathbf{A} + \mathbf{A}' = \mathbf{O}$.

De fato, se $\mathbf{A} + \mathbf{A}' = \mathbf{O}$, então

$$a_{ij} + a'_{ij} = 0 \Rightarrow a'_{ij} = -a_{ij},$$

ou seja, a matriz \mathbf{A}' é formada pelo oposto de cada entrada da matriz \mathbf{A} . Quanto à sua unicidade, esta é provada de modo análogo à unicidade do elemento neutro. Por simplicidade notacional, nos referiremos à **matriz oposta** \mathbf{A}' da matriz \mathbf{A} apenas por $-\mathbf{A}$.

1.1.2 Multiplicação de matrizes

Antes de fornecermos a definição de multiplicação de matrizes, utilizemos o Exemplo 1.2 para auxiliar na compreensão de como as matrizes são multiplicadas.

Se quiséssemos saber quanto Ana gastou em janeiro com as compras de arroz, feijão e carne, deveremos multiplicar cada quantidade por seu respectivo preço. Assim,

$$20 \cdot 2,50 + 10 \cdot 4,30 + 10 \cdot 25,30 = 346 \quad (*)$$

o que mostra que ela gastou R\$ 346,00.

Agora representando as quantidades de alimento compradas por ela em janeiro pela matriz linha $\mathbf{Q}_A = [20 \quad 10 \quad 10]$, os preços desses alimentos (por quilo) pela matriz

coluna $\mathbf{P}_A = \begin{bmatrix} 2,50 \\ 4,30 \\ 25,30 \end{bmatrix}$ e analisando (*), podemos observar que a entrada q_{11} de \mathbf{Q}_A

multiplica a entrada p_{11} de \mathbf{P}_A e da mesma forma q_{12} multiplica p_{21} e q_{13} multiplica p_{31} , o que nos aponta para uma soma de produtos da forma $q_{ij} \cdot p_{ji}$, isto é,

$$q_{11} \cdot p_{11} + q_{12} \cdot p_{21} + q_{13} \cdot p_{31}.$$

Notemos que se alguma das matrizes \mathbf{Q}_A ou \mathbf{P}_A possuíse mais entradas que a outra, a expressão (*) não poderia ser calculada.

Podemos assim fornecer a definição de multiplicação de matrizes da seguinte forma

Definição 1.3. Dadas duas matrizes $\mathbf{A} = [a_{ij}]_{m \times n}$ e $\mathbf{B} = [b_{jk}]_{n \times p}$, tem-se que o produto $\mathbf{A} \cdot \mathbf{B}$ é igual a uma matriz $\mathbf{C} = [c_{ik}]_{m \times p}$, tal que

$$c_{ik} = a_{i1} \cdot b_{1k} + a_{i2} \cdot b_{2k} + a_{i3} \cdot b_{3k} + \dots + a_{in} \cdot b_{nk}$$

Exemplo 1.3. O produto entre as matrizes \mathbf{Q}_A e \mathbf{P}_A , de acordo com a definição acima é dado por

$$\begin{aligned} \mathbf{Q}_A \cdot \mathbf{P}_A &= \begin{bmatrix} 20 & 10 & 10 \end{bmatrix} \cdot \begin{bmatrix} 2,50 \\ 4,30 \\ 25,30 \end{bmatrix} \\ &= \begin{bmatrix} 20 \cdot 2,50 + 10 \cdot 4,30 + 10 \cdot 25,30 \end{bmatrix} \\ &= \begin{bmatrix} 346 \end{bmatrix} \end{aligned}$$

Note que a expressão obtida na segunda igualdade coincide com (*).

Exemplo 1.4. Vamos utilizar a definição acima para calcular o produto entre as matrizes \mathbf{J} e a matriz dos preços dos alimentos do Exemplo 1.2, a qual representaremos por

$$\mathbf{Q} = \begin{bmatrix} 2,5 & 2,3 & 3,2 \\ 4,3 & 4 & 4,8 \\ 25,3 & 28,9 & 36,6 \end{bmatrix}.$$

Daí,

$$\begin{aligned} \mathbf{J} \cdot \mathbf{Q} &= \begin{bmatrix} 20 & 10 & 10 \\ 25 & 12 & 13 \\ 30 & 18 & 15 \end{bmatrix} \cdot \begin{bmatrix} 2,5 & 2,3 & 3,2 \\ 4,3 & 4 & 4,8 \\ 25,3 & 28,9 & 36,6 \end{bmatrix} \\ &= \begin{bmatrix} 50 + 43 + 253 & 46 + 40 + 289 & 64 + 48 + 366 \\ 62,5 + 51,6 + 328,9 & 57,5 + 48 + 375,7 & 80 + 57,6 + 475,8 \\ 75 + 77,4 + 379,5 & 69 + 72 + 433,5 & 96 + 86,4 + 549 \end{bmatrix} \\ &= \begin{bmatrix} 346 & 375 & 478 \\ 443 & 481,2 & 613,4 \\ 531,9 & 574,5 & 731,4 \end{bmatrix} \end{aligned}$$

A interpretação da matriz $\mathbf{J} \cdot \mathbf{Q}$ nos fornece tanto o valor que Ana, Bruna e Carla pagaram nas compras de janeiro (valores da diagonal principal), bem como fornece os valores de suas compras, caso estas tivessem sido realizadas nos outros estabelecimentos. Por exemplo, Bruna pagou R\$ 481,20 comprando no estabelecimento 2, ao passo que pagaria R\$ 443,00, caso tivesse comprado no estabelecimento 1.

A próxima definição nos permitirá conhecer as propriedades da multiplicação em relação às matrizes.

Definição 1.4. Dado um número real c e uma matriz $\mathbf{A} = [a_{ij}]_{m \times n}$, tem-se que o produto $c \cdot \mathbf{A}$ é uma matriz, cujas entradas são iguais a $c \cdot a_{ij}$. Em termos

$$c \cdot \mathbf{A} = c \cdot [a_{ij}] = [c \cdot a_{ij}].$$

Exemplo 1.5. Suponhamos que Ana, Bruna e Carla comprem durante três meses seguidos a mesma quantidade de alimentos que compraram em fevereiro. Quantos quilos de alimento terão comprado no total, ao fim desse período?

Basta multiplicar a matriz \mathbf{F} do Exemplo 1.2. por 3, de modo que

$$3 \cdot \mathbf{F} = 3 \cdot \begin{bmatrix} 23 & 12 & 8 \\ 20 & 13 & 11 \\ 27 & 15 & 17 \end{bmatrix} = \begin{bmatrix} 3 \cdot 23 & 3 \cdot 12 & 3 \cdot 8 \\ 3 \cdot 20 & 3 \cdot 13 & 3 \cdot 11 \\ 3 \cdot 27 & 3 \cdot 15 & 3 \cdot 17 \end{bmatrix} = \begin{bmatrix} 69 & 36 & 24 \\ 60 & 39 & 33 \\ 81 & 45 & 51 \end{bmatrix}.$$

Logo, temos que

- Ana comprou 69 kg de arroz, 36 kg de feijão e 24 kg de carne.
- Bruna comprou 60 kg de arroz, 39 kg de feijão e 33 kg de carne.
- Carla comprou 81 kg de arroz, 45 kg de feijão e 51 kg de carne.

Passemos desde já às propriedades da multiplicação de escalar por matriz, considerando para tanto, dois números reais c, d e duas matrizes $\mathbf{A} = [a_{ij}]_{m \times n}$ e $\mathbf{B} = [b_{ij}]_{m \times n}$. Assim

a) Associatividade: $c \cdot (d \cdot \mathbf{A}) = (c \cdot d) \cdot \mathbf{A}$

Basta aplicar a associatividade dos números reais e a Definição 1.4, de modo que

$$c \cdot (d \cdot \mathbf{A}) = [c \cdot (d \cdot a_{ij})] = [(c \cdot d) \cdot a_{ij}] = (c \cdot d) \cdot [a_{ij}] = (c \cdot d) \cdot \mathbf{A}$$

b) Distributividade:

- $(c + d) \cdot \mathbf{A} = c \cdot \mathbf{A} + d \cdot \mathbf{A}$

Pela distributividade de números reais em relação à soma, temos que:

$$\begin{aligned} (c + d) \cdot \mathbf{A} &= [(c + d)a_{ij}] \\ &= [c \cdot a_{ij} + d \cdot a_{ij}] \\ &= [c \cdot a_{ij}] + [d \cdot a_{ij}] \\ &= c \cdot [a_{ij}] + d \cdot [a_{ij}] \\ &= c \cdot \mathbf{A} + d \cdot \mathbf{A} \end{aligned}$$

- $c \cdot (\mathbf{A} + \mathbf{B}) = c \cdot \mathbf{A} + c \cdot \mathbf{B}$

Novamente iremos aplicar a distributividade dos números reais em relação à soma, de tal modo que

$$\begin{aligned} c \cdot (\mathbf{A} + \mathbf{B}) &= c \cdot ([a_{ij}] + [b_{ij}]) \\ &= c \cdot [a_{ij} + b_{ij}] \\ &= [c \cdot (a_{ij} + b_{ij})] \\ &= [c \cdot a_{ij} + c \cdot b_{ij}] \\ &= [c \cdot a_{ij}] + [c \cdot b_{ij}] \\ &= c \cdot [a_{ij}] + c \cdot [b_{ij}] \\ &= c \cdot \mathbf{A} + c \cdot \mathbf{B} \end{aligned}$$

A seguir, listamos as propriedades principais de multiplicação entre matrizes, as quais não serão provadas, uma vez que suas demonstrações decorrem da Definição 1.3 e das propriedades de produto de um escalar por uma matriz. Além disso, a partir de agora suprimiremos o ponto “.” de multiplicação, de tal modo que, por exemplo, uma expressão que aparecia da forma $\mathbf{A} \cdot \mathbf{B}$ agora será escrita apenas como \mathbf{AB} .

a) Associatividade: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$

b) Distributividade à direita em relação à soma: $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$

c) Distributividade à esquerda em relação à soma: $\mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}$

d) Dada uma constante real c , tem-se que

$$(c\mathbf{A})\mathbf{B} = \mathbf{A}(c\mathbf{B}) = c(\mathbf{AB})$$

e) Potência de uma matriz: $\mathbf{A}^k = \underbrace{\mathbf{A} \cdot \mathbf{A} \cdot \mathbf{A} \cdots \mathbf{A}}_{k \text{ termos}}, k \in \mathbb{N}$.

Em relação aos itens b) e c) logo acima, cabe a seguinte pergunta: “a multiplicação de matrizes não é comutativa assim como a multiplicação de números reais?” A resposta para essa pergunta é: talvez. Consideremos, por exemplo, as matrizes a seguir:

$$\mathbf{X} = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix} \qquad \mathbf{Y} = \begin{bmatrix} 5 & 1 \\ 3 & 4 \end{bmatrix}$$

Pode-se facilmente verificar que

$$\mathbf{XY} = \begin{bmatrix} 14 & 13 \\ 6 & 8 \end{bmatrix} \neq \begin{bmatrix} 5 & 17 \\ 3 & 17 \end{bmatrix} = \mathbf{YX}$$

Por outro lado, existem casos em que a comutatividade é válida, como por exemplo, na multiplicação de uma matriz quadrada $\mathbf{A} = [a_{ij}]_{n \times n}$ pela matriz identidade $\mathbf{I}_{n \times n}$. Temos abaixo um caso particular de tal fato:

$$\mathbf{XI} = \begin{bmatrix} 14 & 13 \\ 6 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 13 \\ 6 & 8 \end{bmatrix}$$

e

$$\mathbf{IX} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 14 & 13 \\ 6 & 8 \end{bmatrix} = \begin{bmatrix} 14 & 13 \\ 6 & 8 \end{bmatrix}$$

Exemplo 1.6. Algumas matrizes apresentam características interessantes quando multiplicadas por si mesma. Uma delas é a *matriz idempotente*, isto é, uma matriz quadrada \mathbf{A} de ordem n , tal que, $\mathbf{A}^2 = \mathbf{A}$. O exemplo mais comum de tais matrizes é \mathbf{I} , porém podemos listar outras a seguir

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Existem também matrizes quadradas \mathbf{A} , tais que $\mathbf{A}^2 = \mathbf{I}$. Tais matrizes são denominadas *involuções*. As matrizes abaixo são exemplos de tais matrizes

$$\begin{bmatrix} \cos \theta & \text{sen } \theta \\ \text{sen } \theta & -\cos \theta \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Temos ainda as matrizes *nilpotentes*, onde estas, por sua vez, são matrizes quadradas \mathbf{A} com $\mathbf{A}^k = \mathbf{O}$, onde k é algum número natural. Temos, por exemplo, que

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -3 \\ -5 & -2 & -3 \\ 2 & 1 & 3 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

são matrizes nilpotentes.

Exemplo 1.7. Desde que $\mathbf{A} = [a_{ij}]_{n \times n}$ seja uma involução, as matrizes

$$\mathbf{B} = \frac{1}{2}(\mathbf{I} + \mathbf{A}) \quad \text{e} \quad \mathbf{C} = \frac{1}{2}(\mathbf{I} - \mathbf{A})$$

são idempotentes e $\mathbf{BC} = \mathbf{O}$. De fato, verifiquemos a idempotência de B , sendo a de C verificada de modo análogo. Temos assim que

$$\begin{aligned} \mathbf{B}^2 &= \frac{1}{4}(\mathbf{I} + \mathbf{A})^2 \\ &= \frac{1}{4}(\mathbf{I}^2 + \mathbf{A}^2 + 2\mathbf{AI}) \\ &= \frac{1}{4}(2\mathbf{I} + 2\mathbf{A}) \\ &= \frac{1}{2}(\mathbf{I} + \mathbf{A}), \end{aligned}$$

onde na terceira igualdade acima foi utilizado o fato de $\mathbf{I}^2 = \mathbf{I}$, $\mathbf{A}^2 = \mathbf{I}$ e $\mathbf{AI} = \mathbf{A}$. Quanto a \mathbf{BC} , temos

$$\begin{aligned} \mathbf{BC} &= \frac{1}{2}(\mathbf{I} + \mathbf{A})\frac{1}{2}(\mathbf{I} - \mathbf{A}) \\ &= \frac{1}{4}(\mathbf{I}^2 - \mathbf{A}^2) \\ &= \frac{1}{4}(\mathbf{I}^2 - \mathbf{I}^2) \\ &= \mathbf{O}. \end{aligned}$$

Exemplo 1.8. Uma área equivalente a 85 km² de certo país é ocupada de forma comercial (62%), industrial (23%) e residencial (15%). Vamos assumir que as probabilidades de transição de um setor para o outro a cada 5 anos são dadas pela matriz

$$\mathbf{T} = \begin{matrix} & \begin{matrix} C & I & R \end{matrix} \\ \begin{matrix} C \\ I \\ R \end{matrix} & \begin{bmatrix} 0.72 & 0.06 & 0.03 \\ 0.19 & 0.91 & 0.05 \\ 0.09 & 0.03 & 0.92 \end{bmatrix} \end{matrix}$$

Note que as entradas de \mathbf{T} indicam a probabilidade de uma determinada passar a ser utilizada por outro setor ou não ser alterada, num período de 5 anos. Por exemplo, a entrada t_{11} aponta que há 72% de chance de uma área comercial continuar sendo comercial após 5 anos, enquanto que t_{23} indica que existe apenas 5% de chance de uma área industrial passar a ser utilizada de forma residencial após o mesmo período. Note também que a soma das entradas de qualquer uma das colunas de \mathbf{T} é igual a 1 e nenhuma das entradas é negativa. Uma tal matriz é denominada *matriz estocástica*. Daí, seja

$$\mathbf{A} = \begin{bmatrix} 62 \\ 23 \\ 15 \end{bmatrix}$$

a matriz das porcentagens de cada área ocupada. Para então determinar quais serão as porcentagens de ocupação após cinco anos, basta então calcular \mathbf{TA} , isto é

$$\mathbf{TA} = \begin{bmatrix} 0.72 & 0.06 & 0.03 \\ 0.19 & 0.91 & 0.05 \\ 0.09 & 0.03 & 0.92 \end{bmatrix} \begin{bmatrix} 62 \\ 23 \\ 15 \end{bmatrix} = \begin{bmatrix} 46.47 \\ 33.46 \\ 20.07 \end{bmatrix}.$$

Temos, portanto, que após 5 anos, as áreas de ocupação comercial, industrial e residencial serão, respectivamente, 46,47%, 33,46% e 20,07%. Como a matriz \mathbf{T} é utilizada em cada intervalo de 5 anos, podemos determinar as porcentagens de ocupação em cada um desses intervalos, apenas utilizando a matriz de ocupações atual e a matriz \mathbf{T} . Por exemplo, se quiséssemos saber qual seriam as porcentagens de ocupação em 10 anos, bastaria calcular $\mathbf{T} \cdot \mathbf{TA} = \mathbf{T}^2 \mathbf{A}$, em 15 anos, $\mathbf{T}^3 \mathbf{A}$, e assim sucessivamente. Um tal processo é chamado *processo de Markov*². Em 50 anos, isto é, 10 intervalos de 5 anos, por exemplo, as áreas de ocupação seriam dadas pela matriz

$$\mathbf{T}^{10} \mathbf{A} = \begin{bmatrix} 0.17 & 0.16 & 0.12 \\ 0.52 & 0.60 & 0.35 \\ 0.31 & 0.24 & 0.53 \end{bmatrix} \begin{bmatrix} 62 \\ 23 \\ 15 \end{bmatrix} = \begin{bmatrix} 16.02 \\ 51.29 \\ 32.69 \end{bmatrix},$$

o que nos permite concluir que a área predominante após esse período será a industrial, com 51,29% do território. Perceba que a soma das entradas da matriz $\mathbf{T}^{10} \mathbf{A}$ continua sendo igual a 100%.

1.1.3 Matriz Transposta

Apresentaremos agora um dos tipos de matrizes que possuem muitas aplicações, tanto na Matemática, como em outros campos de estudo.

Definição 1.5. Dada uma matriz $\mathbf{A} = [a_{ij}]_{m \times n}$, diz-se que a *matriz transposta* \mathbf{A}^T de \mathbf{A} é tal que

$$\mathbf{A}^T = [\bar{a}_{ji}]_{n \times m},$$

tal que $\bar{a}_{ji} = a_{ij}$, para todo i e j .

²Andrei Andreevich Markov foi um talentoso matemático russo, reconhecido pelos seus trabalhos em Teoria dos Números, Análise e Teoria das Probabilidades [8].

De modo resumido, a trasposta de uma matriz tem suas linhas transformadas em colunas, o que por sua vez faz com que as colunas passem a ser linhas. Vejamos a seguir alguns exemplos esclarecedores a respeito deste fato:

- $\mathbf{A} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \Rightarrow \mathbf{A}^T = [1 \ 2 \ 4]$
- $\mathbf{B} = \begin{bmatrix} 10 & 17 \\ -5 & 12 \end{bmatrix} \Rightarrow \mathbf{B}^T = \begin{bmatrix} 10 & -5 \\ 17 & 12 \end{bmatrix}$
- $\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 4 & 9 & 16 \end{bmatrix} \Rightarrow \mathbf{C}^T = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}$
- $\mathbf{D} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \end{bmatrix} \Rightarrow \mathbf{D}^T = \begin{bmatrix} a & e \\ b & f \\ c & g \\ d & h \end{bmatrix}$

Observando o segundo e o terceiro exemplo acima, pode-se imaginar que a diagonal principal de uma matriz quadrada e a diagonal principal de sua transposta são iguais. De fato, isso ocorre diretamente da definição, uma vez que $a_{ii} = \bar{a}_{ii}$, quando $i = j$.

Vamos agora tomar conhecimento das principais propriedades das matrizes transpostas. Para tanto, consideremos duas matrizes \mathbf{A} , \mathbf{B} e uma constante real arbitrária c . Daí:

- $(\mathbf{A}^T)^T = \mathbf{A}$

Decorre diretamente da definição, pois se a_{ij} são as entradas de \mathbf{A} , então a_{ji} serão as entradas de \mathbf{A}^T e, portanto, a_{ij} serão as entradas de $(\mathbf{A}^T)^T$.

- $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$

Tomemos, $\mathbf{A} + \mathbf{B} = \mathbf{C} = [c_{ij}]_{m \times n}$ e $(\mathbf{A} + \mathbf{B})^T = \mathbf{C}^T = [\bar{c}_{ji}]_{n \times m}$. Daí, como $\mathbf{A}^T = [\bar{a}_{ji}]_{n \times m}$ e $\mathbf{B}^T = [\bar{b}_{ji}]_{n \times m}$, segue que

$$\bar{c}_{ji} = c_{ij} = a_{ij} + b_{ij} = \bar{a}_{ji} + \bar{b}_{ji},$$

e portanto, temos $\mathbf{C}^T = \mathbf{A}^T + \mathbf{B}^T$.

- $(c\mathbf{A})^T = c\mathbf{A}^T$

Seja $(c\mathbf{A})^T = [\bar{p}_{ji}]_{n \times m}$. Daí, vem que

$$\bar{p}_{ji} = ca_{ij} = c\bar{a}_{ji},$$

donde segue o resultado.

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

Tomemos $\mathbf{AB} = \mathbf{C} = [c_{ik}]_{m \times p}$ e $(\mathbf{AB})^T = \mathbf{C}^T = [\bar{c}_{ki}]_{p \times m}$. Assim, decorre da Definição 1.3 que

$$\begin{aligned}\bar{c}_{ki} &= c_{ik} = a_{i1}b_{1k} + \cdots + a_{in}b_{nk} \\ &= b_{1k}a_{i1} + \cdots + b_{nk}a_{in} \\ &= \bar{b}_{k1}\bar{a}_{1i} + \cdots + \bar{b}_{kn}\bar{a}_{ni},\end{aligned}$$

o que prova o resultado desejado.

Quanto a esta última propriedade, vejamos um exemplo numérico para dar ênfase à compreensão. Consideremos assim as matrizes

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 0 & 2 & -7 \\ 4 & 11 & 0 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 8 & -9 \\ 3 & 4 \\ 5 & 3 \end{bmatrix}$$

Daí, temos que

$$\mathbf{AB} = \begin{bmatrix} 42 & 18 \\ -29 & -13 \\ 65 & 8 \end{bmatrix} \Rightarrow (\mathbf{AB})^T = \begin{bmatrix} 42 & -29 & 65 \\ 18 & -13 & 8 \end{bmatrix}.$$

Por outro lado, temos que

$$\mathbf{A}^T = \begin{bmatrix} 1 & 0 & 4 \\ 3 & 2 & 11 \\ 5 & -7 & 0 \end{bmatrix} \qquad \mathbf{B}^T = \begin{bmatrix} 8 & 3 & 5 \\ -9 & 4 & 3 \end{bmatrix}$$

que por sua vez, nos fornece

$$\mathbf{B}^T \mathbf{A}^T = \begin{bmatrix} 42 & -29 & 65 \\ 18 & -13 & 8 \end{bmatrix}$$

Voltando agora à observação feita sobre a transposta de uma matriz quadrada preservar a diagonal principal, pode-se então indagar: “existem matrizes cuja transposta é exatamente igual à matriz original?” A resposta para essa pergunta é “sim”, e tais matrizes são chamadas de *simétricas*.

Definição 1.6. *Seja \mathbf{A} uma matriz quadrada de ordem n . Diz-se que \mathbf{A} é uma **matriz simétrica** quando $\mathbf{A}^T = \mathbf{A}$, e nesse caso tem-se $a_{ij} = a_{ji}$, para todo i e j . Já se for $\mathbf{A}^T = -\mathbf{A}$, então diz-se que \mathbf{A} é **antissimétrica**, e portanto $a_{ij} = -a_{ji}$.*

Vejamos alguns exemplos de matrizes simétricas e antissimétricas.

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \qquad \begin{bmatrix} -17 & 26 & 12 & 74 \\ 26 & 34 & 3 & 17 \\ 12 & 3 & 8 & 13 \\ 74 & 17 & 13 & 10 \end{bmatrix} \qquad \begin{bmatrix} a & b & c & d & e \\ b & l & m & n & f \\ c & m & k & o & g \\ d & n & o & j & h \\ e & f & g & h & i \end{bmatrix}$$

Matrizes simétricas

$$\begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & 5 \\ -3 & -5 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & -26 & -12 & -74 \\ 26 & 0 & -3 & -17 \\ 12 & 3 & 0 & -13 \\ 74 & 17 & 13 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & b & c & d & e \\ -b & 0 & m & n & f \\ -c & -m & 0 & o & g \\ -d & -n & -o & 0 & h \\ -e & -f & -g & -h & 0 \end{bmatrix}$$

Matrizes antissimétricas

Note que a diagonal principal de uma matriz antissimétrica de números reais possui todas suas entradas iguais a zero. Tal fato deriva diretamente da definição, pois quando $i = j$, temos $a_{ii} = -a_{ii} \Rightarrow 2a_{ii} = 0 \Rightarrow a_{ii} = 0$.

Teorema 1.1. *Toda matriz quadrada \mathbf{A} pode ser escrita como soma de uma matriz simétrica \mathbf{S} com uma matriz antissimétrica \mathbf{S}' .*

Demonstração: Provemos, inicialmente, que as matrizes $\mathbf{A} + \mathbf{A}^T$ e $\mathbf{A} - \mathbf{A}^T$ são simétrica e antissimétrica, respectivamente. Com efeito, temos que

$$(\mathbf{A} + \mathbf{A}^T)^T = \mathbf{A}^T + (\mathbf{A}^T)^T = \mathbf{A}^T + \mathbf{A} = \mathbf{A} + \mathbf{A}^T,$$

ou seja, $\mathbf{A} + \mathbf{A}^T$ é simétrica de acordo com a definição. A verificação para a antissimetria de $\mathbf{A} - \mathbf{A}^T$ é realizada de modo análogo. Segue então que

$$2\mathbf{A} = \mathbf{A} + \mathbf{A} = \mathbf{A} + \mathbf{A}^T - \mathbf{A}^T + \mathbf{A} = (\mathbf{A} + \mathbf{A}^T) + (\mathbf{A} - \mathbf{A}^T) = \mathbf{S} + \mathbf{S}',$$

o que equivale a

$$\mathbf{A} = \frac{1}{2}(\mathbf{S} + \mathbf{S}'),$$

onde $\mathbf{S} = \mathbf{A} + \mathbf{A}^T$ e $\mathbf{S}' = \mathbf{A} - \mathbf{A}^T$.

Exemplo 1.9. Uma certa companhia aérea faz voos para as cidades A , B e C , de acordo com o seguinte diagrama:

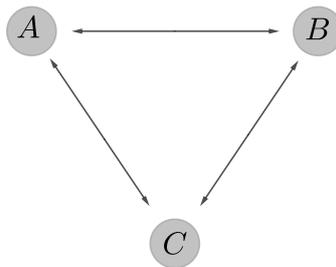


Figura 1.1: Rotas de viagem ABC

Fonte: O autor

As setas \longleftrightarrow indicam que há voos de uma cidade para a outra e vice-versa. Podemos assim criar uma matriz $\mathbf{V} = [v_{ij}]_{3 \times 3}$, seguindo a condição de que $v_{ij} = 1$, se houverem voos da cidade i para a cidade j e $v_{ij} = 0$, caso contrário. Temos assim que

$$\mathbf{V} = \begin{matrix} & \begin{matrix} A & B & C \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

Daí, pode-se observar que as entradas de \mathbf{V} representam a quantidade de maneiras de chegar de uma cidade a outra com apenas um voo. Pode-se notar também que a diagonal principal de \mathbf{V} possui todas as entradas iguais a zero porque de acordo com a Figura 1.1, não são permitidos voos de uma cidade para ela mesma. Ora, mas isto pode ser constatado apenas olhando o diagrama da Figura 1.1. Suponhamos então que a rota $A \longleftrightarrow C$ esteja sob forte tempestade. Quantos voos serão necessárias para chegar em C , partindo de A ? Nesse caso, obviamente serão realizados dois voos, sendo um de A para B e outro de B para C . Até então, não haveria necessidade nenhuma de construir a matriz \mathbf{V} . Porém, sabe-se que as rotas de avião não funcionam apenas em três cidades, mas sim em várias, inclusive de países diferentes. Dessa forma, consideremos agora que essa mesma companhia aérea atenda a cinco cidades A_1, A_2, A_3, A_4 e A_5 , conforme o diagrama a seguir:

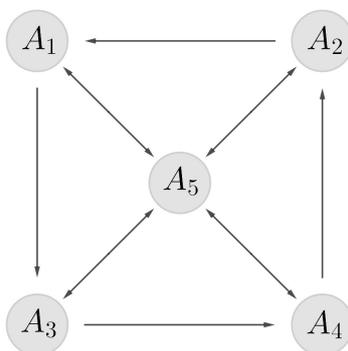


Figura 1.2: Rotas de viagem ABCDE

Fonte: O autor

Vejamos que neste caso, há mais cidades e mais rotas, logo existem mais combinações de voos partindo de uma cidade para outra. É possível determinar visualmente quantas rotas de 2 voos existem para chegar em A_1 , partindo de A_3 , ou quantas rotas de 3 voos permitem partir de A_4 e chegar em A_1 . Porém, pensando matematicamente, vamos criar, de modo análogo à matriz \mathbf{V} , uma matriz $\mathbf{W} = [w_{ij}]_{5 \times 5}$, de modo que

$$\mathbf{W} = \begin{matrix} & \begin{matrix} A_1 & A_2 & A_3 & A_4 & A_5 \end{matrix} \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{matrix} & \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

Agora observemos um fato interessante: note que se multiplicarmos, por exemplo, w_{21} por w_{13} , obtemos 1 como resultado, isto é, existe uma rota de dois voos que vai de A_2 para A_3 passando por A_1 . Já os produtos $w_{22}w_{23}$, $w_{23}w_{33}$ e $w_{24}w_{43}$ são todos iguais a zero, isto porque não existem rotas de dois voos indo de A_2 para A_3 passando por A_2 , nem de A_2 para A_3 passando por A_3 (nesse caso, se fosse possível, seria um voo e não dois como desejado), nem de A_2 para A_3 passando por A_4 . Já $w_{25}w_{53} = 1$, pois existe uma rota de dois voos que vai de A_2 para A_3 passando por A_5 . Generalizando então esta ideia, temos que o produto $w_{ij}w_{jk}$ nos fornece o número de rotas de dois voos que vão

de uma cidade A_i para uma cidade A_k passando pela cidade A_j . Logo a soma desses produtos nos fornece o número total de rotas de dois voos que vão de uma cidade A_i para uma cidade A_j . Com sorte, essa ideia coincide com a multiplicação de matrizes (e.g. $w_{21}w_{13} + w_{22}w_{23} + w_{23}w_{33} + w_{24}w_{43} + w_{25}w_{53} = 2$ é o número total de rotas de dois voos que vão de A_2 para A_3). Temos assim que a matriz simétrica

$$\mathbf{W}^2 = \begin{bmatrix} 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 4 \end{bmatrix}$$

fornece o número exato de todas as rotas de dois voos que vão de uma cidade A_i para uma cidade A_j , onde i e j variam de 1 a 5. Seguindo esse pensamento, temos que a matriz (também simétrica) que fornece o número de rotas de 3 voos de uma cidade A_i para uma cidade A_j é dada por

$$\mathbf{W}^3 = \begin{bmatrix} 2 & 3 & 2 & 2 & 5 \\ 2 & 2 & 2 & 3 & 5 \\ 3 & 2 & 2 & 2 & 5 \\ 2 & 2 & 3 & 2 & 5 \\ 5 & 5 & 5 & 5 & 4 \end{bmatrix}.$$

Se quisermos saber quantas rotas existem com no máximo n , $n \in \mathbb{N}$, voos para ir de uma cidade a outra, basta então somar as matrizes $\mathbf{W}, \mathbf{W}^2, \mathbf{W}^3, \dots, \mathbf{W}^n$.

Exemplo 1.10. Note que, do exemplo anterior, temos

$$\mathbf{W} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} = \mathbf{W}^T,$$

porém

$$\mathbf{W}^2 = \begin{bmatrix} 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 4 \end{bmatrix} = (\mathbf{W}^T)^2$$

$$\mathbf{W}^4 = \begin{bmatrix} 8 & 7 & 7 & 7 & 9 \\ 7 & 8 & 7 & 7 & 9 \\ 7 & 7 & 8 & 7 & 9 \\ 7 & 7 & 7 & 8 & 9 \\ 9 & 9 & 9 & 9 & 20 \end{bmatrix} = (\mathbf{W}^T)^4$$

e assim por diante. Observando que as potências obtidas são matrizes simétricas, podemos concluir que

$$\mathbf{W}^k \text{ é simétrica} \not\Rightarrow \mathbf{W} \text{ é simétrica.}$$

Exemplo 1.11. Note como identificar um determinado número de rotas de n voos de uma cidade a outra pode não ser um trabalho tão simples quando o número de cidades aumenta.

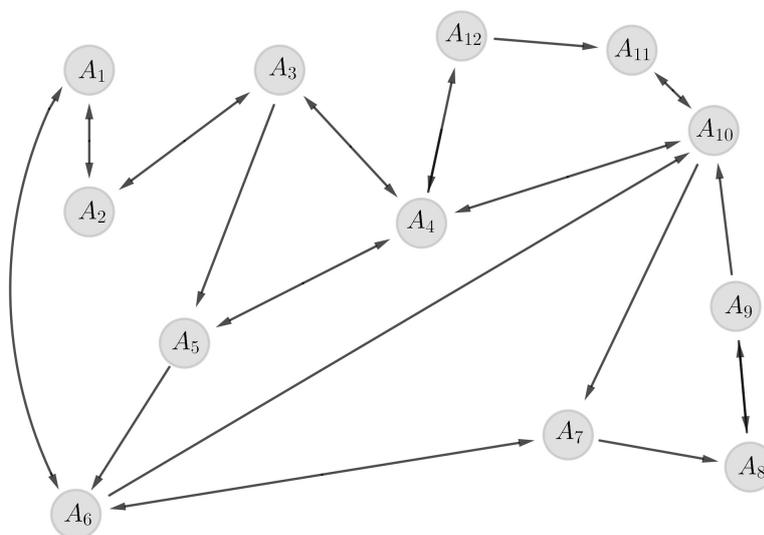


Figura 1.3: Rotas de viagem $A_1 \dots A_{12}$

Fonte: O autor

Porém, se considermos a matriz das rotas deste caso abaixo

$$\mathbf{Z} = \begin{matrix} & \begin{matrix} A_1 & A_2 & A_3 & A_4 & A_5 & A_6 & A_7 & A_8 & A_9 & A_{10} & A_{11} & A_{12} \end{matrix} \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \\ A_7 \\ A_8 \\ A_9 \\ A_{10} \\ A_{11} \\ A_{12} \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{matrix}$$

podemos determinar, por exemplo, a matriz que informa quantas rotas de 4 voos existem

entre a cidade A_i e a cidade A_j , com $i, j = 1, 2, \dots, 12$, como podemos ver a seguir:

$$\mathbf{Z}^4 = \begin{bmatrix} 6 & 0 & 5 & 1 & 2 & 2 & 3 & 1 & 1 & 6 & 0 & 2 \\ 1 & 6 & 1 & 7 & 4 & 5 & 3 & 1 & 0 & 2 & 3 & 1 \\ 5 & 2 & 8 & 7 & 6 & 6 & 4 & 2 & 0 & 9 & 3 & 5 \\ 2 & 7 & 1 & 20 & 6 & 7 & 9 & 1 & 1 & 3 & 11 & 1 \\ 5 & 0 & 7 & 1 & 5 & 3 & 3 & 2 & 1 & 11 & 0 & 5 \\ 1 & 5 & 0 & 8 & 2 & 6 & 5 & 3 & 1 & 2 & 5 & 0 \\ 3 & 0 & 2 & 1 & 1 & 1 & 3 & 1 & 2 & 4 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 2 & 0 & 3 & 0 & 1 \\ 1 & 1 & 0 & 6 & 1 & 1 & 4 & 0 & 2 & 1 & 4 & 0 \\ 2 & 1 & 6 & 2 & 5 & 5 & 2 & 4 & 0 & 10 & 1 & 5 \\ 1 & 1 & 0 & 5 & 1 & 1 & 3 & 0 & 1 & 1 & 3 & 0 \\ 2 & 0 & 6 & 1 & 5 & 3 & 1 & 2 & 0 & 9 & 0 & 5 \end{bmatrix}$$

Os exemplos 1.9 a 1.11 estão diretamente ligados à Teoria dos Grafos. A demonstração do caso geral de tais exemplos, bem como os detalhes adicionais sobre as matrizes de adjacência podem ser encontrados em [7] e [18]. Vale ressaltar que o número de voos não necessariamente implica maior ou menor distância, pois para isso, é necessário que hajam valores para as distâncias. Uma das maneiras de determinar a menor distância entre duas cidades é utilizando o *Algoritmo de Dijkstra*³, o qual pode ser visto em [62].

1.1.4 Matriz Inversa

Como havíamos mencionado no final da seção 1.1, a divisão de matrizes não está definida devido ao fato de que nem toda matriz é *invertível* e, além disso, para que uma matriz possa admitir sua *inversa*, quando é possível, é necessário que ela seja uma matriz quadrada. Formalizando a ideia, temos a seguinte

Definição 1.7. *Dada uma matriz $\mathbf{A} = [a_{ij}]_{n \times n}$, se existir uma matriz \mathbf{B} , tal que $\mathbf{AB} = \mathbf{I}_{n \times n} = \mathbf{BA}$, então essa matriz é chamada de *inversa* de \mathbf{A} . Nesse caso, diz-se que \mathbf{A} é *não singular*.*

Exemplo 1.12. Seguem abaixo alguns exemplos de matrizes e suas respectivas inversas

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & \mathbf{A}^{-1} &= -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} \\ \mathbf{B} &= \begin{bmatrix} -1 & 1 & 3 \\ 2 & 0 & 13 \\ -1 & 2 & 6 \end{bmatrix} & \mathbf{B}^{-1} &= \frac{1}{13} \begin{bmatrix} -26 & 0 & 13 \\ -25 & -3 & 19 \\ 4 & 1 & -2 \end{bmatrix} \\ \mathbf{C} &= \begin{bmatrix} 1 & 0 & -1 & -1 \\ -1 & 0 & -3 & -1 \\ 1 & 3 & -2 & -1 \\ 1 & -2 & 0 & -1 \end{bmatrix} & \mathbf{C}^{-1} &= -\frac{1}{2} \begin{bmatrix} -11 & 1 & 4 & 6 \\ 4 & 0 & -2 & -2 \\ 10 & 0 & -4 & -6 \\ -19 & 1 & 8 & 12 \end{bmatrix} \end{aligned}$$

³Edsger W. Dijkstra (1930-2002), vencedor do Prêmio Turing em 1972, é bem conhecido por suas contribuições à ciência computacional [24]. Dijkstra nasceu em Rotterdam (Holanda) e, em 1959, ele publicou um artigo de três páginas “A note on two problems in connexion with graphs”, o célebre e extremamente simples algoritmo para encontrar o caminho mais curto em um gráfico, agora conhecido como Algoritmo de Dijkstra [4].

Pode-se notar que a Definição 1.5 diz apenas quando a matriz possui ou não uma inversa, porém não informa como encontrá-la. Existem vários métodos para determinar a inversa de uma matriz, entre eles estão aqueles em que se utiliza *sistemas lineares*, determinantes e *eliminação gaussiana*. Este último nos fornece um algoritmo prático para determinar a inversa de uma matriz dada, o qual podemos utilizar, mesmo sem ter estudado a eliminação gaussiana.

Dada uma matriz quadrada

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

o algoritmo consiste em adicionar uma matriz identidade de mesma ordem de \mathbf{A} à direita desta, da seguinte forma

$$\mathbf{A} = \left[\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array} \right]$$

e, a partir daí, realizando-se somas e/ou subtrações de múltiplos das linhas de \mathbf{A} às linhas da mesma, chega-se a matriz

$$\mathbf{A} = \left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & y_{11} & y_{12} & \cdots & y_{1n} \\ 0 & 1 & \cdots & 0 & y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & y_{n1} & y_{n2} & \cdots & y_{nn} \end{array} \right],$$

onde $\mathbf{Y} = [y_{ij}]_{n \times n}$ é a inversa de \mathbf{A} . Resumindo, começa-se com \mathbf{A} do lado esquerdo e \mathbf{I} do lado direito, ao passo que no final do algoritmo deve-se ter \mathbf{I} do lado esquerdo e a matriz obtida do lado direito é a inversa de \mathbf{A} , quando esta existir. O próximo exemplo ilustra como se aplica o algoritmo. Para tanto, utilizaremos expressões como $L_1 + 2L_2$ para indicar que somamos as entradas da linha 1 com o dobro das entradas da linha 2 e, nesse caso, substituímos L_1 por $L_1 + 2L_2$. Expressões dessa forma serão chamadas, de agora em diante, de *combinações lineares*. Estas serão estudadas com mais profundidade na Seção 1.3.

Exemplo 1.13. Seja \mathbf{A} a matriz do Exemplo 1.12. Vamos determinar \mathbf{A}^{-1} seguindo o processo descrito acima.

$$\begin{aligned} & \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array} \right] \xrightarrow{L_2 - 3L_1} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array} \right] \\ & \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array} \right] \xrightarrow{L_1 + L_2} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{array} \right] \\ & \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{array} \right] \xrightarrow{-\frac{1}{2}L_2} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} \right] \end{aligned}$$

Portanto, temos que

$$\mathbf{A}^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$$

Agora, da mesma forma, determinemos \mathbf{B}^{-1} .

$$\begin{aligned} & \left[\begin{array}{ccc|ccc} -1 & 1 & 3 & 1 & 0 & 0 \\ 2 & 0 & 13 & 0 & 1 & 0 \\ -1 & 2 & 6 & 0 & 0 & 1 \end{array} \right] \xrightarrow{\substack{L_1+L_2 \\ L_2+2L_1 \\ L_3-L_1}} \left[\begin{array}{ccc|ccc} 1 & 1 & 16 & 1 & 1 & 0 \\ 0 & 2 & 19 & 2 & 1 & 0 \\ 0 & 1 & 3 & -1 & 0 & 1 \end{array} \right] \\ & \left[\begin{array}{ccc|ccc} 1 & 1 & 16 & 1 & 1 & 0 \\ 0 & 2 & 19 & 2 & 1 & 0 \\ 0 & 1 & 3 & -1 & 0 & 1 \end{array} \right] \xrightarrow{\substack{L_1-L_3 \\ L_2-L_3 \\ \frac{1}{3}L_3}} \left[\begin{array}{ccc|ccc} 1 & 0 & 13 & 2 & 1 & -1 \\ 0 & 1 & 16 & 3 & 1 & -1 \\ 0 & \frac{1}{3} & 1 & -\frac{1}{3} & 0 & \frac{1}{3} \end{array} \right] \\ & \left[\begin{array}{ccc|ccc} 1 & 0 & 13 & 2 & 1 & -1 \\ 0 & 1 & 16 & 3 & 1 & -1 \\ 0 & \frac{1}{3} & 1 & -\frac{1}{3} & 0 & \frac{1}{3} \end{array} \right] \xrightarrow{L_3-\frac{1}{3}L_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 13 & 2 & 1 & -1 \\ 0 & 1 & 16 & 3 & 1 & -1 \\ 0 & 0 & -\frac{13}{3} & -\frac{4}{3} & -\frac{1}{3} & \frac{2}{3} \end{array} \right] \\ & \left[\begin{array}{ccc|ccc} 1 & 0 & 13 & 2 & 1 & -1 \\ 0 & 1 & 16 & 3 & 1 & -1 \\ 0 & 0 & -\frac{13}{3} & -\frac{4}{3} & -\frac{1}{3} & \frac{2}{3} \end{array} \right] \xrightarrow{\substack{L_1+3L_3 \\ -\frac{3}{13}L_3}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & 1 \\ 0 & 1 & 16 & 3 & 1 & -1 \\ 0 & 0 & 1 & \frac{4}{13} & \frac{1}{13} & -\frac{2}{13} \end{array} \right] \\ & \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & 1 \\ 0 & 1 & 16 & 3 & 1 & -1 \\ 0 & 0 & 1 & \frac{4}{13} & \frac{1}{13} & -\frac{2}{13} \end{array} \right] \xrightarrow{L_2-16L_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & 1 \\ 0 & 1 & 0 & -\frac{25}{13} & -\frac{3}{13} & \frac{19}{13} \\ 0 & 0 & 1 & \frac{4}{13} & \frac{1}{13} & -\frac{2}{13} \end{array} \right] \end{aligned}$$

Logo, decorre que

$$\mathbf{B}^{-1} = \begin{bmatrix} -2 & 0 & 1 \\ -\frac{25}{13} & -\frac{3}{13} & \frac{19}{13} \\ \frac{4}{13} & \frac{1}{13} & -\frac{2}{13} \end{bmatrix} = \frac{1}{13} \begin{bmatrix} -26 & 0 & 13 \\ -25 & -3 & 19 \\ 4 & 1 & -2 \end{bmatrix}$$

Embora seja um processo prático para matrizes de pequena ordem, tal algoritmo depende das combinações entre as linhas, o que pode se tornar uma tarefa bastante dispendiosa, à medida em que o algoritmo for aplicado a matrizes de ordens maiores. Quando estudarmos os determinantes no Capítulo 2, iremos conhecer uma outra maneira de determinar a inversa de uma matriz.

Visto os exemplos de matrizes invertíveis e o algoritmo para determinar suas inversas, devemos ter em mente o seguinte resultado, o qual possui grande utilidade na resolução de problemas e demonstrações de outros teoremas relacionados a matrizes. Tal resultado diz que a inversa de uma matriz é única. Para melhor apresentar o resultado, enunciemos o seguinte

Teorema 1.2. *Seja \mathbf{A} uma matriz quadrada de ordem n . Se \mathbf{A} é invertível, então sua inversa é única.*

Demonstração. Seja \mathbf{A} invertível e seja \mathbf{B} sua inversa. Suponhamos que exista uma matriz quadrada \mathbf{C} , de mesma ordem que \mathbf{A} , tal que $\mathbf{AC} = \mathbf{I} = \mathbf{CA}$. Daí, temos que

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

provando assim a unicidade de \mathbf{B} . □

Vejam os a seguir as principais propriedades das matrizes inversas. Para tanto, consideremos duas matrizes não singulares \mathbf{A} e \mathbf{B} . Daí,

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

Tomemos $\mathbf{X} = \mathbf{A}^{-1}$. Assim, temos que

$$\mathbf{XX}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A} \Rightarrow \mathbf{A}^{-1}\mathbf{X}^{-1} = \mathbf{A}^{-1}\mathbf{A}.$$

Em seguida, multiplicando (à esquerda) ambos os membros da última igualdade acima, obtemos

$$\mathbf{AA}^{-1}\mathbf{X}^{-1} = \mathbf{AA}^{-1}\mathbf{A} \Rightarrow \mathbf{IX}^{-1} = \mathbf{IA} \Rightarrow \mathbf{X}^{-1} = \mathbf{A},$$

isto é, $\mathbf{A} = \mathbf{X}^{-1} = (\mathbf{A}^{-1})^{-1}$, o que acaba por provar o resultado procurado.

- \mathbf{AB} é não-singular.

Tomemos $\mathbf{Y} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. Como a multiplicação de matrizes é associativa, segue que

$$(\mathbf{AB})\mathbf{Y} = (\mathbf{AB})\mathbf{B}^{-1}\mathbf{A}^{-1} = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AIA}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}.$$

Segue, por definição, que $\mathbf{Y} = (\mathbf{AB})^{-1}$ e, portanto, \mathbf{AB} é não-singular.

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

Já provada na demonstração da propriedade acima.

- $(\mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1} \cdots \mathbf{A}_2^{-1}\mathbf{A}_1^{-1}$

Basta aplicar a propriedade acima $k - 1$ vezes seguidas.

- $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$

Seja $\mathbf{X} = \mathbf{A}^{-1}$. Assim

$$\mathbf{I} = \mathbf{XX}^{-1} \Rightarrow \mathbf{I} = \mathbf{I}^T = (\mathbf{XX}^{-1})^T = (\mathbf{X}^{-1})^T\mathbf{X}^T = \mathbf{A}^T\mathbf{X}^T.$$

Temos então, por definição, que $\mathbf{X}^T = (\mathbf{A}^T)^{-1}$. Por outro lado, como

$$\mathbf{X} = \mathbf{A}^{-1} \Rightarrow \mathbf{X}^T = (\mathbf{A}^{-1})^T,$$

podemos concluir que

$$(\mathbf{A}^{-1})^T = \mathbf{X}^T = (\mathbf{A}^T)^{-1},$$

como queríamos demonstrar.

Tendo em vista as propriedades acima, é válida a dúvida sobre a veracidade ou inveracidade de $(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1} + \mathbf{B}^{-1}$. Vejamos um contraexemplo para esse caso. Sejam

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

Temos que, por um lado

$$(\mathbf{A} + \mathbf{B})^{-1} = \begin{bmatrix} -\frac{8}{21} & \frac{1}{3} \\ \frac{1}{7} & 0 \end{bmatrix}$$

enquanto que

$$\mathbf{A}^{-1} + \mathbf{B}^{-1} = \begin{bmatrix} -3 & \frac{9}{4} \\ \frac{3}{2} & -\frac{1}{4} \end{bmatrix}.$$

Os casos em que a igualdade é verdadeira existem, porém não os abordaremos aqui, a fim de evitar digressões. Uma abordagem de tais casos pode ser vista em [42].

Em relação à lei do cancelamento para a multiplicação, isto é, $\mathbf{AB} = \mathbf{AC} \Rightarrow \mathbf{B} = \mathbf{C}$, esta também não é válida de modo geral. Por exemplo, se

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 6 & 8 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix},$$

segue que

$$\mathbf{AB} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \mathbf{AC},$$

porém $\mathbf{B} \neq \mathbf{C}$.

1.2 Matrizes de blocos

Esta seção será voltada ao estudo das *submatrizes* e *matrizes de blocos*, onde estas não só facilitam o processo de multiplicar matrizes, bem como possuem aplicabilidade em áreas da informática, como por exemplo a memória de dados.

Uma *submatriz* de uma matriz \mathbf{A} é uma matriz obtida através da eliminação de linhas e/ou colunas de \mathbf{A} . Consideremos, a título de exemplo, a matriz

$$\mathbf{A} = \begin{bmatrix} -1 & 3 & 7 & 6 \\ 4 & 10 & 3 & 19 \\ 0 & 5 & -3 & -13 \\ 1 & 9 & 7 & 2 \end{bmatrix}$$

As matrizes

$$\mathbf{B} = \begin{bmatrix} -1 & 7 \\ 0 & -3 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 3 & 6 \\ 9 & 2 \end{bmatrix} \quad \mathbf{D} = [10 \ 3 \ 19]$$

são exemplos de submatrizes de \mathbf{A} , onde

- **B** é obtida eliminando-se: 2ª e 4ª linha, 2ª e 4ª coluna de **A**.
- **C** é obtida eliminando-se: 2ª e 3ª linha, 1ª e 3ª coluna de **A**.
- **D** é obtida eliminando-se as linhas 1, 3 e 4 e a coluna 1 de **A**.

Dada uma matriz **A**, dizemos que esta é uma *matriz de blocos* ou *matriz particionada* quando é dividida em submatrizes por linhas horizontais e verticais, onde estas ficam entre as linhas e colunas de **A**. Temos a seguir alguns exemplos

$$\left[\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{array} \right] \quad \left[\begin{array}{ccc|c} -5 & 7 & 12 & 4 \\ 0 & 1 & 6 & -5 \\ \hline 13 & 0 & 7 & -8 \end{array} \right] \quad \left[\begin{array}{cc|ccc} a & b & c & d & e \\ f & g & h & i & j \\ k & l & m & n & o \\ \hline p & q & r & s & t \end{array} \right]$$

Exemplo 1.14. Sejam **A** e **B** as matrizes de blocos a seguir:

$$\mathbf{A} = \left[\begin{array}{ccc|cc} 5 & -1 & 2 & 3 & 4 \\ 8 & 3 & 1 & 0 & 6 \\ \hline 1 & -3 & 4 & 5 & 9 \end{array} \right] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \mathbf{B} = \left[\begin{array}{cc} 3 & -5 \\ 8 & 1 \\ \hline -5 & 7 \\ 2 & 0 \\ -1 & 4 \end{array} \right] = \begin{bmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{bmatrix}$$

Calculando **AB**, obtemos

$$\mathbf{AB} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} \end{bmatrix}$$

onde

$$\mathbf{A}_{11}\mathbf{B}_{11} = \begin{bmatrix} 5 & -1 & 2 \\ 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -5 \\ 8 & 1 \\ -5 & 7 \end{bmatrix} = \begin{bmatrix} -3 & -12 \\ 43 & -30 \end{bmatrix}$$

$$\mathbf{A}_{12}\mathbf{B}_{21} = \begin{bmatrix} 3 & -4 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 10 & -16 \\ -6 & 24 \end{bmatrix}$$

$$\mathbf{A}_{21}\mathbf{B}_{11} = \begin{bmatrix} 1 & -3 & 4 \end{bmatrix} \begin{bmatrix} 3 & -5 \\ 8 & 1 \\ -5 & 7 \end{bmatrix} = \begin{bmatrix} -41 & 20 \end{bmatrix}$$

$$\mathbf{A}_{22}\mathbf{B}_{21} = \begin{bmatrix} 5 & 9 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 36 \end{bmatrix}$$

de modo que

$$\mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} = \begin{bmatrix} 7 & -28 \\ 37 & -6 \end{bmatrix}, \quad \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} = \begin{bmatrix} -40 & 56 \end{bmatrix}.$$

Portanto

$$\mathbf{AB} = \begin{bmatrix} 7 & -28 \\ 37 & -6 \\ \hline -40 & 56 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix},$$

com

$$\mathbf{X}_1 = \begin{bmatrix} 7 & -28 \\ 37 & -6 \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} -40 & 56 \end{bmatrix}$$

A partir da definição de matrizes de blocos e do exemplo acima, podemos observar que uma matriz $\mathbf{A} = [a_{ij}]_{m \times n}$ pode ser escrita da forma

$$\mathbf{A} = [\text{col}_1(\mathbf{A}) \quad \text{col}_2(\mathbf{A}) \quad \dots \quad \text{col}_n(\mathbf{A})]$$

ou

$$\mathbf{A} = \begin{bmatrix} \text{lin}_1(\mathbf{A}) \\ \text{lin}_2(\mathbf{A}) \\ \vdots \\ \text{lin}_m(\mathbf{A}) \end{bmatrix},$$

onde $\text{lin}_i(\mathbf{A})$ representa a i -ésima linha de \mathbf{A} e $\text{col}_j(\mathbf{A})$ a j -ésima coluna de \mathbf{A} . Dessa forma, dadas duas matrizes $\mathbf{X} = [x_{ij}]_{m \times n}$ e $\mathbf{Y} = [y_{jk}]_{n \times p}$ e fazendo

$$\mathbf{X} = [\text{col}_1(\mathbf{X}) \quad \text{col}_2(\mathbf{X}) \quad \dots \quad \text{col}_n(\mathbf{X})] \quad \mathbf{Y} = \begin{bmatrix} \text{lin}_1(\mathbf{Y}) \\ \text{lin}_2(\mathbf{Y}) \\ \vdots \\ \text{lin}_m(\mathbf{Y}) \end{bmatrix},$$

podemos escrever, de modo geral

$$\mathbf{XY} = [\text{col}_1(\mathbf{X})\text{lin}_1(\mathbf{Y}) + \text{col}_2(\mathbf{X})\text{lin}_2(\mathbf{Y}) + \dots + \text{col}_n(\mathbf{X})\text{lin}_n(\mathbf{Y})]$$

Exemplo 1.15. Quando uma sonda espacial é lançada, talvez sejam necessárias algumas correções para colocá-la em uma trajetória precisa. A Rádio Telemetria⁴ fornece uma sequência de matrizes $\mathbf{X}_1, \dots, \mathbf{X}_k$, a qual informa, em momentos diferentes, uma comparação entre a posição atual da sonda e a posição planejada. Tomando \mathbf{A}_k como sendo a matriz de blocos $[\mathbf{X}_1 \quad \dots \quad \mathbf{X}_k]$, segue que a matriz $\mathbf{G}_k = \mathbf{A}_k \mathbf{A}_k^T$ é calculada à medida que os dados do radar são analisados, de modo que

$$\mathbf{G}_k = [\mathbf{X}_1 \quad \dots \quad \mathbf{X}_k] \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_k \end{bmatrix} = [\mathbf{X}_1^2 + \dots + \mathbf{X}_k^2]$$

Se uma nova matriz \mathbf{X}_{k+1} aparecer na sequência, então uma outra matriz $\mathbf{G}_{k+1} =$

⁴A Telemetria é uma tecnologia que permite medir e monitorar dados de maneira remota. Ela normalmente se refere à direção a qual uma determinada informação está direcionada, ou seja, a direção do sensor até o sistema de interrogação ou sistema de registro de dados. A Telemetria pode ser definida como uma subclasse das Telecomunicações, uma maneira mais complexa de trocar informações como internet, chamadas telefônicas ou transmissão de vídeo [36].

$\mathbf{A}_{k+1}\mathbf{A}_{k+1}^T$ deve ser calculada. Como as matrizes \mathbf{X}_i são alcançadas em alta velocidade, a carga computacional para realizar o cálculo de uma matriz \mathbf{G}_{i+1} pode ser severo. Porém, a multiplicação de matrizes em blocos reduz extremamente uma tal carga, visto que

$$\mathbf{G}_{k+1} = [\mathbf{X}_1 \quad \dots \quad \mathbf{X}_k \quad \mathbf{X}_{k+1}] \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_k \\ \mathbf{X}_{k+1} \end{bmatrix} = [\mathbf{X}_1^2 + \dots + \mathbf{X}_k^2 + \mathbf{X}_{k+1}^2]$$

ou seja, ao invés da matriz \mathbf{G}_{k+1} ser calculada do início, o computador apenas adiciona uma nova coluna à matriz \mathbf{G}_k , a qual é preenchida pela matriz \mathbf{X}_{k+1}^2 .

As matrizes em blocos possuem grande aplicabilidade computacional. De acordo com GOLUB [27], algoritmos expressos em blocos são especificamente ricos em multiplicação de matrizes, operação esta preferível em muitos ambientes de computação de alta performance. Um dos conceitos ligados ao desempenho de um computador, mais precisamente à memória deste, é a inversão de matrizes, pois este, segundo COSME [14] é um processo que pode exceder a capacidade de memória do computador, quando trata-se de uma matriz de ordem extremamente alta, como ocorre em problemas de regressão⁵.

Vejam os a seguir como a maneira de determinar a inversa de uma matriz de blocos 2×2 pode não ser uma tarefa tão simples. Consideremos assim, duas matrizes de blocos \mathbf{A} e \mathbf{B} , ambas de ordem n

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix},$$

com \mathbf{A}_{11} e \mathbf{A}_{22} quadradas e não-singulares de ordem k e $n - k$, respectivamente. Afim de que seja \mathbf{B} a inversa de \mathbf{A} , devemos calcular \mathbf{AB} , obtendo

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k \times k} & \mathbf{O}_{k \times (n-k)} \\ \mathbf{O}_{(n-k) \times k} & \mathbf{I}_{(n-k) \times (n-k)} \end{bmatrix}.$$

A igualdade acima nos fornece as quatro equações a seguir:

$$\mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} = \mathbf{I}_{k \times k} \tag{1.1}$$

$$\mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} = \mathbf{O}_{k \times (n-k)} \tag{1.2}$$

$$\mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} = \mathbf{O}_{(n-k) \times k} \tag{1.3}$$

$$\mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} = \mathbf{I}_{(n-k) \times (n-k)}. \tag{1.4}$$

Multiplicando 1.2 por \mathbf{A}_{11}^{-1} à esquerda, obtemos

$$\begin{aligned} \mathbf{A}_{11}^{-1}\mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22} &= \mathbf{A}_{11}^{-1}\mathbf{O}_{k \times (n-k)} \\ \mathbf{B}_{12} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22} &= \mathbf{O}_{k \times (n-k)} \\ \mathbf{B}_{12} &= -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22} \end{aligned} \tag{1.5}$$

⁵Problemas de regressão são estudados pela Análise de Regressão, uma das técnicas estatísticas mais amplamente usadas para investigar e modelar problemas que envolvem várias variáveis, onde uma delas é a variável de interesse e as demais são variáveis preditoras [47].

