Universidade Federal de Mato Grosso
Instituto de Ciências Exatas e da Terra
Departamento de Matemática

Matrizes e algumas aplicações

Dalcimar Badio Barbosa de Moraes
Mestrado Profissional em Matemática: PROFMAT/SBM

Orientadora: **Prof[a]. Dra. Eunice Cândida Pereira Rodrigues**

Trabalho financiado pela Capes

Cuiabá - MT
25 de abril de 2017

Matrizes e algumas aplicações

Este exemplar corresponde à redação final da dissertação, devidamente corrigida e defendida por Dalcimar Badio Barbosa de Moraes e aprovada pela comissão julgadora.

Cuiabá, 25 de abril de 2017.

Profª. Dra. Eunice Cândida Pereira Rodrigues
Orientadora

Banca examinadora:

Prof. Dra. Eunice Cândida Pereira Rodrigues
Prof. Dra. Ivonildes Ribeiro Martins Dias
Prof. Dr. Dercio Braga Santos

Dissertação apresentada ao curso de Mestrado Profissional em Matemática – PROFMAT, da Universidade Federal de Mato Grosso, como requisito parcial para obtenção do título **de Mestre em Matemática**.

Dados Internacionais de Catalogação na Fonte.

B238m Moraes, Dalcimar Badio Barbosa de.
Matrizes e algumas aplicações / Dalcimar Badio Barbosa de Moraes. -- 2017
xii, 78 f. : il. color. ; 30 cm.

Orientadora: Eunice Cândida Pereira Rodrigues.
Dissertação (mestrado profissional) - Universidade Federal de Mato Grosso, Instituto de Ciências Exatas e da Terra, Programa de Pós-Graduação em Matemática, Cuiabá, 2017.
Inclui bibliografia.

1. Fatorações. 2. Sequência de Fibonacci. 3. Cadeia de Markov. 4. Material de apoio. I. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Dissertação de Mestrado defendida em 31 de março de 2017 e aprovada pela banca examinadora composta pelos Professores Doutores

---

Prof[a]. Dra. Eunice Cândida Pereira Rodrigues

Prof[a]. Dr. Ivonildes Ribeiro Martins Dias

Prof. Dr. Dercio Braga Santos

À minha mãe Aniversina pela sabedoria dedicada a mim e aos meus irmãos, por sempre acreditar e incentivar a buscar meus objetivos. E à minha amada esposa pela segurança de vida.

Agradecimentos

A Deus pela a vida e por ter colocado estas pessoas no meu caminho: a minha segunda mãe: Coraci (madrinha e irmã); e ao meu segundo pai: Generoso (padrinho e cunhado), pelo apoio nas horas difíceis e pelo abrigo que me deram na sua casa e nos seus corações.

A minha família pelo apoio, que fez sentir-me capaz de construir este trabalho.

Aos colegas: André, Cândido, Marcos Terra, Gileno e Fábio, companheiros de viagens e de estudos.

A minha orientadora Eunice, exemplo e inspiração, pelo seu imenso conhecimento colocado a minha disposição que contribuiu para o meu crescimento profissional.

Aos meus professores que tanto me ensinaram.

Aos companheiros de escola e, segunda casa, a grande família Daniel Martins Moura.

Por fim, os meus agradecimentos também a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela concessão da bolsa durante todo o período de realização deste mestrado.

Muito obrigado a todos.

"A educação é um ato de amor e, portanto,
um ato de coragem. Não pode temer o debate,
a análise da realidade;
não pode fugir à discussão criadora,
sob pena de ser uma farsa. "

Paulo Freire.

Resumo

Com um estudo recente de apenas 150 anos as matrizes ganharam tamanha importância que não se consegue conceber, hoje, a ideia de computadores, engenharia civil, elétrica, mecânica, meteorologia, oceanografia entre outras inúmeras áreas, sem o estudo delas. Dessa forma neste trabalho faz-se uma abordagem sobre matrizes, enfatizando as fatorações LU, cholesky e QR. Além disso, apresentam-se algumas aplicações relacionadas com a Criptografia, sequência de Fibonacci e cadeia de Markov. Visamos com este trabalho a elaboração de um material de apoio pedagógico aos docentes do ensino médio. O nosso suporte teórico foi alicerçado nos autores Anton et al. (2001); Boldrini et al. (1980); Lima (2009); Hefez e Fernandez (2012).

Palavras chave: Fatorações, Sequência de Fibonacci, Cadeia de Markov, Material de apoio.

