Matrizes e resolução de problemas

Alexandre Hartung

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: ______________________

Alexandre Hartung

Matrizes e resolução de problemas

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Programa de Mestrado Profissional em Matemática. *VERSÃO REVISADA*

Área de Concentração: Matemática

Orientadora: Profa. Dra. Janete Crema Simal

USP – São Carlos
Junho de 2017

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

Hartung, Alexandre
H336m Matrizes e resolução de problemas /
Alexandre Hartung; orientadora Janete Crema Simal.
- São Carlos - SP, 2017.
133 p.

Dissertação (Mestrado - Programa de Pós-graduação
em Mestrado Profissional em Matemática em Rede
Nacional) - Instituto de Ciências Matemáticas e de
Computação, Universidade de São Paulo, 2017.

1. Álgebra Linear. 2. Matrizes. 3. Sistemas
Lineares. 4. Modelos Econômicos. I. Simal, Janete
Crema, orient. II. Título.

Alexandre Hartung

Matrices and problem solving

Master dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of Mathematics Professional Master's Program. *FINAL VERSION*

Concentration Area: Mathematics

Advisor: Profa. Dra. Janete Crema Simal

USP – São Carlos
June 2017

A minha esposa Isa e filhos, que compreenderam o tempo sacrificado em função desse trabalho e aos meus pais, por tudo o que eles representam.

AGRADECIMENTOS

Agradeço em primeiro lugar a Deus que iluminou o meu caminho durante esta caminhada.

Aos meus pais, irmãos, minha esposa Isa, meus filhos e toda minha família que, com muito carinho e apoio, não mediram esforços para que eu chegasse até essa etapa de minha vida.

A todos os professores, que foram tão importantes na minha vida acadêmica e no desenvolvimento desta dissertação.

Aos amigos e colegas, pelo incentivo e pelo apoio constante.

A CAPES pelo suporte financeiro.

RESUMO

HARTUNG, A. **Matrizes e resolução de problemas**. 2017. 133 p. Dissertação (Mestrado em Ciências – Programa de Mestrado Profissional em Matemática) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

Álgebra Linear e particularmente a teoria das matrizes e dos sistemas lineares são tópicos da Matemática que têm aplicações, não só dentro da própria Matemática, mas também em várias outras áreas do conhecimento humano. Neste trabalho, além de estudar estas teorias, estudamos algumas de suas aplicações na área da Economia, como em modelos lineares de produção, modelos de Markov para emprego e modelos de benefícios obtidos no pagamento de impostos após realizarmos contribuições filantrópicas.

Palavras-chave: Álgebra Linear, Matrizes, Sistemas Lineares, Modelos Econômicos.

ABSTRACT

HARTUNG, A. **Matrices and problem solving**. 2017. 133 p. Dissertação (Mestrado em Ciências – Programa de Mestrado Profissional em Matemática) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

Linear Algebra and particularly matrices and linear systems theory are topics in Mathematics with many applications in several branches of science. In this work we study this theory and some of its applications in Economy as in linear models of production, Markov models of employment and tax benefits of charitable contributions.

Keywords: Linear Algebra, Matrices, Linear Systems, Economic Models .

SUMÁRIO

INTRODUÇÃO

O estudo das matrizes tem importância singular na resolução de problemas que envolvem sistemas de equações lineares. Muitos são os problemas que são modelados por este tipo de sistemas: problemas geométricos, físicos, químicos, biológicos, etc. Mas neste trabalho nos interessamos particularmente por modelos ligados a Economia. Nossos objetivos foram basicamente dois. O primeiro foi refazer o estudo de sistemas lineares, verificando sua íntima relação com as equações matriciais. Neste caso, estudamos inicialmente sistemas lineares muito simples, e que pudessem ser resolvidos pela técnica da "substituição", para então resolver sistemas mais complexos que pudessem ser transformados nestes sistemas simples e equivalentes aos primeiros. A forma como direcionamos este estudo é também uma proposta para sistematizar o estudo dos sistemas lineares com três ou mais equações e variáveis. Acreditamos que inicialmente os alunos deveriam estudar os sistemas que aparecem na forma triangular, para então motivarmos a apresentação da forma de resolução por escalonamento. Para justificar matematicamente que o novo sistema era equivalente ao primeiro, fizemos o estudo aprofundado de matrizes e de equações matriciais.

O segundo objetivo foi estudar problemas na área de economia que fossem modelados através de sistemas de equações lineares. Aproveitando os conceitos explorados na teoria de matrizes e equações matriciais, fizemos então a análise de vários problemas econômicos.

Por fim, escolhemos o modelo linear de produção de bens, para elaborar um plano de aulas para alunos do segundo ou terceiro ano do ensino médio. Para este problema de produção, foi importante o estudo das matrizes "diagonais dominantes" o qual dedicamos um estudo mais detalhado e que eventualmente pode vir a se tornar um projeto para ser estudado também pelos alunos do ensino médio. Neste plano de aula, os alunos não só modelariam o problema, mas veriam as implicações em se mudar os parâmetros do problema e testar se estes continuariam (ou não) a ter soluções e qual o significado destes fatos. Cabe aqui observar que a existência de solução matemática pode ou não ter significado concreto para o problema, por exemplo, um modelo de produção que gera um sistema cuja solução existe e é negativa, pode não fazer sentido, já que não se pensa em produção negativa de bens.

Esta dissertação se divide da seguinte maneira. No capítulo 1, fizemos o estudo de sistemas lineares, preparando a linguagem que usaríamos para matrizes e equações matriciais. Destacamos a última seção onde estudamos as matrizes diagonais dominantes e o fato de serem sempre invertíveis com matriz inversa de entradas todas não negativas. Fato que será utilizado no modelo de produção. No capítulo 2, fizemos o estudo de determinantes de matrizes e suas propriedades. Neste, entre outras coisas, vimos como a expressão que define o determinante aparece naturalmente em escalonamento de matrizes não singulares e por isso pode ser utilizado como motivação (no caso de matrizes de ordem 2 ou 3) para defini-lo para os alunos do ensino médio. Além é claro, da motivação fundamental de ser o número associado a uma matriz quadrada que nos informa quando o sistema a ela relacionado terá ou não solução única. O capítulo 3 é dedicado ao estudo básico de espaços vetoriais. O capítulo 4 é dedicado ao estudo de três modelos econômicos: Modelo de benefícios de impostos através de contribuições filantrópicas, Modelo de empregos de Markov e o Modelo linear de produção. Por fim, no capítulo 5, elaboramos um plano de aulas sobre o modelo linear de produção, que possa ser aplicado nos anos finais do ensino médio.

CAPÍTULO

1

MATRIZES E SISTEMAS LINEARES

Neste capítulo, apresentamos os conceitos básicos sobre matrizes e sistemas de equações lineares, que foram consultados à partir de (SIMON; BLUME; DOERING, 2004), (BALDIN; FURUYA, 2011), (IEZZI; HAZZAN, 1995), (HEFEZ; FERNANDEZ, 2016) e (BOLDRINI *et al.*, 1980). Esses conceitos surgem naturalmente na resolução de muitos tipos de problemas de diversas áreas do conhecimento, e são essenciais, não somente porque "simplificam e ordenam" o problema, mas também porque fornecem novos métodos de resolução.

Veremos a importância do estudo das matrizes observando sua relação com sistemas de equações lineares. Assim, o sistema linear geral de m equações a n incógnitas pode ser escrito como

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\vdots \qquad \vdots \qquad\qquad & \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned} \tag{1.1}$$

Neste sistema, os a_{ij} e b_i são números reais dados; a_{ij} é o **coeficiente** da incógnita x_j na i-ésima equação.

Podemos simplificar a apresentação do sistema linear geral de m equações a n incógnitas usando seus coeficientes para escrever dois agrupamentos retangulares denominados **matrizes**.

1.1 Definição de matrizes

Dados dois números naturais e não nulos m e n, definimos uma matriz m por n (escreve-se $m \times n$) como uma tabela formada por números reais distribuídos em m linhas e n colunas. Cada elemento da matriz é indicado por a_{ij} e é chamado de entrada da matriz [1]. O índice i indica a linha e o índice j a coluna às quais o elemento pertence. Convenciona-se que as linhas sejam numeradas de cima para baixo (de 1 até m) e as colunas, da esquerda para a direita (de 1 até n).

Uma matriz A do tipo $m \times n$ também pode ser representada por $M = [a_{ij}]_{m\times n}$, onde $i \in \{1,2,3,\dots,m\}$ e $j \in \{1,2,3,\dots,n\}$ ou simplesmente por $A = [a_{ij}]$, quando a ordem da matriz estiver subentendida.

Como exemplo vamos escrever duas matrizes que são usadas para simplificar a apresentação do sistema linear geral de m equações a n incógnitas em (1.1). A primeira é

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

que recebe o nome de **matriz de coeficientes**. Observe que o número de colunas de A corresponde ao número de variáveis do sistema (1.1) e o número de linhas corresponde ao número de equações. Quando acrescentamos a coluna

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

correspondente ao lado direito do sistema geral, obtemos a matriz

$$\hat{A} = \begin{pmatrix} A & | & b \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & | & b_2 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & | & b_m \end{pmatrix}$$

que recebe o nome de **matriz aumentada**. As linhas verticais imediatamente antes da ultima coluna da matriz corresponde aos sinais de igualdade nas equações.

1.2 Operações com matrizes

Igualdade

Dizemos que duas matrizes $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{m\times n}$, de mesma ordem, são iguais, e escrevemos $A = B$ quando $a_{ij} = b_{ij}$ para todo $i \in \{1,2,3\dots,m\}$ e todo $j \in \{1,2,3\dots,n\}$.

Por exemplo, se x e y denotam números reais, temos que as matrizes

[1] As entradas de uma matriz não precisam ser necessariamente números reais, podem ser números complexos

$$\begin{pmatrix} 3 & x \\ 0 & -2 \end{pmatrix} \text{ e } \begin{pmatrix} 1 & -7 \\ y & 9 \end{pmatrix}$$

são iguais quando $x = -7$ e $y = 0$.

Adição

A soma de duas matrizes de mesma ordem, $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{m\times n}$, denotada $A+B$, é uma matriz $C = [c_{ij}]_{m\times n}$ tal que $c_{ij} = a_{ij} + b_{ij}$ para todo $i \in \{1,2,3\ldots,m\}$ e todo $j \in \{1,2,3\ldots,n\}$.

Por exemplo,

$$\begin{pmatrix} 2 & 5 \\ -1 & 6 \\ 4 & 3 \end{pmatrix} + \begin{pmatrix} -8 & 4 \\ 2 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 2-8 & 5+4 \\ -1+2 & 6+0 \\ 4+2 & 3+1 \end{pmatrix} = \begin{pmatrix} -6 & 9 \\ 1 & 6 \\ 6 & 4 \end{pmatrix}$$

Note que a soma de matrizes de ordens diferentes não está definida.

Subtração

Dada a matriz $A = [a_{ij}]_{m\times n}$, chama-se oposta de A e indica-se $-A = [-a_{ij}]$ de modo que $A + (-A) = 0$.

Dadas duas matrizes $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{m\times n}$, definimos $A - B = A + (-B)$.

Por exemplo,

$$\begin{aligned} \begin{pmatrix} 5 & 3 & 9 \\ -1 & 1 & 4 \end{pmatrix} - \begin{pmatrix} 2 & 4 & 0 \\ 6 & 5 & 7 \end{pmatrix} &= \begin{pmatrix} 5 & 3 & 9 \\ -1 & 1 & 4 \end{pmatrix} + \begin{pmatrix} -2 & -4 & 0 \\ -6 & -5 & -7 \end{pmatrix} = \\ &= \begin{pmatrix} 3 & -1 & 9 \\ -7 & -4 & -3 \end{pmatrix} \end{aligned}$$

A adição de matrizes apresenta as seguintes propriedades, como mostra o resultado a seguir.

Proposição 1. Dadas as matrizes reais A, B e C de mesma ordem $m \times n$, temos:

(i) $A + B = B + A$ (comutatividade)

(ii) $A + (B + C) = (A + B) + C$ (associatividade)

(iii) $A + 0 = A$ (a matriz 0, cujas entradas são todas zero, é uma *identidade aditiva*)

(iv) $A + (-A) = 0$.

Demonstração. As propriedades acima decorrem diretamente das definições de igualdade e adição de matrizes e das propriedades dos números reais.

(i) Se $A = [a_{ij}]$ e $B = [b_{ij}]$

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

onde usamos a comutatividade da adição de números reais.

(ii) Se $A = [a_{ij}]$, $B = [b_{ij}]$ e $C = [c_{ij}]$, então

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

onde usamos a associatividade da adição de números reais.

(iii) Se $A = [a_{ij}]$

$$A + 0 = [a_{ij}] + 0 = [a_{ij}] = A$$

onde usamos a existência do elemento neutro da adição de números reais.

(iv) Se $A = [a_{ij}]$

$$A + (-A) = [a_{ij}] + [-a_{ij}] = 0$$

onde usamos a existência do simétrico de um número real qualquer.

□

Produto por escalar

Dado uma matriz $A = [a_{ij}]_{m \times n}$ chama-se produto da matriz A pelo número real k à matriz $B = [b_{ij}]_{m \times n}$ tal que $b_{ij} = ka_{ij}$ para todo $i \in \{1, 2, 3 \ldots, m\}$ e todo $j \in \{1, 2, 3 \ldots, n\}$.

Por exemplo,

$$5 \cdot \begin{pmatrix} 9 & 1 & -4 \\ -5 & -2 & 3 \\ 6 & 0 & 7 \end{pmatrix} = \begin{pmatrix} 45 & 5 & -20 \\ -25 & -10 & 15 \\ 30 & 0 & 35 \end{pmatrix}$$

O produto de uma matriz por um escalar apresenta as seguintes propriedades, como mostra o resultado a seguir.

Proposição 2. Dadas A e B matrizes quaisquer do tipo $m \times n$ e a e b números reais quaisquer.

(i) $a \cdot (b \cdot A) = (ab) \cdot A$

(ii) $a \cdot (A + B) = a \cdot A + a \cdot B$

(iii) $(a + b) \cdot A = a \cdot A + b \cdot A$

(iv) $1 \cdot A = A$

Demonstração. As propriedades acima decorrem diretamente da definição de produto de matriz por escalar e das propriedades dos números reais.

(i) sejam $A = [a_{ij}]$, $a \in \mathbb{R}$ e $b \in \mathbb{R}$

$$a(b \cdot A) = a(b \cdot [a_{ij}]) = a \cdot b \cdot [a_{ij}] = (ab) \cdot [a_{ij}] = (ab) \cdot A$$

onde usamos a associatividade da multiplicação de números reais.

(ii) sejam $A = [a_{ij}]$, $B = [b_{ij}]$ e $a \in \mathbb{R}$

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

onde usamos a distributividade da multiplicação em relação à adição de números reais.

(iii) sejam $A = [a_{ij}]$, $a \in \mathbb{R}$ e $b \in \mathbb{R}$

$$(a + b) \cdot A = (a + b) \cdot [a_{ij}] = a \cdot [a_{ij}] + b \cdot [a_{ij}] = a \cdot A + b \cdot A$$

onde, novamente, usamos a distributividade da multiplicação em relação à adição de números reais.

(iv) seja $A = [a_{ij}]$

$$1 \cdot A = 1 \cdot [a_{ij}] = [a_{ij}] = A$$

onde usamos o elemento neutro da multiplicação de números reais.

$\square$

Produto de matrizes

Sejam as matrizes $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{n\times p}$, chama-se produto AB a matriz $C = [c_{ij}]_{m\times p}$, tal que

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}$$

para todo $i \in \{1,2,3\cdots,m\}$ e todo $j \in \{1,2,3\cdots,n\}$.

Note que para o produto AB estar definido o número de colunas de A deve ser igual ao número de linhas de B pois A é do tipo $m\times n$ e B é do tipo $n\times p$. Além disso, o produto AB é uma matriz que tem o número de linhas de A e o número de colunas de B, pois $C = (c_{ij})_{m\times p}$.

Por exemplo,

$$\begin{pmatrix} 4 & 3 \\ 1 & 5 \\ 6 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 \\ 5 & 4 \end{pmatrix} = \begin{pmatrix} 4+15 & 12+12 \\ 1+25 & 3+20 \\ 6+10 & 18+8 \end{pmatrix} = \begin{pmatrix} 19 & 24 \\ 26 & 23 \\ 16 & 26 \end{pmatrix}$$

Note que, nesse caso, o produto tomado na ordem inversa BA, não está definido.

Retornando ao sistema de equações lineares (1.1) do início deste capítulo dado por

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

Este sistema pode ser expresso de maneira muito mais compacta utilizando-se a notação sugerida pela álgebra das matrizes. Como já dissemos antes, seja A a matriz de coeficientes do sistema:

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

Também tomamos

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{e} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Ambos, $\mathbf{x}$ e $\mathbf{b}$, são matrizes, chamadas matrizes coluna. A matriz $\mathbf{x}$ do tipo $n\times 1$ contém as variáveis e a matriz $\mathbf{b}$ de tamanho $k\times 1$ contém os parâmetros do lado direito do sistema. Então, o sistema de equações pode ser escrito como

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

ou, simplesmente, como

$$A\mathbf{x} = \mathbf{b}$$

onde $A\mathbf{x}$ se refere ao produto matricial da matriz A do tipo $k \times n$ com a matriz $\mathbf{x}$ do tipo $n \times 1$. Esse produto resulta numa matriz de tamanho $k \times 1$, que deve ser igualada à matriz $\mathbf{b}$ do tipo $k \times 1$.

Definimos a **matriz identidade** ou **matriz unidade** $n \times n$ como sendo $I_n = [a_{ij}]_{n \times n}$ tal que $a_{ii} = 1$ para todo $i = 1, \dots, n$ e $a_{ij} = 0$ se $i \neq j = 1, \dots, n$.

Por exemplo:

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{e} \qquad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

O produto matricial apresenta as seguintes propriedades, como mostra o resultado a seguir.

Proposição 3. Desde que o tamanho das matrizes torne as operações possíveis, temos:

(i) $A(B+C) = AB + AC$ (distributividade à esquerda);

(ii) $(A+B)C = AC + BC$ (distributividade à direita);

(iii) $(AB)C = A(BC)$ (associatividade);

(iv) $(aA)B = A(aB) = a(AB)$ quaisquer que sejam o número a e as matrizes $A = [a_{ij}]_{m \times n}$ e $B = [b_{ij}]_{n \times p}$;

(v) $AI = IA = A$ (existência de elemento identidade).

Demonstração. As propriedades acima decorrem diretamente da definição do produto matricial.

(i) Sejam $A=[a_{ij}]_{m\times n}$, $B=[b_{ij}]_{n\times p}$, $C=[c_{ij}]_{n\times p}$, $AB[e_{ij}]_{m\times p}$, $AC=[f_{ij}]_{m\times p}$ e $A(B+C)=[d_{ij}]_{m\times p}$.

$$\begin{aligned}
[d_{ij}] &= \sum_{k=1}^{n} a_{ik}\cdot(b_{kj}+c_{kj}) \\
&= \sum_{k=1}^{n} (a_{ik}\cdot b_{kj}+a_{ik}\cdot c_{kj}) \\
&= \sum_{k=1}^{n} a_{ik}\cdot b_{kj}+\sum_{k=1}^{n} a_{ik}\cdot c_{kj} = [e_{ij}]+[f_{ij}]
\end{aligned}$$

(ii) Sejam $A=[a_{ij}]_{m\times n}$, $B=[b_{ij}]_{m\times n}$, $C=[c_{ij}]_{n\times p}$, $(A+B)C=[d_{ij}]_{m\times p}$, $AC=[e_{ij}]_{m\times p}$ e $BC=[f_{ij}]_{m\times p}$.

$$\begin{aligned}
[d_{ij}] &= \sum_{k=1}^{n} (a_{ik}+b_{ik})\cdot c_{kj} \\
&= \sum_{k=1}^{n} (a_{ik}\cdot c_{kj}+b_{ik}\cdot c_{kj}) \\
&= \sum_{k=1}^{n} a_{ik}\cdot c_{kj}+\sum_{k=1}^{n} b_{ik}\cdot c_{kj} = [e_{ij}]+[f_{ij}]
\end{aligned}$$

(iii) Sejam $A=[a_{ij}]_{m\times n}$, $B=[b_{ij}]_{n\times p}$, $C=[c_{ij}]_{p\times s}$, $(AB)C=[d_{ij}]_{m\times s}$ e $A(BC)=[e_{ij}]_{m\times s}$.

$$\begin{aligned}
[d_{ij}] &= \sum_{k=1}^{p} (AB)_{ik}\cdot c_{kj} \\
&= \sum_{k=1}^{p} \left(\sum_{l=1}^{n} a_{il}\cdot b_{lk}\right)\cdot c_{kj} \\
&= \sum_{l=1}^{n} a_{il}\cdot\left(\sum_{k=1}^{p} b_{lk}\cdot c_{kj}\right) \\
&= \sum_{l=1}^{n} a_{il}\cdot(BC)_{lj} = [e_{ij}]
\end{aligned}$$

(iv) Sejam $A=[a_{ij}]_{m\times n}$, $B=[b_{ij}]_{n\times p}$, $(\alpha A)B=[d_{ij}]_{m\times p}$, $A(\alpha B)=[e_{ij}]_{m\times p}$, $\alpha(AB)=[f_{ij}]_{m\times p}$ e $\alpha\in\mathbb{R}$.

$$[d_{ij}] = \sum_{k=1}^{n} (\alpha\cdot a_{ik})\cdot b_{kj} = \alpha\cdot\sum_{k=1}^{n} a_{ik}\cdot b_{kj} = [f_{ij}]$$

$$[e_{ij}] = \sum_{k=1}^{n} a_{ik}\cdot(\alpha\cdot b_{kj}) = \alpha\cdot\sum_{k=1}^{n} a_{ik}\cdot b_{kj} = [f_{ij}]$$

então, $(\alpha A)B=A(\alpha B)=\alpha(AB)$

(v) Sejam $A = [a_{ij}]_{m\times n}$ e $I_n = [\delta_{ij}]_{n\times n}$.

a) Se $AI = [d_{ij}]_{m\times n}$.

$$\begin{aligned}[d_{ij}] = \sum_{k=1}^{n} a_{ik}\cdot\delta_{kj} &= a_{i1}\cdot\delta_{1j} + \cdots + a_{ij}\cdot\delta_{jj} + \cdots + a_{in}\cdot\delta_{nj}\\ &= a_{i1}\cdot 0 + \cdots + a_{ij}\cdot 1 + \cdots + a_{in}\cdot 0 = [a_{ij}]\end{aligned}$$

para todo i e j, então $A\cdot I = A$.

b) Se $IA = [d_{ij}]_{m\times n}$.

$$\begin{aligned}[d_{ij}] = \sum_{k=1}^{n} \delta_{ik}\cdot a_{kj} &= \delta_{i1}\cdot a_{1j} + \ldots \delta_{ii}\cdot a_{ij} + \cdots + \delta_{in}\cdot a_{nj}\\ &= 0\cdot a_{1j} + \cdots + 1\cdot a_{ij} + \cdots + 0\cdot a_{nj} = [a_{ij}]\end{aligned}$$

para todo i e j, então $I\cdot A = A$.

Portanto $AI = IA = A$.

□

Matriz transposta

Dada uma matriz $A = [a_{ij}]_{m\times n}$, chama-se *transposta* de A à matriz $A^T = [a'_{ij}]_{n\times m}$ tal que $a'_{ij} = a_{ji}$, para todo i e todo j. Isso significa que a (i, j)-ésima entrada de A torna-se a (j, i)-ésima entrada de A^T.

Por exemplo,

$$A = \begin{pmatrix} a & b\\ c & d\\ e & f\end{pmatrix} \Longrightarrow A^T = \begin{pmatrix} a & c & e\\ b & d & f\end{pmatrix}$$

A matriz transposta apresenta as seguintes propriedades, como mostra o resultado a seguir.

Proposição 4. Desde que as operações sejam possíveis, temos:

(i) $(A+B)^T = A^T + B^T$;

(ii) $(A-B)^T = A^T - B^T$;

(iii) $(A^T)^T = A$;

(iv) $(rA)^T = rA^T$;

(v) $(AB)^T = B^T A^T$.

Demonstração. As propriedades acima decorrem diretamente da definição de matriz transposta.

(i) Sejam $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{m\times n}$

$$(A+B)^T = [(a_{ij}+b_{ij})'] = [a_{ji}+b_{ji}] = [a'_{ij}] + [b'_{ij}] = A^T + B^T$$

(ii) Sejam $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{m\times n}$

$$(A-B)^T = [(a_{ij}-b_{ij})'] = [a_{ji}-b_{ji}] = [a'_{ij}] - [b'_{ij}] = A^T - B^T$$

(iii) Seja $A = [a_{ij}]_{m\times n}$

$$(A^T)^T = [(a'_{ij})'] = [(a_{ji})'] = [a_{ij}] = A$$

(iv) Seja $A = [a_{ij}]_{m\times n}$ e $r \in \mathbb{R}$

$$(rA)^T = [r\cdot a_{ij}]' = r[a_{ji}] = r[a_{ij}]' = rA^T$$

(v) Sejam $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{n\times p}$ e $C = AB = [c_{ij}]_{m\times p}$

$$\begin{aligned} (AB)^T &= [c'_{ij}] = [c_{ji}] = \left[\sum_{k=1}^{n} a_{jk}\cdot b_{ki}\right] \\ &= \sum_{k=1}^{n} [b'_{ik}]\cdot[a'_{kj}] = B^T A^T. \end{aligned}$$

□

1.3 Tipos especiais de matrizes

Nesta seção descreveremos algumas importantes classes de matrizes $k\times n$ por apresentarem uma utilidade maior nesta teoria.

Matriz linha

É toda matriz do tipo $1\times n$. Por exemplo,

$$\begin{pmatrix} a & b & c & d \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 5 & -3 & 0 \end{pmatrix}$$

Matriz coluna

É toda matriz do tipo $k \times 1$. Por exemplo,

$$\begin{pmatrix} 9 \\ 7 \\ 1 \\ 4 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Matriz nula

É toda matriz que tem todos os elementos iguais a zero. Por exemplo,

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Matriz quadrada

É toda matriz que tem o mesmo número de linhas e de colunas, ou seja $k = n$. Por exemplo,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 6 & -3 & 1 \\ -9 & 7 & 0 \\ -5 & 4 & 8 \end{pmatrix}$$

Em uma matriz quadrada de ordem n, chamamos de diagonal principal o conjunto dos elementos a_{ij} em que $i = j$ e chamamos de diagonal secundária o conjunto dos elementos a_{ij} em que $i + j = n + 1$.

Matriz diagonal

É toda matriz em que $k = n$ e $a_{ij} = 0$ para $i \neq j$, ou seja, uma matriz quadrada na qual cada entrada fora da diagonal principal é 0. Por exemplo,

$$\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Matriz triangular superior

É a matriz (geralmente quadrada) na qual cada entrada abaixo da diagonal principal é 0, ou seja, a matriz em que $a_{ij} = 0$ se $i > j$. Por exemplo,

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \qquad \text{e} \qquad \begin{pmatrix} -3 & 6 & 7 \\ 0 & 8 & 1 \\ 0 & 0 & 5 \end{pmatrix}$$

Matriz triangular inferior

É a matriz (geralmente quadrada) na qual cada entrada acima da diagonal principal é 0, ou seja, a matriz em que $a_{ij} = 0$ se $i < j$. Por exemplo,

$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \qquad \text{e} \qquad \begin{pmatrix} 5 & 0 & 0 \\ 4 & 1 & 0 \\ 3 & -4 & 7 \end{pmatrix}$$

Matriz simétrica

É toda matriz com $k = n$ em que $A^T = A$, ou seja, uma matriz quadrada onde $a_{ij} = a_{ji}$ para quaisquer i, j. Por exemplo,

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix} \qquad \text{e} \qquad \begin{pmatrix} 5 & 4 & 3 \\ 4 & 1 & 0 \\ 3 & 0 & 7 \end{pmatrix}$$

Matriz idempotente

É a matriz quadrada B tal que $B \cdot B = B$. Por exemplo,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{e} \qquad \begin{pmatrix} 3 & -3 \\ 2 & -2 \end{pmatrix}$$

Matriz de permutação

É uma matriz quadrada de entradas 0 e 1, na qual cada linha e cada coluna contêm exatamente um 1. Por exemplo,

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

Daremos no momento atenção especial às matrizes triangulares superiores e as matrizes de permutação entre outras.

Primeiramente definiremos as matrizes na forma escalonada por linhas e escalonada reduzida por linhas.

1.4 Forma escalonada de uma matriz

Definição 1. Seja k_i o número de zeros que antecedem o primeiro elemento não nulo da linha i, de uma matriz A. Definimos que A está na forma **escalonada por linhas** se $k_i < k_{i+1},\ i = 1,2,\ldots$ Neste caso o primeiro elemento não nulo da linha i é dito pivô.

Definição 2. Seja $A = [a_{ij}]$ matriz escalonada. Denominaremos por pivô o elemento a_{ij} que corresponde ao primeiro elemento não nulo da linha e neste caso o denotamos por k_{ij}. Assim, se $k_{i,j}$ e $k_{i+1,l}$ são pivôs de A então $j < l$.

Por exemplo, as matrizes

$$A = \begin{pmatrix} 7 & 4 & -2 \\ 0 & 5 & 8 \\ 0 & 0 & -9 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} -6 & 4 & 3 & 0 \\ 0 & 5 & -1 & 8 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

estão na forma escalonada por linhas. Na matriz A, $k_{11} = 7$, $k_{22} = 5$, $k_{33} = -9$. Já na matriz B $k_{11} = -6$, $k_{22} = 5$ e $k_{34} = 1$.

Definição 3. Uma matriz está na forma **escalonada reduzida por linhas** se:

1. Estiver na forma escalonada por linhas;
2. O pivô de cada linha não nula é 1;
3. Cada coluna que contém um pivô tem todos os seus outros elementos iguais a zero;

Por exemplo, as matrizes

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 1 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

estão na forma escalonada reduzida por linhas. Já as matrizes seguintes:

$$\begin{pmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

não estão na forma escalonada reduzida por linhas. A primeira não satisfaz a condição 3 e a segunda não satisfaz as condições 1.

Definição 4. Seja A uma matriz na forma escalonada. O posto de A é o número de linhas não nulas de A.

Observação 1. Observe que se acrescentarmos uma coluna a uma matriz A, obtendo a matriz aumentada $\hat{A}$, então posto $A \leq$ posto $\hat{A}$.

Exemplo 1. Sejam as matrizes

$$A = \begin{pmatrix} 7 & 4 & -2 \\ 0 & 5 & 8 \\ 0 & 0 & -9 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{e} \quad \hat{A} = \begin{pmatrix} 7 & 4 & -2 & | & b_1 \\ 0 & 5 & 8 & | & b_2 \\ 0 & 0 & -9 & | & b_3 \\ 0 & 0 & 0 & | & b_4 \end{pmatrix}$$

Se $b_4 = 0 \Rightarrow$ posto $A =$ posto $\hat{A}$, se $b_4 \neq 0 \Rightarrow$ posto $A <$ posto $\hat{A}$

Veremos agora como determinar se um sistema linear tem solução ou não, a partir do posto das matrizes a ele associadas.

Lembramos que para um sistema geral da forma

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

Tem-se

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{e} \quad \hat{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & | & b_2 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & | & b_m \end{pmatrix}$$

respectivamente matriz e matriz aumentada do sistema.

Para o caso específico onde A e $\hat{A}$ sejam da forma escalonada tem-se

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{mn} \end{pmatrix} \quad \text{e} \quad \hat{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & | & b_1 \\ 0 & a_{22} & a_{23} & \dots & a_{2n} & | & b_2 \\ 0 & 0 & a_{33} & \dots & a_{3n} & | & b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & 0 & \dots & a_{mn} & | & b_m \end{pmatrix}$$

e por consequência o sistema é da forma

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n & = & b_1 \\ a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n & = & b_2 \\ a_{33}x_3 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & \ddots & \vdots \\ a_{mn}x_n & = & b_m \end{array}$$

Chamaremos de p o posto de A, q o posto de $\hat{A}$ e $k_{i,j}$ o pivô de cada linha i.

Vejamos os exemplos abaixo:

Exemplo 2. Vamos determinar a solução do sistema de três equações e três variáveis

$$\begin{array}{rcl} 2x_1 + 3x_2 - x_3 & = & -3 \\ 2x_2 + 3x_3 & = & 10 \\ 2x_3 & = & 8 \end{array}$$

cuja matriz de coeficientes e matriz aumentada, associadas a ele são, respectivamente

$$\begin{pmatrix} 2 & 3 & -1 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix} \text{ e } \begin{pmatrix} 2 & 3 & -1 & | & -3 \\ 0 & 2 & 3 & | & 10 \\ 0 & 0 & 2 & | & 8 \end{pmatrix}$$

Como as matrizes já estão na forma escalonada podemos observar que **o posto da matriz de coeficientes é igual ao posto da matriz aumentada e ambos iguais ao número de colunas da matriz de coeficientes**. Para encontrarmos a solução desse sistema podemos utilizar a substituição inversa.

Partindo da última equação, encontramos $x_3 = 4$. Substituindo o valor de x_3 na penúltima equação encontramos $x_2 = -1$. Por fim, substituímos $x_3 = 4$ e $x_2 = -1$ na primeira equação encontrando $x_1 = 2$. Portanto a solução do sistema é única e igual a terna $(2, -1, 4)$.

Exemplo 3. Vamos determinar agora a solução do sistema de duas equações e três variáveis

$$\begin{array}{rcl} 2x_1 + 3x_2 - x_3 & = & -3 \\ 2x_2 + 3x_3 & = & 10 \end{array}$$

cuja matriz de coeficientes e matriz aumentada, associadas a ele são, respectivamente

$$\begin{pmatrix} 2 & 3 & -1 \\ 0 & 2 & 3 \end{pmatrix} \text{ e } \begin{pmatrix} 2 & 3 & -1 & | & -3 \\ 0 & 2 & 3 & | & 10 \end{pmatrix}$$

As matrizes já estão na forma escalonada, logo podemos observar que ambas possuem posto 2. Como no exemplo anterior, vamos usar a substituição inversa partindo da última equação e nela encontramos x_2, em função de x_3, então

$$2x_2 + 3x_3 = 10 \Rightarrow 2x_2 = 10 - 3x_3 \Rightarrow x_2 = \frac{10 - 3x_3}{2} = 5 - \frac{3x_3}{2}.$$

Substituindo x_2 na primeira equação encontramos x_1 em função de x_3

$$\begin{aligned} 2x_1 + 3x_2 - x_3 = -3 \Rightarrow 2x_1 + 3 \cdot \left(5 - \frac{3x_3}{2}\right) - x_3 = -3 \Rightarrow \\ 2x_1 + 15 - \frac{9x_3}{2} - x_3 = -3 \Rightarrow x_1 = \frac{11}{4}x_3 - 9 \end{aligned}$$

Portanto a solução do sistema é $\left(\frac{11}{4}x_3 - 9,\ \frac{10 - 3x_3}{2},\ x_3\right)$. Note que a cada valor de x_3 que escolhemos, temos uma solução particular do sistema e, portanto, o sistema possui infinitas soluções.