Analogamente, multiplicando 1.3 por \mathbf{A}_{22}^{-1} à esquerda, obtemos

$$\mathbf{B}_{21} = -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{11}. \quad (1.6)$$

Agora, substituindo 1.6 em 1.1,

$$\begin{aligned} \mathbf{A}_{11}\mathbf{B}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{11} &= \mathbf{I}_{k \times k} \\ (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})\mathbf{B}_{11} &= \mathbf{I}_{k \times k} \\ \mathbf{B}_{11} &= (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}. \end{aligned} \quad (1.7)$$

Da mesma forma, substituindo 1.5 em 1.4, obtemos

$$\mathbf{B}_{22} = (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}. \quad (1.8)$$

Antes de darmos prosseguimento, observemos que \mathbf{B} depende da invertibilidade das matrizes

$$\mathbf{K} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}, \quad \mathbf{L} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21},$$

as quais são denominadas *complementos de Schur*⁶ de \mathbf{A}_{11} e \mathbf{A}_{22} , respectivamente. Suponhamos então que \mathbf{K} e \mathbf{L} são não-singulares. Daí, substituindo 1.8 em 1.5 e 1.7 em 1.6, obtemos

$$\begin{aligned} \mathbf{B}_{12} &= -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \\ \mathbf{B}_{21} &= -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}, \end{aligned}$$

o que nos leva a

$$\mathbf{B}_1 = \begin{bmatrix} (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \end{bmatrix}.$$

Em termos dos complementos de Schur

$$\mathbf{B}_1 = \begin{bmatrix} \mathbf{L}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{K}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{L}^{-1} & \mathbf{K}^{-1}. \end{bmatrix}$$

Prosseguindo de maneira análoga para a igualdade abaixo

$$\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k \times k} & \mathbf{O}_{k \times (n-k)} \\ \mathbf{O}_{(n-k) \times k} & \mathbf{I}_{(n-k) \times (n-k)} \end{bmatrix}.$$

chegamos a

$$\mathbf{B}_2 = \begin{bmatrix} \mathbf{L}^{-1} & -\mathbf{L}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{K}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{K}^{-1} \end{bmatrix}.$$

Resta agora verificar se $\mathbf{B}_1 = \mathbf{B}_2$, para finalmente comprovar que $\mathbf{B} = \mathbf{A}^{-1}$. Ora, tanto a diagonal principal de \mathbf{B}_1 como a de \mathbf{B}_2 são iguais. Basta então verificar se

$$-\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{K}^{-1} = -\mathbf{L}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \quad \text{e} \quad -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{L}^{-1} = -\mathbf{K}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}$$

⁶Este nome é uma homenagem ao matemático alemão Issai Schur (1875-1941), o qual foi o primeiro a estudar matrizes deste tipo [46]. Suas descobertas matemáticas, particularmente em Álgebra, e sua história de vida, que inclui o terrível tratamento sofrido por ele na Alemanha nazista entre 1933 e 1939, são bem conhecidas [39].

Façamos o primeiro caso, sendo o segundo inteiramente análogo. Notemos que, por um lado

$$\begin{aligned}
(-\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{K}^{-1})^{-1} &= -\mathbf{K}\mathbf{A}_{12}^{-1}\mathbf{A}_{11} \\
&= (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})\mathbf{A}_{12}^{-1}\mathbf{A}_{11} \\
&= -\mathbf{A}_{22}\mathbf{A}_{12}^{-1}\mathbf{A}_{11} + \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{12}^{-1}\mathbf{A}_{11} \\
&= -\mathbf{A}_{22}\mathbf{A}_{12}^{-1}\mathbf{A}_{11} + \mathbf{A}_{21},
\end{aligned}$$

enquanto que

$$\begin{aligned}
(-\mathbf{L}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1})^{-1} &= -\mathbf{A}_{22}\mathbf{A}_{12}^{-1}\mathbf{L}_1 \\
&= -\mathbf{A}_{22}\mathbf{A}_{12}^{-1}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}) \\
&= -\mathbf{A}_{22}\mathbf{A}_{12}^{-1}\mathbf{A}_{11} + \mathbf{A}_{22}\mathbf{A}_{12}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} \\
&= -\mathbf{A}_{22}\mathbf{A}_{12}^{-1}\mathbf{A}_{11} + \mathbf{A}_{21}.
\end{aligned}$$

Com isso, temos que

$$(-\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{K}^{-1})^{-1} = -\mathbf{A}_{22}\mathbf{A}_{12}^{-1}\mathbf{A}_{11} + \mathbf{A}_{21} = (-\mathbf{L}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1})^{-1},$$

o que equivale a

$$-\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{K}^{-1} = -\mathbf{L}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}.$$

Da mesma forma,

$$-\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{L}^{-1} = -\mathbf{K}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}.$$

Logo, podemos concluir que $\mathbf{B}_1 = \mathbf{B}_2$ e, portanto, $\mathbf{B} = \mathbf{A}^{-1}$.

Podemos assim enunciar o seguinte teorema.

Teorema 1.3. *Seja \mathbf{A} uma matriz quadrada de ordem n escrita como matriz de blocos da forma*

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}.$$

com \mathbf{A}_{11} e \mathbf{A}_{22} não-singulares. Sejam ainda

$$\mathbf{K} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}, \quad \mathbf{L} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21},$$

os complementos de Schur de \mathbf{A}_{11} e \mathbf{A}_{22} , respectivamente. Se \mathbf{K} e \mathbf{L} forem invertíveis, então \mathbf{A} também o é, onde

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{L}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{K}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{L}^{-1} & \mathbf{K}^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{L}^{-1} & -\mathbf{L}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{K}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{K}^{-1} \end{bmatrix}.$$

Exemplo 1.16. Consideremos a matriz

$$\mathbf{A} = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 2 & 1 \\ \hline 1 & 4 & 4 & 3 \\ 5 & 1 & 2 & 1 \end{array} \right].$$

Temos que

$$\mathbf{A}_{11} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \Rightarrow \mathbf{A}_{11}^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

$$\mathbf{A}_{22} = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \Rightarrow \mathbf{A}_{22}^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{3}{2} \\ 1 & -2 \end{bmatrix}.$$

Daí, vem

$$\mathbf{K} = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 4 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ \frac{3}{2} & -\frac{7}{2} \end{bmatrix} \Rightarrow \mathbf{K}^{-1} = \frac{1}{26} \begin{bmatrix} 7 & 8 \\ 3 & -4 \end{bmatrix}$$

$$\mathbf{L} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & \frac{3}{2} \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} -6 & \frac{5}{2} \\ -2 & 3 \end{bmatrix} \Rightarrow \mathbf{L}^{-1} = \frac{1}{26} \begin{bmatrix} -6 & 5 \\ -4 & 12 \end{bmatrix}$$

de modo que

$$-\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{K}^{-1} = -\frac{1}{26} \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 7 & 8 \\ 3 & -4 \end{bmatrix} = -\frac{1}{26} \begin{bmatrix} 3 & -4 \\ 2 & 6 \end{bmatrix}$$

$$-\mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{L}^{-1} = -\frac{1}{26} \begin{bmatrix} -\frac{1}{2} & \frac{3}{2} \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} -6 & 5 \\ -4 & 12 \end{bmatrix} = -\frac{1}{26} \begin{bmatrix} -40 & 29 \\ 46 & -21 \end{bmatrix}.$$

Logo, temos que

$$\mathbf{A}^{-1} = -\frac{1}{26} \left[\begin{array}{cc|cc} 6 & -5 & 3 & -4 \\ 4 & -12 & 2 & 6 \\ \hline -40 & 29 & -7 & -8 \\ 46 & -21 & -3 & 4 \end{array} \right].$$

Se tivéssemos calculado a inversa de \mathbf{A} diretamente, por exemplo, utilizando o algoritmo descrito na seção anterior, teríamos encontrado

$$\mathbf{A}^{-1} = -\frac{1}{26} \begin{bmatrix} 6 & -5 & 3 & -4 \\ 4 & -12 & 2 & 6 \\ -40 & 29 & -7 & -8 \\ 46 & -21 & -3 & 4 \end{bmatrix}$$

A seção a seguir mostra uma aplicação interessante das matrizes, na qual poderemos observar algumas das propriedades estudadas até aqui, utilizadas de maneira prática.

1.3 Matrizes elementares

As *matrizes elementares* estão diretamente ligadas ao algoritmo descrito na seção 1.2 utilizado para a inversão de matrizes. Além disso, permitem que algumas demonstrações sejam realizadas sem a necessidade de envolver complicações notacionais e conceitos mais avançados, como por exemplo a *indução finita*.

Do ponto de vista computacional, segundo POOLE [53], o uso de matrizes elementares não é uma boa ideia quando se trata de realizar operações elementares entre as linhas, sendo preferível fazer isso de maneira direta, porém, elas são valiosas tanto na inversão

de matrizes, bem como para a resolução de sistemas lineares.

Definição 1.8. Uma **matriz elementar** \mathbf{E} é uma matriz obtida através de uma das três operações elementares entre linhas de uma matriz identidade:

- (Tipo 1) Troca da linha i pela linha j e vice-versa:

$$L_i \longleftrightarrow L_j.$$

- (Tipo 2) Multiplicação da linha i por uma constante real $c \neq 0$:

$$L_i \longleftrightarrow cL_i.$$

- (Tipo 3) Substituição da linha i pela soma/subtração desta com um múltiplo da linha j (combinação linear):

$$L_i \longleftrightarrow L_i \pm cL_j.$$

Temos a seguir um exemplo de matriz Tipo 1.

$$\mathbf{E}_{L_1 \leftrightarrow L_2} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

a qual é obtida através da troca de posição entre a primeira e a segunda linha da matriz $\mathbf{I}_{3 \times 3}$. Se multiplicarmos uma matriz 3×3 por $\mathbf{E}_{L_1 \leftrightarrow L_2}$, veremos que a mesma operação será realizada na referida matriz. Para comprovar isso, consideremos a matriz abaixo:

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

e então calculemos $\mathbf{E}_{L_1 \leftrightarrow L_2} \mathbf{A}$, de modo que

$$\begin{aligned} & \mathbf{E}_{L_1 \leftrightarrow L_2} \mathbf{A} \\ = & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \\ = & \begin{bmatrix} 0 \cdot a + 1 \cdot d + 0 \cdot g & 0 \cdot b + 1 \cdot e + 0 \cdot h & 0 \cdot c + 1 \cdot f + 0 \cdot i \\ 1 \cdot a + 0 \cdot d + 0 \cdot g & 1 \cdot b + 0 \cdot e + 0 \cdot h & 1 \cdot c + 0 \cdot f + 0 \cdot i \\ 0 \cdot a + 0 \cdot d + 1 \cdot g & 0 \cdot b + 0 \cdot e + 1 \cdot h & 0 \cdot c + 0 \cdot f + 1 \cdot i \end{bmatrix} \\ = & \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix} \end{aligned}$$

Como podemos observar, a primeira e a segunda coluna de \mathbf{A} trocaram de posição. Essa propriedade é importante, pois ela garante que a troca de linhas de uma matriz não a faz deixar de ser matriz, apenas gera uma nova matriz.

As matrizes do Tipo 2 realizam a mesma operação pela qual são formadas, a uma matriz que ela multiplica, como podemos ver a seguir com as matrizes

$$\mathbf{E}_{L_2 \leftrightarrow 5L_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplicando $\mathbf{E}_{L_2 \leftrightarrow 5L_2}$ por \mathbf{A}

$$\begin{aligned} & \mathbf{E}_{L_2 \leftrightarrow 5L_2} \mathbf{A} \\ = & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \\ = & \begin{bmatrix} 1 \cdot a + 0 \cdot d + 0 \cdot g & 1 \cdot b + 0 \cdot e + 0 \cdot h & 1 \cdot c + 0 \cdot f + 0 \cdot i \\ 0 \cdot a + 5 \cdot d + 0 \cdot g & 0 \cdot b + 5 \cdot e + 0 \cdot h & 0 \cdot c + 5 \cdot f + 0 \cdot i \\ 0 \cdot a + 0 \cdot d + 1 \cdot g & 0 \cdot b + 0 \cdot e + 1 \cdot h & 0 \cdot c + 0 \cdot f + 1 \cdot i \end{bmatrix} \\ = & \begin{bmatrix} a & b & c \\ 5d & 5e & 5f \\ g & h & i \end{bmatrix}. \end{aligned}$$

Podemos então observar que a segunda linha de \mathbf{A} foi multiplicada por 5, assim como em $\mathbf{E}_{L_2 \leftrightarrow 5L_2}$.

Consideremos agora a matriz de Tipo 3 a seguir

$$\mathbf{E}_{L_3 \leftrightarrow L_3 - 4L_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}.$$

Multiplicando-a por \mathbf{A} , obtemos

$$\begin{aligned} & \mathbf{E}_{L_3 \leftrightarrow L_3 - 4L_1} \mathbf{A} \\ = & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \\ = & \begin{bmatrix} 1 \cdot a + 0 \cdot d + 0 \cdot g & 1 \cdot b + 0 \cdot e + 0 \cdot h & 1 \cdot c + 0 \cdot f + 0 \cdot i \\ 0 \cdot a + 5 \cdot d + 0 \cdot g & 0 \cdot b + 5 \cdot e + 0 \cdot h & 0 \cdot c + 5 \cdot f + 0 \cdot i \\ -4 \cdot a + 0 \cdot d + 1 \cdot g & -4 \cdot b + 0 \cdot e + 1 \cdot h & -4 \cdot c + 0 \cdot f + 1 \cdot i \end{bmatrix} \\ = & \begin{bmatrix} a & b & c \\ d & e & f \\ g - 4a & h - 4b & i - 4c \end{bmatrix} \end{aligned}$$

Mais uma vez, podemos observar que a operação pela qual $\mathbf{E}_{L_3 \leftrightarrow L_3 - 4L_1}$ foi obtida, é realizada em \mathbf{A} após a multiplicação.

Para generalizarmos a ideia, utilizaremos as matrizes de blocos, pois estas evitarão uma sobrecarga notacional e, além disso, fará com que as ideias fiquem melhor organiza-

das. Vamos então escrever a matriz

$$\mathbf{I}_{n \times n} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

sob a forma

$$\mathbf{I} = \begin{bmatrix} \mathbf{I}_{11} \\ \mathbf{I}_{21} \\ \vdots \\ \mathbf{I}_{n1} \end{bmatrix}$$

e consideraremos uma matriz \mathbf{X} qualquer $n \times p$

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

a qual será escrita da forma

$$\mathbf{X} = [\mathbf{X}_{11} \quad \mathbf{X}_{12} \quad \cdots \quad \mathbf{X}_{1p}].$$

Agora observemos que cada submatriz \mathbf{I}_{i1} é uma matriz linha $1 \times n$, enquanto que cada submatriz \mathbf{X}_{1k} é uma matriz coluna $n \times 1$, de modo que o produto $\mathbf{I}_{i1}\mathbf{X}_{1k}$ nos fornece, pela definição de multiplicação de matrizes, uma matriz 1×1 , a qual é $[a_{ik}]$. Em termos

$$\mathbf{I}_{i1}\mathbf{X}_{1k} = [x_{ik}].$$

Dessa forma, se trocarmos duas linhas de \mathbf{I} (a primeira linha com alguma linha s de \mathbf{I} , por simplicidade), obteremos a matriz elementar abaixo

$$\mathbf{E}_{L_1 \leftrightarrow L_s} = \begin{bmatrix} \mathbf{I}_{s1} \\ \mathbf{I}_{21} \\ \vdots \\ \mathbf{I}_{11} \\ \vdots \\ \mathbf{I}_{n1} \end{bmatrix}$$

Daí, fazendo $\mathbf{E}_{L_1 \leftrightarrow L_s}\mathbf{A}$, obtemos

$$\mathbf{E}_{L_1 \leftrightarrow L_s}\mathbf{X} = \begin{bmatrix} \mathbf{I}_{s1}\mathbf{X}_{11} & \mathbf{I}_{s1}\mathbf{X}_{12} & \cdots & \mathbf{I}_{s1}\mathbf{X}_{1p} \\ \mathbf{I}_{21}\mathbf{X}_{11} & \mathbf{I}_{21}\mathbf{X}_{12} & \cdots & \mathbf{I}_{21}\mathbf{X}_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{I}_{11}\mathbf{X}_{11} & \mathbf{I}_{11}\mathbf{X}_{12} & \cdots & \mathbf{I}_{11}\mathbf{X}_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{I}_{n1}\mathbf{X}_{11} & \mathbf{I}_{n1}\mathbf{X}_{12} & \cdots & \mathbf{I}_{n1}\mathbf{X}_{1p} \end{bmatrix} = \begin{bmatrix} x_{s1} & x_{s2} & \cdots & x_{sp} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}.$$

Logo, podemos comprovar que a mesma operação de $\mathbf{E}_{L_1 \leftrightarrow L_s}$ foi realizada em \mathbf{A} . O processo é análogo para quaisquer outra matriz de tipo 1.

Consideremos agora uma matriz de tipo 2 de ordem n e uma constante real c :

$$\mathbf{E}_{L_s \leftrightarrow cL_s} = \begin{bmatrix} \mathbf{I}_{11} \\ \mathbf{I}_{21} \\ \vdots \\ c\mathbf{I}_{s1} \\ \vdots \\ \mathbf{I}_{n1} \end{bmatrix},$$

e em seguida calculemos $\mathbf{E}_{L_s \leftrightarrow cL_s} \mathbf{X}$, obtendo

$$\mathbf{E}_{L_s \leftrightarrow cL_s} \mathbf{X} = \begin{bmatrix} \mathbf{I}_{11} \mathbf{X}_{11} & \mathbf{I}_{11} \mathbf{X}_{12} & \cdots & \mathbf{I}_{11} \mathbf{X}_{1p} \\ \mathbf{I}_{21} \mathbf{X}_{11} & \mathbf{I}_{21} \mathbf{X}_{12} & \cdots & \mathbf{I}_{21} \mathbf{X}_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ c\mathbf{I}_{s1} \mathbf{X}_{11} & c\mathbf{I}_{s1} \mathbf{X}_{12} & \cdots & c\mathbf{I}_{s1} \mathbf{X}_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{I}_{n1} \mathbf{X}_{11} & \mathbf{I}_{n1} \mathbf{X}_{12} & \cdots & \mathbf{I}_{n1} \mathbf{X}_{1p} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ cx_{s1} & cx_{s2} & \cdots & cx_{sp} \\ \vdots & \vdots & \vdots & \vdots \\ cx_{n1} & cx_{n2} & \cdots & cx_{np} \end{bmatrix},$$

o que mostra que a linha s de \mathbf{X} foi multiplicada por c , assim como em $\mathbf{E}_{L_i \leftrightarrow cL_i}$.

Finalmente, seja $\mathbf{E}_{L_s \leftrightarrow L_s \pm cL_t}$ uma matriz de tipo 3 e de ordem n , escrita como matriz de blocos, tal que

$$\mathbf{E}_{L_s \leftrightarrow L_s \pm cL_t} = \begin{bmatrix} \mathbf{I}_{11} \\ \mathbf{I}_{21} \\ \vdots \\ \mathbf{I}_{s1} \pm c\mathbf{I}_{t1} \\ \vdots \\ \mathbf{I}_{n1} \end{bmatrix}.$$

Daí, fazendo $\mathbf{E}_{L_s \leftrightarrow L_s \pm cL_t} \mathbf{X}$, obtemos

$$\begin{aligned} & \mathbf{E}_{L_s \leftrightarrow L_s \pm cL_t} \mathbf{X} \\ &= \begin{bmatrix} \mathbf{I}_{11} \mathbf{X}_{11} & \mathbf{I}_{11} \mathbf{X}_{12} & \cdots & \mathbf{I}_{11} \mathbf{X}_{1p} \\ \mathbf{I}_{21} \mathbf{X}_{11} & \mathbf{I}_{21} \mathbf{X}_{12} & \cdots & \mathbf{I}_{21} \mathbf{X}_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ (\mathbf{I}_{s1} \pm c\mathbf{I}_{t1}) \mathbf{X}_{11} & (\mathbf{I}_{s1} \pm c\mathbf{I}_{t1}) \mathbf{X}_{12} & \cdots & (\mathbf{I}_{s1} \pm c\mathbf{I}_{t1}) \mathbf{X}_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{I}_{n1} \mathbf{X}_{11} & \mathbf{I}_{n1} \mathbf{X}_{12} & \cdots & \mathbf{I}_{n1} \mathbf{X}_{1p} \end{bmatrix} \\ &= \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{s1} \pm cx_{t1} & x_{s2} \pm cx_{t2} & \cdots & x_{sp} \pm cx_{tp} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}. \end{aligned}$$

Obtivemos acima, como resultado, uma matriz cuja linha s foi substituída pela soma\subtração

desta com um múltiplo de uma linha t qualquer de \mathbf{X} . Portanto, podemos enunciar, tendo em vista a discussão acima, o seguinte teorema:

Teorema 1.4. *Sejam $\mathbf{E}_{L_i \leftrightarrow L_j}$, $\mathbf{E}_{L_i \leftrightarrow cL_i}$ e $\mathbf{E}_{L_i \leftrightarrow L_i \pm cL_j}$ matrizes elementares de ordem n e \mathbf{X} uma matriz de ordem $n \times p$. Tem-se que*

- $\mathbf{E}_{L_i \leftrightarrow L_j} \mathbf{X} = \mathbf{X}_{L_i \leftrightarrow L_j}$
- $\mathbf{E}_{L_i \leftrightarrow cL_i} \mathbf{X} = \mathbf{X}_{L_i \leftrightarrow cL_i}$
- $\mathbf{E}_{L_i \leftrightarrow L_i \pm cL_j} \mathbf{X} = \mathbf{X}_{L_i \leftrightarrow L_i \pm cL_j}$.

Após a enunciação do teorema acima, vamos agora buscar, como prometido, o resultado que valida o algoritmo supracitado para encontrar a inversa de uma matriz invertível. Para tanto, necessitaremos dos lemas a seguir:

Lema 1.1. *Toda matriz elementar é invertível e sua inversa é uma matriz elementar de igual tipo.*

Demonstração. Com efeito, consideremos inicialmente uma matriz elementar de tipo 1 escrita como matriz coluna de blocos, onde, por simplicidade, trocamos a linha s pela primeira linha de \mathbf{I} , sendo o processo análogo para quaisquer duas linhas.

$$\mathbf{E}_{L_1 \leftrightarrow L_s} = \begin{bmatrix} \mathbf{I}_{s1} \\ \mathbf{I}_{21} \\ \vdots \\ \mathbf{I}_{11} \\ \vdots \\ \mathbf{I}_{n1} \end{bmatrix}.$$

Em seguida escrevemos $\mathbf{E}_{L_1 \leftrightarrow L_s}$ como matriz linha de blocos, trocando a coluna s e a primeira coluna de posições:

$$\mathbf{E}_{L_1 \leftrightarrow L_s} = [\mathbf{I}_{1s} \quad \mathbf{I}_{12} \quad \cdots \quad \mathbf{I}_{11} \quad \cdots \quad \mathbf{I}_{1n}].$$

Vamos mostrar que $\mathbf{E}_{L_1 \leftrightarrow L_s}$ é sua própria inversa, o que em palavras significa que para revertermos uma troca de linhas, basta aplicar a troca novamente. Assim

$$\begin{aligned} \mathbf{E}_{L_1 \leftrightarrow L_s} \mathbf{E}_{L_1 \leftrightarrow L_s} &= \begin{bmatrix} \mathbf{I}_{s1} \mathbf{I}_{1s} & \mathbf{I}_{s1} \mathbf{I}_{12} & \cdots & \mathbf{I}_{s1} \mathbf{I}_{11} & \cdots & \mathbf{I}_{s1} \mathbf{I}_{1n} \\ \mathbf{I}_{21} \mathbf{I}_{1s} & \mathbf{I}_{21} \mathbf{I}_{12} & \cdots & \mathbf{I}_{21} \mathbf{I}_{11} & \cdots & \mathbf{I}_{21} \mathbf{I}_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{I}_{11} \mathbf{I}_{1s} & \mathbf{I}_{11} \mathbf{I}_{12} & \cdots & \mathbf{I}_{11} \mathbf{I}_{11} & \cdots & \mathbf{I}_{11} \mathbf{I}_{1n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{I}_{n1} \mathbf{I}_{1s} & \mathbf{I}_{n1} \mathbf{I}_{12} & \cdots & \mathbf{I}_{n1} \mathbf{I}_{11} & \cdots & \mathbf{I}_{n1} \mathbf{I}_{1n} \end{bmatrix} = \\ &= \begin{bmatrix} i_{ss} & i_{s2} & \cdots & i_{s1} & \cdots & i_{sn} \\ i_{2s} & i_{22} & \cdots & i_{21} & \cdots & i_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ i_{1s} & i_{12} & \cdots & i_{11} & \cdots & i_{1n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ i_{ns} & i_{n2} & \cdots & i_{n1} & \cdots & i_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \end{aligned}$$

onde a última igualdade decorre do fato de que $i_{jk} = 1$ quando $j = k$ e $i_{jk} = 0$ quando $j \neq k$.

Vamos agora tomar uma matriz de tipo 2 $\mathbf{E}_{L_s \leftrightarrow cL_s}$, onde c é uma constante real não-nula e consideraremos a matriz elementar de tipo 2 $\mathbf{E}_{L_s \leftrightarrow \frac{1}{c}L_s}$, a qual mostraremos ser a sua inversa. Prosseguindo de maneira análoga ao primeiro caso provado acima, escreveremos $\mathbf{E}_{L_s \leftrightarrow \frac{1}{c}L_s}$ como matriz coluna de blocos, $\mathbf{E}_{L_s \leftrightarrow cL_s}$ como matriz linha de blocos e em seguida desenvolvemos o produto abaixo

$$\begin{aligned}
& \mathbf{E}_{L_s \leftrightarrow \frac{1}{c}L_s} \mathbf{E}_{L_s \leftrightarrow cL_s} \\
= & \begin{bmatrix} \mathbf{I}_{11} \\ \mathbf{I}_{21} \\ \vdots \\ \frac{1}{c} \mathbf{I}_{s1} \\ \vdots \\ \mathbf{I}_{n1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{11} & \mathbf{I}_{12} & \cdots & c \mathbf{I}_{1s} & \cdots & \mathbf{I}_{1n} \end{bmatrix} \\
= & \begin{bmatrix} \mathbf{I}_{11} \mathbf{I}_{11} & \mathbf{I}_{11} \mathbf{I}_{12} & \cdots & \mathbf{I}_{11} c \mathbf{I}_{1s} & \cdots & \mathbf{I}_{11} \mathbf{I}_{1n} \\ \mathbf{I}_{21} \mathbf{I}_{11} & \mathbf{I}_{21} \mathbf{I}_{12} & \cdots & \mathbf{I}_{21} c \mathbf{I}_{1s} & \cdots & \mathbf{I}_{21} \mathbf{I}_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \frac{1}{c} \mathbf{I}_{s1} \mathbf{I}_{11} & \frac{1}{c} \mathbf{I}_{s1} \mathbf{I}_{12} & \cdots & \frac{1}{c} \mathbf{I}_{s1} c \mathbf{I}_{1s} & \cdots & \frac{1}{c} \mathbf{I}_{s1} \mathbf{I}_{1n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{I}_{n1} \mathbf{I}_{11} & \mathbf{I}_{n1} \mathbf{I}_{12} & \cdots & \mathbf{I}_{n1} c \mathbf{I}_{1s} & \cdots & \mathbf{I}_{n1} \mathbf{I}_{1n} \end{bmatrix} \\
= & \begin{bmatrix} i_{11} & i_{12} & \cdots & c \cdot i_{1s} & \cdots & i_{1n} \\ i_{21} & i_{22} & \cdots & c \cdot i_{2s} & \cdots & i_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \frac{1}{c} i_{s1} & \frac{1}{c} i_{s2} & \cdots & \frac{1}{c} \cdot c \cdot i_{ss} & \cdots & \frac{1}{c} i_{sn} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ i_{n1} & i_{n2} & \cdots & c \cdot i_{ns} & \cdots & i_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}.
\end{aligned}$$

Temos portanto, provado que $\mathbf{E}_{L_s \leftrightarrow \frac{1}{c}L_s}$ é inversa à esquerda de $\mathbf{E}_{L_s \leftrightarrow cL_s}$. Em palavras, podemos dizer que para reverter a multiplicação de uma linha por uma constante não-nula, basta multiplicarmos esta linha pelo inverso de uma tal constante.

Para finalizar, provemos que a inversa de uma matriz elementar de tipo 3 $\mathbf{E}_{L_s \leftrightarrow L_s \pm cL_1}$ (onde a linha L_i foi escolhida apenas por simplicidade) é uma matriz também elementar de mesmo tipo, a saber $\mathbf{E}_{L_s \leftrightarrow L_s \mp cL_1}$. Prosseguindo de maneira análoga aos dois casos provados anteriormente, temos que

$$\begin{aligned}
& \mathbf{E}_{L_s \leftrightarrow L_s \mp cL_i} \mathbf{E}_{L_s \leftrightarrow L_s \pm cL_1} \\
= & \begin{bmatrix} \mathbf{I}_{11} \\ \mathbf{I}_{21} \\ \vdots \\ \mathbf{I}_{s1} \mp c\mathbf{I}_{11} \\ \vdots \\ \mathbf{I}_{n1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{11} \pm c\mathbf{I}_{1s} & \mathbf{I}_{12} & \cdots & \mathbf{I}_{1s} & \cdots & \mathbf{I}_{1n} \end{bmatrix} \\
= & \begin{bmatrix} i_{11} \pm i_{1s} & i_{12} & \cdots & i_{1s} & \cdots & i_{1n} \\ i_{21} \pm i_{2s} & i_{22} & \cdots & i_{2s} & \cdots & i_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ i_{s1} \pm ci_{ss} \mp ci_{11} - c^2i_{1s} & i_{s2} \mp ci_{12} & \cdots & i_{ss} \mp ci_{1s} & \cdots & i_{sn} \mp ci_{1n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ i_{n1} \pm ci_{ns} & i_{n2} & \cdots & i_{ns} & \cdots & i_{nn} \end{bmatrix} \\
= & \begin{bmatrix} 1 \pm 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 \pm 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 \pm c \cdot 1 \mp c \cdot 1 - c^2 \cdot 0 & 0 \mp c \cdot 0 & \cdots & 1 \mp c \cdot 0 & \cdots & 0 \mp c \cdot 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 \pm c \cdot 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \\
= & \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}.
\end{aligned}$$

Logo, podemos concluir que toda matriz elementar de tipo 1, 2 ou 3 possui inversa, a qual é também elementar e de mesmo tipo da matriz a qual ela inverte. \square

Exemplo 1.17. As inversas das matrizes elementares abaixo

$$\mathbf{E}_{L_2 \leftrightarrow L_2 - 3L_1} = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}, \quad \mathbf{E}_{L_1 \leftrightarrow L_1 + L_2} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{E}_{L_2 \leftrightarrow -\frac{1}{2}L_2} = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}$$

são

$$\mathbf{E}_{L_2 \leftrightarrow L_2 + 3L_1} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}, \quad \mathbf{E}_{L_1 \leftrightarrow L_1 - L_2} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{E}_{L_2 \leftrightarrow -2L_2} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix},$$

respectivamente.

O Lema 1.1 acima, juntamente com o Teorema 1.4, sustentam a validade do algoritmo para inverter matrizes, pois cada operação realizada nas linhas de \mathbf{A} é ao mesmo tempo realizada nas linhas de \mathbf{I} , porém de maneira reversa, levando \mathbf{A} para \mathbf{I} e \mathbf{I} para \mathbf{A}^{-1} . Formalizaremos tal discussão quando provarmos o Teorema Fundamental das Matrizes Elementares.

Como já sabemos que a inversa de uma matriz, quando existe, é única, precisaremos apenas do próximo lema para provarmos o supracitado teorema.

Lema 1.2. Se \mathbf{A} é uma matriz quadrada invertível de ordem n e \mathbf{B} , \mathbf{X} são matrizes $n \times 1$, tais que $\mathbf{AX} = \mathbf{B}$, então a matriz \mathbf{X} é única, a qual é dada por

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$$

Demonstração. Provemos primeiro a existência de \mathbf{X} , isto é, que existe uma matriz \mathbf{X} que cumpre $\mathbf{AX} = \mathbf{B}$. Para tanto, basta substituir $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$ na referida equação, de modo que

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

Segue então que tal matriz existe. Resta agora provarmos sua unicidade. Suponhamos assim que exista uma matriz \mathbf{Y} , tal que $\mathbf{AY} = \mathbf{B}$. Daí, multiplicando esta última igualdade pela inversa de \mathbf{A} em ambos os membros, decorre

$$\mathbf{A}^{-1}(\mathbf{AY}) = \mathbf{A}^{-1}\mathbf{B} \Rightarrow (\mathbf{A}^{-1}\mathbf{A})\mathbf{Y} = \mathbf{A}^{-1}\mathbf{B} \Rightarrow \mathbf{Y} = \mathbf{A}^{-1}\mathbf{B} = \mathbf{X},$$

provando assim a unicidade de \mathbf{X} . □

Definição 1.9. Dadas duas matrizes \mathbf{A} e \mathbf{B} , diz-se que \mathbf{A} é *equivalente por linhas* à matriz \mathbf{B} se esta última pode ser obtida pela multiplicação de sucessivas matrizes elementares por \mathbf{A} , ou seja,

$$\mathbf{B} = \mathbf{E}_1 \cdots \mathbf{E}_k \mathbf{A}.$$

A definição acima nos permite ir além, pois se uma matriz \mathbf{B} é equivalente por linhas a uma matriz \mathbf{A} , então podemos dizer que \mathbf{A} é equivalente por linhas a \mathbf{B} , pois toda matriz elementar é invertível e sua inversa é uma matriz elementar de mesmo tipo. Dizemos então, de maneira mais simples, que \mathbf{A} e \mathbf{B} são equivalentes por linhas. Por exemplo, as matrizes

$$\mathbf{A} = \begin{bmatrix} 4 & -26 & 20 \\ -8 & 29 & 17 \\ 12 & 3 & 3 \end{bmatrix} \quad \text{e} \quad \mathbf{B} = \begin{bmatrix} 0 & -1 & 5 \\ 2 & 7 & -3 \\ 4 & 1 & 1 \end{bmatrix}$$

são equivalentes por linhas, pois

$$\begin{bmatrix} 4 & -26 & 20 \\ -8 & 29 & 17 \\ 12 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -4 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 5 \\ 2 & 7 & -3 \\ 4 & 1 & 1 \end{bmatrix},$$

isto é, $\mathbf{A} = \mathbf{E}_{L_1 \leftrightarrow L_2} \mathbf{E}_{L_3 \leftrightarrow 3L_3} \mathbf{E}_{L_1 \leftrightarrow L_1 - 4L_2} \mathbf{B}$.

Podemos agora enunciar o Teorema Fundamental das Matrizes Elementares, o qual, de acordo com POOLE [53], está conectado a boa parte dos assuntos da Álgebra Linear, seja no desenvolvimento de suas caracterizações ou aplicações.

Teorema 1.5. Seja $\mathbf{A} = [a_{ij}]$ uma matriz $n \times n$. As seguintes afirmações são equivalentes:

a) \mathbf{A} é invertível.

b) $\mathbf{AX} = \mathbf{B} \Rightarrow \mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$, onde $\mathbf{X} = [x_{i1}]_{n \times 1}$ e $\mathbf{B} = [b_{i1}]_{n \times 1}$.

c) $\mathbf{AX} = \mathbf{0} \Rightarrow \mathbf{X} = \mathbf{0}$, onde $\mathbf{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{n \times 1}$

d) \mathbf{A} pode ser reduzida a \mathbf{I} por meio de operações elementares.

e) \mathbf{A} pode ser escrita como produto de matrizes elementares.

Demonstração. Provaremos as implicações

$$a) \Rightarrow b) \Rightarrow c) \Rightarrow d) \Rightarrow e) \Rightarrow a)$$

a fim de garantir a equivalência entre cada afirmação.

a) \Rightarrow b): Verdadeira, de acordo com o Lema 1.2.

b) \Rightarrow c): Seja $\mathbf{AX} = \mathbf{0}$. Temos, pelo Lema 1.2, que \mathbf{X} é única e é tal que

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{0} = \mathbf{0}.$$

c) \Rightarrow d): Se $\mathbf{AX} = \mathbf{0}$, então podemos escrever tal igualdade sob a forma

$$\begin{aligned} a_{11}x_{11} + a_{12}x_{21} + \cdots + a_{1n}x_{n1} &= 0 \\ a_{21}x_{11} + a_{22}x_{21} + \cdots + a_{2n}x_{n1} &= 0 \\ &\vdots \\ a_{n1}x_{11} + a_{n2}x_{21} + \cdots + a_{nn}x_{n1} &= 0 \end{aligned}$$

Porém, como a única solução é a trivial, segue que o algoritmo de inversão de matrizes pode ser aplicado até obter-se

$$\begin{aligned} 1 \cdot x_{11} + 0 \cdot x_{21} + \cdots + 0 \cdot x_{n1} &= 0 \\ 0 \cdot x_{11} + 1 \cdot x_{21} + \cdots + 0 \cdot x_{n1} &= 0 \\ &\vdots \\ 0 \cdot x_{11} + 0 \cdot x_{21} + \cdots + 1 \cdot x_{n1} &= 0 \end{aligned}$$

o que significa que \mathbf{A} pode ser reduzida à matriz identidade por meio de matrizes elementares. De fato, se \mathbf{A} não pudesse ser reduzida à matriz identidade, então só seria possível multiplicá-la por uma certa quantidade k de matrizes elementares, até obtermos

$$\mathbf{E}_1\mathbf{E}_2 \cdots \mathbf{E}_k\mathbf{A} = \mathbf{B}.$$

Como $\mathbf{B} \neq \mathbf{I}$, então, no pior dos casos, \mathbf{B} é tal que $b_{ii} = 1$, porém existe pelo menos uma entrada b_{ij} , com $i \neq j$, tal que $b_{ij} \neq 0$. Suponhamos, sem perda de generalidade, que seja $b_{12} \neq 0$. Temos assim que

$$x_{11} + b_{12}x_{21} = 0$$

É claro que $\mathbf{X} = \mathbf{0}$ é uma solução, porém, se for $x_{21} \neq 0$, então $x_{11} = -b_{12}x_{21}$ e a matriz

$$\mathbf{X} = \begin{bmatrix} -b_{12}x_{21} \\ x_{21} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

é também uma solução (não-trivial) da equação dada, o que contradiz a hipótese, isto é, de que a única solução é a trivial. Logo, segue que $\mathbf{B} = \mathbf{I}$ e assim,

$$\mathbf{E}_1\mathbf{E}_2 \cdots \mathbf{E}_k\mathbf{A} = \mathbf{I}.$$

Logo, \mathbf{A} pode ser reduzida à matriz identidade por meio de matrizes elementares.

d) \Rightarrow e): Se \mathbf{A} pode ser reduzida a uma matriz identidade, então existe uma sequência de matrizes elementares $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k$, tais que

$$\mathbf{E}_k \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} = \mathbf{I}.$$

Com sorte, o Lema 1.1 nos garante que cada uma dessas matrizes elementares é invertível e suas inversas também são matrizes elementares, de modo que podemos multiplicar ambos os membros da igualdade acima por $\mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1}$, e assim

$$\begin{aligned} \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1} \mathbf{E}_k \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} &= \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1} \mathbf{I} \\ \Rightarrow \mathbf{A} &= \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1} \end{aligned}$$

Portanto, \mathbf{A} pode ser escrita como produto de matrizes elementares.

e) \Rightarrow a): Sabemos que o produto de matrizes inversas é igual à inversa do produto de tais matrizes, em ordem contrária. Logo, temos que

$$\mathbf{A} = \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1} = (\mathbf{E}_k \cdots \mathbf{E}_2 \mathbf{E}_1)^{-1},$$

ou seja, \mathbf{A} é invertível. □

Um exemplo da redução citada no item d) do teorema acima pode ser vista a seguir

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

Provado o TFME, podemos afirmar que o algoritmo descrito no Capítulo 1, o qual recebe o nome de *método de Gauss-Jordan*^{7 8} pode ser aplicado a qualquer matriz \mathbf{A} invertível. De modo geral, dada uma matriz \mathbf{A} , podemos escrever a sua *matriz escalonada*, isto é, uma matriz \mathbf{B} equivalente por linhas a \mathbf{A} dada por

$$\mathbf{B} = \mathbf{E}_1 \cdots \mathbf{E}_k \mathbf{A},$$

a qual atende aos seguintes critérios:

- a) Se \mathbf{B} possui uma ou mais linhas nulas, estas devem ser suas últimas linhas no sentido de cima para baixo.
- b) Em cada linha não-nula, a primeira entrada não-nula está imediatamente à esquerda de qualquer uma das outras entradas não-nulas que estejam abaixo desta.

Podemos então concluir que, se uma matriz é invertível, então ela é escalonável e pode ser reduzida a uma matriz identidade. Caso contrário, sua forma escalonada possui no mínimo uma linha nula, a qual é a última das linhas de \mathbf{A} .

⁷Carl Friederich Gauss (1777-1855), foi um brilhante matemático alemão que ficou conhecido como “Príncipe da Matemática” que desenvolveu trabalhos nas mais diversas áreas da matemática, tais como Álgebra, Teoria dos números, Probabilidade, Estatística, Geometria não-euclidianas, Geometria Diferencial entre muitas outras [56].

⁸Wilhelm Jordan (1842-1899) foi um professor alemão que contribuiu para a resolução de sistemas lineares com um método sistemático de substituição reversa [53].

Exemplo 1.18. A matriz escalonada das matrizes

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 4 & 2 & 3 \\ 7 & 1 & 6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 0 & 5 & 1 & 4 \\ 2 & 7 & 4 & 5 \\ 5 & 2 & 1 & 3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}$$

são

$$\mathbf{D} = \begin{bmatrix} 1 & 3 & 5 \\ 0 & -10 & -17 \\ 0 & 0 & 5 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 2 & 7 & 4 & 5 \\ 0 & 5 & 1 & 4 \\ 0 & 0 & \frac{14}{5} & \frac{6}{5} \\ 0 & 0 & 0 & \frac{93}{14} \end{bmatrix} \quad \text{e} \quad \mathbf{F} = \begin{bmatrix} 1 & 2 \\ 0 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

respectivamente.

O processo utilizado para determinar a matriz escalonada de uma matriz dada é o seguinte:

- (1) Se $a_{11} \neq 0$, então multiplica-se cada linha (a partir da segunda) por $-a_{i1}/a_{11}$, da primeira coluna e abaixo de a_{11} sejam todos nulos.
- (2) Se $a_{11} = 0$, então deve-se buscar uma linha cuja entrada a_{i1} seja não-nula, com $i \geq 2$, e aplicar uma troca de linhas.
- (3) Se porém, todas as entradas da primeira coluna forem nulas, então passa-se à coluna mais próxima onde exista pelo menos uma entrada $\neq 0$. Se tal entrada for a primeira da coluna, opera-se como no passo 1. Do contrário, segue-se o passo 2.
- (4) Quando todas as entradas abaixo daquela citada no passo 3 forem iguais a 0, passa-se então para a próxima linha e aplica-se os passos anteriores.

Tal processo recebe o nome de *eliminação Gaussiana*, o qual possui grande utilidade na determinação das soluções (quando existem) de um sistema linear. Podemos então notar que o método de Gauss-Jordan é a aplicação da eliminação Gaussiana sucessivas vezes, de modo que somente as entradas da diagonal principal sejam não-nulas.

Exemplo 1.18. Vamos aplicar o processo de eliminação Gaussiana à matriz \mathbf{B} do exemplo anterior, a fim de demonstrar sua aplicabilidade na prática.

$$\begin{aligned} \mathbf{B} = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 0 & 5 & 1 & 4 \\ 2 & 7 & 4 & 5 \\ 5 & 2 & 1 & 3 \end{bmatrix} & \xrightarrow{L_1 \leftrightarrow L_3} \begin{bmatrix} 2 & 7 & 4 & 5 \\ 0 & 5 & 1 & 4 \\ 0 & 1 & 3 & 2 \\ 5 & 2 & 1 & 3 \end{bmatrix} \xrightarrow{\begin{matrix} L_3 - \frac{1}{5}L_2 \\ L_4 - \frac{5}{2}L_1 \end{matrix}} \begin{bmatrix} 2 & 7 & 4 & 5 \\ 0 & 5 & 1 & 4 \\ 0 & 0 & \frac{14}{5} & \frac{6}{5} \\ 0 & -\frac{31}{2} & -9 & -\frac{19}{2} \end{bmatrix} \\ & \xrightarrow{L_4 + \frac{31}{10}L_2} \begin{bmatrix} 2 & 7 & 4 & 5 \\ 0 & 5 & 1 & 4 \\ 0 & 0 & \frac{14}{5} & \frac{6}{5} \\ 0 & 0 & -\frac{59}{10} & \frac{41}{10} \end{bmatrix} \xrightarrow{L_4 + \frac{59}{28}L_3} \begin{bmatrix} 2 & 7 & 4 & 5 \\ 0 & 5 & 1 & 4 \\ 0 & 0 & \frac{14}{5} & \frac{6}{5} \\ 0 & 0 & 0 & \frac{93}{14} \end{bmatrix} = \mathbf{E} \end{aligned}$$

Finalizamos esta seção com um teorema que nos será útil na seção seguinte, o qual afirma que se uma matriz $n \times n$ é singular, então o seu produto por qualquer outra matriz

de mesma ordem é também singular.

Teorema 1.6. *Seja \mathbf{A} uma matriz singular de ordem n . Tem-se que \mathbf{AB} é também singular, qualquer que seja a matriz \mathbf{B} de mesma ordem.*

Demonstração. Suponhamos que \mathbf{AB} seja invertível e consideremos uma matriz \mathbf{X} , tal que $\mathbf{BX} = \mathbf{0}$. Multiplicando esta última igualdade à esquerda por \mathbf{A} , obtemos $\mathbf{ABX} = \mathbf{A0}$, o que nos leva a

$$\mathbf{ABX} = \mathbf{0},$$

o que por sua vez, sendo \mathbf{AB} singular, implica, de acordo com o TFME, que $\mathbf{X} = \mathbf{0}$ é a única matriz que satisfaz a equação acima. Consequentemente, esta mesma matriz é a única que satisfaz $\mathbf{BX} = \mathbf{0}$ e, novamente pelo TFME, decorre que \mathbf{B} é invertível. Assim, temos que

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

Portanto, como o produto de matrizes invertíveis é também invertível, decorre que \mathbf{A} é não-singular, provando assim o resultado buscado. \square

1.4 Algumas aplicações matriciais

1.4.1 Matrizes de Leslie

As *matrizes de Leslie*⁹, de acordo com MURRAY [49] são um dos meios de estudo populacional, pois estas podem abranger diferentes faixas etárias, tais como jovem e adulto, bem como quantificar mudanças de uma faixa para outra. Tais matrizes compõem o *modelo de Leslie*¹⁰, onde este têm como principal objetivo descrever o crescimento de uma população de fêmeas, seja de animais ou humanos, uma vez que estas podem dar origem a novos indivíduos. O modelo é estruturado da seguinte maneira:

- 1) Define-se a idade máxima da população como sendo Y anos.
- 2) A população é dividida em n faixas etárias.
- 3) Cada faixa etária está associada a um intervalo de tempo conforme o esquema abaixo

Faixa etária	Intervalo de idade
1	$[t_0, t_1)$
2	$[t_1, t_2)$
3	$[t_2, t_3)$
\vdots	\vdots
n	$[t_{n-1}, t_n]$

⁹Patrick Holt Leslie (1900-1974) foi um dos primeiros cientistas a estudar a dinâmica de crescimento populacional utilizando matrizes [1].

¹⁰De acordo com ANTON [3], este é um dos modelos de crescimento populacional mais comumente usado, o qual foi desenvolvido na década de 1940.

onde $t_k = kY/n$ representa o tempo em que termina a faixa etária k e inicia-se a faixa etária $k + 1$.

- 4) Define-se o número de fêmeas em cada uma das n faixas etárias no instante $t = 0$, com t representando o tempo, de modo que tais quantidades são organizadas na matriz a seguir

$$\mathbf{F}_{11} = \begin{bmatrix} f_{11} \\ f_{21} \\ \vdots \\ f_{n1} \end{bmatrix},$$

onde f_{i1} é o número de fêmeas na faixa etária i . Tal matriz é denominada *matriz de distribuição etária inicial*.

- 5) Dependendo da espécie a ser considerada, pode ser que esta possa gerar novos indivíduos em diferentes faixas etárias, de modo que é necessário definir qual é a *taxa de natalidade* $m_i \geq 0$ de fêmeas reproduzidas na faixa etária i , $i = 1, 2, \dots, n$. É necessário que pelo menos um dos m_i seja não nulo, a fim de garantir que haja reprodução da espécie e, nesse caso, dizemos que i é uma *faixa etária fértil*.
- 6) Define-se a *taxa de sobrevivência* $0 < s_j \leq 1$ de fêmeas durante a transição da faixa etária j para a faixa etária $j + 1$, onde $j = 1, 2, \dots, n - 1$.
- 7) No intervalo $[t_0, t_n]$, todas as fêmeas em faixa etária fértil geram, de acordo com suas respectivas médias de reprodução, um número f_{12} de fêmeas na faixa etária 1, ou seja

$$f_{12} = m_1 f_{11} + m_2 f_{21} + \dots + m_n f_{n1}.$$

Ao término desse intervalo, sobreviverão fêmeas de acordo com a taxa de sobrevivência de cada faixa etária., ou seja, de f_{11} , restarão $f_{22} = s_1 f_{11}$, de f_{21} , restarão $f_{32} = s_2 f_{21}$ e assim sucessivamente até $f_{n2} = s_{n-1} f_{(n-1)1}$, de modo que tais valores podem ser organizados na matriz

$$\mathbf{F}_{12} = \begin{bmatrix} f_{12} \\ f_{22} \\ f_{32} \\ \vdots \\ f_{n2} \end{bmatrix} = \begin{bmatrix} m_1 f_{11} + m_2 f_{21} + \dots + m_n f_{n1} \\ s_1 f_{11} \\ s_2 f_{21} \\ \vdots \\ s_{n-1} f_{(n-1)1} \end{bmatrix}.$$

Escrevendo \mathbf{F}_{12} como produto de matrizes, temos que

$$\mathbf{F}_{12} = \begin{bmatrix} f_{12} \\ f_{22} \\ f_{32} \\ \vdots \\ f_{n2} \end{bmatrix} = \begin{bmatrix} m_1 & m_2 & \dots & m_{n-1} & m_n \\ s_1 & 0 & \dots & 0 & 0 \\ 0 & s_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & s_{n-1} & 0 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ \vdots \\ f_{n1} \end{bmatrix},$$

isto é,

$$\mathbf{F}_{12} = \mathbf{L}\mathbf{F}_{11},$$

onde

$$\mathbf{L} = \begin{bmatrix} m_1 & m_2 & \cdots & m_{n-1} & m_n \\ s_1 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & s_{n-1} & 0 \end{bmatrix}$$

é a *matriz de Leslie*. Com um raciocínio análogo ao descrito acima, temos que

$$\begin{aligned} \mathbf{F}_{13} &= \mathbf{L}\mathbf{F}_{12} = \mathbf{L}\mathbf{L}\mathbf{F}_{11} = \mathbf{L}^2\mathbf{F}_{11} \\ \mathbf{F}_{14} &= \mathbf{L}\mathbf{F}_{13} = \mathbf{L}\mathbf{L}^2\mathbf{F}_{11} = \mathbf{L}^3\mathbf{F}_{11} \\ &\vdots \\ \mathbf{F}_{1n} &= \mathbf{L}\mathbf{F}_{1(n-1)} = \mathbf{L}\mathbf{L}^{n-1}\mathbf{F}_{11} = \mathbf{L}^n\mathbf{F}_{11} \end{aligned}$$

Portanto, a matriz de blocos

$$\mathbf{F} = [\mathbf{F}_{11} \quad \mathbf{F}_{12} \quad \cdots \quad \mathbf{F}_{1n}]$$

fornece o número de fêmeas nos n primeiros períodos Y .

Vejam os a seguir um exemplo, a fim de melhor compreender o modelo acima apresentado.

Exemplo 1.20. Consideremos uma colônia de guaxinins, na qual a população fêmea vive no máximo 5 anos e está dividida em 5 faixas etárias, as quais mudam a cada ano, conforme o quadro a seguir:

Faixa etária (em anos)	f_{i1}	m_i	s_i
[0, 1)	40	0	0,6
[1, 2)	30	5	0,4
[2, 3)	15	3	0,2
[3, 4)	10	2	0,1
[4, 5]	5	1	0

onde $i = 1, \dots, 5$. Temos assim que

$$\mathbf{F}_{11} = \begin{bmatrix} 40 \\ 30 \\ 15 \\ 10 \\ 5 \end{bmatrix} \quad \mathbf{L} = \begin{bmatrix} 0 & 5 & 3 & 2 & 1 \\ 0,6 & 0 & 0 & 0 & 0 \\ 0 & 0,4 & 0 & 0 & 0 \\ 0 & 0 & 0,2 & 0 & 0 \\ 0 & 0 & 0 & 0,1 & 0 \end{bmatrix}.$$

Dessa forma, após um ano, a população de fêmeas será dada pela matriz

$$\mathbf{F}_{12} = \mathbf{L}\mathbf{F}_{11} = \begin{bmatrix} 0 & 5 & 3 & 2 & 1 \\ 0,6 & 0 & 0 & 0 & 0 \\ 0 & 0,4 & 0 & 0 & 0 \\ 0 & 0 & 0,2 & 0 & 0 \\ 0 & 0 & 0 & 0,1 & 0 \end{bmatrix} \begin{bmatrix} 40 \\ 30 \\ 15 \\ 10 \\ 5 \end{bmatrix} = \begin{bmatrix} 220 \\ 24 \\ 12 \\ 3 \\ 1 \end{bmatrix}.$$

Ao longo de 5 anos, temos que

$$\mathbf{F}_{12} = \begin{bmatrix} 220 \\ 24 \\ 12 \\ 3 \\ 1 \end{bmatrix}, \quad \mathbf{F}_{13} = \begin{bmatrix} 163 \\ 132 \\ 9,6 \\ 2,4 \\ 0,3 \end{bmatrix}, \quad \mathbf{F}_{14} = \begin{bmatrix} 693,9 \\ 97,8 \\ 52,8 \\ 1,9 \\ 0,2 \end{bmatrix}, \quad \mathbf{F}_{15} = \begin{bmatrix} 651,48 \\ 416,34 \\ 39,12 \\ 10,56 \\ 0,19 \end{bmatrix},$$

$$\mathbf{F}_{16} = \begin{bmatrix} 2220,37 \\ 390,89 \\ 166,54 \\ 7,82 \\ 1,06 \end{bmatrix}.$$

Observando o gráfico da Figura 1.4 abaixo apenas para os 10 primeiros anos, é possível perceber como o crescimento populacional se dá de forma exponencial

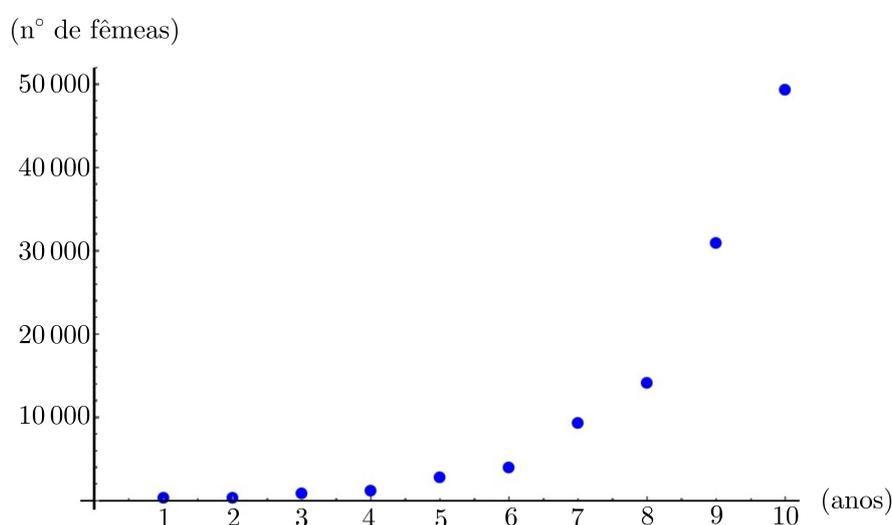


Figura 1.4: Crescimento da população de guaxinins ao longo de 10 anos.

Fonte: O autor

Agora observemos as taxas de crescimento populacional anual ao longo de 50 anos, em ordem cronológica

160%, 18,2%, 175,52%, 32,01%, 149,32%, 43,3%, 131,57%,
52,45%, 119,11%, 59,75%, 110,2%, 65,51%, 103,72%, 70,01%,
98,96%, 73,5%, 95,44%, 76,2%, 92,82%, 78,27%, 90,86%,
79,85%, 89,39%, 81,06%, 88,29%, 81,98%, 87,45%, 82,68%,
86,82%, 83,21%, 86,35%, 83,69%, 86%, 83,93%, 85,73%,
84,16%, 85,52%, 84,33%, 85,37%, 84,47%, 85,25%, 84,57%,
85,16%, 84,65%, 85,10%, 84,7%, 85,04%, 84,75%, 85%,
84,79.%

É possível perceber que, nos anos iniciais, há uma grande divergência de aumento populacional anual entre um ano e outro, porém, à medida que os anos avançam os valores começam a ficar mais próximos, de modo que o crescimento populacional anual

médio é aproximadamente 84,98%. O gráfico abaixo nos fornece uma melhor visão deste fato

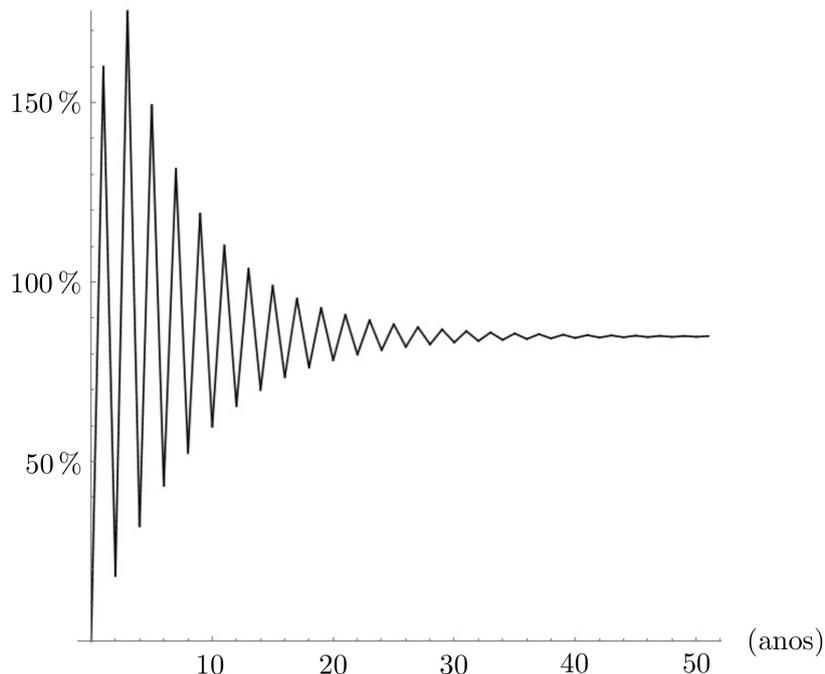


Figura 1.5: Estabilização do crescimento populacional ao longo de 50 anos.