Exemplo 4. Trataremos agora da solução do sistema

$$\begin{aligned} 2x_1 + 3x_2 - x_3 &= -3 \\ 2x_2 + 3x_3 &= 10 \\ 0 &= 8 \end{aligned}$$

cuja matriz de coeficientes e matriz aumentada, associadas a ele são, respectivamente

$$\begin{pmatrix} 2 & 3 & -1 \\ 0 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix} \text{ e } \begin{pmatrix} 2 & 3 & -1 & | & -3 \\ 0 & 2 & 3 & | & 10 \\ 0 & 0 & 0 & | & 8 \end{pmatrix}$$

Observe que **o posto da matriz de coeficiente é menor que o posto da matriz aumentada e portanto o sistema não possui solução**, pois partindo da última equação $0 = 8$ nos deparamos com uma situação impossível.

Concluindo, inspirados nestes exemplos, se $p =$ posto de A, $q =$ posto de $\hat{A}$ e n é o número de incógnitas ou número de colunas de A temos três casos a considerar:

1. $p = q = n$;
2. $p = q < n$;
3. $p \neq q$, isto é, $p < q$.

Caso 1. Se $p = q = n$, a matriz aumentada será do tipo abaixo onde $k_{ii} \neq 0$ para $i = 1, \ldots, n$.

$$\hat{A} = \begin{pmatrix} k_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} & | & b_1 \\ 0 & k_{22} & a_{23} & a_{24} & \cdots & a_{2n} & | & b_2 \\ 0 & 0 & k_{33} & a_{34} & \cdots & a_{3n} & | & b_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & 0 & 0 & \cdots & k_{nn} & | & b_n \\ 0 & 0 & 0 & 0 & \cdots & 0 & | & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & | & 0 \end{pmatrix}$$

Partindo da última equação não nula do sistema associado obtemos $k_{nn}x_n = b_n$ o que nos dá x_n. Em seguida substituímos x_n na penúltima equação obtendo $k_{n-1,n-1}x_{n-1} = b_{n-1} - a_{n-1,n}x_n$ o que nos dá o valor de x_{n-1} e assim sucessivamente.
Concluímos que o sistema associado a $\hat{A}$ tem solução única $(x_1, x_2, \cdots, x_n)$.

Caso 2. Se $p = q < n$, a matriz aumentada associada ao sistema será da forma

$$\hat{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} & \cdots & a_{1j} & \cdots & a_{1n} & | & b_1 \\ 0 & a_{22} & \cdots & a_{2p} & \cdots & a_{2j} & \cdots & a_{2n} & | & b_2 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & k_{pj} & \cdots & a_{pn} & | & b_p \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & | & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & | & 0 \end{pmatrix}$$

Logo a última linha não nula de $\hat{A}$ corresponde a equação

$$k_{pj}x_j + a_{p+1,j}x_{j+1} + \cdots + a_{pn}x_n = b_p$$

De modo geral a equação i, $i = 1, 2, \cdots, p$, do sistema se escreve como

$$k_{i,j_i}x_{j_i} + a_{i,j_{i+1}}x_{j_{i+1}} + \cdots + a_{i,j_p}x_p = b_i$$

Como $k_{i,j_i} \neq 0$ temos

$$x_{j_i} = b_i - \frac{\sum_{j \neq j_i} a_{i,j}x_j}{k_{i,j_i}}$$

Note que as colunas com pivôs correspondem às variáveis que se escrevem em função das demais, e são denominadas variáveis dependentes, enquanto as demais são denominadas livres. Assim dando valores arbitrários para as variáveis livres x_j, $j \neq j_i$, $i = 1, 2, \cdots p$, obtemos um "novo" sistema de p equações e p incógnitas cujo posto da matriz é p bem como o posto da matriz aumentada. Assim pelo **Caso 1**, para cada conjunto de valores atribuídos às variáveis livres obtemos uma solução para o sistema. Portanto o sistema inicial possui infinitas soluções.

Caso 3. Se $p < q$, significa que a matriz ampliada tem a q-ésima linha do tipo

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & | & b_q \end{pmatrix}$$

que corresponde à equação

$$0x_1 + 0x_2 + \cdots + 0x_n = b_q$$

que não possui solução, já que $b_q \neq 0$ portanto o sistema não possui solução.
Observe que $p \neq q$ só ocorre com $q \geq p + 1$.

Acabamos de demonstrar:

Teorema 1. (Existência de Solução para Sistemas Escalonados) Seja um sistema linear de m equações a n incógnitas na forma $A\mathbf{x} = \mathbf{b}$, cuja matriz de coeficientes $A_{(m\times n)}$ está na forma escalonada por linhas, tem posto p e a respectiva matriz aumentada $\hat{A}_{(m\times(n+1))}$ tem posto q.

Então

1. Se $p = q = n$ o sistema têm solução única;
2. Se $p = q < n$ o sistema têm infinitas soluções;
3. Se $p < q$ o sistema não tem solução.

Logo concluímos que sempre que um sistema se apresente na forma escalonada respondemos facilmente as questões:

1. O sistema tem solução?
2. Se tiver, quantas soluções tem?

Mas nem todo sistema se apresenta nesta forma. Porém veremos adiante que todo sistema tem forma escalonada equivalente.

Para isso vamos estudar inicialmente as **matrizes elementares**.

1.5 Matrizes elementares

Definimos três operações elementares que aplicamos sobre as linhas da matriz identidade de ordem n. São elas:

1. Permutação de linhas;
2. Multiplicação de uma linha por um escalar não-nulo;

3. Soma de um múltiplo de uma linha a uma outra linha.

Denominamos de matriz elementar a toda matriz que for obtida a partir da matriz identidade, após sofrer uma ou mais operações elementares. No caso específico da matriz identidade sofrer apenas uma operação elementar ela terá uma notação específica. Vejamos cada caso.

Quando aplicamos a operação (1) a uma matriz identidade, trocando-se a linha i com a linha j, obtemos a matriz elementar E_{ij}. Por exemplo,

$$E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \text{onde as linhas 2 e 3 foram permutadas,}$$

Note que os índices 2 e 3 fazem referência às linhas que foram permutadas.

Ao aplicar a operação (2) a uma matriz identidade, isto é, multiplicando-se a linha i por um número r qualquer obtemos a matriz elementar $E_i(r)$. Por exemplo,

$$E_2(5) = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix} \qquad \text{onde a linha 2 foi multiplicada por 5}$$

Por fim, ao aplicar a operação (3) a uma matriz identidade, isto é trocamos a linha j por r vezes a linha i somada a linha j, obtemos a matriz elementar $E_{ij}(r)$. Por exemplo,

$$E_{34}(2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix} \qquad \text{onde a linha 4 foi acrescida do dobro da linha 3}$$

Enunciamos assim o resultado a seguir.

Teorema 2. Seja E uma matriz elementar $N \times N$. Se A é uma matriz $N \times M$ qualquer, então EA é a matriz obtida executando-se a mesma operação elementar que E representa sobre A.

Demonstração. Os três tipos de matrizes elementares são E_{ij}, $E_i(r)$ e $E_{ij}(r)$.

(i) (Caso E_{ij}) Para verificar isso, denotemos por e_{hk} uma entrada qualquer de E_{ij}. Observe que as linhas n, com $n \neq i, j$ não sofreram alteração. Assim

$$\begin{aligned} e_{ij} &= e_{ji} = 1; \\ e_{ii} &= e_{jj} = 0; \\ e_{hh} &= 1 \qquad \text{se} \qquad h \neq i, j; \\ e_{hk} &= 0 \qquad \text{caso contrário.} \end{aligned}$$

O elemento na linha k e coluna n de $E_{ij}A$ é

$$\sum_{m=1}^{n} e_{km}a_{mn} = \begin{cases} a_{jn} & k = i \\ a_{in} & k = j \\ a_{kn} & k \neq i, j \end{cases}$$

por (1). Portanto, $E_{ij}A$ é simplesmente A com as linhas i e j permutadas.

(ii) (Caso $E_i(r)$) Para verificar isso, denotemos por e_{hj} uma entrada qualquer de $E_i(r)$:

$$\begin{aligned} e_{hj} &= 0 \quad \text{se} \quad h \neq j; \\ e_{hh} &= 1 \quad \text{se} \quad h \neq i; \\ e_{ii} &= r. \end{aligned}$$

O elemento na linha k e coluna m de $E_i(r)A$ é

$$\sum_{j=1}^{n} e_{kj}a_{jm} = \begin{cases} a_{km} & k \neq i \\ ra_{km} & k = i \end{cases}$$

Assim $E_i(r)A$ é A com sua linha i multiplicada por r.

(iii) (Caso $E_{ij}(r)$) Para verificar isso, denotemos por e_{jh} uma entrada qualquer de $E_{ij}(r)$:

$$\begin{aligned} e_{hh} &= 1 \quad \text{para todo} \quad h; \\ e_{ji} &= r; \\ e_{hk} &= 0 \quad \text{para} \quad h \neq k \quad \text{e} \quad (h,k) \neq (j,i). \end{aligned}$$

O elemento na linha k e coluna m de $E_i(r)A$ é

$$\sum_{l=1}^{n} e_{kl}a_{lm} = \begin{cases} a_{km} & k \neq j \\ e_{jj}a_{jm} + e_{ji}a_{im} = a_{jm} + ra_{im} & k = j \end{cases}$$

Logo a linha j de $E_{ij}(r)A$ é (linha j de A)+r(linha i de A).

□

Vejamos os exemplos a seguir.

Exemplo 5. Seja A a matriz $\begin{pmatrix} 2 & 5 & -4 \\ 7 & 1 & 3 \\ -2 & -8 & 6 \end{pmatrix}$ e a matriz elementar $E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, então

$$E_{12} \cdot A = \begin{pmatrix} 7 & 1 & 3 \\ 2 & 5 & -4 \\ -2 & -8 & 6 \end{pmatrix}$$

Ou seja, $E_{12} \cdot A$ é a matriz A com as linhas 1 e 2 permutadas.

Exemplo 6. Seja A a matriz $\begin{pmatrix} 2 & 5 & -4 \\ 7 & 1 & 3 \\ -2 & -8 & 6 \end{pmatrix}$ e a matriz elementar $E_1(2) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, então

$$E_1(2) \cdot A = \begin{pmatrix} 4 & 10 & -8 \\ 7 & 1 & 3 \\ -2 & -8 & 6 \end{pmatrix}$$

Ou seja, $E_1(2) \cdot A$ é a matriz A com a linha 1 multiplicada por 2.

Exemplo 7. Seja A a matriz $\begin{pmatrix} 2 & 5 & -4 \\ 7 & 1 & 3 \\ -2 & -8 & 6 \end{pmatrix}$ e a matriz elementar $E_{12}(3) = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, então

$$E_{12}(3) \cdot A = \begin{pmatrix} 2 & 5 & -4 \\ 13 & 11 & -9 \\ -2 & -8 & 6 \end{pmatrix}$$

Ou seja, $E_{12}(3) \cdot A$ é a matriz A com a linha 2 somada com o produto da linha 1 por 3

O resultado que apresentaremos a seguir garantir-nos-á que toda equação da forma $A\mathbf{x} = \mathbf{b}$ tem uma correspondente equação da forma $U\mathbf{x} = \mathbf{c}$ onde U e a correspondente matriz aumentada $\hat{U}$ estão na forma escalonada por linhas.

Antes de mostrar tal fato, ilustremos com um exemplo.

Seja a equação matricial $A\mathbf{x} = \mathbf{b}$ dada por

$$\begin{pmatrix} 0 & 2 & 1 \\ 1 & 3 & -1 \\ 3 & -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \\ -5 \end{pmatrix}$$

onde

$$A = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 3 & -1 \\ 3 & -1 & -2 \end{pmatrix} \quad \text{e} \quad \hat{A} = \begin{pmatrix} 0 & 2 & 1 & | & 7 \\ 1 & 3 & -1 & | & 4 \\ 3 & -1 & -2 & | & -5 \end{pmatrix}$$

Vamos usar matrizes elementares para transformar a matriz $\hat{A}$ na forma escalonada $\hat{U}$ seguindo os seguintes passos.

(**Passo 1)** Como a primeira entrada da primeira linha é zero temos que permutar a primeira linha com outra na qual a primeira entrada é diferente de zero, digamos a segunda linha. Para tanto calculamos o produto $E_{12} \cdot \hat{A}$.

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 & 1 & | & 7 \\ 1 & 3 & -1 & | & 4 \\ 3 & -1 & -2 & | & -5 \end{pmatrix} = \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 2 & 1 & | & 7 \\ 3 & -1 & -2 & | & -5 \end{pmatrix}$$

(Passo 2) Queremos que o único elemento não nulo da primeira coluna seja o a_{11}, para isso multiplicamos $E_{13}(-3)$ pela matriz obtida no passo anterior.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 2 & 1 & | & 7 \\ 3 & -1 & -2 & | & -5 \end{pmatrix} = \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 2 & 1 & | & 7 \\ 0 & -10 & 1 & | & -17 \end{pmatrix}$$

(Passo 3) A próxima etapa será zerar o terceiro elemento da segunda coluna, para isso multiplicamos $E_{23}(5)$ pela matriz obtida no passo anterior.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 2 & 1 & | & 7 \\ 0 & -10 & 1 & | & -17 \end{pmatrix} = \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 2 & 1 & | & 7 \\ 0 & 0 & 6 & | & 18 \end{pmatrix}$$

Com o último resultado podemos escrever o novo sistema matricial escalonado $U\mathbf{x} = \mathbf{c}$

$$\begin{pmatrix} 1 & 3 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 6 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 18 \end{pmatrix}$$

correspondente ao sistema $A\mathbf{x} = \mathbf{b}$.

Observação 2. Caso se queira obter a forma escalonada reduzida por linhas $\bar{U}$ da matriz $\hat{A}$ basta prosseguir com os seguintes passos:

(Passo 4) Cada pivô têm que ser igual a 1 e como o pivô da primeira linha já o é, faltam apenas controlar os pivôs da segunda e da terceira linha. Para isso vamos multiplicar a matriz obtida no passo anterior pelas matrizes elementares $E_2\left(\frac{1}{2}\right)$ e $E_3\left(\frac{1}{6}\right)$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 2 & 1 & | & 7 \\ 0 & 0 & 6 & | & 18 \end{pmatrix} = \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 1 & \frac{1}{2} & | & \frac{7}{2} \\ 0 & 0 & 6 & | & 18 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{6} \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 1 & \frac{1}{2} & | & \frac{7}{2} \\ 0 & 0 & 6 & | & 18 \end{pmatrix} = \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 1 & \frac{1}{2} & | & \frac{7}{2} \\ 0 & 0 & 1 & | & 3 \end{pmatrix}$$

(Passo 5) Cada coluna que contém um pivô deve ter todos os seus outros elementos iguais a zero, isso não é observado nas colunas 2 e 3, portanto devemos multiplicar a matriz obtida no passo anterior inicialmente pela matriz elementar $E_{21}(-3)$.

$$\begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & -1 & | & 4 \\ 0 & 1 & \frac{1}{2} & | & \frac{7}{2} \\ 0 & 0 & 1 & | & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\frac{5}{2} & | & -\frac{13}{2} \\ 0 & 1 & \frac{1}{2} & | & \frac{7}{2} \\ 0 & 0 & 1 & | & 3 \end{pmatrix}$$

Agora multipliquemos a matriz obtida acima pelas matrizes elementares $E_{32}\left(-\frac{1}{2}\right)$ e $E_{31}\left(\frac{5}{2}\right)$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -\frac{5}{2} & | & -\frac{13}{2} \\ 0 & 1 & \frac{1}{2} & | & \frac{7}{2} \\ 0 & 0 & 1 & | & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\frac{5}{2} & | & -\frac{13}{2} \\ 0 & 1 & 0 & | & 2 \\ 0 & 0 & 1 & | & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \frac{5}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -\frac{5}{2} & | & -\frac{13}{2} \\ 0 & 1 & 0 & | & 2 \\ 0 & 0 & 1 & | & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & | & 1 \\ 0 & 1 & 0 & | & 2 \\ 0 & 0 & 1 & | & 3 \end{pmatrix}$$

De forma geral dado um sistema $A\mathbf{x} = \mathbf{b}$ com matriz de coeficientes A e matriz aumentada $\hat{A}$, podemos obter a matriz escalonada $\hat{U}$ correspondente à matriz $\hat{A}$ aplicando os seguintes passos:

(**Passo 1)** Seja k_1 a primeira coluna da matriz $\hat{A}$ com algum elemento não nulo. Troque as linhas entre si de modo que esse elemento não nulo apareça na primeira linha, isto é, de modo que na nova matriz $a_{1k_1} \neq 0$. Para tanto calculamos o produto $E_{1k_1} \cdot \hat{A}$.

(**Passo 2)** Anulamos todas as entradas da coluna k_1 (com exceção da primeira linha), calculando-se os produtos $E_{1j}(-a_{ik_1})$.

(**Passo 3)** Deixamos de lado a primeira equação e aplicamos os dois primeiros passos nas linhas restantes.

(**Passo 4)** Deixamos de lado a primeira e segunda equações e aplicamos os dois primeiros passos nas linhas restantes, e assim por diante, até a matriz ficar na forma escalonada.
Para alcançarmos a forma escalonada reduzida por linhas $\bar{U}$ da matriz $\hat{A}$ prosseguimos com os seguintes passos:

(**Passo 5)** Seja a_{ik_i} o pivô da linha i e multiplicamos a matriz obtida nos passos anteriores pelas matrizes elementares $E_i\left(\frac{1}{a_{ik_i}}\right)$ até atingir a última linha que contém pivô.

(**Passo 6)** Para terminar multiplicamos a matriz obtida no passo anterior pelas matrizes elementares $E_{il}(-a_{lk_i})$, $l = 1, 2, \cdots, i-1$, para $i = 2$. Depois para $i = 3$, e assim por diante até a última linha que contém pivô.

Com isso podemos enunciar o seguinte resultado.

Teorema 3. Dada qualquer matriz $A_{m\times n} \quad \exists \quad E_1, E_2, \cdots, E_k$ matrizes elementares tais que

$$E_k \cdot E_{k-1} \cdots E_1 \cdot A = U$$

onde U é a forma escalonada de A.

Observe que este mesmo resultado pode ser obtido em relação a matriz escalonada reduzida por linhas $\bar{U}$ de A, uma vez que para obtê-la continuamos os passos dados anteriormente. De modo que vale o seguinte resultado:

Corolário 1. Dada qualquer matriz $A_{m\times n} \quad \exists \quad E_1, E_2, \cdots, E_l$ matrizes elementares tais que

$$E_l \cdots E_{k+1} \cdot E_k \cdot E_{k-1} \cdots E_1 \cdot A = \bar{U}$$

onde $\bar{U}$ é a matriz escalonada reduzida por linhas de A.

A existência de solução do sistema $U\mathbf{x} = \mathbf{c}$ está determinada pelo Teorema 1. A pergunta que precisamos fazer agora é: soluções de $U\mathbf{x} = \mathbf{c}$ e $A\mathbf{x} = \mathbf{b}$ são as mesmas?

Antes de responder esta pergunta precisamos do conceito de matrizes invertíveis.

1.6 Matrizes invertíveis

De agora em diante representaremos o conjunto das matrizes quadradas reais de ordem n por $\mathbf{M_n}$

Definição 5. Seja A uma matriz em $\mathbf{M_n}$. Uma matriz B em $\mathbf{M_n}$ é uma **inversa** para A se $AB = BA = I$.

Se existir a matriz B, dizemos que A é **invertível** e representaremos a matriz B por A^{-1}. Essa definição deixa aberta a possibilidade de uma matriz A ter várias inversas. Isso não é verdade, como veremos no teorema a seguir.

Teorema 4. **(Unicidade da matriz inversa)** Uma matriz A em $\mathbf{M_n}$ pode ter, no máximo, uma única inversa.

Demonstração. Suponha que B e C sejam, inversas de A. Então,

$$C = CI = C(AB) = (CA)B = IB = B$$

□

Definição 6. Seja A uma matriz de tamanho $k \times n$. A matriz B de tamanho $n \times k$ é uma **inversa à direita** de A se $AB = I$. A matriz C de tamanho $n \times k$ é uma **inversa à esquerda** de A se $CA = I$.

Lema 1. Se uma matriz A tem uma inversa à direita B e uma inversa à esquerda C. Então A é invertível e $B = C = A^{-1}$.

Demonstração. Exatamente a mesma do Teorema 4 □

Vejamos alguns exemplos:

Exemplo 8. Suponha $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_n \neq 0$ e sejam as matrizes diagonais

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad \text{e} \quad A^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1} & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n^{-1} \end{pmatrix}$$

Então

$$A \cdot A^{-1} = A^{-1} \cdot A = I$$

Exemplo 9. Seja a matriz elementar E_{ij} de ordem n, vimos que E_{ij} é a matriz identidade após sofrer troca da linha i com a linha j. Vimos também que se A é uma matriz arbitrária, então $E_{ij}A$ é a matriz obtida à partir de A trocando-se a linha i com a linha j. Logo $E_{ij}E_{ij}$ é a matriz E_{ij} após sofrer troca da linha i com a linha j, retornando, portanto a matriz identidade. Logo concluímos que E_{ij} é invertível e $E_{ij}^{-1} = E_{ij}$, isto é

$$E_{ij} \cdot E_{ij} = I$$

Exemplo 10. Seja a matriz elementar $E_i(r)$ de ordem n, vimos que se A é uma matriz arbitrária, então $E_i(r)A$ é a matriz obtida à partir de A multiplicando-se a linha i por r. Logo inferimos que $E_i(r)^{-1} = E_i(\frac{1}{r})$ onde $E_i(r)^{-1}$ desfaz o que $E_i(r)$ faz. Logo

$$E_i\left(\frac{1}{r}\right) \cdot E_i(r) = I = E_i(r) \cdot E_i\left(\frac{1}{r}\right)$$

Exemplo 11. Seja a matriz elementar $E_{ij}(r)$ de ordem n, vimos que se A é uma matriz arbitrária, então $E_{ij}(r)A$ é a matriz obtida à partir de A trocando-se a linha j com a linha j adicionada a r vezes a linha i. Logo inferimos que $E_{ij}(r)^{-1} = E_{ij}(-r)$ onde $E_{ij}(r)^{-1}$ desfaz o que $E_{ij}(r)$ faz. De fato

$$E_{ij}(-r) \cdot E_{ij}(r) = I = E_{ij}(r) \cdot E_{ij}(-r)$$

Os exemplos acima mostram que:

Proposição 5. Todas as matrizes elementares são invertíveis e

(a) $E_{ij}^{-1} = E_{ji}$

(b) $E_i(r)^{-1} = E_i\left(\frac{1}{r}\right)$

(c) $E_{ij}(r)^{-1} = E_{ij}(-r)$

para cada $r \in \mathbb{R}$

Proposição 6. Se A e B são invertíveis, o produto AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$.

Demonstração. Vamos mostrar apenas que o produto $(AB) \cdot (B^{-1}A^{-1}) = I$ pois o produto $(B^{-1}A^{-1}) \cdot (AB) = I$ é análogo.

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

□

Corolário 2. Seja $E = E_1 \cdot E_2 \cdots E_k$ onde E_i é matriz elementar, $i = 1, 2, \cdots, k$. Então E é invertível.

Teorema 5. Seja A uma matriz quadrada invertível. Então,

(i) $\left(A^{-1}\right)^{-1} = A$,

(ii) $\left(A^{T}\right)^{-1} = \left(A^{-1}\right)^{T}$

Demonstração. (i) A é invertível, logo $A^{-1}A = AA^{-1} = I$. Então A^{-1} é invertível. Como a inversa é única

$$A = (A^{-1})^{-1}.$$

(ii) Temos que $I = I^T = (AA^{-1})^T = (A^{-1})^T \cdot A^T$. Mas também temos $I = I^T = (A^{-1}A)^T = A^T(A^{-1})^T$. Então $(A^T)^{-1} = (A^{-1})^T$.

□

Definição 7. Seja A matriz quadrada. Definimos para m inteiro positivo $A^m = \underbrace{A\cdots A}_{m \text{ vezes}}$ e $A^0 = I$

Definição 8. Seja A matriz invertível. Então definimos $A^{-m} = (A^{-1})^m$ para $m = 1, 2, 3$

Teorema 6. Se A é invertível:

(i) A^m é invertível para qualquer inteiro $m \neq 0$ e $(A^m)^{-1} = (A^{-1})^m = A^{-m}$,

(ii) para quaisquer inteiros r e s, $A^r A^s = A^{r+s}$, e

(iii) para qualquer escalar $r \neq 0$, rA é invertível e $(rA)^{-1} = \frac{1}{r}A$

Demonstração. (i) Para $m > 0$ temos

$$\begin{aligned}
A^m \cdot (A^{-1})^m &= \underbrace{(A\cdot A\cdots A)}_{m \text{ vezes}}\underbrace{(A^{-1}\cdot A^{-1}\cdots A^{-1})}_{m \text{ vezes}} \\
&= \underbrace{(A\cdot A\cdots A)}_{(m-1) \text{ vezes}}(AA^{-1})\underbrace{(A^{-1}\cdot A^{-1}\cdots A^{-1})}_{(m-1) \text{ vezes}} \\
&= \underbrace{(A\cdot A\cdots A)}_{(m-1) \text{ vezes}}\underbrace{(A^{-1}\cdot A^{-1}\cdots A^{-1})}_{(m-1) \text{ vezes}} \\
&= \underbrace{(A\cdot A\cdots A)}_{(m-2) \text{ vezes}}(AA^{-1})\underbrace{(A^{-1}\cdot A^{-1}\cdots A^{-1})}_{(m-2) \text{ vezes}} \\
&= \underbrace{(A\cdot A\cdots A)}_{(m-2) \text{ vezes}}\underbrace{(A^{-1}\cdot A^{-1}\cdots A^{-1})}_{(m-2) \text{ vezes}} \\
&= \cdots = AA^{-1} = I
\end{aligned}$$

Analogamente obtemos $(A^{-1})^m \cdot A^m = I$.

Assim, para $m > 0$ temos que $A^{-m} = (A^m)^{-1}$. Da unicidade da matriz inversa segue que $(A^{-1})^{-1} = A$ o que na nossa situação nos dá $[(A^m)^{-1}]^{-1} = A^m$. Como $A^{-m} = (A^m)^{-1}$ segue que também para potências negativas temos que a inversa de A^m é A^{-m}.

(ii) Sejam $m, n \in \{0, 1, 2, \dots\}$, vamos dividir em quatro casos:

a) Fazendo $r = m \geq 0$ e $s = n \geq 0$

$$\begin{aligned} A^m \cdot A^n &= \underbrace{(A \cdots A)}_{m \text{ vezes}} \underbrace{(A \cdots A)}_{n \text{ vezes}} \\ &= \underbrace{(A \cdots A \cdot A \cdots A)}_{m+n \text{ vezes}} \\ &= A^{m+n} \end{aligned}$$

b) Supondo A invertível, fazendo $r = -m \leq 0$ e $s = -n \leq 0$ e usando (i) temos

$$\begin{aligned} A^{-m} \cdot A^{-n} &= \underbrace{(A^{-1} \cdots A^{-1})}_{m \text{ vezes}} \underbrace{(A^{-1} \cdots A^{-1})}_{n \text{ vezes}} \\ &= \underbrace{(A^{-1} \cdots A^{-1} \cdot A^{-1} \cdots A^{-1})}_{m+n \text{ vezes}} \\ &= (A^{-1})^{m+n} = A^{-(m+n)} \\ &= A^{-m-n} \end{aligned}$$

c) Fazendo $r = m \geq 0$, $s = -n \leq 0$ e $m - n \geq 0$ e usando (a)

$$A^m = A^{(m-n)+n} = \underbrace{(A \cdots A)}_{m-n \text{ vezes}} \underbrace{(A \cdots A)}_{n \text{ vezes}}$$

Logo

$$\begin{aligned} A^m \cdot A^{-n} &= \underbrace{(A \cdots A)}_{m-n \text{ vezes}} \underbrace{(A \cdots A)}_{n \text{ vezes}} \underbrace{(A^{-1} \cdots A^{-1})}_{n \text{ vezes}} \\ &= \underbrace{(A \cdots A)}_{m-n \text{ vezes}} \cdot I \\ &= A^{m-n} = A^{m+(-n)} \end{aligned}$$

Por outro lado

$$\begin{aligned} A^{-n} \cdot A^m &= \underbrace{(A^{-1} \cdots A^{-1})}_{n \text{ vezes}} \underbrace{(A \cdots A)}_{m \text{ vezes}} \\ &= A^{m-n} \end{aligned}$$

d) Fazendo $r = m \geq 0$, $s = -n \leq 0$ e $m - n < 0$, então $n - m > 0$ e assim por (i) temos

$$\begin{aligned} A^{m-n} &= (A^{-1})^{n-m} \overset{(c)}{=} (A^{-1})^n (A^{-1})^{-m} \overset{(i)}{=} A^{-n} [(A^{-1})^{-1}]^m \\ &= A^{-n} A^m \end{aligned}$$

(iii) $(rA) \cdot \left(\frac{1}{r}A^{-1}\right) = r \cdot \frac{1}{r}A \cdot A^{-1} = 1 \cdot I = I$

□

Teorema 7. Qualquer matriz A pode ser escrita como um produto

$$A = F_1 \cdots F_l \cdot U$$

no qual as F_i são matrizes elementares e U está na foma escalonada ou escalonada reduzida por linhas.

Demonstração. Pela Proposição 5 cada matriz elementar é invertível e pela Proposição 6 qualquer produto de matrizes elementares também é invertível. Pelo Corolário 1

$$E_l \cdots E_1 \cdot A = U$$

e observando-se que inversas de matrizes elementares também o são temos que se $F_i = E_i^{-1}$

$$A = F_1 \cdot F_2 \cdots F_l \cdot U$$

□

Estamos prontos para responder a pergunta:

Se A é uma matriz $m \times n$ então resolver $A\mathbf{x} = \mathbf{b}$ equivale a resolver $U\mathbf{x} = \mathbf{c}$ onde $\begin{pmatrix} U & | & \mathbf{c} \end{pmatrix}$ é a forma escalonada da matriz aumentada $\begin{pmatrix} A & | & \mathbf{b} \end{pmatrix}$?

Se $\mathbf{x}$ é solução de $A\mathbf{x} = \mathbf{b}$ e como $U = \underbrace{E_k \cdot E_{k-1} \cdot E_1}_{E} A = EA$ então,

$$E \cdot Ax = \underbrace{E \cdot b}_{c} \Rightarrow Ux = c$$

Analogamente se $Ux = c$ onde $U = EA$ e $c = Eb$, como E é invertível temos

$$E^{-1} \cdot Ux = E^{-1}c \Rightarrow Ax = b$$

Deste modo concluímos:

Proposição 7. Seja o sistema dado por $A\mathbf{x} = \mathbf{b}$ e seja U a matriz escalonada de A isto é, $U = EA$ onde E é produto de matrizes elementares. Então x é solução de $Ax = b$ se e somente se x é solução de $Ux = c$ onde $c = Eb$.

Logo para resolver o sistema $A\mathbf{x} = \mathbf{b}$ basta encontrar o correspondente sistema escalonado.

Com base nesta proposição temos a definição abaixo:

Definição 9. Seja a matriz $A_{m\times n}$. Definimos o posto de A como sendo o posto da matriz escalonada U correspondente.

Logo, desta definição e da Proposição 7 segue naturalmente a generalização do Teorema 1

Teorema 8. (Existência de Solução para Sistemas) Seja um sistema linear de m equações a n incógnitas que matricialmente é dado por $A\mathbf{x} = \mathbf{b}$, cuja matriz de coeficientes $A_{(m\times n)}$ tem posto p e a respectiva matriz aumentada $\hat{A}_{(m\times(n+1))}$ tem posto q.

Então

1. Se $p = q = n$ o sistema têm solução única;
2. Se $p = q < n$ o sistema têm infinitas soluções;
3. Se $p < q$ o sistema não tem solução.

Vamos agora continuar o estudo das matrizes invertíveis.

Definição 10. Uma matriz quadrada cujo posto é igual ao número de suas linhas (ou colunas) é chamada de **matriz não-singular**. Caso contrário ela será chamada de **matriz singular**.

Teorema 9. Se uma matriz A de tamanho $n\times n$ é invertível, então A é não-singular e a única solução do sistema de equações lineares $A\mathbf{x} = \mathbf{b}$ é $\mathbf{x} = A^{-1}\mathbf{b}$.

Demonstração. (a) Se A é invertível, do Teorema 7 existe E invertível tal que $U = EA$ e portanto Ué invertível.

Por contradição se o posto de $A < n$ então a última linha de U é nula. Como U é invertível, existe U^{-1} matriz $n\times n$ tal que $UU^{-1} = I$.

Em particular se $\begin{pmatrix} \mu_{1n} \\ \mu_{2n} \\ \vdots \\ \mu_{nn} \end{pmatrix}^{-1}$ denota a última coluna de U^{-1} temos por um lado

$$U\begin{pmatrix} \mu_{1n} \\ \mu_{2n} \\ \vdots \\ \mu_{nn} \end{pmatrix}^{-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Por outro lado,

$$U\begin{pmatrix} \mu_{1n} \\ \mu_{2n} \\ \vdots \\ \mu_{nn} \end{pmatrix}^{-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

pois a última linha de U é nula, o que nos dá uma contradição.
Logo A invertível $\Rightarrow$ posto $A = n$, portanto A é não-singular.

(b) Já vimos que se A têm posto $n \Rightarrow U$ têm posto n, logo $Ux = c = Eb$ têm solução única e consequentemente $Ax = b$ têm solução única.
Além disso, como A é invertível

$$\begin{aligned} A^{-1}(Ax) &= A^{-1}b \\ (A^{-1}A)x &= A^{-1}b \\ x &= A^{-1}b \end{aligned}$$

□

Corolário 3. Se uma matriz $A_{n\times n}$ é não-singular então a sua forma escalonada reduzida por linhas $\bar{U}$ é a matriz identidade.

Demonstração. Sabemos que a matriz escalonada reduzida por linhas de uma matriz quadrada é composta por zeros nas posições $(i,j), i \neq j$ e por zeros ou 1 nas posições (i,i). Como A é não singular, seu posto é máximo, isto é, n. Consequentemente todos os elementos (i,i) de $\bar{U}$ são 1. Portanto $\bar{U} = I$. □

Teorema 10. Se uma matriz $A_{n\times n}$ é não-singular então A é invertível.

Demonstração. Suponha que A é não-singular. Provaremos que essa matriz tem uma inversa mostrando como calculá-la. Denotemos por $\mathbf{e_i}$ a i-ésima coluna de I. Por exemplo, se $n = 4$,

$$\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{e} \quad \mathbf{e_4} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Para destacar as colunas da matriz identidade escrevemos $I = (\mathbf{e_1}, \mathbf{e_2}, \cdots, \mathbf{e_n})$. Sendo A não-singular, a equação $A\mathbf{x} = \mathbf{e_i}$ tem uma única solução $\mathbf{x} = \mathbf{c_i}$ ($\mathbf{c_i}$ é uma matriz $n \times 1$). Seja C a matriz cujas n colunas são as respectivas soluções $\mathbf{c_1}, \mathbf{c_2}, \cdots, \mathbf{c_n}$. Como multiplicamos cada linha de A pela j-ésima coluna de C para obter a j-ésima coluna de AC, podemos escrever

$$\begin{aligned} AC &= A(\mathbf{c_1}, \cdots, \mathbf{c_n}) \\ &= (A\mathbf{c_1}, \cdots, A\mathbf{c_n}) \\ &= (\mathbf{e_1}, \cdots, \mathbf{e_n}) \\ &= I \end{aligned} \tag{1.2}$$

Assim, C é uma inversa à direita de A.

Para mostrar que A também tem uma inversa à esquerda, usamos o Teorema 3 para escrever $EA = \bar{U}$, onde E é um produto de matrizes elementares e $\bar{U}$ é a forma escalonada reduzida por linhas de A. Como A é não-singular, pelo Corolário 3, $\bar{U} = I$. Portanto, E é uma inversa à esquerda de A. Como A tem uma inversa à direita e uma inversa à esquerda, ela é invertível e $C = E = A^{-1}$. □

A prova do Teorema 10 realmente mostra como calcular a inversa de uma matriz não-singular. Para encontrar a i-ésima coluna $\mathbf{c_i}$ de A^{-1}, resolvemos o sistema

$$A\mathbf{x} = \mathbf{e_i}$$

para encontrar a solução $\mathbf{x} = \mathbf{c_i}$. Uma sequência de transformações elementares pode ser usada para resolver esse sistema para cada i. Neste caso, a matriz aumentada é $\begin{pmatrix} A & | & \mathbf{e_i} \end{pmatrix}$. As transformações elementares que reduzirão este sistema dependem tão somente das n primeiras colunas da matriz aumentada, ou seja, somente da matriz A. Nunca utilizamos a última coluna de uma matriz aumentada para determinar quais transformações elementares usamos para reduzi-la à sua forma escalonada reduzida. Portanto, as mesmas operações sobre linhas que reduzirão $\begin{pmatrix} A & | & \mathbf{e_i} \end{pmatrix}$ a $\begin{pmatrix} I & | & \mathbf{c_i} \end{pmatrix}$ também reduzirão $\begin{pmatrix} A & | & \mathbf{e_j} \end{pmatrix}$ a $\begin{pmatrix} I & | & \mathbf{c_j} \end{pmatrix}$. Podemos ser mais eficientes aglutinando toda essa informação em uma matriz aumentada $\begin{pmatrix} A & | & \mathbf{e_i} & \cdots & \mathbf{e_n} \end{pmatrix} = \begin{pmatrix} A & | & I \end{pmatrix}$ e aplicar a sequência de transformações elementares somente uma única vez. Neste processo, a matriz aumentada $\begin{pmatrix} A & | & I \end{pmatrix}$ transforma-se em $\begin{pmatrix} I & | & A^{-1} \end{pmatrix}$.

Exemplo 12. Vamos aplicar esse método para encontrar a inversa da matriz A no exemplo a seguir:

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

Primeiro, aumente A com a matriz identidade:

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

Vamos agora reduzir a matriz A a sua forma escalonada reduzida por linhas, para isso aplicamos os passos de 1 a 4 descritos no Teorema 3.

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

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & -1 & | & -2 & 1 & 0 \\ 3 & 1 & 0 & | & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & -1 & | & -2 & 1 & 0 \\ 0 & 1 & -6 & | & -3 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & -1 & | & -2 & 1 & 0 \\ 0 & 1 & -6 & | & -3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & -1 & | & -2 & 1 & 0 \\ 0 & 0 & -5 & | & -1 & -1 & 1 \end{pmatrix}$$

Agora vamos fazer o pivô da terceira linha ficar igual a 1.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{5} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & -1 & | & -2 & 1 & 0 \\ 0 & 0 & -5 & | & -1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & -1 & | & -2 & 1 & 0 \\ 0 & 0 & 1 & | & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{pmatrix}$$

Agora cada coluna que contém um pivô deve ter todos os outros elementos iguais a zero.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & -1 & | & -2 & 1 & 0 \\ 0 & 0 & 1 & | & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & -\frac{9}{5} & \frac{6}{5} & -\frac{1}{5} \\ 0 & 0 & 1 & | & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & -\frac{9}{5} & \frac{6}{5} & -\frac{1}{5} \\ 0 & 0 & 1 & | & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & | & \frac{3}{5} & -\frac{2}{5} & \frac{2}{5} \\ 0 & 1 & 0 & | & -\frac{9}{5} & \frac{6}{5} & -\frac{1}{5} \\ 0 & 0 & 1 & | & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{pmatrix}$$

Como indica a prova do Teorema 10, o lado direito desta matriz aumentada, ou seja,

$$\begin{pmatrix} \frac{3}{5} & -\frac{2}{5} & \frac{2}{5} \\ -\frac{9}{5} & \frac{6}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{pmatrix}$$

é a inversa de A.

Teorema 11. Uma matriz arbitrária A de tamanho 2×2 dada por

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

é não-singular (e, portanto, invertível) se, e somente se, $ad - bc \neq 0$. E sua inversa é

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Demonstração. Vamos aplicar o método descrito no Teorema 10 e usado no exemplo anterior na matriz A para obter sua inversa.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Primeiro, aumentamos A com a matriz identidade:

$$\begin{pmatrix} A & | & I \end{pmatrix} = \begin{pmatrix} a & b & | & 1 & 0 \\ c & d & | & 0 & 1 \end{pmatrix}$$

Se a e c são ambos 0, A claramente é singular. vamos supor, portanto, que $a \neq 0$. Assim fazemos a seguinte transformação:

$$\begin{pmatrix} 1 & 0 \\ -\frac{c}{a} & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b & | & 1 & 0 \\ c & d & | & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b & | & 1 & 0 \\ 0 & \frac{ad-bc}{a} & | & -\frac{c}{a} & 1 \end{pmatrix}$$

Este curto cálculo já nos diz que, quando $a \neq 0$, A é não-singular (e portanto invertível) se, e somente se, $ad - bc \neq 0$. Agora continuamos fazendo os pivôs ficarem iguais a 1 com as seguintes transformações:

$$\begin{pmatrix} 1 & 0 \\ 0 & \frac{a}{ad-bc} \end{pmatrix} \cdot \begin{pmatrix} a & b & | & 1 & 0 \\ 0 & \frac{ad-bc}{a} & | & -\frac{c}{a} & 1 \end{pmatrix} = \begin{pmatrix} a & b & | & 1 & 0 \\ 0 & 1 & | & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{a} & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b & | & 1 & 0 \\ 0 & \frac{ad-bc}{a} & | & -\frac{c}{a} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{b}{a} & | & \frac{1}{a} & 0 \\ 0 & 1 & | & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$$

Para completar a redução fazemos

$$\begin{pmatrix} 1 & -\frac{b}{a} \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & \frac{b}{a} & | & \frac{1}{a} & 0 \\ 0 & 1 & | & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} = \begin{pmatrix} 1 & 0 & | & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & | & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$$

Lendo a partir da segunda metade da matriz aumentada, vemos que

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

De modo análogo, se $a = 0$ como A é não singular tem-se $b \neq 0 \neq c$. Escalonando $\begin{pmatrix} 0 & b & | & 1 & 0 \\ c & d & | & 0 & 1 \end{pmatrix}$ concluímos que

$$A^{-1} = -\frac{1}{bc} \begin{pmatrix} d & -b \\ -c & 0 \end{pmatrix}.$$

□

1.6.1 Aplicação: Matrizes diagonais dominantes

Vamos agora investigar um tipo especial de matrizes invertíveis. Este tipo diz respeito a um caso particular de Matriz Diagonal Dominante. Como não trabalharemos com o caso geral, vamos defini-la somente para o caso particular de nosso interesse.

Definição 11. Seja matriz B satisfazendo:

$$B = \begin{pmatrix} b_{11} & -b_{12} & \cdots & -b_{1n} \\ -b_{21} & b_{22} & \cdots & -b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -b_{n1} & -b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

onde

- $b_{ii} > 0$ para $i = 1, \ldots, n$.
- $b_{ij} \geq 0$ para $i, j = 1, \ldots, n$ e $i \neq j$.
- $b_{jj} - \sum_{h=1,\ h\neq j}^{n} b_{hj} > 0$ para $j = 1, \ldots, n$.

Tal matriz será denominada **matriz diagonal dominante**.

As matrizes diagonais dominantes serão utilizadas adiante em modelos lineares de produção.

Teorema 12. Se B é uma matriz diagonal dominante, então B é invertível e todas as entradas de B^{-1} serão não-negativas.

Antes de demonstrar este Teorema, consideremos inicialmente o sistema $B\mathbf{X} = \mathbf{C}$ onde $\mathbf{C} = (c_1, c_2, \ldots, c_n)$ e $\mathbf{X} = (x_1, x_2, \ldots, x_n)$.

Para resolver este sistema utilizaremos eliminação gaussiana na matriz aumentada $\begin{pmatrix} B & | & \mathbf{C} \end{pmatrix}$.

Ao invés de multiplicarmos à esquerda por matrizes elementares, operaremos diretamente com as linhas da matriz aumentada.

Observemos primeiramente que $b_{11} > 0$. Assim a fim de anularmos os elementos $(i,1)$ da matriz aumentada multiplicaremos a linha 1 por $\dfrac{b_{i1}}{b_{11}}$ e a somaremos à linha i, obtendo:

$$\begin{pmatrix} b_{11} & -b_{12} & \cdots & -b_{1n} & | & c_1 \\ 0 & b_{22}-\dfrac{b_{21}}{b_{11}}b_{12} & \cdots & -b_{2n}-\dfrac{b_{21}}{b_{11}}b_{1n} & | & c_2+\dfrac{b_{21}}{b_{11}}c_1 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & -b_{n2}-\dfrac{b_{n1}}{b_{11}}b_{12} & \cdots & b_{nn}-\dfrac{b_{n1}}{b_{11}}b_{1n} & | & c_n+\dfrac{b_{n1}}{b_{11}}c_1 \end{pmatrix} \tag{1.3}$$
$$= \begin{pmatrix} b_{11} & * & | & c_1 \\ \mathbf{0} & \hat{\mathbf{B}} & | & \hat{\mathbf{C}} \end{pmatrix}$$

Observe que $\hat{\mathbf{B}}$ também é uma matriz diagonal dominante. De fato, a exemplo da notação usada para os elementos de B, vamos denotar os elementos de $\hat{B}$ por:

$$\hat{b}_{ii}, \text{ para } i = 2,3,\cdots n \quad \text{e} \quad -\hat{b}_{ij}, \text{ para } i \neq j = 2,3,\cdots n,$$

vamos mostrar que

- $\hat{b}_{ii} > 0$ para $i = 2,\ldots,n$.
- $\hat{b}_{ij} \geq 0$ para $i,j = 2,\ldots,n$ e $i \neq j$.
- $\hat{b}_{jj} - \sum\limits_{h=2,\ h\neq j}^{n} \hat{b}_{hj} > 0$ para $j = 2,\ldots,n$.

(i)

$$\hat{b}_{jj} = b_{jj} - \frac{b_{j1}}{b_{11}}b_{1j} = \frac{b_{11}b_{jj} - b_{j1}b_{1j}}{b_{11}}.$$

Mas para todo j temos que $\sum\limits_{h=1}^{n} b_{hj} > 0 \ \forall \ j$ logo

$$b_{jj} > \sum_{h=1,\ h\neq j}^{n} b_{hj} \geq b_{1j} \geq 0,\ j = 2,\ldots,n$$
$$b_{11} > \sum_{h\neq 1} b_{h1} \geq b_{j1} \geq 0.$$

Logo $b_{11}b_{jj} > b_{j1}b_{1j}$ o que nos dá $\hat{b}_{jj} > 0$ para $j = 2,\ldots,n$.

(ii)

$$-\hat{b}_{ij} = -b_{ij} - \frac{b_{i1}}{b_{11}}b_{1j} = \frac{-b_{ij}b_{11} - b_{i1}b_{1j}}{b_{11}} \leq 0.$$

(iii)

$$\begin{aligned}
\hat{b}_{jj} - \sum_{h\neq 1,j} \hat{b}_{hj} &= \left(b_{jj} - \frac{b_{j1}}{b_{11}} b_{1j}\right) - \sum_{h\neq 1,j} \left(b_{hj} + \frac{b_{h1}}{b_{11}} b_{1j}\right) \\
&= \left[b_{jj} - \sum_{h\neq 1,j} b_{hj}\right] - \frac{b_{1j}}{b_{11}} \sum_{h\neq 1} b_{h1}.
\end{aligned}$$

A primeira parcela é positiva pois $b_{jj} > \sum_{h\neq j} b_{hj}$.

Estimando a segunda parcela, como $b_{11} > \sum_{h\neq 1} b_{h1}$, então, $0 < \left(\frac{\sum b_{h1}}{b_{11}}\right) b_{1j} < b_{1j}$. Logo,

$$\hat{b}_{jj} - \sum_{h\neq 1,j} \hat{b}_{hj} > \left[b_{jj} - \sum_{h\neq 1,j} b_{hj}\right] - b_{1j} = b_{jj} - \sum_{h\neq j} b_{hj} > 0.$$

Logo, $\hat{\mathbf{B}}$ é diagonal dominante.

Acabamos de demonstrar o seguinte lema:

Lema 2. Seja $B_{n\times n}$ diagonal dominante e seja $\hat{\mathbf{B}}$ a submatriz de ordem $n-1$ obtida após realizarmos o primeiro passo para escalonar B. Então, $\hat{\mathbf{B}}$ também é diagonal dominante.

Ainda analisando (1.3) vejamos o que é possível deduzir de $\hat{\mathbf{C}}$ quando conhecemos $\mathbf{C} = (c_1, c_2, \ldots, c_n)$.

Assim, voltando a (1.3), vemos que no caso específico de:

$$\mathbf{C} = e_i = (0, \ldots, 0, \underbrace{1}_{i\text{-ésima}}, 0, \ldots) \in \mathbb{R}^n \text{ e } i = 2, 3, \ldots, n$$

então

$$\hat{\mathbf{C}} = (0, \ldots, 0, \underbrace{1}_{(i-1)\text{-ésima}}, 0, \ldots) \in \mathbb{R}^{n-1}.$$

Mas para $\mathbf{C} = e_1 = (1, 0, 0, \ldots) \in \mathbb{R}^n$ então

$$\hat{\mathbf{C}} = \left(\frac{b_{21}}{b_{11}}, \frac{b_{31}}{b_{11}}, \ldots, \frac{b_{n1}}{b_{11}}\right) \geq 0.$$

Observe que se para algum $i \geq 2, b_{i1} > 0$ então $\hat{\mathbf{C}}$ terá a respectiva entrada positiva.

Já se $b_{i1} = 0 \ \forall\ i = 2, \ldots, n$ é porque a primeira etapa do escalonamento é desnecessária.

Assim se, em particular, $\mathbf{C}$ for algum dos vetores da base canônica, $e_1, \ldots, e_n$, o vetor

$$\begin{pmatrix} c_1 \\ c_2 + \frac{b_{21}}{b_{11}} c_1 \\ \vdots \\ c_n + \frac{b_{n1}}{b_{11}} c_1 \end{pmatrix}$$

terá as suas coordenadas não negativas e pelo menos uma delas será positiva.

Destes fatos e do Lema 2 concluímos que continuando o processo de escalonamento a matriz $\begin{pmatrix} B & | & \mathbf{C} \end{pmatrix}$, para $\mathbf{C} = e_i$ para qualquer $i = 1, \ldots, n$, tornar-se-á:

$$\begin{pmatrix} + & - & - & \cdots & - & | & \geq 0 \\ 0 & + & - & \cdots & - & | & \geq 0 \\ 0 & 0 & + & \cdots & - & | & \geq 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & 0 & \cdots & + & | & \geq 0 \end{pmatrix}$$

onde pelo menos uma das entradas da última coluna será positiva. Consequentemente na resolução de $B\mathbf{X} = e_i$ a solução, que denominaremos $\mathbf{X}_i$, será única, não-negativa e com pelo menos uma de suas entradas positiva.

Logo, B será invertível e como a i-ésima coluna de B^{-1} será $\mathbf{X}_i$, concluímos que B^{-1} só terá entradas não-negativas, como queríamos.

Corolário 4. Seja A uma matriz $n \times n$, tal que cada uma de suas entradas é não-negativa e a soma das entradas em cada coluna é menor do que 1. Então, a matriz $(I - A)^{-1}$ existe e contém somente entradas não-negativas.

Demonstração. Por hipótese temos $a_{ij} \geq 0$ e $\sum_{i=1}^{n} a_{ij} < 1$. Então

$$B = (I - A) = \begin{pmatrix} (1-a_{11}) & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & (1-a_{22}) & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & -a_{n2} & \cdots & (1-a_{nn}) \end{pmatrix} \begin{pmatrix} b_{11} & -b_{12} & \cdots & -b_{1n} \\ -b_{21} & b_{22} & \cdots & -b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -b_{n1} & -b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

$$\begin{aligned} \text{Logo} \quad & b_{jj} = 1 - a_{jj} > 0, \\ & b_{ij} = a_{ij} \geq 0 \text{ para } i \neq j, \\ & b_{jj} - \sum_{i \neq j}^{n} b_{ij} = (1 - a_{jj}) - \sum_{i \neq j}^{n} a_{ij} = 1 - \sum_{i=1}^{n} a_{ij} > 0, \end{aligned}$$

para cada j, isto é, B é matriz diagonal dominante.

Pelo Teorema 12, B^{-1} existe e contém somente entradas não-negativas. $\square$

CAPÍTULO

2

DETERMINANTES

Os sistemas de equações lineares com n equações e n incógnitas

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\vdots \qquad \vdots \qquad\qquad \vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned} \tag{2.1}$$

tem notação matricial $A\mathbf{x} = \mathbf{b}$, onde A é matriz quadrada, e aparecem com frequência em modelos econômicos, que será nosso objeto de estudo no capítulo 4.

É conveniente quando, de antemão, sabemos se a matriz A é não-singular porque determina se o sistema associado tem solução única. Descreveremos neste capítulo um teste direto para determinar se uma dada matriz é não-singular.

Veremos que o **determinante** é o número associado a uma matriz quadrada que informará se ela é não-singular. Também utilizaremos o determinante para desenvolver uma fórmula explícita para a solução de um sistema de equações lineares com n equações e n incógnitas. Tal fórmula será denominada regra de Cramer.

2.1 Definição do Determinante

Nesta seção a ideia é definir um número associado a uma matriz quadrada A que mostre se ela é não-singular e, portanto, invertível. Este número será denominado determinante de A, que indicaremos pelo símbolo $\begin{vmatrix} A \end{vmatrix}$; ou seja, colocando uma barra vertical de cada lado de A ou simplesmente $\det A$.

Para uma matriz 1×1

$$A = \begin{pmatrix} a_{11} \end{pmatrix}$$

a condição necessária para A ser não-singular é $a_{11} \neq 0$. Observe que o inverso deste número $\frac{1}{a_{11}}$ existe se, e somente se, a_{11} é não-nulo. Então, a matriz A será invertível e, portanto, não-singular se, e somente se, $a_{11} \neq 0$. Definimos, então, $\det\begin{pmatrix} a_{11} \end{pmatrix} = a_{11}$.

Para uma matriz 2×2

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

pelo Teorema 11 A é não singular se, e somente se $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Assim, definimos o determinante de uma matriz A de tamanho 2×2:

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{2.2}$$

Para uma matriz 3×3 tomemos

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

não-singular e vejamos uma condição numérica necessária para ela ser não-singular.

Suponha A não-singular, aplicando operações elementares sobre as linhas de A obtemos U triangular superior, equivalente a A, também não-singular. Suponha que U foi obtida:

Caso 1. Sem permuta de linhas.

Então para reduzir A à forma escalonada por linhas U, inicialmente sabemos que a_{11} é diferente de zero. Vamos então pivotar em a_{11} para tornar nulas todas as outras entradas na coluna 1, ou seja vamos multiplicar a matriz A pela esquerda sucessivamente pelas matrizes elementares $E_{12}\left(-\dfrac{a_{21}}{a_{11}}\right)$ e $E_{13}\left(-\dfrac{a_{31}}{a_{11}}\right)$. O resultado é

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} & \dfrac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}} \\ 0 & \dfrac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}} & \dfrac{a_{11}a_{33} - a_{31}a_{13}}{a_{11}} \end{pmatrix} \tag{2.3}$$

Para continuarmos sem trocas de linhas temos que a $(2,2)$-ésima, ou seja, o elemento $\dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}}$ é não-nulo, multiplicamos pela matriz elemantar $E_{23}\left(-\dfrac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}a_{22} - a_{21}a_{12}}\right)$ para zerar a $(3,2)$-ésima entrada de (2.3). Após alguns cálculos o resultado é

$$U = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} & \dfrac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}} \\ 0 & 0 & (*) \end{pmatrix} \tag{2.4}$$

onde $(*)$ é igual a

$$\frac{1}{a_{11}a_{22}-a_{21}a_{12}}(a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31} \qquad (2.5)$$
$$+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31})$$

Como a matriz é não singular, e supomos que seu escalonamento não envolveu permuta de linhas, U é não-singular e o fator entre parênteses em (2.5) é não-nulo. Vamos chamar esse número de D, então

$$\begin{aligned} D &= a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31} \\ &+ a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31} \end{aligned} \qquad (2.6)$$

Observe que no caso especial em que $a_{11}a_{32}-a_{31}a_{12}=0$ em (2.3) então a matriz obtida está na forma escalonada e por ser não singular temos $a_{11}a_{22}-a_{21}a_{12}\neq 0$ e $a_{11}a_{33}-a_{31}a_{13}\neq 0$. Voltando a expressão (2.6), fazendo $a_{11}a_{32}-a_{31}a_{12}=0$, obtemos que $D=a_{11}a_{22}a_{33}-a_{21}a_{12}a_{33}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}$ e portanto não-nulo.

Observação 3. Caso $a_{11}a_{22}-a_{21}a_{12}=0$ em (2.3) deveríamos trocar as linhas 2 e 3 para obter a matriz não singular escalonada U. Assim obteríamos que $a_{11}a_{32}-a_{31}a_{12}\neq 0$ e $a_{11}a_{23}-a_{21}a_{13}\neq 0$. Esta obervação será útil no estudo do caso a seguir.

Observação 4. Observe que em (2.4) a matriz é triangular superior e que o número que chamamos de D é obtido pelo produto de suas entradas na diagonal.

Caso 2. Com 1 permuta de linhas.

(i) Considere que na matriz A $a_{11}=0$ e $a_{21}\neq 0$.

Então para reduzir A à forma escalonada por linhas U vamos permutar a primeira e a segunda linha, para isso multiplicamos A pela esquerda pela matriz elementar E_{12}, obtendo

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} \\ 0 & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Agora pivotamos na $(1,1)$-ésima entrada, ou seja a_{21}, para tornar nula a $(3,1)$-ésima entrada, para isso vamos multiplicar pela matriz elementar $E_{13}\left(-\dfrac{a_{31}}{a_{21}}\right)$. O resultado é

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} \\ 0 & a_{12} & a_{13} \\ 0 & \dfrac{a_{32}a_{21}-a_{31}a_{22}}{a_{21}} & \dfrac{a_{33}a_{21}-a_{31}a_{23}}{a_{21}} \end{pmatrix} \qquad (2.7)$$

Como estamos admitindo uma única permuta e esta já foi feita, temos que a $(2,2)$-ésima entrada de (2.7), ou seja, o elemento a_{12} é não-nulo, então basta zerar o elemento $\dfrac{a_{32}a_{21}-a_{31}a_{22}}{a_{21}}$ multiplicando pela esquerda pela matriz elementar $E_{23}\left(-\dfrac{a_{32}a_{21}-a_{31}a_{22}}{a_{12}a_{21}}\right)$. Depois de alguns cálculos chegamos ao resultado

$$U=\begin{pmatrix} a_{21} & a_{22} & a_{23} \\ 0 & a_{12} & a_{13} \\ 0 & 0 & (*) \end{pmatrix} \tag{2.8}$$

onde $(*)$ é igual a

$$\frac{1}{a_{12}a_{21}}(a_{12}a_{21}a_{33}-a_{12}a_{23}a_{31}-a_{13}a_{21}a_{32}$$
$$+a_{13}a_{22}a_{31})$$

Como o termo a_{21} e o termo a_{12} são não-nulos, e U é não-singular, então o fator entre parenteses em (2.8) é não-nulo.

Comparando esse número com o número que chamamos de D em (2.5) verificamos que trata-se do mesmo número, a menos do sinal, onde $a_{11}=0$. Ou seja

$$\begin{aligned} D &= a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31} \\ &= 0\cdot a_{22}a_{33}-0\cdot a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31} \\ &= -a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31} \\ &= -(a_{12}a_{21}a_{33}-a_{12}a_{23}a_{31}-a_{13}a_{21}a_{32}+a_{13}a_{22}a_{31}) \end{aligned}$$

(ii) Consideremos agora que os elementos $a_{11}=0$, $a_{21}=0$ e $a_{31}\neq 0$.

Para reduzir A a U, inicialmente vamos permutar primeira e a terceira linha, para isso multiplicamos a matriz A pela esquerda pela matriz elementar E_{13}, obtendo

$$\begin{pmatrix} a_{31} & a_{32} & a_{33} \\ 0 & a_{22} & a_{23} \\ 0 & a_{12} & a_{13} \end{pmatrix} \tag{2.9}$$

Novamente como admitimos uma única permuta e U é não singular concluímos que $(2,2)$-ésima entrada de (2.9), ou seja, o elemento a_{22} é não-nulo. E então basta zerar o elemento a_{12} multiplicando pela esquerda pela matriz elementar $E_{23}\left(-\dfrac{a_{12}}{a_{22}}\right)$, chegando ao resultado

$$U=\begin{pmatrix} a_{31} & a_{32} & a_{33} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & \dfrac{a_{13}a_{22}-a_{12}a_{23}}{a_{22}} \end{pmatrix} \tag{2.10}$$

Como U é não-singular, então o numerador da $(3,3)$-ésima entrada de U é naõ nulo, ou seja, $a_{13}a_{22} - a_{12}a_{23} \neq 0$.

Note que esse número é obtido a partir do número D em (2.5) fazendo $a_{11} = 0$ e $a_{21} = 0$, ou seja,

$$\begin{aligned}
D &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \\
&= 0 \cdot a_{22}a_{33} - 0 \cdot a_{23}a_{32} - a_{12}0 \cdot a_{33} + a_{12}a_{23}a_{31} + a_{13}0 \cdot a_{32} - a_{13}a_{22}a_{31} \\
&= +a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} = -a_{31}(a_{13}a_{22} - a_{12}a_{23})
\end{aligned}$$

(iii) No caso de $a_{11} \neq 0$ a permuta ocorrerá na segunda etapa do escalonamento pois $\dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} = 0$. Conforme vimos na Observação (3) $\dfrac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}} \neq 0$ e $\dfrac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}} \neq 0$. Voltando a (2.6) e anulando o termo $\dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}}$ ou seja, fazendo $a_{11}a_{22} - a_{21}a_{12} = 0$ vemos que $D = a_{11}a_{23}a_{32} - a_{12}a_{23}a_{31} - a_{13}a_{21}a_{32} + a_{13}a_{22}a_{31}$ e portanto não-nulo.

Caso 3. Com duas permutas de linhas.

Para não cansarmos o leitor, omitiremos os cálculos, que seguirão de modo análogo aos anteriores. Também aqui, concluiremos que a expressão dada em (2.6) será não nula quando A é não singular.

Conclusão: em todos os casos dada matriz $A_{3\times3}$ não-singular concluímos que $D \neq 0$. Assim temos que $D \neq 0$ é condição necessária para que A seja não singular. Vamos definir D como sendo o determinante de A. Veremos adiante que $D \neq 0$ será também condição suficiente para que A seja não-singular, isto é, $\det(A) \neq 0$ é condição necessária e suficiente para A ser não-singular.

Definição 12. Definimos o determinante de uma matriz $A = [a_{i,j}]_{3\times3}$ por:

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

Vejamos agora a seguinte definição:

Definição 13. Dada A uma matriz $n \times n$ definimos A_{ij} a submatriz $(n-1) \times (n-1)$ obtida suprimindo-se a i-ésima linha e a j-ésima coluna de A. O escalar

$$M_{ij} \equiv \det A_{ij}$$

é denominado o (i,j)-ésimo **menor** de A e o escalar

$$C_{ij} \equiv (-1)^{i+j} M_{ij}$$

é denominado o **co-fator** do elemento $a_{i,j}$ de A. Um co-fator é um menor com sinal. Observe que $M_{ij} = C_{ij}$ se $(i+j)$ é par e $M_{ij} = -C_{ij}$ se $(i+j)$ é ímpar.

Note que no caso 2×2 o determinante de A pode ser reescrito por

$$\begin{aligned} \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= a_{11}a_{22} - a_{12}a_{21} \\ &= a_{11}\cdot(-1)^{1+1}M_{11} + a_{12}\cdot(-1)^{1+2}M_{12} \\ &= a_{21}(-1)^{2+1}M_{21} + a_{22}(-1)^{2+2}\cdot M_{22} \\ &= a_{11}(-1)^{1+1}M_{11} + a_{21}(-1)^{2+1}\cdot M_{21} \\ &= a_{12}(-1)^{1+2}M_{12} + a_{22}(-1)^{2+2}\cdot M_{22} \end{aligned} \tag{2.11}$$

Da mesma forma no caso 3×3

$$\begin{aligned} \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \\ &+ a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \\ = a_{11}\cdot(-1)^{1+1}M_{11} &+ a_{12}\cdot(-1)^{1+2}M_{12} + a_{13}\cdot(-1)^{1+3}M_{13} \\ = a_{21}\cdot(-1)^{2+1}M_{21} &+ a_{22}\cdot(-1)^{2+2}M_{22} + a_{23}\cdot(-1)^{2+3}M_{23} \\ = a_{31}\cdot(-1)^{3+1}M_{31} &+ a_{32}\cdot(-1)^{3+2}M_{32} + a_{33}\cdot(-1)^{3+3}M_{33} \\ = a_{11}\cdot(-1)^{1+1}M_{11} &+ a_{21}\cdot(-1)^{2+1}M_{21} + a_{31}\cdot(-1)^{3+1}M_{31} \\ = a_{12}\cdot(-1)^{1+2}M_{12} &+ a_{22}\cdot(-1)^{2+2}M_{22} + a_{32}\cdot(-1)^{3+2}M_{32} \\ = a_{13}\cdot(-1)^{1+3}M_{13} &+ a_{23}\cdot(-1)^{2+3}M_{23} + a_{33}\cdot(-1)^{3+3}M_{33} \end{aligned} \tag{2.12}$$

Não mostraremos aqui que tal definição é consistente, isto é, expandindo-se o cálculo do determinante ao longo de qualquer linha ou coluna o resultado será sempre o mesmo.

Baseados nas duas últimas observações, podemos definir o determinante de uma matriz $n\times n$.

Definição 14. Definimos o determinante de uma matriz A de tamanho $n\times n$ por:

$$\begin{aligned} \det A &= \sum_{k=1}^{n}(-1)^{i+k}a_{ik}M_{ik} \qquad \text{(expansão ao longo da linha i)} \\ &= \sum_{h=1}^{n}(-1)^{h+j}a_{hj}M_{hj} \qquad \text{(expansão ao longo da coluna j)} \end{aligned} \tag{2.13}$$

Exemplo 13. O determinante da matriz identidade I de tamanho $n \times n$ é 1.

Por indução. A afirmação é claramente verdadeira para matrizes 1×1. Supondo verdadeira para uma matriz identidade $(n-1) \times (n-1)$ (Hipótese de Indução), vamos mostrar que também é verdadeira para uma matriz identidade $n \times n$. Seja I_{ij} como na definição 13. Fazendo o cálculo do determinante de I pela expansão através da linha 1, temos

$$\det I = 1 \cdot \det I_{11} + 0 \cdot \det I_{12} + \cdots + 0 \cdot \det I_{1n} = 1 \cdot \det I_{11}.$$

Mas I_{11} é uma matriz identidade de tamanho $(n-1) \times (n-1)$ e pela hipótese de indução $\det A_{11} = 1$ então $\det I_n = 1$.

Ainda não mostramos que uma matriz é não-singular se, e somente se, seu determinante é não-nulo. Para fazê-lo vamos primeiramente desenvolver algumas propriedades do determinante que nos ajudarão nessa tarefa.

2.2 Propriedades do determinante

Como visto na Definição 14 o determinante de uma matriz é invariante pela escolha da linha ou coluna que utilizarmos como referência, para expansão. Assim os resultados desta seção serão provados usando-se a expansão em linhas, mas poderiam ser tratados pela expansão em colunas.

Propriedade 1. Para qualquer matriz A de tamanho $n \times n$, $\det A = \det A^T$.