Fonte: O autor

Observemos ainda que, nesse exemplo, o número médio de filhas que uma única guaxinim fêmea é obtido da seguinte forma:

- A fêmea nasce e, durante seu primeiro ano de vida não é capaz de reproduzir, pois $m_1 = 0$.
- Se ela sobreviver, reproduzirá $m_2 s_1 = 5 \cdot 0,6 = 3$ fêmeas, pois há 60% de chance dela sobreviver. Até aqui temos que o *valor esperado* de fêmeas reproduzidas durante seu tempo de vida (até o dado momento) é $0 + 3 = m_1 + m_2 s_1$.
- No terceiro ano de vida, caso sobreviva, produzirá $0,4 \cdot 0,6 \cdot 3 = 0,72$ fêmeas, pois há $0,4 \cdot 0,6 = 0,24$ de probabilidade que ela sobreviva até o terceiro ano. Dessa forma, temos que, até agora, ela produziu, em média, $m_1 + m_2 s_1 + m_3 s_1 s_2$.
- Seguindo o raciocínio descrito acima, temos que, atingindo o quarto ano de vida, essa fêmea produzirá $0,2 \cdot 0,4 \cdot 0,6 \cdot 2 = 0,096$ fêmeas e no quinto ano, $0,1 \cdot 0,2 \cdot 0,4 \cdot 0,6 \cdot 1 = 0,0048$, totalizando assim

$$0 + 3 + 0,72 + 0,096 + 0,0048 = 3,8202$$

fêmeas reproduzidas, em média, ao longo de toda a vida do animal. Pondo os dados em termos de m_i e s_j , com $i = 1, \dots, 5$ e $j = 1, \dots, 4$, obtemos

$$m_1 + m_2 s_1 + m_3 s_1 s_2 + m_4 s_1 s_2 s_3 + m_5 s_1 s_2 s_3 s_4.$$

Podemos assim definir a média de filhas fêmeas de uma única fêmea ao longo de n faixas etárias como sendo

$$m_1 + m_2s_1 + m_3s_1s_2 + \dots + m_ns_1 \cdots s_{n-1}.$$

A expressão acima é denominada **taxa líquida de reprodução** da população fêmea, a qual será representada por T_L .

Observemos agora que as matrizes de Leslie podem apresentar comportamento periódico, desde que exista uma relação entre a taxa de reprodução e as taxas de sobrevivência. Para iniciar, consideremos a matriz de Leslie

$$\mathbf{L} = \begin{bmatrix} 0 & m \\ s_1 & 0 \end{bmatrix}$$

e notemos que

$$\mathbf{L}^2 = \begin{bmatrix} ms_1 & 0 \\ 0 & ms_1 \end{bmatrix} = ms_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = ms_1 \mathbf{I}_2.$$

Para a matriz

$$\mathbf{L} = \begin{bmatrix} 0 & 0 & m \\ s_1 & 0 & 0 \\ 0 & s_2 & 0 \end{bmatrix}$$

temos que

$$\mathbf{L}^3 = \begin{bmatrix} ms_1s_2 & 0 & 0 \\ 0 & ms_1s_2 & 0 \\ 0 & 0 & ms_1s_2 \end{bmatrix} = ms_1s_2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = ms_1s_2 \mathbf{I}_3.$$

Seguindo essa sequência, obtemos, para algum $n \in \mathbb{N}$

$$\mathbf{L}^n = ms_1 \cdots s_{n-1} \mathbf{I}_n.$$

Daí, para que seja $\mathbf{L} = \mathbf{I}_n$, devemos ter

$$m = \frac{1}{s_1 \cdots s_{n-1}}.$$

Nesse caso, seja qual for nossa matriz inicial \mathbf{F}_{11} , teremos

$$\mathbf{F}_{11} = \mathbf{F}_{1(k+1)} = \mathbf{F}_{1(2k+1)} = \mathbf{F}_{1(3k+1)} = \dots$$

significando assim que o número de fêmeas de uma população que constituem uma matriz de Leslie sob as condições acima, oscila durante k unidades de tempo e em seguida retorna ao valor inicial, onde estas oscilações, de acordo com ANTON [3], recebem o nome de **ondas populacionais**. Quando uma população volta a atingir o mesmo valor inicial de modo periódico, diz-se que ela possui **crescimento populacional nulo**. Como nesse caso temos $m_1 = \dots = m_{n-1} = 0$ e $m_n = m$, decorre que

$$T_L = ms_1 \cdots s_{n-1} = 1,$$

ou seja, o crescimento populacional é nulo se, e somente se, $T_L = 1$.

Exemplo 1.21. A *semelparidade* é, de acordo com HILL [30], a condição em que indivíduos são fisiologicamente programados para se reproduzirem uma única vez em toda sua

vida. Esse é o caso do Rei Salmão (*Oncorhynchus tshawytscha*) que, segundo QUINN [55], é o maior dentre as espécies de salmão do Oceano Pacífico e possui uma taxa de reprodução de aproximadamente 5400 ovos. Já o seu ciclo de vida pode atingir mais de 4 anos, aponta Lewis [41]. Dividindo o tempo de vida dessa espécie em 4 faixas etárias $[0, 1)$, $[1, 2)$, $[2, 3)$ e $[3, 4)$ onde a passagem de uma para outra se dá a cada ano, consideremos as seguintes matrizes de Leslie \mathbf{S} e de distribuição etária inicial \mathbf{F}_{11}

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & 0 & 5400 \\ \frac{1}{8} & 0 & 0 & 0 \\ 0 & \frac{1}{25} & 0 & 0 \\ 0 & 0 & \frac{1}{27} & 0 \end{bmatrix} \quad \text{e} \quad \mathbf{F}_{11} = \begin{bmatrix} 40 \\ 30 \\ 20 \\ 10 \end{bmatrix}.$$

tais que

$$\mathbf{F}_{12} = \begin{bmatrix} 54000 \\ 5 \\ 1,2 \\ 0,74 \end{bmatrix}, \quad \mathbf{F}_{13} = \begin{bmatrix} 4000 \\ 6750 \\ 0,2 \\ 0,04 \end{bmatrix}, \quad \mathbf{F}_{14} = \begin{bmatrix} 240 \\ 500 \\ 270 \\ 0,007 \end{bmatrix}, \quad \mathbf{F}_{15} = \begin{bmatrix} 40 \\ 30 \\ 20 \\ 10 \end{bmatrix}, \dots$$

Supondo, apenas para uma análise gráfica, que essa população mantivesse o mesmo ritmo reprodutivo durante 20 anos, o gráfico obtido seria o que se pode observar na Figura 1.6 abaixo

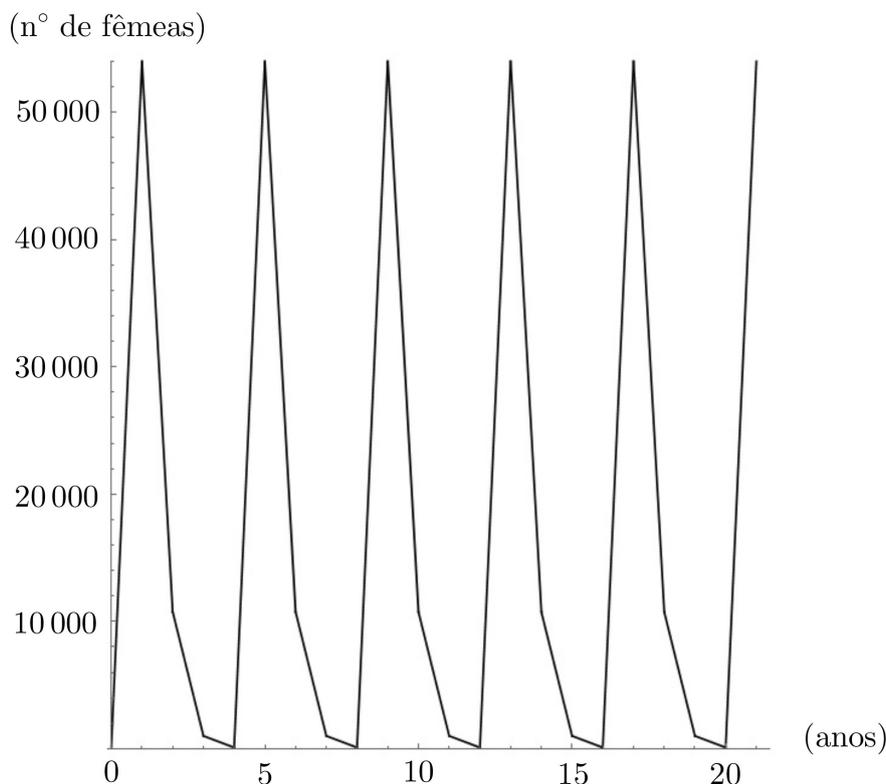


Figura 1.6: Periodicidade do crescimento populacional do Rei Salmão.

Fonte: O autor

mostrando assim que a população fêmea oscila durante 4 anos e retorna ao valor inicial.

1.4.2 Criptografia

A Criptografia, conforme [45], é o estudo de técnicas matemáticas relacionadas aos aspectos de segurança da informação tais como, confidencialidade, integridade de dados, autenticação de entidades e autenticação de origem de dados [45]. De acordo com ISMAIL [32], a *cifra de Hill*¹¹ é um algoritmo que consiste em tomar m letras e substituí-las por m letras de texto cifrado. Vamos aqui tomar os números correspondentes às posições das letras no alfabeto para substituí-las como podemos ver abaixo

A	B	C	D	E	F	G	H	I	J	K	L	M
1	2	3	4	5	6	7	8	9	10	11	12	13
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
14	15	16	17	18	19	20	21	22	23	24	25	26

O objetivo de fazer isso é criptografar uma mensagem por meio de uma *matriz de cifragem*, a qual deve ser invertível, a fim de que seja possível decodificar a mensagem utilizando uma *matriz de decifragem*, isto é, a matriz inversa da matriz de cifragem. Dependendo das palavras existentes na mensagem, pode ser necessário utilizar mais números para representar, por exemplo, letras acentuadas, vírgulas, exclamações, etc. Vamos codificar a mensagem “SÓ SEI QUE NADA SEI” utilizando a matriz de cifragem

$$\mathbf{M}_C = \begin{bmatrix} -1 & 1 & 3 \\ -1 & 4 & 1 \\ 1 & 1 & -4 \end{bmatrix}.$$

Vamos dividir a frase em blocos de 3. Os espaços entre palavras serão representado por 0 e a letra Ó será representada por 27. Dessa forma, nosso quadro de codificação será

	A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12	13
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Ó
14	15	16	17	18	19	20	21	22	23	24	25	26	27

Daí, escrevendo a frase dada conforme o esquema acima, obtemos

S	Ó		S	E	I		Q	U	E		N
19	27	0	19	5	9	0	17	21	5	0	14

A	D	A		S	E	I		
1	4	1	0	19	5	9	0	0

Organizando os blocos em matrizes 3×1 , tem-se

$$\mathbf{P}_1 = \begin{bmatrix} 19 \\ 27 \\ 0 \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} 19 \\ 5 \\ 9 \end{bmatrix}, \quad \mathbf{P}_3 = \begin{bmatrix} 0 \\ 17 \\ 21 \end{bmatrix}, \quad \mathbf{P}_4 = \begin{bmatrix} 5 \\ 0 \\ 14 \end{bmatrix}$$

$$\mathbf{P}_5 = \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}, \quad \mathbf{P}_6 = \begin{bmatrix} 0 \\ 19 \\ 5 \end{bmatrix}, \quad \mathbf{P}_7 = \begin{bmatrix} 9 \\ 0 \\ 0 \end{bmatrix}.$$

¹¹Lester Sanders Hill (1890-1961) foi um matemático americano que ficou conhecido por seus trabalhos em criptografia e criptoanálise [64]. Hill publicou, em 1929, a cifra que leva seu nome em seu livro intitulado “*Cryptography in an algebraic alphabet*” [31].

Daí, multiplicando cada uma destas por \mathbf{M}_C , obtém-se matrizes $\mathbf{Q}_i = \mathbf{M}_C \mathbf{P}_i$, com $i = 1, \dots, 7$, tais que

$$\mathbf{Q}_1 = \begin{bmatrix} 8 \\ 89 \\ 46 \end{bmatrix}, \quad \mathbf{Q}_2 = \begin{bmatrix} 13 \\ 10 \\ -12 \end{bmatrix}, \quad \mathbf{Q}_3 = \begin{bmatrix} 80 \\ 89 \\ -67 \end{bmatrix}, \quad \mathbf{Q}_4 = \begin{bmatrix} 37 \\ 9 \\ -51 \end{bmatrix}$$

$$\mathbf{Q}_5 = \begin{bmatrix} 6 \\ 16 \\ 1 \end{bmatrix}, \quad \mathbf{Q}_6 = \begin{bmatrix} 34 \\ 81 \\ -1 \end{bmatrix}, \quad \mathbf{Q}_7 = \begin{bmatrix} -9 \\ -9 \\ 9 \end{bmatrix}.$$

Agora note que algumas entradas das matrizes imediatamente acima são menores que 0 ou maiores que 28. Logo, para obtermos seus correspondentes no quadro de codificação, devemos formar novas matrizes, onde as entradas destas serão os restos das divisões das entradas anteriores por 28. Por exemplo,

$$\begin{bmatrix} 8 \\ 89 \\ 46 \end{bmatrix} \text{ corresponde a } \begin{bmatrix} 8 \\ 5 \\ 18 \end{bmatrix}, \text{ que corresponde a } \begin{bmatrix} \mathbf{H} \\ \mathbf{E} \\ \mathbf{R} \end{bmatrix},$$

pois

$$\begin{aligned} 8 &= 0 \cdot 28 + \mathbf{8} \text{ e } \mathbf{H} \text{ está na posição } \mathbf{8} \\ 89 &= 3 \cdot 28 + \mathbf{5} \text{ e } \mathbf{E} \text{ está na posição } \mathbf{5} \\ 46 &= 1 \cdot 28 + \mathbf{18} \text{ e } \mathbf{R} \text{ está na posição } \mathbf{18} \end{aligned}$$

Seguindo a ideia acima, obtemos as seguintes correspondências

$$\begin{aligned} \mathbf{Q}_1 &= \begin{bmatrix} 8 \\ 89 \\ 46 \end{bmatrix} \longrightarrow \mathbf{R}_1 = \begin{bmatrix} 8 \\ 5 \\ 18 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{H} \\ \mathbf{E} \\ \mathbf{R} \end{bmatrix} \\ \mathbf{Q}_2 &= \begin{bmatrix} 13 \\ 10 \\ -12 \end{bmatrix} \longrightarrow \mathbf{R}_2 = \begin{bmatrix} 13 \\ 10 \\ 16 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{M} \\ \mathbf{J} \\ \mathbf{P} \end{bmatrix} \\ \mathbf{Q}_3 &= \begin{bmatrix} 80 \\ 89 \\ -67 \end{bmatrix} \longrightarrow \mathbf{R}_3 = \begin{bmatrix} 24 \\ 5 \\ 17 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{X} \\ \mathbf{E} \\ \mathbf{Q} \end{bmatrix} \\ \mathbf{Q}_4 &= \begin{bmatrix} 37 \\ 9 \\ -51 \end{bmatrix} \longrightarrow \mathbf{R}_4 = \begin{bmatrix} 9 \\ 9 \\ 5 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \\ \mathbf{E} \end{bmatrix} \\ \mathbf{Q}_5 &= \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} \longrightarrow \mathbf{R}_5 = \begin{bmatrix} 6 \\ 16 \\ 1 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{F} \\ \mathbf{P} \\ \mathbf{A} \end{bmatrix} \\ \mathbf{Q}_6 &= \begin{bmatrix} 34 \\ 81 \\ -1 \end{bmatrix} \longrightarrow \mathbf{R}_6 = \begin{bmatrix} 6 \\ 25 \\ 27 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{F} \\ \mathbf{Y} \\ \mathbf{O} \end{bmatrix} \\ \mathbf{Q}_7 &= \begin{bmatrix} -9 \\ -9 \\ 9 \end{bmatrix} \longrightarrow \mathbf{R}_7 = \begin{bmatrix} 19 \\ 19 \\ 9 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{S} \\ \mathbf{S} \\ \mathbf{I} \end{bmatrix} \end{aligned}$$

Portanto, a mensagem codificada adquire a forma

“HERMJXPXEQIIEFPAYÓSSI”

Logo, para decifrar a mensagem, utiliza-se a matriz de decifragem $\mathbf{M}_D = \mathbf{M}_C^{-1}$, onde

$$\mathbf{M}_D = \begin{bmatrix} 17 & -7 & 11 \\ 3 & -1 & 2 \\ 5 & -2 & 3 \end{bmatrix}.$$

Daí, dividindo a mensagem codificada em blocos de 3, o que acaba por fornecer as matrizes \mathbf{R}_i , basta fazer as multiplicações $\mathbf{M}_D \mathbf{R}_i$, o que fornecerá matrizes \mathbf{S}_i , as quais serão correspondentes à matrizes \mathbf{P}_i . Por exemplo, observe que

$$\begin{aligned} \mathbf{S}_1 = \mathbf{M}_D \mathbf{R}_1 &= \begin{bmatrix} 299 \\ 55 \\ 84 \end{bmatrix} \longrightarrow \mathbf{P}_1 = \begin{bmatrix} 19 \\ 27 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} \text{S} \\ \text{Ó} \\ \end{bmatrix} \\ \mathbf{S}_2 = \mathbf{M}_D \mathbf{R}_2 &= \begin{bmatrix} 327 \\ 61 \\ 93 \end{bmatrix} \longrightarrow \mathbf{P}_2 = \begin{bmatrix} 19 \\ 5 \\ 9 \end{bmatrix} \longrightarrow \begin{bmatrix} \text{S} \\ \text{E} \\ \text{I} \end{bmatrix} \\ \mathbf{S}_3 = \mathbf{M}_D \mathbf{R}_3 &= \begin{bmatrix} 560 \\ 101 \\ 161 \end{bmatrix} \longrightarrow \mathbf{P}_3 = \begin{bmatrix} 0 \\ 17 \\ 21 \end{bmatrix} \longrightarrow \begin{bmatrix} \text{Q} \\ \text{U} \\ \end{bmatrix} \\ \mathbf{S}_4 = \mathbf{M}_D \mathbf{R}_4 &= \begin{bmatrix} 145 \\ 28 \\ 42 \end{bmatrix} \longrightarrow \mathbf{P}_4 = \begin{bmatrix} 5 \\ 0 \\ 14 \end{bmatrix} \longrightarrow \begin{bmatrix} \text{E} \\ \text{N} \\ \end{bmatrix} \\ \mathbf{S}_5 = \mathbf{M}_D \mathbf{R}_5 &= \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} \longrightarrow \mathbf{P}_5 = \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} \longrightarrow \begin{bmatrix} \text{A} \\ \text{D} \\ \text{A} \end{bmatrix} \\ \mathbf{S}_6 = \mathbf{M}_D \mathbf{R}_6 &= \begin{bmatrix} 224 \\ 47 \\ 61 \end{bmatrix} \longrightarrow \mathbf{P}_5 = \begin{bmatrix} 0 \\ 19 \\ 5 \end{bmatrix} \longrightarrow \begin{bmatrix} \text{S} \\ \text{E} \\ \end{bmatrix} \\ \mathbf{S}_7 = \mathbf{M}_D \mathbf{R}_7 &= \begin{bmatrix} 289 \\ 56 \\ 84 \end{bmatrix} \longrightarrow \mathbf{P}_5 = \begin{bmatrix} 9 \\ 0 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} \text{I} \\ \end{bmatrix} \end{aligned}$$

o que acaba por fornecer a mensagem original.

Observe que podemos dificultar o processo de cifragem com auxílio de funções, por exemplo, por meio de combinação dos processos. Consideremos, como ilustração, uma função afim $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = ax + b$, $a, b \in \mathbb{R}$ com $a \neq 0$. Dessa forma, para cada valor atribuído a cada um dos caracteres utilizados no alfabeto do qual se quer enviar a mensagem, toma-se sua imagem pela função escolhida, a fim de determinar o caractere cifrado correspondente. Por exemplo, se considerarmos a tabela do exemplo anterior e a função $m : \mathbb{R} \rightarrow \mathbb{R}$, $m(x) = 3x + 5$, temos que a letra “S” (com valor 19) será representada por “F” (com valor 6), uma vez que

$$m(19) = 3 \cdot 19 + 5 = 62 = 2 \cdot 28 + 6.$$

Com posse de tal ideia, seja α o número de caracteres a serem utilizados. Daí, temos que o remetente cifrará a mensagem por meio de

$$f(x) = ax + b = Q \cdot \alpha + r,$$

onde Q é o quociente da divisão de $f(x)$ por $\alpha \neq 0$ e r é o resto. Note que, sendo $x \leq \alpha$, decorre que $Q \leq a$. Por conseguinte, o destinatário receberá a função g , inversa de f , dada por

$$g(x) = \frac{x - b}{a},$$

para decodificar a mensagem cifrada. Porém ele não receberá as imagens de m correspondentes a cada caractere, mas sim os restos. Assim, nem todos os valores recebidos poderão ser calculados por meio de ng de maneira direta. Como ilustração disso, note que a inversa da função m é $n : \mathbb{R} \rightarrow \mathbb{R}$, com

$$n(x) = \frac{x - 5}{3}.$$

Calculando $n(6)$, obtemos $\frac{1}{3}$ como resposta, valor este que não está relacionado a nenhum dos caracteres do esquema utilizada. Para contornar o problema, devemos tomar a imagem de $6 + 28k$, pois foram utilizados 28 caracteres e k representa o número de “voltas” dadas no alfabeto. Assim

$$n(6 + 28k) = \frac{6 + 28k - 5}{3} = \frac{28k + 1}{3}.$$

Notemos então que o único resultado inteiro de $n(6 + 28k)$ se dá quando $k = 2$, resultando em 19 (compreendido entre 0 e 28).

Generalizando a ideia acima, devemos tomar a imagem de $r + \alpha \cdot Q$ por g , de modo a obtermos

$$g(Q \cdot \alpha + r) = \frac{Q \cdot \alpha + r - b}{a}$$

Vamos agora cifrar a mensagem “SÓ SEI QUE NADA SEI” utilizando a função m , dada por $m(x) = 3x + 5$.

Letra	x	$m(x)$	$Q \cdot \alpha + r$	Letra criptografada
S	19	62	$2 \cdot 28 + 6$	F
Ó	27	86	$3 \cdot 28 + 2$	B
	0	5	$0 \cdot 28 + 5$	E
E	5	20	$0 \cdot 28 + 20$	T
I	9	32	$1 \cdot 28 + 4$	D
Q	17	56	$2 \cdot 28 + 0$	
U	21	68	$2 \cdot 28 + 12$	L
N	14	47	$1 \cdot 28 + 19$	S
A	1	8	$0 \cdot 28 + 8$	H
D	4	17	$0 \cdot 28 + 17$	Q

Logo a mensagem criptografada é

“FBEFTDE LTESHQHEFTD”.

Para descriptografá-la, utilizaremos a função n , com

$$n(x) = \frac{x - 5}{3}.$$

Já sabemos que “F” corresponde a “S” pelo que foi visto acima. A letra “B”, por sua vez, possui valor 2. Assim

$$n(28Q + 2) = \frac{28Q + 2 - 5}{3} = \frac{28Q}{3} - 1.$$

Segue então que $n(28Q + 2) = -1$ para $Q = 0$, porém -1 não corresponde a nenhum caractere da tabela acima. Por outro lado $n(28Q + 2) = 27$ para $Q = 3$, fornecendo, portanto, a letra “Ó” como correspondente de “B”.

Prosseguindo de maneira análoga para os demais caracteres, segue que

- “E” tem valor 5:

$$\begin{aligned} n(28Q + 5) &= 0, & \text{para } Q = 0 &\Rightarrow \text{“E”} \longrightarrow \text{Espaço} \\ n(28Q + 5) &= 28, & \text{para } Q = 3 &\Rightarrow \text{“E”} \longrightarrow \text{“Ó”} \end{aligned}$$

- “T” tem valor 20:

$$\begin{aligned} n(28Q + 20) &= 5, & \text{para } Q = 0 &\Rightarrow \text{“T”} \longrightarrow \text{“E”} \\ n(28Q + 20) &= 33, & \text{para } Q = 3 &\Rightarrow \text{Sem caractere correspondente} \end{aligned}$$

- “D” tem valor 4:

$$n(28Q + 4) = 9, \quad \text{para } Q = 1 \Rightarrow \text{“D”} \longrightarrow \text{“I”}$$

- “Espaço” tem valor 0:

$$n(28Q + 0) = 17, \quad \text{para } Q = 2 \Rightarrow \text{“Espaço”} \longrightarrow \text{“Q”}$$

- “L” tem valor 12:

$$n(28Q + 12) = 21, \quad \text{para } Q = 2 \Rightarrow \text{“L”} \longrightarrow \text{“U”}$$

- “S” tem valor 19:

$$n(28Q + 19) = 14, \quad \text{para } Q = 1 \Rightarrow \text{“S”} \longrightarrow \text{“N”}$$

- “H” tem valor 8:

$$\begin{aligned} n(28Q + 8) &= 1, & \text{para } Q = 0 &\Rightarrow \text{“H”} \longrightarrow \text{“A”} \\ n(28Q + 8) &= 29, & \text{para } Q = 3 &\Rightarrow \text{Sem caractere correspondente} \end{aligned}$$

- “Q” tem valor 17:

$$\begin{aligned} n(28Q + 17) &= 4, & \text{para } Q = 0 &\Rightarrow \text{“Q”} \longrightarrow \text{“D”} \\ n(28Q + 17) &= 32, & \text{para } Q = 3 &\Rightarrow \text{Sem caractere correspondente} \end{aligned}$$

Daí, fazendo as respectivas correspondências na ordem dos caracteres da mensagem criptografada, percebe-se que a letra “E” de fato equivale a um espaço entre palavras e não à letra “O”. Portanto, podemos observar o resultado no quadro abaixo

F	B	E	F	T	D	E		L	T	E	S	H	Q	H	E	F	T	D
S	Ó		S	E	I		Q	U	E		N	A	D	A		S	E	I

Note ainda que, nos casos sem caractere correspondente, basta aplicar a divisão do resultado por α novamente, e o resto irá fornecer o caractere descryptografado. Por exemplo, a letra “T” forneceu 33 como resultado para $Q = 3$, mas, se dividirmos 33 por 28, obtemos resto 5, o qual corresponde à letra “E”.

Vejamos um caso adicional, agora utilizando funções quadráticas. Sendo assim, tomemos a função $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = ax^2 + bx + c$, com a, b e c sendo reais e $a \neq 0$. Tomando $y = f(x)$ temos que f^{-1} é dada por

$$f^{-1}(x) = \frac{-b \pm \sqrt{b^2 - 4(c - y)}}{2a}.$$

Utilizando o mesmo raciocínio do exemplo anterior, escreveremos

$$y = f(x) = Q\alpha + r.$$

Segue daí que

$$f^{-1}(x) = \frac{-b \pm \sqrt{b^2 - 4c + 4Q\alpha + 4r}}{2a}.$$

Criptografaremos a mensagem “SÓ SEI QUE NADA SEI” por intermédio de $f(x) = x^2 - 5x + 6$, obtemos o quadro a seguir

Letra	x	$f(x)$	$Q \cdot \alpha + r$	Letra criptografada
S	22	272	$9 \cdot 28 + 20$	T
Ó	27	600	$21 \cdot 28 + 12$	L
	0	6	$0 \cdot 28 + 6$	F
E	5	6	$0 \cdot 28 + 6$	F
I	9	42	$1 \cdot 28 + 14$	N
Q	17	210	$7 \cdot 28 + 14$	N
U	21	342	$12 \cdot 28 + 6$	F
N	14	132	$4 \cdot 28 + 20$	T
A	1	2	$0 \cdot 28 + 2$	B
D	4	2	$0 \cdot 28 + 2$	B

a qual nos fornece a mensagem cifrada

“TLFTFNFNFFF'TBBBFTFN”

Agora, sabendo que o vértice da parábola que representa f é $(\frac{5}{4}, -\frac{1}{4})$, decorre que

$$D(f) = \left(-\infty, \frac{5}{2}\right] \cup \left[\frac{5}{2}, +\infty\right) \quad \text{e} \quad Im(f) = \left[-\frac{1}{4}, +\infty\right)$$

Assim, temos que o $D(f^{-1}) = [-\frac{1}{4}, +\infty)$. Daí, podemos observar que os valores atribuídos aos caracteres (de 0 a 28) pertencem a $D(f^{-1})$ e, além disso, a parábola possui a propriedade de simetria, o que evita o risco de, durante o processo de decifragem, obter-se um valor sem correspondentes. Notemos então que

$$f^{-1}(x) = \frac{5 \pm \sqrt{1 + 112Q + 4r}}{2}.$$

Para a decifragem, tomaremos

$$f^{-1}(x) = \frac{5 + \sqrt{1 + 112Q + 4r}}{2},$$

a fim de evitar valores negativos, aos quais deve-se somar 28 para obter caracteres correspondentes. Logo, tomando $g(x) = f^{-1}(x)$, temos que

- “T” tem valor 20:

$$\begin{aligned} g(20) &= 7, & \text{para } Q = 0 & \Rightarrow \text{“T”} \longrightarrow \text{“G”} \\ g(20) &= 14, & \text{para } Q = 4 & \Rightarrow \text{“T”} \longrightarrow \text{“N”} \\ g(20) &= 19, & \text{para } Q = 9 & \Rightarrow \text{“T”} \longrightarrow \text{“S”} \\ g(20) &= 26, & \text{para } Q = 19 & \Rightarrow \text{“T”} \longrightarrow \text{“Z”} \end{aligned}$$

- “L” tem valor 12:

$$\begin{aligned} g(12) &= 6, & \text{para } Q = 0 & \Rightarrow \text{“L”} \longrightarrow \text{“F”} \\ g(12) &= 27, & \text{para } Q = 21 & \Rightarrow \text{“L”} \longrightarrow \text{“Ó”} \end{aligned}$$

- “F” tem valor 6:

$$\begin{aligned} g(6) &= 5, & \text{para } Q = 0 & \Rightarrow \text{“F”} \longrightarrow \text{“E”} \\ g(6) &= 12, & \text{para } Q = 3 & \Rightarrow \text{“F”} \longrightarrow \text{“L”} \\ g(6) &= 21, & \text{para } Q = 12 & \Rightarrow \text{“F”} \longrightarrow \text{“U”} \\ g(6) &= 28 = 1 \cdot 28 + \mathbf{0}, & \text{para } Q = 23 & \Rightarrow \text{“T”} \longrightarrow \text{Espaço} \end{aligned}$$

- “N” tem valor 14:

$$\begin{aligned} g(14) &= 9, & \text{para } Q = 1 & \Rightarrow \text{“N”} \longrightarrow \text{“I”} \\ g(14) &= 16, & \text{para } Q = 6 & \Rightarrow \text{“N”} \longrightarrow \text{“P”} \\ g(14) &= 17, & \text{para } Q = 7 & \Rightarrow \text{“N”} \longrightarrow \text{“Q”} \\ g(14) &= 24, & \text{para } Q = 16 & \Rightarrow \text{“N”} \longrightarrow \text{“X”} \end{aligned}$$

- “B” tem valor 2:

$$\begin{aligned} g(2) &= 4, & \text{para } Q = 0 & \Rightarrow \text{“B”} \longrightarrow \text{“D”} \\ g(2) &= 8, & \text{para } Q = 1 & \Rightarrow \text{“B”} \longrightarrow \text{“H”} \\ g(2) &= 25, & \text{para } Q = 18 & \Rightarrow \text{“B”} \longrightarrow \text{“Y”} \\ g(2) &= 29 = 1 \cdot 28 + \mathbf{1}, & \text{para } Q = 25 & \Rightarrow \text{“B”} \longrightarrow \text{“A”} \end{aligned}$$

Obtemos então a seguinte tabela

T	L	F	T	F	N	F	N	F	F	F	T	B	B	B	F	T	F	N
G	F	E	G	E	I	E	I	E	E	E	G	D	D	D	E	G	E	I
N	Ó	L	N	L	P	L	P	L	L	L	N	H	H	H	L	N	L	P
S	⊗	U	S	U	Q	U	Q	U	U	U	S	Y	Y	Y	U	S	U	Q
Z	⊗		Z		X		X				Z	A	A	A		Z		X

a qual permite visualizar a mensagem descriptografada.

2 Determinantes e Sistemas Lineares

2.1 Determinantes

Neste capítulo iremos conhecer a definição de *determinante* de uma matriz, bem como suas principais propriedades. Embora, de acordo com LARSON [38], a teoria dos determinantes emergiu a partir de padrões encontrados em soluções de *sistemas lineares*, seguiremos aqui a ordem contrária, somente por uma questão de organização. Quanto às aplicações dos determinantes, estas serão melhor estudadas ao longo dos próximos capítulos, pois eles contêm informações importantes para o entendimento de tais aplicações, as quais podem ser, segundo KAZUNGA [33], cálculo de área de polígonos e volume de poliedros, determinar a solução de sistemas lineares com n equações e n incógnitas, inversão de matrizes e determinar se um conjunto de *vetores* é ou não *linearmente dependente*.

O determinante de uma matriz é um número real. Somente as matrizes quadradas possuem determinante, de modo que este é indefinido para matrizes que possuem mais linhas que colunas ou vice-versa. Para melhor organizar as ideias, apresentamos a definição a seguir.

Definição 2.1. *Seja \mathbf{A} uma matriz quadrada de ordem n , com $n \leq 2$. Se $\mathbf{A} = [a_{11}]$, então o *determinante* de \mathbf{A} é igual a a_{11} e escrevemos*

$$\det(\mathbf{A}) = |a_{11}| = a_{11}.$$

Se for

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

então

$$\det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

É importante lembrar que, no caso $n = 1$, não se deve confundir as barras $||$, com a notação de valor absoluto de um número, de modo que, por exemplo, a matriz $\mathbf{X} = [-7]$ é tal que $\det(\mathbf{X}) = -7$ e não 7 , onde este é o valor absoluto de -7 . Vejamos mais alguns exemplos de matrizes e seus respectivos determinantes:

- $\mathbf{A} = [-15] \Rightarrow \det(\mathbf{A}) = |-15| = -15$
- $\mathbf{B} = \begin{bmatrix} 4 & 7 \\ 1 & 3 \end{bmatrix} \Rightarrow \det(\mathbf{B}) = \begin{vmatrix} 4 & 7 \\ 1 & 3 \end{vmatrix} = 4 \cdot 3 - 7 \cdot 1 = 5$
- $\mathbf{C} = \begin{bmatrix} \sin(x) & \cos(x) \\ \sin(x) & \cos(x) \end{bmatrix} \Rightarrow \det(\mathbf{C}) = \begin{vmatrix} \sin(x) & \cos(x) \\ \sin(x) & \cos(x) \end{vmatrix} = \sin^2(x) - \cos^2(x).$

Para as matrizes de ordem $n \geq 3$, vamos precisar das noções de *menor complementar* e *cofator*, onde estas são fornecidas nas definições abaixo.

Definição 2.2. *Seja $\mathbf{A} = [a_{ij}]_{n \times n}$ uma matriz, e seja \mathbf{A}_{ij} a submatriz obtida a partir da supressão da linha e da coluna que contém o elemento a_{ij} . O determinante de \mathbf{A}_{ij} é chamado **menor complementar** de a_{ij} . O **cofator** A_{ij} de a_{ij} será definido por*

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

Exemplo 2.1. Dada a matriz

$$\mathbf{A} = \begin{bmatrix} -2 & 3 & 7 \\ 1 & 9 & 4 \\ -5 & 8 & -1 \end{bmatrix},$$

temos que

- $\det(\mathbf{A}_{11}) = \begin{vmatrix} 9 & 4 \\ 8 & -1 \end{vmatrix} = -41$ é o menor complementar de a_{11} .
 $A_{11} = (-1)^{1+1} \det(\mathbf{A}_{11}) = -41$ é o cofator de a_{11} .
- $\det(\mathbf{A}_{12}) = \begin{vmatrix} 1 & 4 \\ -5 & -1 \end{vmatrix} = 19$ é o menor complementar de a_{12} .
 $A_{12} = (-1)^{1+2} \det(\mathbf{A}_{12}) = -19$ é o cofator de a_{12} .
- $\det(\mathbf{A}_{13}) = \begin{vmatrix} 1 & 9 \\ -5 & 8 \end{vmatrix} = 53$ é o menor complementar de a_{13} .
 $A_{13} = (-1)^{1+3} \det(\mathbf{A}_{12}) = 53$ é o cofator de a_{13} .
- \vdots
- $\det(\mathbf{A}_{33}) = \begin{vmatrix} -2 & 3 \\ 1 & 9 \end{vmatrix} = -21$ é o menor complementar de a_{33} .
 $A_{33} = (-1)^{3+3} \det(\mathbf{A}_{33}) = -21$ é o cofator de a_{33} .

Podemos agora fornecer a definição para determinantes de matrizes quadradas com ordem maior ou igual a 3.

Definição 2.3. *O **determinante** de uma matriz $\mathbf{A} = [a_{ij}]_{n \times n}$ é um número real associado à esta, o qual é definido indutivamente como*

$$\det(\mathbf{A}) = \begin{cases} a_{11} & , n = 1 \\ a_{i1}A_{11} + a_{i2}A_{12} + \cdots + a_{in}A_{1n} & , n > 1 \end{cases}$$

A definição acima é de grande praticidade, uma vez que ela define o determinante de uma dada matriz como sendo a soma dos produtos entre o menor complementar e o cofator de cada entrada da primeira linha.

Exemplo 2.2. Seja \mathbf{A} a matriz do exemplo anterior. Segue, pela Definição 2.3, que

$$\begin{aligned}\det(\mathbf{A}) &= a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \\ &= -2 \cdot (-41) + 3 \cdot (-19) + 7 \cdot 53 \\ &= 396.\end{aligned}$$

Ainda a respeito da definição, perceba que o sinal de cada cofator depende somente da soma dos índices, de modo que quando a soma é par, o sinal é positivo e quando é ímpar, negativo. Com isso, podemos escrever o determinante de uma matriz \mathbf{A} como somas e subtrações alternadas, ou seja,

$$\det(\mathbf{A}) = a_{i1} \det(\mathbf{A}_{i1}) - a_{i2} \det(\mathbf{A}_{i2}) + a_{i3} \det(\mathbf{A}_{i3}) - \dots + (-1)^{i+j} a_{in} \det(\mathbf{A}_{in}).$$

Exemplo 2.3. A troca de linhas de uma matriz identidade fornece uma matriz elementar de tipo I. Note que, seu determinante é igual a -1 , independente da ordem e de quais linhas foram trocadas de posição. A demonstração formal desse fato pode ser encontrada em [53] por meio de indução ou utilizando análise combinatória em [9]. Observemos apenas alguns exemplos para convenvermo-nos sobre tal fato:

$$\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 0 \cdot 0 - 1 \cdot 1 = -1.$$

$$\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 0 \cdot \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} - 0 \cdot \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1.$$

$$\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 0 \cdot \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + 0 \cdot \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} = -1.$$

$$\begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 0 \cdot \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} - 0 \cdot \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} - 0 \cdot \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = -1$$

Quanto ao determinante de uma matriz elementar de tipo II $\mathbf{E}_{L_i \leftrightarrow cL_i}$ é igual a c . A prova disto será dada na segunda propriedade da lista de propriedades da subseção 2.1.1. Já as matrizes elementares de tipo III possuem determinante igual a 1. A prova deste fato também pode ser encontrada nas referências logo acima citadas. Vejamos aqui apenas um exemplo para o caso 2×2 , considerando para tanto, a matriz a seguir $\mathbf{E}_{L_2 \leftrightarrow cL_1 + L_2}$, c constante, de modo que

$$\begin{vmatrix} 1 & 0 \\ k \cdot 1 + 0 & k \cdot 0 + 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ k & 1 \end{vmatrix} = 1 \cdot 1 - 0 \cdot k = 1.$$

Dessa forma, podemos então enunciar o seguinte teorema

Teorema 2.1. *Sejam \mathbf{E}_I , \mathbf{E}_{II} e \mathbf{E}_{III} matrizes elementares de tipos I, II e III, respectivamente. Tem-se que*

- ▷ $\det(\mathbf{E}_I) = -1$.
- ▷ $\det(\mathbf{E}_{II}) = c$, $c \neq 0$.

$$\triangleright \det(\mathbf{E}_{III}) = 1.$$

O teorema acima, juntamente com o Teorema 1.4, nos permite enunciar o seguinte resultado

Teorema 2.2 *Sejam \mathbf{E} uma matriz elementar qualquer e \mathbf{A} uma matriz $n \times n$. Então*

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

Provemos a validade do teorema para o caso 2×2 , sendo análoga a demonstração para casos de ordens superiores. Tomando as matrizes

$$\mathbf{E}_I = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{E}_{II} = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{E}_{III} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \quad \text{e} \quad \mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

temos que

- $\det(\mathbf{E}_I \mathbf{A}) = \begin{vmatrix} c & d \\ a & b \end{vmatrix} = bc - ad$
 $\det(\mathbf{E}_I) \det(\mathbf{A}) = -1 \cdot (ad - bc) = bc - ad$
- $\det(\mathbf{E}_{II} \mathbf{A}) = \begin{vmatrix} ak & bk \\ c & d \end{vmatrix} = akd - bkc = k(ad - bc)$
 $\det(\mathbf{E}_{II}) \det(\mathbf{A}) = k(ad - bc)$
- $\det(\mathbf{E}_{III} \mathbf{A}) = \begin{vmatrix} a + kc & b + kd \\ c & d \end{vmatrix} = ad + cdk - bc - ckd = ad - bc$
 $\det(\mathbf{E}_{III}) \det(\mathbf{A}) = 1 \cdot (ad - bc) = ad - bc$

Como consequência da aplicação sucessivas vezes do Teorema 2.1, temos o

Corolário 2.1. *Dadas matrizes elementares $\mathbf{E}_1, \dots, \mathbf{E}_k$ e uma matriz \mathbf{A} , tem-se que*

$$\det(\mathbf{E}_1 \cdots \mathbf{E}_k \mathbf{A}) = \det(\mathbf{E}_1) \cdots \det(\mathbf{E}_k) \det(\mathbf{A}).$$

O resultado carregado pelo Teorema 2.1, de maneira simples, nos diz que

- ★ *Se duas linhas de uma matriz são trocadas de posição, então seu determinante é multiplicado por -1 .*
- ★ *Se uma das linhas de uma matriz é multiplicada por uma constante $k \in \mathbb{R}$, então seu determinante é multiplicado por k .*
- ★ *Se uma combinação linear for realizada entre as linhas de uma matriz, então seu determinante não se altera.*

A seguir, provaremos um dos principais resultados deste capítulo, o qual é, segundo Poole [30], conhecido como *caracterização de invertibilidade em termos de determinantes*

. **Teorema 2.3.** *Uma matriz $n \times n$ é invertível se, e somente se, $\det(\mathbf{A}) \neq 0$.*

Demonstração. Se \mathbf{A} é invertível, então segue pelo TFME que

$$\mathbf{E}_1 \cdots \mathbf{E}_n \mathbf{A} = \mathbf{I}.$$

daí, pelo Corolário 2.1, temos que

$$\det(\mathbf{E}_1) \cdots \det(\mathbf{E}_k) \det(\mathbf{A}) = 1.$$

Como $\det(\mathbf{E}_i) \neq 0$ com $0, i = 1, \dots, k$, de acordo com o Teorema 2.1, a igualdade acima implica que $\det(\mathbf{A}) \neq 0$. Reciprocamente, seja $\det(\mathbf{A}) \neq 0$. Então, escalonando \mathbf{A} , temos

$$\det(\mathbf{E}_1) \cdots \det(\mathbf{E}_k) \det(\mathbf{A}) = \det(\mathbf{B}),$$

onde \mathbf{B} é a matriz escalonada de \mathbf{A} . Novamente, pelo Teorema 2.1, temos que $\det(\mathbf{E}_i) \neq 0$ para todo i , e como $\det(\mathbf{A}) \neq 0$, decorre que $\det(\mathbf{B}) \neq 0$. Dessa forma, temos que \mathbf{B} não pode possuir linhas nulas, conforme a propriedade 2 da subseção 2.1.1 e portanto $\mathbf{B} = \mathbf{I}$ (se não fosse, então \mathbf{A} não seria invertível, de acordo com o TFME) pela definição de matriz escalonada. Logo,

$$\mathbf{A} = (\mathbf{E}_1 \cdots \mathbf{E}_k)^{-1},$$

ou seja, \mathbf{A} é invertível. □

2.1.1 Propriedades dos determinantes

Esta subseção é voltada à listagem e discussão das principais propriedades dos determinantes, as quais são de grande utilidade para diversos cálculos e demonstrações, como veremos ao longo do presente trabalho. Inciamos assim com o fato de que

1) *O determinante de uma matriz e o determinante de sua transposta são iguais.*

$$\det(\mathbf{A}) = \det(\mathbf{A}^T)$$

Demonstração: Façamos os casos em que \mathbf{A} tem ordem 1, 2, e 3, respectivamente, sendo análogos os demais casos. Quando \mathbf{A} tem ordem 1, temos o caso trivial, de modo que

$$\mathbf{A} = [a_{11}] \Rightarrow \mathbf{A}^T = [a_{11}] \Rightarrow \det(\mathbf{A}) = a_{11} = \det(\mathbf{A}^T).$$

Notemos que se \mathbf{A} possui ordem 2, então

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Rightarrow \det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21} \\ \mathbf{A}^T &= \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \Rightarrow \det(\mathbf{A}^T) = a_{11}a_{22} - a_{12}a_{21}. \end{aligned}$$

Logo, $\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21} = \det(\mathbf{A})^T$. Podemos assim escrever

$$\begin{aligned} \det(\mathbf{A}) &= a_{11} \det(\mathbf{A}_{11}) - a_{12} \det(\mathbf{A}_{12}) \\ &= a_{11} \det(\mathbf{A}_{11}^T) - a_{21} \det(\mathbf{A}_{21}^T) \\ &= \det(\mathbf{A}^T) \end{aligned}$$

Para compreender o que foi dito acima, basta observar que quando transpomos uma matriz, são também transpostas as suas submatrizes. Façamos isso observando o caso em que \mathbf{A} possui ordem 3. Sendo $\mathbf{B} = \mathbf{A}^T$, temos que

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}.$$

Daí, podemos observar que

$$\begin{aligned} \mathbf{A}_{11} &= \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} & \mathbf{B}_{11} &= \begin{bmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} a_{22} & a_{32} \\ a_{23} & a_{33} \end{bmatrix} = \mathbf{A}_{11}^T \\ \mathbf{A}_{12} &= \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} & \mathbf{B}_{12} &= \begin{bmatrix} b_{21} & b_{23} \\ b_{31} & b_{33} \end{bmatrix} = \begin{bmatrix} a_{12} & a_{32} \\ a_{13} & a_{33} \end{bmatrix} = \mathbf{A}_{21}^T \\ \mathbf{A}_{13} &= \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} & \mathbf{B}_{13} &= \begin{bmatrix} b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} = \mathbf{A}_{31}^T \end{aligned}$$

Agora, comparando os determinantes a seguir

$$\begin{aligned} a_{11} \det(\mathbf{A}_{11}) &= a_{11} a_{22} a_{33} - a_{11} a_{23} a_{32} \\ b_{11} \det(\mathbf{A}_{11}^T) &= a_{11} a_{22} a_{33} - a_{11} a_{32} a_{23} \\ -a_{12} \det(\mathbf{A}_{12}) &= -a_{12} a_{21} a_{33} + a_{12} a_{23} a_{31} \\ -b_{12} \det(\mathbf{A}_{21}^T) &= -a_{21} a_{12} a_{33} + a_{21} a_{32} a_{13} \\ a_{13} \det(\mathbf{A}_{13}) &= a_{13} a_{21} a_{32} - a_{13} a_{22} a_{31} \\ b_{13} \det(\mathbf{A}_{31}^T) &= a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13} \end{aligned}$$

podemos observar que

i)

$$a_{11} \det(\mathbf{A}_{11}) = a_{11} \det(\mathbf{A}_{11}^T)$$

ii)

$$\begin{aligned} &-a_{12} \det(\mathbf{A}_{12}) + a_{13} \det(\mathbf{A}_{13}) \\ &= -a_{12}(a_{21} a_{33} - a_{23} a_{31}) + a_{13}(a_{21} a_{32} - a_{22} a_{31}) \end{aligned}$$

iii)

$$\begin{aligned} &-b_{12} \det(\mathbf{A}_{21}^T) + b_{13} \det(\mathbf{A}_{31}^T) \\ &= -a_{21} a_{12} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13} \\ &= -a_{21} a_{12} a_{33} + a_{31} a_{12} a_{23} + a_{21} a_{32} a_{13} - a_{31} a_{22} a_{13} \\ &= -a_{12}(a_{21} a_{33} - a_{23} a_{31}) + a_{13}(a_{21} a_{32} - a_{22} a_{31}) \end{aligned}$$

Segue então, a partir de i), ii) e iii) que

$$\begin{aligned} \det(\mathbf{A}) &= a_{11} \det(\mathbf{A}_{11}) - a_{12} \det(\mathbf{A}_{12}) + a_{13} \det(\mathbf{A}_{13}) \\ &= b_{11} \det(\mathbf{A}_{11}^T) - b_{12} \det(\mathbf{A}_{21}^T) + b_{13} \det(\mathbf{A}_{31}^T) \\ &= b_{11} \det(\mathbf{B}_{11}) - b_{12} \det(\mathbf{B}_{12}) + b_{13} \det(\mathbf{B}_{13}) \\ &= \det(\mathbf{B}) \\ &= \det(\mathbf{A}^T). \end{aligned}$$

Logo, pelo que foi discutido acima, podemos generalizar a ideia para uma matriz quadrada de ordem n , de tal modo que

$$\begin{aligned}\det(\mathbf{A}) &= a_{11} \det(\mathbf{A}_{11}) - a_{12} \det(\mathbf{A}_{12}) + \cdots + (-1)^{1+n} a_{1n} \det(\mathbf{A}_{1n}) \\ &= a_{11} \det(\mathbf{A}_{11}^T) - a_{21} \det(\mathbf{A}_{21}^T) + \cdots + (-1)^{1+n} a_{n1} \det(\mathbf{A}_{n1}^T) \\ &= \det(\mathbf{A}^T).\end{aligned}$$

2) Se \mathbf{A} tiver uma linha ou uma coluna multiplicada por uma constante c , então seu determinante é igual a $c \cdot \det(\mathbf{A})$.

Demonstração. Se for \mathbf{B} a matriz obtida por meio da multiplicação da linha i de \mathbf{A} por c , então o determinante dessa nova matriz é, por definição,

$$\det(\mathbf{B}) = ca_{i1} \mathbf{A}_{i1} + \cdots + ca_{in} \mathbf{A}_{in} = c(a_{i1} \mathbf{A}_{i1} + \cdots + a_{in} \mathbf{A}_{in}) = c \det(\mathbf{A}).$$

Caso seja \mathbf{C} a matriz obtida por meio da multiplicação da coluna j de \mathbf{A} por c , então, pela propriedade anterior, temos que $\det(\mathbf{C}) = \det(\mathbf{C}^T)$ e pela primeira parte provada logo acima, decorre que $\det(\mathbf{C}^T) = c \det(\mathbf{A})$. Como consequência, se \mathbf{A} possuir uma linha ou coluna nula, então seu determinante é igual a 0. Além disso, considerando \mathbf{C} com $c = 1$, a propriedade anterior garante que o *determinante de uma matriz pode ser calculado pela expansão dos cofatores de qualquer uma de suas colunas*. \square

3) **(Teorema de Cauchy-Binet)**^{12 13} O determinante do produto de duas matrizes \mathbf{A} e \mathbf{B} é igual ao produto dos determinantes destas duas matrizes.

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

Demonstração. Se \mathbf{A} é não invertível, então de acordo com o Teorema 2.6, decorre que \mathbf{AB} também não é invertível. Segue daí que, conforme o Teorema 2.9, $\det(\mathbf{AB}) = 0 = \det(\mathbf{A}) \det(\mathbf{B})$. Por outro lado, se for \mathbf{A} não-singular, então podemos escrever, de acordo com o TFME

$$\mathbf{A} = \mathbf{E}_1 \cdots \mathbf{E}_k,$$

onde $\mathbf{E}_1, \dots, \mathbf{E}_k$ são matrizes elementares. Daí, $\mathbf{AB} = \mathbf{E}_1 \cdots \mathbf{E}_k \mathbf{B}$, que pelo Corolário 2.1 nos fornece

$$\begin{aligned}\det(\mathbf{AB}) &= \det(\mathbf{E}_1) \cdots \det(\mathbf{E}_k) \det(\mathbf{B}) \\ &= \det(\mathbf{E}_1 \cdots \mathbf{E}_k) \det(\mathbf{B}) \\ &= \det(\mathbf{A}) \det(\mathbf{B})\end{aligned}$$

\square

4) Se uma matriz quadrada possui duas linhas (ou colunas) iguais, então seu determinante é igual a 0.

¹²Jacques Philippe Marie Binet (1776-1856) foi um matemático francês que desenvolveu trabalhos em Mecânica, Matemática e Astronomia. Em 1812 ele descobriu a fórmula para calcular o determinante da multiplicação de matrizes, feito este que glorificou seu nome mais que qualquer outro de seus trabalhos [59].

¹³Augustin-Louis Cauchy (1789-1857) foi um matemático francês que obteve resultados importantes em diversas áreas, tais como Aritmética, Álgebra, Geometria, Estatística, Mecânica, Análises Real e Complexa e Física Matemática. Publicou diversos resultados sobre determinantes, incluindo a fórmula do produto de determinantes de duas matrizes, a qual ele publicou em 1812 com alto rigor matemático [9].

Demonstração. Suponhamos que as linhas L_i e L_j de uma matriz quadrada \mathbf{A} sejam iguais. Consideremos então uma matriz elementar $\mathbf{E}_{L_i \leftrightarrow L_i - L_j}$ e uma matriz \mathbf{B} tal que

$$\mathbf{B} = \mathbf{E}_{L_i \leftrightarrow L_i - L_j} \mathbf{A}.$$

Assim, temos que \mathbf{B} possui uma linha nula, de modo que $\det(\mathbf{B}) = 0$, pela propriedade 2. Daí,

$$0 = \det(\mathbf{B}) = \det(\mathbf{E}_{L_i \leftrightarrow L_i - L_j} \mathbf{A}) = \det(\mathbf{E}_{L_i \leftrightarrow L_i - L_j}) \det(\mathbf{A}).$$

Como $\det(\mathbf{E}_{L_i \leftrightarrow L_i - L_j}) = 1$, devemos ter $\det(\mathbf{A}) = 0$. A demonstração para o caso de colunas iguais é análogo. \square

5) Se \mathbf{A} é invertível, então

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}.$$

Demonstração. De fato, desde que $\mathbf{A}^{-1} \mathbf{A} = \mathbf{I}$, segue, pelo Teorema de Cauchy-Binet, que

$$\det(\mathbf{A}^{-1} \mathbf{A}) = \det(\mathbf{I}) = 1 \Rightarrow \det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}.$$

\square

6) O determinante de uma matriz diagonal é igual ao produto dos elementos de sua diagonal principal.

Demonstração. Consideremos uma matriz diagonal \mathbf{D} de ordem n , cujas entradas da diagonal principal sejam d_{ii} e seja \mathbf{D}_{11} o menor complementar de d_{11} . Como $d_{1j} = 0$ com $j \geq 2$, decorre que

$$\det(\mathbf{D}) = d_{11} \det(\mathbf{D}_{11}).$$

A matriz \mathbf{D}_{11} é também diagonal. Então sendo \mathbf{D}_{22} o cofator de d_{22} e $d_{2j} \neq 0$, $j \geq 3$, decorre que $\det(\mathbf{D}_{11}) = d_{22} \det(\mathbf{D}_{22})$ e assim

$$\det(\mathbf{D}) = d_{11} d_{22} \det(\mathbf{D}_{22}).$$

Prosseguindo com o argumento acima sucessivas vezes, obtemos

$$\det(\mathbf{D}) = d_{11} d_{22} \cdots d_{(n-1)(n-1)} \det(\mathbf{D}_{nn}),$$

porém, $\mathbf{D}_{nn} = d_{nn}$, o que implica $\det(\mathbf{D}_{nn}) = d_{nn}$ e portanto

$$\det(\mathbf{D}) = d_{11} d_{22} \cdots d_{nn}.$$

\square

7) O determinante de uma matriz de Vandermonde $n \times n$ é dado por

$$\det(\mathbf{V}_{n \times n}) = (x_n - x_1)(x_n - x_2) \cdots (x_j - x_i)(x_j - x_{i-1}) \cdots (x_2 - x_1),$$

onde $1 \leq i < n \leq j$.

Demonstração. Vamos fazer a demonstração dos casos iniciais, sendo as demonstrações dos demais casos inteiramente análoga.

- $n = 2$: É imediato, visto que

$$\det(\mathbf{V}_{2 \times 2}) = \begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = x_2 - x_1$$

- $n = 3$: Como já sabemos, o determinante não se altera por transposição ou combinações lineares. Logo, temos que

$$\det(\mathbf{V}_{3 \times 3}) = \begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{vmatrix}$$

Daí, multiplicando a segunda linha por x_1 e subtraindo da terceira linha, segue que

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ 0 & x_2^2 - x_1x_2 & x_3^2 - x_1x_3 \end{vmatrix}$$

Agora, multiplicando a primeira linha por x_1 e subtraindo da segunda linha, vem que

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ 0 & x_2^2 - x_1x_2 & x_3^2 - x_1x_3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 \\ 0 & x_2^2 - x_1x_2 & x_3^2 - x_1x_3 \end{vmatrix}$$

que por sua vez fornece

$$\begin{vmatrix} 1 & 1 & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 \\ 0 & x_2^2 - x_1x_2 & x_3^2 - x_1x_3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_1x_2 \\ 1 & x_3 - x_1 & x_3^2 - x_1x_3 \end{vmatrix}.$$

Note agora que

$$\begin{vmatrix} 1 & 0 & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_1x_2 \\ 1 & x_3 - x_1 & x_3^2 - x_1x_3 \end{vmatrix} = 1 \cdot \begin{vmatrix} x_2 - x_1 & x_2^2 - x_1x_2 \\ x_3 - x_1 & x_3^2 - x_1x_3 \end{vmatrix},$$

onde a matriz do lado direito da igualdade acima possui a segunda linha como sendo múltipla de $x_2 - x_1$ e a terceira linha múltipla de $x_3 - x_1$, o que implica, de acordo com o Teorema 2.2, que

$$\begin{aligned} \begin{vmatrix} x_2 - x_1 & x_2^2 - x_1x_2 \\ x_3 - x_1 & x_3^2 - x_1x_3 \end{vmatrix} &= (x_3 - x_1)(x_2 - x_1) \begin{vmatrix} 1 & x_2 \\ 1 & x_3 \end{vmatrix} \\ &= (x_3 - x_1)(x_3 - x_2)(x_2 - x_1). \end{aligned}$$

Observe que os passos acima podem ser resumidos como

$$\begin{aligned} \det(\mathbf{V}_{3 \times 3}) &= \det(\mathbf{V}_{3 \times 3}^T) = \det(\mathbf{E}_{L_3 \leftrightarrow L_3 - x_1 L_2} \mathbf{V}_{3 \times 3}^T) \\ &= \det(\mathbf{E}_{L_2 \leftrightarrow L_2 - x_1 L_1} \mathbf{E}_{L_3 \leftrightarrow L_3 - x_1 L_2} \mathbf{V}_{3 \times 3}^T) \\ &= \det(\mathbf{E}_{L_2 \leftrightarrow L_2 - x_1 L_1} \mathbf{E}_{L_3 \leftrightarrow L_3 - x_1 L_2} \mathbf{V}_{3 \times 3}) \\ &= k_1 k_2 \det(\widehat{\mathbf{V}}_{2 \times 2}), \end{aligned}$$

onde $k_1 = x_2 - x_1$, $k_2 = x_3 - x_1$ e $\widehat{\mathbf{V}}_{2 \times 2} = \begin{bmatrix} 1 & x_2 \\ 1 & x_3 \end{bmatrix}$.