Demonstração. Usando indução, vemos facilmente que a afirmação é verdadeira para matrizes 1×1. Supondo-a verdadeira para matrizes $(n-1)(n-1)$, vamos mostrar que também é verdadeira para a matriz arbitrária A de tamanho $n \times n$. Seja A_{ij} como na Definição 13 matriz cofatora do elemento a_{ij} de A. Então $a^T_{ij} = a_{ji}$ e $A^T_{ij} = A_{ji}$. Calculando o determinante de pela expansão através da linha 1, temos

$$\begin{aligned}
\det A^T &= \sum_j (-1)^{1+j} a^T_{1j} \det A^T_{1j} && \text{hipótese de indução} \\
&= \sum_j (-1)^{1+j} a_{j1} \det A_{j1} = \det A && \text{expansão através da coluna 1}
\end{aligned}$$

$\square$

Propriedade 2. Seja B a matriz obtida da matriz arbitrária A $n \times n$ pela permuta de duas linhas (ou colunas). Então $\det B = -\det A$.

Demonstração. Provaremos essa propriedade usando o Princípio da Indução Matemática. Inicialmente mostraremos que a propriedade vale para $n = 2$.

$$\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix} = a_{21}a_{12} - a_{11}a_{22} = - \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

Vamos supor por, hipótese de indução, que a propriedade seja válida para matrizes de ordem $(n-1)$ e provemos que ela também será válida para matrizes de ordem n.

Formemos a matriz B pela permuta das linhas i e j de A. Vamos expandir $\det A$ e $\det B$ ao longo de uma linha diferente da i-ésima ou j-ésima, digamos a linha l. Então

$$\det A = \sum_{k=1}^{n} (-1)^{l+k} a_{lk} \cdot \det A_{lk}$$

Mas se B_{lk} é a submatriz obtida suprimindo-se a linha l e coluna k de B então tal matriz tem tamanho $(n-1) \times (n-1)$ e como $l \neq \{i, j\}$ esta submatriz difere de A_{lk} apenas pela permutação da linha i e linha j de A. Assim pela hipótese de indução, $\det B_{lk} = -\det A_{lk}$. Portanto,

$$\begin{aligned} \det B &= \sum_{k=1}^{n} (-1)^{l+k} a_{lk} \cdot \det B_{lk} \\ &= \sum_{k=1}^{n} (-1)^{l+k} a_{lk} \cdot (-\det A_{lk}) \\ &= -\sum_{k=1}^{n} (-1)^{l+k} a_{lk} \cdot \det A_{lk} \\ &= -\det A. \end{aligned}$$

□

Propriedade 3. Se duas linhas (colunas) de A são iguais, então $detA = 0$.

Demonstração. Suponha que as linhas de índices i e j são iguais. Pela Propriedade 2, se permutarmos as linhas i e j estamos trocando o sinal de $\det A$. Contudo não modificamos A. Em outras palavras, $\det A = -\det A$. como 0 é o único número igual ao seu oposto, $\det A = 0$. □

Propriedade 4. Se a matriz B é obtida a partir da matriz A pela multiplicação de cada entrada na linha (coluna) i pelo escalar r, então $\det B = r \cdot \det A$.

Demonstração. Como A e B coincidem fora da linha (coluna) i, calculando o determinante de B expandido ao longo da linha (coluna) i vemos que o (i,k)-ésimo menor M_{ik} de B é também o (i,k)-ésimo menor de A, para qualquer k. Usando (2.13) na Definição 14, resulta

$$\begin{aligned} \det B &= \sum_{k=1}^{n} (-1)^{i+k} (r a_{ik}) \cdot \det A_{ik} \\ &= r \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \cdot \det A_{ik} = r \cdot \det A. \end{aligned}$$

□

Propriedade 5. Se uma matriz A tem uma linha (coluna) inteiramente constituída de zeros, então $\det A = 0$.

Demonstração. Suponha que cada entrada na linha i de A seja nula. Calculando-se o determinante de A expandido em relação a linha i e observando-se que $a_{ij} = 0$ para todo j, teremos $\det A = 0$. □

Propriedade 6. Sejam A e B duas matrizes $n \times n$ que diferem somente em suas i-ésimas linhas, para algum i fixo. Seja C a matriz cuja i-ésima linha é a soma matricial das i-ésimas linhas de A e B e cujas demais linhas são as mesmas que as de A e B. Então,

$$\det A + \det B = \det C$$

Noutra notação, se $\mathbf{a_i} = (a_{i1} \cdots a_{in})$ e $\mathbf{b_i} = (b_{i1} \cdots b_{in})$ são as linhas i de A e B respectivamente, então

$$\det \begin{pmatrix} \mathbf{a_1} \\ \mathbf{a_2} \\ \vdots \\ \mathbf{a_i} \\ \vdots \\ \mathbf{a_n} \end{pmatrix} + det \begin{pmatrix} \mathbf{a_1} \\ \mathbf{a_2} \\ \vdots \\ \mathbf{b_i} \\ \vdots \\ \mathbf{a_n} \end{pmatrix} = det \begin{pmatrix} \mathbf{a_1} \\ \mathbf{a_2} \\ \vdots \\ \mathbf{a_i} + \mathbf{b_i} \\ \vdots \\ \mathbf{a_n} \end{pmatrix}.$$

Demonstração. Calculando-se o determinante de C pela expansão ao longo da i-ésima linha e observando-se que os menores (i,k) de A, B e C são todos iguais, o que denotaremos por $M_{i,k}$, então

$$\begin{aligned} \det C &= \sum_k (-1)^{i+k} (a_{ik} + b_{ik}).M_{ik} \\ &= \sum_k (-1)^{i+k} a_{ik} M_{ik} + (-1)^{i+k} b_{ik} M_{ik} \\ &= \det A + \det B. \end{aligned}$$

□

Propriedade 7. Se B é a matriz obtida de uma matriz A, somando-se r vezes a linha i à j-ésima linha da matriz A, então

$$\det B = \det A$$

Demonstração. Usando a mesma notação da Propriedade 6, se

$$A = \begin{pmatrix} \mathbf{a_1} \\ \vdots \\ \mathbf{a_i} \\ \vdots \\ \mathbf{a_j} \\ \vdots \\ \mathbf{a_n} \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} \mathbf{a_1} \\ \vdots \\ \mathbf{a_i} \\ \vdots \\ \mathbf{ra_i} + \mathbf{a_j} \\ \vdots \\ \mathbf{a_n} \end{pmatrix}$$

então,

$$\begin{aligned}
\det B &= det \begin{pmatrix} \mathbf{a_1} \\ \vdots \\ \mathbf{a_i} \\ \vdots \\ \mathbf{ra_i} \\ \vdots \\ \mathbf{a_n} \end{pmatrix} + \det \begin{pmatrix} \mathbf{a_1} \\ \vdots \\ \mathbf{a_i} \\ \vdots \\ \mathbf{a_j} \\ \vdots \\ \mathbf{a_n} \end{pmatrix} \quad \text{(pela Propriedade 6)} \\
&= r \cdot \det \begin{pmatrix} \mathbf{a_1} \\ \vdots \\ \mathbf{a_i} \\ \vdots \\ \mathbf{a_i} \\ \vdots \\ \mathbf{a_n} \end{pmatrix} + \det \begin{pmatrix} \mathbf{a_1} \\ \vdots \\ \mathbf{a_i} \\ \vdots \\ \mathbf{a_j} \\ \vdots \\ \mathbf{a_n} \end{pmatrix} \quad \text{(pela Propriedade 4)} \\
&= r \cdot 0 + \det A \quad \text{(pela Propriedade 3)} \\
&= \det A
\end{aligned}$$

□

Propriedade 8. Seja U a forma escalonada por linhas da matriz A. Então, $\det U = \pm \det A$

Demonstração. A transformação de A para U envolve uma sequência finita das seguintes duas operações elementares sobre linhas:

(a) permutando duas linhas, e

(b) somando um múltiplo de uma linha a outra.

Pela Propriedade 2, sempre que efetuarmos a operação (a), somente alteramos o sinal do determinante de nossa matriz, mais precisamente se permutarmos k vezes as linhas de A, $\det U = (-1)^k \det A$. Pela Propriedade 7, sempre que efetuarmos a operação (b) não alteramos o determinante. Assim, o resultado final U do processo de eliminação gaussiana terá ou o mesmo determinante que A ou $\det U$ diferirá de $\det A$ somente pelo sinal. □

Propriedade 9. Para

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \ \det A = a_{11}a_{22}\cdots a_{nn}.$$

Demonstração. Usando a Propriedade 4 n vezes, temos:

$$\begin{aligned}
\det A &= a_{11}\cdot\det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \\
&= a_{11}a_{22}\cdots\det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \\
&\ \ \vdots \qquad\qquad\qquad \vdots \\
&= a_{11}a_{22}\cdots a_{nn}\cdots\det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \\
&= a_{11}a_{22}\cdot a_{nn}\cdots 1 \qquad \text{pelo exemplo (13)}
\end{aligned}$$

□

Observação 5. Vimos no exemplo 13 que $\det I = 1$.

Propriedade 10. O determinante de uma matriz triangular inferior ou superior é o produto das entradas da sua diagonal.

Demonstração. Como a transposta de uma matriz triangular inferior é uma matriz triangular superior, faremos a demonstração apenas para o caso triangular superior.

Vemos que a afirmação é claramente verdadeira para matrizes 1×1. Como hipótese de indução vamos supor verdadeira para matriz triangular superior $(n-1)\times(n-1)$ e mostrar que também é verdadeira para o caso $n \times n$. Seja A_{ij} como na Definição 13. Calculando o determinante pela expansão através da coluna 1, temos

$$\det A = a_{11} \cdot \det A_{11} + 0 \cdot \det A_{21} + \cdots + 0 \cdot \det A_{n1} = a_{11} \cdot \det A_{11}.$$

Mas A_{11} é uma matriz triangular superior $(n-1) \times (n-1)$ com $a_{22}, \cdots, a_{nn}$ em sua diagonal. Pela Hipótese de Indução $\det A_{11} = a_{22} \cdots a_{nn}$ o que nos dá $\det A = a_{11} \cdot a_{22} \cdots a_{nn}$. □

Corolário 5. Uma matriz A triangular superior (ou inferior) é não-singular se, e somente se, $a_{ii} \neq 0 \; \forall \; i = 1, \ldots, n$ se, e somente se, $\det A \neq 0$.

Finalmente chegamos ao nosso principal objetivo. O determinante, como definido em 14, realmente determina se uma matriz quadrada é não-singular.

Teorema 13. Uma matriz quadrada A é não-singular se, e somente se, $\det A \neq 0$.

Demonstração. Seja U a matriz escalonada por linhas de A. Sabemos que o posto de A é igual ao posto de U e portanto A é não-singular se, e somente se, U é não-singular. Mas pelo Corolário 5 U é não-singular se, e somente se, $\det U \neq 0$. Como $\det A = \pm \det U$ concluímos o resultado. □

Vejamos um exemplo para o cálculo do determinante das matrizes elementares.

Exemplo 14. Sejam E_{ij}, $E_i(r)$ e $E_{ij}(r)$ as três classes de matrizes elementares, conforme definidas na seção 1.5. Então,

(a) $\det E_{ij} = -1$

(b) $\det E_i(r) = r$, e

(c) $\det E_{ij}(r) = 1$.

De fato, chamemos por l_i a i-ésima linha da matriz identidade I.

No caso (a), como E_{ij} é a matriz formada pela permuta de l_i com l_j pelo exemplo 13 $\det(I) = 1$ a Propriedade 2 nos dá $\det E_{ij} = -\det I = -1$.

No caso (b) Como $E_i(r)$ consiste em trocar l_i por rl_i o resultado segue da Propriedade 4.

Em (c) como $E_{ij}(r)$ consiste em trocar l_i por $rl_i + l_j$ o resultado segue da Propriedade 7.

Veremos agora como se comporta o determinante em relação ao produto de matrizes, mas antes de provar o resultado geral, vamos tratar de um caso especial.

Lema 3. Seja E uma matriz elementar $n \times n$ e seja B uma matriz $n \times n$ arbitrária. Então

$$\det(E \cdot B) = \det E \cdot \det B.$$

Demonstração. Inicialmente vamos supor $E = E_{ij}$. Então EB é a matriz B com linhas i e j permutadas. Pela Propriedade 2 e exemplo 14, temos $\det EB = -\det B = (-1)\det B = \det E_{ij} \cdot \det B$.

Se $E = E_i(r)$, então EB é igual a matriz B com a linha i multiplicada por r. Pela propriedade 4 e pelo exemplo 14 temos $\det EB = r \cdot \det B = \det E \cdot \det B$.

Se $E = E_{ij}(r)$. Então EB é B com r vezes linha i adicionada a linha j. Pela propriedade 7 e pelo exemplo 14 temos $\det(E_{ij}(r) \cdot B) = \det B = \det E_{ij}(r) \cdot \det B$.

□

Corolário 6. Se $E_1, \ldots, E_k$ são matrizes elementares de ordem n e B matriz qualquer também de ordem n, então $\det(E_1 \cdots E_k \cdot B) = \det E_1 \cdots \det E_k \cdot \det B$.

Demonstração. Segue por indução sobre o número de matrizes E_j. □

Teorema 14. Para matrizes $n \times n$ arbitrárias A e B, vale

$$\det(AB) = \det A \cdot \det B. \tag{2.14}$$

Demonstração. Supondo que A é não-singular, pelo Teorema 7, A pode ser escrita como o produto de matrizes elementares: $A = E_1 \cdot E_2 \cdots E_k$, isto implica que

$$A \cdot B = E_1 \cdot E_2 \cdots E_k \cdot B \tag{2.15}$$

Do Corolário acima

$$\begin{aligned} \det AB &= \det(E_1 \cdot E_2 \cdots E_k \cdot B) \\ &= \det(E_1 \cdot E_2 \cdots E_k) \cdot \det B \\ &= \det A \cdot \det B. \end{aligned}$$

Se A é singular, podemos escrever $A = E_1 \cdots E_k \cdot U$, onde $E_1, \cdots, E_k$ são matrizes elementares e U é a forma escalonada por linhas de A. Agora,

$$AB = E_1 \cdots E_k \cdot U \cdot B$$

e pelo argumento que acabamos de apresentar,

$$\det(AB) = \det(E_1 \cdots E_k) \cdot \det(UB).$$

Como A é singular, cada entrada na última linha de U é nula. Pela maneira que é definido o produto matricial, cada entrada na última linha de UB é igualmente nula. Pela Propriedade 5, segue $\det(UB) = 0$. Portanto, $\det(AB) = 0$. Como A é singular, temos $\det A = 0$. Portanto, $\det(AB) = 0 = \det A \cdot \det B$. □

2.3 Aplicações do determinante

Na seção anterior vimos que o determinante, dado pela Definição 14 é uma ferramenta efetiva para conferir se uma dada matriz quadrada é ou não é singular. Nesta seção descreveremos algumas outras aplicações do determinante.

Definição 15. Seja M_{ij} o (i,j)-ésimo menor da matriz A de tamanho $n \times n$ e seja $C_{ij} = (-1)^{i+j}M_{ij}$ o cofator do elemento a_{ij} de A, como na Definição 13. Definimos a **matriz adjunta** de A e denotamos por $adjA$ de tamanho $n \times n$ a transposta da matriz dos cofatores de A, ou seja, a matriz cuja (i,j)-ésima entrada é o (i,j)-ésimo cofator C_{ji} de A.

Lema 4. Para qualquer matriz A de tamanho $n \times n$, temos $A \cdot adjA = adjA \cdot A = \det A \cdot I$.

Demonstração. Faremos a demonstração apenas para o caso $A \cdot adjA = \det A \cdot I$ pois o caso $adjA \cdot A = \det A \cdot I$ é análogo. Assim devemos mostrar que

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{j1} & \cdots & C_{jn} \\ \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{nn} \end{pmatrix}^T =$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} C_{11} & \cdots & C_{j1} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots & & \\ C_{1n} & \cdots & C_{jn} & \cdots & C_{nn} \end{pmatrix} =$$

$$\begin{pmatrix} \det(A) & 0 & \cdots & 0 \\ 0 & \det(A) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det(A) \end{pmatrix}$$

A entrada na linha i e coluna j de $A \cdot adjA$ é

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots a_{in}C_{jn}. \tag{2.16}$$

Se $i = j$, a expressão (2.16) é simplesmente a expansão de $\det A$ ao longo da i-ésima linha, como na Definição 14, e assim os elementos da diagonal da matriz produto são sempre $\det A$. Para

$i \neq j$, note que se B é a matriz obtida de A ao se substituir a linha j pela linha i

$$B = \begin{array}{cl} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \cdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} & \begin{array}{l} \\ \\ \text{Linha } i \\ \\ \text{Linha } j \\ \\ \\ \end{array} \end{array}$$

então a expressão (2.16) é simplesmente a expansão de $\det B$ ao longo da j-ésima linha. Mas como B tem duas linhas iguais, a Propriedade 3 garante que $\det B = 0$.

Logo $A \cdot adjA$ é matriz diagonal onde todos os elementos da diagonal são $\det A$ como queríamos. □

Teorema 15. Se A é não-singular, então $A^{-1} = \dfrac{1}{\det A} \cdot adjA$.

Demonstração. Se A é não-singular $\det A \neq 0$. Logo do Lema 4 $\dfrac{1}{\det A}(A \cdot adjA) = \dfrac{1}{\det A}(adjA \cdot A) = I$. Logo

$$A^{-1} = \frac{1}{\det A} \cdot adjA.$$

□

Teorema 16. (Regra de Cramer) Seja A uma matriz não-singular de tamanho $n \times n$. A única solução $\mathbf{x} = (x_1, \cdots x_n)$ do sistema $A\mathbf{x} = \mathbf{b}$ é dada por

$$x_j = \frac{\det B_j}{\det A}, \qquad \text{para } j = 1, \cdots n$$

onde B_j é a matriz obtida de A substituindo a j-ésima coluna de A por $\mathbf{b}$.

Demonstração. Usando o Teorema 15 e lembrando que C_{ij} é o (i, j)-ésimo co-fator de A teremos

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{\det A} \begin{pmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1j} & \cdots & C_{nj} \\ \vdots & \ddots & \vdots \\ C_{1n} & \cdots & C_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Portanto se $x = (x_1, x_2, \ldots, x_n)$ temos

$$x_j = \frac{b_1 C_{1j} + b_2 C_{2j} + \cdots + b_n C_{nj}}{\det A}. \tag{2.17}$$

Assim, o que resta mostrar é que o numerador de (2.17) é $\det B_j$. Como B_j e A são iguais exceto pela coluna j, então para cada k o (k, j)-ésimo co-fator de B_j e de A são iguais, isto é C_{kj}. Se expandimos $\det B_j$ ao longo da j-ésima coluna obtemos o numerador (2.17). Assim $x_j = \dfrac{\det B_j}{\det A}$. □

Mostraremos agora mais uma aplicação do determinante de uma matriz. Trata-se de um teste para verificar se é necessária ou não alguma permuta de linhas para obtermos sua forma escalonada. Vimos na demonstração do Teorema 11 que não é necessário permuta alguma de linhas para reduzir uma matriz 2×2 arbitrária se, e somente se, $a \neq 0$. Já para uma matriz 3×3 arbitrária, das expressões (2.3), (2.4) concluímos que não será necessária permuta de linhas para obtermos sua forma escalonada se

$$a_{11} \neq 0 \;\text{ e }\; \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11}} \neq 0.$$

No resultado a seguir, veremos como este fato se generaliza para o caso de matrizes de ordem n. Observe que quando A é 2×2 ou 3×3 a condição para que não haja permuta de linhas é que os determinantes das matrizes abaixo sejam não nulas.

$$A_1 = \begin{pmatrix} a_{11} \end{pmatrix}, A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Antes de ver o resultado definimos:

Definição 16. Dada matriz $A = [a_{ij}]$ de ordem n definimos as matrizes de ordem k, onde $k = 1, \cdots, n$,

$$A_k = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{pmatrix}.$$

Note que $A_1 = a_{11}$ e $A_n = A$.

Teorema 17. Seja A uma matriz de ordem m e seja A_k como na definição anterior, $k \leq m$. Suponha que $\det A_k \neq 0$ para $k = 1,2,\dots,n \leq m$. Então para obtermos a forma U, escalonada por linhas de A, não será necessário realizar nenhuma permuta de suas n primeiras linhas. Além disso os n primeiros pivôs de U serão $p_1 = a_{11}$ e $p_k = \dfrac{\det A_k}{\det A_{k-1}}$ para $k = 2,3,\dots,n$.

Demonstração. Faremos a prova por indução sobre n. Suponha $\det A_1 \neq 0$ então $a_{11} \neq 0$ e podemos escalonar A sem permutar a linha 1 e neste caso $p_1 = a_{11} = \det A_1$.

Suponha, como hipótese de indução, que se $\det A_k \neq 0$ para $k = 1,2,\dots,n-1$ então podemos escalonar A sem permutar suas $(n-1)$ primeiras linhas, e seus $(n-1)$ primeiros pivôs serão $p_1 = \det A_1$, $p_k = \dfrac{\det A_k}{\det A_{k-1}}$ para $2 \leq k \leq n-1$.

Suponha agora $\det A_k \neq 0$, $k = 1, 2, \ldots, n$. Pela hipótese de indução até a etapa $n-1$ realizam o escalonamento sem permutar as $n-1$ primeiras linhas. Logo existem matrizes elementares $E_1, E_2, \ldots, E_n$, nenhuma delas matriz de permutação, tais que

$$E_1 \cdot E_2 \cdot \cdots \cdot E_n \cdot A = \begin{pmatrix} U_{k-1} & * & * \\ 0 & p_k & * \\ 0 & * & * \end{pmatrix}.$$

Observe também que segundo a definição e notação dada em 16 temos que

$$[E_1]_n \cdot [E_2]_n \ldots A_n = \begin{pmatrix} U_{k-1} & * \\ 0 & p_k \end{pmatrix} = U_n.$$

Em primeiro lugar, como nenhuma das matrizes elementares é de permutação segue que $\det U_n = \det A_n \neq 0$. Note que pelo mesmo motivo conclui-se que $\det A_k = \det U_k$, $k = 1, \ldots, n$. Além disso U_n é matriz diagonal superior, então

$$\det U_n = (\det U_{n-1}) \cdot p_n$$

portanto

$$p_n = \frac{\det A_n}{\det A_{n-1}}.$$

□

Corolário 7. Se A tem ordem n e $\det A_k \neq 0$, $k = 1, 2, \ldots, n-1$ então a forma escalonada U de A é obtida sem permuta de quaisquer linhas de A e $\det A = p_1 \cdots p_n$ onde $p_k = \dfrac{\det A_k}{\det A_{k-1}}$.

Demonstração. $A = A_n$ logo pelo resultado anterior

$$\det A = \det A_n = (\det A_{n-1}) \cdot p_n = (\det A_{n-2}) \cdot p_{n-1} \cdot p_n = \cdots = p_1 \ldots p_n.$$

□

CAPÍTULO

3

ESPAÇOS VETORIAIS

3.1 Noções preliminares

Nesta seção reunimos fatos básicos sobre Espaços Vetoriais que foram consultados à partir de (ZANI, 2007), (LADEIRA, 2007) e (SIMON; BLUME; DOERING, 2004). Isso por que, em particular, o conjunto das matrizes $n \times m$ é um exemplo de Espaço vetorial. Assumiremos conhecido o conjunto $\mathbb{R}$ dos números reais e suas propriedades algébricas elementares: suas operações de adição e multiplicação são associativas, comutativas, têm elemento neutro, cada número tem seu oposto aditivo e cada número não nulo tem seu inverso multiplicativo.

3.1.1 *Espaço euclidiano n-dimensional*

As noções de par ordenado (x,y) e terna ordenada (x,y,z) de números reais têm uma extensão natural ao conceito de n-**upla** $(x_1,\ldots,x_n)$, que é uma sucessão ordenada de n números reais, Denotaremos as n-uplas por letras em negrito. Se $\mathbf{x} = (x_1,\ldots,x_n)$, cada um dos números $x_1,\ldots,x_n$ é chamado uma **componente** ou **coordenada** de $\mathbf{x}$. Duas n-uplas $(x_1,\ldots,x_n)$ e $(y_1,\ldots,y_n)$ são ditas **iguais** (indicamos $(x_1,\ldots,x_n) = (y_1,\ldots,y_n)$) se e somente se $x_1 = y_1,\ldots,x_n = y_n$. O conjunto de todas n-uplas de números reais é denotado por $\mathbb{R}^n$, isto é,

$$\mathbb{R}^n = (x_1,\ldots,x_n) : x_k \in \mathbb{R}, k = 1,\ldots,n.$$

Recordemos da Geometria Analítica que $\mathbb{R}^3$ pode ser identificado com o conjunto V^3 dos *vetores geométricos* (definidos pelos segmentos orientados) por meio da correspondência que cada $\mathbf{v} = a\mathbf{i}+b\mathbf{j}+c\mathbf{k}$ de V_3 associa a terna $(a,b,c) \in \mathbb{R}^3$:

$$\mathbf{v} = a\mathbf{i}+b\mathbf{j}+c\mathbf{k} \in V_3 \longleftrightarrow (a,b,c) \in \mathbb{R}^3. \tag{3.1}$$

É claro que ao vetor $\mathbf{i}$ corresponde a terna $\mathbf{e}_1 = (1,0,0)$ ao vetor $\mathbf{j}$ corresponde a terna $\mathbf{e}_2 = (0,1,0)$ e a $\mathbf{k}$ corresponde a terna $\mathbf{e}_3 = (0,0,1)$. O **módulo** (ou **comprimento**) do vetor $\mathbf{v}$ é

$\|\mathbf{v}\| = \sqrt{a^2+b^2+c^2}$. A correspondência (3.1) é importante, pois permite caracterizar elementos geométricos, tais como retas, planos, etc, em termos de equações algébricas.

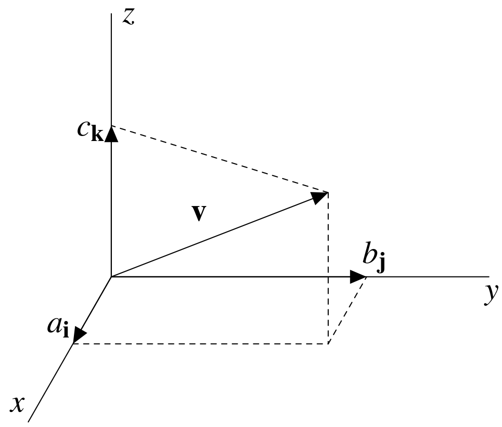

Figura 1

Por causa dessa identificação com os vetores geométricos, as ternas ordenadas também são chamadas de vetores; por extensão, as n-uplas também são chamadas de **vetores**; neste contexto, os números reais serão chamados **escalares**. Lembremos também que, se $\alpha \in \mathbb{R}$, temos $\alpha\mathbf{v} = \alpha a\mathbf{i} + \alpha b\mathbf{j} + \alpha c\mathbf{k}$, ou seja, ao vetor αv associamos a terna $(\alpha a, \alpha b, \alpha c)$. Da mesma maneira, se (a_1,b_1,c_1) e (a_2,b_2,c_2) forem as ternas associadas aos vetores $\mathbf{w_1}$ e $\mathbf{w_2}$, respectivamente (ou seja, $\mathbf{w_1} = a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k}$ e $\mathbf{w_2} = a_2\mathbf{i} + b_2\mathbf{j} + c_2\mathbf{k}$), então temos $\mathbf{w_1} + \mathbf{w_2} = (a_1+a_2)\mathbf{i} + (b_1+b_2)\mathbf{j} + (c_1+c_2)\mathbf{k}$; assim, ao vetor $\mathbf{w_1} + \mathbf{w_2}$ fica associado a terna $(a_1+a_2, b_1+b_2, c_1+c_2)$.

Essas observações mostram a importância de se definir adição de ternas e multiplicação de ternas por números reais: dadas as ternas (a_1,b_1,c_1) e (a_2,b_2,c_2) e o número real α, definimos:

$$(a_1,b_1,c_1) + (a_2,b_2,c_2) = (a_1+a_2, b_1+b_2, c_1+c_2)$$
$$\alpha(a_1,b_1,c_1) = (\alpha a_1, \alpha b_1, \alpha c_1)$$

Pode-se mostrar que, quaisquer que sejam $\mathbf{u},\mathbf{v},\mathbf{w} \in \mathbb{R}^3$ e $\alpha,\beta \in \mathbb{R}$, temos:

(A1) $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$.

(A2) $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$.

(A3) se $\mathbf{0}$ designa a terna $(0,0,0)$, então $\mathbf{u}+\mathbf{0} = \mathbf{u}$, $\forall\, \mathbf{u} \in \mathbb{R}^3$.

(A4) para qualquer $\mathbf{u} = (a,b,c) \in \mathbb{R}^3$, a terna $\mathbf{v} = (-a,-b,-c)$ satisfaz $\mathbf{u}+\mathbf{v} = \mathbf{0}$.

(M1) $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$.

(M2) $(\alpha+\beta)\mathbf{u} = (\alpha\mathbf{u}+\beta\mathbf{u})$.

(M3) $\alpha(\mathbf{u}+\mathbf{v}) = (\alpha\mathbf{u}+\alpha\mathbf{v})$.

(M4) $1\mathbf{u} = \mathbf{u}$.

As operações acima estendem-se de modo natural ao $\mathbb{R}^n$. Dados $\mathbf{u} = (a_1, \ldots, a_n)$ e $\mathbf{v} = (b_1, \ldots, b_n)$ em $\mathbb{R}^n$ e $\alpha \in \mathbb{R}$, definimos a **soma** $\mathbf{u}+\mathbf{v}$ e o **produto por escalar** $\alpha\mathbf{u}$ por

$$\mathbf{u}+\mathbf{v} = (a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_1 + b_n) \tag{3.2}$$

$$\alpha\mathbf{u} = \alpha(a_1, \ldots, a_n) = (\alpha a_1, \ldots, \alpha a_n) \tag{3.3}$$

Como no caso das ternas ordenadas, pode-se verificar que em $\mathbb{R}^n$ estão satisfeitas as propriedades **(A 1)** a **(A 4)** e **(M 1)** a **(M 4)**. Por estarem satisfeitas essas propriedades, dizemos que $\mathbb{R}^n$ é um **espaço vetorial**.

A igualdade (3.2) define a soma de dois vetores. Para somar três vetores $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$, podemos considerar as combinações $\mathbf{u}+(\mathbf{v}+\mathbf{w})$ e $(\mathbf{u}+\mathbf{v})+\mathbf{w}$. A propriedade associativa afirma que esses vetores são iguais. Mais geralmente, dados p vetores $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p$ e p escalares $\alpha_1, \alpha_2, \ldots, \alpha_p$, podemos definir o vetor

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p,$$

que chamaremos **combinação linear** de $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p$.

Exemplo 15. O vetor $(5,0,6,1)$ de $\mathbb{R}^4$ é combinação linear de $(1,4,7,2)$, $(3,8,9,5)$ e $(6,0,1,2)$ pois

$$2 \cdot (1,4,7,2) + (-1) \cdot (3,8,9,5) + 1 \cdot (6,0,1,2) = (5,0,6,1).$$

Exemplo 16. Já o vetor $(1,2,2)$ de $\mathbb{R}^3$ não é combinação linear de $(1,2,-3)$, $(3,1,1)$ e $(8,1,6)$: de fato, se $(1,2,2)$ fosse combinação linear de $(1,2,-3)$, $(3,1,1)$ e $(8,1,6)$ existiriam escalares x,y,z tais que

$$x \cdot (1,2,-3) + y \cdot (3,1,1) + z \cdot (8,1,6) = (1,2,2),$$

ou seja,

$$(x+3y+8z, 2x+y+z, -3x+y+6z) = (1,2,2).$$

Dessa igualdade vemos que x,y,z deveriam satisfazer o sistema de equações

$$\begin{cases} x+3y+8z = 1 \\ 2x+y+z = 2 \\ -3x+y+6z = 2 \end{cases}$$

Que possui matriz de coeficientes A e matriz aumentada $\hat{A}$ dadas por

$$A = \begin{pmatrix} 1 & 3 & 8 \\ 2 & 1 & 1 \\ -3 & 1 & 6 \end{pmatrix} \text{ e } \hat{A} = \begin{pmatrix} 1 & 3 & 8 & | & 1 \\ 2 & 1 & 1 & | & 2 \\ -3 & 1 & 6 & | & 2 \end{pmatrix}$$

cuja forma escalonada de $\hat{A}$ é

$$\left(\begin{array}{ccc|c} 1 & 3 & 8 & 1 \\ 0 & -5 & -15 & 0 \\ 0 & 0 & 0 & 5 \end{array}\right)$$

Observando a forma escalonada de $\hat{A}$ notamos que o posto da matriz aumentada $\hat{A}$ é maior que o posto da matriz de coeficientes A, então pelo Teorema 8 o sistema não possui solução. Logo não existem tais escalares x, y, z. Então $(1,2,2)$ não é combinação linear de $(1,2,-3)$, $(3,1,1)$ e $(8,1,6)$.

Exemplo 17. Consideremos em $\mathbb{R}^n$ os vetores

$$\mathbf{e}_1 = (1,0,\ldots,0),\ \mathbf{e}_2 = (0,1,\ldots,0),\ \ldots,\ \mathbf{e}_n = (0,0,\ldots,1)$$

Podemos escrever $\mathbf{x} = (x_1,\ldots,x_n)$ se escreve, de modo único, como combinação linear dos vetores $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. De fato,

$$\begin{aligned} (x_1,\ldots,x_n) &= (x_1,0,\ldots,0) + \cdots + (0,0,\ldots,x_n) \\ &= x_1(1,0,\ldots,0) + \cdots + x_n(0,0,\ldots,1) \\ &= x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n \end{aligned}$$

Logo, $\mathbf{x}$ é combinação linear de $\mathbf{e}_1,\ldots,\mathbf{e}_n$. Para ver que essa é a única maneira de escrever $\mathbf{x}$ como combinação linear de $\mathbf{e}_1,\ldots,\mathbf{e}_n$, suponhamos que $\mathbf{x}$ também se escreva como $\mathbf{x} = t_1\mathbf{e}_1 + \cdots + t_n\mathbf{e}_n$. Então

$$\begin{aligned} (x_1,\ldots,x_n) &= \mathbf{x} = t_1\mathbf{e}_1 + \cdots + t_n\mathbf{e}_n \\ &= t_1(1,0,\ldots,0) + \cdots + t_n(0,0,\ldots,1) \\ &= (t_1,\ldots,t_n) \end{aligned}$$

Logo, $t_1 = x_1, \ldots, t_n = x_n$.

Por causa desta propriedade, diremos que os vetores $\mathbf{e}_1,\ldots,\mathbf{e}_n$ formam uma base de $\mathbb{R}_n$, chamada **base canônica de** $\mathbb{R}_n$.

Além das operações de adição de n-upla e multiplicação de n-upla por número real, podemos definir em $\mathbb{R}^n$ o chamado *produto interno* de n-uplas, que estende a noção de produto escalar visto nos cursos de Física e Geometria Analítica. Lembremos que o produto escalar de vetores não nulos, $\mathbf{u}$ e $\mathbf{v}$, e que formam entre si um ângulo θ é definido por

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta \tag{3.4}$$

No caso de um deles ser nulo, o produto escalar também será. É conveniente escrever o produto escalar em termos das componentes dos vetores $\mathbf{u} = (a,b,c)$ e $\mathbf{v} = (x,y,z)$. Aplicando a lei dos cossenos ao triângulo cujos lados são $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{u}-\mathbf{v}$ (Figura 2), temos

$$\|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\,\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta. \tag{3.5}$$

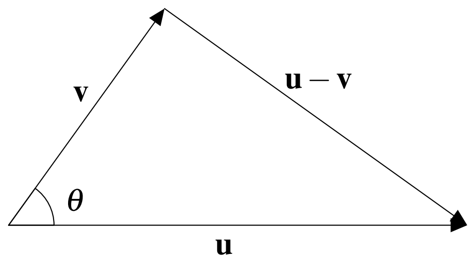

Figura 2

Substituindo em (3.5): $\|\mathbf{u}\|^2 = a^2+b^2+c^2$, $\|\mathbf{v}\|^2 = x^2+y^2+z^2$, $\|\mathbf{u}-\mathbf{v}\|^2 = (x-a)^2+(y-b)^2+(z-c)^2 = \|\mathbf{u}\|^2+\|\mathbf{v}\|^2-2(ax+by+cz)$ e $\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta = \mathbf{uv}$, obtemos

$$\mathbf{uv} = ax+by+cz \tag{3.6}$$

Uma vantagem da relação (3.6) sobre (3.4) é que ela não depende do apelo geométrico e portanto permite estender a $\mathbb{R}^n$, com $n \geq 4$, essa noção de produto escalar, que chamaremos produto interno.

Inspirados neste fato dados $\mathbf{u} = (x_1,\ldots,x_n)$, $\mathbf{v} = (y_1,\ldots,y_n) \in \mathbb{R}^n$, definimos o **produto interno** de $\mathbf{u}$ e $\mathbf{v}$ (ou $\langle \mathbf{u},\mathbf{v}\rangle$), como sendo

$$\mathbf{u}\cdot\mathbf{v} = x_1y_1+\cdots+x_ny_n \tag{3.7}$$

(notemos que o produto interno de dois vetores de $\mathbb{R}^n$ é um número real por isso em muitos textos também é chamado de produto escalar). O espaço vetorial $\mathbb{R}^n$, munido do produto interno, é chamado **espaço euclidiano**. O produto interno tem as seguintes propriedades

$$\mathbf{u}\cdot\mathbf{u} \geq 0 \ \ \forall\, \mathbf{u} \text{ e } \mathbf{u}\cdot\mathbf{u} > 0 \text{ somente se } \mathbf{u} \neq 0. \tag{3.8}$$

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u} \tag{3.9}$$

$$(\mathbf{u}+\alpha\mathbf{w})\cdot\mathbf{v} = \mathbf{u}\cdot\mathbf{v}+\alpha(\mathbf{w}\cdot\mathbf{v}) \tag{3.10}$$

Definimos a **norma** de um vetor $\mathbf{u}$ como sendo $\|\mathbf{u}\| = \sqrt{\mathbf{u}\cdot\mathbf{u}}$. Note que a norma de um vetor é obtida a partir do produto interno.

Exemplo 18. Se $\mathbf{u} = (5,1,0)$, $\mathbf{v} = (\sqrt{2},-1,1)$ e $\mathbf{w} = (0,6,8)$, então $\|\mathbf{u}\| = \sqrt{26}$, $\|\mathbf{v}\|$ =2, $\|\mathbf{w}\|$ =10, $\mathbf{u}\cdot\mathbf{v} = -1+5\sqrt{2}$, $\mathbf{u}\cdot\mathbf{w} = 6$, $\mathbf{v}\cdot\mathbf{w} = 2$.

Existe uma importante desigualdade relacionando norma e produto interno, conhecida como *desigualdade de Cauchy-Schwarz*

$$|\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|\,. \tag{3.11}$$

Se $\mathbf{v} = \mathbf{0}$, temos $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, e a desigualdade (3.11) é trivial. Para mostrar essa desigualdade quando $\mathbf{v} \neq \mathbf{0}$, notemos que, para qualquer $t \in \mathbb{R}$, temos $\|\mathbf{u} + t\mathbf{v}\| \geq 0$. Usando as propriedades (3.9) e (3.10), temos

$$\begin{aligned} 0 &\leq \|\mathbf{u} + \mathbf{v}t\|^2 = (\mathbf{u} + \mathbf{v}t) \cdot (\mathbf{u} + \mathbf{v}t) = \mathbf{u} \cdot \mathbf{u} + 2t\mathbf{u} \cdot \mathbf{v} + t^2 \mathbf{v} \cdot \mathbf{v} \\ &= \|\mathbf{u}\|^2 + 2t\mathbf{u} \cdot \mathbf{v} + t^2 \|\mathbf{v}\|^2 \end{aligned}$$

donde

$$\|\mathbf{v}\|^2 t^2 + 2(\mathbf{u} \cdot \mathbf{v})t + \|\mathbf{u}\|^2 \geq 0. \tag{3.12}$$

O primeiro membro dessa desigualdade é uma função quadrática em t. Para que essa função quadrática seja sempre não negativa, seu discriminante não pode ser positivo, isto é,

$$(\mathbf{u} \cdot \mathbf{v})^2 - 4\|\mathbf{v}\|^2 \|\mathbf{u}\|^2 \leq 0. \tag{3.13}$$

A desigualdade (3.13) implica (3.11).

Como $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos\theta$ dizemos que dois vetores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ são **ortogonais** quando $\mathbf{u} \cdot \mathbf{v} = 0$. Por exemplo, os vetores $\mathbf{u} = (7, 2, -3)$ e $\mathbf{v} = (0, 6, 4)$ são ortogonais, pois $\mathbf{u} \cdot \mathbf{v} = 7 \times 0 + 2 \times 6 + (-3) \times 4 = 0$. Um conjunto de vetores $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ é dito um **conjunto ortogonal** se os seus vetores são dois a dois ortogonais, isto é, $\mathbf{u}_i \cdot \mathbf{u}_j = 0$, quaisquer que sejam i, j com $1 \leq i, j \leq m$ e $i \neq j$; se, além disso, $\|\mathbf{u}_1\| = \cdots = \|\mathbf{u}_m\| = 1$, dizemos que esse conjunto é **ortonormal**. A base canônica $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ é um conjunto ortonormal em $\mathbb{R}^n$.

Exemplo 19. Vamos encontrar todos os vetores de $\mathbb{R}^2$ que são ortogonais a $\mathbf{v} = (3, -1)$. Procuramos então, os vetores $\mathbf{u} = (x, y)$ tais que $\mathbf{u} \cdot \mathbf{v} = 0$, isto é, $3x - y = 0$. Logo, $\mathbf{u} = (x, 3x)$. Notemos que $y = 3x$ é a equação da reta que passa pela origem e tem $\mathbf{v}$ como vetor normal. Observe a figura 3.

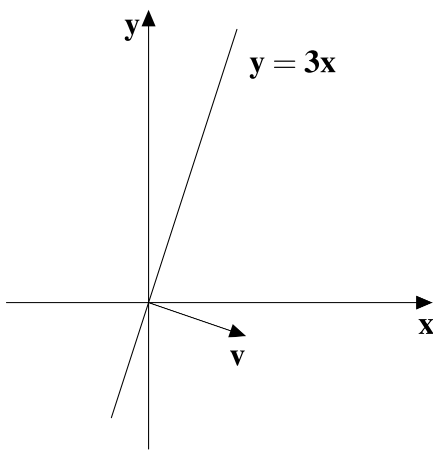

Figura 3

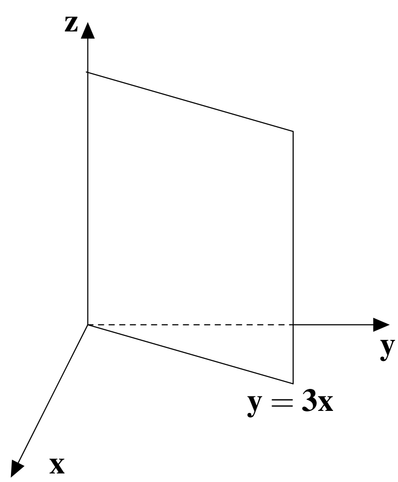

Figura 4

Exemplo 20. Encontraremos agora todos os vetores de $\mathbb{R}^3$ que são ortogonais a $\mathbf{n} = (3, -1, 0)$. Devemos procurar os vetores $\mathbf{u} = (x, y, z)$ tais que $\mathbf{u} \cdot \mathbf{n} = 0$, isto é, $3x - y = 0$. Logo $\mathbf{u} = (x, 3x, z) = x(3, 1, 0) + z(0, 0, 1)$. Notemos que $y = 3x$ é a equação do plano que contém a origem e tem $\mathbf{n}$ como vetor normal. Observe a figura 4.

Exemplo 21. Por fim, vamos encontrar todos os vetores de $\mathbb{R}^3$ que são ortogonais a $\mathbf{v} = (1, 3, 1)$ e $\mathbf{w} = (-1, 1, 0)$. Estamos procurando os vetores $\mathbf{u} = (x, y, z)$ tais que $\mathbf{u} \cdot \mathbf{v} = 0$ e $\mathbf{u} \cdot \mathbf{w} = 0$, ou seja, $x + 3y + z = 0$ e $-x + y = 0$. resolvendo este sistema obtemos$y = x$ e $z = -4x$. Portanto $\mathbf{u} = (x, x, -4x) = x(1, 1, -4)$.

3.2 Definição e exemplos

Inspirados pelas propriedades dos espaços euclidianos vamos definir espaços vetoriais reais em geral.

Definição 17. Um conjunto não vazio V é dito um **espaço vetorial real** (ou simplesmente, um **espaço vetorial**) quando estão definidas em V duas operações

$$\begin{array}{l} V \times V \longrightarrow V \\ (x, y) \longmapsto x + y \in V \end{array} \quad \text{e} \quad \begin{array}{l} \mathbb{R} \times V \longrightarrow V \\ (\alpha, y) \longmapsto \alpha y \in V, \end{array}$$

chamadas **adição** e **multiplicação por escalar**, respectivamente, satisfazendo as seguintes condições:

(EV1) $x + (y + z) = (x + y) + z, \forall\, x, y, z \in V$;

(EV2) $x + y = y + x, \forall\, x, y \in V$;

(EV3) existe um elemento, chamado vetor nulo e denotado por 0, tal que $x + 0 = x, \forall\, x \in V$;

(EV4) para cada $x \in V$, existe $y \in V$ (chamado oposto de x) tal que $x + y = 0$;

(EV5) $\alpha(\beta x) = (\alpha\beta)x, \forall\, \alpha, \beta \in \mathbb{R}, x \in V$;

(EV6) $(\alpha + \beta)x = \alpha x + \beta x, \forall\, \alpha, \beta \in \mathbb{R}, x \in V$;

(EV7) $\alpha(x + y) = \alpha x + \alpha y, \forall\, \alpha \in \mathbb{R}, x, y \in V$;

(EV8) $1x = x, \forall\, x \in V$.

Os elementos de V são chamados **vetores** e os números reais, **escalares**.

O conjunto $V = \mathbb{R}$, com as operações usuais de adição e multiplicação, é um espaço vetorial real: as propriedades acima são as propriedades associativas e comutativas da adição e multiplicação, elemento neutro para adição, elemento unidade para multiplicação, elemento oposto para adição e elemento inverso para multiplicação. Do mesmo modo, o conjunto $\mathbb{C}$ dos números complexos, com as operações usuais de adição e de multiplicação de escalar real por número complexo, é um espaço vetorial real.

Exemplo 22. O conjunto V^3 dos vetores geométricos no espaço (definidos por meio dos segmentos orientados), munido das operações usuais de adição de vetores e multiplicação de vetores escalar real (como indicadas na Figura 5), é um espaço vetorial real.

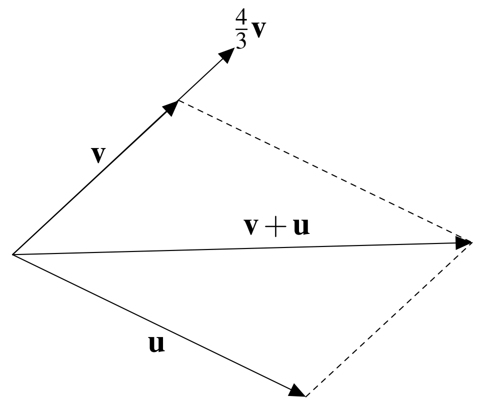

Figura 5

Seja $\mathbb{R}^2 = \{(x,y) : x,y \in \mathbb{R}\}$. Dados $\mathbf{u} = (x,y)$ e $\mathbf{v} = (s,t)$ em $\mathbb{R}^2$ e $\alpha \in \mathbb{R}$ definimos

$$\mathbf{u}+\mathbf{v} = (x+s, y+t)$$
$$\alpha\mathbf{u} = (\alpha x, \alpha y).$$

Com as operações assim definidas, $\mathbb{R}^2$ é um espaço vetorial. Verifiquemos, por exemplo, a condição (**EV** 1): dados $\mathbf{u} = (x,y)$, $\mathbf{v} = (s,t)$, $\mathbf{w} = (p,q) \in \mathbb{R}^2$, usando em cada componente, o fato que a adição de números reais é associativa, temos:

$$\begin{aligned}\mathbf{u}+(\mathbf{v}+\mathbf{w}) &= (x,y)+(s+p,t+q) = (x+(s+p), y+(t+q)) \\ &= ((x+s)+p, (y+t)+q) = (\mathbf{u}+\mathbf{v})+\mathbf{w}.\end{aligned}$$

É fácil ver que o vetor nulo em $\mathbb{R}^2$ é o par $(0,0)$, o oposto de $\mathbf{u} = (x,y)$ é o vetor $(-x,-y)$. As outras propriedades são facilmente verificadas. Os vetores de $\mathbb{R}^2$ podem ser representados geometricamente por segmentos orientados e a adição definida acima corresponde a adição de segmentos orientados, como na Figura 5.

Exemplo 23. O conjunto $\mathbb{R}^n = \{(x_1,\ldots,x_n) : x_1,\ldots,x_n \in \mathbb{R}\}$, com as operações definidas por (3.2) e (3.3), é um espaço vetorial real.

Exemplo 24. O conjunto $V = M_{m\times n}(\mathbb{R})$ das matrizes $m \times n$ é um espaço vetorial real com a adição e a multiplicação por escalar definida na seção 1.2.

Dados $\mathbf{u}$, $\mathbf{v} \in V$, definimos a **diferença** de $\mathbf{u}$ por $\mathbf{v}$ como sendo

$$\mathbf{u}-\mathbf{v} = \mathbf{u}+(-\mathbf{v}).$$

O teorema seguinte contém algumas propriedades que decorrem diretamente da definição de espaço vetorial

Teorema 18. Seja V um espaço vetorial. Então:

1. O vetor nulo é o único elemento neutro da adição em V, isto é, se $z \in V$ é tal que $z+u=u$, $\forall u \in V$, então $z=0$.

2. Para cada $x \in V$ existe um único elemento oposto $(-x)$.

3. Dados $a,b \in V$, e $\gamma \in \mathbb{R}$, $\gamma \neq 0$, a equação $\gamma x + a = b$ tem uma única solução, que é $x = \gamma^{-1}(b-a)$.

4. $\forall \alpha \in \mathbb{R}$, temos $\alpha \cdot 0 = 0$.

5. $\forall u \in V$, temos $0 \cdot u = 0$.

6. Se $\alpha \cdot u = 0$, então $\alpha = 0$ ou $u = 0$.

7. (Regra de sinais) $\forall \alpha \in \mathbb{R}$, $u \in V$ temos
$$(-\alpha)\cdot u = \alpha \cdot (-u) = -(\alpha u).$$

8. $\forall \alpha, \beta \in \mathbb{K}$, $u \in V$ temos $(\alpha - \beta)u = \alpha u - \beta u$.

9. $\forall \alpha \in \mathbb{K}$, $u, v \in V$ temos $\alpha(u-v) = \alpha u - \alpha v$.

Demonstração.

1. De fato, suponha 0 e $\widetilde{0}$ elementos neutros de V. Então $0 + \widetilde{0} = \widetilde{0}$ bem como $\widetilde{0} + 0 = 0$. De (**EV2**) concluímos que $\widetilde{0} = 0$.

2. De fato, suponha que $x+y=0$. Então $(-x)+x+y=(-x)$ de (**EV1**) temos $[(-x)+x]+y=(-x)$. Logo de (**EV4**) $0+y=(-x) \Rightarrow y=-x$.

3. $\gamma x + a = b$ e $\gamma \neq 0$ logo $\gamma x = b + (-a) = b - a$. Então $\gamma^{-1}\gamma x = x = \gamma^{-1}(b-a)$.

4. De fato, $0+0=0$, logo $\alpha 0 = \alpha(0+0) = \alpha 0 + \alpha 0 \Rightarrow 0 = \alpha 0 + (-\alpha 0) = \alpha 0 + \alpha 0 + (-\alpha 0) = \alpha 0 + (\alpha 0 - \alpha 0) = \alpha 0$.

5. Segue de modo análogo a 4.

6. Suponha $\alpha u = 0$. Se $\alpha = 0$ segue de 5. que $\alpha u = 0$. Se $\alpha \neq 0$ então $\alpha^{-1}(\alpha u) = \alpha^{-1} 0 = 0$ pelo item 4.. Logo $u = (\alpha^{-1}\alpha)u = 0$.

7. Mostraremos inicialmente que $-(\alpha u) = (-\alpha)u$.

   $\alpha u + (-\alpha)u \stackrel{(\mathbf{EV6})}{=} (\alpha - \alpha)u = 0u \stackrel{5.}{=} 0$.

   Por outro lado $\alpha u + \alpha(-u) \stackrel{(\mathbf{EV7})}{=} \alpha(u+(-u)) = \alpha 0 \stackrel{4.}{=} 0$. Logo $\alpha(-u) = (-\alpha)u = -(\alpha u)$.

8. $(\alpha - \beta)u = (\alpha + (-\beta))u \stackrel{(\mathbf{EV6})}{=} \alpha u + (-\beta)u \stackrel{7.}{=} \alpha u - \beta u$.

9. $\alpha(u-v)=\alpha(u+(-v))=\alpha u+\alpha(-v)\overset{7.}{=}\alpha u+(-\alpha v)=\alpha u-\alpha v.$

$\square$

Observação 6. Em muitas situações, é conveniente considerar multiplicação de vetores por escalar complexo. Quando, na Definição 17 a multiplicação $(\alpha,x)\mapsto\alpha x$ for definida para todo $\alpha\in\mathbb{C}$, diremos que V é um **espaço vetorial complexo**. Quando quisermos nos referir indistintamente a um espaço vetorial real ou um espaço vetorial complexo usaremos a expressão **espaço vetorial sobre** $\mathbb{K}$.

Exemplo 25. Pelas mesmas razões mencionadas anteriormente, o conjunto $\mathbb{C}$ dos números complexos, com as operações usuais de adição e multiplicação, é um espaço vetorial complexo.

3.3 Subespaços vetoriais

Definição 18. Um subconjunto U de um espaço vetorial V é dito um **subespaço vetorial de** V se:

$$\text{dados } u,v\in U \text{ e } \alpha\in\mathbb{K}, \text{ temos } u+v\in U \text{ e } \alpha u\in U.$$

Como V é um espaço vetorial, as propriedades (**EV1**) a (**EV8**) da Definição 17, estão satisfeitas para todos elementos de V; como $U\subset V$, elas estão satisfeitas também para todos elementos de U. Logo, em particular, U também é um espaço vetorial.

Se V é um espaço vetorial qualquer, então os subconjuntos $U=\{0\}$ e $U=V$ são chamados subespaços vetoriais triviais de V.

Exemplo 26. O conjunto $U=\{(x,u):x-2y=0\}$ é um subespaço vetorial de $\mathbb{R}^2$. De fato, em primeiro lugar, U é não vazio, pois, por exemplo, $(0,0)\in U$. Além disso, se $(x,y),(s,t)\in U$ e $\alpha\in\mathbb{R}$, temos $x=2y$ e $s=2t$, donde $x+s=2(y+t)$ e $\alpha x=2\alpha y$ e portanto $(x+s,y+t)\in U$ e $(\alpha x,\alpha y)\in U$.

Da mesma maneira, mostramos que qualquer reta passando pela origem é um subespaço de $\mathbb{R}^2$.

Exemplo 27. Em $V=\mathbb{R}^3$, os seguintes subconjuntos:

- a origem $\{(0,0,0)\}$,
- o próprio $\mathbb{R}^3$,
- as retas passando pela origem $(0,0,0)$,
- os planos contendo a origem

são subespaços vetoriais. Pode-se mostrar que esses são os únicos subespaços de $\mathbb{R}^3$.

Exemplo 28. Seja $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$. O conjunto U de todas as soluções $v = (x_1, x_2, \ldots, x_n)^T$ do sistema linear homogêneo

$$A\mathbf{v} = \mathbf{0}. \tag{3.14}$$

é um subespaço vetorial de $\mathbb{R}^n$.

É claro que a n-upla $\mathbf{0} = (0, 0, \ldots,)^T$ é solução de (3.14), portanto pertence a U. Se $\mathbf{v}_1, \mathbf{v}_2 \in U$ e $\alpha \in \mathbb{R}$, temos $A\mathbf{v}_1 = 0$ e $A\mathbf{v}_2 = 0$, donde

$$A(\mathbf{v}_1 + \mathbf{v}_2) = A\mathbf{v}_1 + A\mathbf{v}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0};$$

portanto $\mathbf{v}_1 + \mathbf{v}_2 \in U$. Analogamente, $A(\alpha\mathbf{v}_1) = \alpha A\mathbf{v}_1 = \alpha\mathbf{0} = \mathbf{0}$; logo $\alpha\mathbf{v}_1 \in U$.

De modo geral, todo conjunto solução de um sistema linear homogêneo no $\mathbb{R}^n$ é um subespaço do $\mathbb{R}^n$.

Exemplo 29. Seja $V = M_{m\times n}(\mathbb{R})$ e seja $U = A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$; $a_{ij} = 0$ para $i \geq j$, então U é subespaço de V e também é conhecido como o espaço das matrizes triangulares superiores.

3.4 Combinações lineares

Sejam $u_1, \ldots, u_n \in V$, $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$. Como nos espaços euclideanos denominamos o vetor

$$v = \alpha_1 u_1 + \cdots + \alpha_n u_n$$

de **combinação linear** de $u_1, \ldots, u_n$.

Por exemplo, conforme visto nos Exemplos 15 e 16.

Exemplo 30. O vetor $(-1, 2, 0)$ é combinação linear de $(1, 3, 1)$, $(-2, -1, -1)$ e $(1, -2, -1)$: procuremos α, β, γ tais que

$$\alpha(1, 3, 1) + \beta(-2, -1, -1) + \gamma(1, -2, -1) = (-1, 2, 0);$$

então α, β, γ devem satisfazer o sistema de equações

$$\begin{cases} 1\alpha - 2\beta + \gamma = -1 \\ 3\alpha - \beta - 2\gamma = 2 \\ \alpha - \beta - \gamma = 0 \end{cases}$$

Como esse sistema é equivalente ao sistema

$$\begin{cases} \alpha - 2\beta + \gamma = -1 \\ 5\beta - 5\gamma = 5 \\ -\gamma = 0 \end{cases}$$

ele terá a solução, $\alpha = 1$, $\beta = 1$, $\gamma = 0$, segue-se que $(-1, 2, 0)$ é combinação linear de $(1, 3, 1)$, $(-2, -1, -1)$ e $(1, -2, -1)$.

Teorema 19. Seja V um espaço vetorial e sejam $u_1, \dots, u_n \in V$. O conjunto U de todas combinações lineares de $u_1, \dots, u_n$ é um subespaço vetorial de V.

Demonstração. Em primeiro lugar, U é não vazio, pois o vetor nulo é combinação linear de $u_1, \dots, u_n$; de fato, $0 = 0u_1 + \cdots + 0u_n$. Além disso, dados $v, w \in U$,

$$v = \alpha_1 u_1 + \cdots + \alpha_n u_n, \quad w = \beta_1 u_1 + \cdots + \beta_n u_n,$$

e $\alpha \in \mathbb{R}$, temos

$$\begin{aligned} v + w &= (\alpha_1 + \beta_1)u_1 + \cdots + (\alpha_n + \beta_n)u_n \\ \alpha v &= (\alpha\alpha_1)u_1 + \cdots + (\alpha\alpha_n)u_n \end{aligned}$$

o que mostra que $v + w$ e αv são combinações lineares de $u_1, \dots, u_n$, ou seja, $v + w \in U$ e $\alpha v \in U$. Logo, U é um subespaço vetorial de V. $\square$

O subespaço W dado no Teorema 19 chama-se **subespaço gerado** por $u_1, \dots, u_n$ e é denotado por $[u_1, \dots, u_n]$; os vetores $u_1, \dots, u_n$ são então chamados **geradores** de W. Um espaço vetorial é dito **finitamente gerado** quando possui um número finito de geradores. Neste texto, estaremos interessados somente nos espaços vetoriais finitamente gerados.

Exemplo 31. Considere em $\mathbb{R}^3$ os vetores $\mathbf{a} = (1,0,0)$, $\mathbf{b} = (0,1,0)$ e $\mathbf{c} = (1,1,0)$. Então $[\mathbf{a}] = \{(x,0,0) : x \in \mathbb{R}\} =$ eixo x, $[\mathbf{c}] = \{(y,y,0) : y \in \mathbb{R}\} =$ reta passando pela origem paralela a $\mathbf{c}$, $[\mathbf{a},\mathbf{c}] = [\mathbf{b},\mathbf{a}] = [\mathbf{a},\mathbf{b},\mathbf{c}] = \{(x,y,0) : y \in \mathbb{R}^3 : x, y, \in \mathbb{R}\}$ é o plano $z = 0$.

Exemplo 32. Vamos encontrar um conjunto de geradores para o subespaço $U = \{(x,y,z,w) \in \mathbb{R}^4 : x + y - z = 0, \quad y - z + w = 0\}$.
Que pode ser escrito como

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

Note que esse sistema já está na forma escalonada por linhas. Temos $(x,y,z,w) \in U \iff z = x + y, \quad w = x$. Portanto

$$(x,y,z,w) = (x,y,x+y,x) = x(1,0,1,1) + y(0,1,1,0).$$

Logo $U = [(1,0,1,1) + (0,1,1,0)]$.

Exemplo 33. O espaço vetorial $\mathbb{R}^n$ é finitamente gerado: todo vetor $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ é combinação linear dos vetores

$$\mathbf{e}_1 = (1,0,\dots,0), \dots, \mathbf{e}_n = (0,\dots,0,1).$$

De fato, temos $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$

A definição de subespaço gerado estende-se aos seguintes casos:

1. Se $S = 0$, então $[S] = \{0\}$.

2. Se S for o conjunto infinito isto é, $S = \{v_i;\ i = 1, 2, \ldots\}$ definimos o subespaço gerado $[S]$ do seguinte modo: $[S]$ é o conjunto de todas as combinações lineares de elementos de S, isto é, $u \in [S]$ se, e somente se, existem $v_1, \ldots, v_r \in S$, $\alpha_1, \ldots, \alpha_r \in \mathbb{K}$ tais que $u = \alpha_1 v_1 + \cdots + \alpha_r v_r$.

3.5 Dependência linear

Sejam V um espaço vetorial e $S = \{v_1, \ldots, v_n\} \subset V$. Dizemos que os vetores $v_1, \ldots, v_n$ são **linearmente dependentes**, ou que S é um conjunto **linearmente dependente** (escreveremos abreviadamente LD) quando existem escalares *não todos nulos* $\alpha_1, \ldots, \alpha_n$ tais que

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0. \tag{3.15}$$

Caso contrário, isto é, se uma igualdade do tipo $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$ só for possível quando $\alpha_1 = \cdots = \alpha_n = 0$, dizemos que os vetores $v_1, \ldots, v_n$ são **linearmente independentes**, ou que S é um conjunto **linearmente independente** (abreviadamente **LI**).

Notemos que, quaisquer que sejam os vetores $v_1, \ldots, v_n$, os escalares $\alpha_1 = 0, \ldots, \alpha_n = 0$ satisfazem a igualdade (3.15). O que realmente interessa nessa definição é saber se também é possível escrever (3.15) com escalares não todos nulos (quando dizemos que $v_1, \ldots, v_n$ são **LD**) ou se **a única** maneira possível de escrever (3.15) é pondo $\alpha_1 = 0, \ldots, \alpha_n = 0$ (neste caso, $v_1, \ldots, v_n$ são **LI**).

Exemplo 34. Em $\mathbb{R}^3$ os vetores $(2, -1, 3)$, $(-1, 0, -2)$ e $(2, -3, 1)$ são LD, pois podemos escrever

$$3(2, -1, 3) + 4(-1, 0, -2) - (2, -3, 1) = (0, 0, 0).$$

Exemplo 35. Em $\mathbb{R}^4$ os vetores $(2, 2, 3, 4)$, $(0, 5, -3, 1)$ e $(0, 0, 4, -2)$ são LI

De fato, se escrevermos

$$\alpha(2, 2, 3, 4) + \beta(0, 5, -3, 1) + \gamma(0, 0, 4, -2) = (0, 0, 0, 0),$$

temos

$$(2\alpha, 2\alpha + 5\beta, 3\alpha - 3\beta + 4\gamma, 4\alpha + \beta - 2\gamma) = (0, 0, 0, 0),$$

ou seja,

$$\begin{cases} 2\alpha = 0 \\ 2\alpha + 5\beta = 0 \\ 3\alpha - 3\beta + 4\gamma = 0 \\ 4\alpha + \beta - 2\gamma = 0 \end{cases}$$

que implica $\alpha = \beta = \gamma = 0$.

Exemplo 36. Os vetores $\mathbf{e_1} = (1, 0, \ldots, 0), \ldots, \mathbf{e_n} = (0, \ldots, 0, 1)$ são linearmente independentes. De fato, se os números $x_1, \ldots, x_n$ são tais que $x_1\mathbf{e_1} + \ldots, x_n\mathbf{e_n} = 0$, temos $(x_1, \ldots, x_n) = (0, \ldots, 0)$, ou seja, $x_1 = 0, \ldots, x_n = 0$, donde concluímos que os vetores $e_1, \ldots, e_n$ são linearmente independentes.

Exemplo 37. Se um dos vetores $v_1, \ldots, v_n$ for combinação linear dos outros, então $v_1, \ldots, v_n$ são LD.

Seja v_k o vetor que é combinação linear dos demais:

$$v_k = \alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1} + \alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n.$$

Podemos então escrever

$$\alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1} + (-1) v_k + \alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n = 0.$$

Como o coeficiente de v_k é não nulo, temos que $v_1, \ldots, v_n$ são LD.

A recíproca desse fato também é verdadeira: se uma sequência de vetores $v_1, \ldots, v_n \in V$ é LD e se $v_1 \neq 0$, então ao menos um desses vetores é combinação linear dos precedentes. Mais precisamente.

Teorema 20. Se $v_1, \ldots, v_n \in V$ são vetores LD e se $v_1 \neq 0$, então existe $k \geq 2$ tal que v_k é combinação linear de $v_1, \ldots, v_{k-1}$.

Demonstração. Como $v_1, \ldots, v_n$ são linearmente dependentes, existem escalares não todos nulos $\alpha_1, \ldots, \alpha_n$ tais que

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0. \tag{3.16}$$

Seja k o maior dentre esses índices tal que $\alpha_k \neq 0$; como $v_1 \neq 0$,. Temos que $k \geq 2$, de fato, se tivéssemos $\alpha_1 \neq 0$ e $\alpha_2 = 0, \ldots, \alpha_n \geq 0$, a igualdade (3.16) ficaria $\alpha_1 v_1 = 0$, o que é impossível, pois $\alpha_1 \neq 0$ e $v_1 \neq 0$). Visto isso, da codição sobre k temos $\alpha_{k+1} = \cdots = \alpha_n = 0$, podemos então escrever a igualdade (3.16) na forma

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0.$$

Como $\alpha_k \neq 0$ segue que

$$v_k = -\frac{\alpha_1}{\alpha_k} v_1 - \cdots - \frac{\alpha_{k-1}}{\alpha_k} v_{k-1},$$

o que mostra que v_k é combinação linear dos vetores $v_1, \ldots, v_{k-1}$. □

Corolário 8. Todo conjunto finito que contém um subconjunto LD é LD, isto é, se os vetores $v_1, \ldots, v_p \in V$ são LD e $v_{p+1}, \ldots, v_n$ são vetores quaisquer em V, então $v_1, \ldots, v_p, v_{p+1}, \ldots, v_n$ são LD.

Demonstração. Como os vetores $v_1,\ldots,v_p$ são LD, um deles, digamos, v_k, $k<p$, é combinação linear de $v_1,\ldots,v_{k-1}$; então é claro que v_k, também se escreve como combinação linear de $v_1,\ldots,v_n$. $\square$

Corolário 9. Se os vetores $v_1,\ldots,v_n$ são LI e $v_1,\ldots,v_n,x$ são LD, então x é combinação linear de $v_1,\ldots,v_n$.

Demonstração. Como nenhum dos v_j pode ser combinação linear dos precedentes (pois os vetores $v_1,\ldots,v_n$ são LI), segue-se que x é combinação linear de $v_1,\ldots,v_n$. $\square$

Observação 7. O conceito de independência linear também pode ser definido para conjuntos infinitos de vetores: um conjunto S é dito **linearmente independente** quando todo subconjunto finito de S for LI (de acordo com a definição acima).

3.6 Base e dimensão

Uma **base** de um espaço vetorial V é um conjunto de vetores LI que geram V.

Exemplo 38. O conjunto $\{\mathbf{u}=(1,1),\mathbf{v}=(-1,0)\}$ é base de $\mathbb{R}^2$

Os vetores $\mathbf{u}$ e $\mathbf{v}$ geram $\mathbb{R}^2$. Dado $\mathbf{w}=(a,b)\in\mathbb{R}^2$, procuramos escalaras x,y tais que $x\mathbf{u}+y\mathbf{v}=\mathbf{w}$, ou seja $(x-y,x)=(a,b)$. Então x,y precisam ser soluções do sistema

$$\begin{cases} x-y=a \\ x=b \end{cases} \tag{3.17}$$

Portando $x=b, y=b-a$. Logo, $\mathbf{w}=b\mathbf{u}+(b-a)\mathbf{v}$.