- $n = 4$: Seguindo a ideia do caso anterior, temos

$$\det(\mathbf{V}_{4 \times 4}) = k_1 k_2 k_3 \det(\widehat{\mathbf{V}}_{3 \times 3}),$$

com $k_1 = x_2 - x_1$, $k_2 = x_3 - x_1$, $k_3 = x_4 - x_1$ e

$$\widehat{\mathbf{V}}_{3 \times 3} = \begin{bmatrix} 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \end{bmatrix}$$

Logo, aplicando a ideia descrita acima sucessivas vezes, obtém-se o resultado procurado. \square

2.1.2 Matriz Adjunta e a Regra de Cramer

Seja $\mathring{\mathbf{A}}$ a matriz formada por todos os cofatores de cada uma das entradas de uma matriz $\mathbf{A}_{n \times n}$. Diz-se que $\mathring{\mathbf{A}}^T$ é a *matriz adjunta* de \mathbf{A} . Denotaremos $\mathring{\mathbf{A}}^T$ apenas por $\overline{\mathbf{A}}$. Por exemplo, os cofatores da matriz

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 5 \\ 1 & 7 & -2 \\ 6 & 3 & 1 \end{bmatrix}$$

são

$$\begin{aligned} A_{11} &= + \begin{vmatrix} 7 & -2 \\ 3 & 1 \end{vmatrix} = 13, & A_{12} &= - \begin{vmatrix} 1 & -2 \\ 6 & 1 \end{vmatrix} = -13, & A_{13} &= + \begin{vmatrix} 1 & 7 \\ 6 & 3 \end{vmatrix} = -39, \\ A_{21} &= + \begin{vmatrix} 0 & 5 \\ 3 & 1 \end{vmatrix} = -15, & A_{22} &= - \begin{vmatrix} 3 & 5 \\ 6 & 1 \end{vmatrix} = -27, & A_{23} &= + \begin{vmatrix} 3 & 0 \\ 6 & 3 \end{vmatrix} = 9 \\ A_{31} &= + \begin{vmatrix} 0 & 5 \\ 7 & -2 \end{vmatrix} = -35, & A_{32} &= - \begin{vmatrix} 3 & 5 \\ 1 & -2 \end{vmatrix} = 11, & A_{33} &= + \begin{vmatrix} 3 & 0 \\ 1 & 7 \end{vmatrix} = 21. \end{aligned}$$

Portanto

$$\overline{\mathbf{A}} = \begin{bmatrix} 13 & -15 & -35 \\ -13 & -27 & 11 \\ -39 & 11 & 21 \end{bmatrix}.$$

Podemos assim definir, de modo geral, que a matriz adjunta de uma matriz quadrada $n \times n$ é dada por

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

Definido o conceito de matriz adjunta, devemos entender como substituir uma determinada coluna de uma matriz. Para entender o processo, vejamos o exemplo a seguir.

Exemplo 2.4. Consideremos as matrizes

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad \text{e} \quad \mathbf{K} = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}.$$

Vamos substituir a segunda coluna de \mathbf{A} por \mathbf{K} . Seja então \mathbf{A}_{12} a segunda coluna de \mathbf{A} e \mathbf{I}_{21} a segunda linha da matriz $\mathbf{I}_{3 \times 3}$. Daí,

$$\begin{aligned} \mathbf{A} + (\mathbf{K} - \mathbf{A}_{12})\mathbf{I}_{21} &= \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} + \left(\begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} - \begin{bmatrix} b \\ e \\ h \end{bmatrix} \right) \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} + \begin{bmatrix} 0 & k_1 - b & 0 \\ 0 & k_2 - e & 0 \\ 0 & k_3 - h & 0 \end{bmatrix} \\ &= \begin{bmatrix} a & k_1 & c \\ d & k_2 & f \\ g & k_3 & i \end{bmatrix} \end{aligned}$$

Como a coluna substituída foi a segunda, podemos representar a nova matriz da seguinte forma

$$\mathbf{A}_2(\mathbf{K}) = \begin{bmatrix} a & k_1 & c \\ d & k_2 & f \\ g & k_3 & i \end{bmatrix}.$$

Generalizando a ideia do exemplo acima, se quisermos substituir a coluna j de uma matriz $\mathbf{A} = [a_{ij}]_{m \times n}$ por uma matriz coluna $\mathbf{K} = [k_i]_{m \times 1}$, devemos proceder da seguinte forma. Primeiro, escrevemos \mathbf{A} sob a forma

$$\mathbf{A} = [\mathbf{A}_{11} \quad \cdots \quad \mathbf{A}_{1n}]$$

e em seguida calculamos

$$\begin{aligned} &\mathbf{A} + (\mathbf{K} - \mathbf{A}_{1j})\mathbf{I}_{j1} \\ &= [\mathbf{A}_{11} \quad \cdots \quad \mathbf{A}_{1j} \quad \cdots \quad \mathbf{A}_{1n}] - [\mathbf{O}_{11} \quad \cdots \quad \mathbf{O}_{1(j-1)} \quad \mathbf{K} - \mathbf{A}_{1j} \quad \mathbf{O}_{1(j+1)} \quad \cdots \quad \mathbf{O}_{1n}], \\ &= [\mathbf{A}_{11} \quad \cdots \quad \mathbf{A}_{1(j-1)} \quad \mathbf{K} \quad \mathbf{A}_{1(j+1)} \quad \cdots \quad \mathbf{A}_{1n}] \end{aligned}$$

onde \mathbf{A}_{1j} é a coluna j de \mathbf{A} e \mathbf{I}_{j1} é a linha j da matriz identidade $\mathbf{I}_{n \times n}$. Escrevemos então

$$\mathbf{A}_j(\mathbf{K}) = [\mathbf{A}_{11} \quad \cdots \quad \mathbf{A}_{1(j-1)} \quad \mathbf{K} \quad \mathbf{A}_{1(j+1)} \quad \cdots \quad \mathbf{A}_{1n}].$$

Partiremos agora para a demonstração da *regra de Cramer*¹⁴, a qual estabelece uma ligação entre os determinantes à solução de sistemas lineares, bem como a matriz inversa de tais sistemas.

Teorema 2.3. (Regra de Cramer) *Sejam \mathbf{A} uma matriz invertível $n \times n$, $\mathbf{B} = [b_{i1}]_{n \times 1}$ e $\mathbf{X} = [x_{i1}]_{m \times 1}$ tais que $\mathbf{AX} = \mathbf{B}$. Então*

$$x_{i1} = \frac{\det(\mathbf{A}_i(\mathbf{B}))}{\det(\mathbf{A})}$$

¹⁴Gabriel Cramer nasceu em Geneva (1704) e morreu em Bagnols (1752). O trabalho pelo qual ele é mais conhecido é o seu tratado sobre curvas algébricas (*Introduction à l'Analyse des Lignes Courbes Algebriques*), publicado em 1750 [6]. Embora alguns autores atribuam a descoberta da regra de Cramer a Colin Maclaurin (1698-1754), onde este, segundo TWEEDIE [61], foi um dos mais distintos do brilhante círculo de matemáticos escoceses que marcaram o curso do século XVIII, é assegurado por KOSINSKI [34] que tal regra foi obtida devido a Cramer.

Demonstração. Inicialmente, iremos escrever \mathbf{A} como uma matriz de blocos 1×1 , ou seja, $[\mathbf{A}]$. Em seguida, escrevendo a matriz identidade $\mathbf{I}_{n \times n}$ da forma

$$\mathbf{I} = [\mathbf{I}_{11} \quad \cdots \quad \mathbf{I}_{1i} \quad \cdots \quad \mathbf{I}_{n1}],$$

segue que

$$\begin{aligned} \mathbf{A}\mathbf{I}_i(\mathbf{X}) &= [\mathbf{A}] [\mathbf{I}_{11} \quad \cdots \quad \mathbf{X} \quad \cdots \quad \mathbf{I}_{n1}] \\ &= [\mathbf{A}\mathbf{I}_{11} \quad \cdots \quad \mathbf{A}\mathbf{X} \quad \cdots \quad \mathbf{A}\mathbf{I}_{n1}] \\ &= [\mathbf{A}_{11} \quad \cdots \quad \mathbf{B} \quad \cdots \quad \mathbf{A}_{1n}] \\ &= \mathbf{A}_i(\mathbf{B}). \end{aligned}$$

Daí, temos que

$$\det(\mathbf{A}) \det(\mathbf{I}_i(\mathbf{X})) = \det(\mathbf{A}_i(\mathbf{B})),$$

porém,

$$\det(\mathbf{I}_i(\mathbf{X})) = \begin{vmatrix} 1 & 0 & \cdots & x_{11} & \cdots & 0 \\ 0 & 1 & \cdots & x_{21} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & x_{i1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & x_{n1} & \cdots & 1 \end{vmatrix} = x_{i1}$$

de acordo com a propriedade 2 da subseção anterior. Assim, como \mathbf{A} é invertível, temos que $\det(\mathbf{A}) \neq 0$ e portanto

$$\det(\mathbf{A})x_{i1} = \det(\mathbf{A}_i(\mathbf{B})) \Rightarrow x_{i1} = \frac{\det(\mathbf{A}_i(\mathbf{B}))}{\det(\mathbf{A})},$$

finalizando a demonstração. □

Corolário 2.2. *Seja \mathbf{A} uma matriz $n \times n$ invertível. Então*

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \overline{\mathbf{A}}.$$

Demonstração. Seja \mathbf{X}_{1j} a j -ésima coluna de \mathbf{A}^{-1} . Segue então que

$$\mathbf{A}\mathbf{X}_{1j} = \mathbf{I}_{1j}.$$

Pela regra de Cramer, temos que

$$x_{ij} = \frac{\det(\mathbf{A}_i(\mathbf{I}_{1j}))}{\det(\mathbf{A})}.$$

Resta agora observar que $\mathbf{A}_i(\mathbf{I}_{1j})$ é a matriz obtida de \mathbf{A} , substituindo sua coluna i pela coluna j da matriz identidade, de modo que

$$\det(\mathbf{A}_i(\mathbf{I}_{1j})) = 0 \cdot A_{1i} + 0 \cdot A_{2i} + \cdots + 1 \cdot A_{ji} + \cdots + 0 \cdot A_{(n-1)i} + 0 \cdot A_{ni} = A_{ji},$$

de modo que

$$x_{ij} = \frac{1}{\det(\mathbf{A})} A_{ji},$$

ou seja,

$$\mathbf{A}^{-1} = \mathbf{X} = \frac{1}{\det(\mathbf{A})} \overline{\mathbf{A}}.$$

□

Embora, de acordo com POOLE [53], a utilização da regra de Cramer para fornecer a inversa de uma matriz de ordem grande (e por sua vez as soluções de um sistema linear) é computacionalmente ineficiente, sua aplicação na obtenção da inversa de matrizes 2×2 se mostra eficaz, pois, dada uma matriz invertível

$$\mathbf{Y} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

temos que

$$\mathbf{Y}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

2.2 Sistemas Lineares

Definição 2.4. *Uma equação sob a forma*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

é chamada de equação linear, onde x_1, \dots, x_n são as variáveis, a_1, \dots, a_n são os coeficientes e b é o termo constante.

Dessa forma, temos que, equações como

- $-2x_1 + 3x_2 = 6$
- $10p - 31q + 9r = 0$
- $-\frac{2}{3}n = 5$

são exemplos de equações lineares, enquanto que

- $\sqrt{5p} - 2q = 7$
- $\sin\left(\frac{x}{2}\right) - 1 = 4x^2$
- $\frac{2}{\pi} \sqrt[3]{x} - 2y + \cos(z) = 1$

são exemplos de equações não-lineares.

Uma **solução** de uma equação linear $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ é uma lista de números (s_1, \dots, s_n) , tal que $x_1 = s_1, \dots, x_n = s_n$, ou seja,

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

Por exemplo, uma solução para a equação $-2x_1 + 3x_2 = 6$ é $(0, 2)$, uma vez que $-2 \cdot 0 + 3 \cdot 2 = 0 + 6 = 6$. Já, quando temos duas ou mais equações relacionadas pelas mesmas variáveis, então temos o que se chama de **sistema de equações lineares**, onde, neste

caso, a solução deve satisfazer cada uma das equações do sistema. Vejamos abaixo um sistema formado por duas equações lineares

$$\begin{aligned} -2x_1 + x_2 &= 1 \\ x_1 + 3x_2 &= 10 \end{aligned}$$

Notemos que $(1, 3)$ satisfaz ambas as equações, pois

$$\begin{aligned} -2 \cdot 1 + 1 \cdot 3 &= -2 + 3 = 1 \\ 1 \cdot 1 + 3 \cdot 3 &= 1 + 9 = 10 \end{aligned}$$

Agora observemos que tal sistema pode ser escrito em termos matriciais, obtendo o seguinte aspecto

$$\begin{bmatrix} -2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

Uma vez que $\begin{vmatrix} -2 & 1 \\ 1 & 3 \end{vmatrix} = -7 \neq 0$, decorre pela regra de Cramer que

$$\begin{aligned} x_1 &= \frac{\begin{vmatrix} 1 & 1 \\ 10 & 3 \end{vmatrix}}{\begin{vmatrix} -2 & 1 \\ 1 & 3 \end{vmatrix}} = \frac{-7}{-7} = 1 \\ x_2 &= \frac{\begin{vmatrix} -2 & 1 \\ 1 & 10 \end{vmatrix}}{\begin{vmatrix} -2 & 1 \\ 1 & 3 \end{vmatrix}} = \frac{-21}{-7} = 3 \end{aligned}$$

Dessa forma, de modo geral, dado um sistema linear

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned},$$

podemos escrevê-lo sob a forma matricial a seguir

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

ou simplesmente

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

onde $\mathbf{A} = [a_{ij}]_{m \times n}$ (chamada *matriz do sistema*), $\mathbf{X} = [x_{i1}]_{i \times 1}$ e $\mathbf{B} = [b_{i1}]_{i \times 1}$. Dessa forma, desde que a matriz \mathbf{A} seja invertível, temos, de acordo com o TFME, que a solução do sistema acima é única e é dada por $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$, podendo assim ser solucionado utilizando a regra de Cramer. Porém, como adverte LIMA [43], não é necessário determinar a inversa de \mathbf{A} para resolver um sistema linear, uma vez que para isso é necessário resolver

n sistemas lineares. Nesta mesma linha de pensamento temos POOLE [53], o qual afirma que a eliminação gaussiana se sobrepõe até à forma mais eficaz de computar soluções de sistemas lineares pela regra de Cramer, uma vez que para tal, é necessário calcular uma grande quantidade de determinantes.

Para resolver sistemas lineares por meio da eliminação gaussiana, basta tomarmos a *matriz aumentada* do sistema, ou seja, uma matriz de blocos da forma

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array} \right]$$

e em seguida aplicar operações elementares às linhas da matriz acima, de modo que a matriz do sistema, isto é, o bloco esquerdo, fique sob a forma escalonada. Isso nos permite determinar a solução ou as soluções do sistema, quando esta(s) existir(em).

Exemplo 2.5. (Ajuste de Curvas) Os sistemas lineares podem ser utilizados para determinar um polinômio conhecendo alguns pontos pelos quais o seu gráfico passa, ou determinar um polinômio que mais se aproxima de um gráfico de dados. Façamos isso para um polinômio de 2º grau $p(x) = ax^2 + bx + c$, onde $a, b, c \in \mathbb{R}$ e $a \neq 0$, cujo gráfico passa pelos pontos $(1, 15)$, $(2, 19)$ e $(3, 33)$. Notemos que

$$\begin{aligned} 15 &= p(1) = a \cdot 1^2 + b \cdot 1 + c = a + b + c \\ 19 &= p(2) = a \cdot 2^2 + b \cdot 2 + c = 4a + 2b + c \\ 33 &= p(3) = a \cdot 3^2 + b \cdot 3 + c = 9a + 3b + c \end{aligned}$$

Obtemos assim o seguinte sistema linear

$$\begin{aligned} a + b + c &= 15 \\ 4a + 2b + c &= 19 \\ 9a + 3b + c &= 33 \end{aligned}$$

cuja matriz aumentada é

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 4 & 2 & 1 & 19 \\ 9 & 3 & 1 & 33 \end{array} \right]$$

Daí, aplicando a eliminação gaussiana, temos

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 4 & 2 & 1 & 19 \\ 9 & 3 & 1 & 33 \end{array} \right] \xrightarrow[\underline{L_3-9L_1}]{\underline{L_2-4L_1}} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & -2 & -3 & -41 \\ 0 & -6 & -8 & -102 \end{array} \right] \xrightarrow{L_3-3L_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & -2 & -3 & -41 \\ 0 & 0 & 1 & 21 \end{array} \right]$$

Como a matriz do sistema já está escalonada, ou seja, sob a forma

$$\left[\begin{array}{ccc} 1 & 1 & 1 \\ 0 & -2 & -3 \\ 0 & 0 & 1 \end{array} \right]$$

podemos reescrever o sistema inicial sob a forma

$$\begin{aligned} a + b + c &= 15 \\ -2b - 3c &= -41 \\ c &= 21 \end{aligned}$$

Agora basta fazer a **substituição regressiva**, ou seja, substituir as variáveis determinadas nas equações que ainda possuem variáveis indeterminadas, de baixo para cima. Pois, substituindo $c = 21$ em $-2b - 3c = -41$, obtemos

$$-2b - 63 = -41 \Rightarrow b = -11$$

e substituindo $c = 21$, $b = 11$ em $a + b + c = 15$, decorre que

$$a - 11 + 21 = 15 \Rightarrow a = 5.$$

Logo, temos que o polinômio $p(x)$ é dado por

$$p(x) = 5x^2 - 11x + 21.$$

Exemplo 2.6 Em relação ao exemplo anterior, podemos observar que a situação nele descrita poderia ser representada como produto de matrizes, isto é,

$$\begin{bmatrix} 1 & 1 & 1^2 \\ 1 & 2 & 2^2 \\ 1 & 3 & 3^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 15 \\ 19 \\ 33 \end{bmatrix},$$

onde a matriz

$$\begin{bmatrix} 1 & 1 & 1^2 \\ 1 & 2 & 2^2 \\ 1 & 3 & 3^2 \end{bmatrix}$$

é claramente de Vandermonde. Tal fato não é mera coincidência. A aproximação de uma curva por meio de polinômios é realizada utilizando matrizes de vandermonde, cujas linhas são progressões geométricas de bases x_0, x_1, \dots, x_n , com $x_i \neq x_j$ para todo $i \neq j$. Tal matriz multiplica uma matriz coluna de coeficientes a_0, \dots, a_n , o que acaba por resultar em uma matriz coluna com entradas y_0, y_1, \dots, y_n . Em termos

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Daí, de acordo com a propriedade 7 da subseção 2.1.1, segue que o determinante da matriz de Vandermonde é sempre diferente de zero, desde que $x_i \neq x_j$ para todo $i \neq j$ e, portanto, tal sistema sempre admite soluções via Regra de Cramer.

O exemplo acima nos fornece um resultado útil, pois desde que os elementos da matriz de vandermonde sejam distintos, é possível resolver o sistema. Porém, como já apontado anteriormente, resolver um sistema até mesmo por meio da regra de Cramer pode ser um trabalho laborioso dependendo da ordem da matriz de seu sistema. Para contornar tal problema, apresentamos aqui o processo de **interpolação polinomial de Lagrange**¹⁵. Para compreendê-lo, utilizaremos um caso particular para, em seguida, fornecer sua generalização.

¹⁵Joseph Louis Lagrange (1736-1813), nascido em Turin, Itália, é considerado um dos dois maiores matemáticos do século XVIII - o outro é Euler [46].

Consideremos os pontos $(x_0, y_0), (x_1, y_1), (x_2, y_2)$. Vamos obter um polinômio $P(x)$ de grau 2, cujo gráfico passe por tais pontos. Inicialmente precisaremos construir polinômios $P_0(x), P_1(x)$ e $P_2(x)$ de grau 2, tais que

$$P_i(x_j) = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases}$$

Em seguida, tomamos os produtos da forma $y_i P_i$, com $i = 0, 1, 2$. Daí, o polinômio $P(x)$ buscado será dado por

$$P(x) = y_0 P_0 + y_1 P_1 + y_2 P_2.$$

Essa construção de $P(x)$ permite que a curva descrita por tal polinômio passe pelos pontos fornecidos inicialmente. Por exemplo, $P(x_0) = y_0 P_0(x_0) + y_1 P_1(x_0) + y_2 P_2(x_0) = y_0 \cdot 1 + y_1 \cdot 0 + y_2 \cdot 0 = y_0$. Por construção, temos que cada polinômio $P_i(x)$ possui duas raízes. Podemos então escrever

$$P_0(x) = r_0(x - x_1)(x - x_2)$$

$$P_1(x) = r_1(x - x_0)(x - x_2)$$

$$P_2(x) = r_2(x - x_0)(x - x_1)$$

onde $r_i \in \mathbb{R}$, com $i = 0, 1, 2$. Agora observemos que

$$1 = P_0(x_0) = r_0(x_0 - x_1)(x_0 - x_2) \Rightarrow r_0 = \frac{1}{(x_0 - x_1)(x_0 - x_2)}$$

$$1 = P_1(x_1) = r_1(x_1 - x_0)(x_1 - x_2) \Rightarrow r_1 = \frac{1}{(x_1 - x_0)(x_1 - x_2)}$$

$$1 = P_2(x_2) = r_2(x_2 - x_0)(x_2 - x_1) \Rightarrow r_2 = \frac{1}{(x_2 - x_0)(x_2 - x_1)}$$

Logo, podemos escrever $P(x)$ sob a forma

$$P(x) = y_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + y_2 \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}$$

Note que, aplicando a expressão acima aos pontos do Exemplo 2.7, obtemos

$$15 \frac{(x - 2)(x - 3)}{(1 - 2)(1 - 3)} = \frac{15x^2 - 75x + 19}{2}$$

$$19 \frac{(x - 1)(x - 3)}{(2 - 1)(2 - 3)} = -19x^2 + 76x - 57$$

$$33 \frac{(x - 1)(x - 2)}{(3 - 1)(3 - 2)} = \frac{33x^2 - 99x + 66}{2}$$

e portanto, somando os resultados obtidos, segue que

$$P(x) = 5x^2 - 11x + 21,$$

como esperado.

Assim, generalizando a ideia descrita acima, temos que, dados os pontos $(x_0, y_0), \dots, (x_n, y_n)$, existe um polinômio de grau n

$$P(x) = y_0P_0 + \dots + y_nP_n$$

onde

$$P_i = \frac{(x - x_0)(x - x_1) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n)}{(x_i - x_0)(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)}$$

para todo $i = 0, 1, \dots, n$. Tal polinômio é único. De fato, se não fosse, então existiria um outro polinômio $Q(x)$ que interpola os mesmos pontos. Daí, teríamos que $P(x) - Q(x)$ seria um polinômio de grau n que $P(x_i) = 0$ para $i = 0, 1, \dots, n$, o que é absurdo, uma vez que um polinômio de grau n possui no máximo n raízes reais. Logo, segue que $P(x) = Q(x)$, provando assim a unicidade.

Realizada a discussão acima, vejamos mais um exemplo de situação que pode ser resolvida por meio de sistemas lineares.

Exemplo 2.7. Uma exposição cultural é realizada em um espaço formado por 4 salas conforme a Figura 2.1 abaixo:

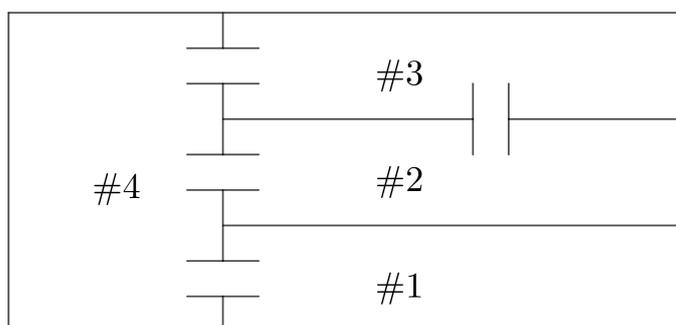


Figura 2.1: Planta do local da exposição cultural.

Fonte: O autor

O número de visitantes em um determinado dia é igual a 100, onde estes estão distribuídos de modo que há x_i visitantes na sala $\#i$, com $i = 1, 2, 3, 4$. A cada dez minutos os participantes se redistribuem e, além disso, nenhum participante visita mais que um quarto nesse intervalo de tempo. Terminados os dez minutos, 40% dos visitantes permanecem na sala em que estavam, enquanto que os demais se distribuem de maneira uniforme para os quartos diretamente acessíveis ao que eles inicialmente ocupava, como por exemplo, de $\#2$, metade vai para $\#3$ e a outra metade para $\#4$. Supondo que ao final de dez minutos permaneceram 12, 25, 26 e 37 visitantes nos quarto $\#1$, $\#2$, $\#3$ e $\#4$, respectivamente, podemos montar o sistema linear que nos fornece quantas pessoas haviam inicialmente em cada sala. Seguindo o diagrama da Figura 2.2 abaixo,

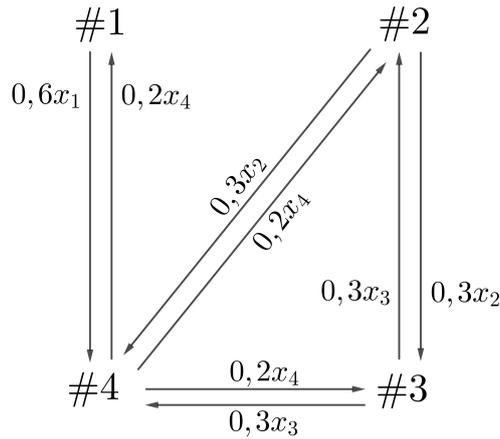


Figura 2.2: Diagrama de troca de salas.

Fonte: O autor

podemos montar o sistema linear que descreve a situação, o qual é dado por

$$\begin{aligned} 0.4x_1 + 0x_2 + 0x_3 + 0.2x_4 &= 12 \\ 0x_1 + 0.4x_2 + 0.3x_3 + 0.2x_4 &= 25 \\ 0x_1 + 0.3x_2 + 0.4x_3 + 0.2x_4 &= 26 \\ 0.6x_1 + 0.3x_2 + 0.3x_3 + 0.4x_4 &= 37 \end{aligned}$$

Escalonando tal sistema, obtemos

$$\begin{aligned} \frac{4}{10}x_1 + 0x_2 + 0x_3 + \frac{2}{10}x_4 &= 120 \\ \frac{4}{10}x_2 + \frac{3}{10}x_3 + \frac{2}{10}x_4 &= 250 \\ \frac{7}{40}x_3 + \frac{1}{20}x_4 &= \frac{29}{4} \\ -\frac{5}{70}x_4 &= -\frac{20}{7} \end{aligned}$$

de modo que, fazendo a substituição regressiva, obtemos como solução

$$x_1 = 10, \quad x_2 = 20, \quad x_3 = 30, \quad x_4 = 40.$$

Quando a matriz aumentada de um sistema é escalonada, os pivôs de cada linha não nula são os coeficientes das *variáveis líderes*, enquanto que os demais são os coeficientes das *variáveis livres*. O número de soluções de um sistema depende de vários fatores. Por exemplo, os *sistemas homogêneos*, isto é, os sistemas da forma $\mathbf{AX} = \mathbf{O}$ possuem apenas a solução *trivial* $\mathbf{X} = \mathbf{O}$ quando \mathbf{A} é quadrada e invertível, fato comprovado pelo TFME.

Note que o número de variáveis livres de um sistema homogêneo está ligado ao número de linhas não nulas de sua forma escalonada. Com efeito, seja \mathbf{A}' a matriz escalonada do sistema $\mathbf{AX} = \mathbf{O}$. Temos que \mathbf{A}' possui um número r de entradas não nulas. Se $r = 0$, então o sistema possui infinitas soluções possíveis, uma vez que infinitas matrizes \mathbf{B} satisfazem a equação $\mathbf{OB} = \mathbf{O}$. Se $r > 0$, então \mathbf{A}' possui r variáveis líderes e $n - r$ variáveis livres. Dessa forma, podemos concluir que, se um sistema homogêneo possui mais variáveis que equações, então o seu número de soluções é infinito. Mais formalmente, se $\mathbf{AX} = \mathbf{O}$ possui m equações e n variáveis, com $m < n$, então o dado sistema possui

infinitas soluções, pois sendo $m \geq r$, decorre que $n > m \geq r$, o que por sua vez implica $n - r > 0$, o que caracteriza a infinitude de soluções do referido sistema. Com isso, podemos enunciar os dois teoremas que seguem.

Teorema 2.4. *Se um sistema linear homogêneo $\mathbf{AX} = \mathbf{O}$ possui n variáveis e sua matriz \mathbf{A}' possui r linhas não nulas, então o seu número de variáveis livres é igual a $n - r$.*

Teorema 2.5. *Se um sistema linear homogêneo possui mais variáveis do que equações, então o mesmo possui infinitas soluções.*

Perceba que, quando $m \geq n$, então o sistema tanto pode apresentar infinitas soluções, como também a única solução. Para analisarmos isso, consideremos o sistema

$$\begin{aligned} ax + by &= 0 \\ x + y &= 0 \end{aligned}$$

com $a, b \neq 0$. Escalonando sua matriz aumentada obtemos

$$\begin{bmatrix} a & b & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

ou seja, $ax + by = 0$, onde y é a única variável livre, o que nos permite escrever

$$x = -\frac{b}{a}y$$

Dessa forma, podemos escrever y como um **parâmetro** t real, o qual pode assumir infinitos valores, e portanto

$$x = -\frac{b}{a}t$$

também assume infinitos valores e conseqüentemente o sistema possui infinitas soluções. Já se fosse $a = 0$ e $b \neq 0$, por exemplo, então teríamos

$$\begin{aligned} by &= 0 \\ x + y &= 0 \end{aligned}$$

ou seja, um sistema homogêneo com duas equações e duas variáveis, porém $by = 0 \Rightarrow y = 0 \Rightarrow x = 0$, de modo que o sistema possui apenas a solução trivial.

Já para o caso $m > n$, consideremos, por exemplo, os sistemas

$$\begin{array}{l} 2x = 0 \\ 3x = 0 \end{array} \quad \text{e} \quad \begin{array}{l} 3x + 3y = 0 \\ 2x + 2y = 0 \\ x + y = 0 \end{array}$$

onde ambos possuem, respectivamente, a solução trivial e infinitas soluções.

A seguir, provaremos que, independente de qual seja o sistema linear, só existem três alternativas para a sua quantidade de soluções: infinitas, uma ou nenhuma.

Teorema 2.6. Dado um sistema linear, temos que só existem três possibilidades no que se refere ao seu número de soluções, as quais são

- O sistema possui a única solução.
- O sistema possui infinitas soluções.

- O sistema não possui soluções.

Demonstração. Consideremos um sistema linear $\mathbf{AX} = \mathbf{B}$. Se tal sistema possui uma única solução ou não possui nenhuma solução, nada há para ser provado. Porém, é necessário provar que, se um sistema admite duas soluções diferentes, então na verdade ele possui infinitas soluções. Com efeito, suponhamos que as matrizes \mathbf{X}_1 e \mathbf{X}_2 são duas soluções do dado sistema, com $\mathbf{X}_1 \neq \mathbf{X}_2$. Temos então que $\mathbf{AX}_1 = \mathbf{B} = \mathbf{AX}_2$ e assim

$$\mathbf{AX}_1 - \mathbf{AX}_2 = \mathbf{O} \Rightarrow \mathbf{A}(\mathbf{X}_1 - \mathbf{X}_2) = \mathbf{O}$$

□

No próximo capítulo vamos obter uma interpretação geométrica para o Teorema 2.6.

3 Transformações Lineares e Matrizes

3.1 O plano \mathbb{R}^2 e o espaço tridimensional \mathbb{R}^3

O plano é formado por duas retas perpendiculares x e y , tais que $x \cap y = \{0\}$, ou seja, intersectam-se em um único ponto, ao qual é atribuído o nome de **origem** do plano. A reta x , também chamada de **eixo das abcissas** fica na posição horizontal, enquanto que a reta y , também chamada de **eixo das ordenadas**, fica esta na posição vertical. Esse sistema de eixos recebe o nome de **plano cartesiano**, também representado por \mathbb{R}^2 (lê-se “r dois”). Cada ponto do plano cartesiano é localizado por duas coordenadas x e y , sendo que x é a **abscissa** e y é a **ordenada**. Dessa forma, um ponto P qualquer do plano pode ser representado por um par (x, y) . A origem é representada pelo ponto $O = (0, 0)$.

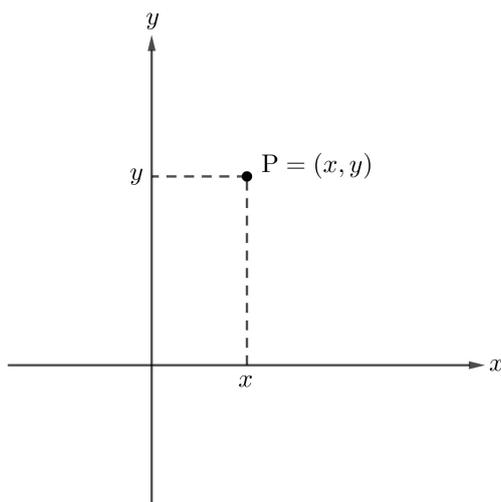


Figura 3.1

Fonte: O autor

O plano é dividido em quatro **quadrantes**, de modo que um ponto $P = (x, y)$ pertence a um desses quadrantes de acordo com as seguintes condições:

- $P = (x, y)$ pertence ao 1º quadrante se $x \geq 0$ e $y \geq 0$.
- $P = (x, y)$ pertence ao 2º quadrante se $x \leq 0$ e $y \geq 0$.
- $P = (x, y)$ pertence ao 3º quadrante se $x \leq 0$ e $y \leq 0$.
- $P = (x, y)$ pertence ao 4º quadrante se $x \geq 0$ e $y \leq 0$.

A imagem abaixo ilustra um caso em que $P_i = (x_i, y_i)$ pertence ao quadrante i , com $i = 1, 2, 3, 4$.

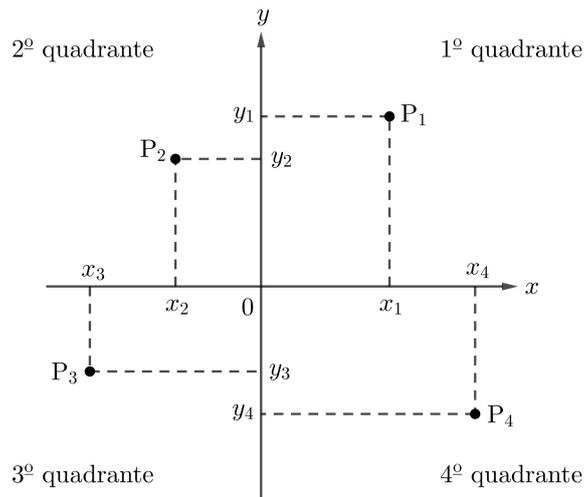


Figura 3.2

Fonte: O autor

Tendo definido o plano cartesiano, podemos agora definir o conceito de **vetor**. Um vetor no plano \mathbb{R}^2 é uma matriz que indica o deslocamento de um ponto do plano para outro. Tendo em vista tal definição, um vetor que, por exemplo, resulta do deslocamento de um ponto A para um ponto B é geralmente representado como \overrightarrow{AB} , conforme podemos observar na Figura 3.3.

A representação de um vetor em termos matemáticos é apresentada de diferentes maneiras em diversos livros, seja como uma matriz linha ou como matriz coluna, indicada por uma letra minúscula e em negrito. Utilizaremos aqui a notação $\mathbf{v} = (x, y)$ para nos referirmos a um vetor de coordenadas x e y . Em alguns momentos, escreveremos \mathbf{v} em sua forma de matriz

$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Ao longo deste capítulo ficará mais claro o porquê de adotar tais notações. O importante é manter a ordem das coordenadas, a fim de evitar ambiguidades, pois $(x, y) \neq (y, x)$, como se pode observar na figura 3.4.

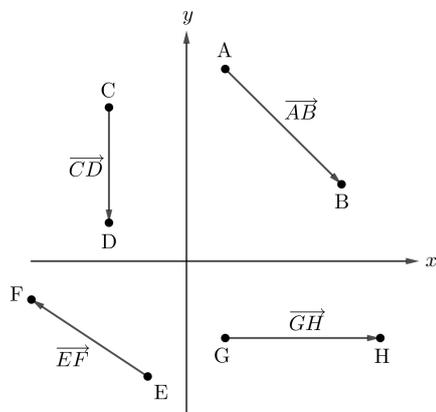


Figura 3.3

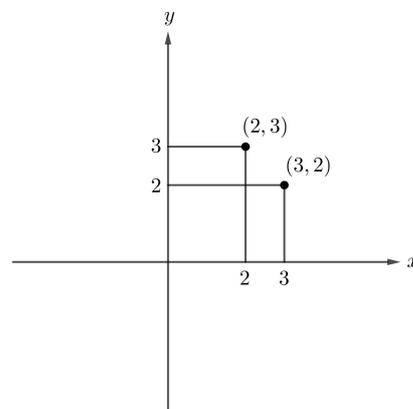


Figura 3.4

Fonte: O autor

As operações de soma, subtração e multiplicação por constante real (também há a multiplicação de vetor por vetor, mas falaremos da mesma mais a diante) são válidas, uma vez que os vetores de \mathbb{R}^2 são matrizes. Assim, para determinar as coordenadas de um dado vetor no plano, devemos subtrair as coordenadas do ponto de partida das coordenadas do ponto de chegada do vetor, de modo que, se um vetor parte da origem, suas coordenadas são exatamente aquelas do ponto de chegada, ou seja, um vetor \overrightarrow{OA} , com $A = (x, y)$ é tal que $\overrightarrow{OA} = (x, y) - (0, 0) = (x - 0, y - 0) = (x, y)$. Já no caso de dois pontos arbitrários $A = (x, y)$ e $B = (z, w)$, temos que

$$\overrightarrow{AB} = (z, w) - (x, y) = (z - x, w - y).$$

Observe que as coordenadas $z - x$ e $w - y$ indicam que o vetor \overrightarrow{AB} é obtido pelo deslocamento horizontal do ponto A por $z - x$ unidades de medida, e em seguida um deslocamento vertical de $w - y$ unidades de medida, tal como apresenta a imagem a seguir:

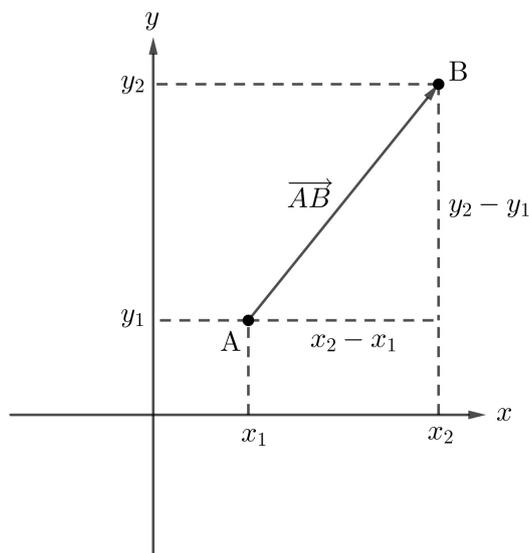


Figura 3.5

Fonte: O autor

A observação acima nos permite, por exemplo, observar graficamente a soma de dois vetores \mathbf{u} e \mathbf{v} . Consideremos então pontos $A = (x_1, y_1)$, $B = (x_2, y_2)$ e $C = (x_3, y_3)$ tais que $\mathbf{u} = \overrightarrow{AB}$, $\mathbf{v} = \overrightarrow{BC}$ e $\mathbf{w} = \overrightarrow{AC}$ conforme a figura abaixo

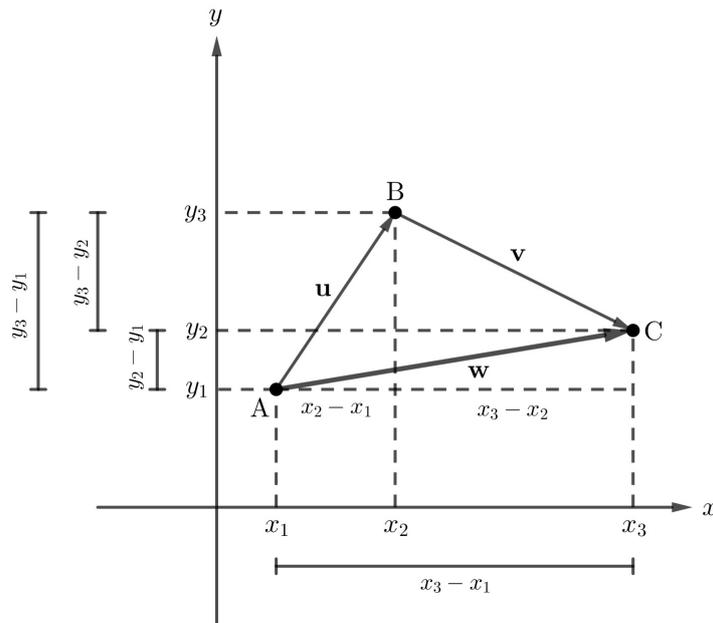


Figura 3.6

Fonte: O autor

daí, é fácil ver que $\mathbf{u} + \mathbf{v} = \mathbf{w}$, pois

$$\mathbf{u} + \mathbf{v} = (x_2 - x_1, y_2 - y_1) + (x_3 - x_2, y_3 - y_2) = (x_3 - x_1, y_3 - y_1) = \mathbf{w}.$$

É claro que, se escritos em forma de matrizes, o resultado não é diferente, como podemos ver a seguir

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix} + \begin{bmatrix} x_3 - x_2 \\ y_3 - y_2 \end{bmatrix} = \begin{bmatrix} x_3 - x_1 \\ y_3 - y_1 \end{bmatrix} = \mathbf{w}.$$

Quanto à multiplicação de um vetor \mathbf{v} por um escalar (número real), temos que esta também é válida, uma vez que matrizes podem ser multiplicadas por números reais. A título de exemplo, vamos considerar um vetor $\mathbf{v} = \overrightarrow{AB}$, onde $A = (x_1, y_1)$ e $B = (x_2, y_2)$. Pelo que já foi visto anteriormente, podemos escrever $\mathbf{v} = (x_2 - x_1, y_2 - y_1)$. Tomando $v_1 = x_2 - x_1$ e $v_2 = y_2 - y_1$, segue que

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

Se quiséssemos multiplicar $\mathbf{v} = \overrightarrow{AB}$ por, digamos, 4, obteríamos um novo vetor $\overrightarrow{AC} = 4\overrightarrow{AB}$, como é possível notar na imagem que segue

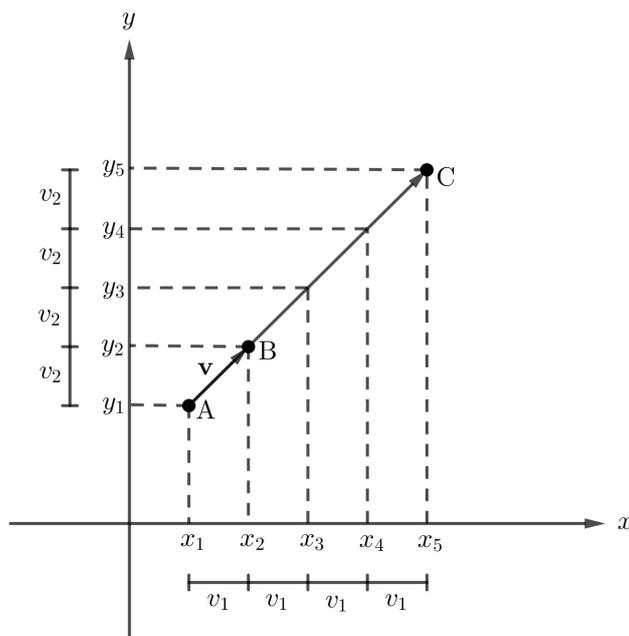


Figura 3.7

Fonte: O autor

Assim, dada uma constante real c arbitrária, temos que

$$c\mathbf{v} = c \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} cv_1 \\ cv_2 \end{bmatrix}$$

Com as ferramentas definidas acima, podemos enunciar um teorema sobre as propriedades dos vetores, onde estas são garantidas pelas propriedades já provadas no Capítulo 1 a respeito das matrizes, uma vez que os vetores no plano \mathbb{R}^2 são também matrizes. Desta forma, podemos enunciar o seguinte teorema

Teorema 3.1. *Sejam \mathbf{u} , \mathbf{v} e \mathbf{w} vetores em \mathbb{R}^2 e $c, d \in \mathbb{R}$. Então*

- ▷ $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- ▷ $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
- ▷ $\mathbf{u} + \mathbf{0} = \mathbf{u}$
- ▷ $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
- ▷ $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
- ▷ $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
- ▷ $c(d\mathbf{u}) = (cd)\mathbf{u}$
- ▷ $1 \cdot \mathbf{u} = \mathbf{u}$

Exemplo 3.1. A figura abaixo mostra um paralelogramo $ABCD$, cujos lados são formados pelos vetores $\mathbf{u} = (u_1, u_2) = (x_2 - x_1, y_3 - y_1)$ e $\mathbf{v} = (v_1, v_2) = (x_3 - x_1, y_2 - y_1)$.

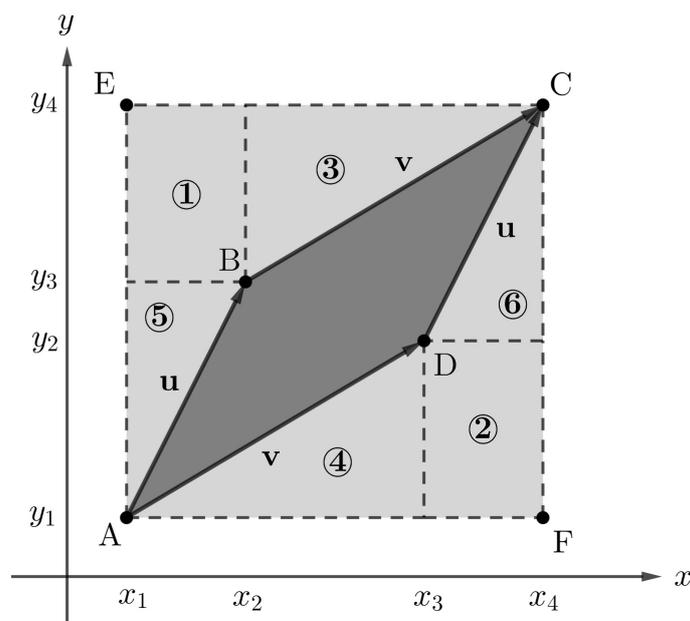


Figura 3.8

Fonte: O autor

Afirmamos que

$$\text{Área}_{ABCD} = \left| \det \left(\begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} \right) \right| = |u_1v_2 - u_2v_1|$$

De fato, para determinarmos a área de $ABCD$ basta calcular a área do retângulo $AECF$ e em seguida subtrair as áreas ①, ②, ..., ⑥. Portanto, notemos que

$$\text{Área}_{AECF} = (u_1 + v_1)(u_2 + v_2) = u_1u_2 + u_1v_2 + u_2v_1 + v_1v_2$$

e que

$$\text{Área}_{\textcircled{1}} = u_1v_2 = \text{Área}_{\textcircled{2}}$$

$$\text{Área}_{\textcircled{3}} = \frac{v_1v_2}{2} = \text{Área}_{\textcircled{4}}$$

$$\text{Área}_{\textcircled{5}} = \frac{u_1u_2}{2} = \text{Área}_{\textcircled{6}}$$

o que nos leva a

$$\begin{aligned} \text{Área}_{ABCD} &= \text{Área}_{AECF} - \text{Área}_{\textcircled{1}} - \text{Área}_{\textcircled{2}} - \dots - \text{Área}_{\textcircled{6}} \\ &= \text{Área}_{AECF} - 2\text{Área}_{\textcircled{1}} - 2\text{Área}_{\textcircled{3}} - \text{Área}_{\textcircled{5}} \\ &= u_1u_2 + u_1v_2 + u_2v_1 + v_1v_2 - 2u_1v_2 - v_1v_2 - u_1u_2 \\ &= u_1v_2 - u_2v_1 \\ &= \det \left(\begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} \right) \end{aligned}$$

Como $AECF$ está no primeiro quadrante, temos que $u_1v_2 - u_2v_1 > 0$, porém em outros quadrantes o resultado poderia ser negativo, mas, uma vez que a área negativa de uma figura plana não está definida, consideramos o valor absoluto de $u_1v_2 - u_2v_1 > 0$ e portanto

$$\text{Área}_{ABCD} = \left| \det \left(\begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} \right) \right| = |u_1v_2 - u_2v_1|$$

Uma consequência disso é que, a área de um triângulo formado por dois vetores é igual à metade do módulo do determinante acima.

O **produto interno** entre dois vetores \mathbf{u} e \mathbf{v} é representado por $\langle \mathbf{u}, \mathbf{v} \rangle$, sendo igual à soma dos produtos entre as entradas correspondentes desses vetores. Ou seja, dados dois vetores

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

temos

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

Note que $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T \mathbf{v} = \mathbf{v}^T \mathbf{u} = \langle \mathbf{v}, \mathbf{u} \rangle$, ou seja, o produto interno é comutativo. A distributividade da multiplicação em relação à soma e a multiplicação por escalar também estão definidos para produto interno, pois tais propriedades são válidas para matrizes. Agora notemos que

$$\langle \mathbf{u}, \mathbf{u} \rangle = u_1^2 + u_2^2 + \cdots + u_n^2,$$

o que nos permite concluir que $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$, onde a igualdade ocorre se, e somente se, $\mathbf{u} = \mathbf{0}$.

O conceito de produto interno nos permite calcular o comprimento de um vetor $\mathbf{u} = (u_1, u_2)$, sendo denotado por $\|\mathbf{u}\|$. Observando a figura abaixo,

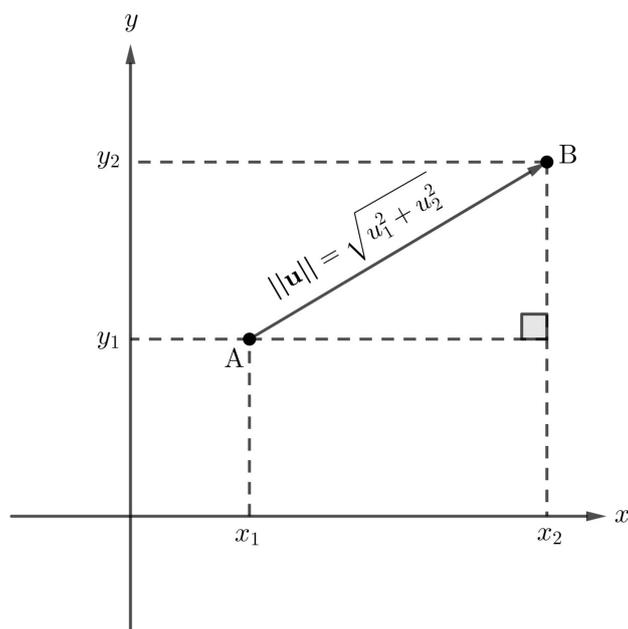


Figura 3.9

Fonte: O autor

podemos observar um triângulo retângulo, o qual, pelo teorema de Pitágoras, nos permite deduzir que

$$\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2} = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$

Exemplo 3.2. Seja h a altura do paralelogramo $ABCD$ do exemplo 3.1 e $\widehat{BAD} = \theta$ na figura abaixo

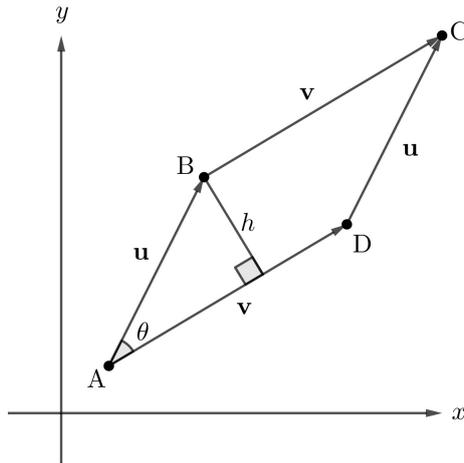


Figura 3.10

Fonte: O autor

Temos que

$$\text{sen}(\theta) = \frac{h}{\|\mathbf{u}\|}$$

e assim

$$\text{Área}_{ABCD} = \|\mathbf{v}\| \cdot h = \|\mathbf{u}\| \|\mathbf{v}\| \text{sen}(\theta)$$

Exemplo 3.3. o produto interno de dois vetores $\mathbf{u} = (u_1, u_2)$ e $\mathbf{v} = (v_1, v_2)$ pode ser calculado multiplicando seus comprimentos pelo cosseno do ângulo θ entre eles. Para tanto, basta considerar o vetor $\mathbf{u} - \mathbf{v}$.

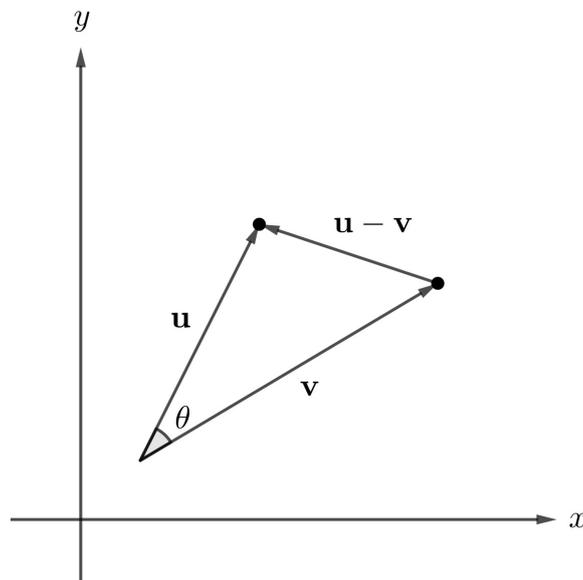


Figura 3.11

Fonte: O autor

Daí, notemos que, pela lei dos cossenos

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

Por outro lado, a definição de produto interno nos fornece

$$\begin{aligned}\|\mathbf{u} - \mathbf{v}\|^2 &= \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle - \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle, \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\langle \mathbf{u}, \mathbf{v} \rangle\end{aligned}$$

o que implica

$$\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos(\theta) = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\langle \mathbf{u}, \mathbf{v} \rangle$$

e portanto

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\|\|\mathbf{v}\|\cos(\theta)$$

Deriva, da expressão acima, um resultado bastante importante, chamado *desigualdade de Cauchy-Schwarz*. Para tanto, consideremos a função cosseno $f: \mathbb{R} \rightarrow \mathbb{R}$, tal que $f(x) = \cos(x)$, para todo $x \in \mathbb{R}$. Como é sabido da trigonometria, a função cosseno é periódica, de modo que oscila de modo uniforme no intervalo $[0, 1]$, isto é,

$$-1 \leq \cos(x) \leq 1.$$

Como $\|\mathbf{u}\|\|\mathbf{v}\| \geq 0$, podemos multiplicar a desigualdade acima por tal valor, sem mudar os sinais, de modo que

$$-\|\mathbf{u}\|\|\mathbf{v}\| \leq \|\mathbf{u}\|\|\mathbf{v}\|\cos(x) \leq \|\mathbf{u}\|\|\mathbf{v}\|,$$

ou seja,

$$|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\|\|\mathbf{v}\|\cos(x) \leq \|\mathbf{u}\|\|\mathbf{v}\|,$$

provando assim a desigualdade de Cauchy-Schwarz

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|.$$

Apresentamos agora o conceito de *ortogonalidade* entre vetores. Consideremos então dois vetores \mathbf{u} e \mathbf{v} . Dizemos que \mathbf{u} é *ortogonal* a \mathbf{v} , o que representaremos por $\mathbf{u} \perp \mathbf{v}$, quando

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0.$$

Dessa forma, dizemos que um conjunto $R = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ é ortogonal se, e somente se,

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0, \quad \forall i \neq j.$$

Conhecendo tal definição, provemos o teorema a seguir.

Teorema 3.2. *Seja $R = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ um conjunto ortogonal de vetores não todos nulos. Então tais vetores são **linearmente independentes**, isto é, nenhum deles pode ser escrito como combinação linear dos demais.*

Demonstração. Suponhamos que um dos vetores $\mathbf{v}_i, i = 1, \dots, n$, de R seja uma combinação linear dos demais. Assim

$$\mathbf{v}_i = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n.$$

Daí, tomando o produto interno da igualdade acima por \mathbf{v}_i , em ambos os membros, obtemos

$$\begin{aligned}\langle \mathbf{v}_i, \mathbf{v}_i \rangle &= \langle \mathbf{v}_i, c_1 \mathbf{v}_1 + \cdots + c_n \mathbf{v}_n \rangle \\ &= \langle \mathbf{v}_i, c_1 \mathbf{v}_1 \rangle + \cdots + \langle \mathbf{v}_i, c_n \mathbf{v}_n \rangle \\ &= c_1 \langle \mathbf{v}_i, \mathbf{v}_1 \rangle + \cdots + c_n \langle \mathbf{v}_i, \mathbf{v}_n \rangle \\ &= 0.\end{aligned}$$

Como $\|\mathbf{v}_i\|^2 = \langle \mathbf{v}_i, \mathbf{v}_i \rangle$, segue que $\|\mathbf{v}_i\|^2 = 0 \iff \mathbf{v}_i = \mathbf{0}$. Porém, \mathbf{v}_i foi tomado de maneira arbitrária, o que nos leva a concluir que $\mathbf{v}_i = \mathbf{0}$, para todo $i = 1, \dots, n$, o que é absurdo, uma vez que nem todos os vetores de R são nulos. \square

Agora, levando a ideia de independência linear para as colunas de uma matriz quadrada $\mathbf{A} = [a_{ij}]_{n \times n}$, podemos dizer que $\mathbf{Ax} = \mathbf{0}$ apenas quando $\mathbf{x} = \mathbf{0}$. Para entendermos um tal fato, consideremos um conjunto $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ de vetores linearmente independentes e uma combinação linear

$$x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n = \mathbf{0},$$

onde nem todos os escalares x_i são todos nulos. Digamos, por simplicidade, que seja $x_1 \neq 0$. Assim, poderíamos escrever

$$\mathbf{a}_1 = - \left(\frac{x_2}{x_1} \mathbf{a}_2 + \cdots + \frac{x_n}{x_1} \mathbf{a}_n \right),$$

isto é, \mathbf{a}_1 seria uma combinação linear dos demais vetores, o que acarretaria em um absurdo, uma vez que tais vetores são linearmente independentes. Bilateralmente, se tais vetores não fossem linearmente independentes, então pelo menos um deles poderia ser escrito como combinação linear dos outros, digamos

$$\mathbf{a}_1 = k_2 \mathbf{a}_2 + \cdots + k_n \mathbf{a}_n,$$

o que por sua vez nos daria

$$-1 \mathbf{a}_1 + k_2 \mathbf{a}_2 + \cdots + k_n \mathbf{a}_n = \mathbf{0},$$

isto é, uma combinação linear dos vetores $\mathbf{a}_1, \dots, \mathbf{a}_n$, onde pelo menos o primeiro escalar é diferente de 0, neste caso, igual a -1 . Podemos assim afirmar que *uma combinação linear de vetores linearmente independentes é igual ao vetor nulo se, e somente se, todos os escalares são iguais a 0*.