Os vetores $\mathbf{u}$ e $\mathbf{v}$ são LI pois, se $x\mathbf{u}+y\mathbf{v}=0$, então, x e y são soluções do sistema (3.17) com $a=b=0$, e portanto, $x=y=0$. Logo, $\mathbf{u}$ e $\mathbf{v}$ são LI.

Exemplo 39. O conjunto $B=\{(1,0,0),(0,1,0),(0,0,1)\}$ é base de $\mathbb{R}^3$: de fato, já vimos que os vetores de B são LI e geram $\mathbb{R}^3$.

Mais geralmente, temos

Exemplo 40. O conjunto $B=\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$, em que $\mathbf{e}_1=(1,0,\ldots,0),\mathbf{e}_2=(0,1,\ldots,0),\ldots,\mathbf{e}_n=(0,\ldots,0,1)$, é base de $\mathbb{R}^n$.

Sejam V um espaço vetorial, $B=\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ uma base de V e seja $x\in V$. Como $\mathbf{e}_1,\ldots,\mathbf{e}_n$ geram V, existem escalares $\alpha_1,\ldots,\alpha_n$ tais que

$$x=\alpha_1\mathbf{e}_1+\cdots+\alpha_n\mathbf{e}_n. \tag{3.18}$$

Além disso, como $\mathbf{e}_1,\ldots,\mathbf{e}_n$ são LI, os escalares são determinados de modo único, no sentido que, se

$$x=\beta_1\mathbf{e}_1+\cdots+\beta_n\mathbf{e}_n,$$

então

$$\alpha_1 = \beta_1, \cdots, \alpha_n = \beta_n.$$

Reciprocamente, se todo vetor $x \in V$ se escreve de modo único como combinação linear de $\mathbf{e}_1, \ldots, \mathbf{e}_n$, então eles são geradores de V. Além disso, como o vetor nulo se escreve de modo único como combinação linear de $\mathbf{e}_1, \ldots, \mathbf{e}_n$, segue-se que esses vetores são LI. Logo, $B = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ é base de V.

Exemplo 41. Seja $V = M_{m\times n}(\mathbb{R})$. Então $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ é base de V.

Esses fatos mostram a importância do conceito de base e, por isso, vamos enunciá-lo como um teorema.

Teorema 21. Seja $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ uma base de um espaço vetorial V. Então todo $x \in V$ se escreve de modo único como combinação linear de $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Reciprocamente, se todo vetor $x \in V$ se escreve de modo único como combinação linear de $\mathbf{e}_1, \ldots, \mathbf{e}_n$, então $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ é uma base de V.

Os números $\alpha_1, \alpha_2 \ldots, \alpha_n$ chama-se **coordenadas** de x em relação à base B. A partir deste ponto, é conveniente considerar uma base como sendo um conjunto ordenado de vetores: isto significa que neste ponto é importante a ordem em que os vetores $\mathbf{e}_1, \ldots, \mathbf{e}_n$ são relacionados (com isto queremos dizer, por exemplo, que $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ e $\mathbf{e}_2, \mathbf{e}_1, \ldots, \mathbf{e}_n$ são bases distintas de V). Podemos então escrever os escalares de (3.18) como uma matriz coluna (ou como uma n-upla, se for conveniente), chamada **matriz de coordenadas** de x na base B.

$$\begin{bmatrix} x \end{bmatrix}_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}. \tag{3.19}$$

Deve ficar entendido que α_1 é o coeficiente de $\mathbf{e}_1, \ldots, \alpha_n$ é o coeficiente de $\mathbf{e}_n$ em (3.18). Para simplificar a notação vamos indicar a matriz em (3.19) por $[\alpha_1, \ldots, \alpha_n]^T$: o símbolo T indica a transposta da matriz.

Exemplo 42. Consideremos em $\mathbb{R}^4$ os vetores $\mathbf{v}_1 = (1,0,-1,0)$, $\mathbf{v}_2 = (0,0,0,1)$, $\mathbf{v}_3 = (0,0,1,2)$, $\mathbf{v}_4 = (0,1,0,1)$ e $\mathbf{w} = (a,b,c,d)$.
Dado $\mathbf{w} = (a,b,c,d) \in \mathbb{R}^4$, procuremos $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ tais que

$$\alpha\mathbf{v}_1 + \beta\mathbf{v}_2 + \gamma\mathbf{v}_3 + \delta\mathbf{v}_4 = \mathbf{w}, \tag{3.20}$$

isto é,

$$(\alpha, \delta, -\alpha+\gamma, \beta+2\gamma+\delta) = (a,b,c,d).$$

Dessa igualdade temos $\alpha = a, \beta = d - 2a - b - 2c, \gamma = a + c$ e $\delta = b$, o que mostra que todo $\mathbf{w} \in \mathbb{R}^4$ se escreve, de modo único, como combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$, ou seja, os vetores $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ geram $\mathbb{R}^4$.

Além disso, como a única solução de (3.20) para $\mathbf{w} = (0,0,0,0)$ é $(\alpha, \beta, \gamma, \delta) = (0,0,0,0)$, temos que os vetores $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ são LI. Logo. B é base de $\mathbb{R}^4$.

É claro que, como podemos escrever

$$(a,b,c,d) = a(1,0,0,0) + b(0,1,0,0) + c(0,0,1,0) + d(0,0,0,1);$$

temos $[\mathbf{w}]_C = [a,b,c,d]^T$.

Para obter as coordenadas de $\mathbf{w}$ em relação à base B, procuramos x, y, z, t tais que

$$(a,b,c,d) = x(1,0,-1,0) + y(0,0,0,1) + z(0,0,1,2) + t(0,1,0,1).$$

Então x, y, z, t devem satisfazer

$$x = a, \quad t = b, \quad -x + z = c, \quad y + 2z + t = d,$$

donde $x = a$, $y = d - 2a - 2c - b$, $z = c + a$, $t = b$. Logo,

$$[\mathbf{w}]_B = [a, d - 2a - 2c - b, c + a, b]^T.$$

Teorema 22. Suponhamos que o espaço vetorial V tenha uma base com n vetores. Então qualquer subconjunto de V contendo mais de n vetores é LD.

Demonstração. Seja $B = \{v_1, \ldots, v_n\}$ uma base de V e sejam $x_1, \ldots, x_m$ vetores quaisquer em V, com $m > n$. Suponhamos por absurdo que os vetores $x_1, \ldots, x_n$ são LI. Como $V = [v_1, \ldots, v_n]$, temos que x_1 é combinação linear de $v_1, \ldots, v_n$:

$$x_1 = \alpha_1 v_1 + \cdots + \alpha_n v_n;$$

nessa representação, pelo menos um dos escalares α_j é não nulo (se todos os α_j fosses nulos, teríamos $x_1 = 0$, o que contradiz o fato que os vetores $x_1, \ldots, x_n$ são LI); para simplificar a notação, vamos supor que $\alpha_n \neq 0$; então

$$v_n = \frac{1}{\alpha_n}(x_1 - \alpha_1 v_1 - \cdots - \alpha_{n-1} v_{n-1})$$

donde segue que $V = [x_1, v_1, \ldots, v_{n-1}]$. Agora, como $x_2 \in V$, temos

$$x_2 = \gamma x_1 + \beta_1 v_1 + \cdots + \beta_{n-1} v_{n-1},$$

em que, ao menos um dos β_j é não nulo (se todos os β_j fossem nulos, teríamos $x_2 = \gamma x_1$, o que contradiz o fato que os vetores $x_1, \ldots, x_n$ são LI); para simplificar a notação, vamos supor que $\beta_{n-1} \neq 0$; então

$$v_{n-1} = \frac{1}{\beta_{n-1}}(x_2 - \gamma x_1 - [\beta_1 v_1 + \cdots + \beta_{n-2} v_{n-2}])$$

donde temos $V = [x_1, x_2, v_1, \ldots, v_{n-2}]$. Repetindo esse procedimento, chegaremos a $V = [x_1, x_2, \ldots, x_n]$. Como $x_{n+1}, \ldots, x_m \in V$, esses vetores são combinações lineares de $x_1, x_2, \ldots, x_n$. Pelo Corolário 8 $x_1, \ldots, x_n, x_{n+1}, \ldots, x_m$ são LD. O que nos dá a contradição. □

Teorema 23. Seja V um espaço vetorial finitamente gerado. Então todas as bases de V têm a mesma quantidade de elementos.

Demonstração. Sejam $\{e_1, \ldots, e_n\}$ e $\{v_1, \ldots, v_p\}$ duas bases de V. Como $V = [e_1, \ldots, e_n]$ e $\{v_1, \ldots, v_p\}$ é um conjunto LI em V temos, pelo Teorema 22, $p \leq n$. Trocando os papéis de $e_1, \ldots, e_n$ e $v_1, \ldots, v_p$, obtemos $n \leq p$. Logo, $n = p$. □

Definição 19. Seja V um espaço vetorial finitamente gerado: o número de vetores de uma base qualquer de V chama-se **dimensão** de V. Se um espaço vetorial V não é finitamente gerado, diz-se que ele tem **dimensão infinita**.

Exemplo 43. $\dim \mathbb{R}^n = n$

Apresentamos a seguir um método prático para estudar a dependência linear em $\mathbb{R}^n$. O método baseia-se no seguinte lema.

Lema 5. Suponhamos $W = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m] \subset \mathbb{R}^n$. Definamos

$$\mathbf{v}_1 = \mathbf{u}_1, \quad \mathbf{v}_2 = \mathbf{u}_2 + k_1\mathbf{v}_1, \ldots, \mathbf{v}_m = \mathbf{u}_m + k_{m-1}\mathbf{v}_1, \tag{3.21}$$

em que $k_1, \ldots, k_{m-1} \in \mathbb{R}$. Então $W = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m]$.

Demonstração. De fato, a partir das igualdades (3.21) é fácil ver que cada vetor $\mathbf{u}_j$, $j = 1, \ldots, m$ é combinação linear de $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Segue que $W = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m]$. □

É óbvio que se $W = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m]$ então $W = [\mathbf{u}_2, \mathbf{u}_1, \ldots, \mathbf{u}_m]$, o mesmo valendo para qualquer permutação de vetores $\mathbf{u}_i, \mathbf{u}_j$. Uma consequência imediata destes fatos é um método fácil para decidir se um dado conjunto de vetores é LI ou LD. Formamos a matriz A de ordem $m \times n$ cujos vetores linhas são $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ e escalonamos a matriz A. O processo de escalonamento consiste precisamente em efetuar convenientemente as operações indicadas em (3.21) ou por permutações de linhas.

Exemplo 44. Vamos verificar se os se os vetores $\mathbf{u}_1 = (1, 2, 3, -5)$, $\mathbf{u}_2 = (2, 4, 1, 2)$, $\mathbf{u}_3 = (1, 3, 4, 0)$ e $\mathbf{u}_4 = (3, 5, 8, -10)$ são LI ou LD.

Formemos a matriz A cujos vetores linhas são $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ e $\mathbf{u}_4$ e encontremos sua correspondente forma escalonada por linhas

$$A = \begin{pmatrix} 1 & 2 & 3 & -5 \\ 2 & 4 & 1 & 2 \\ 1 & 3 & 4 & 0 \\ 3 & 5 & 8 & -10 \end{pmatrix}, \qquad \hat{A} = \begin{pmatrix} 1 & 2 & 3 & -5 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & -5 & 12 \\ 0 & 0 & 0 & 10 \end{pmatrix}$$

Como todas as linhas da matriz escalonada $\hat{A}$ são não nulas, o correspondente sistema $\sum_{i=1}^{4} \alpha_i \hat{\mathbf{u}}_i = \mathbf{0}$ tem solução única $(\alpha_1, \cdots, \alpha_4)^T = (0,0,0,0)^T$ logo seus vetores linhas são LI. Pelo Lema 5 os vetores $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ também são LI.

Exemplo 45. Vamos verificar se os vetores $\mathbf{u}_1 = (1,2,1,3)$, $\mathbf{u}_2 = (4,0,2,-1)$, $\mathbf{u}_3 = (3,1,-2,3)$ e $\mathbf{u}_4 = (-3,2,-1,4)$ são LI ou LD.

Formemos a matriz A cujos vetores linhas são $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ e $\mathbf{u}_4$ e encontremos sua correspondente forma escalonada por linhas $\hat{A}$.

$$A = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 4 & 0 & 2 & -1 \\ 3 & 1 & -2 & 3 \\ -3 & 2 & -1 & 4 \end{pmatrix}, \qquad \hat{A} = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & -8 & -2 & -13 \\ 0 & 0 & -\frac{15}{4} & \frac{17}{8} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Como a matriz escalonada $\hat{A}$ tem apenas 3 linhas não nulas o correspondente sistema $\sum_{i=1}^{4} \alpha_i \hat{\mathbf{u}}_i = \mathbf{0}$ não tem solução única, já que tem um grau de liberdade. Portando os vetores $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ são LD.

Teorema 24. Seja V um espaço vetorial, $\dim V = n$, e sejam $v_1, \ldots, v_p$ (com $p < n$) vetores LI em V. Então existem $n - p$ vetores $v_{p+1}, \ldots, v_n$ em V tais que $v_1, \ldots, v_n$ é base de V.

Demonstração. Vamos supor por absurdo que para qualquer $v_{p+1} \in V$ tal que $v_{p+1} \notin [v_1, \ldots, v_p]$ tenhamos $\{v_1, v_2, v_3, \ldots, v_p, v_{p+1}\}$ seja linearmente dependente. Então $\sum_{1}^{p+1} \alpha_i v_i = \vec{0}$ para algum $\alpha_i \neq 0$. Então v_i é combinação linear dos demais. Se $i = p+1$ é absurdo, pois $v_{p+1} \notin [v_1, \ldots, v_p]$. Logo $i \in \{1, 2, \ldots, p\}$ logo $v_j = \sum_{j=1, j\neq i}^{p+1} \beta_j v_j$. Novamente $\beta_{p+1} = 0$ caso contrário $v_{p+1} \in [v_1, \cdots, v_p]$.

Portanto $\beta_{p+1} = 0 \rightarrow v_j \in [v_1, \cdots, v_{p+1}]$ o que contradiz $\{v_1, \ldots, v_p\}$ é LI.

□

Exemplo 46. Vamos retomar o exemplo 45. Se completarmos o conjunto formado pelos vetores $\mathbf{u}_1 = (1,2,1,3)$, $\mathbf{u}_2 = (4,0,2,-1)$ e $\mathbf{u}_3 = (3,1,-2,3)$ com o vetor $\mathbf{u}_5 = (0,0,0,1)$ teremos uma base no $\mathbb{R}^4$.

3.7 Bases ortogonais em $\mathbb{R}^n$

Consideremos em $\mathbb{R}^n$ o seu produto interno usual: se $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$, então

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n.$$

Teorema 25. Se $X = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ é um conjunto de vetores ortogonais não nulos, então X é um conjunto LI.

Demonstração. Suponhamos que os números $\alpha_1, \alpha_2, \ldots, \alpha_m$ são tais que

$$\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \ldots, \alpha_m \mathbf{u}_m = 0. \tag{3.22}$$

Efetuando o produto escalar dos dois membros de (3.22) com $\mathbf{u}_1$ e notando que $\mathbf{u}_1 \cdot \mathbf{u}_1 = \|\mathbf{u}_1\|^2$ e $\mathbf{u}_j \cdot \mathbf{u}_1 = 0$; $\forall j \neq 1$, obtemos $\alpha_1 \|\mathbf{u}_1\|^2 = 0$. Como $\|\mathbf{u}_1\|^2 \neq 0$, temos $\alpha_1 = 0$. Analogamente obtemos $\alpha_2 = 0, \alpha_3 = 0, \ldots, \alpha_m = 0$. Logo, os vetores $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ são LI. □

Uma base $B = \{\mathbf{u}_1, , \ldots, \mathbf{u}_n\}$ de $\mathbb{R}^n$ formada por vetores 2 a 2 ortogonais é chamada uma **base ortogonal**. Se além disso, todos os vetores forem unitários (isto é, $\|\mathbf{u}_j\| = 1, \forall\ j$) dizemos que B é uma **base ortonormal.**

Exemplo 47. A base canônica de $\mathbb{R}^n$ é ortonormal.

Exemplo 48. Em $\mathbb{R}^3$, o conjunto $\left\{(1,0,0), \left(0, \frac{1}{2}, \frac{\sqrt{3}}{2}\right), \left(0, \frac{\sqrt{3}}{2}, -\frac{1}{2}\right)\right\}$ é uma base ortonormal.

O próximo teorema mostra que fica mais simples obter as coordenadas de um vetor quando trabalhamos com uma base ortonormal.

Teorema 26. Se $B = \{v_1, v_2, \ldots, v_n\}$ é uma base ortonormal de $\mathbb{R}^n$, então, para todo $\mathbf{x} \in \mathbb{R}^n$, temos

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{v}_1)\, \mathbf{v}_1 + (\mathbf{x} \cdot \mathbf{v}_2)\, \mathbf{v}_2 + \cdots + (\mathbf{x} \cdot \mathbf{v}_n)\, \mathbf{v}_n, \tag{3.23}$$

isto é, se $\mathbf{x} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n$, então $\alpha_1 = (\mathbf{x} \cdot \mathbf{v}_1), \cdots, \alpha_n = (\mathbf{x} \cdot \mathbf{v}_n)$.

Demonstração. Como B é base de $\mathbb{R}^n$ existem números $\alpha_1, \alpha_2, \ldots, \alpha_n$ tais que

$$\mathbf{x} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n \tag{3.24}$$

Efetuando o produto escalar dos dois membros de (3.24) com $\mathbf{v}_1$ e notando que $\mathbf{v}_1 \cdot \mathbf{v}_1 = 1$ e $\mathbf{v}_j \cdot \mathbf{v}_1 = 0, \forall\ j \neq 1$, obtemos $\mathbf{x} \cdot \mathbf{v}_1 = \alpha_1$. De modo análogo, obtemos $\mathbf{x} \cdot \mathbf{v}_2 = \alpha_2, \ldots, \mathbf{x} \cdot \mathbf{v}_n = \alpha_n$. □

Teorema 27. Sejam $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ um conjunto ortonormal em $\mathbb{R}^n$ e $\mathbf{x} \in \mathbb{R}^n$. Mostrar que o vetor $\mathbf{z} = \mathbf{x} - (\mathbf{x} \cdot \mathbf{v}_1)\, \mathbf{v}_1 - \cdots - (\mathbf{x} \cdot \mathbf{v}_p)\, \mathbf{v}_p$ é ortogonal a cada um dos vetores $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

Demonstração. De fato, como $\mathbf{v}_j \cdot \mathbf{v}_j = 1$ e $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, se $i \neq j$, temos

$$\mathbf{z} \cdot \mathbf{v}_j = \mathbf{x} \cdot \mathbf{v}_j - (\mathbf{x} \cdot \mathbf{v}_1)\left(\mathbf{v}_1 \cdot \mathbf{v}_j\right) - \cdots - (\mathbf{x} \cdot \mathbf{v}_p)\left(\mathbf{v}_p \cdot \mathbf{v}_j\right) = \mathbf{x} \cdot \mathbf{v}_j - \mathbf{x} \cdot \mathbf{v}_j = 0.$$

□

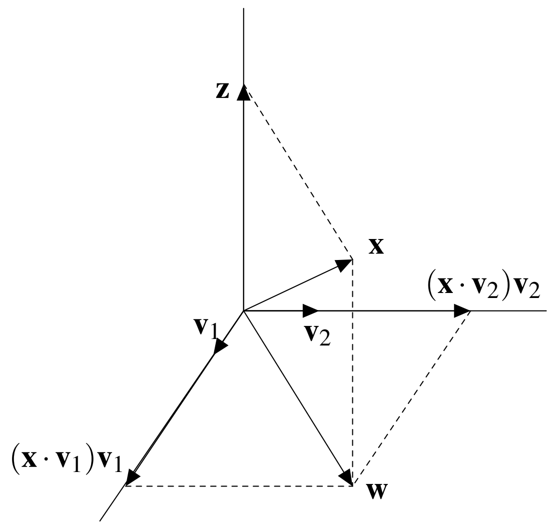

Figura 6

O vetor $\mathbf{w} = (\mathbf{x}\cdot\mathbf{v}_1)\,\mathbf{v}_1 + \cdots + (\mathbf{x}\cdot\mathbf{v}_p)\,\mathbf{v}_p$, dado no Teorema 27, chama-se **projeção ortogonal** de $\mathbf{x}$ sobre o subespaço $[\mathbf{v}_1,\ldots,\mathbf{v}_p]$.

Usando o Teorema 27 podemos construir uma base ortonormal de um subespaço vetorial W de $\mathbb{R}^n$ a partir de uma dada base de W. Seja $\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n$ uma base de W. Definamos os vetores $\mathbf{w}_2,,\ldots,\mathbf{w}_m$ e $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_m$ do seguinte modo:

$$\begin{aligned}
\mathbf{v}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\
\mathbf{w}_2 &= \mathbf{u}_2 - (\mathbf{v}_1\cdot\mathbf{u}_2)\,\mathbf{v}_1 && \mathbf{v}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \\
\mathbf{w}_3 &= \mathbf{u}_3 - (\mathbf{v}_1\cdot\mathbf{u}_3)\,\mathbf{v}_1 - (\mathbf{v}_2\cdot\mathbf{u}_3)\,\mathbf{v}_2 && \mathbf{v}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} \\
&\vdots \\
\mathbf{w}_m &= \mathbf{u}_m - (\mathbf{v}_1\cdot\mathbf{u}_2)\,\mathbf{v}_1 - (\mathbf{v}_2\cdot\mathbf{w}_m)\,\mathbf{v}_2 - \cdots - (\mathbf{v}_{m-1}\cdot\mathbf{u}_m)\,\mathbf{v}_{m-1} \\
\mathbf{v}_m &= \frac{\mathbf{w}_m}{\|\mathbf{w}_m\|}
\end{aligned}$$

De acordo com o Teorema 27, cada vetor $\mathbf{v}_k$ é ortogonal ao conjunto de vetores $\{\mathbf{v}_1,\ldots,\mathbf{v}_{k-1}\}$, consequentemente o conjunto $\{v_1,\cdots,v_n\}$ é conjunto ortogonal e portanto LI e portanto, formam uma base de W. Este método de obter uma base ortonormal chama-se **Método de Ortonormalização de Gram-Schimidt**.

Exemplo 49. Usando o Método de Gram-Schimidt vamos ortonormalizar a base $\mathbf{u}_1 = (1,0,0)$, $\mathbf{u}_2 = (2,3,0)$, $\mathbf{u}_3 = (1,7,5)$ de $\mathbb{R}^3$.

$$\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{(1,0,0)}{\sqrt{1^2+0^2+0^2}} = \frac{(1,0,0)}{1} = (1,0,0).$$

Como

$$\begin{aligned}(\mathbf{v}_1 \cdot \mathbf{u}_2) &= \langle \mathbf{v}_1, \mathbf{u}_2 \rangle = 1\cdot 2 + 0\cdot 3 + 0\cdot 0 = 2 \text{ e}\\ \mathbf{w}_2 &= \mathbf{u}_2 - (\mathbf{v}_1 \cdot \mathbf{u}_2)\,\mathbf{v}_1 = (2,3,0) - 2\cdot(1,0,0) = (0,3,0)\end{aligned}$$

então

$$\mathbf{v}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \frac{(0,3,0)}{\sqrt{0^2+3^2+0^2}} = \frac{(0,3,0)}{\sqrt{0+9+0}} = \frac{(0,3,0)}{\sqrt{9}} = \frac{(0,3,0)}{3} = (0,1,0)$$

e como

$$\begin{aligned}(\mathbf{v}_1 \cdot \mathbf{u}_3) &= \langle \mathbf{v}_1, \mathbf{u}_3 \rangle = 1\cdot 1 + 0\cdot 7 + 0\cdot 5 = 1,\\ (\mathbf{v}_2 \cdot \mathbf{u}_3) &= \langle \mathbf{v}_2, \mathbf{u}_3 \rangle = 0\cdot 1 + 1\cdot 7 + 0\cdot 5 = 7 \text{ e}\\ \mathbf{w}_3 &= \mathbf{u}_3 - (\mathbf{v}_1 \cdot \mathbf{u}_3)\,\mathbf{v}_1 - (\mathbf{v}_2 \cdot \mathbf{u}_3)\,\mathbf{v}_2 = (1,7,5) - 1(1,0,0) - 7(0,1,0) = (0,0,5)\end{aligned}$$

então

$$v_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \frac{(0,0,5)}{\sqrt{0^2+0^2+5^2}} = \frac{(0,0,5)}{\sqrt{0+0+25}} = \frac{(0,0,5)}{\sqrt{25}} = \frac{(0,0,5)}{5} = (0,0,1)$$

Assim, a base procurada é $\{(1,0,0),(0,1,0),(0,0,1)\}$.

CAPÍTULO

4

APLICAÇÕES NA ECONOMIA

Neste capítulo, trataremos de alguns problemas ligados à Economia que podem ser utilizados como contextualização no estudo de matrizes e sistemas lineares em sala de aula, sobretudo no Ensino Médio. Entendemos que neste nível de escolaridade os alunos já possuem maturidade suficiente para o entendimento de tais problemas e, visto que na Educação Básica pouco se fala sobre Economia, assunto esse que é fundamental para a formação do cidadão.

Vamos então descrever alguns problemas estudados que foram retirados de (SIMON; BLUME; DOERING, 2004).

No próximo capítulo, faremos sugestão de um plano de aulas onde exploramos um desses problemas com o entendimento de que o público alvo são alunos do ensino médio.

4.1 Benefícios de Impostos com Contribuições Beneficentes

Entendemos que esse problema é um bom motivador para o estudo dos sistemas de equações lineares, pois praticamente não requer pré-requisitos para sua modelagem e compreensão, visto que trabalharemos com alunos do Ensino Médio, que é nosso alvo nesse trabalho.

Exemplo 50. Suponha uma empresa que reside num país, onde são cobrados impostos estaduais e federais sobre seu lucro. Estes impostos são cobrados da seguinte forma:

- os impostos estaduais são de 4% do lucro bruto, (lucro sem o pagamento de impostos),
- os impostos federais são de 30% de seu lucro após o pagamento dos impostos estaduais.

Caso essa empresa resolva fazer uma doação beneficente, essa doação será totalmente deduzida de seus impostos. Num determinado ano seu lucro bruto foi de $200.000,00. Suponha que a empresa decida fazer uma doação de 10% do lucro líquido (lucro após pagamento dos impostos estaduais e federais), para a organização humanitária internacional Médicos sem Fronteiras.

(a) Qual o montante pago em impostos estaduais, impostos federais e para doação aos Médicos sem Fronteiras?

(b) Qual o montante pago em impostos se a empresa não fizer a doação?

(c) compare o lucro líquido nos dois casos acima.

Resolução

(a) Vamos modelar esse problema chamando de C o valor da doação beneficente, S o valor pago do imposto estadual e F o valor pago do imposto federal. O lucro da empresa, após os impostos, é $200.000-(S+F)$, então $C=(0,10)\cdot[200.000-(S+F)]$, ou seja,

$$C+(0,1)S+(0,1)F=20.000,$$

Da afirmação de que o imposto estadual é 4% do lucro descontada a doação, obtemos a equação $S=(0,04)\cdot(200.000-C)$, ou seja,

$$(0,04)C+S=8.000.$$

A informação de que o imposto federal é 30% do lucro após a dedução de C e F nos dá a equação $F=(0,30)\cdot[200.000-(C+S)]$, ou

$$(0,3)C+(0,3)S+F=60.000.$$

Reunindo esses pagamentos obtemos o sistema de equações lineares

$$\begin{array}{lcl} C+(0,1)S+(0,1)F & = & 20.000 \\ (0,04)C+S & = & 8.000 \\ (0,3)C+(0,3)S+F & = & 60.000 \end{array}$$

Para resolver esse sistema vamos usar a regra de Cramer, então

$$C=\frac{\det B_1}{\det A}=\frac{\begin{vmatrix} 20.000 & 0,10 & 0,10 \\ 8.000 & 1,00 & 0,00 \\ 60.000 & 0,30 & 1,00 \end{vmatrix}}{\begin{vmatrix} 1,00 & 0,10 & 0,10 \\ 0,04 & 1,00 & 0,00 \\ 0,30 & 0,30 & 1,00 \end{vmatrix}}=\frac{13440}{0,967}=13895,782$$

$$S = \frac{\det B_2}{\det A} = \frac{\begin{vmatrix} 1,00 & 20.000 & 0,10 \\ 0,04 & 8.000 & 0,00 \\ 0,30 & 60.000 & 1,00 \end{vmatrix}}{\begin{vmatrix} 1,00 & 0,10 & 0,10 \\ 0,04 & 1,00 & 0,00 \\ 0,30 & 0,30 & 1,00 \end{vmatrix}} = \frac{7200}{0,967} = 7444,169$$

$$F = \frac{\det B_3}{\det A} = \frac{\begin{vmatrix} 1,00 & 0,10 & 20.000 \\ 0,04 & 1,00 & 8.000 \\ 0,30 & 0,30 & 60.000 \end{vmatrix}}{\begin{vmatrix} 1,00 & 0,10 & 0,10 \\ 0,04 & 1,00 & 0,00 \\ 0,30 & 0,30 & 1,00 \end{vmatrix}} = \frac{51840}{0,967} = 53598,015$$

Logo a solução do sistema é $C = 13.896,782$, $S = 7.444,169$ e $F = 53.598,015$

Portanto, arredondando para unidade mais próxima, os montantes pagos em impostos estaduais, federais e a contribuição, obtemos respectivamente, $R\$$ $7.444,00$, $R\$$ $53.598,00$ e $R\$$ $13.897,00$.

(b) Se a empresa não optar por fazer a contribuição beneficente, o montante pago em impostos estaduais e federais são:

$$\begin{aligned} S &= (0,04) \cdot 200.000 \\ F &= (0,30) \cdot (200.000 - S) \end{aligned}$$

que resulta no sistema

$$\begin{aligned} S &= 8.000 \\ F - 0,3S &= 60.000 \end{aligned}$$

que pode ser facilmente resolvido substituindo o valor de S da primeira equação na segunda, encontrando $F = 62.400$.

Portando os montantes pagos em impostos estaduais e federais são respectivamente $R\$$ 8.000 e $R\$$ 62.400.

(c) No item (a) o lucro líquido é obtido subtraindo do lucro bruto os impostos e a contribuição beneficente, então:

$$200.000 - (53.598 + 13.896 + 7.444) = 125.062$$

Já no item (b), para obter o lucro líquido basta subtrair os impostos do lucro bruto, então:

$$200.000 - (8.000 + 62.400) = 129.600$$

Comparando os dois resultados podemos perceber que a doação de $R\$$ 13.897 aos Médicos sem Fronteira realmente custou $R\$$ 4.538 ($= R\$$ 129.600 $- R\$$ 125.062).

Exemplo 51. Agora, vamos modelar o problema anterior de modo genérico. Suponha uma empresa como do exemplo 50 com lucro bruto P, porcentagem de contribuição beneficente c e porcentagens de impostos estaduais s e federais f. Calcule o montante pago em impostos estaduais (S), impostos federais (F) e contribuição (C) em termos de P, c, s e f.

Resolução

O lucro da empresa, após os impostos, é $P-(S+F)$, então $C = c\cdot[P-(S+F])$, ou seja,

$$C + cS + cF = cP,$$

colocando todas as variáveis do lado esquerdo da igualdade. O imposto estadual é obtido através da equação $S = s\cdot(P-C)$, ou seja,

$$sC + S = sP.$$

O imposto federal é obtido através da equação $F = f\cdot[P-(C+S)]$, ou

$$fC + fS + F = fP.$$

Reunindo esses pagamentos futuros no sistema de equações lineares

$$\begin{array}{lcl} C+cS+cF & = & cP \\ sC+S & = & sP \\ fC+fS+F & = & fP \end{array}$$

Para resolver esse sistema vamos usar a regra de Cramer, então

$$C = \frac{\det B_1}{\det A} = \frac{\begin{vmatrix} cP & c & c \\ sP & 1 & 0 \\ fP & f & 1 \end{vmatrix}}{\begin{vmatrix} 1 & c & c \\ s & 1 & 0 \\ f & f & 1 \end{vmatrix}} = \frac{(1+sf-f-s)cP}{1+sfc-fc-sc}$$

$$S = \frac{\det B_2}{\det A} = \frac{\begin{vmatrix} 1 & cP & c \\ s & sP & 0 \\ f & fP & 1 \end{vmatrix}}{\begin{vmatrix} 1 & c & c \\ s & 1 & 0 \\ f & f & 1 \end{vmatrix}} = \frac{sP(1-c)}{1+sfc-fc-sc}$$

$$F = \frac{\det B_3}{\det A} = \frac{\begin{vmatrix} 1 & c & cP \\ s & 1 & sP \\ f & f & fP \end{vmatrix}}{\begin{vmatrix} 1 & c & c \\ s & 1 & 0 \\ f & f & 1 \end{vmatrix}} = \frac{(1+sc-s-c)fP}{1+sfc-fc-sc}$$

Portanto os montantes pagos são $C = \dfrac{(1+sf-f-s)cP}{1+sfc-fc-sc}$, $S = \dfrac{sP(1-c)}{1+sfc-fc-sc}$ e $F = \dfrac{(1+sc-s-c)fP}{1+sfc-fc-sc}$.

4.2 Modelos de Emprego de Markov

Entender o comportamento do desemprego é fundamental para o estudo de uma economia. Dessa forma pode-se direcionar políticas convenientes para reduzir o número de desempregados. Os modelos probabilísticos em geral utilizados neste estudo, são os **processos de Markov**.

Analisemos a situação seguinte. Se um indivíduo não estiver empregado em um dado mês, então no próximo mês ele poderá encontrar um emprego ou continuar desempregado. Chamaremos de p a probabilidade do indivíduo encontrar um emprego no próximos mês e, portanto, de $1-p$ a probabilidade dele continuar desempregado. Analogamente, se um indivíduo estiver empregado em um dado mês, seja q a probabilidade desse indivíduo permanecer empregado e, portanto, de $1-q$ a probabilidade dele perder o emprego. As probabilidades p, $1-p$, q e $1-q$ são as chamadas **probabilidades de transição** e que comporão a chamada matriz de transição

$$P = \begin{pmatrix} q & p \\ 1-q & 1-p \end{pmatrix}.$$

As duas possibilidades, empregado ou desempregado, são os **estados** do processo.

Suponha que estamos estudando uma dada população em idade produtiva, num período em que ela se mantem constante. Com esta hipótese não admitimos que no período estudado pessoas deixem de ser produtivas ou que novas entrem nesta população. Das pessoas desta população, vamos supor que x delas estão atualmente empregadas e y delas estão atualmente desempregadas. Das x pessoas atualmente empregadas, em média qx permanecerão empregadas e $(1-q)x$ ficarão desempregadas. Das y pessoas atualmente desempregadas, em média py encontrarão emprego e $(1-p)y$ permanecerão desempregadas. Assim o número de empregados no mês seguinte será de $qx+py$ e o de desempregados será $(1-q)x+(1-p)y$. Então chegamos ao seguinte sistema de equações lineares

$$\begin{aligned} x_{t+1} &= qx_t + py_t \\ y_{t+1} &= (1-q)x_t + (1-p)y_t \end{aligned} \tag{4.1}$$

onde x_t e y_t são os números médios do emprego e desemprego, respectivamente, no mês t.

Observe que de (4.1) concluímos que

$$x_{t+1} + y_{t+1} = x_t + y_t,$$

visto que a população em idade de trabalho não muda no nosso exemplo. Em particular, se começarmos com dados percentuais, de modo que x_0 e y_0 somam 1, pois é uma distribuição da população, então x_t e y_t sempre somam 1, para todo t, pois também serão uma distribuição da população.

Os economistas estão interessados em duas questões de sistemas de Markov:

1. Existe uma distribuição da população entre os dois estados que será replicada, ou seja, x_t e y_t chegarão a ser constante no decorrer do tempo? Em outras palavras, será que existe um par não-negativo (x,y) tal que

$$\begin{aligned} x &= qx + py, \\ y &= (1-q)x + (1-p)y, \\ 1 &= x + y. \end{aligned} \tag{4.2}$$

   Caso exista tal par ele é denominado uma **distribuição estacionária** ou um **estado contínuo** de (4.1). Se tal distribuição ocorrer, a menos que p e q mudem, ela continuará a ocorrer permanentemente.

2. Começando a partir de uma distribuição inicial qualquer de estados, o sistema convergirá a uma distribuição estacionária? Se isso ocorrer, dizemos que o sistema é **globalmente estável**.

Podemos responder essas questões utilizando-se de técnicas de Álgebra Linear.

Começando do sistema em (4.2) podemos escrever as duas primeiras equações como

$$\begin{aligned} 0 &= (q-1)x + py \\ 0 &= (1-q)x - py. \end{aligned}$$

Note que as duas equações são equivalentes, logo podemos eliminar uma delas, digamos a segunda, e assim ficamos com

$$\begin{aligned} 0 &= (q-1)x + py \\ 1 &= x + y. \end{aligned}$$

Ou equivalentemente,

$$\begin{pmatrix} q-1 & p \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Resolvendo o sistema acima para x e y temos:

$$x = \frac{p}{1+p-q} \text{ e } y = \frac{1-q}{1+p-q}.$$

Vamos analisar o exemplo específico abaixo:

Exemplo 52. Em certo país, uma pessoa com ensino médio completo que esteja empregada, em determinado mês, tem a probabilidade $q_m = 0,983$ de continuar empregada. Já se estiver desempregada tem probabilidade $p_m = 0,1$ de arrumar um emprego no mês seguinte. Já uma pessoa que possua ensino superior completo e que esteja empregada tem uma probabilidade $q_s = 0,997$ de continuar empregada e, se estiver desempregada, tem uma probabilidade $p_m = 0,106$ de arrumar um emprego, no mês seguinte.

(a) Supondo que 13% dos indivíduos com idade de trabalhar que possuam ensino médio completo estejam desempregadas no mês de janeiro. Encontre o modelo de Markov para esse segmento da população e calcule a porcentagem desses indivíduos que continuarão desempregados no mês de fevereiro.

(b) Supondo que 9% dos indivíduos com idade de trabalhar que possuam ensino superior completo estejam desempregadas no mês de janeiro. Encontre o modelo de Markov para esse segmento da população e calcule a porcentagem desses indivíduos que estarão empregados no mês de fevereiro.

(c) Calcule a distribuição estacionária para cada segmento da população desse pais.

Resolução

(a) Do enunciado do problema temos $1-q_m = 0,017$ e $1-p_m = 0,9$, então de (4.1) o modelo de Markov para a situação descrita na forma matricial é:

$$\begin{pmatrix} x_{t+1_m} \\ y_{t+1_m} \end{pmatrix} = \begin{pmatrix} 0,983 & 0,1 \\ 0,017 & 0,9 \end{pmatrix} \cdot \begin{pmatrix} x_{t_m} \\ y_{t_m} \end{pmatrix}.$$

Como no mês anterior, 13% da população com ensino médio estava desempregada (e consequentemente 87% estava empregada) para calcularmos o número de indivíduos com ensino médio completo desempregados em fevereiro basta fazer

$$y_{t+1_m} = 0,017x_t + 0,9y_t \Rightarrow y_{t+1_m} = 0,017 \cdot 0,87 + 0,9 \cdot 0,13 = 0,13179$$

Ou seja $13,179\%$.

(b) Do enunciado do problema as probabilidades de emprego e desemprego da população com ensino superior são $1-q_s = 0,003$ e $1-p_s = 0,894$, então o modelo de Markov para a situação descrita é:

$$\begin{aligned} x_{t+1_m} &= 0,997x_{t_m} + 0,106y_{t_m} \\ y_{t+1_m} &= 0,003x_{t_m} + 0,894y_{t_m} \end{aligned}$$

Como 9% dos indivíduos com idade de trabalhar que possuem ensino superior completo estão desempregados no mês de janeiro, então $91\% = 0,91$ estão empregados. Para calcularmos o número de indivíduos com ensino superior completo que estarão empregados em fevereiro basta fazer

$$x_{t+1_m} = 0,997x_t + 0106y_t \Rightarrow y_{t+1_m} = 0,997 \cdot 0,91 + 0,106 \cdot 0,09 = 0,91681$$

Ou seja $91,681\%$.

(c) Para calcularmos a distribuição estacionária devemos fazer $x = x_{t+1} = x_t$ e $y = y_{t+1} = y_t$ para cada segmento da população. Para os indivíduos com ensino médio completo utilizamos o sistema do item (a). Assim formamos o sistema

$$\begin{aligned} x &= 0,983x + 0,1y \\ y &= 0,017x + 0,9y \end{aligned}$$

que pode ser reescrito como

$$\begin{aligned} 0,017x - 0,1y &= 0 \\ -0,017x + 0,1y &= 0. \end{aligned}$$

Observe que a segunda equação é equivalente a primeira e pode ser suprimida. Mas $x + y = 1$.

$$\begin{aligned} 0,017x - 0,1y &= 0 \\ x + y &= 1. \end{aligned}$$

Resolvendo esse sistema encontramos como solução $x = 0,8547$ e $y = 0,1453$ aproximadamente. Isso significa que na população com ensino médio a porcentagem de indivíduos que permanacerá empregada e desempregada, respectivamente, será de aproximadamente $85,47\%$ e $14,53\%$.

Para a população com ensino superior, de modo inteiramente análogo chegamos a

$$\begin{aligned} x &= 0,997x + 0,106y \\ y &= 0,003x + 0,894y \end{aligned}$$

que junto com $x+y=1$ nos dá

$$\begin{aligned} 0,003x-0,106y &= 0 \\ x+y &= 1. \end{aligned}$$

Resolvendo esse sistema encontramos como solução $x=0,9725$ e $y=0,0275$ aproximadamente. Isso significa o número de empregados e desempregados que a distribuição dos indivíduos com ensino superior completo vai estabilizar-se em aproximadamente $97,25\%$ e $2,75\%$ respectivamente.

4.3 Modelos Lineares de Produção

Abordaremos aqui os modelos lineares de produção. No capítulo seguinte apresentaremos um plano de aulas usando tais modelos.

Suponhamos que uma economia tenha $n+1$ bens onde cada um deles é produzido por um processo. Destes, chamaremos de bem de número 0 o trabalho. Este bem não é produzido por processo algum, mas é utilizado por todos os outros processos. Chamamos de **processo de produção** uma lista de quantidades de bens, isto é, quantidade do bem i, onde $i=0,1,\cdots n$. Estas representam a quantidade de itens necessária para produzir determinado produto, isto é, são o montante de insumo necessário para produzir uma unidade do produto do processo. Podemos citar como exemplo a produção de bicicletas, que necessita de uma quantidade de aço, outra de plástico, outra de borracha, outra de trabalho, de eletricidade, e assim por diante. Na realidade, alguns processos de produção, como o do aço para bicicletas, utilizam parte de seu próprio produto para auxiliar na produção subsequente.

Os modelos lineares de produção serão aqui considerados por poderem ser analisados com o auxílio de sistemas lineares de equações ou equivalentemente, matricialmente. Por exemplo, na produção de k bicicletas, são necessários k vezes a quantidade de insumo necessária para a produção de uma bicicleta. Esta produção não pode ser aumentada utilizando-se mais de um só fator; precisa-se mais de todos os fatores, e sempre na mesma proporção, ou seja, a produção de 100 bicicletas não precisa ser calculada, é simplesmente a combinação de 100 vezes os insumos necessários para a produção de uma bicicleta.

Vamos analisar uma economia que produz n bens onde a produção de cada bem, digamos o bem j, pode ser descrito por um conjunto de coeficientes que chamaremos de **coeficientes de insumo-produto** e que denotaremos por $a_{0j},a_{1j},\ldots,a_{nj}$. Aqui a_{ij} denota o insumo do bem i necessário para produzir uma unidade do bem j. A produção de x_j unidades do bem j requer $a_{0j}x_j$ unidades do bem 0 (que é o trabalho), $a_{1j}x_j$ unidades do bem 1, $a_{2j}x_j$ unidades do bem 2, e assim por diante. O produto total do bem i deve ser alocado entre as atividades de produção e de consumo. Vamos denotar por c_i a demanda de consumo para o bem i. Essa demanda não é resolvida dentro do modelo, ela é dada de forma externa. Por exemplo, o mercado demandando a produção de 1000 bicicletas, independendo da forma como ela é produzida. Seja c_0 o suprimento de trabalho do consumidor. Como o bem 0 é suprido pelos consumidores, em vez de ser demandado pelos consumidores, no modelo c_0 será um número negativo. Uma n-upla $(c_0, c_1, \ldots, c_n)$ é denominada uma n-upla de **demanda de consumo admissível** se c_0 é negativo, mas todos os demais c_i são não-negativos. Queremos que cada processo produza um produto que é suficiente para atender tanto à demanda do consumo quanto às exigências de insumo das n indústrias. Para nossa economia linear simples essa é a lei da oferta e demanda: o produto produzido deve ser usado na produção ou no consumo.

Seja x_i a quantidade total do bem i necessária para suprir a demanda de consumo deste bem e para a produção de cada bem j. Sendo $a_{ij}x_j$ a quantidade do bem i necessária para a produção de x_j unidades do bem j, resulta que a demanda para o bem i será: $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + c_i$. A lei de oferta e demanda então requer

$$x_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + c_1.$$

É conveniente rearranjar essa equação para dizer que a demanda de consumo deve igualar o produto bruto menos a quantidade de bens necessária para insumo do processo de produção. Para o bem 1, significa

$$(1 - a_{11})x_1 - a_{12}x_2 - \cdots - a_{1n}x_n = c_1.$$

A equação análoga para o bem i é

$$-a_{i1}x_1 - \cdots - a_{ii-1}x_{i-1} + (1 - a_{ii})x_i - a_{ii+1}x_{i+1} - \cdots - a_{in}x_n = c_i.$$

Como o trabalho não é produzido mas apenas demandado pela indústria, tem-se que $x_0 = 0$ e portanto

$$0 = a_{01}x_1 + \cdots + a_{0n}x_n + c_0 \text{ isto é, } -a_{01}x_1 - \cdots - a_{0n}x_n = c_0.$$

Isso leva ao seguinte sistema de $n+1$ equações em n incógnitas, que resume os níveis de equilíbrio de produção para toda essa economia de n indústrias:

$$\begin{array}{rcrcccrcl}
(1-a_{11})x_1 & - & a_{12}x_2 & - & \ldots & - & a_{1n}x_n & = & c_1 \\
-a_{21}x_1 & + & (1-a_{22})x_2 & - & \ldots & - & a_{2n}x_n & = & c_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
-a_{n1}x_1 & - & a_{n2}x_2 & - & \ldots & + & (1-a_{nn})x_n & = & c_n \\
-a_{01}x_1 & - & a_{02}x_2 & - & \ldots & - & a_{0n}x_n & = & c_0
\end{array}$$

Esse sistema linear é denominado **sistema de Leontief aberto**, em homenagem a Wassily Leontief, que primeiro estudou esse tipo de sistema na década de 1930. Esse sistema é conhecido como **aberto** porque as demandas são dadas de modo externo ao sistema, enquanto a oferta de bens é determinada pelas equações em questão. Nesse sistema de equações, os a_{ij} e c_i, são fornecidos e devemos resolver em x_i, os produtos brutos das indústrias.

Note que as quantidades a_{ij} demandadas para produção de x_j unidades do bem j não podem ser maiores que 1, isto é a quantidade de bens j usada na produção de todos os bens, não pode ser maior que a quantidade produzida do bem j. Escrevendo o sistema acima na forma matricial, ignorando a ultima equação, temos:

$$\begin{pmatrix} (1-a_{11}) & -a_{12} & -a_{1n} \\ -a_{21} & (1-a_{22}) & -a_{2n} \\ \vdots & \vdots & \vdots \\ -a_{n1} & -a_{n2} & (1-a_{nn}) \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Observe que a matriz de coeficientes é uma **matriz diagonal dominante**, ver Definição 11. Logo sabemos pelo Teorema 12 que a solução $X = (x_1, x_2, \ldots, x_n)$ deste sistema será sempre não negativa, o que faz sentido no nosso problema já que buscamos quantidades demandadas para uma certa produção.

Exemplo 53. Suponha a economia de uma cidade onde as indústrias foram agrupadas de acordo com o que produzem em quatro setores como descritos na Tabela 1 e as quantidades de insumo de cada uma estão listadas na Tabela 2. As unidades são milhões de reais. Portanto, o elemento $0,173$ na linha 3 e coluna 2 significa que a produção no valor de 1 milhão de reais de produtos de metal final requer o gasto de 173 mil reais em bens de metal básico, a Tabela 3 lista as estimativas das demandas de consumo da economia dessa cidade para determinado ano.

	Setor	**Exemplos**
NF	Não-metal final	Couro, móveis, alimentação
MF	Metal final	Maquinário de construção, eletrodomésticos
MB	Metal básico	Mineração, peças de oficina
NB	Não-metal básico	Vidro, madeira, têxteis, produtos agrícolas

Tabela 1 – Os quatro setores

	NF	MF	MB	NB
NF	0,170	0,004	0,000	0,029
MF	0,003	0,295	0,018	0,002
MB	0,025	0,173	0,460	0,007
NB	0,348	0,037	0,021	0,403

Tabela 2 – Insumo dos setores

(a) Determine quantas unidades teriam que ser produzidas em cada um dos quatro setores para manter funcionando a economia dessa cidade.

(b) Suponha que no ano seguinte as estimativas das demandas de consumo são as listadas na Tabela 4. Determine as novas quantidades a serem produzidas em cada um dos quatro setores.

NF	R$ 99.640
MF	R$ 75.548
MB	R$ 14.444
NB	R$ 33.501

Tabela 3 – Demandas de consumo

NF	R$ 120.500
MF	R$ 87.485
MB	R$ 23.136
NB	R$ 45.209

Tabela 4 – Demandas de consumo para o segundo ano

Resolução

(a) Para resolver o problema, vamos utilizar as entradas da Tabela 2 para formar a matriz dos coeficientes de insumo-produto A e as Tabela 3 para formar a matriz coluna de demanda de consumo $\mathbf{C}$. Nosso objetivo é resolver $(I-A)\mathbf{X} = \mathbf{C}$ para encontrar $\mathbf{X}$ que é a matriz coluna cujas entradas são as produções totais de cada setor:

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

Inicialmente vamos calcular $I-A$

$$\begin{aligned}
(I-A) &= \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0,170 & 0,004 & 0,000 & 0,029 \\ 0,003 & 0,295 & 0,018 & 0,002 \\ 0,025 & 0,173 & 0,460 & 0,007 \\ 0,348 & 0,037 & 0,021 & 0,403 \end{pmatrix} = \\
&= \begin{pmatrix} 0,830 & -0,004 & 0,000 & -0,029 \\ -0,003 & 0,705 & -0,018 & -0,002 \\ -0,025 & -0,173 & 0,540 & -0,007 \\ -0,348 & -0,037 & -0,021 & 0,597 \end{pmatrix}
\end{aligned}$$

Note que a matriz $(I-A)$ é diagonal dominante pois: cada entrada fora da diagonal é não-negativa, cada entrada diagonal é positiva, e a soma das entradas de cada coluna é positiva. Logo, pelo Teorema 12, $(I-A)^{-1}$ existe e contém apenas entradas não negativas. A solução do sistema $(I-A)\mathbf{X} = \mathbf{C}$ é $\mathbf{X} = (I-A)^{-1} \cdot \mathbf{C}$; como $\mathbf{C}$ contém apenas entradas não-negativas podemos prever uma solução não-negativa para o sistema e, consequentemente, admissível para o problema de produção.

Para calcular $(I-A)^{-1}$ podemos utilizar o método descrito no exemplo 12 do Capítulo 1 e, então, utilizá-la para resolver

$$\begin{aligned}
\mathbf{X} &= (I-A)^{-1}\cdot\mathbf{C} = \\
&= \begin{pmatrix} 1,230 & 0,011 & 0,003 & 0,060 \\ 0,009 & 1,431 & 0,048 & 0,006 \\ 0,069 & 0,460 & 1,868 & 0,027 \\ 0,720 & 0,111 & 0,070 & 1,711 \end{pmatrix} \cdot \begin{pmatrix} 99.640 \\ 75.548 \\ 14.444 \\ 33.501 \end{pmatrix} = \\
&= \begin{pmatrix} 125.416 \\ 109.862 \\ 69.546 \\ 138.477 \end{pmatrix}
\end{aligned}$$

Concluímos que são necessários 125.416 milhões em produtos da indústria de não-metal final, 109.862 milhões em produtos da indústria de metal final, 69.546 milhões em produtos da indústria de metal básico e 138.477 milhões em produtos da indústria de não-metal básico para atender as necessidades de insumo das indústrias assim como a demanda de consumo dessa economia no primeiro ano considerado.

(b) Vamos utilizar as entradas da Tabela 4 para formar a matriz $\mathbf{C}'$ de demanda de consumo e assim resolver o sistema $\mathbf{X} = (I-A)^{-1}\mathbf{C}'$

$$\begin{aligned}
\mathbf{X} &= (I-A)^{-1}\cdot\mathbf{C} = \\
&= \begin{pmatrix} 1,230 & 0,011 & 0,003 & 0,060 \\ 0,009 & 1,431 & 0,048 & 0,006 \\ 0,069 & 0,460 & 1,868 & 0,027 \\ 0,720 & 0,111 & 0,070 & 1,711 \end{pmatrix} \cdot \begin{pmatrix} 120.500 \\ 87.485 \\ 23.136 \\ 45.209 \end{pmatrix} = \\
&= \begin{pmatrix} 151.926 \\ 127.611 \\ 93.035 \\ 175468 \end{pmatrix}
\end{aligned}$$

Concluímos que são necessários 151.926 milhões em produtos da indústria de não-metal final, 127.611 milhões em produtos da indústria de metal final, 93.035 milhões em produtos da indústria de metal básico e 175468 milhões em produtos da indústria de não-metal básico para atender as necessidades de insumo das indústrias assim como a demanda de consumo dessa economia no segundo ano considerado.

CAPÍTULO

5

PLANO DE AULAS - MODELOS LINEARES DE PRODUÇÃO

5.1 Objetivos

O plano de aulas descrito a seguir tem como objetivo contextualizar a resolução de sistemas de equações lineares quadrados de ordem 2 e de ordem 3, dar sentido ao estudo da matriz inversa usando-a na resolução de tais sistemas e estudo das propriedades da matriz diagonal dominante (Definição 11).

5.2 Público alvo

De acordo com (SÃO PAULO (Estado), Secretaria da Educação, 2010) o estudo de matrizes e sistemas de equações lineares deve ser realizado no segundo ano do Ensino Médio; portanto, esse plano de aulas destina-se aos alunos que cursam o referido ano do Ensino Médio.

5.3 Pré-requisitos

Para aplicação do plano de aulas descrito a seguir o professor já deve ter trabalhado o conceito de matrizes, operações com matrizes, cálculo de determinante, existência de matrizes inversas, cálculo de matrizes inversas e a relação entre sistemas lineares de equações com equações matriciais. Neste plano, usaremos a técnica de escalonamento para cálculo da matriz inversa, mas nada impede que se use outras técnicas.

5.4 Materiais e tecnologias

Material impresso, calculadora e lousa.

5.5 Recomendações metodológicas

A metodologia adotada envolve leitura de texto, pesquisa individual para se obter dados sobre a produção de cana-de-açúcar e etanol, resolução de exercícios exemplares, aula expositiva, resolução de problemas em grupos de alunos e, além disso, fazer uso de discussões acerca do contexto dos problemas de produção.

5.6 Dificuldades previstas

Ao utilizar modelos lineares de produção para contextualizar a resolução de sistemas de equações lineares, nos deparamos com matrizes cujos elementos são números não inteiros, dessa forma os alunos podem encontrar dificuldades para realizar as operações com tais matrizes. Sendo assim, sugerimos a utilização de calculadora para auxiliá-los nos cálculos.

5.7 Descrição Geral

Aula 1

O objetivo dessa primeira aula é introduzir para os alunos os modelos lineares de produção através da leitura de um texto que exemplificará um problema de produção de uma fazenda. Sugerimos que o professor faça uma leitura compartilhada do texto com os alunos, fazendo comentários sempre que julgar necessário de forma a assegurar que todos os alunos o entendam. Após a leitura do texto propomos uma pesquisa para os alunos realizarem extraclasse com o objetivo de adiantar o exemplo que será trabalhado na aula seguinte.

Leitura e Análise de Texto

Consideremos inicialmente uma fazenda que produza apenas milho. Suponha que parte do milho será destinado para comercialização e outra parte será reservada para seu plantio. Suponha também que para uma produção de certa quantidade x de milho seja necessária um décimo desta quantidade em sementes para o plantio.

Assim se na próxima safra quisermos comercializar 3 toneladas de milho, qual a quantidade total de milho que será necessária para este objetivo?

$$x = \frac{x}{10} + 3 \Rightarrow x = \frac{10}{3}$$

Logo será necessário $\frac{10}{3} = 3,333333$ de toneladas de milho.

Consideremos agora, uma fazenda orgânica que produz dois bens: milho e fertilizante. O milho é produzido utilizando-se milho e fertilizante. O fertilizante é preparado a partir de safras anteriores do próprio milho.

Suponha que a produção de 1 tonelada de milho requeira $0,1$ tonelada de milho e $0,8$ tonelada de fertilizante. E que a produção de uma tonelada de fertilizante requeira $0,5$ tonelada de milho e nenhum fertilizante.

Logo serão usados,

(i) $0,1$ tonelada de milho para produção de 1 tonelada de milho.

(ii) $0,5$ tonelada de milho para a produção de 1 tonelada de fertilizante.

(iii) $0,8$ tonelada de fertilizante para a produção de 1 tonelada milho.

(iv) 0 tonelada de fertilizante para a produção de 1 tonelada de fertilizante.

Logo, se produzirmos um total de x toneladas de milho e y toneladas de fertilizante serão necessárias proporcionalmente as quantidades de:

(i) $0,1x$ tonelada de milho para produção de x toneladas de milho.

(ii) $0,5y$ tonelada de milho para a produção de y toneladas de fertilizante.

(iii) $0,8x$ tonelada de fertilizante para a produção de x toneladas milho.

(iv) $0y$ tonelada de fertilizante para a produção de y tonelada de fertilizante.

Em economia, a quantidade de produtos necessários para a produção de milho e fertilizante, pode ser representada através de pares ordenados (a,b), onde no nosso exemplo, a representa a quantidade de milho e b representa a quantidade de fertilizante necessárias para produção de cada um destes itens. Os números a e b são os chamados insumos para a produção de milho e de fertilizante nessa fazenda. De modo geral, esses números são chamados pelos economistas de **coeficientes de insumo-produto**. Chamaremos de **processo de produção** de uma **economia**, uma lista de quantidades de bens necessários para produzir determinado produto.

Novamente, no nosso exemplo, a economia seria a fazenda e os bens serão o milho e o fertilizante. O processo de produção do milho será descrito pelo par de números $(0,1;\ 0,8)$ e o processo de produção de fertilizante será descrito pelo par de números $(0,5;\ 0)$.

De forma geral, numa economia que tenha n bens onde cada um deles é produzido por um processo, o processo de produção do bem i será então uma lista de quantidades de insumos dada pela n-upla $(a_{i1},\ a_{i2},\ \ldots,\ a_{in})$.

Suponha que a fazenda acima deseje abastecer o mercado com 4 toneladas de milho e 2 toneladas de fertilizante. Qual será a quantidade de milho e fertilizante necessárias para conseguir tal produção? Isto é, como modelar este problema e desta forma determinar as quantidades de insumos necessárias para isso?

As respostas a estas perguntas podem ser encontradas examinando-se um sistema de equações lineares. Suponha que os dois processos de produção são conduzidos de tal modo que produzam x toneladas de milho e y toneladas de fertilizante. Assim, da produção total de milho parte será destinada às produções de milho e fertilizante e a outra parte, 4 toneladas, para a comercialização. O mesmo raciocínio temos para a quantidade total de fertilizante produzido.

Pela análise feita anteriormente vemos que da produção total de x toneladas de milho esta deverá satisfazer:

$$x = 0,1x + 0,5y + 4$$

Com o mesmo raciocínio tem-se que

$$y = 0,8x + 2$$

Assim o sistema resultante nos dá:

$$\begin{cases} 0,9x - 0,5y & = & 4 \\ -0,8x + y & = & 2 \end{cases}.$$

Pode ser mostrado que a solução deste sistema é $x = 10$ e $y = 10$, isto é, $x = 10$ e $y = 10$ toneladas de milho e fertilizante respectivamente deve ser a produção total da fazenda para alcançar o objetivo desejado.

Pesquisa

Cana-de-açúcar é um grupo de espécies de gramíneas perenes altas do gênero Saccharum, tribo Andropogoneae, nativas das regiões tropicais do sul da Ásia e da Melanésia e utilizadas principalmente para a produção de açúcar e etanol. A cana-de-açúcar foi introduzida no Brasil no início do século XVI, quando foi iniciada a instalação de engenhos de açúcar, a primeira indústria implantada na nova possessão de Portugal, que em pouco tempo substituiu a indústria extrativa do pau-brasil. O Brasil é, hoje, o principal produtor de cana-de-açúcar do mundo. Seus produtos são largamente utilizados na produção de açúcar, álcool combustível (etanol) e, mais recentemente, biodiesel.

Faça uma pesquisa ou uma entrevista com alguém que trabalhe no setor de produção de cana-de-açúcar ou de etanol e descubra qual a quantidade de cana-de-açúcar é necessária para se produzir uma tonelada de cana-de-açúcar e qual a quantidade de cana-de-açúcar é necessária para se produzir um litro de etanol.

Aulas 2 e 3

O objetivo dessa aula dupla é levar os alunos a entenderem a modelagem e a resolução de problemas de produção, verem a vantagem de optarem pelo método de resolução via matriz inversa e também chamar a atenção deles para a admissibilidade das soluções matemáticas encontradas.

O professor deve fazer os exemplos com a participação dos alunos. Ao final de cada exemplo o professor deve fazer os comentários com os questionamentos sobre as soluções matemáticas do problema apresentado.

Exemplo 54. Uma cidade do interior do estado de São Paulo possui dois setores econômicos principais: o setor agrícola, responsável pela produção de cana-de-açúcar e uma usina, responsável pela produção de etanol combustível. Para produção de uma tonelada de cana-de-açúcar são necessários $0,05$ toneladas de cana-de-açúcar e $0,3$ litros de etanol. Para produção de um litro de etanol são necessários $0,01$ tonelada de cana-de-açúcar e $0,1$ litro de etanol. Suponha que num determinado ano as demandas externas por cana-de-açúcar e etanol sejam de 50.000 toneladas e 150.000 litros respectivamente.

Quais devem ser as quantidades brutas a serem produzidas de cana-de-açúcar e etanol para satisfazer as demandas de produção e as demandas de consumo para aquele ano?

Resolução

De acordo com as informações descritas no enunciado do problema, as quantidades de insumo necessárias para produção de 1 tonelada de cana-de-açúcar e 1 litro de etanol são:

(i) $0,05$ tonelada de cana-de-açúcar para produção de 1 tonelada de cana-de-açúcar.

(ii) $0,01$ tonelada de cana-de-açúcar para produção de 1 litro de etanol.

(iii) $0,3$ litro de etanol para produção de 1 tonelada de cana-de-açúcar.

(iv) $0,1$ litro de etanol para produção de 1 litro de etanol.

Assim para produção de x toneladas de cana-de-açúcar e y litros de etanol serão necessários:

(i) $0,05x$ tonelada de cana-de-açúcar para produção de x toneladas de cana-de-açúcar.

(ii) $0,01y$ tonelada de cana-de-açúcar para produção de y litros de etanol.

(iii) $0,3x$ litro de etanol para produção de x toneladas de cana-de-açúcar.

(iv) $0,1y$ litro de etanol para produção de y litros de etanol.

Como temos uma demanda externa por cana-de-açúcar de 50.000 toneladas, a produção total de cana-de-açúcar deve satisfazer as necessidades internas de produção destes bens mais a externa. Assim a equação que descreve a produção total de cana-de-açúcar é

$$x = 0,05x + 0,01y + 50000$$

Já com demanda externa por etanol de 150.000 litros, então a produção total de etanol deve ser

$$y = 0,3x + 0,1y + 150000$$

O que nos dá o sistema

$$\begin{cases} x &= 0,05x + 0,01y + 50000 \\ y &= 0,3x + 0,1y + 150000 \end{cases}$$

Rearranjando o sistema temos

$$\begin{cases} 0,95x - 0,01y &= 50000 \\ -0,3x + 0,9y &= 150000 \end{cases}$$

Deste sistema temos associadas as seguintes matrizes: **matriz de coeficientes** A, a **matriz dos termos independentes** C e a **matriz de incógnitas** X dadas por

$$A = \begin{pmatrix} 0,95 & -0,01 \\ -0,3 & 0,9 \end{pmatrix}, C = \begin{pmatrix} 50000 \\ 150000 \end{pmatrix} \text{ e } X = \begin{pmatrix} x \\ y \end{pmatrix}$$

As entradas da matriz C representam a demanda de consumo por cana-de-açúcar e etanol e, por isso, são todas não-negativas. Esta matriz é denominada **matriz de demanda de consumo admissível**. Não seria admissível se tivesse algum termo negativo, pois o mercado demanda um produto ou não, e se não demanda C seria a matriz nula ou teria alguma entrada nula, mas nunca teria entradas negativas.

Usando o produto de matrizes vemos que esse sistema pode ser escrito na forma da equação $AX = C$, ou seja,

$$\begin{pmatrix} 0,95 & -0,01 \\ -0,3 & 0,9 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 50000 \\ 150000 \end{pmatrix}.$$

Para resolver a equação $AX = C$ vamos usar um método que faz o uso da matriz inversa da matriz de coeficientes A, que é chamada A^{-1}.

Lembrando que para resolver uma equação numérica $ax = c$ vemos que esta terá solução se $a \neq 0$ e neste caso multiplicando-se ambos os lados por a^{-1} teremos $a^{-1}ax = x = a^{-1}c$. A única restrição para o número a é que ele não pode ser nulo, pois este não possui inverso multiplicativo.

Numa equação matricial $AX = C$, a solução X pode ser obtida da mesma forma, desde que A tenha matriz inversa A^{-1}. Então, multiplicando-se ambos os lados da igualdade $AX = C$

por A^{-1}, obtemos

$$A^{-1}AX = A^{-1}C \Rightarrow IX = A^{-1}C \Rightarrow X = A^{-1}C.$$

Lembrando que nem toda matriz quadrada possui inversa, somente aquelas cujo determinante é não-nulo.

Como no nosso exemplo $\det(A) = 0,95 \cdot 0,9 - (-0,01) \cdot (-0,3) = 0,855 - 0,003 = 0,852 \neq 0$, então a matriz A é invertível. Mostraremos, agora, como calcular a inversa da matriz A.

Observação: Caso haja interesse em exercitar o processo de escalonamento, seguiríamos os procedimento abaixo. Caso o interesse maior seja na aplicação em si, seria recomendável utilizar o fato que para matrizes $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sua inversa é obtida através de $\frac{1}{detA}\begin{pmatrix} d & -b \\ c & -a \end{pmatrix}$.
Usando a técnica de escalonamento, inicialmente aumentamos a matriz A com a matriz identidade I

$$\begin{pmatrix} A & | & I \end{pmatrix} = \begin{pmatrix} 0,95 & -0,01 & | & 1 & 0 \\ -0,3 & 0,9 & | & 0 & 1 \end{pmatrix}$$

Lembrando que, conforme visto em aulas anteriores, o processo consiste em utilizar operações elementares, com as linhas desta nova matriz, até que seu lado esquerdo se transforme na matriz identidade. Tais operações elementares são:

(a) permutação de linhas;

(b) multiplicação de uma linha por um escalar não nulo;

(c) soma de um múltiplo de uma linha a uma outra linha.