Dessa forma, se $\mathbf{Ax} = \mathbf{0}$, onde as colunas de \mathbf{A} , que representaremos por vetores $\mathbf{a}_1, \dots, \mathbf{a}_n$ são linearmente independentes e $\mathbf{x} = [x_{1j}], j = 1, \dots, n$. Assim, podemos escrever esta última igualdade sob a forma

$$x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n = \mathbf{0}, \tag{3.1}$$

o que, de acordo com o que foi provado acima, equivale a $x_i = 0$, ou seja, $\mathbf{x} = \mathbf{0}$. Sendo assim, a recíproca é imediata, isto é, se $\mathbf{Ax} = \mathbf{0}$ apenas quando $\mathbf{0}$, então a combinação linear (3.1) ocorre apenas se $x_i = 0$ para todo i e, portanto, os vetores $\mathbf{a}_1, \dots, \mathbf{a}_n$ são linearmente independentes.

Logo, pelo que foi provado acima juntamente com a equivalência dos itens a) e c) do TFME, podemos concluir que

Uma matriz quadrada \mathbf{A} é invertível se, e somente se, suas colunas são linearmente independentes.

Vamos agora passar ao estudo do espaço tridimensional, ou mais simplesmente, \mathbb{R}^3 , para o qual todas as definições e propriedades sobre o \mathbb{R}^2 são também válidas. A diferença aqui é que, ao invés de duas coordenadas tanto para pontos, bem como para vetores, estes apresentam três coordenadas, tal como podemos visualizar na imagem a seguir o vetor \vec{OP} , onde $O = (0, 0, 0)$ e $P = (x_1, y_1, z_1)$.

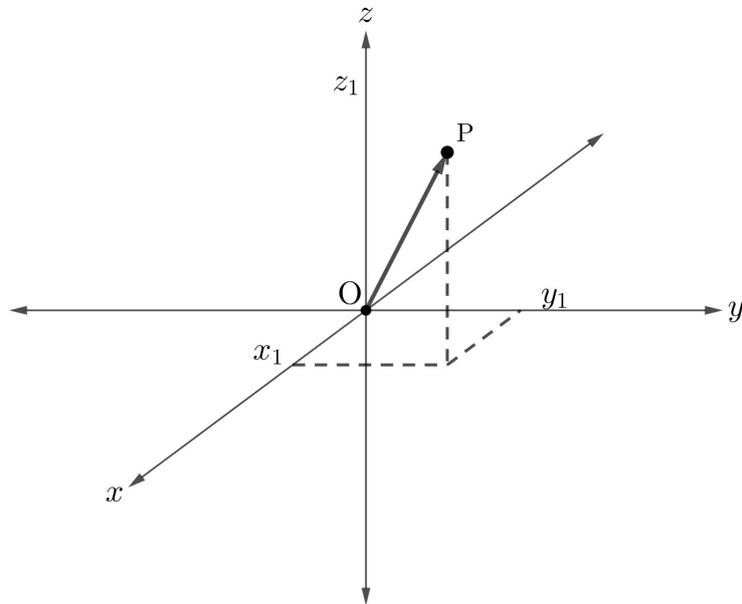


Figura 3.12

Fonte: O autor

A fim de melhor observar o vetor \vec{OP} , podemos imaginá-lo como a diagonal de um paralelepípedo, conforme o esboço abaixo

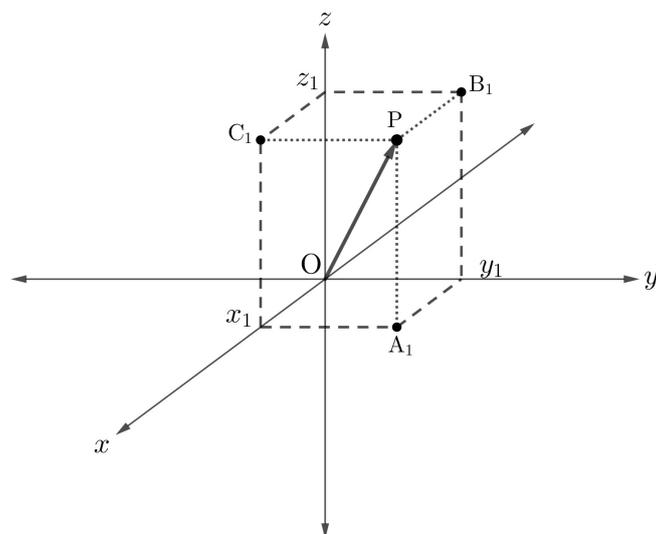


Figura 3.13

Fonte: O autor

As retas perpendiculares aos pontos A_1 , B_1 e C_1 exatamente nos planos xy , yz e xz , respectivamente, intersecta-se no ponto P . Em relação ao espaço tridimensional,

passaremos a nos referir ao mesmo utilizando os termos “espaço xyz ”, “espaço \mathbb{R}^3 ” ou simplesmente \mathbb{R}^3 .

O espaço \mathbb{R}^3 permite definir o conceito de **produto vetorial**, onde este é um vetor perpendicular a outros dois vetores \mathbf{u} e \mathbf{v} , ambos de mesma origem. O produto vetorial neste caso é representado por $\mathbf{u} \times \mathbf{v}$, tal como mostra a figura abaixo

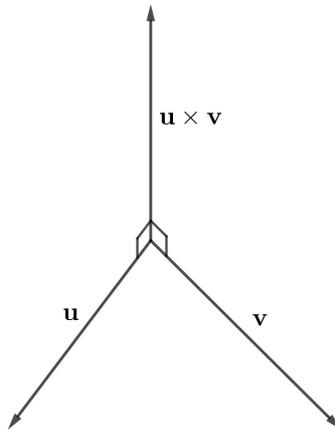


Figura 3.14

Fonte: O autor

Para determinar as coordenadas do produto vetorial entre dois vetores \mathbf{u} e \mathbf{v} , devemos tomar um vetor \mathbf{w} , tal que $\mathbf{w} \perp \mathbf{u}$ (lê-se \mathbf{w} perpendicular a \mathbf{u}) e $\mathbf{w} \perp \mathbf{v}$. Vale lembrar que quando dois vetores são perpendiculares, então seu produto interno é igual a 0, como podemos constatar de acordo com a fórmula obtida no Exemplo 3.3, uma vez que

$$\langle \mathbf{u}, \mathbf{w} \rangle = \|\mathbf{u}\| \|\mathbf{w}\| \cos(90^\circ) = \|\mathbf{u}\| \|\mathbf{w}\| \cdot 0 = 0,$$

sendo análogo o processo para $\langle \mathbf{v}, \mathbf{w} \rangle$. Assim, sejam $\mathbf{u} = (u_x, u_y, u_z)$, $\mathbf{v} = (v_x, v_y, v_z)$ e $\mathbf{w} = (w_x, w_y, w_z)$. Daí, temos que

$$\begin{aligned} 0 &= \langle \mathbf{u}, \mathbf{w} \rangle = u_x w_x + u_y w_y + u_z w_z \\ 0 &= \langle \mathbf{v}, \mathbf{w} \rangle = v_x w_x + v_y w_y + v_z w_z \end{aligned}$$

Como o sistema acima é homogêneo e possui aís variáveis que equações, decorre, pelo Teorema 2.9 que o mesmo possui infinitas soluções. Suponhamos que $u_x v_y - u_y v_x \neq 0$ e vamos então escrevê-lo sob a forma

$$\begin{aligned} u_x w_x + u_y w_y &= -u_z w_z \\ v_x w_x + v_y w_y &= -v_z w_z \end{aligned}$$

Pela regra de Cramer, obtemos

$$\begin{aligned} w_x &= \frac{\begin{vmatrix} -u_z w_z & u_y \\ -v_z w_z & v_y \end{vmatrix}}{\begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix}} = \frac{u_y v_z - u_z v_y}{u_x v_y - u_y v_x} w_z \\ w_y &= \frac{\begin{vmatrix} u_x & -u_z w_z \\ v_x & -v_z w_z \end{vmatrix}}{\begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix}} = \frac{u_z v_x - u_x v_z}{u_x v_y - u_y v_x} w_z \end{aligned}$$

Dáí, como w_z pode assumir qualquer valor real, tomemos $w_z = u_x v_y - u_y v_x$ e assim

$$\mathbf{w} = (w_x, w_y, w_z) = (u_y v_z - u_z v_y, u_z v_x - u_x v_z, u_x v_y - u_y v_x).$$

Se fosse $u_x v_y - u_y v_x = 0$, então passamos a $u_x v_z - u_z v_x$ e, se este for diferente de zero, então basta seguir de modo análogo o processo descrito acima e tomar $w_y = u_x v_z - u_z v_x$. Porém, se for também $u_x v_z - u_z v_x = 0$, então passamos a $u_y v_z - u_z v_y$ e repetimos o processo. Dessa forma, podemos enunciar a seguinte definição

Definição 3.1. O *produto vetorial* $\mathbf{u} \times \mathbf{v}$ de dois vetores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ é tal que

$$\mathbf{u} \times \mathbf{v} = (u_y v_z - u_z v_y, u_z v_x - u_x v_z, u_x v_y - u_y v_x).$$

Uma mnemônica bastante utilizada na literatura utilizando determinantes, permite calcular o produto vetorial de dois vetores de forma mais prática. O processo se baseia em calcular o “determinante” de uma matriz 3×3 formada pelas coordenadas dos vetores $\mathbf{u} = (u_x, u_y, u_z)$ e $\mathbf{v} = (v_x, v_y, v_z)$, juntamente com três vetores unitários $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$, da seguinte forma

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} = \begin{vmatrix} u_y & u_z \\ v_y & v_z \end{vmatrix} \mathbf{i} - \begin{vmatrix} u_x & u_z \\ v_x & v_z \end{vmatrix} \mathbf{j} + \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} \mathbf{k}.$$

O uso das aspas se deve ao fato de que o processo acima é utilizado de modo didático, não sendo um determinante de fato, uma vez que este é definido como um número real, enquanto que o produto vetorial entre dois vetores é também um vetor.

Outra forma de representar o produto vetorial será mostrada a seguir, a qual possui uma interpretação geométrica a respeito dos determinantes, porém, ao invés de área, esta refere-se ao volume. Para tanto, inicialmente devemos calcular

$$\|\mathbf{u} \times \mathbf{v}\|^2 = (u_y v_z - u_z v_y)^2 + (u_z v_x - u_x v_z)^2 + (u_x v_y - u_y v_x)^2.$$

A fim de evitar operações com carga notacional, tomemos

$$\begin{aligned} a &= u_y v_z & b &= u_z v_y & g &= u_x v_x \\ c &= u_z v_x & d &= u_x v_z & h &= u_y v_y, \\ e &= u_x v_y & f &= u_y v_x & i &= u_z v_z \end{aligned}$$

de modo que

$$\begin{aligned} \|\mathbf{u} \times \mathbf{v}\|^2 &= (a - b)^2 + (c - d)^2 + (e - f)^2 \\ &= a^2 + b^2 + c^2 + d^2 + e^2 + f^2 - 2(ab + cd + ef). \end{aligned}$$

Por outro lado, temos que

$$\begin{aligned} \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 &= (u_x^2 + u_y^2 + u_z^2)(v_x^2 + v_y^2 + v_z^2) \\ &= a^2 + b^2 + c^2 + d^2 + e^2 + f^2 + g^2 + h^2 + i^2 \end{aligned}$$

e

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle^2 &= (u_x v_x + u_y v_y + u_z v_z)^2 \\ &= g^2 + h^2 + i^2 + 2(ab + cd + ef), \end{aligned}$$

de modo que

$$\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - \langle \mathbf{u}, \mathbf{v} \rangle^2.$$

Como $\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\|\|\mathbf{v}\|\cos(\theta)$, onde θ é o ângulo formado por \mathbf{u} e \mathbf{v} , decorre da igualdade acima que

$$\begin{aligned} \|\mathbf{u} \times \mathbf{v}\|^2 &= \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - \|\mathbf{u}\|^2\|\mathbf{v}\|^2 \cos^2(\theta) \\ &= \|\mathbf{u}\|^2\|\mathbf{v}\|^2(1 - \cos^2(\theta)) \\ &= \|\mathbf{u}\|^2\|\mathbf{v}\|^2 \text{sen}^2(\theta). \end{aligned}$$

Daí, extraindo a raiz quadrada em ambos os membros da igualdade $\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 \text{sen}^2(\theta)$, obtemos

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\|\text{sen}(\theta).$$

Comparando o resultado acima com aquele obtido no Exemplo 3.2, podemos concluir que o comprimento do produto vetorial de dois vetores \mathbf{u} e \mathbf{v} é igual à área de um paralelogramo, cujos lados opostos são os vetores \mathbf{u} e \mathbf{v} . Além disso, podemos concluir que, a área de um triângulo formado por \mathbf{u} e \mathbf{v} , tal como no Exemplo 3.3, é igual à metade de $\|\mathbf{u} \times \mathbf{v}\|$.

Vamos finalizar nossa discussão a respeito das propriedades do \mathbb{R}^3 com o exemplo a seguir, o qual mostra a interpretação geométrica de determinantes de matrizes 3×3 .

Exemplo 3.4. Sejam

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \quad e \quad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

vetores arbitrários de \mathbb{R}^3 e chamemos de \mathbf{V} a matriz

$$\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}.$$

Daí, como

$$\mathbf{v} \times \mathbf{w} = \begin{bmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{bmatrix},$$

segue que

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \times \mathbf{w} \rangle &= u_1(v_2w_3 - v_3w_2) + u_2(v_3w_1 - v_1w_3) + u_3(v_1w_2 - v_2w_1) \\ &= u_1(v_2w_3 - v_3w_2) - u_2(v_1w_3 - v_3w_1) + u_3(v_1w_2 - v_2w_1) \\ &= u_1 \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} - u_2 \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} + u_3 \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} \\ &= \det(\mathbf{V}^T) \\ &= \det(\mathbf{V}). \end{aligned}$$

Agora consideremos o paralelepípedo formado pelos vetores \mathbf{u} , \mathbf{v} e \mathbf{w} na figura a seguir

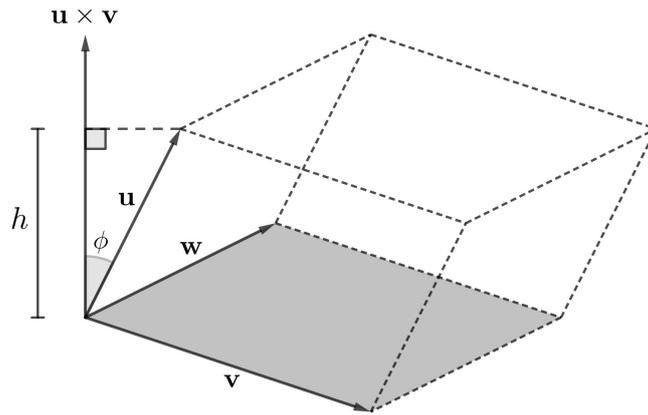


Figura 3.15

Fonte: O autor

Como já é sabido, a área do paralelogramo sombreado (base do paralelepípedo) é igual a $\|\mathbf{u} \times \mathbf{v}\|$, enquanto que $h = \|\mathbf{u}\| \cos(\phi)$, onde ϕ é o ângulo entre $\mathbf{u} \times \mathbf{v}$ e \mathbf{u} . Portanto, temos que o volume Vol_p do paralelepípedo é dado por

$$\text{Vol}_p = \|\mathbf{u} \times \mathbf{v}\| \cdot \|\mathbf{u}\| \cos(\phi) = \langle \mathbf{u}, \mathbf{v} \times \mathbf{w} \rangle$$

Logo, podemos concluir que

$$\text{Vol}_p = \langle \mathbf{u}, \mathbf{v} \times \mathbf{w} \rangle = \det(\mathbf{V}),$$

ou seja, o determinante de uma matriz 3×3 formada por três vetores do \mathbb{R}^3 , todos de mesma origem, é igual ao volume de um paralelepípedo, cujas arestas são formadas por tais vetores.

Como prometido, vamos apresentar aqui uma interpretação geométrica do Teorema 2.10. Consideremos para tanto um sistema linear

$$\begin{cases} ax + by = j \\ cx + dy = k \end{cases}$$

Sabemos que o mesmo só possui solução se suas equações forem linearmente independentes, pois só assim a mesma será invertível. Nesse caso, o sistema possui a única solução. Geometricamente, significa dizer que as retas que representam cada equação não possuem mesma inclinação, portanto são concorrentes em um único ponto (x_0, y_0) , com

$$x_0 = \frac{dj - bk}{ad - bc} \quad y_0 = \frac{ak - cj}{ad - bc}$$

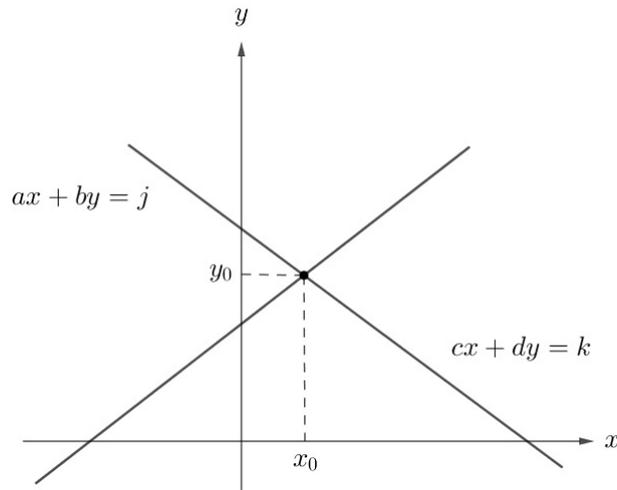


Figura 3.16

Fonte: O autor

No caso em que uma das linhas da matriz do sistema é múltipla da outra, ou seja, são linearmente dependentes, o sistema possui infinitas soluções, pois é possível partir de uma e chegar a outra por meio de combinações elementares. Assim, o sistema reduz-se a uma só equação. Por exemplo, dados $\alpha, \beta \in \mathbb{R}$ não nulos, se tivermos o sistema

$$\begin{cases} \alpha ax + \alpha by = \alpha j \\ \beta ax + \beta by = \beta j \end{cases}$$

temos que a matriz escalonada do sistema é

$$\begin{bmatrix} a & b & j \\ 0 & 0 & 0 \end{bmatrix}$$

ou seja, reduz-se a uma só equação, a qual possui infinitas soluções (uma reta possui infinitos pontos).

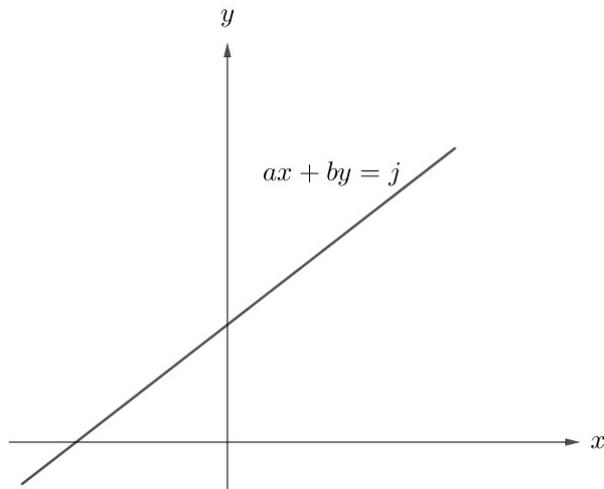


Figura 3.17

Fonte: O autor

Existem ainda os casos em que se tem um sistema impossível, o qual não possui soluções. Sabemos que duas retas só podem se intersectar em único ponto (solução

única) ou em todos os pontos, quando são coincidentes (infinitas soluções). Logo, no caso de sistemas impossíveis, temos que suas equações representam retas paralelas, não possuindo assim pontos em comum e, conseqüentemente, nenhuma solução (x_0, y_0) para o sistema. Tais sistemas possuem a forma

$$\begin{cases} ax + by = j \\ ax + by = k \end{cases}$$

com $j \neq k$.

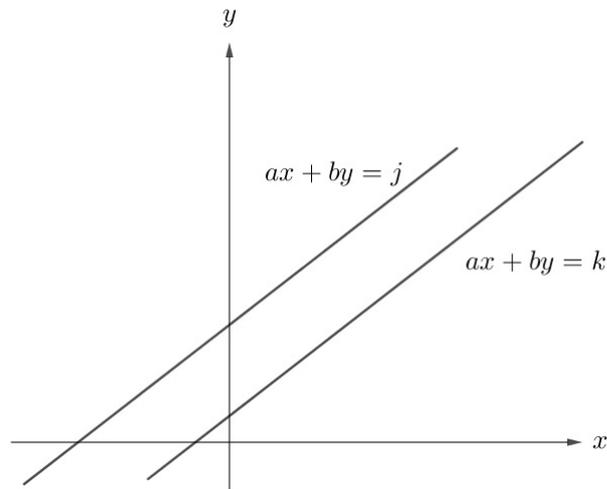


Figura 3.18

Fonte: O autor

Com raciocínio análogo podemos pensar a respeito de sistemas lineares com três incógnitas. Considerando um sistema

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}$$

onde $a_1x + b_1y + c_1z = d_1$ é a equação do plano Γ_1 e $a_2x + b_2y + c_2z = d_2$ é a equação do plano Γ_2 , temos duas possibilidades:

a) o sistema é possível e determinado, tendo infinitas soluções, ou seja, os planos se intersectam em uma reta como mostra a figura a seguir

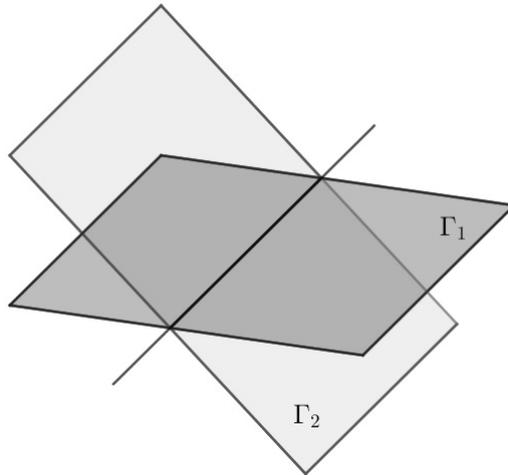


Figura 3.19

Fonte: O autor

b) o sistema é impossível, ou seja, os planos são paralelos, não possuindo pontos em comum.

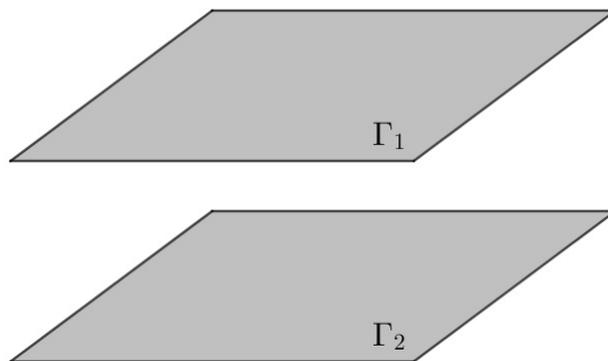


Figura 3.20

Fonte: O autor

3.2 Transformações Lineares

3.2.1 Conceito e exemplos

Definição 3.2. Uma *transformação linear* é uma aplicação do tipo $T : X \rightarrow Y$, tal que para quaisquer vetores $\mathbf{u}, \mathbf{v} \in X$ e qualquer escalar c , são válidas as seguintes propriedades

- ▷ $T(c\mathbf{v}) = cT(\mathbf{v})$
- ▷ $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

A definição acima nos permite concluir um fato importante sobre as transformações lineares, que é a preservação da origem, isto é, $T(\mathbf{0}) = \mathbf{0}$, isto porque

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

A propriedade acima é bastante útil para identificar quando uma aplicação é uma transformação linear. Porém, vale ressaltar que $T_{\mathbf{v}}(\mathbf{0}) = \mathbf{0}$ não implica a linearidade de uma aplicação, pois, tomando, por exemplo, a aplicação $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ com

$$T \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} xy \\ x + y \end{bmatrix},$$

temos que

$$T(\mathbf{0}) = T \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 0 \cdot 0 \\ 0 + 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0},$$

porém

$$\begin{aligned} T \left(\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} w \\ z \end{bmatrix} \right) &= T \left(\begin{bmatrix} x+w \\ y+z \end{bmatrix} \right) = \begin{bmatrix} (x+w)(y+z) \\ (x+z) + (y+w) \end{bmatrix} \\ &= \begin{bmatrix} xy + xz + wy + wz \\ (x+y) + (w+z) \end{bmatrix} \\ &= \begin{bmatrix} xy + xz \\ x + y \end{bmatrix} + \begin{bmatrix} wz + wy \\ w + z \end{bmatrix} \\ &\neq \begin{bmatrix} xy \\ x + y \end{bmatrix} + \begin{bmatrix} wz \\ w + z \end{bmatrix} \\ &= T \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) + T \left(\begin{bmatrix} w \\ z \end{bmatrix} \right), \end{aligned}$$

contrariando assim a propriedade de soma da Definição 3.2, e portanto T não é linear.

Exemplo 3.5. Uma *translação* no plano é uma aplicação $T_{\mathbf{v}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, tal que

$$T_{\mathbf{v}}(P) = P + \mathbf{v},$$

o que significa que $T_{\mathbf{v}}$ desloca qualquer ponto P do plano na direção do vetor \mathbf{v} , como podemos observar na imagem 3.16, onde os vértices (e conseqüentemente todos os outros pontos) do triângulo ABC , com $A = (x_1, y_1)$, $B = (x_2, y_2)$ e $C = (x_3, y_3)$, são transladados na direção do vetor $\mathbf{v} = (\alpha, \beta)$, de modo a obtermos o triângulo $A'B'C'$, com $A' = (x_1 + \alpha, y_1 + \beta)$, $B' = (x_2 + \alpha, y_2 + \beta)$ e $C' = (x_3 + \alpha, y_3 + \beta)$.

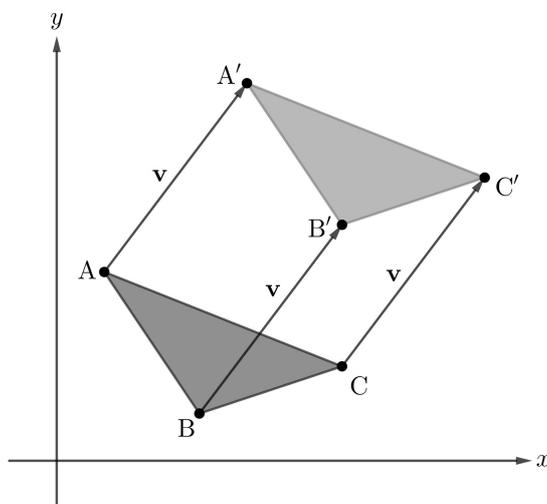


Figura 3.21

Fonte: O autor

Note agora que se \mathbf{v} não for o vetor nulo, então

$$T_{\mathbf{v}}(\mathbf{0}) = T_{\mathbf{v}}(0, 0) = (0, 0) + \mathbf{v} = (0, 0) + (\alpha, \beta) = (0 + \alpha, 0 + \beta) = (\alpha, \beta) \neq \mathbf{0},$$

o que contraria a propriedade de preservação da origem, nos permitindo concluir que $T_{\mathbf{v}}$ é não-linear. Todavia, *as translações possuem a propriedade de preservar distâncias*. Provemos isso utilizando os lados \overline{AB} e $\overline{A'B'}$, sendo análoga a demonstração para os outros lados. Primeiro, observe que

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

ao passo que

$$\begin{aligned} d(A', B') &= \sqrt{(x_2 + \alpha - x_1 - \alpha)^2 + (y_2 + \beta - y_1 - \beta)^2} \\ &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \\ &= d(A, B). \end{aligned}$$

Portanto, $AB = A'B'$, $AC = A'C'$ e $BC = B'C'$, ou seja, os triângulos ABC e $A'B'C'$ são congruentes, pelo caso de congruência LLL. Analogamente, podemos observar que um paralelepípedo Π pode ser transladado no \mathbb{R}^3 por uma translação $T_{\mathbf{v}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, de modo que $T_{\mathbf{v}}(\Pi)$ é um paralelepípedo com as mesmas medidas (e conseqüentemente o mesmo volume) de Π .

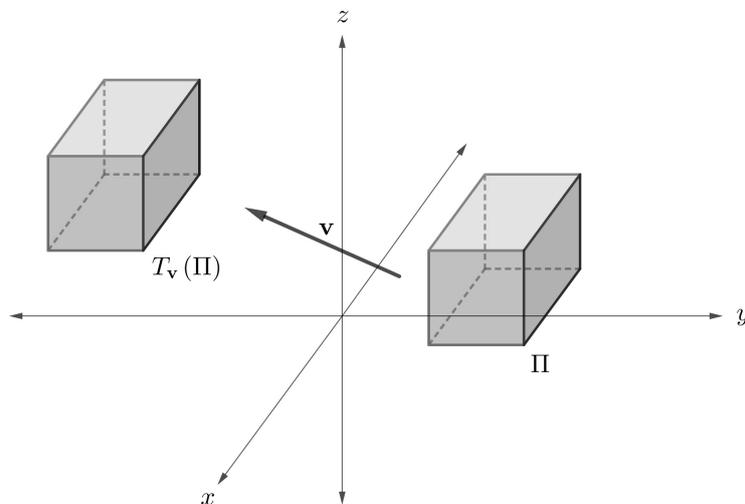


Figura 3.22

Fonte: O autor

Um exemplo importante de transformação linear é a **rotação**. Uma rotação no \mathbb{R}^2 é uma aplicação $R_{\theta} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, a qual possui a propriedade de rotacionar um vetor \mathbf{u} qualquer do plano por um ângulo θ , tal como podemos observar na figura a seguir

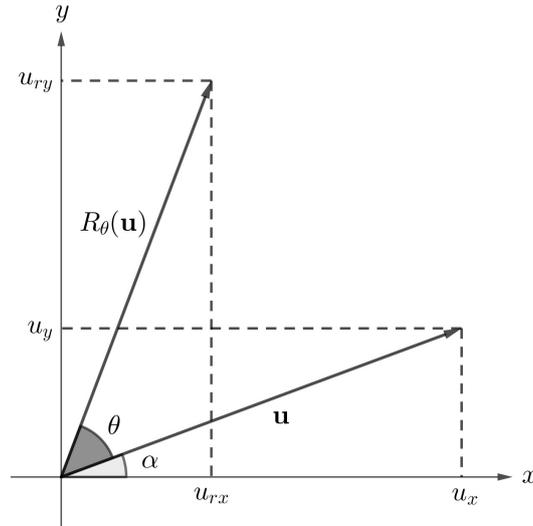


Figura 3.23

Fonte: O autor

Observe que o ângulo α formado por \mathbf{u} e o eixo x nos permite concluir que

$$\cos(\alpha) = \frac{u_x}{\|\mathbf{u}\|} \Rightarrow u_x = \|\mathbf{u}\| \cos(\alpha)$$

$$\text{sen}(\alpha) = \frac{u_y}{\|\mathbf{u}\|} \Rightarrow u_y = \|\mathbf{u}\| \text{sen}(\alpha),$$

ao passo que

$$\cos(\alpha + \theta) = \frac{u_{rx}}{\|R_\theta(\mathbf{u})\|} \Rightarrow u_{rx} = \|R_\theta(\mathbf{u})\| \cos(\alpha + \theta)$$

$$\text{sen}(\alpha + \theta) = \frac{u_{ry}}{\|R_\theta(\mathbf{u})\|} \Rightarrow u_{ry} = \|R_\theta(\mathbf{u})\| \text{sen}(\alpha + \theta).$$

Sendo $\|\mathbf{u}\|$ e $\|R_\theta(\mathbf{u})\|$ raios de um arco de circunferência de centro na origem, segue que $\|R_\theta(\mathbf{u})\| = \|\mathbf{u}\|$. Daí, segue que

$$\begin{aligned} u_{rx} &= \|R_\theta(\mathbf{u})\| \cos(\alpha + \theta) \\ &= \|\mathbf{u}\| \cos(\alpha + \theta) \\ &= \|\mathbf{u}\| \cos(\alpha) \cos(\theta) - \|\mathbf{u}\| \text{sen}(\alpha) \text{sen}(\theta) \\ &= u_x \cos(\theta) - u_y \text{sen}(\theta). \end{aligned}$$

e

$$\begin{aligned} u_{ry} &= \|R_\theta(\mathbf{u})\| \text{sen}(\alpha + \theta) \\ &= \|\mathbf{u}\| \text{sen}(\alpha + \theta) \\ &= \|\mathbf{u}\| \text{sen}(\alpha) \cos(\theta) + \|\mathbf{u}\| \text{sen}(\theta) \cos(\alpha) \\ &= u_x \text{sen}(\theta) + u_y \cos(\theta). \end{aligned}$$

Utilizando a multiplicação de matrizes, podemos escrever os dois resultados obtidos acima da seguinte maneira

$$R_\theta(\mathbf{u}) = \begin{bmatrix} u_{rx} \\ u_{ry} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\text{sen}(\theta) \\ \text{sen}(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix},$$

onde a matriz

$$\begin{bmatrix} \cos(\theta) & -\text{sen}(\theta) \\ \text{sen}(\theta) & \cos(\theta) \end{bmatrix}$$

é denominada *matriz da transformação* R_θ . Tomando-a por \mathbf{R} , podemos simplesmente escrever

$$R_\theta(\mathbf{u}) = \mathbf{R}\mathbf{u}.$$

Vamos agora provar que R_θ é linear. Consideremos para tanto dois vetores $\mathbf{u} = (u_1, u_2)$, $\mathbf{v} = (v_1, v_2)$ de \mathbb{R}^2 e uma constante real arbitrária c . Daí,

$$R_\theta(\mathbf{u} + \mathbf{v}) = \mathbf{R}(\mathbf{u} + \mathbf{v}) = \mathbf{R}\mathbf{u} + \mathbf{R}\mathbf{v} = R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$$

e

$$R_\theta(c\mathbf{v}) = \mathbf{R}(c\mathbf{v}) = c\mathbf{R}\mathbf{v} = cR_\theta(\mathbf{v}),$$

o que prova a linearidade de R_θ .

Por ser menos intuitivo, o fato de que $R_\theta(\mathbf{u} + \mathbf{v}) = R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$ pode ser observado visualmente na figura abaixo

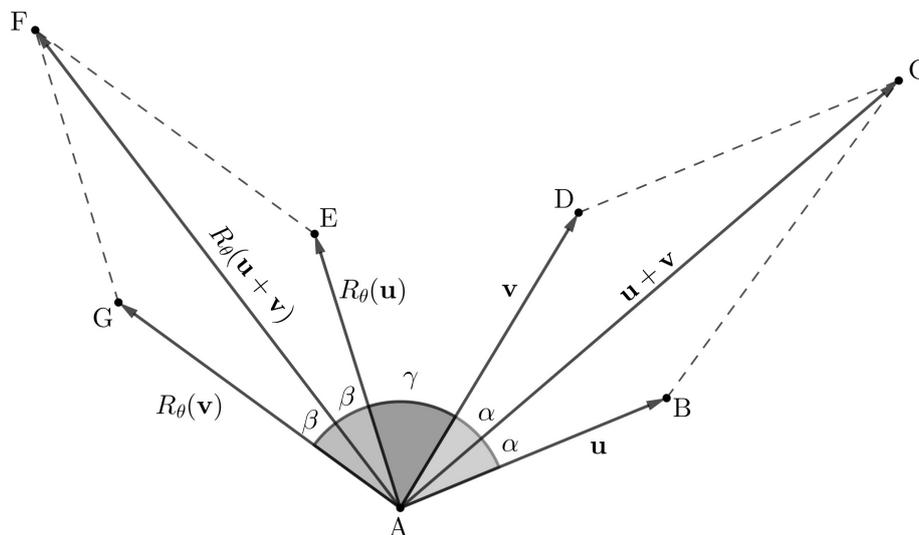


Figura 3.24

Fonte: Adaptado de LIMA [43], p.42

Note que os vetores \overrightarrow{AC} e \overrightarrow{AF} estão sobre as diagonais dos paralelogramos ABCD e AEFG, respectivamente, o que justifica os ângulos \widehat{BAC} e \widehat{EAG} serem bisectados. Assim, temos que

$$2\alpha + \gamma = \theta = 2\beta + \gamma \Rightarrow \alpha = \beta.$$

Decorre daí que $\widehat{CAF} = \alpha + \beta + \gamma = \theta$, ou seja, rotacionar o vetor $\mathbf{u} + \mathbf{v}$ por um ângulo θ é o mesmo que a somar as rotações de \mathbf{u} e \mathbf{v} por este mesmo ângulo.

Vale ressaltar que a rotação é invertível, uma vez que $\det(\mathbf{R}) = \cos^2(\theta) + \text{sen}^2(\theta) = 1 \neq 0$. Portanto, para reverter uma rotação de ângulo θ , basta realizar a rotação em sentido contrário $R_{-\theta}$, sendo sua matriz de transformação

$$\mathbf{R}^{-1} = \begin{bmatrix} \cos(-\theta) & -\text{sen}(-\theta) \\ \text{sen}(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \text{sen}(\theta) \\ -\text{sen}(\theta) & \cos(\theta) \end{bmatrix} = \mathbf{R}^T,$$

o que acaba por mostrar que *a inversa da matriz de uma rotação é igual à transposta desta mesma matriz.*

Assim como podem ser realizadas rotações no plano, também podem ser realizadas no espaço. A diferença neste caso é que, a rotação deve ser em torno de um ou mais eixos. Por exemplo, a matriz de rotação de ângulo ψ (no sentido anti-horário) em torno do eixo z é a seguinte

$$\mathbf{R}_{z,\psi} = \begin{bmatrix} \cos(\psi) & -\text{sen}(\psi) & 0 \\ \text{sen}(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

pois uma rotação de ângulo ψ de um vetor $\mathbf{v} = (v_x, v_y, v_z)$ em torno do eixo z equivale a rotacionar as coordenadas v_x, v_y por um ângulo ψ , deixando a coordenada v_z intacta, o que justifica o fato da terceira linha e terceira coluna de $R_{z,\psi}$ serem, respectivamente, iguais à terceira linha e à terceira coluna da matriz $\mathbf{I}_{3 \times 3}$.

De maneira análoga, as matrizes de rotação (no sentido anti-horário) de ângulo θ e ϕ , em torno dos eixos x e y , respectivamente, são

$$\mathbf{R}_{x,\theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\text{sen}(\theta) \\ 0 & \text{sen}(\theta) & \cos(\theta) \end{bmatrix} \quad \text{e} \quad \mathbf{R}_{y,\phi} = \begin{bmatrix} \cos(\phi) & 0 & -\text{sen}(\phi) \\ 0 & 1 & 0 \\ \text{sen}(\phi) & 0 & \cos(\phi) \end{bmatrix}.$$

Dessa forma, a combinação, por exemplo, $z \rightarrow x \rightarrow y$ das rotações acima aplicadas a um vetor arbitrário do \mathbb{R}^3 é dada por uma matriz \mathbf{R} , tal que

$$\mathbf{R} = \mathbf{R}_{z,\psi} \mathbf{R}_{x,\theta} \mathbf{R}_{y,\phi},$$

ou seja,

$$\mathbf{R} = \begin{bmatrix} c(\phi)c(\psi) + s(\phi)s(\psi)s(\theta) & -c(\theta)s(\psi) & -c(\psi)s(\phi) + c(\phi)s(\psi)s(\theta) \\ c(\phi)s(\psi) - c(\psi)s(\phi)s(\theta) & c(\psi)c(\theta) & -s(\phi)s(\psi) - c(\phi)c(\psi)s(\theta) \\ c(\theta)s(\phi) & s(\theta) & c(\phi)s(\theta) \end{bmatrix},$$

onde $s() = \text{sen}()$ e $c() = \text{cos}()$.

Exemplo 3.6. A *projeção ortogonal* $P : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ de um vetor $v = (v_x, v_y)$ sobre uma reta $y = ax$ (vide Figura 3.20. abaixo) também é uma transformação linear. Mas antes de provarmos sua linearidade, vamos obter sua matriz de transformação.

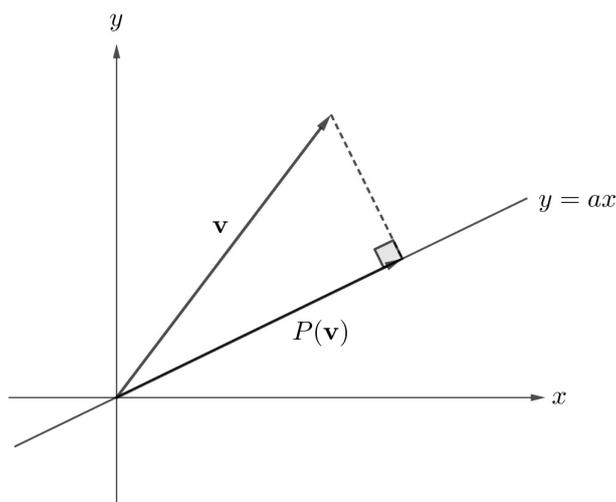


Figura 3.25

Fonte: Adaptado de LIMA [43], p.44

Notemos que $P(\mathbf{v}) = (v'_x, av'_x)$. Daí, pelo teorema de Pitágoras, temos que

$$d(\mathbf{v}, \mathbf{0})^2 = d(\mathbf{v}, P(\mathbf{v}))^2 + d(P(\mathbf{v}), \mathbf{0})^2,$$

o que equivale a

$$\begin{aligned} v_x^2 + v_y^2 &= \{(v'_x - v_x)^2 + (av'_x - v_y)^2\} + \{(v'_x)^2 + a^2(v'_x)^2\} \\ &= 2(v'_x)^2 + v_x^2 + 2a^2(v'_x)^2 + v_y^2 - 2v'_x v_x - 2av'_x v_y. \end{aligned}$$

Decorre, a partir daí, que

$$\begin{aligned} 2v'_x v_x + 2av'_x v_y &= 2(v'_x)^2 + 2a^2(v'_x)^2 \\ \iff v'_x(v_x + av_y) &= (1 + a^2)(v'_x)^2 \\ \iff v'_x &= \frac{v_x + av_y}{1 + a^2} \end{aligned}$$

Portanto

$$v'_x = \left(\frac{1}{1 + a^2} \right) v_x + \left(\frac{a}{1 + a^2} \right) v_y$$

e

$$v'_y = \left(\frac{a}{1 + a^2} \right) v_x + \left(\frac{a^2}{1 + a^2} \right) v_y.$$

Logo, temos que a matriz da projeção P é

$$\mathbf{P} = \begin{bmatrix} \frac{1}{1+a^2} & \frac{a}{1+a^2} \\ \frac{a}{1+a^2} & \frac{a^2}{1+a^2} \end{bmatrix}$$

de modo que, utilizando as propriedades de multiplicação de matrizes, podemos mostrar facilmente a linearidade de P . Vamos então tomar $\alpha = \frac{1}{1+a^2}$, o que nos permite escrever

$$\mathbf{P} = \alpha \begin{bmatrix} 1 & a \\ a & a^2 \end{bmatrix}.$$

Assim, considerando dois vetores $\mathbf{u} = (u_1, u_2), \mathbf{v} = (v_1, v_2)$ arbitrários de \mathbb{R}^2 e uma constante real k , temos que

$$\begin{aligned}
 P(\mathbf{u} + k\mathbf{v}) &= P\left(\begin{bmatrix} u_1 + kv_1 \\ u_2 + kv_2 \end{bmatrix}\right) \\
 &= \alpha \begin{bmatrix} 1 & a \\ a & a^2 \end{bmatrix} \begin{bmatrix} u_1 + kv_1 \\ u_2 + kv_2 \end{bmatrix} \\
 &= \alpha \begin{bmatrix} u_1 + au_2 + kv_1 + akv_2 \\ au_1 + a^2u_2 + kv_1 + a^2kv_2 \end{bmatrix} \\
 &= \alpha \left(\begin{bmatrix} u_1 + au_2 \\ au_1 + a^2u_2 \end{bmatrix} + k \begin{bmatrix} v_1 + av_2 \\ av_1 + a^2v_2 \end{bmatrix} \right), \\
 &= \alpha \begin{bmatrix} 1 & a \\ a & a^2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + k\alpha \begin{bmatrix} 1 & a \\ a & a^2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\
 &= \mathbf{P} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + k\mathbf{P} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\
 &= P(\mathbf{u}) + kP(\mathbf{v})
 \end{aligned}$$

o que acaba por provar a linearidade de P .

Exemplo 3.7. Uma *reflexão* $S : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ em torno de uma reta que passa pela origem é uma transformação linear que leva um vetor $\mathbf{v} = (v_1, v_2)$ em um vetor $S(\mathbf{v})$, de modo que $P(\mathbf{v}) = P(S(\mathbf{v}))$, como podemos observar na Figura 3.21. Dessa forma, temos que o vetor $P(\mathbf{v})$ coincide com a metade da diagonal do paralelogramo formado pela soma dos vetores \mathbf{v} e $S(\mathbf{v})$, o que justifica o fato de $\mathbf{v} + S(\mathbf{v}) = 2P(\mathbf{v})$.

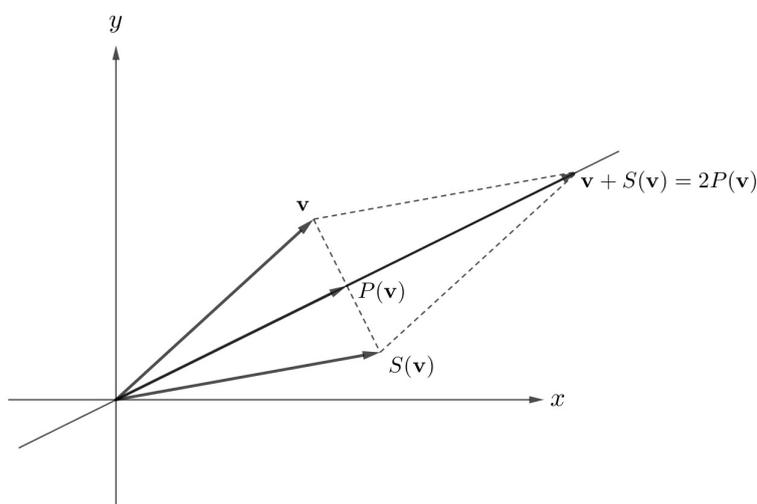


Figura 3.26

Fonte: Adaptado de LIMA [43], p.45

Daí, temos que

$$S(\mathbf{v}) = 2P(\mathbf{v}) - \mathbf{v} = 2\mathbf{P}\mathbf{v} - \mathbf{I}\mathbf{v} = (2\mathbf{P} - \mathbf{I})\mathbf{v}.$$

Assim, temos que a matriz da transformação S é dada por

$$\mathbf{S} = 2\mathbf{P} - \mathbf{I} = 2\alpha \begin{bmatrix} 1 & a \\ a & a^2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2\alpha - 1 & 2\alpha a \\ 2\alpha a & 2\alpha a^2 - 1 \end{bmatrix}$$

Portanto, as coordenadas do vetor $S(\mathbf{v}) = (v'_1, v'_2)$ são dadas por

$$\begin{aligned}v'_1 &= 2\alpha(v_1 + av_2) - v_1 \\v'_2 &= 2\alpha a(v_1 + av_2) - v_2\end{aligned}$$

Quanto à linearidade de S , esta pode ser facilmente verificada, seguindo passos análogos aqueles mostrados no exemplo anterior.

Os casos em que a reflexão ocorre em torno de uma reta $y = ax + b$, com $b \neq 0$ podem ser mais facilmente resolvidos com o auxílio das translações. Por exemplo, se quiséssemos refletir o ponto $(4, 3)$ em torno da reta $y = 2x + 5$ da figura abaixo

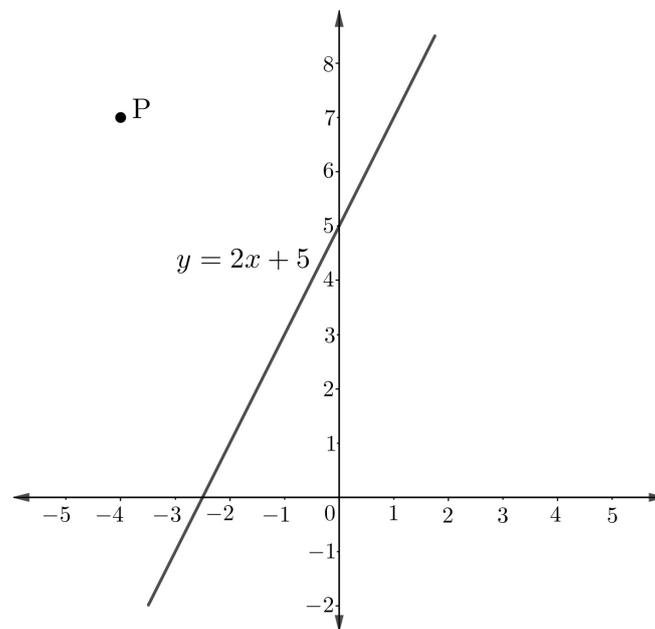


Figura 3.27

Fonte: O autor

bastaria seguir os seguintes passos

- Aplicar a translação $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2, T(x, y) = (x, y) + (0, -5)$ à reta $y = 2x + 5$ e ao ponto $P = (-4, 7)$, de modo a obtermos, respectivamente, a reta $y_T = 2x$ e o ponto $P_T = (-4, 2)$

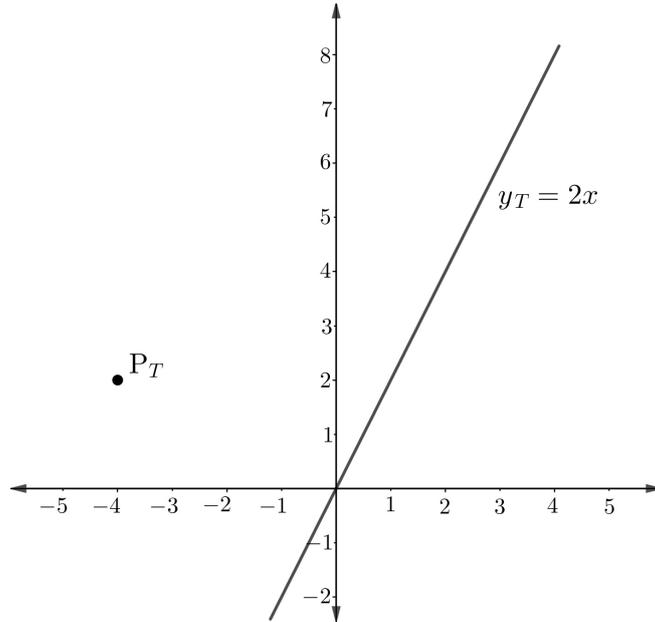


Figura 3.28

Fonte: O autor

- Aplicar a reflexão $S : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ do ponto P_T em torno da reta $y = 2x$, de modo a obter o ponto $P_{ST} = (p_1, p_2)$. Sabendo que $\alpha = \frac{1}{1+2^2} = \frac{1}{5}$, segue que

$$p_1 = 2 \cdot \frac{1}{5}(-4 + 2 \cdot 2) - (-4) = 4$$

$$p_2 = 2 \cdot \frac{1}{5}(-4 + 2 \cdot 2) - 2 = -2$$

ou seja, $P_{ST} = (4, -2)$.

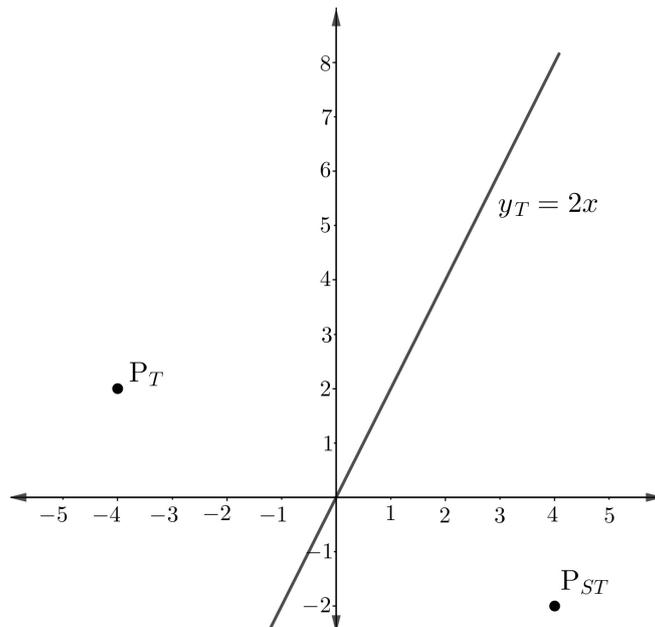


Figura 3.29

Fonte: O autor

- Aplicar a inversa de T , isto é, $T^{-1} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, T^{-1}(x, y) = (x, y) + (0, 5)$, tanto à reta $y = 2x$, como aos pontos P_T e P_{ST} , de modo a obtermos novamente a reta $y = 2x + 5$, o ponto P e a reflexão deste em torno desta última, ou seja, o ponto $P_{TST} = (p'_1, p'_2)$, onde

$$P_{TST} = T^{-1}(P_{ST}) = T^{-1}(4, -2) = (4, -2) + (0, 5) = (4, 3),$$

o que finaliza o processo de reflexão, podendo assi ser observado na imagem abaixo

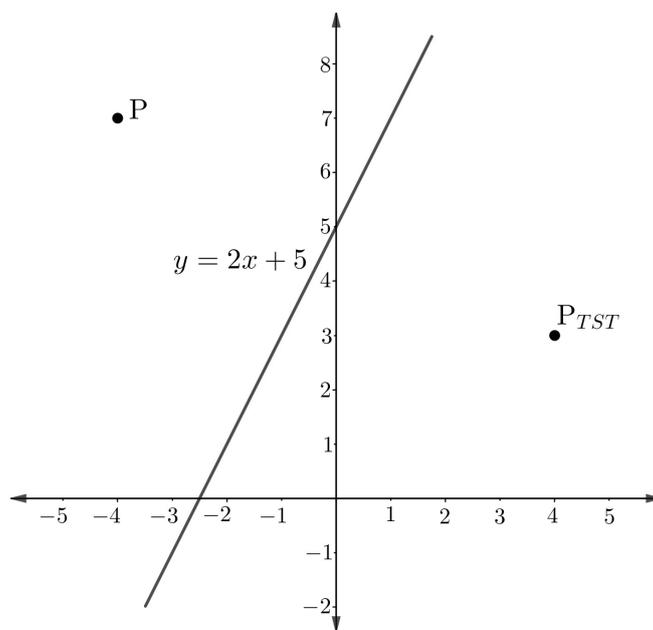


Figura 3.30

Fonte: O autor

A subseção a seguir está diretamente voltada à aplicação das translações, rotações e reflexões às cônicas, um dos principais foco deste trabalho.

3.2.2 Translação, rotação e reflexão de cônicas

Será admitido, a partir desta subseção, que o leitor conhece as cônicas, de modo que não será apresentada aqui a forma como a equação de cada uma delas é obtida, pois acabaríamos digredindo do real foco, que são as aplicações das matrizes em diferentes áreas do conhecimento, tanto dentro como fora da Matemática.

Exemplo 3.8. Considere uma circunferência C , de centro $P = (a, b)$ e raio r . Vamos transladar C por um vetor $\mathbf{v} = (\alpha, \beta)$ de modo a obtermos a circunferência C_T , de mesmo raio, porém com centro $P_T = (c, d)$. Seja então $T_{\mathbf{v}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ uma translação, com $T_{\mathbf{v}}(x, y) = (x + \alpha, y + \beta)$, para todo $x, y \in \mathbb{R}^2$. Agora notemos que, desenvolvendo a equação de C , obtemos

$$\begin{aligned} (x - a)^2 + (y - b)^2 &= r^2 \\ \Rightarrow x^2 + y^2 - 2ax - 2by &= r^2 - (a^2 + b^2) \end{aligned}$$

Em seguida, sabendo que as coordenadas de cada ponto de C_T serão dadas por (x', y') , temos que

$$\begin{aligned}x' = x + \alpha &\iff x = x' - \alpha \\y' = y + \beta &\iff y = y' - \beta\end{aligned}$$

Daí, substituindo x por $x' - \alpha$ e y por $y' - \beta$ na expansão da equação de C , obtemos

$$(x')^2 + (y')^2 - 2(\alpha + a)x' - 2(\beta + b)y' = k, \quad (3.1)$$

onde $k = r^2 - (a^2 + b^2) - 2(\alpha a + \beta b)$. Por outro lado, a equação de C_T nos fornece

$$(x')^2 + (y')^2 - 2c(x') - 2d(y') = r^2 - (c^2 + d^2) \quad (3.2)$$

Assim, para que (3.1) e (3.2) coincidam, é necessário que

$$\begin{aligned}\alpha + a = c &\iff \alpha = c - a \\ \beta + b = d &\iff \beta = d - b\end{aligned}$$

o que nos permite concluir que $\mathbf{v} = (c - a, d - b)$ e, portanto, $T_{\mathbf{v}}$ é dada por

$$T_{\mathbf{v}}(x, y) = (x + c - a, y + d - b)$$

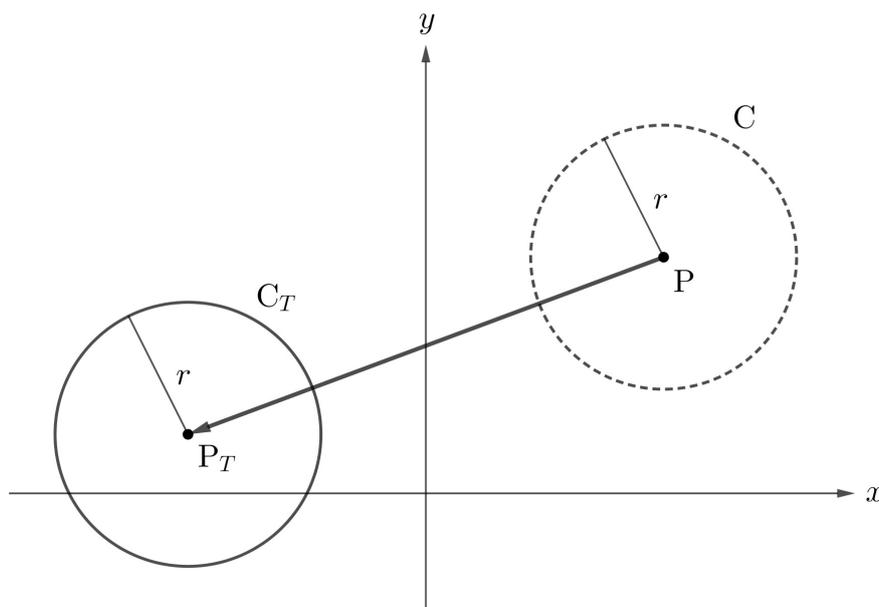


Figura 3.31

Fonte: O autor

O exemplo acima nos mostra um resultado que já é esperado das translações, o qual nos diz, de modo mais simples, que para transladar uma circunferência por um vetor \mathbf{v} , basta aplicarmos tal translação a cada um dos pontos da circunferência.

Vejam os mais um exemplo, apenas para reforçar essa ideia.

Exemplo 3.9. Consideremos a função $f : \mathbb{R} \rightarrow \mathbb{R}$, com $f(x) = ax^2 + bx + c$, $a \neq 0$. Sabemos que seu gráfico $G = \{(x, y) \in \mathbb{R}^2; y = ax^2 + bx + c\}$ é uma parábola. Vamos supor, sem perda de generalidade, que $a > 0$ e $\Delta < 0$. Aplicando a translação $T_{\mathbf{v}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\mathbf{v} = (\alpha, \beta)$, aos pontos $P = (x, y) \in G$

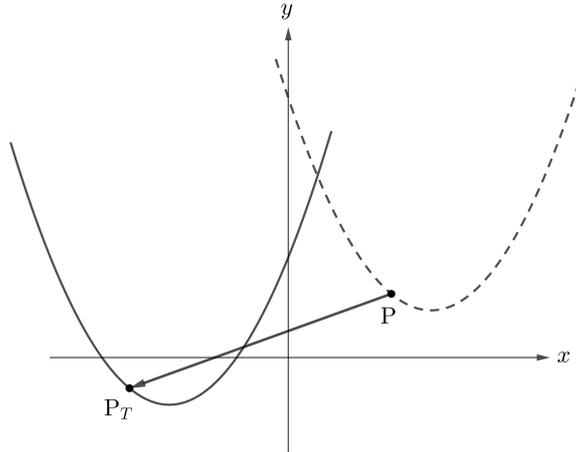


Figura 3.32

Fonte: O autor

temos que as novas coordenadas da parábola serão dadas por pontos $P_T = (x', y') = (x + \alpha, y + \beta)$, que equivale a

$$\begin{aligned}x &= x' - \alpha \\ y &= y' - \beta\end{aligned}$$

e substituindo as expressões acima em $y = ax^2 + bx + c$, obtemos

$$\begin{aligned}\Leftrightarrow \quad y' - \beta &= a(x' - \alpha)^2 + b(x' - \alpha) + c \\ y' &= a(x')^2 + (-2\alpha a + b)x' + (a\alpha^2 - b\alpha + c + \beta) \\ &= a'(x')^2 + b'x' + c',\end{aligned}$$

onde

$$\begin{aligned}a' &= a \\ b' &= -2\alpha a + b \\ c' &= a\alpha^2 - b\alpha + c + \beta.\end{aligned}$$

Algo que podemos observar é que, os valores de c e β podem interferir no número de raízes reais da equação $a'(x')^2 + b'x' + c' = 0$, uma vez que o discriminante Δ' da parábola $G' = \{(x', y') \in \mathbb{R}^2; y' = a'(x')^2 + b'x' + c'\}$ é

$$\begin{aligned}\Delta' &= (b')^2 - 4a'c' \\ &= (-2\alpha a + b)^2 - 4a(a\alpha^2 - b\alpha + c + \beta) \\ &= b^2 - 4a(c + \beta).\end{aligned}$$

e sendo o número de raízes reais um ou dois apenas quando $\Delta' \geq 0$, decorre que

$$b^2 - 4a(c + \beta) \geq 0 \Leftrightarrow b \geq \pm 2\sqrt{c + \beta},$$

isto é, só existem raízes reais quando $c \geq -\beta$. Nos casos em que $b = 0$, para que seja $\Delta' \geq 0$, é necessário que $a < 0$ e $c + \beta \geq 0$ ou $a > 0$ e $c + \beta \leq 0$.

É importante lembrar que uma cônica pode assumir qualquer posição no plano, de modo que, por exemplo, uma parábola de vértice V pode não apresentar o seu eixo de simetria paralelo aos eixos x ou y , mas concorrente a ambos, como mostra a figura a seguir

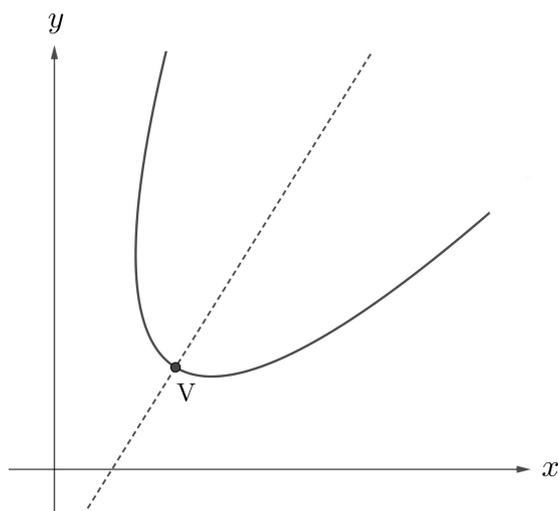


Figura 3.33

Fonte: O autor

Vejamos que sempre é possível obter uma parábola como a que aparece na figura acima, partindo de uma parábola na posição padrão, isto é, com eixo de simetria paralelo a um dos eixos x ou y , por meio da composição duas translações e uma rotação. Devemos antes lembrar que, uma parábola rotacionada não apresenta uma lei de formação padrão, isto é, a forma $ax^2 + bx + c$. A forma geral de uma função quadrática possui duas variáveis e é dada por $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, tal que

$$\varphi(x, y) = Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F,$$

onde A, B, C, D, E e F são constantes reais. De modo mais geral, uma função como a citada acima é chamada **forma quadrática**, porém não trataremos aqui os detalhes sobre tais formas. Para uma abordagem minuciosa, consultar LIMA [44].

Vamos tomar, sem perda de generalidade, uma parábola $G = \{(x, y); y = ax^2 + bx + c, a \neq 0\}$ que possua eixo de simetria paralelo ao eixo y e seja $a > 0$. Procederemos de forma similar ao que foi realizado no exemplo de reflexão realizado na seção anterior. Primeiro, transladamos a parábola de modo que seu vértice $(h, k) = \left(-\frac{b}{2a}, -\frac{\Delta}{4a}\right)$ coincida com a origem (Figura 3.27). Para tanto, basta fazer $\alpha = -h$ e $\beta = -k$ no resultado obtido no Exemplo 3.9, de modo a obtermos a parábola G_T , tal que

$$y' = a(x')^2 + (2ah + b)x' + (ah^2 + bh + c - k) \quad (3.3)$$

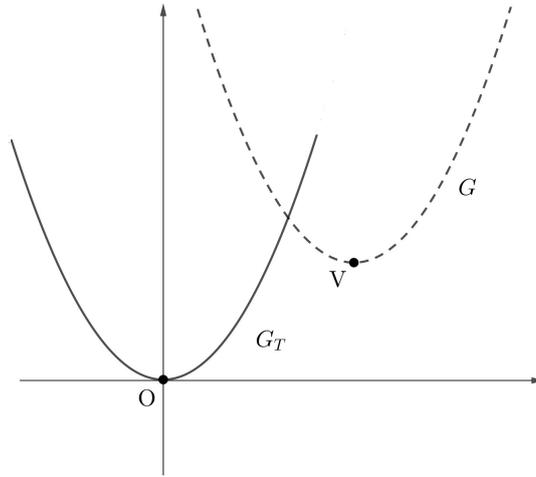


Figura 3.34

Fonte: O autor

A expressão acima se torna mais simples, uma vez que $2ah + b = 0$ e $ah^2 + bh + c - k = 0$, reduzindo-se então a

$$y' = a(x')^2$$

Agora, aplicamos uma rotação de ângulo θ no sentido horário, obtendo assim a parábola G_{RT} , a qual pode ser visualizada na figura a seguir

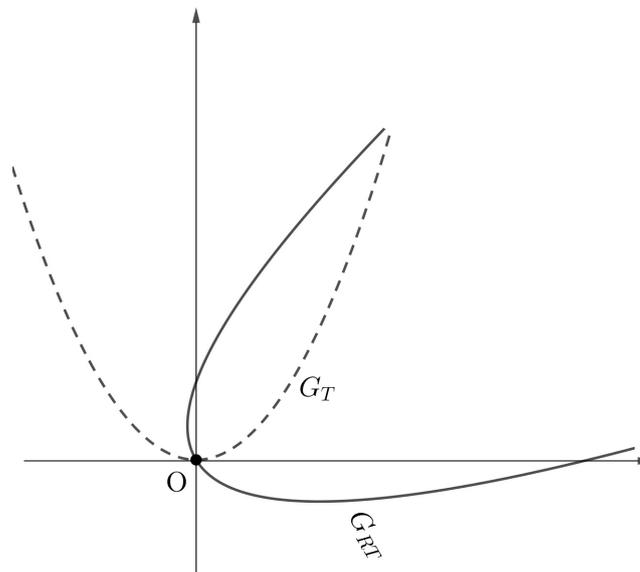


Figura 3.35

Fonte: O autor

onde suas coordenadas são dadas por

$$\begin{aligned} x'' &= x' \cos(\theta) + y' \sin(\theta) \\ y'' &= -x' \sin(\theta) + y' \cos(\theta) \end{aligned}$$

Uma vez que $\begin{vmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{vmatrix} = 1 \neq 0$, temos que o sistema acima possui solução única,

a qual, pela Regra de Cramer, é

$$x' = \frac{\begin{vmatrix} x'' & \text{sen}(\theta) \\ y'' & \text{cos}(\theta) \end{vmatrix}}{\begin{vmatrix} \text{cos}(\theta) & \text{sen}(\theta) \\ -\text{sen}(\theta) & \text{cos}(\theta) \end{vmatrix}} = x'' \text{cos}(\theta) - y'' \text{sen}(\theta)$$

$$y' = \frac{\begin{vmatrix} \text{cos}(\theta) & x'' \\ -\text{sen}(\theta) & y'' \end{vmatrix}}{\begin{vmatrix} \text{cos}(\theta) & \text{sen}(\theta) \\ -\text{sen}(\theta) & \text{cos}(\theta) \end{vmatrix}} = x'' \text{sen}(\theta) + y'' \text{cos}(\theta)$$

Substituindo os valores acima em (3.3), obtemos uma expressão da forma

$$A(x'')^2 + 2Bx''y'' + C(y'')^2 + Dx'' + Ey'' + F = 0, \quad (3.4)$$

onde

$$\begin{aligned} A &= a \text{cos}^2(\theta) & D &= -\text{cos}(\theta) \\ B &= -a \text{sen}(\theta) \text{cos}(\theta) & E &= -\text{sen}(\theta) \\ C &= a \text{sen}^2(\theta) & F &= 0 \end{aligned}$$

OBS.: Note que se, ao invés de termos aplicado a translação sobre o vértice da parábola, mas sobre um ponto (p, q) qualquer da mesma, os coeficientes A , B e C seriam os mesmos, ao passo que

$$\begin{aligned} D &= (2ap + b - 1) \text{cos}(\theta) \\ E &= -(2ap + b + 1) \text{sen}(\theta) \\ F &= ap^2 + bp + c - q \end{aligned}$$

Por fim, aplicamos a translação $T_{-\mathbf{v}}$, com $\mathbf{v} = (h, k)$, o que nos fornece a parábola G_{TRT} (Figura 3.29), cujos pontos possuem coordenadas

$$(\bar{x}, \bar{y}) = (x'' + h, y'' + k)$$

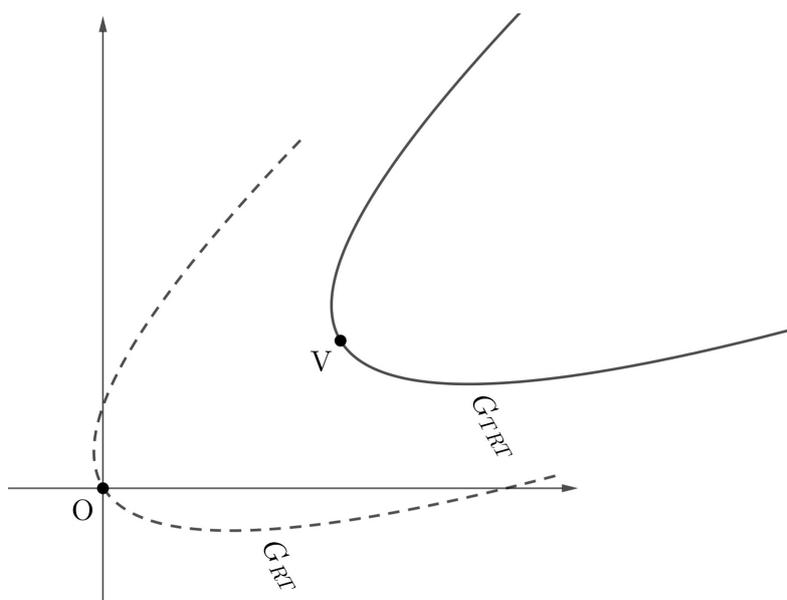


Figura 3.36

Fonte: O autor

Daí, substituindo $x'' = \bar{x} - h$ e $y'' = \bar{y} - k$ em (3.4) e reagrupando os termos semelhantes, obtemos

$$A\bar{x}^2 + 2B\bar{x}\bar{y} + C\bar{y}^2 + \bar{D}\bar{x} + \bar{E}\bar{y} + \bar{F} = 0,$$

onde

$$\begin{aligned}\bar{D} &= -2Ah - 2Bk + D \\ \bar{E} &= -2Bh - 2Ck + E \\ \bar{F} &= Ah^2 + 2Bhk + Ck^2 - Dh - Ek + F\end{aligned}$$

Apesar de termos utilizado uma parábola como exemplo, os resultados obtidos acima são análogos para uma forma quádrlica qualquer $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, com

$$\varphi(x, y) = Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F.$$

Desse modo, se realizarmos translações a esta, os seus coeficientes A , B e C não se alteram, o que nos permite dizer que estes são invariantes por translações, enquanto que os demais coeficientes se tornam

$$\begin{aligned}D' &= -2Ah - 2Bk + D \\ E' &= -2Bh - 2Ck + E \\ F' &= Ah^2 + 2Bhk + Ck^2 - Dh - Ek + F\end{aligned}$$

Exemplo 3.10. Vamos aplicar à parábola $G = \{(x, y); y = x^2 - 5x + 6\}$ uma rotação de 45° no sentido horário em torno de seu vértice (h, k) utilizando os passos acima. Sabemos que

$$(h, k) = \left(-\frac{b}{2a}, -\frac{\Delta}{4a}\right) = \left(\frac{5}{2}, -\frac{1}{4}\right).$$

Transladando G para origem pelo vetor $\mathbf{v} = \left(-\frac{5}{2}, \frac{1}{4}\right)$, obtemos G_T , onde

$$y' = (x')^2.$$

Aplicando a rotação de 45° à G_T , obtemos a parábola G_{RT} , cujos coeficientes são

$$\begin{aligned}A &= 1 \cdot \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2} & D &= -\frac{1}{\sqrt{2}} = -\frac{\sqrt{2}}{2} \\ B &= -1 \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = -\frac{1}{2} & E &= -\frac{1}{\sqrt{2}} = -\frac{\sqrt{2}}{2} \\ C &= 1 \cdot \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2} & F &= 0,\end{aligned}$$

ou seja, sua equação é dada por

$$\frac{1}{2}x^2 - xy + \frac{1}{2}y^2 - \frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y = 0.$$

Transladando G_{RT} por $\mathbf{v} = \left(\frac{5}{2}, -\frac{\Delta}{4a}\right)$, obtemos G_{TRT} , cujos coeficientes invariantes por translações são

$$\begin{aligned}D' &= -2 \cdot \frac{1}{2} \cdot \frac{5}{2} - 2 \left(-\frac{1}{2}\right) \left(-\frac{1}{4}\right) - \frac{\sqrt{2}}{2} = \frac{-11-2\sqrt{2}}{4} \\ E' &= -2 \left(-\frac{1}{2}\right) \cdot \frac{5}{2} - 2 \cdot \frac{1}{2} \left(-\frac{1}{4}\right) - \frac{\sqrt{2}}{2} = \frac{11-2\sqrt{2}}{4} \\ F' &= \frac{25}{8} + \frac{5}{8} + \frac{1}{32} + \frac{5\sqrt{2}}{4} - \frac{\sqrt{2}}{8} + 0 = \frac{121+36\sqrt{2}}{32},\end{aligned}$$

o que finalmente nos fornece sua equação

$$\frac{1}{2}x^2 - xy + \frac{1}{2}y^2 - \frac{11+2\sqrt{2}}{4}x + \frac{11-2\sqrt{2}}{4}y + \frac{121+36\sqrt{2}}{32} = 0$$

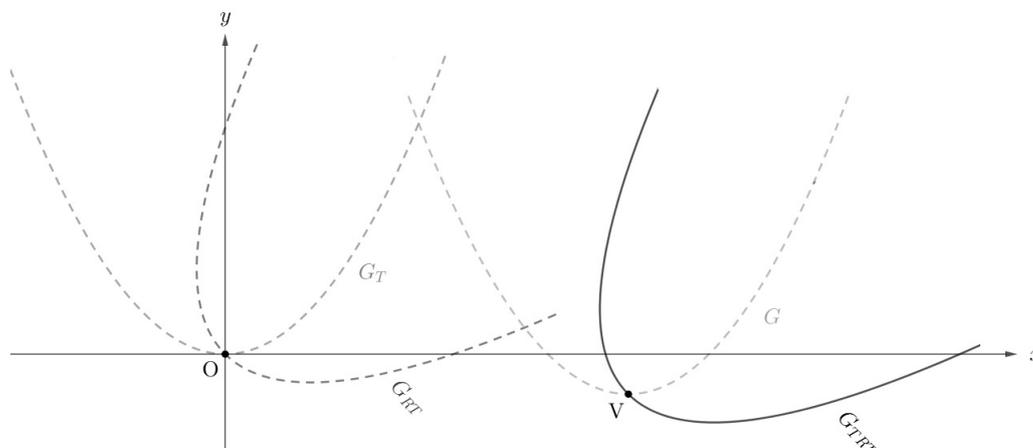


Figura 3.37

Fonte: O autor

Feita a discussão acima, uma pergunta pode ocorrer: “o que acontece com as raízes de nova parábola?” Bem, as raízes de G são 2 e 3. Ao aplicarmos a translação sobre o vértice de G , a parábola G_T obtida possui apenas uma raiz, a qual é 0. Os pontos $(2, 0)$ e $(3, 0)$ se tornam $(-\frac{1}{2}, \frac{1}{4})$ e $(\frac{1}{2}, \frac{1}{4})$, deixando de assumir o papel de raízes, tornando-se apenas soluções da equação $y = x^2$. Quando a rotação é realizada, podemos observar que a parábola G_{RT} obtida novamente intercepta o eixo x em dois pontos, os quais fornecem as raízes de tal parábola. Uma delas obviamente é $x = 0$, mas vamos determiná-las de maneira formal. Para tanto, vamos multiplicar a equação abaixo por 2 em ambos os membros

$$\frac{1}{2}x^2 - xy + \frac{1}{2}y^2 - \frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y = 0.$$

e em seguida reescrevê-la em termos de x , de forma a obtermos

$$x^2 + (-2y - \sqrt{2})x + y^2 - \sqrt{2}y = 0.$$

Agora notemos que, resolvendo a equação quadrática acima, obtemos

$$x = y + \frac{\sqrt{2}}{2} \pm \frac{1}{2}\sqrt{8\sqrt{2}y + 2}$$

e fazendo $y = 0$, segue que $x = 0$ ou $x = \sqrt{2}$, que são as raízes de G_{RT} .

A fim de determinar as raízes de G_{TRT} , basta realizar o mesmo processo realizado acima, ou seja, reescrevemos a equação de G_{TRT} em termos de x , ou seja

$$x^2 + \left[-2y - \left(\frac{11+2\sqrt{2}}{2}\right)\right]x + y^2 + \frac{11-2\sqrt{2}}{2}y + \frac{121+36\sqrt{2}}{16} = 0.$$

Resolvendo-a, obtemos

$$x = y + \frac{11+2\sqrt{2}}{4} \pm \frac{1}{2}\sqrt{8\sqrt{2}y + 2\sqrt{2} + 2}.$$

Tomando $y = 0$ e $\sqrt{2} \cong 1,414$, obtemos

$$x = \frac{11 + 2\sqrt{2}}{4} + \frac{1}{2}\sqrt{2\sqrt{2} + 2} \cong 4,56$$

ou

$$x = \frac{11 + 2\sqrt{2}}{4} - \frac{1}{2}\sqrt{2\sqrt{2} + 2} \cong 2,36,$$

onde tais valores são as respectivas raízes de G_{TRT} .

Vejamos a seguir que, assim como é possível rotacionar e transladar cônicas, também é possível reflexioná-las em torno de uma reta qualquer.

Exemplo 3.11. Vamos reflexionar a parábola $G = \{(x, y); y = x^2 - 4x + 5\}$ em torno da reta r de equação $y = -2x + 1$. Para tanto, devemos determinar o foco e a reta diretriz de G . Ora, completando quadrados em $x^2 - 4x + 5 = y$, obtemos

$$(x - 2)^2 = 4 \cdot \frac{1}{4}(y - 1),$$

o que nos fornece o foco $F = \left(2, \frac{5}{4}\right)$ e a reta diretriz $y_d = \frac{3}{4}$ de G . Daí,

- Transladando r , F e $y_d = \frac{3}{4}$ pelo vetor $\mathbf{v} = (0, -1)$, obtemos

$$y' = -2x, \quad F' = \left(2, \frac{1}{4}\right) \quad \text{e} \quad y'_d = -\frac{1}{4}.$$

- Reflexionando F' em torno de y' , obtemos

$$F'' = \left(-\frac{7}{5}, -\frac{29}{20}\right).$$

Agora, reflexionando $y'_d = -\frac{1}{4}$, isto é, pontos da forma $\left(x'_d, -\frac{1}{4}\right)$, em torno de y' , obtemos pontos (x''_d, t''_d) , tais que

$$x''_d = 2 \cdot \frac{1}{5} \cdot \left(x'_d + (-2) \cdot \left(-\frac{1}{4}\right)\right) - x'_d = -\frac{3}{5}x'_d + \frac{1}{5} \Rightarrow x'_d = -\frac{5}{3}x''_d + \frac{1}{3}$$

$$y''_d = 2 \cdot \frac{1}{5} \cdot (-2) \cdot \left(-\frac{5}{3}x''_d + \frac{1}{3} + (-2) \cdot \left(-\frac{1}{4}\right)\right) - \left(-\frac{1}{4}\right) = \frac{4}{3}x''_d - \frac{5}{12}$$

- Agora, transladando F'' e $y''_d = \frac{4}{3}x''_d - \frac{5}{12}$ pelo vetor $\mathbf{w} = (0, 1)$, temos

$$F''' = \left(-\frac{7}{5}, -\frac{9}{20}\right) \quad \text{e} \quad y'''_d = \frac{4}{3}x'''_d + \frac{7}{12}.$$

Obtivemos com os passos acima, tanto o foco F''' , que chamaremos simplesmente de F_X , como a reta diretriz y'''_d da parábola G_X , ou seja, a reflexão de G em torno de r . Por uma questão de simplicidade, vamos escrever a reta diretriz de G_X apenas como $y = \frac{4}{3}x + \frac{7}{12}$, a qual chamaremos de y_{dX} . Logo, para determinar a equação de G_X , basta

utilizarmos a definição de parábola, ou seja, encontrarmos o conjunto de pontos (x, y) que são equidistantes de F''' e y_d''' . Para tanto, notemos que

$$\begin{aligned} d_1 = d((x, y), F''') &= \sqrt{\left(x + \frac{7}{5}\right)^2 + \left(y + \frac{9}{20}\right)^2} \\ &= \sqrt{x^2 + y^2 + \frac{14}{5}x + \frac{9}{10}y + \frac{173}{80}} \end{aligned}$$

e

$$d_2 = d((x, y), r) = \frac{|16x - 12y + 7|}{\sqrt{16^2 + 12^2}}$$

Daí, como devemos ter $d_1 = d_2$, elevando esta última ao quadrado e reordenando os termos, obtemos

$$\frac{9}{25}x^2 + \frac{24}{25}xy + \frac{16}{25}y^2 + \frac{56}{25}x + \frac{33}{25}y + \frac{51}{25} = 0,$$

que é exatamente a equação de G_X .

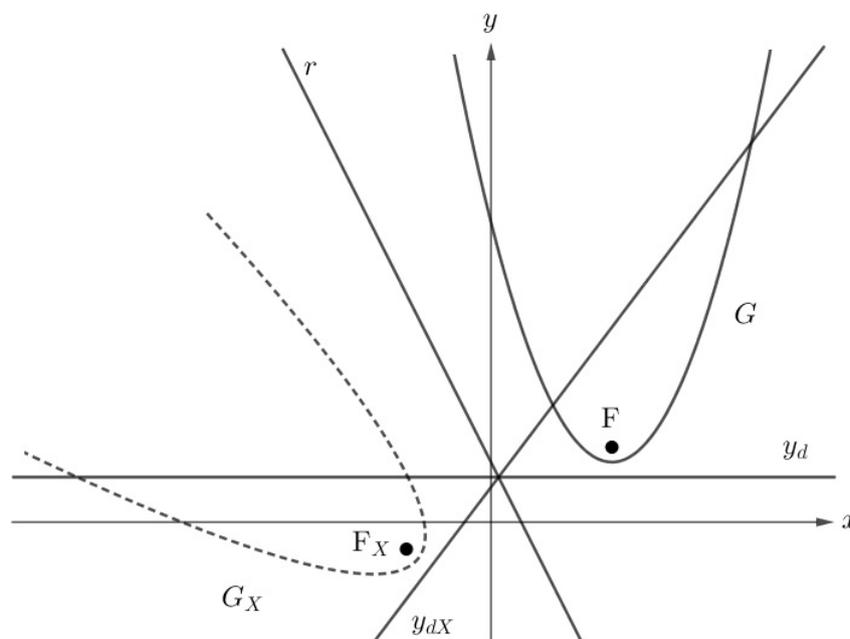


Figura 3.38

Fonte: O autor

3.2.3 Identificando formas quadráticas

Anteriormente, vimos como realizar uma translação de vetor \mathbf{v} e rotação de ângulo θ a uma forma quadrática (a qual, em seus casos mais interessantes, assumem a forma de ua cônica). Porém, como fazer o processo reverso? isto é, como partir de uma forma quadrática qualquer e descobrir se ela é uma elipse, uma parábola ou uma hipérbole? Bem, sabemos que uma forma quadrática pode ser facilmente identificada, desde que esteja sob a forma

$$\varphi(x, y) = Mx^2 + Ny^2$$

onde M, N são números reais não ambos nulos. Logo, dependendo dos valores de M e N , φ e tomando $\varphi(x, y) = 0$ (pois este é o caso que nos interessa) pode assumir a forma de uma cônica (parábola, elipse ou hipérbole). Neste caso, dada uma forma quadrática

$$\varphi(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$$

temos de encontrar uma forma de colocá-la em sua forma padrão, isto é, centrada na origem, e então aplicar as transformações lineares convenientes (translação e rotação) a fim de identificar sua posição original.

Como já discutido na subseção anterior, se transladarmos uma forma quadrática como a que temos acima por um vetor $\mathbf{v} = (\alpha, \beta)$, temos que os coeficientes A, B e C não se alteram, enquanto que

$$\begin{aligned} D' &= -2A\alpha - 2B\beta + D \\ E' &= -2B\alpha - 2C\beta + E \\ F' &= A\alpha^2 + 2B\alpha\beta + C\beta^2 - D\alpha - E\beta + F \end{aligned}$$

Note que devemos nos preocupar em eliminar os termos D' e E' , pois estes são termos (também chamados de **termos lineares**) diretamente ligados a x e y (ou seja, podem alterar o tipo de forma quadrática) respectivamente, enquanto que F' influi apenas na posição que a forma quadrática vai assumir. Neste caso, φ se reduzirá a

$$\bar{\varphi}(x, y) = Ax^2 + Bxy + Cy^2 + F'$$

Note que ainda temos o termo B , que também deverá ser eliminado, mas já sabemos que este surge da rotação, então vamos atentar inicialmente para $D' = E' = 0$. Ora, se isso ocorre, então obtemos o seguinte sistema

$$\begin{cases} A\alpha + B\beta = \frac{D'}{2} \\ B\alpha + C\beta = \frac{E'}{2} \end{cases}$$

cujas matriz é

$$\mathcal{M} = \begin{bmatrix} A & B \\ B & C \end{bmatrix},$$

a qual será chamada de **matriz de φ** . Logo, desde que $AC - B^2 \neq 0$, o sistema acima possui uma única solução, a qual é, pela regra de Cramer

$$\alpha = \frac{CD - BE}{2(AC - B^2)} \quad \beta = \frac{AE - BD}{2(AC - B^2)}$$

Exemplo 3.12 Considere a forma quadrática $\varphi(x, y) = 7x^2 - 6xy + 7y^2 + 2x + 10y - 4$. Vamos eliminar os termos lineares da mesma. Para tanto, temos que a matriz de φ é

$$\mathcal{M} = \begin{bmatrix} 7 & -3 \\ -3 & 7 \end{bmatrix},$$

onde $\det(\mathcal{M}) = 40 \neq 0$. Daí, temos que

$$\begin{aligned} \alpha &= \frac{2 \cdot 7 - (-3) \cdot 10}{2 \cdot 40} = \frac{11}{20} \\ \beta &= \frac{7 \cdot 10 - (-3) \cdot 2}{2 \cdot 40} = \frac{19}{20} \end{aligned}$$

Portanto, o vetor que elimina os termos $2x$ e $10y$ é $\mathbf{v} = (\frac{11}{20}, \frac{19}{20})$. Calculando F' , obtemos $F' = -\frac{93}{10}$, de modo que a nova forma quadrática possui a forma

$$\overline{\varphi}(x, y) = 7x^2 - 6xy + 7y^2 - \frac{93}{10}.$$

O exemplo acima nos mostra, na prática, como eliminar os termos lineares. Agora precisamos encontrar uma forma de eliminar o termo que contém xy . Como já sabemos, este surge após uma rotação. Então, sem perda de generalidade, tomemos uma forma quádrlica com $D = E = F = 0$, ou seja, sob a forma

$$\varphi(x, y) = Ax^2 + Bxy + Cy^2.$$

Aplicando uma rotação de ângulo θ , podemos fazer a substituição

$$\begin{aligned} x &\longmapsto ax - by \\ y &\longmapsto bx + ay \end{aligned}$$

onde $a = \cos(\theta)$, $b = \sin(\theta)$, tal que

$$\varphi(x, y) = \varphi(ax - by, bx + ay) = \overline{\overline{\varphi}}(x, y) = A'x^2 + B'xy + C'y^2,$$

com

$$\begin{aligned} A' &= Aa^2 + 2Bab + Cb^2 \\ B' &= -Aab + B(a^2 - b^2) + Cab \\ C' &= Ab^2 - 2Bab + Ca^2 \end{aligned}$$

Note que B' pode ser reescrito da forma

$$B' = a(Ba + Cb) - b(Aa + Bb) = \begin{vmatrix} a & Aa + Bb \\ b & Ba + Cb \end{vmatrix} = \begin{vmatrix} a & b \\ Aa + Bb & Ba + Cb \end{vmatrix}.$$

Para que seja $B' = 0$, então a matriz

$$\begin{bmatrix} a & b \\ Aa + Bb & Ba + Cb \end{bmatrix}$$

deve possuir uma linha (ou coluna) nula ou uma linha (ou coluna) seja múltipla da outra. Se for o primeiro caso, então, temos as seguintes possibilidades

- $(a, b) = (0, 0)$: absurdo, uma vez que (a, b) é um vetor unitário.
- $(Aa + Bb, Ba + Cb) = (0, 0)$: Nesse caso, teríamos obviamente $B' = 0$, porém

$$A' = a(Aa + Bb) + b(Ba + Cb) = 0.$$

Além disso, escrevendo C' da seguinte forma

$$C' = b(Ab - Ba) - a(Bb - Ca)$$

e notando que $Ba = -Cb$ e $Bb = -Aa$, segue que

$$C' = b(Ab + Cb) + a(Aa + Ca) = (A + C)(a^2 + b^2) = A + C,$$

reduzindo $\overline{\overline{\varphi}}$ a

$$\overline{\overline{\varphi}}(x, y) = (A + C)y^2.$$

Daí, teríamos $\overline{\overline{\varphi}}(x, y) = 0$ apenas se $A + C = 0$, e neste caso $\overline{\overline{\varphi}}$ seria reduzida a um ponto, o que é absurdo. Se $A + C \neq 0$, então $y = 0$ e $\overline{\overline{\varphi}}$ novamente se reduz a um ponto.

- $(a, Aa + Bb) = (0, 0)$: Aqui temos $a = 0$, o que implica $b = 1$ (pois (a, b) é unitário), de modo que

$$Aa + Bb = 0 \Rightarrow Bb = 0 \Rightarrow B = 0.$$

Assim, a matriz em estudo se condensa em

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

onde a segunda linha é múltipla da primeira, bastando multiplicá-la por C .

- $(b, Ba + Cb) = (0, 0)$: análogo ao item anterior.

Logo, de acordo com o que foi discutido acima, podemos constatar que o vetor $(Aa + Bb, Ba + Cb)$ é múltiplo do vetor (a, b) , ou seja, existe uma constante λ , tal que

$$Aa + Bb = \lambda a \quad \text{e} \quad Ba + Cb = \lambda b$$

fornecendo-nos assim o seguinte sistema

$$\begin{cases} (A - \lambda)a + Bb = 0 \\ Ba + (C - \lambda)b = 0 \end{cases}$$

onde (a, b) deve ser uma solução não-trivial para que se tenha $B' = 0$, portanto, o determinante $(A - \lambda)(C - \lambda) - B^2$ da matriz do sistema acima deve ser igual a zero, ou seja,

$$\lambda^2 - (A + C)\lambda + AC - B^2 = 0.$$

A equação acima é chamada, devido à sua importância, de **equação característica** da matriz de φ . Note que esta sempre possui raízes reais λ_1 e λ_2 , pois seu discriminante é não negativo, como podemos ver abaixo

$$\begin{aligned} \Delta &= (A + C)^2 - 4 \cdot 1 \cdot (AC - B^2) \\ &= A^2 + 2AC + C^2 - 4AC + 4B^2 \\ &= A^2 - 2AC + C^2 + 4B^2 \\ &= (A - C)^2 + 4B^2 \\ &\geq 0. \end{aligned}$$

Além disso, como $\text{tr}(\mathcal{M}) = A + C$ e $\det(\mathcal{M}) = AC - B^2$, podemos escrever a equação característica simplesmente como

$$\lambda^2 - \text{tr}(\mathcal{M})\lambda + \det(\mathcal{M}) = 0$$

Dito isto, resta agora vermos como determinar o vetor (a, b) , que por sua vez fornece o ângulo de rotação necessário para eliminar o termo B' . Para tanto, uma vez que o sistema acima possui infinitas soluções, podemos tomar, por exemplo, $x = 1$, o que implica $y = (\lambda - A)/B$. Daí, como o vetor (a, b) é unitário, basta então tomarmos

$$a = \frac{x}{\|(x, y)\|} = \frac{x}{\sqrt{x^2 + y^2}} \quad \text{e} \quad b = \frac{y}{\|(x, y)\|} = \frac{y}{\sqrt{x^2 + y^2}}.$$

Note que (a, b) continua sendo solução do sistema com $\lambda = \lambda_1$, pois

$$Aa + Bb = \frac{Ax + By}{\|(x, y)\|} = \frac{\lambda_1 x}{\|(x, y)\|} = \lambda_1 \frac{x}{\|(x, y)\|} = \lambda_1 a$$

$$Ba + Cb = \frac{Bx + Cy}{\|(x, y)\|} = \frac{\lambda_1 y}{\|(x, y)\|} = \lambda_1 \frac{y}{\|(x, y)\|} = \lambda_1 b,$$

que podemos escrever de forma resumida como

$$\begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \lambda_1 \begin{bmatrix} a \\ b \end{bmatrix},$$

ou seja, o vetor (a, b) está relacionado à raiz λ_1 . Agora, notemos que o vetor $(-b, a)$, perpendicular a (a, b) está associado à raiz λ_2 . De fato, como

$$\lambda_1 + \lambda_2 = -\frac{-(A+C)}{1} = A+C \Rightarrow \lambda_1 = A+C - \lambda_2,$$

decorre que

$$\begin{aligned} Ba + Cb = \lambda_1 b = (A+C - \lambda_2)b &\Rightarrow A(-b) + Ba = \lambda_2(-b) \\ Aa + Bb = \lambda_1 a = (A+C - \lambda_2)a &\Rightarrow B(-b) + Ca = \lambda_2 a, \end{aligned}$$

ou seja,

$$\begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} -b \\ a \end{bmatrix} = \lambda_2 \begin{bmatrix} -b \\ a \end{bmatrix}.$$

Uma vez encontrado o vetor (a, b) (e consequentemente $(-b, a)$), o termo B' é eliminado, de modo a termos

$$\varphi(x, y) = A'x^2 + C'y^2,$$

porém, sendo

$$\begin{aligned} A' &= a(Aa + Bb) + b(Ba + Cb) = \lambda_1 a^2 + \lambda_1 b^2 = \lambda_1 \\ C' &= b(Ab - Ba) - a(Bb - Ca) = \lambda_2 b^2 + \lambda_2 a^2 = \lambda_2 \end{aligned}$$

decorre

$$\varphi(x, y) = \lambda_1 x^2 + \lambda_2 y^2.$$

Podemos então concluir que, ao eliminar o termo B' , os coeficientes A' e C' não precisam mais ser calculados, pois estes são iguais a λ_1 e λ_2 , respectivamente.

Exemplo 3.13. Vamos aplicar o processo à forma quadrática

$$\overline{\varphi}(x, y) = 7x^2 - 6xy + 7y^2 - \frac{93}{10},$$

obtida no Exemplo 3.12. Temos que sua equação característica é dada por

$$\lambda^2 - 14\lambda + 40 = 0,$$

cujas raízes são $\lambda_1 = 4$ e $\lambda_2 = 10$. Estas informações já são suficientes para identificar a forma quadrática em questão, pois após a rotação, obtemos

$$\overline{\overline{\varphi}}(x, y) = 4x^2 + 10y^2 - \frac{93}{10}.$$

Assim, quando $\overline{\overline{\varphi}}(x, y) = 0$, temos

$$4x^2 + 10y^2 = \frac{93}{10}$$

que pode ser reescrita como

$$\frac{x^2}{\left(\sqrt{\frac{93}{40}}\right)^2} + \frac{y^2}{\left(\sqrt{\frac{93}{100}}\right)^2} = 1,$$

o que nos leva a concluir que $\overline{\varphi}(x, y) = 0$ é uma elipse centrada na origem.

Daí, para determinar o ângulo de rotação, devemos calcular as coordenadas dos vetores (a, b) e $(-b, a)$. Já sabemos que (a, b) (associado a λ_1) deve ser solução do sistema

$$\begin{cases} (7 - \lambda_1)x - 3y = 0 \\ -3x + (7 - \lambda_1)y = 0 \end{cases}$$

Uma vez que $\det(\mathcal{M}) \neq 0$, temos que o mesmo possui infinitas soluções, de modo que tomando $x = 1$, decorre $y = (7 - 4)/3 = 1$, o que nos leva a

$$\cos(\theta) = a = \frac{1}{\sqrt{1^2 + 1^2}} = \frac{\sqrt{2}}{2}$$

$$\text{sen}(\theta) = b = \frac{1}{\sqrt{1^2 + 1^2}} = \frac{\sqrt{2}}{2}$$

e assim

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

Logo, a matriz de rotação é dada por

$$\begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

ou seja, uma rotação de $\theta = 45^\circ$ no sentido horário elimina o termo $-6xy$.

Temos, de acordo com os exemplos 3.11 e 3.12 que $\varphi(x, y) = 0$, $\overline{\varphi}(x, y) = 0$ e $\overline{\overline{\varphi}}(x, y) = 0$ são elipses em diferentes posições no plano xy , as quais chamaremos, por simplicidade, de E_1 , E_2 e E_3 , respectivamente. Se quisermos obter informações a cerca de E_1 , como por exemplo, as coordenadas de seus focos, deveremos primeiro encontrar os focos de E_3 , em seguida os focos de E_2 e por último, os de E_1 , ou seja, de modo resumido, basta reverter o processo (Figura 3.31). Como sabemos, a distância focal de E_3 é dada por

$$c = \sqrt{\frac{93}{40} - \frac{93}{100}} = \frac{3\sqrt{62}}{20},$$

logo, seus focos são dados por

$$\overline{\overline{F}}_1 = \left(-\frac{3\sqrt{62}}{20}, 0\right) \quad \text{e} \quad \overline{\overline{F}}_2 = \left(\frac{3\sqrt{62}}{20}, 0\right).$$

Para determinar os focos de E_2 , basta aplicar a rotação de 45° no sentido anti-horário a $\overline{\overline{F}}_1$ e $\overline{\overline{F}}_2$, ou seja,

$$\overline{F}_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{3\sqrt{62}}{20} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3\sqrt{31}}{20} \\ \frac{3\sqrt{31}}{20} \end{bmatrix}$$

e

$$\bar{F}_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} -\frac{3\sqrt{62}}{20} \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{3\sqrt{31}}{20} \\ -\frac{3\sqrt{31}}{20} \end{bmatrix}$$

Finalmente, aplicando a translação pelo vetor $\mathbf{v} = (-\frac{11}{20}, -\frac{19}{20})$ a \bar{F}_1 e \bar{F}_2 , obtemos F_1 e F_2 abaixo

$$F_1 = \left(\frac{3\sqrt{31} - 11}{20}, \frac{3\sqrt{31} - 19}{20} \right) \quad \text{e} \quad F_2 = \left(\frac{-3\sqrt{31} - 11}{20}, \frac{-3\sqrt{31} - 19}{20} \right)$$

que são os focos de E_1 .

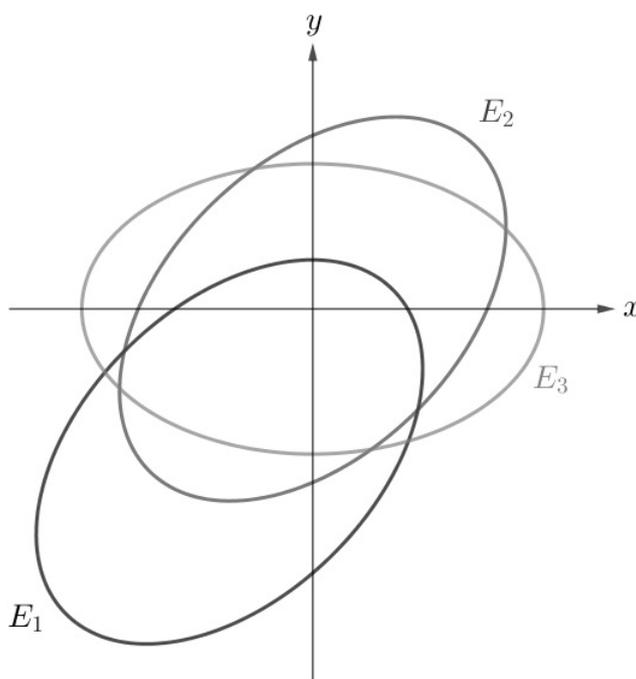


Figura 3.39

Fonte: O autor

A identificação de uma forma quadrática φ , como aponta LIMA [44], pode ser realizada por mera inspeção dos coeficientes A , B e C , mais especificamente pelo determinante de sua matriz, já que este está diretamente ligado ao processo de eliminação dos termos lineares e do termo que contém xy e, além disso, $AC - B^2 = \lambda_1 \cdot \lambda_2$. Tal observação nos permite analisar o caso em que $\det(\mathcal{M}) = 0$, pois sendo este o caso, então teríamos $\lambda_1 \cdot \lambda_2 = 0$, logo, uma das raízes da equação característica é igual a 0. Porém, ambas não poderiam ser iguais a 0, uma vez que isso levaria a $\bar{\varphi}(x, y) = 0$ para todo $x, y \in \mathbb{R}$, isto é, após realizar a rotação e a translação da cônica, esta simplesmente sumiria, o que não faz sentido. Dessa forma, supondo $\lambda_1 \neq 0$, teríamos

$$\bar{\varphi}(x, y) = \lambda_1 x^2.$$

Daí, observando o significado de $\bar{\varphi}(x, y) = c \neq 0$, temos que

$$x = \pm \sqrt{\frac{c}{\lambda_1}}$$

o que nos fornece um par de retas paralelas, desde que c e λ_1 possuam o mesmo sinal.

4 Diagonalização de matrizes

4.1 O conceito

Definição 4.1 Uma matriz $\mathbf{A} = [a_{ij}]_{n \times n}$ é dita **diagonalizável** quando existe uma matriz invertível $\mathbf{X} = [x_{ij}]_{n \times n}$ e uma matriz diagonal $\mathbf{D} = [d_{ij}]_{n \times n}$, tal que

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1},$$

ou, equivalentemente

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}.$$

Neste último caso, dizemos que \mathbf{X} **diagonaliza** \mathbf{A} .

Consideremos, por exemplo, a matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$$

Temos que a matriz \mathbf{X} que a diagonaliza \mathbf{A} é

$$\begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$$

e portanto

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

O processo aplicado para determinar a matriz \mathbf{X} que diagonaliza \mathbf{A} surge da ampliação/contração de um vetor. Para entendê-lo, vamos considerar uma matriz $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,

um vetor coluna $\mathbf{X}_1 = \begin{bmatrix} x \\ y \end{bmatrix}$ e um número real λ , tal que $\mathbf{A}\mathbf{X}_1 = \lambda\mathbf{X}_1$, ou seja,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix},$$

onde a equação acima nos fornece o sistema homogêneo

$$\begin{cases} (a - \lambda)x + by = 0 \\ cx + (d - \lambda)y = 0 \end{cases}$$

com o qual já somos familiarizados e sabemos que seu determinante deve ser igual a zero para que possamos ter soluções não-triviais, de modo que

$$\lambda^2 - \text{tr}(\mathbf{A})\lambda + \det(\mathbf{A}) = 0$$

é a *equação característica* de \mathbf{A} . Neste caso, não poderemos assegurar que sempre haverão raízes de tal equação como no caso das formas quadráticas, uma vez que $\Delta = (a - d)^2 + 4bc$, pois b e c não necessariamente precisam ser iguais, muito menos ter o mesmo sinal.

Vejamos como aplicar o processo de diagonalização á matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$$

Primeiro calculamos as raízes de sua equação característica

$$\lambda^2 - 4\lambda - 5 = 0$$

que são $\lambda_1 = -1$ e $\lambda_2 = 5$. Daí, tomando $\lambda_1 = -1$ no sistema

$$\begin{cases} (1 - \lambda)x + 2y = 0 \\ 4x + (3 - \lambda)y = 0 \end{cases}$$

obtemos

$$(I) \quad \begin{cases} 2x + 2y = 0 \\ 4x + 4y = 0 \end{cases}$$

cuja solução é $x = -y$, de modo que tomando $y = 1$, o vetor

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

está associado a $\lambda_1 = -1$. Já se tomarmos $\lambda = \lambda_2$ nesse mesmo sistema, obtemos

$$(II) \quad \begin{cases} -4x + 2y = 0 \\ 4x - 2y = 0 \end{cases}$$

o que nos leva a $y = 2x$. Daí, tomando $x = 1$, segue que o vetor

$$\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

está associado a $\lambda_2 = 5$. Temos assim que a matrix \mathbf{X} que diagonaliza \mathbf{A} é

$$\mathbf{X} = \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix},$$

cuja inversa é dada por

$$\mathbf{X}^{-1} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

e assim

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1},$$

onde

$$\mathbf{D} = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}.$$

Note que se tivéssemos tomado $\lambda_1 = 5$ e $\lambda_2 = -1$, então seriam

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{X}^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

mas ainda assim

$$\mathbf{XDX}^{-1} = \mathbf{A}.$$

Podemos assim perceber que a ordem que tomamos as raízes da equação característica não importa e, além disso, fica claro que a matriz que diagonaliza uma determinada matriz não é única. Se tivéssemos tomado $x = 1$ ao invés de $y = 1$ no sistema (I) e $y = 1$ ao invés de $x = 1$ no sistema (II), a matriz \mathbf{X} seria

$$\mathbf{X} = \begin{bmatrix} -1 & -\frac{1}{2} \\ 1 & 1 \end{bmatrix}$$

Assim, não precisamos nos preocupar com a ordem das raízes e nem mesmo, a qual das variáveis se atribui valor. Ainda assim, a Definição 4.1 abre portas para a seguinte pergunta: “Uma matriz precisa ser invertível para ser diagonalizável?” Não. Tomemos por exemplo a matriz singular

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

Sua forma diagonalizada é

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

enquanto que a matriz $\mathbf{B} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ não é diagonalizável, pois o sistema

$$\begin{cases} -\lambda x + y = 0 \\ -\lambda y = 0 \end{cases}$$

possui equação característica $\lambda^2 = 0$, o que implica $\lambda_1 = \lambda_2 = 0$. Daí, teríamos \mathbf{D} como sendo uma matriz nula e assim, se existisse uma matriz invertível \mathbf{X} que diagonalizasse \mathbf{B} , então seria

$$\mathbf{B} = \mathbf{XDX}^{-1} = \mathbf{O},$$

o que é absurdo, uma vez que \mathbf{B} é não-nula.

Outra questão importante acerca da diagonalização de uma matriz é: “Quais as vantagens de diagonalizar uma matriz?” A resposta para essa pergunta é: Depende do problema que se trata, como por exemplo, de acordo com LEON [40], problemas que envolvem a aplicação de uma transformação linear sucessivas vezes. Mas antes de apresentarmos um tal problema, vamos conhecer uma propriedade interessante (que também é uma vantagem) da diagonalização, a qual é, dada uma matriz \mathbf{A} diagonalizável, então

$$\mathbf{A} = \mathbf{XDX}^{-1} \Rightarrow \mathbf{A}^n = \mathbf{XD}^n\mathbf{X}^{-1},$$

para todo $n \in \mathbb{N}$. De fato, sabemos que para $n = 1$ o caso é verdadeiro por definição. Suponhamos então que seja válido para algum $k \in \mathbb{N}$, isto é,

$$\mathbf{A}^k = \mathbf{XD}^k\mathbf{X}^{-1}.$$

Agora notemos que

$$\mathbf{A}^{k+1} = \mathbf{A}\mathbf{A}^k = (\mathbf{X}\mathbf{D}\mathbf{X}^{-1})(\mathbf{X}\mathbf{D}^k\mathbf{X}^{-1}) = \mathbf{X}\mathbf{D}\mathbf{D}^k\mathbf{X}^{-1} = \mathbf{X}\mathbf{D}^{k+1}\mathbf{X}^{-1},$$

logo, podemos concluir que tal propriedade vale para qualquer que seja $n \in \mathbb{N}$.

Observemos agora um exemplo em que as duas vantagens acima podem ser utilizadas para analisar uma determinada matriz.

Exemplo 4.1. Uma determinada cidade possui 78768 habitantes, onde 21535 destes praticam exercícios, enquanto que 57233 são sedentários. A cada ano, 35% dos praticantes de atividades físicas passam a ser sedentários, ao passo que 25% da população sedentária começa a praticar exercícios. Vamos representar por \mathbf{A} a matriz de alternância entre os dois estilos de vida citados acima, onde

$$\mathbf{A} = \begin{bmatrix} 0.65 & 0.25 \\ 0.35 & 0.75 \end{bmatrix}.$$

Seja $\mathbf{v}_0 = \begin{bmatrix} 21535 \\ 57233 \end{bmatrix}$ o vetor inicial de estilo de vida dessa cidade. Daí, para saber quantas pessoas estarão em um desses dois estilos de vida após um ano, basta fazer

$$\mathbf{v}_1 = \mathbf{A}\mathbf{v}_0 = \begin{bmatrix} 28306 \\ 50462 \end{bmatrix}.$$

No ano seguinte, temos

$$\mathbf{v}_2 = \mathbf{A}\mathbf{v}_1 = \mathbf{A}^2\mathbf{v}_0 \cong \begin{bmatrix} 31014 \\ 47754 \end{bmatrix}.$$

Podemos assim perceber que após n anos, tem-se $\mathbf{v}_n = \mathbf{A}^n\mathbf{v}_0$. Porém, se utilizarmos um software matemático, podemos observar que o vetor de estilos de vida para $n \geq 14$ se estabiliza, ou seja,

$$\mathbf{v}_n = \begin{bmatrix} 32820 \\ 45948 \end{bmatrix}, \quad \forall n \geq 14.$$

Tal processo pode ser realizado sem a necessidade de um software matemático, utilizando para tanto apenas a diagonalização de \mathbf{A} . Ora, aplicando o processo de diagonalização a tal matriz, obtemos

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & \frac{7}{5} \end{bmatrix} \begin{bmatrix} \frac{2}{5} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{7}{12} & -\frac{5}{12} \\ \frac{5}{12} & \frac{5}{12} \end{bmatrix},$$

o que por sua vez, implica

$$\mathbf{A}^n = \begin{bmatrix} 1 & 1 \\ -1 & \frac{7}{5} \end{bmatrix} \begin{bmatrix} \left(\frac{2}{5}\right)^n & 0 \\ 0 & 1^n \end{bmatrix} \begin{bmatrix} \frac{7}{12} & -\frac{5}{12} \\ \frac{5}{12} & \frac{5}{12} \end{bmatrix}.$$

Como $\frac{2}{5}$ está entre 0 e 1, sabemos que quanto maior for n , mais próximo de zero será o valor de $\left(\frac{2}{5}\right)^n$, enquanto que $1^n = 1$ para todo $n \in \mathbb{N}$. Logo, para n grande (nesse caso, $n \geq 14$), podemos escrever

$$\mathbf{A}^n = \begin{bmatrix} 1 & 1 \\ -1 & \frac{7}{5} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{7}{12} & -\frac{5}{12} \\ \frac{5}{12} & \frac{5}{12} \end{bmatrix} = \frac{1}{12} \begin{bmatrix} 5 & 5 \\ 7 & 7 \end{bmatrix}$$

e assim

$$\mathbf{v}_n = \mathbf{A}^n \mathbf{v}_0 = \frac{1}{12} \begin{bmatrix} 5 & 5 \\ 7 & 7 \end{bmatrix} \begin{bmatrix} 21535 \\ 57233 \end{bmatrix} = \begin{bmatrix} 32820 \\ 45948 \end{bmatrix},$$

confirmando assim a matriz encontrada com software matemático. Dessa forma, podemos concluir que, embora a taxa de pessoas que se tornam sedentárias seja maior que a taxa de pessoas que passam a praticar exercícios físicos, o número de sedentários diminui e o de praticantes de atividades físicas aumenta durante um determinado período, até então se estabilizar.

O exemplo acima permite-nos perceber que a diagonalização de uma matriz pode facilitar problemas que, em outras circunstâncias, só poderiam ser resolvidos com auxílio de um software matemático. Obviamente, não está sendo aqui colocada de lado a importância do uso de tais ferramentas, uma vez que, de acordo com, AYUB [5], o uso de softwares dinâmicos pode ser benéfico tanto para o ensino como para a aprendizagem de matemática. Além disso, segundo GLADCHEFF [26], o computador pode ser um grande aliado do desenvolvimento cognitivo dos alunos, permitindo que este possa aprender com seus erros de acordo com o seu ritmo de aprendizagem. Porém, a utilização de recursos tecnológicos nem sempre é possível, pois, como aponta PACHECO [52], o emprego de tais ferramentas depende da disponibilidade das mesmas no ambiente escolar, bem como do preparo didático dos professores para utilizá-las.

Com relação à utilização de ferramentas tecnológicas em sala de aula por meio dos professores, é algo que depende não somente de seu preparo, mas também de sua linha pedagógica de pensamento, como por exemplo, de acordo com o estudo realizado por NIEDERHAUSER [50], professores construtivistas permitem que os alunos explorem ferramentas manipulativas aplicando suas próprias estratégias, enquanto que professores behavioristas permitem o uso de tais ferramentas, porém utilizando passos específicos.

Pedagogias baseadas em orientações teóricas behaviorista e construtivista representam visões dramaticamente diferentes de ensino e aprendizagem e dão origem a concepções fundamentalmente diferentes do uso de computadores na instrução como uma máquina de ensino didático ou como ferramentas construtivas de pensamento e reflexão. (NIEDERHAUSER [50], 2001).

Estando ciente das observações realizadas logo acima, vamos aplicar o processo de diagonalização a uma matriz 3×3 . Sabemos que o processo é o mesmo e que, no caso das matrizes de ordem 2 o processo de determinação da equação característica pode ser simplificado, uma vez que tal equação pode ser facilmente obtida por meio do traço e do determinante da matriz, o que acaba por fornecer uma equação quadrática, cujas raízes são obtidas pelo então chamado método de Bháskara. Antes de iniciarmos, precisaremos de três definições que serão úteis para que no caso das matrizes de ordem 3, o processo seja também facilitado. Vale ressaltar que tais definições não foram encontradas na literatura, porém estas serão baseadas em definições já existentes na matemática, o que por sua vez mantém a validade das mesmas.

Definição 4.2. Seja $\mathbf{A} = [a_{ij}]$ uma matriz quadrada de ordem n . Diz-se que a matriz

$$\mathbf{A}_i^* = \begin{bmatrix} a_{ii} & a_{i(i+1)} & \cdots & a_{in} \\ a_{(i+1)i} & a_{(i+1)(i+1)} & \cdots & a_{(i+1)n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{ni} & a_{n(i+1)} & \cdots & a_{nn} \end{bmatrix},$$

com $i = 1, 2, \dots, n-1$ é o i -ésimo **menor complementar diagonal** da matriz \mathbf{A} .

Exemplo 4.2. Os menores complementares diagonais da matriz

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

são

$$\mathbf{A}_1^* = \begin{bmatrix} 0 & 1 & 3 \\ 5 & 4 & 5 \\ -6 & 7 & 2 \end{bmatrix}, \quad \mathbf{A}_2^* = \begin{bmatrix} 4 & 5 \\ 7 & 2 \end{bmatrix}, \quad \text{e} \quad \mathbf{A}_3^* = [2]$$

Definição 4.3. Seja $\mathbf{A} = [a_{ij}]$ uma matriz quadrada de ordem n . A expressão

$$a_{ij}a_{ji} + a_{i(j+1)}a_{(j+1)i} + \cdots + a_{(n-1)n}a_{n(n-1)},$$

com $i \neq j$ para todo $i, j = 1, 2, \dots, n$, é chamada **sutura** de \mathbf{A} e é representada por $\text{st}(\mathbf{A})$.

O termo “sutura” foi escolhido devido à forma com que é realizada a multiplicação, onde os saltos sobre a diagonal principal são semelhantes ao processo de costura, como podemos observar na imagem abaixo

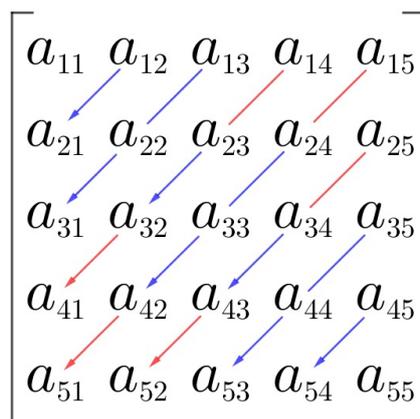


Figura 4.1

Fonte: O autor

Por exemplo, a sutura da matriz \mathbf{A} dada no Exemplo 4.2 é

$$\text{st}(\mathbf{A}) = 7 \cdot (-1) + 4 \cdot 9 + 2 \cdot 1 + 1 \cdot 5 + 3 \cdot (-6) + 5 \cdot 7 = 53.$$

A definição a seguir depende das duas últimas e será utilizada para simplificar a determinação da equação característica para matrizes quadradas de ordem 3.