Para facilitar os cálculos vamos escrever as entradas da matriz A na forma fracionária.

$$\begin{pmatrix} \frac{19}{20} & -\frac{1}{100} & | & 1 & 0 \\ -\frac{3}{10} & \frac{9}{10} & | & 0 & 1 \end{pmatrix}.$$

Como o objetivo é transformar sua primeira metade na matriz identidade, começamos multiplicando a primeira linha por $\frac{20}{19}$, obtendo o resultado

$$\begin{pmatrix} 1 & -\frac{1}{95} & | & \frac{20}{19} & 0 \\ -\frac{3}{10} & \frac{9}{10} & | & 0 & 1 \end{pmatrix}.$$

Substituindo a segunda linha pela soma de $\frac{3}{10}$ da primeira com ela mesma obtemos:

$$\begin{pmatrix} 1 & -\frac{1}{95} & | & \frac{20}{19} & 0 \\ 0 & \frac{426}{475} & | & \frac{6}{19} & 1 \end{pmatrix}.$$

Multiplicamos agora a segunda linha por $\frac{475}{426}$, obtendo

$$\left(\begin{array}{cc|cc} 1 & -\frac{1}{95} & \frac{20}{19} & 0 \\ 0 & 1 & \frac{25}{71} & \frac{475}{426} \end{array}\right).$$

Por fim, multiplicamos a segunda linha por $\frac{1}{95}$ e adicionamos o resultado à primeira linha.

$$\left(\begin{array}{cc|cc} 1 & 0 & \frac{75}{71} & \frac{5}{426} \\ 0 & 1 & \frac{25}{71} & \frac{475}{426} \end{array}\right).$$

Logo, obtemos:

$$A^{-1} = \begin{pmatrix} \frac{75}{71} & \frac{5}{426} \\ \frac{25}{71} & \frac{475}{426} \end{pmatrix}.$$

Observação 8. Note que as entradas de A^{-1} são todas não-negativas.

Continuando, como as entradas de A^{-1} e da matriz demanda C são positivas podemos prever que a solução X terá apenas entradas positivas já que se $AX = C$ então $X = A^{-1}C$. Fazendo os cálculos, temos

$$X = \begin{pmatrix} \frac{75}{71} & \frac{5}{426} \\ \frac{25}{71} & \frac{475}{426} \end{pmatrix} \cdot \begin{pmatrix} 50.000 \\ 150.000 \end{pmatrix} = \begin{pmatrix} \frac{3875000}{71} \\ \frac{13125000}{71} \end{pmatrix}.$$

Portanto $x = \dfrac{3875000}{71} \cong 54.577$ e $y = \dfrac{13125000}{71} \cong 184.859$. Como os valores encontrados para x e y são não-negativos, $X = (x, y)$ não só é solução da equação matricial mas de fato, solução admissível para nosso problema de produção. De fato, as quantidades brutas a serem produzidas para satisfazer as demandas de produção e de consumo de cana-de-açúcar e etanol devem ser aproximadamente 54.577 toneladas e 184.859 litros, respectivamente, para o ano considerado.

Comentários

Como vimos no exemplo 54, o sistema de equações lineares que modela o problema da produção de cana-de-açúcar e de etanol tem solução para aquela demanda de consumo. Vimos também que a solução obtida é perfeitamente aceitável para o problema, pois é não-negativa, ou seja, são quantidades positivas, ou eventualmente nulas, para a produção total de cana-de-açúcar e etanol. Não teria sentido uma produção negativa, porque não há como uma indústria produzir uma quantidade negativa de determinado produto. Nesse sentido temos alguns questionamentos a fazer:

(i) Uma vez que o sistema tem solução para determinada demanda de consumo, terá solução para qualquer demanda de consumo?

(ii) Se existir solução para outras demandas de consumo, as soluções obtidas serão sempre admissíveis, isto é, não-negativas?

(iii) Existe alguma maneira de se conhecer as soluções para qualquer demanda de consumo C, uma vez que A seja sempre a mesma?

Antes de respondermos a essas questões vamos analisar mais um exemplo.

Exemplo 55. Usando os dados do exemplo 54, imagine que se queira planejar a produção de etanol e cana de açúcar para os próximos dois anos da seguinte maneira:

(a) 45000 toneladas de cana-de-açúcar e 200000 litros de etanol no primeiro ano e;

(b) 60000 toneladas de cana-de-açúcar e 230000 litros de etanol no segundo ano.

Calcule as quantidades brutas a serem produzidas para obtermos tais produções.

Resolução

(a) Como a matriz de coeficientes é a mesma do exemplo 54, o sistema pode ser escrito na forma de equação matricial como $AX = C'$, onde C' é a nova matriz demanda de consumo, ou seja, para o primeiro ano temos:

$$\begin{pmatrix} 0,95 & -0,01 \\ -0,3 & 0,9 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 45000 \\ 200000 \end{pmatrix}$$

já sabemos que esse sistema possui solução pois $\det(A) \neq 0$ e existe a matriz A^{-1}. Como a matriz A^{-1} já foi calculada na resolução do exemplo 54, para resolver a equação matricial acima basta fazer $X = A^{-1}C$. Então

$$X = \begin{pmatrix} \frac{75}{71} & \frac{5}{426} \\ \frac{25}{71} & \frac{475}{426} \end{pmatrix} \cdot \begin{pmatrix} 45000 \\ 200000 \end{pmatrix} = \begin{pmatrix} \frac{10625000}{213} \\ \frac{50875000}{213} \end{pmatrix}.$$

Portanto, as quantidades brutas para satisfazer as necessidades de insumo e de consumo, de cana-de-açúcar e etanol, devem ser de aproximadamente 49883 toneladas e 238850 litros respectivamente para o primeiro ano considerado.

(b) Para o segundo ano repetimos o procedimento anterior trocando a matriz de demanda de consumo pela matriz $\begin{pmatrix} 60000 \\ 230000 \end{pmatrix}$. Então

$$X = \begin{pmatrix} \frac{75}{71} & \frac{5}{426} \\ \frac{25}{71} & \frac{475}{426} \end{pmatrix} \cdot \begin{pmatrix} 60000 \\ 230000 \end{pmatrix} = \begin{pmatrix} \frac{14075000}{213} \\ \frac{59125000}{213} \end{pmatrix}.$$

Logo, as quantidades brutas de açúcar e etanol devem ser de aproximadamente 66080 toneladas e 277582 litros respectivamente para o segundo ano considerado.

Comentários

Como vimos no exemplo 54 e no exemplo 55 o sistema associado ao problema da produção de cana-de-açúcar e etanol apresentou solução para três demandas de consumo diferentes. Mas será que encontraremos solução para qualquer demanda de consumo, como questionado em (1)?

A resposta a essa pergunta é sim, pois sabemos que A^{-1} existe logo a solução do sistema $AX = C$ é $X = A^{-1}C$ qualquer que seja a demanda de consumo C.

Para responder as perguntas em (2) e (3) vamos resolver o sistema $AX = C$, onde $C = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, e α e β representam as demandas de consumo de cana-de-açúcar e etanol, respectivamente. Já sabemos que a solução de $AX = C$ é dada por $X = A^{-1}C$; como já calculamos A^{-1}, a solução X será dada por

$$X = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{75}{71} & \frac{5}{426} \\ \frac{25}{71} & \frac{475}{426} \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \frac{450\alpha+5\beta}{426} \\ \frac{150\alpha+475\beta}{426} \end{pmatrix}.$$

Do resultado acima vemos que

$$x = \frac{450\alpha + 25\beta}{426} \text{ e } y = \frac{150\alpha + 475\beta}{426}. \tag{5.1}$$

Como α e β são não-negativas, x e y também o são, o que responde a pergunta em (2).

Além disso, veja que para uma demanda de consumo qualquer (α, β) basta substituir na fórmula acima que encontramos a produção total necessária para tal demanda, respondendo assim a pergunta em (3).

Aula 4

O objetivo dessa aula é trabalhar os conceitos utilizados na aula anterior, para um exemplo modelado por um sistema de ordem 3.

Exemplo 56. Considere uma economia simples de três indústrias, onde a primeira produz um produto A, a segunda produz um produto B e a terceira produz um produto C. Para produzir uma unidade do produto A são necessários $0,2$ unidades do produto A, $0,3$ unidades do produto

B e $0,2$ unidades do produto C. Para produzir uma unidade do produto B são necessários $0,3$ unidades do produto A, $0,1$ unidade do produto B e $0,4$ unidades do produto C. Para produzir uma unidade do produto C são necessários $0,4$ unidades do produto A $0,2$ unidades do produto B e $0,3$ unidades do produto C. No mês de janeiro de determinado ano a demanda de consumo dos produtos A, B e C são 15, 20 e 25 unidades, respectivamente. Qual deve ser produção das indústrias para o mês de janeiro, para tal demanda?

Resolução

De acordo com as informações descritas no enunciado do problema, as quantidades de insumo necessárias para produção de 1 unidade do produto A, 1 unidade do produto B e 1 unidade do produto C são:

(i) $0,2$unidade do produto A para produção de 1 unidade do produto A.

(ii) $0,3$ unidade do produto A para produção de 1 unidade do produto B.

(iii) $0,4$ unidade do produto A para produção de 1 unidade do produto C.

(iv) $0,3$ unidade do produto B para produção de 1 unidade do produto A.

(v) $0,1$ unidade do produto B para produção de 1 unidade do produto B.

(vi) $0,2$ unidade do produto B para produção de 1 unidade do produto C.

(vii) $0,2$ unidade do produto C para produção de 1 unidade do produto A.

(viii) $0,4$ unidade do produto C para produção de 1 unidade do produto B.

(ix) $0,3$ unidade do produto C para produção de 1 unidade do produto C.

Assim para produção de x unidades do produto A, y unidades do produto B e z unidades do produto C são necessárias:

(i) $0,2x$unidade do produto A para produção de x unidades do produto A.

(ii) $0,3y$ unidade do produto A para produção de y unidades do produto B.

(iii) $0,4z$ unidade do produto A para produção de z unidades do produto C.

(iv) $0,3x$ unidade do produto B para produção de x unidades do produto A.

(v) $0,1y$ unidade do produto B para produção de y unidades do produto B.

(vi) $0,2z$ unidade do produto B para produção de z unidades do produto C.

(vii) $0,2x$ unidade do produto C para produção de x unidades do produto A.

(viii) $0,4y$ unidade do produto C para produção de y unidades do produto B.

(ix) $0,3z$ unidade do produto C para produção de z unidades do produto C.

Como temos uma demanda externa pelo produto A de 15 unidades a produção total desse produto deve satisfazer as necessidades de insumo das indústrias assim como a demanda externa; então, a equação que descreve a produção total do produto A é

$$x = 0,2x + 0,3y + 0,4z + 15.$$

A demanda externa pelo produto B é 20 unidades; então, para satisfazer as necessidades de insumo das indústrias assim como a demanda externa, a produção total do produto B deve ser

$$y = 0,3x + 0,1y + 0,2z + 20.$$

A demanda externa pelo produto C é 25 unidades, então para satisfazer as necessidades de insumo das indústrias assim como a demanda externa a produção total do produto C deve ser

$$y = 0,2x + 0,4y + 0,3z + 25.$$

Então, temos o sistema

$$\begin{cases} x &= 0,2x + 0,3y + 0,4z + 15 \\ y &= 0,3x + 0,1y + 0,2z + 20 \\ z &= 0,2x + 0,4y + 0,3z + 25 \end{cases}.$$

Rearranjando essas equações, temos

$$\begin{cases} 0,8x - 0,3y - 0,4z &= 15 \\ -0,3x + 0,9y - 0,2z &= 20 \\ -0,2x - 0,4y + 0,7z &= 25 \end{cases}.$$

Usando o produto de matrizes podemos escrever esse sistema na forma $AX = C$, ou seja,

$$\begin{pmatrix} 0,8 & -0,3 & -0,4 \\ -0,3 & 0,9 & -0,2 \\ -0,2 & -0,4 & 0,7 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 15 \\ 20 \\ 25 \end{pmatrix}.$$

Para resolver usando o método da matriz inversa aumentamos a matriz A com a matriz identidade I e escrevemos as entradas de A na forma fracionária para facilitar os cálculos; ou seja,

$$\left(\begin{array}{ccc|ccc} \frac{4}{5} & -\frac{3}{10} & -\frac{4}{10} & 1 & 0 & 0 \\ -\frac{3}{10} & \frac{9}{10} & -\frac{1}{5} & 0 & 1 & 0 \\ -\frac{1}{5} & -\frac{2}{5} & \frac{7}{10} & 0 & 0 & 1 \end{array}\right).$$

Agora, realizamos operações elementares sobre as linhas da matriz acima necessárias para transformar seu lado esquerdo na matriz identidade. Após alguns cálculos obtemos o resultado

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{110}{49} & \frac{74}{49} & \frac{12}{7} \\ 0 & 1 & 0 & \frac{50}{49} & \frac{96}{49} & \frac{8}{7} \\ 0 & 0 & 1 & \frac{60}{49} & \frac{76}{49} & \frac{18}{7} \end{array}\right).$$

Para encontrarmos a solução do sistema $AX = C$ basta calcular $X = A^{-1}C$, ou seja

$$X = \begin{pmatrix} \frac{110}{49} & \frac{74}{49} & \frac{12}{7} \\ \frac{50}{49} & \frac{96}{49} & \frac{8}{7} \\ \frac{60}{49} & \frac{76}{49} & \frac{18}{7} \end{pmatrix} \cdot \begin{pmatrix} 15 \\ 20 \\ 25 \end{pmatrix} = \begin{pmatrix} \frac{5230}{49} \\ \frac{4070}{49} \\ \frac{5570}{49} \end{pmatrix}.$$

Portanto, a solução do sistema é $x = \dfrac{5230}{49} \cong 107$, $y = \dfrac{4070}{49} \cong 83$ e $z = \dfrac{5570}{49} \cong 114$. Então, a produção total das indústria deve ser: 107 unidades do produto A, 83 unidades do produto B e 114 unidades do produto C.

Comentários

Assim como no exemplo 54, o sistema de equações lineares que modela o problema da produção dos produtos A, B e C, do exemplo 56 tem solução para aquela demanda de consumo e a solução encontrada também é admissível para o problema, pois é não-negativa. Nesse sentido vamos refazer os questionamentos do exemplo 54.

(i) Uma vez que o sistema tem solução para determinada demanda de consumo, terá solução para qualquer demanda de consumo?

(ii) Se existir solução para outras demandas de consumo, as solução obtidas serão sempre não-negativas?

(iii) Existe alguma maneira de se conhecer as soluções para qualquer demanda de consumo C, uma vez que A seja sempre a mesma?

Para responder essas questões vamos proceder da mesma forma que fizemos anteriormente.

Se o determinante de A for não nulo a solução do sistema $AX = C$ é $X = A^{-1}C$. Com a existência de A^{-1}, o sistema terá solução para qualquer C.

Para encontrarmos as soluções do sistema para uma demanda de consumo genérica, digamos, $C = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ onde α, β e γ representam as demandas de consumo dos produtos A, B e C respectivamente, vamos calcular o produto $A^{-1}C$. Como já conhecemos A^{-1} obtemos

$$X = \begin{pmatrix} \frac{110}{49} & \frac{74}{49} & \frac{12}{7} \\ \frac{50}{49} & \frac{96}{49} & \frac{8}{7} \\ \frac{60}{49} & \frac{76}{49} & \frac{18}{7} \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \frac{110\alpha+74\beta+84\gamma}{49} \\ \frac{50\alpha+96\beta+56\gamma}{49} \\ \frac{60\alpha+76\beta+126\gamma}{49} \end{pmatrix}.$$

Do resultado acima vemos que as soluções

$$x = \frac{110\alpha + 74\beta + 84\gamma}{49}, y = \frac{50\alpha + 96\beta + 56\gamma}{49} \text{ e } z = \frac{60\alpha + 76\beta + 126\gamma}{49}$$

são não-negativas, pois α, β e γ também o são.

Com o que fizemos, se quisermos encontrar a solução do sistema para uma demanda de consumo qualquer, basta substituir os valores referentes a α, β e γ em

$$x = \frac{110\alpha + 74\beta + 84\gamma}{49}, y = \frac{50\alpha + 96\beta + 56\gamma}{49} \text{ e } z = \frac{60\alpha + 76\beta + 126\gamma}{49}.$$

Concluímos que para qualquer outra demanda C cujas entradas são não negativas, a correspondente solução X também terá apenas entradas não negativas. Sendo assim, as soluções do sistema para demandas não negativas serão de fato, soluções admissíveis do problema.

Aulas 5 e 6

Nessa aula dupla, o objetivo é fazer com que os alunos modelem e analisem outros problemas de produção e exercitem os conceitos desenvolvidos anteriormente. Para isso, sugerimos alguns exercícios que possam ser resolvidos em grupos de 4 ou 5 alunos.

Como sugestão para formação do grupo, o professor escolherá alunos para serem monitores dos grupos (de preferência alunos que estejam acompanhando melhor as aulas) e, então, cada monitor de grupo escolherá, um de cada vez, os outros integrantes dos grupos.

Enquanto os alunos realizam os exercícios o professor deve percorrer os grupos para orientar sobre possíveis dúvidas dos alunos e também para avaliar o desempenho de cada grupo e de cada aluno.

Exercícios propostos

Exercício 1. Considere-se o gerente da fazenda orgânica apresentada na seção leitura e análise de texto, descubra qual deve ser a produção total de milho e de fertilizante para que a fazenda consiga abastecer o mercado com 4 toneladas de milho e 2 toneladas de fertilizante.

Resolução

O problema já foi modelado denotando de x a produção total de milho em toneladas e de y a produção total de fertilizante em toneladas pelo sistema de equações lineares

$$\begin{cases} 0,9x - 0,5y = 4 \\ -0,8x + y = 2 \end{cases}.$$

Usando o produto de matrizes podemos escrever esse sistema na forma de equações $AX = C$, ou seja,

$$\begin{pmatrix} 0,9 & -0,5 \\ -0,8 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

Vamos usar o método da matriz inversa para resolver o sistema; então, aumentamos A com a matriz identidade e para facilitar os cálculos vamos escrever todas as entradas na forma fracionária

$$\begin{pmatrix} \frac{9}{10} & -\frac{5}{10} & | & 1 & 0 \\ -\frac{8}{10} & 1 & | & 0 & 1 \end{pmatrix}.$$

Agora, usamos as operações elementares para transformar o lado esquerdo da matriz acima na matriz identidade, após alguns cálculos temos:

$$\begin{pmatrix} 1 & 0 & | & 2 & 1 \\ 0 & 1 & | & \frac{8}{5} & \frac{9}{5} \end{pmatrix}.$$

A solução do sistema é dada por

$$X = \begin{pmatrix} 2 & 1 \\ \frac{8}{5} & \frac{9}{5} \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \end{pmatrix}.$$

Portanto, a solução do sistema é $x = 10$ e $y = 10$, como a solução do sistema é positiva ela é admissível para o problema; então, a produção total de milho é 10 toneladas e 10 toneladas de fertilizantes.

Exercício 2. Retomando o problema 56 para produção dos produtos A, B e C e usando a mesma matriz de insumo-produto, encontre a produção necessária para o mês de fevereiro do mesmo ano, sabendo-se que as demandas de consumo para os produtos A, B e C são 10, 12 e 15 respectivamente.

Resolução

Usando a nova matriz de demanda de consumo, o sistema pode ser escrito usando produto de matrizes $AX = C'$; ou seja,

$$\begin{pmatrix} 0,8 & -0,3 & -0,4 \\ -0,3 & 0,9 & -0,2 \\ -0,2 & -0,4 & 0,7 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \\ 15 \end{pmatrix}.$$

Usando a matriz inversa da matriz A, calculada no exemplo 56 podemos encontrar a solução do sistema fazendo $X = A^{-1}C'$; ou seja,

$$X = \begin{pmatrix} \frac{110}{49} & \frac{74}{49} & \frac{12}{7} \\ \frac{50}{49} & \frac{96}{49} & \frac{8}{7} \\ \frac{60}{49} & \frac{76}{49} & \frac{18}{7} \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 12 \\ 15 \end{pmatrix} = \begin{pmatrix} \frac{464}{7} \\ \frac{356}{7} \\ \frac{486}{7} \end{pmatrix}.$$

Portanto, a solução do sistema é $x = \dfrac{464}{7}$, $y = \dfrac{356}{7}$ e $z = \dfrac{486}{7}$, como a solução do sistema é positiva ela é admissível para o problema; então, a produção total dos bens A, B e C é de aproximadamente 66 unidades, 51 unidades e 69 unidades respectivamente.

Exercício 3. A economia na ilha Baco produz somente uvas e vinho. A produção de 1 quilo de uvas requer $\dfrac{1}{2}$ quilo de uvas, e nenhum vinho. A produção de 1 litro de vinho requer $\dfrac{1}{2}$ quilo de uvas e $\dfrac{1}{4}$ litro de vinho. Os moradores da ilha exigem 1 quilo de uvas e 3 litros de vinho para consumo próprio. Equacione o sistema de insumo-produto da economia desta ilha e descubra qual deve ser a produção total de uvas e de vinho.

Resolução

As quantidades de insumo necessárias para produção de 1 quilo de uva e 1 litro de vinho são:

(i) $\dfrac{1}{2}$ quilo de uvas para produção de 1 quilo de uva.

(ii) $\dfrac{1}{2}$ quilo de uvas para produção de 1 litro de vinho.

(iii) $\dfrac{1}{4}$ litro de vinho para produção de 1 litro de vinho.

Assim para produção de x quilos de uva e y litros de vinho são necessárias:

(i) $\dfrac{1}{2}x$ quilo de uvas para produção de x quilos de uvas.

(ii) $\dfrac{1}{2}y$ quilo de uvas para produção de y litros de vinho.

(iii) $\dfrac{1}{4}y$ litro de vinho para produção de y litros de vinho.

Como os moradores da ilha desejam 1 quilo de uvas e 3 litros de vinho para consumo próprio, as equações que descrevem a produção total de uvas e vinho são:

$$\begin{cases} x &= \frac{1}{2}x + \frac{1}{2}y + 1 \\ y &= \qquad\;\; \frac{1}{4}y + 3 \end{cases}.$$

Rearranjando o sistema temos

$$\begin{cases} \frac{1}{2}x - \frac{1}{2}y &= 1 \\ \qquad\;\; \frac{3}{4}y &= 3 \end{cases}.$$

Na forma matricial $AX = C$ temos

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{3}{4} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Calculando o determinante da matriz de coeficientes $\det(A) = \dfrac{1}{2} \cdot \dfrac{3}{4} - \dfrac{1}{2} \cdot 0 = \dfrac{3}{8} \neq 0$; então, o sistema possui solução.

Usando o método descrito no exemplo 54 para determinar a inversa da matriz de coeficientes A encontramos

$$A^{-1} = \begin{pmatrix} 2 & \frac{4}{3} \\ 0 & \frac{4}{3} \end{pmatrix}.$$

Para obtermos a solução do sistema basta calcular o produto

$$X = \begin{pmatrix} 2 & \frac{4}{3} \\ 0 & \frac{4}{3} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}.$$

Portanto, a produção total de uvas será de 6 quilos e de vinho será de 4 litros.

Exercício 4. Suponha, agora, que a produção de 1 quilo de uvas requeira $\dfrac{7}{8}$ litro de vinho. Sem alterar os demais coeficientes de insumo-produto, escreva o novo sistema e encontre a produção necessária para satisfazer tal demanda.

Resolução

As novas equações de insumo-produto e de demanda de consumo serão:

$$\begin{cases} x &= \frac{1}{2}x+\frac{1}{2}y+1 \\ y &= \frac{7}{8}x+\frac{1}{4}y+3 \end{cases}.$$

Rearranjando o sistema temos

$$\begin{cases} \frac{1}{2}x-\frac{1}{2}y &= 1 \\ -\frac{7}{8}x+\frac{3}{4}y &= 3 \end{cases}.$$

Na forma matricial $AX=C$ temos

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{7}{8} & \frac{3}{4} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Calculando o determinante da matriz de coeficientes $\det A = \frac{1}{2}\cdot\frac{3}{4} - \left(-\frac{1}{2}\right)\cdot\left(-\frac{7}{8}\right) = -\frac{1}{16} \neq 0$, então o sistema possui solução.

Calculando a matriz a inversa de A encontramos

$$A^{-1} = \begin{pmatrix} -12 & -8 \\ -14 & -8 \end{pmatrix}.$$

Para obter a solução do sistema basta calcular o produto

$$X = \begin{pmatrix} -12 & -8 \\ -14 & -8 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -36 \\ -38 \end{pmatrix}.$$

A solução do sistema é $x=-36$ e $y=-38$, mas esses valores não são admissíveis como solução do problema, pois não é possível produzir quantidades negativas de uvas e de vinho. Portanto. o problema de produção não tem solução admissível.

Aula 7

Essa aula é destinada à correção dos exercícios. Como sugestão propomos que seja feita na forma de seminário, onde cada grupo apresentará a resolução de um exercício.

Observação 9. Neste momento o professor poderia analisar se valeria a pena discutir se as soluções admissíveis encontradas foram fruto do mero acaso ou se as matrizes envolvidas em tais problemas tinham alguma propriedade especial em comum. Se julgar que sim, aconselhamos o estudo que vem a seguir e que fala das matrizes diagonais dominantes.

Aula 8

Comentários

Todos os sistemas exemplificados até aqui, posuem solução, já que a matriz de coeficientes associada a cada um deles é invertível.

Pergunta: será que tivemos sorte em nos deparar com sistemas cuja matriz de coeficientes associada é invertível? Esse fato é uma mera coincidência?

Notamos também que o fato do sistema ter solução não quer dizer que o problema tenha solução, pois o problema requer uma solução não-negativa e encontramos sistemas com solução positiva e outros, como no último exemplo, que são negativas.

Pergunta: será que é possível prever quando um sistema terá solução não-negativa e assim faça sentido como solução para os problemas de processos de produção?

Para responder a essas perguntas vamos falar de um tipo especial de matriz invertível, a **matriz diagonal dominante**.

Chamamos de matriz diagonal dominante uma matriz quadrada que satisfaz as três propriedades seguintes:

(i) cada entrada fora da diagonal é ≤ 0,

(ii) cada entrada diagonal é positiva, e

(iii) a soma das entradas em cada coluna é positiva.

Como exemplo vamos analisar as matrizes de coeficientes dos sistemas de equações lineares que modelaram os problemas de processos de produção que vimos até aqui.

No exemplo 54 e 55 a matriz de coeficientes

$$\begin{pmatrix} 0,95 & -0,01 \\ -0,3 & 0,9 \end{pmatrix}$$

é diagonal dominante pois satisfaz as três propriedades:

(i) cada entrada fora da diagonal é ≤ 0;

(ii) cada entrada diagonal é positiva e

(iii) a soma das entradas em cada coluna é positiva pois $0,95+(-0,3)=0,65$ e $-0,01+0,9=0,89$.

Nesses exemplos vimos que as soluções dos sistemas são positivas.

No exemplo 56 e exercício proposto 2 a matriz de coeficientes

$$\begin{pmatrix} 0,8 & -0,3 & -0,4 \\ -0,3 & 0,9 & -0,2 \\ -0,2 & -0,4 & 0,7 \end{pmatrix}$$

é matriz diagonal dominante pois satisfaz as propriedades (*i*) e (*ii*) e como $0,8+(-0,3)+(-0,2)=0,3>0$, $-0,3+0,9+(-0,4)=0,2>0$ e $-0,4+(-0,2)+0,7=0,1>0$, satisfaz também (*iii*). Nesses casos as soluções dos sistemas também são positivas.

No exercício 1 a matriz de coeficientes

$$\begin{pmatrix} 0,9 & -0,5 \\ -0,8 & 1 \end{pmatrix}$$

também é diagonal dominante pois satisfaz as propriedades (*i*) e (*ii*) e como $0,9+(-0,8)=0,1>0$ e $-0,5+1=0,5>0$, satisfaz também (*iii*). Nesse caso a solução do sistema também é positiva.

No exercício 3 a matriz de coeficientes

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{3}{4} \end{pmatrix}$$

também é diagonal dominante pois satisfaz as propriedades (*i*) e (*ii*) e como $\dfrac{1}{2}+0=\dfrac{1}{2}$ e $-\dfrac{1}{2}+\dfrac{3}{4}=\dfrac{1}{4}$, satisfaz também (*ii*). Nesse caso a solução do sistema também é positiva.

No exercício 4 a matriz de coeficientes

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{7}{8} & \frac{3}{4} \end{pmatrix}$$

não é diagonal dominante pois não satisfaz a propriedade (*iii*). De fato, a soma das entradas na primeira coluna é $\dfrac{1}{2}+\left(-\dfrac{7}{8}\right)=-\dfrac{3}{8}$. Lembremos que neste caso uma demanda positiva resultou em solução negativa e portanto não admissível.

Para nosso estudo vamos considerar apenas sistemas 2×2, isto é, um sistema com duas equações e duas incógnitas

$$AX=C,$$

em que A satisfaz as três condições descritas acima, ou seja, A é diagonal dominante e C é uma demanda de consumo admissível para o problema; ou seja, apresenta todas as entradas não-negativas. Para frisar o sinal de cada elemento da matriz A e facilitar o entendimento vamos escrevê-la como

$$A = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix}$$

onde os números $a, b > 0$ e $c, d \geq 0$.

Para resolver esse sistema vamos usar o método da matriz inversa. Para tanto vamos calcular a matriz inversa da matriz A.

Inicialmente aumentamos a matriz A com a matriz identidade I

$$\begin{pmatrix} A & | & I \end{pmatrix} = \begin{pmatrix} a & -b & | & 1 & 0 \\ -c & d & | & 0 & 1 \end{pmatrix}.$$

Lembrando que nosso objetivo é transformar a primeira metade da matriz acima na matriz identidade; para isso vamos aplicar operações elementares sobre suas linhas.

Para começar supomos $a \neq 0$, multiplicamos a primeira linha por $\frac{c}{a}$ e somamos o resultado na segunda linha, obtendo

$$\begin{pmatrix} a & -b & | & 1 & 0 \\ 0 & \frac{ad-bc}{a} & | & \frac{c}{a} & 1 \end{pmatrix}.$$

Este simples cálculo já nos diz que, quando $a \neq 0$, A é invertível se, e somente se, $ad - bc \neq 0$. Agora, continuamos multiplicando a primeira linha por $\frac{1}{a}$ e a segunda linha por $\frac{a}{ad - bc}$ obtendo

$$\begin{pmatrix} 1 & -\frac{b}{a} & | & \frac{1}{a} & 0 \\ 0 & 1 & | & \frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$$

Para completar multiplicamos a segunda linha por $\frac{b}{a}$ e somamos o resultado na primeira linha obtendo

$$\begin{pmatrix} 1 & 0 & | & \frac{d}{ad-bc} & \frac{b}{ad-bc} \\ 0 & 1 & | & \frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$$

Concluímos, então, que se

$$A = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix} \text{ então}$$

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & b \\ c & a \end{pmatrix}.$$

Lembrando que $a, b > 0$ e $c, d \geq 0$. Analisemos o sinal de $ad - bc$.

Como A é matriz diagonal dominante, temos:

$$\begin{cases} a-c>0 \Rightarrow a>c \\ d-b>0 \Rightarrow d>b \end{cases} \Rightarrow ad-bc>0.$$

Assim concluímos que as entradas de A^{-1} são todas não negativas. Isto é:

Teorema 28. Seja A matriz 2×2 diagonal dominante. Então, A é invertível e as entradas de A^{-1} são todas não negativas.

Lembrando que a solução do sistema $AX=C$ é $X=A^{-1}C$ obtemos

$$X=\begin{pmatrix} \frac{d}{ad-bc} & \frac{b}{ad-bc} \\ \frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}=\begin{pmatrix} \frac{d\alpha+b\beta}{ad-bc} \\ \frac{c\alpha+a\beta}{ad-bc} \end{pmatrix}.$$

Deste cálculo concluímos:

Teorema 29. Se A é matriz 2×2 diagonal dominante e se C é matriz com entradas não negativas então a solução X de $AX=C$ terá apenas entradas não negativas.

É possível mostrar que um sistema de três equações e três incógnitas $AX=C$, com A diagonal dominante e C apresentando somente entradas não-negativas possui solução não-negativa. Na realidade este resultado vale para qualquer sistema envolvendo matriz diagonal dominante de ordem n.

Logo, todos os problemas referentes a processos de produção que vimos envolveram matrizes diagonais dominantes, com exceção do 4. Logo, não foi sorte e nem acaso que as soluções referentes a demandadas não negativas gerassem soluções admissíveis para os problemas, isto é, soluções não negativas.

5.8 Possíveis continuações ou desdobramentos

Nesse plano de aulas foram utilizados problemas de produção que são modelados por sistemas de equações lineares 2×2 ou 3×3, pois analisamos economias como de um fazenda ou até mesmo de alguns setores de uma cidade pequena, mas se estudarmos a economia de uma cidade inteira de forma mais abrangente, de um estado ou até mesmo de um país, vamos nos deparar com sistemas de equações lineares mais gerais. Nesse sentido caso o professor perceba interesse dos alunos no estudo de tais problemas e se for possível o uso de computadores com conexão a internet, pode-se trabalhar problemas de produção em economias maiores como do exemplo 53 do capítulo anterior, utilizando softwares para auxiliar na resolução dos mesmos. Como sugestão o professor pode usar o site <https://matrixcalc.org/pt/> que oferece uma calculadora de matrizes online, que além de realizar operações entre matrizes, calcula: determinantes de diversas formas, matriz inversa, matriz transposta, posto de uma matriz, escalona matrizes e também soluciona sistemas de equações lineares de diversas formas, inclusive usando o método da matriz inversa.

REFERÊNCIAS

BALDIN, Y. Y.; FURUYA, Y. K. **Geometria Analítica para todos e atividades com Octave e GeoGebra**. [S.l.]: EdUFSCar, 2011. Citado na página 17.

BOLDRINI, J. L.; COSTA, S. I.; FIGUEREDO, V.; WETZLER, H. G. **Álgebra linear**. [S.l.]: Harper & Row, 1980. Citado na página 17.

HEFEZ, A.; FERNANDEZ, C. d. S. **Introdução à álgebra linear**. [S.l.]: SBM, 2016. Citado na página 17.

IEZZI, G.; HAZZAN, S. **Fundamentos de matemática elementar, 4: sequências, matrizes, determinantes, sistemas**. [S.l.]: Atual, 1995. v. 4. Citado na página 17.

LADEIRA, L. A. C. **Álgebra Linear e Equações Diferenciais**. [S.l.]: ICMC - USP, 2007. Citado na página 73.

SIMON, C. P.; BLUME, L.; DOERING, C. I. **Matemática para economistas**. [S.l.]: Bookman, 2004. Citado nas páginas 17, 73 e 95.

SÃO PAULO (Estado), Secretaria da Educação. **Currículo do Estado de São Paulo: Matemática e suas tecnologias / Secretaria da Educação**. coordenação geral, Maria Inês Fini; coordenação de área, Nilson José Machado: São Paulo - SEE, 2010. Citado na página 109.

ZANI, S. L. **Álgebra Linear**. [S.l.]: ICMC - USP, 2007. Citado na página 73.