Definição 4.4. Seja \mathbf{A} uma matriz quadrada de ordem n . A **grade** de \mathbf{A} , representada por $\text{gr}(\mathbf{A})$, é definida como sendo

$$\text{gr}(\mathbf{A}) = a_{11} \cdot \text{tr}(\mathbf{A}_1^*) + \cdots + a_{(n-1)(n-1)} \cdot \text{tr}(\mathbf{A}_{n-1}^*) - \text{st}(\mathbf{A}).$$

Exemplo 4.3. Consideremos a matriz

$$\mathbf{C} = \begin{bmatrix} 11 & 5 & -8 \\ -4 & -2 & 4 \\ 10 & 5 & -7 \end{bmatrix}.$$

Temos que sua sutura é $\text{st}(\mathbf{C}) = -80$, e assim

$$\text{gr}(\mathbf{C}) = 11 \cdot (-9) + (-2) \cdot (-7) - (-80) = -5.$$

Fornecidas as três definições acima, tomemos agora uma matriz

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

e um vetor

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Daí, a equação $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, onde $\lambda \in \mathbb{R}$, nos fornece o sistema

$$\begin{cases} (a - \lambda)x + by + cz = 0 \\ dx + (e - \lambda)y + fz = 0 \\ gx + hy + (i - \lambda)z = 0 \end{cases}$$

onde o determinante de sua matriz é dado por

$$\begin{aligned} & (a - \lambda) \begin{vmatrix} e - \lambda & f \\ h & i - \lambda \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i - \lambda \end{vmatrix} + c \begin{vmatrix} d & e - \lambda \\ g & h \end{vmatrix} \\ &= -\lambda^3 + (a + e + i)\lambda^2 - (ae + ai + ei - bd - cg - fh)\lambda + \det(\mathbf{A}) \\ &= -\lambda^3 + \text{tr}(\mathbf{A})\lambda^2 - \text{gr}(\mathbf{A})\lambda + \det(\mathbf{A}). \end{aligned}$$

Logo, a equação característica de uma matriz 3×3 é

$$-\lambda^3 + \text{tr}(\mathbf{A})\lambda^2 - \text{gr}(\mathbf{A})\lambda + \det(\mathbf{A}) = 0.$$

Exemplo 4.4. Vamos diagonalizar a matriz do Exemplo 4.3. Como já sabemos, $\text{gr}(\mathbf{C}) = -5$ e, além disso, podemos obter facilmente $\text{tr}(\mathbf{C}) = 2$ e $\det(\mathbf{C}) = -6$. Daí, decorre que a equação característica de \mathbf{C} é

$$-\lambda^3 + 2\lambda^2 + 5\lambda - 6 = 0,$$

cujas raízes são $\lambda_1 = 1$, $\lambda_2 = -2$ e $\lambda_3 = 3$. Dessa forma, para $\lambda = \lambda_1$, obtemos o sistema

$$\begin{cases} 10x + 5y - 8z = 0 \\ -4x - 3y + 4z = 0, \\ 10x + 5y - 8z = 0 \end{cases}$$

cuja solução é $x = \frac{2}{5}z, y = \frac{4}{5}z, z$. Logo, tomando $z = 5$, obtemos

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}.$$

Para $\lambda = \lambda_2$, o sistema obtido é

$$\begin{cases} 13x + 5y - 8z = 0 \\ -4x + 0y + 4z = 0, \\ 10x + 5y - 5z = 0 \end{cases}$$

que ao ser solucionado, nos fornece $x = z, y = -z, z$. Assim, tomando $z = 1$, segue que

$$\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Finalmente, para $\lambda = \lambda_3$, temos

$$\begin{cases} 8x + 5y - 8z = 0 \\ -4x - 5y + 4z = 0, \\ 10x + 5y - 10z = 0 \end{cases}$$

cuja solução é $x = z, y = 0, z$. Daí, para $z = 1$, decorre que

$$\mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

e portanto

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

Para o caso $n = 4$, veremos que a grade da matriz também aparece na equação característica (sendo o coeficiente do termo de grau 2), porém novos termos aparecem também. Dessa forma, não se deve esperar uma fórmula para determinar qualquer polinômio característico, a menos da expansão dos determinantes por meio dos cofatores, o que pode ser laborioso, para matrizes de ordem superior.

Consideremos então

$$\mathbf{A} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix} \quad \text{e} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

Desenvolvendo $\det(\mathbf{A} - \mathbf{I}x)$, teremos

- O coeficiente do termo de grau 4 é 1.
- O coeficiente do termo de grau 3 é igual a $-\text{tr}(\mathbf{A})$.

- O coeficiente do termo de grau 2 é $\text{gr}(\mathbf{A})$.
- O coeficiente do termo de grau 1 é

$$-(\det(\mathbf{A}_{11}) + \det(\mathbf{A}_{22}) + \det(\mathbf{A}_{33}) + \det(\mathbf{A}_{44})).$$

- O coeficiente do termo de grau 0 é $\det(\mathbf{A})$.

Podemos assim observar que, a partir de $n = 4$, aparece um novo termo, cujo cálculo pode ser ainda mais longo que o de $\text{gr}(\mathbf{A})$, a menos que a diagonal principal de \mathbf{A} contenha alguma(s) entrada(s) nula(s). Todavia, um padrão padrão polinomial pode ser observado na determinação da equação característica de uma matriz \mathbf{A} de ordem n , o qual é

$$(-1)^n(\lambda^n - \text{tr}(\mathbf{A})\lambda^{n-1} + \text{gr}(\mathbf{A})\lambda^{n-2} + \dots + (-1)^n \det(\mathbf{A})) = 0.$$

Dessa forma, podemos extrair uma propriedade bastante interessante acerca das raízes de tal equação, uma vez que, pelas relações de Viète, a soma e o produto das raízes do polinômio

$$p(\lambda) = (-1)^n(\lambda^n - \text{tr}(\mathbf{A})\lambda^{n-1} + \text{gr}(\mathbf{A})\lambda^{n-2} + \dots + (-1)^n \det(\mathbf{A}))$$

são, respectivamente $\text{tr}(\mathbf{A})$ e $\det(\mathbf{A})$. Isto porque,

$$\lambda_1 + \dots + \lambda_n = (-1)^n \frac{\text{tr}(\mathbf{A})}{(-1)^n} = \text{tr}(\mathbf{A}).$$

e

$$\lambda_1 \dots \lambda_n = \frac{(-1)^n \det(\mathbf{A})}{(-1)^n} = \det(\mathbf{A}).$$

Observe, por exemplo, que na matriz \mathbf{C} do Exemplo 4.4

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= 1 + (-2) + 3 = 2 = \text{tr}(\mathbf{A}) \\ \lambda_1 \cdot \lambda_2 \cdot \lambda_3 &= 1 \cdot (-2) \cdot 3 = -6 = \det(\mathbf{A}) \end{aligned}$$

Um comentário adicional, o qual será importante na compreensão do Teorema Espectral (Apêndice) é que, as raízes do polinômio característico de uma matriz triangular estão dispostos em sua diagonal principal. De fato, seja $\mathbf{T} = [t_{ij}]_{n \times n}$ uma matriz triangular. Como sabemos, o determinante de uma matriz triangular é igual ao produto das entradas de sua diagonal principal. Como \mathbf{T} é triangular, então a matriz $\mathbf{T} - \lambda\mathbf{I}$ é também triangular, onde seu determinante coincide com o polinômio característico de \mathbf{T} , isto é,

$$0 = \det(\mathbf{T} - \lambda\mathbf{I}) = (t_{11} - \lambda_1) \dots (t_{nn} - \lambda_n),$$

o que nos fornece $\lambda_1 = t_{11}, \dots, \lambda_n = t_{nn}$.

Tendo em vista tudo o que foi apresentado nesta seção, uma pergunta pode ser levantada: “*Existe alguma condição necessária e suficiente para que uma matriz seja diagonalizável?*” Felizmente, a resposta é sim e a prova disso se encontra no

Teorema 4.1. *Seja \mathbf{A} uma matriz quadrada de ordem n . Então \mathbf{A} é diagonalizável se, e somente se, os vetores associados às raízes do polinômio característico de \mathbf{A} forem linearmente independentes.*

Demonstração. Suponhamos que \mathbf{A} seja diagonalizável. Existem então uma matriz \mathbf{X} invertível e uma matriz diagonal \mathbf{D} , tais que $\mathbf{AX} = \mathbf{XD}$. Esta última igualdade pode ser escrita como

$$[\mathbf{Ax}_1 \ \cdots \ \mathbf{Ax}_n] = [\lambda_1 \mathbf{x}_1 \ \cdots \ \lambda_n \mathbf{x}_n],$$

onde $\mathbf{x}_1, \dots, \mathbf{x}_n$ são os vetores colunas de \mathbf{X} associados às raízes $\lambda_1, \dots, \lambda_n$ nesta mesma ordem. Como \mathbf{X} é invertível, segue do Teorema 3.3 que os vetores coluna $\mathbf{x}_1, \dots, \mathbf{x}_n$ são linearmente independentes. Reciprocamente, se os vetores associados às raízes do polinômio característico de \mathbf{A} são linearmente independentes, então as igualdades $\mathbf{Ax}_1 = \lambda_1 \mathbf{x}_1, \dots, \mathbf{Ax}_n = \lambda_n \mathbf{x}_n$ implicam $\mathbf{AX} = \mathbf{XD}$. Pela independência linear dos vetores coluna de \mathbf{X} , decorre pelo Teorema 3.3 que \mathbf{X} é invertível, o que nos permite escrever $\mathbf{A} = \mathbf{XDX}^{-1}$, o que significa que \mathbf{A} é diagonalizável. \square

4.2 A diagonalização de matrizes na rotação de cônicas

Antes de iniciarmos aplicarmos o processo de diagonalização para realizar translações e/ou rotações de cônicas, necessitaremos do conceito de *matriz ortogonal*, o qual é definido a seguir.

Definição 4.5. Uma matriz \mathbf{A} quadrada de ordem n é dita **ortogonal** quando

$$\mathbf{A}^T = \mathbf{A}^{-1},$$

o que equivale a

$$\mathbf{A}^T \mathbf{A} = \mathbf{AA}^T = \mathbf{I}.$$

Em outras palavras, a definição acima nos diz que uma matriz é ortogonal se, e somente se, esta for igual a sua inversa. A matriz

$$\begin{bmatrix} \frac{\sqrt{5}}{5} & -\frac{2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix}$$

é um exemplo de matriz ortogonal, pois

$$\begin{bmatrix} \frac{\sqrt{5}}{5} & -\frac{2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ -\frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Definição 4.6. Diz-se que uma matriz quadrada \mathbf{A} é **ortogonalmente diagonalizável** se existe uma matriz ortogonal \mathbf{P} e uma matriz diagonal \mathbf{D} , tal que

$$\mathbf{A} = \mathbf{PDP}^T.$$

Nesse caso, diz-se que \mathbf{P} **diagonaliza ortogonalmente** a matriz \mathbf{A} . A definição acima é de grande utilidade para o processo de rotação, uma vez que, permite que provemos um resultado que associa matrizes ortogonais e matrizes simétricas, como por exemplo a

matriz \mathcal{M} de uma forma quadrática φ . Antes disso, observemos abaixo um exemplo de matriz ortogonalmente diagonalizável

$$\begin{bmatrix} -1 & 2 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{5}}{5} & -\frac{2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ -\frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix}.$$

Teorema 4.2. *Uma matriz \mathbf{A} é simétrica e diagonalizável se, e somente se, é ortogonalmente diagonalizável.*

Demonstração. Se \mathbf{A} é diagonalizável, então, temos que, existe uma matriz \mathbf{X} e uma matriz diagonal \mathbf{D} , tal que

$$\mathbf{A} = \mathbf{XDX}^{-1}.$$

Agora, notemos que, sendo \mathbf{A} simétrica, tem-se

$$\mathbf{A} = \mathbf{A}^T = (\mathbf{XDX}^{-1})^T = (\mathbf{X}^{-1})^T \mathbf{DX}^T.$$

Porém, como sabemos, a forma de diagonalizar a matriz \mathbf{A} não é única, de modo que

$$\mathbf{A} = (\mathbf{X}^{-1})^T \mathbf{DX}^T$$

é também uma forma diagonalizada de \mathbf{A} . Daí, \mathbf{A} será ortogonalmente diagonalizada se, e somente se, $(\mathbf{X}^{-1})^T = \mathbf{X}$. Multiplicando ambos os membros desta última igualdade por \mathbf{X}^{-1} , obtemos

$$\begin{aligned} \mathbf{X}^{-1}(\mathbf{X}^{-1})^T &= \mathbf{X}^{-1}\mathbf{X} \\ \Leftrightarrow \mathbf{X}^{-1}(\mathbf{X}^T)^{-1} &= \mathbf{I} \\ \Leftrightarrow (\mathbf{X}^T\mathbf{X})^{-1} &= \mathbf{I} \\ \Leftrightarrow \mathbf{X}^T\mathbf{X} &= \mathbf{I}, \end{aligned}$$

ou seja, \mathbf{X} é ortogonal e assim $\mathbf{A} = \mathbf{XDX}^T$. Reciprocamente, se \mathbf{A} é diagonalmente ortogonalizável, então, existe \mathbf{P} e \mathbf{D} diagonal, tais que $\mathbf{A} = \mathbf{XDX}^T$ e assim

$$\mathbf{A}^T = (\mathbf{XDX}^T)^T = (\mathbf{X}^T)^T \mathbf{DX}^T = \mathbf{XDX}^T = \mathbf{A},$$

isto é, \mathbf{A} é simétrica. □

Tendo conhecimento das definições e resultados acima, consideremos agora uma função quadrática $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$\varphi(x, y) = Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F.$$

Note que a mesma pode ser escrita sob a forma matricial

$$\varphi(x, y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} D & E \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [F].$$

Daí, fazendo $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\mathcal{N} = \begin{bmatrix} D & E \end{bmatrix}$ e $\mathcal{P} = [F]$, a igualdade acima se resume a

$$\varphi(x, y) = \mathbf{x}^T \mathcal{M} \mathbf{x} + \mathcal{N} \mathbf{x} + \mathcal{P}. \tag{4.1}$$

Agora vamos supor que \mathcal{M} é diagonalizável. Sendo esta simétrica, segue do Teorema 4.1 que a mesma é também ortogonalmente diagonalizável. Dessa forma, por definição, existe \mathbf{P} ortogonal e \mathbf{D} diagonal, tais que

$$\mathcal{M} = \mathbf{PDP}^T,$$

de modo que

$$\mathbf{x}^T \mathcal{M} \mathbf{x} = \mathbf{x}^T \mathbf{PDP}^T \mathbf{x} = (\mathbf{P}^T \mathbf{x})^T \mathbf{D} (\mathbf{P}^T \mathbf{x}).$$

Tomando $\mathbf{y} = \mathbf{P}^T \mathbf{x}$, podemos multiplicar (à esquerda) ambos os membros desta última por \mathbf{P}^T e, sendo \mathbf{P} ortogonal, temos que $\mathbf{P}^T = \mathbf{P}^{-1}$ e assim obtemos $\mathbf{x} = \mathbf{P} \mathbf{y}$. Em seguida, antes de reescrevermos (4.1), observemos que

•

$$\mathbf{y} = \mathbf{P}^T \mathbf{x} = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix},$$

onde $\bar{x} = p_{11}x + p_{21}y$ e $\bar{y} = p_{12}x + p_{22}y$.

•

$$\mathbf{x}^T \mathcal{M} \mathbf{x} = \mathbf{y}^T \mathbf{D} \mathbf{y} = [\lambda_1 \bar{x}^2 + \lambda_2 \bar{y}^2],$$

com λ_1 e λ_2 sendo as respectivas entradas não-nulas de \mathbf{D} , ou seja, as raízes da equação característica de \mathcal{M} .

•

$$\mathcal{N} \mathbf{x} = \mathcal{N} \mathbf{P} \mathbf{y} = [D' \bar{x} + E' \bar{y}],$$

onde

$$\begin{aligned} D' &= Dp_{11} + Ep_{21} \\ E' &= Dp_{12} + Ep_{22} \end{aligned}$$

Agora, utilizaremos um abuso de notação, omitindo os colchetes e escrevendo F ao invés de $\mathcal{P} = [F]$, de modo que a forma quádrlica em (4.1) passa a ser

$$\bar{\varphi}(\bar{x}, \bar{y}) = \lambda_1 \bar{x}^2 + \lambda_2 \bar{y}^2 + D' \bar{x} + E' \bar{y} + F. \quad (4.2)$$

Observe que o termo contendo xy desapareceu, o que nos fornece a forma quádrlica rotacionada, contendo seus eixos de simetria paralelos aos eixos x e y , respectivamente. Resta então determinar o vetor de translação, o qual será determinado pelo método de completar quadrados, onde este, sendo aplicado à igualdade acima, nos fornece

$$\bar{\varphi}(\bar{x}, \bar{y}) = \lambda_1 \left(\bar{x} + \frac{D'}{2\lambda_1} \right)^2 + \lambda_2 \left(\bar{y} + \frac{E'}{2\lambda_2} \right)^2 - F'. \quad (4.3)$$

onde $F' = -F + \frac{(D')^2}{4\lambda_1} + \frac{(E')^2}{4\lambda_2}$. Note que os termos lineares também “desaparecem” após a completção de quadrados. Dessa forma, a mudança de coordenadas realizada nos três pontos acima elimina o termo que contém xy , mas os termos lineares permanecem, porém de maneira implícita. Daí, para $\bar{\varphi}(\bar{x}, \bar{y}) = 0$, decorre que

$$\frac{\left(\bar{x} + \frac{D'}{2\lambda_1} \right)^2}{\left(\sqrt{\frac{F'}{\lambda_1}} \right)^2} + \frac{\left(\bar{y} + \frac{E'}{2\lambda_2} \right)^2}{\left(\sqrt{\frac{F'}{\lambda_2}} \right)^2} = 1, \quad (4.4)$$

onde o vetor de translação é

$$\mathbf{v} = \left(\frac{D'}{2\lambda_1}, \frac{E'}{2\lambda_2} \right).$$

Dessa forma, para eliminar os termos lineares, basta aplicarmos uma translação de vetor $-\mathbf{v}$ a $\bar{\varphi}$, obtendo assim $\overline{\bar{\varphi}}$, cuja equação quando $\overline{\bar{\varphi}}(\bar{x}, \bar{y}) = 0$ é dada por

$$\frac{\bar{x}^2}{\left(\sqrt{\frac{F'}{\lambda_1}}\right)^2} + \frac{\bar{y}^2}{\left(\sqrt{\frac{F'}{\lambda_2}}\right)^2} = 1. \quad (4.5)$$

Exemplo 4.5. Vamos identificar a forma quádrlica $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, com $\varphi(x, y) = 7x^2 - 6xy + 7y^2 + 2x + 10y - 4$, quando $\varphi(x, y) = 0$. Inicialmente, vamos diagonalizar a matriz \mathcal{M} de φ . Já sabemos que $\lambda_1 = 4$ e $\lambda_2 = 10$. Daí, para $\lambda = \lambda_1$, obtemos o sistema

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

cuja solução é $x = y$ e, tomando $y = 1$, obtemos $x = 1$, de modo que o vetor $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ está associado a λ_1 . Já para $\lambda = \lambda_2$, o sistema obtido é

$$\begin{cases} -3x - 3y = 0 \\ -3x - 3y = 0 \end{cases},$$

com solução $x = -y$. Assim, pondo $y = 1$, segue que $x = -1$ e assim $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ está associado a λ_2 . Obtemos assim a matriz

$$\mathbf{X} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Temos assim que

$$\mathcal{M} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Porém, notemos que $\mathbf{X}^{-1} \neq \mathbf{X}^T$, ou seja, \mathbf{X} não é ortogonal. Para resolver isso, basta normalizarmos os vetores \mathbf{x}_1 e \mathbf{x}_2 , isto é, tomarmos os vetores

$$\mathbf{u} = \frac{1}{\|\mathbf{x}_1\|} \mathbf{x}_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$\mathbf{v} = \frac{1}{\|\mathbf{x}_2\|} \mathbf{x}_2 = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

Obtemos assim uma matriz ortogonal

$$\mathbf{P} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix},$$

tal que

$$\mathcal{M} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \mathbf{PDP}^T.$$

Assim, decorre que

$$\mathbf{x}^T \mathcal{M} \mathbf{x} = 4\bar{x}^2 + 10\bar{y}^2.$$

Agora, resta determinar D' , E' e F' , o que já é conhecido como fazer. Logo

$$\begin{aligned} D' &= 2 \cdot \frac{\sqrt{2}}{2} + 10 \cdot \frac{\sqrt{2}}{2} = 6\sqrt{2} \\ E' &= 2 \cdot \left(-\frac{\sqrt{2}}{2}\right) + 10 \cdot \frac{\sqrt{2}}{2} = 4\sqrt{2} \\ F' &= -(-4) + \frac{(6\sqrt{2})^2}{4 \cdot 4} + \frac{(4\sqrt{2})^2}{4 \cdot 10} = \frac{93}{10}, \end{aligned}$$

o que finalmente nos permite escrever

$$\frac{\left(\bar{x} + \frac{3\sqrt{2}}{4}\right)^2}{\left(\sqrt{\frac{93}{40}}\right)^2} + \frac{\left(\bar{y} + \frac{\sqrt{2}}{5}\right)^2}{\left(\sqrt{\frac{93}{100}}\right)^2} = 1,$$

onde a equação acima caracteriza uma elipse de centro $\left(-\frac{3\sqrt{2}}{4}, -\frac{\sqrt{2}}{5}\right)$, a qual chamaremos de \bar{E} . Para escrevê-la de uma maneira mais simples, basta aplicarmos uma translação a \bar{E} pelo vetor $\left(\frac{3\sqrt{2}}{4}, \frac{\sqrt{2}}{5}\right)$, de modo a obtermos a elipse $\bar{\bar{E}}$, com equação

$$\frac{\bar{\bar{x}}^2}{\left(\sqrt{\frac{93}{40}}\right)^2} + \frac{\bar{\bar{y}}^2}{\left(\sqrt{\frac{93}{100}}\right)^2} = 1.$$

Logo, podemos concluir que a equação $7x^2 - 6xy + 7y^2 + 2x + 10y - 4 = 0$ representa uma elipse, a qual já havíamos obtido no Exemplo 3.12 e a chamamos de E .

Quanto aos focos de E , estes podem ser encontrados de duas maneiras:

- A primeira é partindo dos focos $\bar{\bar{F}}_1$ e $\bar{\bar{F}}_2$ de $\bar{\bar{E}}$, os quais já conhecemos do Exemplo 3.12. Aplicamos então uma translação de vetor $\mathbf{v} = \left(-\frac{3\sqrt{2}}{4}, -\frac{\sqrt{2}}{5}\right)$, e assim

$$\begin{aligned} \bar{F}_1 &= \bar{\bar{F}}_1 + \mathbf{v} = \begin{bmatrix} \frac{-3(5\sqrt{2} + \sqrt{62})}{20} \\ -\frac{\sqrt{2}}{5} \end{bmatrix} \\ \bar{F}_2 &= \bar{\bar{F}}_2 + \mathbf{v} = \begin{bmatrix} \frac{3(-5\sqrt{2} + \sqrt{62})}{20} \\ -\frac{\sqrt{2}}{5} \end{bmatrix} \end{aligned}$$

Em seguida, aplicamos uma rotação de 45° no sentido anti-horário a \bar{F}_1 e \bar{F}_2 , de modo que

$$F_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{-3(5\sqrt{2}+\sqrt{62})}{20} \\ -\frac{\sqrt{2}}{5} \end{bmatrix} = \begin{bmatrix} \frac{-3\sqrt{31}-11}{20} \\ \frac{-3\sqrt{31}-19}{20} \end{bmatrix}$$

$$F_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{3(-5\sqrt{2}+\sqrt{62})}{20} \\ -\frac{\sqrt{2}}{5} \end{bmatrix} = \begin{bmatrix} \frac{3\sqrt{31}-11}{20} \\ \frac{3\sqrt{31}-19}{20} \end{bmatrix}$$

resultado este que já havia sido encontrado no Exemplo 3.12. Se considerarmos $\sqrt{31} \cong 5,56776$, obtemos

$$F_1 = (-1,38516, -1,78516) \quad F_2 = (0,285165, -0,114835)$$

- A segunda maneira é, partindo do centro de \bar{E} , o qual já sabemos que é $\left(-\frac{3\sqrt{2}}{4}, -\frac{\sqrt{2}}{5}\right)$. Como é sabido, se uma elipse possui centro (h, k) , então seus focos são $(h \pm c, k)$, onde $c = \sqrt{a^2 - b^2}$, $a > b$. Como $c = \frac{3\sqrt{62}}{20}$, decorre que

$$\bar{F}_1 = \left(-\frac{3\sqrt{2}}{4} - \frac{3\sqrt{62}}{20}, -\frac{\sqrt{2}}{5}\right) = \left(\frac{-3(5\sqrt{2}+\sqrt{62})}{20}, -\frac{\sqrt{2}}{5}\right)$$

$$\bar{F}_2 = \left(-\frac{3\sqrt{2}}{4} + \frac{3\sqrt{62}}{20}, -\frac{\sqrt{2}}{5}\right) = \left(\frac{3(-5\sqrt{2}+\sqrt{62})}{20}, -\frac{\sqrt{2}}{5}\right),$$

Daí, aplicando uma rotação de 45° no sentido anti-horário, obtemos F_1 e F_2 , como já foi visto acima.

Vale ressaltar que a elipse \bar{E} obtida no Exemplo 4.5 não é aquela obtida no Exemplo 3.12, uma vez que, neste último, a translação é realizada primeiro e em seguida é realizada a rotação. Já no Exemplo 4.5, a ordem é contrária, ou seja, primeiro é realizada a rotação e, por fim, a translação (Figura 4.2). Tais exemplos mostram por si que, independente da ordem que sejam aplicadas as transformações lineares (rotação e translação), o resultado obtido é o mesmo.

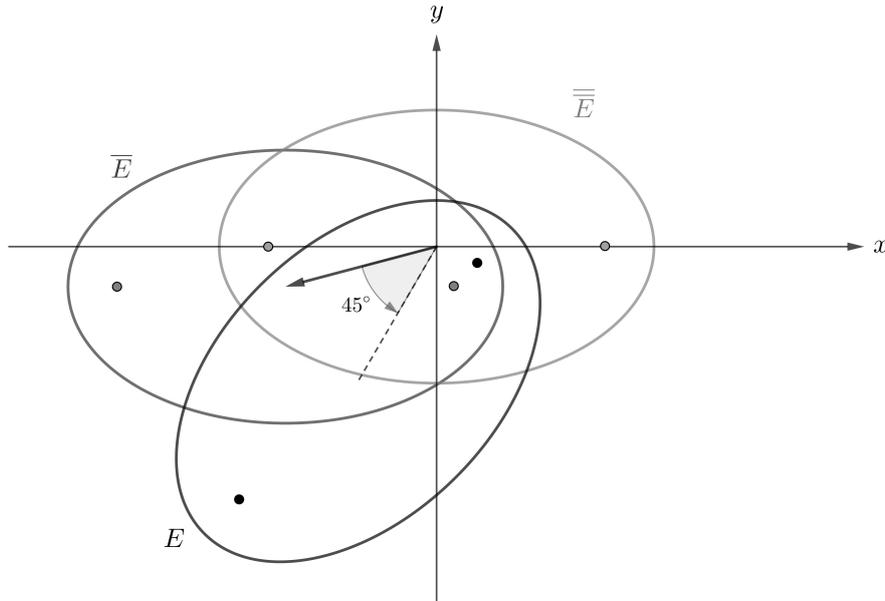


Figura 4.2

Fonte: O autor

Antes de finalizarmos esta seção, é necessário lembrar que, quando um dos autovalores é igual a 0, não se pode utilizar a forma (4.4), muito menos (4.5). Em casos assim, como adverte CAMPOLINO [12], é possível que a forma quadrática em questão seja uma parábola. Em casos assim, utilizaremos (4.2), a qual se resume, tomando $\lambda_2 = 0$ (por exemplo), a

$$\bar{\varphi}(\bar{x}, \bar{y}) = \lambda_1 \bar{x}^2 + D' \bar{x} + E' \bar{y} + F.$$

Para entender o que fazer em situações como esta, vamos utilizar a técnica de diagonalização para identificar a forma quadrática φ , tal que

$$\varphi(x, y) = \frac{1}{2}x^2 - xy + \frac{1}{2}y^2 - \frac{11+2\sqrt{2}}{4}x + \frac{11-2\sqrt{2}}{4}y + \frac{121+36\sqrt{2}}{32},$$

a qual já sabemos, pelo Exemplo 3.10, que quando $\varphi(x, y) = 0$ temos a parábola G_{TRT} . Mas utilizaremos a diagonalização ortogonal para verificar tal fato. Temos então que a matriz de φ é

$$\mathcal{M} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

cuja forma ortogonalmente diagonalizada é

$$\mathcal{M} = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \mathbf{PDP}^T,$$

com $\lambda_1 = 1$ e $\lambda_2 = 0$. Daí, vem que

$$\mathbf{x}^T \mathcal{M} \mathbf{x} = [\bar{x}^2]$$

e, além disso

$$D' = \frac{11\sqrt{2}}{4} \quad \text{e} \quad E' = -1,$$

o que por sua vez implica

$$\mathcal{N}_{\mathbf{x}} = \left[\frac{11\sqrt{2}}{4}\bar{x} - \bar{y} \right].$$

Temos assim

$$\bar{\varphi}(\bar{x}, \bar{y}) = \bar{x}^2 + \frac{11\sqrt{2}}{4}\bar{x} - \bar{y} + \frac{121+36\sqrt{2}}{32}.$$

Como podemos observar, não é possível completar quadrados para \bar{y} , uma vez que a forma quadrática acima não possui termo com \bar{y}^2 . Porém, notemos que $\bar{\varphi}(\bar{x}, \bar{y}) = 0$ equivale a

$$\bar{y} = \bar{x}^2 + \frac{11\sqrt{2}}{4}\bar{x} + \frac{121+36\sqrt{2}}{32},$$

equação esta que, como já se sabe, é de segundo grau e representa uma parábola (a qual chamaremos de \bar{P}), assim como já era de se esperar, visto que a expressão de φ já havia sido obtida no Exemplo 3.10. Completando quadrados e reescrevendo a equação acima em sua forma padrão, obtemos

$$\left(x - \left(-\frac{11\sqrt{2}}{8} \right) \right)^2 = 4 \cdot \frac{1}{4} \left(y - \frac{9\sqrt{2}}{8} \right),$$

o que nos permite concluir que \bar{P} é uma parábola cujo vértice é

$$\left(-\frac{11\sqrt{2}}{8}, \frac{9\sqrt{2}}{8} \right).$$

Daí, fazendo uma translação de vetor $\mathbf{v} = \left(\frac{11\sqrt{2}}{8}, -\frac{9\sqrt{2}}{8} \right)$, obtemos $\bar{\bar{\varphi}}$, tal que

$$\bar{\bar{y}} = \bar{\bar{x}}^2,$$

ou seja, uma parábola com vértice na origem, que será chamada de $\bar{\bar{P}}$. Temos então comprovado que $\varphi(x, y) = 0$ e G_{TRT} são ambas a mesma parábola.

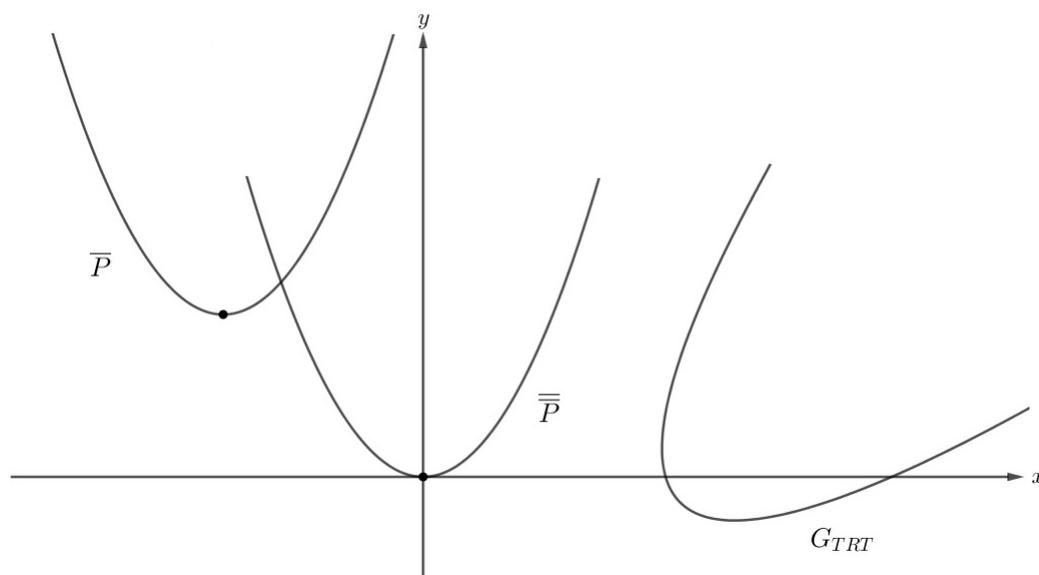


Figura 4.3

Fonte: O autor

Note que, se quisermos determinar o foco de G_{TRT} com as informações obtidas nesse exemplo, precisaremos seguir mais alguns passos. Com efeito, como visto no Exemplo 3.10, sabemos que φ é resultado de translações e rotações sobre a parábola de equação $y = x^2 - 5x + 6$. Então, cabe a pergunta: “É possível reverter o processo?”. A resposta é “sim”. Para tanto, observemos que a diretriz de \bar{P} é

$$\bar{y} = \frac{9\sqrt{2}-2}{8},$$

enquanto que a reta diretriz de $\bar{\bar{P}}$ é $\bar{\bar{y}} = \frac{1}{4}$. Precisamos então determinar a parábola, cujo vértice tenha ordenada $-\frac{1}{4}$ e cuja distância à origem seja igual à distância do vértice de \bar{P} à origem. Ora, a distância de \bar{P} à origem é

$$\sqrt{\left(0 + \frac{11\sqrt{2}}{8}\right)^2 + \left(0 - \frac{9\sqrt{2}}{8}\right)^2} = \frac{\sqrt{101}}{4}.$$

Devemos então determinar o ponto $\left(h, -\frac{1}{4}\right)$ que possui distância igual ao valor obtido logo acima. Assim, devemos ter

$$\sqrt{(0-h)^2 + \left(0 + \frac{1}{4}\right)^2} = \frac{\sqrt{101}}{4} \Rightarrow h = \frac{5}{2}.$$

Portanto, a parábola procurada deve possuir vértice $\left(\frac{5}{2}, -\frac{1}{4}\right)$, ou seja, com equação

$$\left(x - \frac{5}{2}\right)^2 = 4 \cdot \frac{1}{4} \left(y + \frac{1}{4}\right),$$

que pode ser reescrita como

$$y = x^2 - 5x + 6,$$

que é, de fato, a equação da parábola G do Exemplo 3.10.

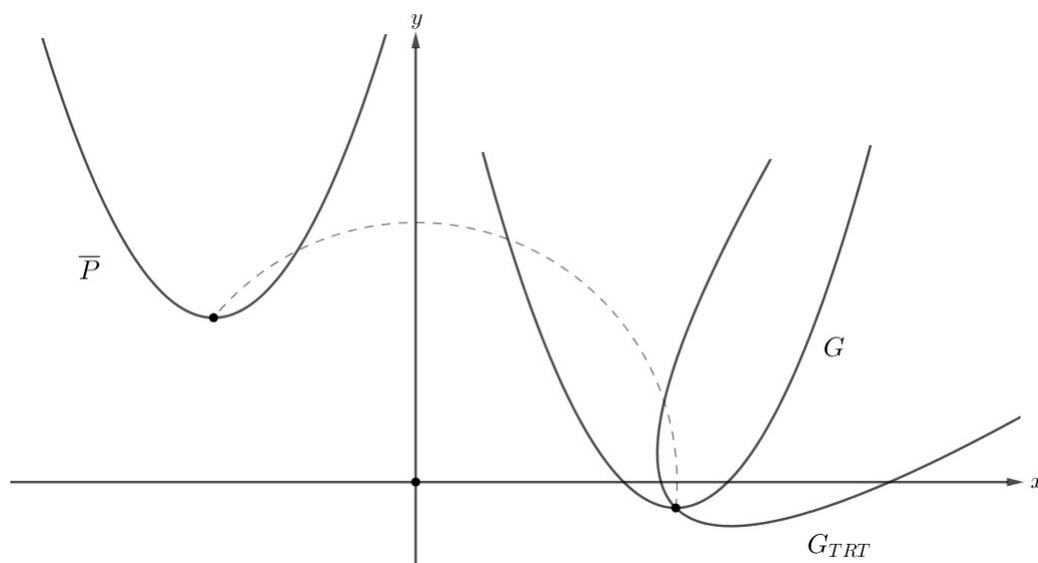


Figura 4.4

Fonte: O autor

Tais informações nos permitem agora determinar:

- O foco de G_{TRT} : Note que o foco de \overline{P} é $\left(0, \frac{1}{4}\right)$. Daí, aplicando rotação de 45° no sentido horário, obtemos

$$P' = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{8} \\ \frac{\sqrt{2}}{8} \end{bmatrix}$$

Em seguida, aplica-se uma translação de vetor $\mathbf{u} = \left(\frac{5}{2}, -\frac{1}{4}\right)$ a P' (note que este é o foco da parábola G_{RT} do Exemplo 3.10) de modo a obtermos o foco P'' de P

$$P'' = \begin{bmatrix} \frac{\sqrt{2}+20}{8} \\ \frac{\sqrt{2}-2}{8} \end{bmatrix}.$$

- A reta diretriz de G_{TRT} : Sabemos que P' pertence à reta $y = x$. Logo, seu simétrico em relação à esta reta, isto é, $\left(-\frac{\sqrt{2}}{8}, -\frac{\sqrt{2}}{8}\right)$, está sob a reta diretriz de G_{RT} . Logo, para determinar a equação de tal reta, basta aplicarmos uma translação de vetor $\left(-\frac{\sqrt{2}}{8}, -\frac{\sqrt{2}}{8}\right)$ à reta $y = -x$ (perpendicular a $y = x$), de modo a obtermos

$$\begin{aligned} y' + \frac{\sqrt{2}}{8} &= -x' - \frac{\sqrt{2}}{8} \\ \Leftrightarrow y' &= -x' - \frac{\sqrt{2}}{4} \end{aligned}$$

Por fim, aplicamos uma translação de vetor $\left(\frac{5}{2}, -\frac{1}{4}\right)$ à reta obtida logo acima, o que por sua vez resulta em

$$y'' = -x'' + \frac{18-2\sqrt{2}}{8},$$

a qual é a reta diretriz de G_{TRT} .

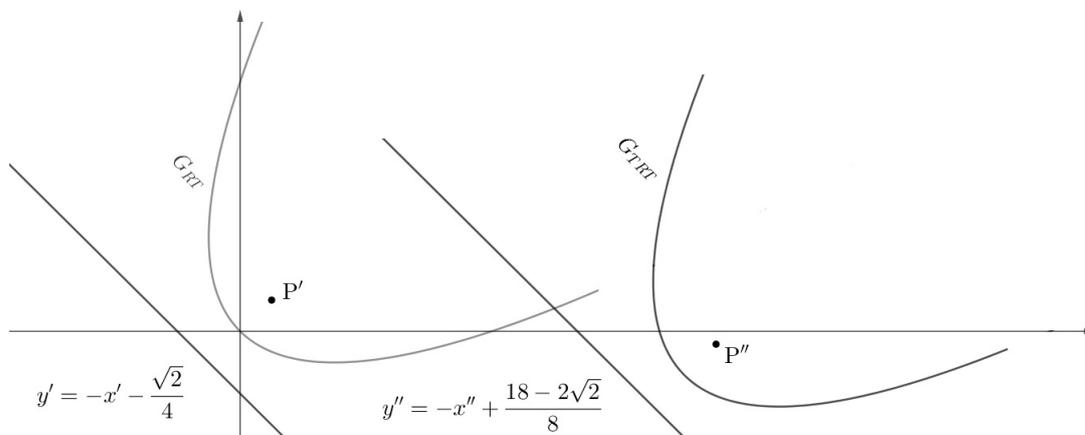


Figura 4.5

Fonte: O autor

Observe que poderíamos ter determinado o vértice de G_{TRT} apenas multiplicando a matriz \mathbf{P} pelo vetor $\left(-\frac{11\sqrt{2}}{8}, \frac{9\sqrt{2}}{8}\right)$, ou seja,

$$\begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} -\frac{11\sqrt{2}}{8} \\ \frac{9\sqrt{2}}{8} \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ -\frac{1}{4} \end{bmatrix}.$$

Finalizamos assim esta seção, tendo como objetivo seguinte, a apresentação de uma aplicação de matrizes, onde estas, assumem a característica de um número, embora não deixem de ser matrizes, pois como já vimos, um número real e uma matriz não são a mesma coisa.

4.3 Exponencial de Matriz

Antes de abordarmos o foco principal desta seção, faz-se necessário definir a noção de *exponencial escalar* de um número real a , isto é, um número da forma

$$e^a = 1 + a + \frac{1}{2!}a^2 + \frac{1}{3!}a^3 + \frac{1}{4!}a^4 + \dots,$$

o qual representa uma potência do famoso *número de Euler*¹⁶. Tal definição é importante, visto que a soma acima é infinita, logo, ao realizarmos uma soma infinita de determinadas matrizes, poderemos obter determinar o resultado da soma infinita das respectivas entradas de cada uma delas. Dito isto, podemos então seguir para a definição chave desta seção.

Definição 4.7. *Seja \mathbf{A} uma matriz quadrada. A exponencial da matriz \mathbf{A} é dada por*

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \frac{1}{3!}\mathbf{A}^3 + \frac{1}{4!}\mathbf{A}^4 + \dots$$

Embora a matriz \mathbf{A} apareça como um expoente, o resultado de $e^{\mathbf{A}}$ não é um número, como podemos notar na definição acima, uma vez que a soma de matrizes é uma matriz. Para melhor compreender, vamos calcular $e^{\mathbf{D}}$, onde \mathbf{D} é uma matriz diagonal de ordem

¹⁶Leonhard Euler, um dos maiores e mais prolíficos matemáticos que o mundo já produziu, nasceu em Basel (Suíça), em 15 de abril de 1707 e morreu em São Ptesburgo (Rússia), em 18 de novembro de 1783 [20].

n .

$$\begin{aligned}
e^{\mathbf{D}} &= \mathbf{I} + \mathbf{D} + \frac{1}{2!}\mathbf{D}^2 + \dots \\
&= \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} + \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix} + \frac{1}{2!} \begin{bmatrix} d_1^2 & 0 & \dots & 0 \\ 0 & d_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^2 \end{bmatrix} + \dots \\
&= \begin{bmatrix} 1 + d_1 + \frac{1}{2!}d_1^2 + \dots & 0 & \dots & 0 \\ 0 & 1 + d_2 + \frac{1}{2!}d_2^2 + \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 + d_n + \frac{1}{2!}d_n^2 + \dots \end{bmatrix} \\
&= \begin{bmatrix} e^{d_1} & 0 & \dots & 0 \\ 0 & e^{d_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{d_n} \end{bmatrix}
\end{aligned}$$

Vale ressaltar que a exponencial de uma matriz não apresenta todas as propriedades da exponencial de um escalar, como por exemplo, não é verdade que $e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{A}+\mathbf{B}}$ para quaisquer matrizes \mathbf{A} e \mathbf{B} . Neste caso, é necessário que $\mathbf{AB} = \mathbf{BA}$, porém não provaremos tal fato, a fi de não fugir do escopo do presente trabalho. Porém, será provado adiante um teorema de grande importância para o cálculo de exponenciais de matrizes diagonalizáveis.

Definição 4.8. Duas matrizes \mathbf{A} e \mathbf{B} são ditas *conjugadas* ou *semelhantes* se existe uma matriz \mathbf{Q} invertível, tal que

$$\mathbf{AQ} = \mathbf{QB},$$

o que é equivalente a

$$\mathbf{A} = \mathbf{QBQ}^{-1}.$$

A definição acima é uma generalização da Definição 4.1, uma vez que a matriz \mathbf{B} não necessariamente precisa ser diagonal. Temos, por exemplo,

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

As matrizes conjugadas apresentam a seguinte propriedade: *Sejam \mathbf{A} e \mathbf{B} duas matrizes conjugadas por uma matriz \mathbf{Q} . Então é verdade que*

$$\mathbf{A}^n \mathbf{Q} = \mathbf{QB}^n,$$

para todo $n \in \mathbb{N}$. De fato, observemos que, dado um número natural qualquer, temos que

$$\mathbf{A}^n \mathbf{Q} = \mathbf{A}^{n-1} \mathbf{AQ} = \mathbf{A}^{n-1} \mathbf{QB}.$$

Repetindo novamente o argumento acima, obtemos

$$\mathbf{A}^{n-1} \mathbf{QB} = \mathbf{A}^{n-2} \mathbf{AQB} = \mathbf{A}^{n-2} \mathbf{QBB} = \mathbf{A}^{n-2} \mathbf{QB}^2.$$

Continuando a aplicar o argumento acima sucessivas vezes, decorre que

$$\mathbf{A}^n \mathbf{Q} = \mathbf{Q} \mathbf{B}^n.$$

Tal propriedade nos permite enunciar o seguinte teorema:

Teorema 4.3. *Se \mathbf{A} e \mathbf{B} são matrizes conjugadas por uma matriz invertível \mathbf{Q} , então as matrizes $e^{\mathbf{A}}\mathbf{Q}$ e $\mathbf{Q}e^{\mathbf{B}}$ são também conjugadas. De forma resumida,*

$$\mathbf{A}\mathbf{Q} = \mathbf{Q}\mathbf{B} \Rightarrow e^{\mathbf{A}}\mathbf{Q} = \mathbf{Q}e^{\mathbf{B}}.$$

Demonstração. Temos que

$$\begin{aligned} e^{\mathbf{A}}\mathbf{Q} &= \left(\mathbf{I} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \dots \right) \mathbf{Q} \\ &= \mathbf{Q} + \mathbf{A}\mathbf{Q} + \frac{1}{2!}\mathbf{A}^2\mathbf{Q} + \dots \\ &= \mathbf{Q} + \mathbf{Q}\mathbf{B} + \frac{1}{2!}\mathbf{Q}\mathbf{B}^2 + \dots \\ &= \mathbf{Q} \left(\mathbf{I} + \mathbf{B} + \frac{1}{2!}\mathbf{B} + \dots \right) \\ &= \mathbf{Q}e^{\mathbf{B}}, \end{aligned}$$

onde a terceira igualdade acima ocorre devido ao Teorema 4.2. □

Note que o Teorema acima é equivalente a

$$\mathbf{A}\mathbf{Q} = \mathbf{Q}\mathbf{B} \Rightarrow e^{\mathbf{A}} = \mathbf{Q}e^{\mathbf{B}}\mathbf{Q}^{-1}.$$

Dessa forma, se \mathbf{A} é diagonalizável com $\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{D}$, decorre que

$$e^{\mathbf{A}} = \mathbf{X}e^{\mathbf{D}}\mathbf{X}^{-1}.$$

Exemplo 4.6. Sabemos do Capítulo 4 que

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

Assim, tem-se que

$$\begin{aligned} e^{\mathbf{A}} &= \mathbf{X}e^{\mathbf{D}}\mathbf{X}^{-1} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} e^{-1} & 0 \\ 0 & e^5 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \\ &\cong \begin{bmatrix} 99,065 & -49,348 \\ 197,762 & -99,065 \end{bmatrix} \end{aligned}$$

Agora que já sabemos como resolver exponenciais de matrizes diagonalizáveis, vamos aplicar tal conhecimento na resolução de *problemas de valor inicial*, isto é, problemas que envolvem uma matriz $\mathbf{Y}'(t)$ da forma

$$\mathbf{Y}'(t) = \mathbf{A}\mathbf{Y}(t), \tag{4.6}$$

onde $\mathbf{Y}(t) = e^{t\mathbf{A}}\mathbf{Y}(0) = e^{t\mathbf{A}}\mathbf{Y}_0$ é uma matriz que depende de $t \in \mathbb{R}$, $\mathbf{Y}(0) = \mathbf{Y}_0$ é um vetor coluna chamado *condição inicial* e \mathbf{A} é uma matriz diagonalizável. Como \mathbf{A} é diagonalizável, então existe uma matriz \mathbf{X} invertível e uma matriz diagonal \mathbf{D} , tais que $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1}$. Assim, podemos escrever $\mathbf{Y}(t)$ da forma

$$\mathbf{Y}(t) = \mathbf{X}e^{t\mathbf{D}}\mathbf{X}^{-1}\mathbf{Y}_0.$$

Agora notemos que $\mathbf{X}^{-1}\mathbf{Y}_0$ é uma matriz coluna, logo podemos escrevê-la como um vetor \mathbf{c} , isto é,

$$\mathbf{X}^{-1}\mathbf{Y}_0 = \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Daí, escrevendo \mathbf{X} como uma matriz de blocos $[\mathbf{X}_1 \cdots \mathbf{X}_n]$, segue que

$$\mathbf{Y}(t) = [\mathbf{X}_1 \cdots \mathbf{X}_n] e^{t\mathbf{D}} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = [c_1 e^{\lambda_1 t} \mathbf{X}_1 + \cdots + c_n e^{\lambda_n t} \mathbf{X}_n].$$

Com abuso de notação, podemos escrever simplesmente

$$\mathbf{Y}(t) = c_1 e^{\lambda_1 t} \mathbf{X}_1 + \cdots + c_n e^{\lambda_n t} \mathbf{X}_n.$$

Agora vamos utilizar um exemplo para entender de fato o que as matrizes $\mathbf{Y}'(t)$ e $\mathbf{Y}(t)$ representam.

Exemplo 4.7. Dois tanques A e B possuem a mesma capacidade de 240 L. No interior do Tanque A foram dissolvidos 80g de sal, formando assim uma salmoura, a qual circula pelos canos que conectam os tanques A e B conforme a figura a seguir

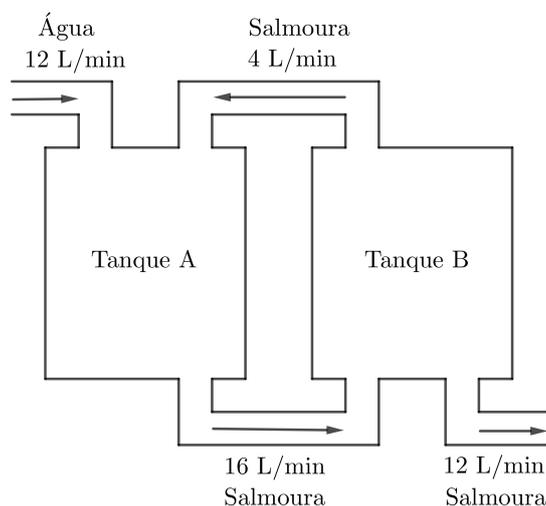


Figura 4.6

Fonte: O autor

Observe que há um fluxo de 15 L/min de água pura entrando no Tanque A, enquanto há um fluxo de saída igual a 15 L/min de salmoura saindo do Tanque B. Além disso, ambos os tanques recebem 16 L de mistura por minuto, o que nos leva a concluir que o volume ocupado por ambos continua sendo de 240 L. Agora, chamando de $y_1(t)$ e $y_2(t)$ as respectivas quantidades de salmoura nos tanques A e B no instante t , temos que

- Taxa de entrada de salmoura no Tanque A

$$(4 \text{ L/min}) \left(\frac{y_2(t)}{240} \text{ g/L} \right) = \frac{y_2(t)}{60} \text{ g/min}$$

- Taxa de saída de salmoura do Tanque A

$$(16 \text{ L/min}) \left(\frac{y_1(t)}{240} \text{ g/L} \right) = \frac{y_1(t)}{15} \text{ g/min}$$

- Taxa de variação de quantidade de salmoura para o Tanque A

$$y_1'(t) = \frac{y_2(t)}{60} - \frac{y_1(t)}{15}$$

- Taxa de variação de quantidade de salmoura para o Tanque B

$$y_2'(t) = \frac{y_1(t)}{15} - \frac{y_2(t)}{15}$$

Obtemos assim

$$\mathbf{Y}'(t) = \begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \begin{bmatrix} -\frac{1}{15} & \frac{1}{60} \\ \frac{1}{15} & -\frac{1}{15} \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \mathbf{A}\mathbf{Y}(t).$$

Agora, sabendo que

$$\mathbf{Y}_0 = \begin{bmatrix} 80 \\ 0 \end{bmatrix},$$

e diagonalizando \mathbf{A} , o que nos fornece

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} -\frac{1}{30} & 0 \\ 0 & -\frac{1}{10} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ -\frac{1}{2} & \frac{1}{4} \end{bmatrix},$$

segue que

$$\mathbf{Y}(t) = c_1 e^{-\frac{1}{30}t} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 e^{-\frac{1}{10}t} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} c_1 e^{-\frac{1}{30}t} - c_2 e^{-\frac{1}{10}t} \\ 2c_1 e^{-\frac{1}{30}t} + 2c_2 e^{-\frac{1}{10}t} \end{bmatrix}.$$

Por fim, para determinar os valores das constantes c_1 e c_2 , basta notar que

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ -\frac{1}{2} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 80 \\ 0 \end{bmatrix} = \begin{bmatrix} 40 \\ -40 \end{bmatrix}$$

e assim

$$\mathbf{Y}(t) = \begin{bmatrix} 40e^{-\frac{1}{30}t} + 40e^{-\frac{1}{10}t} \\ 80e^{-\frac{1}{30}t} - 80e^{-\frac{1}{10}t} \end{bmatrix}.$$

O exemplo acima nos mostra que a matriz $\mathbf{Y}'(t)$ está relacionada com taxas de variação de uma determinada quantidade de salmoura ao longo do tempo, ao passo que a matriz $\mathbf{Y}(t)$ fornece as quantidades de salmoura de cada tanque no momento t . Vejamos abaixo

os gráficos de das quantidades de salmoura nos dois tanques ao longo dos 100 primeiros minutos.

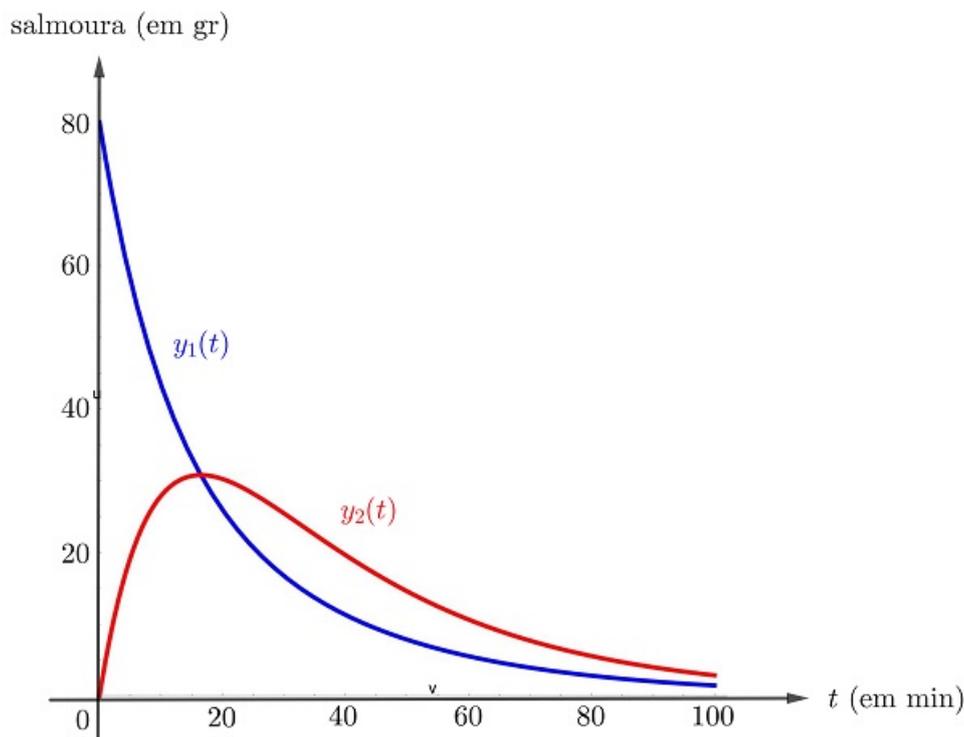


Figura 4.7

Fonte: O autor

Note que em um dado momento t , as quantidades de salmoura em ambos os tanques é exatamente igual. Para determinar um tal ponto, basta resolvermos a equação a seguir

$$\begin{aligned}
 40e^{-\frac{1}{30}t} + 40e^{-\frac{1}{10}t} &= 80e^{-\frac{1}{30}t} - 80e^{-\frac{1}{10}t} \\
 \Leftrightarrow 40e^{-\frac{t}{10}}(e^{\frac{t}{15}} + 1) &= 80e^{-\frac{t}{10}}(e^{\frac{t}{15}} - 1) \\
 \Leftrightarrow e^{\frac{t}{15}} &= 3 \\
 \Leftrightarrow t &= 15 \ln 3
 \end{aligned}$$

Daí, considerando $\ln 3 = ,0986$, decorre que no momento $t \cong 16,48$ min (aproximadamente 16 minutos e 29 segundos), as quantidades de salmoura nos Tanques A e B coincidem.

Ainda com base no Exemplo 4.7, vamos entender o significado das raízes $\lambda_1 = -\frac{1}{30}$ e $\lambda_2 = -\frac{1}{10}$ do polinômio característico da matriz \mathbf{A} . Observemos, para tanto, que o decrescimento percentual da quantidade de salmoura no tanque A após 1 minuto é dado por

$$\frac{y_1(1) - y_1(0)}{y_1(0)} \cong -6,397\%$$

Note que este é também o decrescimento percentual médio no intervalo $[0, 1]$. Para calcular o decrescimento percentual médio no intervalo $[0, 2]$, basta fazermos

$$\frac{1}{2} \left(\frac{y_1(1) - y_1(0)}{y_1(0)} + \frac{y_1(2) - y_1(1)}{y_1(1)} \right) \cong -6,345\%$$

Agora, utilizando a ideia acima e com auxílio de um software matemático, obtemos a seguinte aproximação

$$\frac{1}{1184} \left(\frac{y_1(1) - y_1(0)}{y_1(0)} + \dots + \frac{y_1(1184) - y_1(1183)}{y_1(1183)} \right) \cong -3,334\%,$$

ou seja, o decrescimento percentual médio no intervalo $[0, 1184]$ é aproximadamente $-3,334\%$. Por outro lado, temos que

$$-\frac{1}{30} \cong -3,333\%,$$

o que nos leva a concluir que a raiz $\lambda_1 = -\frac{1}{30}$ nos fornece o decrescimento percentual médio de $y_1(t)$ ao longo do tempo t , quando t assume valores altos. Quanto à quantidade $y_2(t)$, é de se esperar que sua taxa de decrescimento seja 10% , tendo em vista que $\lambda_2 = -\frac{1}{10}$, o que não é verdade, como podemos constatar até mesmo por mera inspeção do gráfico, pois $y_2(t)$ aumenta durante um certo período e em seguida passa a decrescer. Porém, uma aproximação para tal média de decrescimento percentual é obtida de maneira análoga para $y_1(t)$, contudo, a mesma somente é atingida para $t > 5000$. Por exemplo, para $t = 5500$, obtemos

$$\frac{1}{5500} \left(\frac{y_2(2) - y_1(1)}{y_1(1)} + \dots + \frac{y_1(5500) - y_1(5499)}{y_1(5499)} \right) \cong 3,222\%.$$

Note que a diferença percentual entre $3,333\%$ e $3,222\%$ é aproximadamente igual a $3,333\%$.

O Exemplo 4.7 enquadra-se como um problema de *Análise de Redes*¹⁷, por se tratar de uma situação que envolve “fluxo de entrada igual ao fluxo de saída”.

Uma última evidência quanto ao significado das raízes do polinômio característico de uma matriz está presente no exemplo 1.22, onde foi calculado o crescimento populacional anual médio. Utilizando um software matemático, pode-se encontrar uma dessas raízes, a saber $\lambda = 1,84886$, a qual indica o crescimento anual médio de aproximadamente $84,9$, o que está de acordo com a aproximação obtida no referido exemplo.

¹⁷A palavra “rede” pode ter vários significados, uma vez que trata-se da relação entre dados. Dessa forma, a análise de redes possui aplicações em várias áreas, tais como planejamento de projetos, sistemas complexos, circuitos elétricos, redes sociais, sistemas de transporte, redes de comunicação, epidemiologia, bioinformática, sistemas de hipertexto, análise de texto, bibliometria, teoria organizacional, pesquisa genealógica e análise de eventos [11].

5 Sobre números complexos e quatérnios

5.1 Números complexos

O conjunto dos números complexos é definido como sendo

$$\mathbb{C} = \{a + bi; i = \sqrt{-1} \text{ e } a, b \in \mathbb{R}\}.$$

Tais números demoraram bastante tempo para serem estudados, pois, segundo ANDRE-ESCU [2], até o século XVIII, matemáticos evitavam equações quadráticas que não possuíssem raízes reais, sendo $x^2 + 1 = 0$ a mais simples entre elas. A letra i acima, de acordo com DUNHAM [29], foi utilizada pela primeira vez por Euler em seu livro *Elements of Algebra* para representar $\sqrt{-1}$.

Um número complexo $z = a + bi$ qualquer é dividido em duas partes, sendo uma delas a parte real $\Re(z) = a$ e a outra, sua parte imaginária $\Im(z) = b$. Dessa forma, dois números complexos $z_1 = a + bi$ e $z_2 = c + di$ são iguais, se e somente se, $a = \Re(z_1) = \Re(z_2) = c$ e $b = \Im(z_1) = \Im(z_2) = d$.

As operações de soma e multiplicação de números complexos são bem definidas, uma vez que, dados dois números complexos $z_1 = a + bi$ e $z_2 = c + di$, tem-se que a soma e o produto destes são ainda números complexos, como podemos observar logo abaixo

- $z_1 + z_2 = a + bi + c + di = (a + c) + (b + d)i$
- $z_1 \cdot z_2 = (a + bi)(c + di) = (ac - bd) + (ad + bc)i$

Os números complexos apresentam as propriedades a seguir, as quais podem ser facilmente verificadas: Sejam z_1, z_2 e z_3 números complexos, tem-se

- ★ *Comutatividade:* $z_1 + z_2 = z_2 + z_1$
- ★ *Associatividade:* $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$
- ★ *Distributividade:* $z_1(z_2 + z_3) = z_1z_2 + z_1z_3$
- ★ *Elemento neutro (soma):* $z_1 + 0 = z_1$
- ★ *Elemento neutro (multiplicação):* $z_1 \cdot 1 = z_1$
- ★ *Inverso aditivo:* Existe um único elemento $u = -z_1$, tal que $z_1 + u = u + z_1 = 0$, para qualquer que seja $z_1 \in \mathbb{C}$.
- ★ *Inverso multiplicativo:* Existe um único elemento $w = z_1^{-1}$, tal que $z_1w = wz_1 = 1$, para qualquer que seja $z_1 \in \mathbb{C}$.

Tendo em mente o conceito de números complexos, bem como suas principais propriedades, buscaremos agora uma maneira de relacionar números complexos com matrizes. Para tanto, vamos inicialmente considerar a equação

$$\mathbf{X}^2 = -\mathbf{I}, \quad (5.1)$$

onde $\mathbf{X} = \begin{bmatrix} m & n \\ p & q \end{bmatrix}$, com $m, n, p, q \in \mathbb{R}$ e \mathbf{I} é a matriz identidade de ordem 2. Para determinar as entradas da matrix \mathbf{X} , notemos que (5.1) nos fornece as seguintes equações

$$m^2 + np = -1 \quad (5.2)$$

$$n(m + q) = 0 \quad (5.3)$$

$$p(m + q) = 0 \quad (5.4)$$

$$q^2 + np = -1 \quad (5.5)$$

Note que, se $m + q \neq 0$, então segue de (5.2) e (5.3) que $n = 0$ e $p = 0$. Daí, por (5.2) decorre que $m^2 = -1$, de modo que m não assumiria valores reais. Já, se $m + q = 0$, então $q = -m$ e assim, segue de (5.2) que

$$p = -\frac{1 + m^2}{n},$$

o que nos fornece

$$\mathbf{X} = \begin{bmatrix} m & n \\ -\frac{1+m^2}{n} & -m \end{bmatrix}. \quad (5.6)$$

As matrizes \mathbf{X} que satisfazem (5.1), isto é, da forma (5.6), são involuções. Tomando $m = 0$ e $n = -1$ em (5.6), obtemos uma matriz, a qual chamaremos de \mathcal{I} , dada por

$$\mathcal{I} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

A matriz acima será de grande importância no estabelecimento da relação existente entre os números complexos, isto porque será através dela que faremos a correspondência de um número complexo com uma matriz. Tomemos assim um número complexo $z = a + bi$ e uma matriz $\mathbf{Z} = a\mathbf{I} + b\mathcal{I}$. Notemos que

$$\mathbf{Z} = a\mathbf{I} + b\mathcal{I} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Como sabemos, dois números complexos são iguais se suas respectivas partes reais e imaginárias forem iguais e, além disso, duas matrizes são iguais se suas respectivas entradas forem iguais. Dessa forma, podemos estabelecer a seguinte sentença: *Dado um número complexo $z = a + bi$, existe uma única matriz*

$$\mathbf{Z} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

associada a z . Tal associação nos permite escrever os elementos de \mathbb{C} como sendo matrizes $\mathbf{Z} = a\mathbf{I} + b\mathcal{I}$, ou seja,

$$\mathbb{C} = \left\{ \mathbf{Z} = a\mathbf{I} + b\mathcal{I} ; a, b \in \mathbb{R}, \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ e } \mathcal{I} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right\}$$

Podemos assim afirmar que tanto as operações como as propriedades enunciadas no início desta seção podem ser provadas utilizando matrizes. Porém, como estas foram provadas no Capítulo 1, podemos então considerá-las provadas. Vejamos, por exemplo, que a multiplicação de números complexos do ponto de vista matricial se dá da mesma maneira. Tomemos, para tanto, números complexos $z_1 = a + bi$ e $z_2 = c + di$. Sabemos que estes correspondem, respectivamente, a duas matrizes

$$\mathbf{Z}_1 = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{e} \quad \mathbf{Z}_2 = \begin{bmatrix} c & -d \\ d & c \end{bmatrix}.$$

Daí, temos que

$$\begin{aligned} \mathbf{Z}_1 \cdot \mathbf{Z}_2 &= \begin{bmatrix} ac - bd & -(ad + bc) \\ ad + bc & ac - bd \end{bmatrix} \\ &= (ac - bd) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (ad + bc) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\ &= (ac - bd)\mathbf{I} + (ad + bc)\mathcal{I}. \end{aligned}$$

Outra propriedade interessante conecta o módulo de um número complexo e o determinante da matriz associada a este. O módulo de um número complexo $z = a + bi$ é definido como sendo

$$|z| = \sqrt{a^2 + b^2}.$$

Agora notemos que, sendo $\mathbf{Z} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ a matriz associada a z , segue

$$\det(\mathbf{Z}) = a^2 + b^2 = |z|^2,$$

isto é, o módulo de z pode ser obtido por meio do determinante de \mathbf{Z} .

5.2 Quatérnios

O termo *quatérnio* foi utilizado pela primeira vez, de acordo com KRAMER [35], por William Rowan Hamilton¹⁸. Em seu livro *Elements of Quaternions* [28], Hamilton definiu um *quatérnio* como sendo uma expressão quadrimomial da forma

$$q = w + ix + jy + kz,$$

onde $w, x, y, z \in \mathbb{R}$ e i, j, k tais que

$$i^2 = j^2 = k^2 = ijk = -1$$

e

$$\begin{array}{lll} ij = k & jk = i & ki = j \\ ji = -k & kj = -i & ik = -j \end{array}$$

¹⁸William Rowan Hamilton (1805-1865), indiscutivelmente o maior cientista da Irlanda, nasceu em Dublin (Irlanda). Ele realizou descobertas importantes tanto em Óptica, como em Mecânica Analítica, porém considerava a descoberta dos quatérnios o seu maior resultado [19]. Segundo VINCE [63], Hamilton é reconhecido como criador da Álgebra de Quatérnios, que se tornou a primeira álgebra não-comutativa a ser descoberta.

É importante lembrar que os quatérnios levaram algum tempo para assumirem a forma acima. Como atesta VINCE [65], Hamilton inicialmente imaginou que assim como os números complexos no plano possuem a forma $a + bi$, no espaço teriam a forma $a + bi + cj$. porém, o problema encontrado por Hamilton foi que, ao tomar dois termos $q_1 = a_1 + b_1i + c_1j$ e $q_2 = a_2 + b_2i + c_2j$, o produto entre eles resultava em

$$q_1q_2 = (a_1a_2 - b_1b_2 - c_1c_2) + (a_1b_2 + b_1a_2)i + (a_1c_2 + c_1a_2)j + b_1c_2ij + c_1b_2ji,$$

significando assim que a multiplicação não estava bem definida, devido à presença do termo $b_1c_2ij + c_1b_2ji$. Ainda de acordo com VINCE [65], Hamilton levou cerca de dez anos para resolver tal problema, chegando a conclusão de que seriam necessários três termos imaginários i, j, k obedecendo às regras supracitadas.

Agora, com uma breve noção histórica dos quatérnios em mente, vamos estudar algumas entre as várias propriedades que estes apresentam. Antes disso, uma observação. Assim como na vasta literatura sobre o assunto, vamos escrever, apenas por uma questão estética, um quatérnio sobre a forma

$$q = a + bi + cj + dk.$$

Já sabemos que $\mathbb{C} = \{a + bi; a, b \in \mathbb{R} \text{ e } i = \sqrt{-1}\}$ representa o conjunto dos números complexos. De forma parecida, podemos definir o conjunto dos quatérnio como sendo

$$\mathbb{H} = \{a + bi + cj + dk; a, b, c, d \in \mathbb{R} \text{ e } i^2 = j^2 = k^2 = ijk = -1\}.$$

Tanto a soma como a multiplicação dos elementos de \mathbb{H} , bem como as propriedades destas operações (comutatividade, distributividade, elemento neutro, etc) estão bem definidas. De fato, sejam $q_1 = a_1 + b_1i + c_1j$ e $q_2 = a_2 + b_2i + c_2j$ dois quatérnios. Então

- Soma: $q_1 + q_2 = (a_1 + a_2) + (b_1 + b_2)i + (c_1 + c_2)j + (d_1 + d_2)k$
- Produto: $q_1 \cdot q_2 = A + Bi + Cj + Dk$

$$\begin{aligned} A &= a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2 \\ B &= a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2 \\ C &= a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2 \\ D &= a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2 \end{aligned}$$

onde as expressões acima decorrem diretamente das propriedades de soma e multiplicação de números reais.

Cabe agora verificarmos se é possível representar quatérnios via matrizes, assim como ocorre com os números complexos. Felizmente, tal possibilidade existe. Para enxergá-la, basta observarmos a expressão do produto de dois quatérnios. Note a seguir que esta pode ser escrita, com um pequeno abuso de notação, como produto de matrizes

$$q_1 \cdot q_2 = \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} \begin{bmatrix} 1 & i & j & k \end{bmatrix} = \begin{bmatrix} a_1 & -b_1 & -c_1 & -d_1 \\ b_1 & a_1 & -d_1 & c_1 \\ c_1 & d_1 & a_1 & -b_1 \\ d_1 & -c_1 & b_1 & a_1 \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \end{bmatrix} \begin{bmatrix} 1 & i & j & k \end{bmatrix}$$

Daí, tomando $q_2 = 1 + 0 \cdot i + 0 \cdot j + 0 \cdot k = 1$, segue que

$$q_1 = \begin{bmatrix} a_1 & -b_1 & -c_1 & -d_1 \\ b_1 & a_1 & -d_1 & c_1 \\ c_1 & d_1 & a_1 & -b_1 \\ d_1 & -c_1 & b_1 & a_1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & i & j & k \end{bmatrix}.$$

Como q_1 foi tomado arbitrariamente, podemos concluir que *a todo quatérnio* $q = a + bi + cj + dk$ corresponde uma matriz da forma

$$\begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix}$$

Perceba que não foi mencionada a unicidade, pois uma tal representação não é única assim como para os números complexos. Vejamos que, por exemplo, podemos escrever q sob a forma

$$q = \begin{bmatrix} a & d & -b & -c \\ -d & a & c & -b \\ b & -c & a & -d \\ c & b & d & a \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} [1 \quad -k \quad i \quad j].$$

É possível representar um quatérnio $q = a + bi + cj + dk$ por uma matriz 2×2 . Afim de fazê-lo, notemos que um quatérnio pode ser escrito como um número complexo. Bastando para isso observar que, sendo $k = ij$

$$q = a + bi + cj + dk = (a + bi) + (c + di)j = z + wj,$$

onde $z = a + bi$ e $w = c + di$. Daí, tomando dois quatérnios

$$\begin{aligned} q_1 &= a_1 + b_1i + c_1j + d_1k = z_1 + w_1j \\ q_2 &= a_2 + b_2i + c_2j + d_2k = z_2 + w_2j \end{aligned}$$

e calculando o produto entre os mesmos, obtemos

$$q_1 \cdot q_2 = (z_1z_2 - w_1\overline{w_2}) + (z_1w_2 + w_1\overline{z_2}),$$

onde o resultado obtido deriva de uma reordenação dos termos (semelhantes a A, B, C, D) obtidos inicialmente na multiplicação de dois quatérnios. Notemos então que tal expressão pode ser relacionada ao produto de matrizes abaixo

$$\begin{bmatrix} z_1 & w_1 \\ -\overline{w_1} & \overline{z_1} \end{bmatrix} \begin{bmatrix} z_2 & w_2 \\ -\overline{w_2} & \overline{z_2} \end{bmatrix}$$

Daí, tomando $q_2 = 1$, decorre que q_1 está relacionado a matriz

$$\begin{bmatrix} z_1 & w_1 \\ -\overline{w_1} & \overline{z_1} \end{bmatrix}$$

Porém, como q_1 foi tomado arbitrariamente, segue que todo quatérnio $q = a + bi + cj + dk = z + wj$ pode ser relacionado à uma matriz 2×2 da forma

$$\begin{aligned} \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} &= \begin{bmatrix} a + bi & c + di \\ -c + di & a - bi \end{bmatrix} \\ &= a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + c \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \\ &= a\mathbf{I} + b\mathcal{I} + c\mathcal{J} + d\mathcal{K}, \end{aligned}$$

onde

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathcal{I} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad \mathcal{J} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \mathcal{K} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

Além disso, é facilmente verificável que $\mathcal{I}^2 = \mathcal{J}^2 = \mathcal{K}^2 = -\mathbf{I}$ e

$$\begin{array}{lll} \mathcal{I}\mathcal{J} = \mathcal{K} & \mathcal{J}\mathcal{K} = \mathcal{I} & \mathcal{K}\mathcal{I} = \mathcal{J} \\ \mathcal{J}\mathcal{I} = -\mathcal{K} & \mathcal{K}\mathcal{J} = -\mathcal{I} & \mathcal{I}\mathcal{K} = -\mathcal{J} \end{array}$$

Exemplo 5.1. O quatérnio $q = 1 + 2i + 3j + 4k$ pode ser representado tanto pela matriz

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

como pela matriz

$$\begin{bmatrix} 1 + 2i & 3 + 4i \\ -3 + 4i & 1 - 2i \end{bmatrix}.$$

O módulo de um quatérnio $q = a + bi + cj + dk$ é calculado de maneira semelhante ao de um número complexo, sendo representado por

$$|q| = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

Temos assim que, por exemplo, o módulo de $q = 1 + 2i + 3j + 4k$ é dado por

$$|q| = \sqrt{1^2 + 2^2 + 3^2 + 4^2} = \sqrt{30}.$$

Por outro lado, se calcularmos o determinante da matriz associada a q , segue que

$$\begin{vmatrix} 1 & -2 & -3 & -4 \\ 2 & 1 & -4 & 3 \\ 3 & 4 & 1 & -2 \\ 4 & -3 & 2 & 1 \end{vmatrix} = 900 = (\sqrt{30})^4.$$

Podemos assim suspeitar que existe alguma relação entre o determinante da matriz associada a um quatérnio e o seu módulo. De fato, dado um quatérnio $q = a + bi + cj + dk$, decorre que

$$\begin{vmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{vmatrix} = (a^2 + b^2 + c^2 + d^2)^2 = (\sqrt{a^2 + b^2 + c^2 + d^2})^4,$$

ou seja, *o determinante da matriz associada é igual à quarta potência do módulo de um quatérnio.*

Uma outra maneira de associar um quatérnio a uma matriz 2×2 é escrevendo sua matriz 4×4 como uma matriz de blocos

$$\left[\begin{array}{cc|cc} a & -b & -c & -d \\ b & a & -d & c \\ \hline c & d & a & -b \\ d & -c & b & a \end{array} \right]$$

Agora notemos que

$$\begin{bmatrix} -c & -d \\ -d & c \end{bmatrix} = \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Em seguida, tomando

$$\mathbf{Z} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} c & -d \\ d & c \end{bmatrix}, \quad \mathcal{L} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

segue que podemos escrever a matriz associada a q simplesmente como

$$\begin{bmatrix} \mathbf{Z} & \mathbf{W}\mathcal{L} \\ -\mathbf{W}\mathcal{L} & \mathbf{Z} \end{bmatrix},$$

onde as matrizes \mathbf{Z} e \mathbf{W} são associadas, respectivamente, aos números complexos

$$z = a + bi \quad \text{e} \quad w = c + di.$$

Considerações Finais

A realização do presente trabalho traz uma abordagem que, espera-se ser significativa tanto para o ensino, bem como para a aprendizagem não só de Álgebra Linear, mas também de outros ramos do conhecimento onde as matrizes e suas propriedades são utilizadas.

A abordagem das cônicas sob posições incomuns no plano, por exemplo, realizada no Capítulo 3, gera a possibilidade de tratar de temas como transformações lineares no plano para alunos de Ensino Médio, ainda que de maneira elementar. Este é um ponto de partida para apresentar aos alunos uma das tantas aplicações das matrizes. Além disso, quando e se possível, trabalhar de tal tema utilizando softwares matemáticos, pois estes proporcionam a visualização geométrica das transformações lineares.

As diferentes maneiras de tratar de matrizes utilizando suas aplicações são diversas, uma vez que são deixados vários exemplos de teor didático, tal como o Exemplo 1.15, o qual mostra como criptografar uma mensagem utilizando matrizes, o que pode proporcionar uma aula diferenciada sem fugir do tema principal. Até mesmo os tópicos que geralmente são vistos em graduações que possuem Álgebra Linear na matriz curricular, tais como Exponencial de Matriz, podem ser trabalhados em sala de aula, claro, dependendo do nível de aprendizado em que a turma se encontre, pois como é sabido, o ensino de Matemática, de acordo com COIMBRA [13], apresenta os mesmos problemas e queixas por parte dos alunos em relação ao aprendizado dos conteúdos matemáticos, desde as séries mais elementares até o curso superior.

O que se espera deste trabalho é que, de alguma forma, ele possa contribuir positivamente para o processo de ensino-aprendizagem tanto para o Ensino Médio, como também para o Ensino Superior. No que compete ao professor, é esperado que o material aqui apresentado leve fornecça alguma nova alternativa para expor os conteúdos voltado às matrizes, de uma maneira mais dinâmica, principalmente nos casos das turmas que possuem mais dificuldades em compreender a teoria por trás das matrizes e suas propriedades. A busca de novas ideias é de grande importância, pois, como ressalta PRADO [54] apud MORAES [48], o professor precisa aprender a construir e a comparar novas estratégias de ações, novas teorias e novas formas de enfrentar um problema.

Espera-se, de um modo desafiador, que este material possa levantar ideias de trabalhos, agora direcionado ao ensino de matrizes no Ensino Fundamental, de modo que os alunos possam estabelecer familiaridade com o conteúdo desde cedo, mesmo na eventualidade de que sejam apenas as ideias básicas, como por exemplo, as quatro operações entre matrizes. Portanto, espera-se que de algum modo, o trabalho aqui apresentado possa contribuir de maneira auspiciosa para a Educação.

Referências

- [1] ALLEN, Linda JS. **An Introduction to Mathematical Biology**. 2007. ISBN, v. 10, p. 0-13.
- [2] ANDREESCU, Titu et al. **Complex Numbers from A to... Z**. Boston: Birkhäuser, 2006.
- [3] ANTON, Howard; RORRES, Chris. **Elementary linear algebra: applications version**. John Wiley & Sons, 2013.
- [4] APT, Krzysztof R. Edsger Wybe Dijkstra (1930–2002): A portrait of a genius. **Formal Aspects of Computing**, p. 92-98, 2002.
- [5] AYUB, A. et al. Perceived ease of use and usefulness of dynamic mathematical software: Experiences of Malaysian secondary students. In: **Proceeding of the 7th WSEAS International Conference on Education and Educational Technology**. Venice, Nov. 2008. p. 2008.
- [6] BALL, Walter William Rouse. **A short account of the history of mathematics**. Courier Corporation, 1960.
- [7] BAPAT, Ravindra B. **Graphs and matrices**. London: Springer, 2010.
- [8] BASHARIN, Gely P.; LANGVILLE, Amy N.; NAUMOV, Valeriy A. The life and work of AA Markov. **Linear algebra and its applications**, v. 386, p. 3-26, 2004.
- [9] BELHOSTE, Bruno. **Augustin-Louis cauchy: a biography**. Springer Science & Business Media, 2012.
- [10] BOURBAKI, Nicolas. **Elements of the History of Mathematics**. Springer Science & Business Media, 1998.
- [11] BRANDES, Ulrik. **Network analysis: methodological foundations**. Springer Science & Business Media, 2005.
- [12] CAMPOLINO, Marcio Lopes. **Translação e rotação de cônicas em \mathbb{R}^2** . Dissertação (Mestrado em Matemática) - PROFMAT, Universidade de Brasília. Brasília, 2014.
- [13] COIMBRA, Jarbas L. **Alguns aspectos problemáticos relacionados ao ensino-aprendizagem da álgebra linear**. Tese de Doutorado. Dissertação de mestrado, Universidade Federal do Pará, Pará, 2008.

- [14] COSME, Iria Caline Saraiva. **Um método para o cálculo da inversa de matrizes em blocos com uso limitado de memória**. 2018. Tese(Doutorado em Engenharia Elétrica e de Computação) - Universidade Regional do Rio Grande do Norte,2018.
- [15] CUNHA, Simone Miguez; CARRILHO, Denise Madruga. O processo de adaptação ao ensino superior e o rendimento acadêmico. **Psicologia escolar e educacional**, v. 9, n. 2, p. 215-224, 2005.
- [16] DANTE, Luiz Roberto. **Matemática: contexto e aplicações, vol. 3**. Ensino Médio. São Paulo: Ática, 2010.
- [17] DENCKER, Ada de Freitas Maneti. **A pesquisa ea interdisciplinaridade no ensino superior: uma experiência no curso de turismo**. Tese de Doutorado. 2000.
- [18] DIESTEL, Reinhard; SCHRIJVER, Alexander; SEYMOUR, Paul. **Graph theory**. Oberwolfach Reports, v. 7, n. 1, p. 521-580, 2010.
- [19] DIMITRIĆ, Radoslav M.; GOLDSMITH, Brendan. Sir William Rowan Hamilton. **Pokroky matematiky, fyziky a astronomie**, v. 35, n. 5, p. 277-279, 1990.
- [20] DUNHAM, William (Ed.). **The genius of Euler: reflections on his life and work**. MAA, 2007.
- [21] DUNHAM, William. **Euler: The master of us all**. American Mathematical Soc., 2020.
- [22] FAVARÃO, Neide Rodrigues Lago; ARAÚJO, Cíntia de Souza Alferes. Importância da interdisciplinaridade no ensino superior. **Educere-Revista da Educação da UNIPAR**, v. 4, n. 2, 2004.
- [23] FIORENTINI, Dario et al. Uma reflexão sobre o uso de materiais concretos e jogos no Ensino da Matemática. **Boletim da SBEM-SP**, v. 4, n. 7, 1990.
- [24] FRANA, Philip L.; MISA, Thomas J. An interview with Edsger W. Dijkstra. **Communications of the ACM**, v. 53, n. 8, p. 41-47, 2010.
- [25] GENTLE, James E. **Matrix algebra**. Springer texts in statistics, Springer, New York, NY, doi, v. 10, p. 978-0, 2007.
- [26] GLADCHEFF, Ana Paula; ZUFFI, Edna Maura; SILVA, DM da. Um instrumento para avaliação da qualidade de softwares educacionais de matemática para o ensino fundamental. In: **Anais do XXI Congresso da Sociedade Brasileira de Computação**. 2001.
- [27] GOLUB, Gene H.; VAN LOAN, Charles F. **Matrix computations**. JHU press, 2012.
- [28] HAMILTON, William Rowan. **Elements of quaternions**. Longmans, Green, & Company, 1866.
- [29] HERMITE, Charles. Oeuvres de Charles Hermite. Gauthier-Villars, 1905.

- [30] HILL, Richard W.; WYSE, Gordon A.; ANDERSON, Margareth. **Animal Physiology**. 3rd ed. - Sinauer Associates Inc, 2012.
- [31] HILL, Lester S. **Cryptography in an algebraic alphabet**. The American Mathematical Monthly, v. 36, n. 6, p. 306-312, 1929.
- [32] ISMAIL, I. A.; AMIN, Mohammed; DIAB, Hossam. How to repair the Hill cipher. **Journal of Zhejiang University-Science A**, v. 7, n. 12, p. 2022-2030, 2006.
- [33] KAZUNGA, Cathrine; BANSILAL, Sarah. Misconceptions about determinants. In: **Challenges and strategies in teaching linear algebra**. Springer, Cham, 2018. p. 127-145.
- [34] KOSINSKI, A. A. Cramer's rule is due to Cramer. **Mathematics Magazine**, v. 74, n. 4, p. 310-312, 2001.
- [35] KRAMER, Jürg; VON PIPPICH, Anna-Maria. **From Natural Numbers to Quaternions**. Springer, 2017.
- [36] KREJCAR, Ondrej (Ed.). **Modern telemetry**. BoD–Books on Demand, 2011
- [37] KREYSZIG, E. et al. **Advanced Engineering Mathematics**. Wiley, Hoboken, NJ (2011).
- [38] LARSON, Ron. **Elementary linear algebra**. 8 ed. Nelson Education, 2016.
- [39] LEDERMANN, Walter; NEUMANN, Peter M. **The life of Issai Schur through letters and other documents**. PROGRESS IN MATHEMATICS-BOSTON-, v. 210, p. xlv-xc, 2003.
- [40] LEON, Steven J.; BICA, Ion; HOHN, Tiina. **Linear algebra with applications**. Upper Saddle River, NJ: Prentice Hall, 1998.
- [41] LEWIS, Bert et al. Changes in size and age of Chinook salmon *Oncorhynchus tshawytscha* returning to Alaska. **PLoS One**, v. 10, n. 6, 2015.
- [42] LIM, Yongdo. The inverse mean problem of geometric mean and contraharmonic means. **Linear algebra and its applications**, v. 408, p. 221-229, 2005.
- [43] LIMA, Elon Lages. **Algebra linear**. 8 ed - IMPA, Rio de Janeiro, 2012.
- [44] LIMA, Elon Lages. **Geometria Analítica e Álgebra Linear**, 2ª ed. IMPA, Rio de Janeiro, 2012.
- [45] MENEZES, Alfred J. et al. **Handbook of applied cryptography**. CRC press, 1996.
- [46] MEYER, Carl D. **Matrix analysis and applied linear algebra**. Siam, 2000.

- [47] MONTGOMERY, Douglas C.; PECK, Elizabeth A.; VINING, G. Geoffrey. **Introduction to linear regression analysis**. John Wiley & Sons, 2012.
- [48] MORAES, Maria Candida. **O paradigma educacional emergente: implicações na formação do professor e nas práticas pedagógicas**. Em aberto, v. 16, n. 70, 2008.
- [49] MURRAY, James D. **Mathematical biology I: An introduction**. Springer Science & Business Media, 2007.
- [50] NIEDERHAUSER, Dale S.; STODDART, Trish. Teachers' instructional perspectives and use of educational software. **Teaching and teacher education**, v. 17, n. 1, p. 15-31, 2001.
- [51] NÓBREGA, Luciano Xavier Gomes da. **Princípio da Indução Matemática no Ensino Médio**. Dissertação de Mestrado. Universidade Federal do Rio Grande do Norte, 2013.
- [52] PACHECO, José Adson D.; BARROS, Janaina V. O uso de softwares educativos no ensino de matemática. **Revista Diálogos**, v. 8, p. 5-13, 2013.
- [53] POOLE, David. **Linear algebra: A modern introduction**. Cengage Learning, 2014.
- [54] PRADO, Maria Elisabette Brisola Brito. O uso do computador nos cursos de formação de professor: um enfoque reflexivo da prática pedagógica. **Revista Brasileira de Informática na Educação**, v. 3, n. 1, p. 63-64, 1998.
- [55] QUINN, Thomas P. **The Behavior and Ecology of Pacific Salmon and Trout**. 2nd ed. - University of Washington Press, 2018.
- [56] REICH, Karin. Carl Friederich Gauss (1777-1855). **Conferències FME**, p. 75, 1977.
- [57] RICARDO, Henry. **A modern introduction to linear algebra**. CRC Press, 2009.
- [58] SOARES, Ana Paula; ALMEIDA, Leandro S. Transição para a universidade: Apresentação e validação do Questionário de Expectativas Acadêmicas (**QEA**). 2001.
- [59] STAKHOV, Alexey. **The mathematics of harmony: From Euclid to contemporary mathematics and computer science**. World Scientific, 2009.
- [60] TERRADAS, Rodrigo Donizete. A importância da interdisciplinaridade na educação matemática. **Revista da Faculdade de Educação**, v. 14, n. 16, p. 95-114, 2019.
- [61] TWEEDIE, Charles. A study of the life and writings of Colin MacLaurin. **The mathematical gazette**, v. 8, n. 119, p. 133-151, 1915.

- [62] VASUDEV, C. **Graph theory with applications**. New Age International, 2006.
- [63] VINCE, John. **Quaternions for computer graphics**. Springer Science & Business Media, 2011.
- [64] WILLIAM, Mary E. "Hill, Lester Sanders". **Complete Dictionary of Scientific Biography**. Encyclopedia.com. 12 Aug. 2020 <<https://www.encyclopedia.com>>, acessado em 22 de setembro de 2020, 16h34min.

Apêndice

Embora muitos itens ainda possam ser provados, a ligação entre números complexos e matrizes vai além de propriedades. Quando uma matriz apresenta números complexos em suas entradas, esta pode acabar formando matrizes de grande interesse de estudo. Um exemplo disso pode ser dado por meio do conceito de *conjugado* de um número complexo $z = a + bi$, o qual é definido como

$$\bar{z} = a - bi,$$

ou seja, o conjugado de um número complexo é obtido por meio da preservação de sua parte real e da troca do sinal de sua parte imaginária. Vale ressaltar que nos casos em que $\bar{z} = z$, tem-se $z = a \in \mathbb{R}$. Em outras palavras, se um número complexo é igual ao seu conjugado, então ele é real.

Dessa maneira, podemos definir o conceito de *conjugado* de uma matriz $\mathbf{M} = \mathbf{A} + i\mathbf{B}$, onde \mathbf{A}, \mathbf{B} são matrizes com entradas reais, como sendo

$$\overline{\mathbf{M}} = \mathbf{A} - i\mathbf{B}.$$

Por exemplo, o conjugado da matriz

$$\mathbf{M} = \begin{bmatrix} 3 + i & -5 \\ 1 - 2i & 4 + 4i \end{bmatrix}$$

é dado por

$$\overline{\mathbf{M}} = \begin{bmatrix} 3 - i & -5 \\ 1 + 2i & 4 - 4i \end{bmatrix}.$$

Denotaremos por \mathbf{M}^H a transposta do conjugado de uma matriz $\mathbf{M} = \mathbf{A} + i\mathbf{B}$, ou seja,

$$\mathbf{M}^H = \overline{\mathbf{M}}^T.$$

Nos casos em que ocorre $\mathbf{M} = \mathbf{M}^H$, diz-se que \mathbf{M} é uma *matriz hermitiana*¹⁹. Obter exemplos de matrizes hermitianas é uma tarefa simples. Para comprovar isso, considere-mos uma matriz

$$\mathbf{M} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix}$$

com $m_i \in \mathbb{C}$, $i \in \{1, 2, 3, 4\}$, e suponhamos que a mesma é hermitiana. Desse modo, devemos ter

$$\begin{bmatrix} \overline{m_1} & \overline{m_3} \\ \overline{m_2} & \overline{m_4} \end{bmatrix} = \mathbf{M}^H = \mathbf{M} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix}.$$

¹⁹Charles Hermite (1822-1901) foi um matemático francês que realizou grandes contribuições para a matemática pura, em especial, Teoria dos Números e Álgebra. Em 1958 ele resolveu a equação de grau 5 por funções elípticas e, em 1873, ele provou que o número de Euler e é transcendental. O legado de seu trabalho pode ser demonstrado no grande número de termos matemáticos que levam o adjetivo “hermitiano(a)” [29].

Daí, podemos concluir que m_1 e m_4 devem ser números reais, ao passo que m_2 e m_3 devem ser um o conjugado do outro. Logo, com base em tais informações, podemos obter facilmente uma matriz hermitiana, como por exemplo

$$\begin{bmatrix} 2 & 3 - i \\ 3 + i & 5 \end{bmatrix}.$$

Note ainda que, se as entradas de uma matriz hermitiana são todas reais, então tal matriz é simétrica, isto é, $\mathbf{M}^H = \mathbf{M}^T$, o que por sua vez equivale dizer que *se \mathbf{M} é simétrica, então é hermitiana*. Já no caso em que \mathbf{M} é hermitiana e triangular, então \mathbf{M} é uma matriz diagonal, visto que as entradas a_{ij} , com $i \neq j$ deverão ter conjugado igual a 0, o que força-as serem iguais a zero, restando somente os valores da diagonal principal.

Antes de prosseguirmos devemos aqui definir a norma de um vetor $\mathbf{z} = (z_1, \dots, z_n)$, $z_i \in \mathbb{C}$, a qual é dada por

$$\|\mathbf{z}\| = \sqrt{|z_1|^2 + \dots + |z_n|^2}.$$

Temos, por exemplo, que

$$\|(7 - 2i, 7 + 2i)\| = \sqrt{\sqrt{7^2 + 2^2} + \sqrt{7^2 + 2^2}} = \sqrt{2\sqrt{53}}.$$

Note que, se

$$\mathbf{z} = \begin{bmatrix} 7 - 2i \\ 7 + 2i \end{bmatrix}$$

então

$$\mathbf{z}^H = [7 + 2i \quad 7 - 2i]$$

e assim

$$\|\mathbf{z}^H \mathbf{z}\| = \left\| \begin{bmatrix} 53 \\ 53 \end{bmatrix} \right\| = \sqrt{2\sqrt{53}}.$$

Isso acontece devido ao fato de que, dado um número complexo $z = a + bi$, tem-se que

$$|z|^2 = \bar{z}z,$$

o que nos permite escrever

$$\|\mathbf{z}\| = \sqrt{\|\mathbf{z}^H \mathbf{z}\|}.$$

Como $\|\mathbf{z}\| = \sqrt{\langle \mathbf{z}, \mathbf{z} \rangle}$, tal expressão nos unduz a definir o produto interno para vetores complexos da forma

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H \mathbf{z}.$$

De fato, dados três vetores $\mathbf{u} = (u_1, \dots, u_n)$, $\mathbf{w} = (w_1, \dots, w_n)$, $\mathbf{z} = (z_1, \dots, z_n)$ e escalares reais α, β , podemos comprovar isso vendo que tal produto possui

- *Positividade:* $\langle \mathbf{z}, \mathbf{z} \rangle \geq 0$ e $\langle \mathbf{z}, \mathbf{z} \rangle = 0 \iff \mathbf{z} = \mathbf{0}$.

De fato,

$$\langle \mathbf{z}, \mathbf{z} \rangle = \|\mathbf{z}\|^2 \geq 0.$$

Além disso, se $\langle \mathbf{z}, \mathbf{z} \rangle = 0$, então $\|\mathbf{z}\| = 0 \iff \mathbf{z} = \mathbf{0}$.

- *Simetria:* $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$

Com efeito, temos por um lado que

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H \mathbf{z} = z_1 \overline{w_1} + \cdots + z_n \overline{w_n},$$

enquanto que

$$\overline{\langle \mathbf{z}, \mathbf{w} \rangle} = \overline{\mathbf{z}^H \mathbf{w}} = \overline{z_1 w_1 + \cdots + z_n w_n} = z_1 \overline{w_1} + \cdots + z_n \overline{w_n} = \langle \mathbf{z}, \mathbf{w} \rangle$$

- *Linearidade:* $\langle \alpha \mathbf{z} + \beta \mathbf{w}, \mathbf{u} \rangle = \alpha \langle \mathbf{z}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{u} \rangle$

Basta notar que

$$\langle \alpha \mathbf{z} + \beta \mathbf{w}, \mathbf{u} \rangle = \mathbf{u}^H (\alpha \mathbf{z} + \beta \mathbf{w}) = \alpha \mathbf{u}^H \mathbf{z} + \beta \mathbf{u}^H \mathbf{w} = \alpha \langle \mathbf{z}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{u} \rangle$$

Podemos agora provar o seguinte resultado.

Teorema A.1. *Os vetores associados às raízes do polinômio característico de uma matriz hermitiana são todos reais. Além disso, se tais raízes são distintas, então os vetores associados são ortogonais.*

Demonstração. Dada uma matriz hermitiana \mathbf{A} , seja λ a raiz do polinômio característico associada ao vetor \mathbf{x} , tomando $k = \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle$, segue que

$$\overline{k} = \overline{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle} = \langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle = (\mathbf{A}\mathbf{x})^H \mathbf{x} = \mathbf{x}^H \mathbf{A}^H \mathbf{x} = \mathbf{x}^H \mathbf{A}\mathbf{x} = \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = k,$$

Daí, temos que k é real. Agora note que $k = \mathbf{x}^H \mathbf{A}\mathbf{x} = \mathbf{x}^H \lambda \mathbf{x}$, pois $\mathbf{A}\mathbf{x} = \lambda \mathbf{x}$, e portanto

$$k = \lambda \mathbf{x}^H \mathbf{x} = \lambda \|\mathbf{x}\|^2 \iff \lambda = \frac{k}{\|\mathbf{x}\|^2}.$$

Logo, como k e $\|\mathbf{x}\|^2$ são ambos números reais, decorre que λ é real. Agora, consideremos dois vetores \mathbf{x}_1 e \mathbf{x}_2 associados a duas raízes λ_1 e λ_2 do polinômio característico de \mathbf{A} , tais que $\lambda_1 \neq \lambda_2$. Assim, notemos que, por um lado

$$\langle \mathbf{x}_2, \mathbf{A}\mathbf{x}_1 \rangle = \langle \mathbf{x}_2, \lambda_1 \mathbf{x}_1 \rangle = \lambda_1 \langle \mathbf{x}_2, \mathbf{x}_1 \rangle,$$

ao passo que

$$\langle \mathbf{x}_2, \mathbf{A}\mathbf{x}_1 \rangle = \mathbf{x}_1^H \mathbf{A}\mathbf{x}_2 = \mathbf{x}_1^H \lambda_2 \mathbf{x}_2 = \lambda_2 \mathbf{x}_1^H \mathbf{x}_2 = \lambda_2 \langle \mathbf{x}_2, \mathbf{x}_1 \rangle.$$

Obtemos assim,

$$\lambda_1 \langle \mathbf{x}_2, \mathbf{x}_1 \rangle = \lambda_2 \langle \mathbf{x}_2, \mathbf{x}_1 \rangle \iff (\lambda_1 - \lambda_2) \langle \mathbf{x}_2, \mathbf{x}_1 \rangle = 0.$$

Logo, como $\lambda_1 \neq \lambda_2 \iff \lambda_1 - \lambda_2 \neq 0$, decorre que

$$\langle \mathbf{x}_2, \mathbf{x}_1 \rangle = 0,$$

ou seja, \mathbf{x}_1 e \mathbf{x}_2 são ortogonais. □

Definição A.1. Dois vetores complexos \mathbf{u}_1 e \mathbf{u}_2 são ditos **ortonormais** quando $\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = 1$ e $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = 0$. Mais geralmente, um conjunto de vetores complexos $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ é dito **ortonormal** se $\|\mathbf{u}_1\| = \dots = \|\mathbf{u}_n\| = 1$ e $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$, para todo $i \neq j$, com $i, j = 1, 2, \dots, n$. Dessa forma, uma matriz complexa \mathbf{U} é dita **unitária** quando seus vetores coluna formam um conjunto ortonormal. Além disso, \mathbf{U} é dita **ortogonal** quando $\mathbf{U}^H = \mathbf{U}^{-1}$.

Teorema A.2. Uma matriz complexa \mathbf{U} é ortogonal se, e somente se, é unitária.

Demonstração. Com efeito, escrevendo \mathbf{U} como uma matriz de blocos

$$[\mathbf{u}_1 \quad \dots \quad \mathbf{u}_n],$$

decorre que as entradas da matriz $\mathbf{U}^H \mathbf{U}$ serão dadas por $\mathbf{u}_i^H \mathbf{u}_j$. Sendo assim, se \mathbf{U} for unitária, então $\mathbf{u}_i^H \mathbf{u}_j = 1$ para $i = j$ e $\mathbf{u}_i^H \mathbf{u}_j = 0$ para $i \neq j$, de modo que $\mathbf{U}^H \mathbf{U} = \mathbf{I}$, o que equivale a $\mathbf{U}^H = \mathbf{U}^{-1}$ e portanto $|\mathbf{U}|$ é ortogonal. Reciprocamente, se \mathbf{U} é ortogonal, então as entradas da matriz $\mathbf{U}^H \mathbf{U}$, a saber, $\mathbf{u}_i^H \mathbf{u}_j$ são iguais a 1 para $i = j$ e iguais a 0 para $i \neq j$, ou seja, o conjunto das colunas de \mathbf{U} é ortonormal e, por conseguinte, tal matriz é unitária. \square

A Definição A.1 em conjunto com o Teorema A.1 garantem que se os vetores associados às raízes do polinômio característico de uma matriz \mathbf{A} são distintos, então existe uma matriz unitária \mathbf{U} que diagonaliza \mathbf{A} . De fato, se tais raízes são distintas, então os vetores associados são ortogonais e portanto, de acordo com o Teorema 3.2, estes são linearmente independentes. Logo, segue pelo Teorema 4.1 que \mathbf{A} é diagonalizável e portanto, ortogonalmente diagonalizável. Assim, existe uma matriz \mathbf{X} que ortogonaliza \mathbf{A} , ou seja, $\mathbf{A} = \mathbf{X} \mathbf{D} \mathbf{X}^T$. Daí, basta normalizarmos os vetores coluna \mathbf{x}_i de \mathbf{X} , tais que

$$\mathbf{u}_i = \frac{1}{\|\mathbf{x}_i\|} \mathbf{x}_i,$$

obtendo assim uma matriz matriz \mathbf{U} com colunas \mathbf{u}_i , as quais formam um conjunto ortonormal, de modo que \mathbf{U} é unitária e portanto

$$\mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{U}^H.$$

Exemplo A.1. Considere a matriz hermitiana

$$\begin{bmatrix} 2 & 3-i \\ 3+i & 5 \end{bmatrix}.$$

Temos que as raízes do polinômio característico dessa matriz são $\lambda_1 = 0$, associada ao vetor $\mathbf{x}_1 = (-3+i, 2)$ e $\lambda_2 = 7$, associada ao vetor $\mathbf{x}_2 = (3-i, 5)$. Note que tais raízes são reais e distintas, o que implica, de acordo com o Teorema 5.1, que \mathbf{x}_1 e \mathbf{x}_2 são ortogonais. Normalizando tais vetores, obtemos vetores ortonormais \mathbf{u}_1 e \mathbf{u}_2 , tais que

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{x}_1\|} \mathbf{x}_1 = \begin{bmatrix} \frac{-3+i}{\sqrt{14}} \\ \frac{2}{\sqrt{14}} \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2\|} \mathbf{x}_2 = \begin{bmatrix} \frac{3-i}{\sqrt{35}} \\ \frac{5}{\sqrt{35}} \end{bmatrix}.$$

Finalmente, os vetores acima compo em uma matriz unitária, a qual diagonaliza ortogonalmente a matriz informada inicialmente, pois

$$\begin{bmatrix} \frac{-3+i}{\sqrt{14}} & \frac{3-i}{\sqrt{35}} \\ \frac{2}{\sqrt{14}} & \frac{5}{\sqrt{35}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 7 \end{bmatrix} \begin{bmatrix} \frac{-3+i}{\sqrt{14}} & \frac{2}{\sqrt{14}} \\ \frac{3-i}{\sqrt{35}} & \frac{5}{\sqrt{35}} \end{bmatrix} = \begin{bmatrix} 2 & 3-i \\ 3+i & 5 \end{bmatrix}.$$

Antes de provarmos o próximo resultado, devemos apresentar aqui uma ferramenta de grande utilidade quando o assunto é vetores ortonormais. Para isso, vamos chamar de **projeção ortogonal do vetor \mathbf{w} sobre o eixo que contém \mathbf{x}** o vetor $\langle \mathbf{x}, \mathbf{w} \rangle \mathbf{x}$. Observemos então que os vetores \mathbf{x} e $\mathbf{u} = \frac{1}{\|\mathbf{x}\|} \mathbf{x}$ estão alinhados e na mesma direção. Daí, escrevendo $\mathbf{x} = \|\mathbf{x}\| \mathbf{u}$, segue que

$$\langle \mathbf{x}, \mathbf{w} \rangle \mathbf{x} = \langle \|\mathbf{x}\| \mathbf{u}, \mathbf{w} \rangle \|\mathbf{x}\| \mathbf{u} = \|\mathbf{x}\|^2 \langle \mathbf{u}, \mathbf{w} \rangle \mathbf{u} \iff \langle \mathbf{u}, \mathbf{w} \rangle \mathbf{u} = \frac{\langle \mathbf{x}, \mathbf{w} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} \mathbf{x},$$

ou seja, podemos representar a projeção do vetor \mathbf{w} sobre o eixo que contém \mathbf{x} como sendo

$$\text{pr}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{x}, \mathbf{w} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} \mathbf{x}.$$

Podemos, a partir daí, obter um vetor \mathbf{v} perpendicular a \mathbf{x} , desde que

$$\mathbf{v} = \mathbf{w} - \frac{\langle \mathbf{x}, \mathbf{w} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} \mathbf{x}.$$

Com efeito,

$$\langle \mathbf{x}, \mathbf{v} \rangle = \left\langle \mathbf{x}, \mathbf{w} - \frac{\langle \mathbf{x}, \mathbf{w} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} \mathbf{x} \right\rangle = \langle \mathbf{x}, \mathbf{w} \rangle - \frac{\langle \mathbf{x}, \mathbf{w} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} \langle \mathbf{x}, \mathbf{x} \rangle = 0,$$

o que comprova a ortogonalidade entre \mathbf{x} e \mathbf{v} . Vejamos a situação descrita acima na figura a seguir

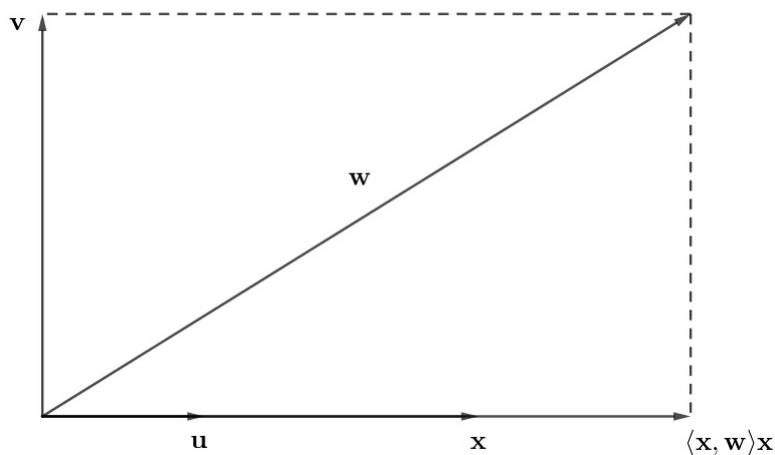


Figura A.1

Fonte: Adaptado de LIMA [43], p.123

Seguindo a ideia descrita acima, podemos obter, a partir de um conjunto de vetores $\mathbf{x}_1, \dots, \mathbf{x}_n$, um conjunto ortogonal $\mathbf{v}_1, \dots, \mathbf{v}_n$, bastando para isso, tomar

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{x}_1 \\ \mathbf{v}_2 &= \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 \\ \mathbf{v}_3 &= \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 \\ &\vdots \\ \mathbf{v}_n &= \mathbf{x}_n - \frac{\langle \mathbf{x}_n, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \dots - \frac{\langle \mathbf{x}_n, \mathbf{v}_{n-1} \rangle}{\langle \mathbf{v}_{n-1}, \mathbf{v}_{n-1} \rangle} \mathbf{v}_{n-1} \end{aligned}$$

Em seguida, normalizando cada um dos vetores \mathbf{v}_i sob a forma $\mathbf{y}_i = \frac{1}{\|\mathbf{v}_i\|} \mathbf{v}_i$, obtemos um conjunto $\mathbf{y}_1, \dots, \mathbf{y}_n$ de vetores ortonormais. Tal procedimento é denominado **processo de ortonormalização de Gram-Schmidt**.

Teorema A.3. (Teorema de Schur) *Seja \mathbf{A} uma matriz complexa de ordem n . Existe uma matriz unitária \mathbf{U} tal que $\mathbf{U}^H \mathbf{A} \mathbf{U}$ é uma matriz triangular superior.*

Demonstração. Para realizar a demonstração, vamos entender o processo utilizado na mesma, de modo que este poderá ser repetido sucessivas vezes até chegar-se ao resultado desejado.

- Para o caso $n = 1$ temos $\mathbf{A} = [a]$ e portanto a $\mathbf{U} = [1]$, de modo que $[1]^H \mathbf{A} [1] = [a]$.
- Para $n = 2$: Sejam \mathbf{x}_1 e \mathbf{x}_2 os vetores associados à respectivas raízes λ_1, λ_2 do polinômio característico da matriz \mathbf{A} . Pelo processo de ortonormalização de Gram-Schmidt podemos obter vetores ortonormais \mathbf{u}_1 e \mathbf{u}_2 (e assim $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = 0$, mas não necessariamente $\langle \mathbf{u}_2, \mathbf{u}_1 \rangle = 0$) associados a λ_1 e λ_2 , respectivamente. Daí, fazendo $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2]$, segue que

$$\begin{aligned} \mathbf{U}^H \mathbf{A} \mathbf{U} &= \begin{bmatrix} \mathbf{u}_1^H \\ \mathbf{u}_2^H \end{bmatrix} \mathbf{A} [\mathbf{u}_1 \ \mathbf{u}_2] \\ &= \begin{bmatrix} \mathbf{u}_1^H \\ \mathbf{u}_2^H \end{bmatrix} [\mathbf{A} \mathbf{u}_1 \ \mathbf{A} \mathbf{u}_2] \\ &= \begin{bmatrix} \mathbf{u}_1^H \mathbf{A} \mathbf{u}_1 & \mathbf{u}_1^H \mathbf{A} \mathbf{u}_2 \\ \mathbf{u}_2^H \mathbf{A} \mathbf{u}_1 & \mathbf{u}_2^H \mathbf{A} \mathbf{u}_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \mathbf{u}_1^H \mathbf{u}_1 & \lambda_2 \mathbf{u}_1^H \mathbf{u}_2 \\ \lambda_2 \mathbf{u}_2^H \mathbf{u}_1 & \lambda_2 \mathbf{u}_2^H \mathbf{u}_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & \lambda_2 \mathbf{u}_1^H \mathbf{u}_2 \\ 0 & \lambda_2 \end{bmatrix} \end{aligned}$$

o que resulta numa matriz triangular superior. Note que a última igualdade acima ocorre pelo fato de que, sendo \mathbf{u}_1 e \mathbf{u}_2 ortonormais, tem-se $\lambda_1 \mathbf{u}_1^H \mathbf{u}_1 = \lambda_1 \|\mathbf{u}_1\|^2 = \lambda_1$, $\lambda_2 \mathbf{u}_2^H \mathbf{u}_2 = \lambda_2 \|\mathbf{u}_2\|^2 = \lambda_2$ e $\lambda_2 \mathbf{u}_2^H \mathbf{u}_1 = \lambda_2 \langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \lambda_2 \cdot 0 = 0$.

- Para $n = 3$: Analogamente ao caso anterior, dados os vetores $\mathbf{x}_1, \mathbf{x}_2$ e \mathbf{x}_3 associados às respectivas raízes $\lambda_1, \lambda_2, \lambda_3$ do polinômio característico de \mathbf{A} , podemos utilizar

Gram-Schmidt para obter vetores ortonormais $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$. Daí, considerando $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{U}_1]$, onde $\mathbf{U}_1 = [\mathbf{u}_2 \ \mathbf{u}_3]$, temos que

$$\begin{aligned} \mathbf{U}^H \mathbf{A} \mathbf{U} &= [\mathbf{u}_1 \ \mathbf{U}_1]^H \mathbf{A} [\mathbf{u}_1 \ \mathbf{U}_1] \\ &= \begin{bmatrix} \mathbf{u}_1^H \\ \mathbf{U}_1^H \end{bmatrix} [\mathbf{A} \mathbf{u}_1 \ \mathbf{A} \mathbf{U}_1] \\ &= \begin{bmatrix} \mathbf{u}_1^H \mathbf{A} \mathbf{u}_1 & \mathbf{u}_1^H \mathbf{A} \mathbf{U}_1 \\ \mathbf{U}_1^H \mathbf{A} \mathbf{u}_1 & \mathbf{U}_1^H \mathbf{A} \mathbf{U}_1 \end{bmatrix} \end{aligned}$$

Agora observe que:

- i) $\mathbf{u}_1^H \mathbf{A} \mathbf{u}_1 = \lambda_1$
 ii) $\mathbf{u}_1^H \mathbf{A} \mathbf{U}_1$ é uma matriz linha, a saber,

$$[\lambda_2 \mathbf{u}_1^H \mathbf{u}_2 \quad \lambda_3 \mathbf{u}_1^H \mathbf{u}_3]$$

- iii) $\mathbf{U}_1^H \mathbf{A} \mathbf{u}_1$ é uma matriz coluna nula, pois

$$\mathbf{U}_1^H \mathbf{A} \mathbf{u}_1 = \begin{bmatrix} \lambda_1 \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \\ \lambda_1 \langle \mathbf{u}_1, \mathbf{u}_3 \rangle \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- iv) Pelo caso $n = 2$, temos

$$\mathbf{U}_1^H \mathbf{A} \mathbf{U}_1 = \begin{bmatrix} \lambda_2 & \lambda_2 \mathbf{u}_2^H \mathbf{u}_3 \\ 0 & \lambda_3 \end{bmatrix}$$

Logo,

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \left[\begin{array}{c|cc} \lambda_1 & \lambda_2 \mathbf{u}_1^H \mathbf{u}_2 & \lambda_3 \mathbf{u}_1^H \mathbf{u}_3 \\ \hline 0 & \lambda_2 & \lambda_2 \mathbf{u}_2^H \mathbf{u}_3 \\ 0 & 0 & \lambda_3 \end{array} \right]$$

obtendo assim uma matriz triangular superior.

- A ideia é a mesma para os demais casos, ou seja, dada uma matriz \mathbf{A} de ordem n , pode-se obter vetores ortonormais $\mathbf{u}_1, \dots, \mathbf{u}_n$ associados às respectivas raízes $\lambda_1, \dots, \lambda_n$ do polinômio característico de \mathbf{A} , tais que

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \left[\begin{array}{c|ccc} \lambda_1 & * & \dots & * \\ \hline 0 & & & \\ \vdots & & \mathbf{U}_1^H \mathbf{A} \mathbf{U}_1 & \\ 0 & & & \end{array} \right]$$

onde $\mathbf{U}_1^H \mathbf{A} \mathbf{U}_1$ é uma matriz triangular superior de ordem $n - 1$ e, portanto, $\mathbf{U}^H \mathbf{A} \mathbf{U}$ também o é.

□

Note que no teorema acima não foi mencionado que as raízes do polinômio característico de \mathbf{A} devem ser distintos, o que nos fornece um resultado ainda mais forte que aquele obtido anteriormente a ele. Consideremos, por exemplo, a matriz

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

Temos que as raízes do seu polinômio característico são $\lambda_1 = \lambda_2 = 1$. Daí, os vetores associados a estas são da forma $\mathbf{x} = \begin{bmatrix} 0 \\ y \end{bmatrix}$. Tomemos então

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Para obtermos um vetor $\mathbf{x}_2 = \begin{bmatrix} x \\ y \end{bmatrix}$ ortogonal a \mathbf{x}_1 , devemos ter

$$0 = \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = y.$$

Logo, $\mathbf{x}_2 = \begin{bmatrix} x \\ 0 \end{bmatrix}$. Tomemos $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Como \mathbf{x}_1 e \mathbf{x}_2 são unitários, não é necessário normalizá-los, de modo que a matriz unitária \mathbf{U} é dada por

$$\mathbf{U} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

o que nos fornece a matriz

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \begin{bmatrix} 2 & -3 \\ 0 & 2 \end{bmatrix}$$

a qual é triangular superior. Note que, desde que

$$\mathbf{U} = \begin{bmatrix} 0 & \frac{1}{\|\mathbf{x}_2\|} \\ \frac{1}{\|\mathbf{x}_1\|} & 0 \end{bmatrix}$$

tem-se que $\mathbf{U}^H \mathbf{A} \mathbf{U}$ é triangular superior. Vejamos a seguir um exemplo em que obteremos uma matriz unitária utilizando o processo de ortonormalização de Gram-Schmidt.

Exemplo A.2. Consideremos a matriz

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

Temos que as raízes de seu polinômio característico são $\lambda_1 = 1, \lambda_2 = \lambda_3 = -2$. Para $\lambda = \lambda_1$, obtemos vetores associados da forma

$$\begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}.$$

Já para $\lambda = \lambda_2$, obtemos vetores associados da forma

$$\begin{bmatrix} -3y \\ y \\ 0 \end{bmatrix}.$$

Tomando $y = 1$, obtemos

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{x}_2 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}.$$

Precisaremos de um vetor \mathbf{x}_3 tal que o conjunto $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ seja linearmente independente. Ora, seja $\mathbf{x}_3 = (x, y, z)$ tal que $\langle \mathbf{x}_2, \mathbf{x}_3 \rangle = 0$. Temos daí que

$$\mathbf{x}_3 = \begin{bmatrix} x \\ 3x \\ z \end{bmatrix}$$

Se $x = 1$ e $z = 0$, obtemos $\mathbf{x}_3 = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$. Com sorte, temos que \mathbf{x}_1 e \mathbf{x}_3 são linearmente independentes. Assim, como \mathbf{x}_1 e \mathbf{x}_2 também o são (pelo Teorema 4.1), segue que $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ é linearmente independente, o que nos permite aplicar Gram-Schmidt. Assim

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\ \mathbf{v}_2 &= \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 = \begin{bmatrix} -3 \\ 0 \\ 0 \end{bmatrix} \\ \mathbf{v}_3 &= \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$

Como \mathbf{v}_3 é nulo, logo não poderá ser coluna de uma matriz unitária. Vamos então substituí-lo pelo vetor $\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$, o qual é claramente linearmente independente em relação a \mathbf{v}_1 e \mathbf{v}_2 . Tomando $z = 1$, obtemos

$$\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Daí, como \mathbf{v}_1 e \mathbf{y} são unitários, segue que

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2\|} \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{u}_3 = \mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

fornecendo-nos assim

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

Finalmente

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

Vamos agora provar um dos teoremas mais belos deste trabalho, que é o

Teorema A.3. (Teorema Espectral) *Se \mathbf{A} é uma matriz hermitiana, então existe uma matriz unitária \mathbf{U} que diagonaliza \mathbf{A} .*

Demonstração. Sabemos, pelo Teorema de Schur que existe uma matriz unitária \mathbf{U} que diagonaliza \mathbf{A} , tal que

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{S},$$

onde \mathbf{S} é uma matriz triangular superior. Além disso, note que

$$\mathbf{S}^H = (\mathbf{U}^H \mathbf{A} \mathbf{U})^H = \mathbf{U}^H \mathbf{A}^H \mathbf{U} = \mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{S},$$

ou seja, \mathbf{S} é hermitiana. Além disso, como \mathbf{S} é triangular superior, decorre que a mesma é uma matriz diagonal. \square

O Teorema Espectral nos proporciona uma afirmação muito especial:

Toda matriz real simétrica é diagonalizável.

Isto porque toda matriz simétrica cujas entradas são números reais, é também hermitiana, logo pode ser diagonalizada.