Abstract

In a recent study around only 150 years, the Matrices have become so important that the idea of computers, civil engineering, electrical engineering, meteorology, oceanography and many other areas can not be conceived without their study. In that way, the work in question has made approaching on matrices, emphasizing LU, cholesky and QR. In addition, some applications related to Cryptography, Fibonacci Sequence and Markov chain are presented. It is our goal with this work to develop a pedagogical support material for high school teachers. Our theoretical support was based on the authors Anton et al. (2001); Boldrini et al. (1980); Lima (2009); Hefez e Fernandez (2012).

Keywords: Fibonacci, Fibonacci sequence, Markov chain, support material.

Sumário

Lista de Figuras

Lista de Tabelas

Introdução

Segundo Cruz et al. (2012), Arthur Cayley (1821-1895) foi um dos pioneiros no estudo das matrizes e, por volta de 1850, divulgou esse nome e passou a demonstrar sua aplicação. As matrizes, inicialmente, eram aplicadas quase que exclusivamente na resolução de sistemas lineares e apenas há pouco mais de 150 anos tiveram sua importância detectada. No entanto, o primeiro uso implícito da noção de matriz se deve a Joseph Louis Lagrange (1736-1813), em 1790; e o primeiro a lhe dar um nome parece ter sido Augustin-Louis Cauchy (1789-1857), que as chamava de "tabelas"; e o nome "matriz"só veio com James Joseph Sylvester (1814-1897) em 1850. Sylvester observava as matrizes como mero ingrediente dos determinantes. Somente com Cayley elas passaram a ter vida própria e, gradativamente, começaram a suplantar os determinantes em importância.

Uma preocupação na educação é como escolher um conteúdo que crie uma situação complexa, para que o aluno adquira habilidade e, por consequência, passe a ter novas competências, utilizando-se de uma situação que faça parte de seu cotidiano, de sua vida, para que ele se sinta parte da situação, provocando nele curiosidade e observando o que nos diz Médio (2006),

> Os objetivos do Ensino Médio em cada área do conhecimento devem envolver, de forma combinada, o desenvolvimento de conhecimentos práticos, contextualizados, que respondam às necessidades da vida contemporânea, e o desenvolvimento de conhecimentos mais amplos e abstratos, que correspondam a uma cultura geral e a uma visão de mundo.

logo para o que professor tenha clareza de que habilidades os alunos terão que atingir, onde são usadas e como serão atingidas, faz-se necessário que o professor tenha consciência de que precisa inserir novas metodologias no ensino dos conteúdos que compõem a matemática básica. Um destes conteúdos, por exemplo, são as matrizes.

Diante do emposto, neste trabalho fez-se uma abordagem sobre matrizes. Especificamente, no primeiro capítulo, apresentamos os conceitos básicos de matrizes, os quais se espera que um aluno do ensino médio tenha competência, ou seja, tenha habilidade de fazer as operações com matrizes, calcular determinantes, fazer escalonamento através do método de Gauss, resolver sistemas lineares, entre outros.

Já no segundo capítulo, temos a justificativa do porquê se ensina os conceitos abstratos ao aluno advindo do ensino médio. Essas competências serão utilizadas para

uma boa assimilação matricial de vetores, combinação linear, transformação linear e decomposições lineares.

Por fim, no último capítulo trazemos três aplicações de matrizes, quer sejam: na Criptografia, sequência de Fibonacci e cadeia de Markov, para que as habilidades definidas no primeiro capítulo sejam construídas de modo significativo para os alunos. Dessa forma, espera-se que o estudo de matrizes seja interessante e contextualizado. Sendo que, Criptografia faz parte da vida cotidiana dos alunos estando presente em celulares, bancos entre outros; já a sequência de Fibonacci é uma aplicação interessante que faz parte da história da matemática entre outras peculiaridades; e a cadeia de Markov é aplicada em várias áreas do conhecimento, sendo que a mesma é utilizada em agricultura, previsão do tempo, entre outras situações que faz parte do dia a dia do aluno.

Capítulo 1

Noções básicas de matrizes

Neste capítulo, apresentaremos algumas definições e resultados preliminares sobre matrizes, os quais podem ser encontrados em Boldrini et al. (1980); Anton et al. (2001); Hefez e Fernandez (2012).

1.1 Matrizes

Definição 1.1.1 *Matriz é uma lista de números representado em forma retangular onde as filas horizontais são chamadas de linhas e as filas verticais são chamadas de colunas.*

Uma matriz com m linhas e n colunas ($m, n \in \mathbb{N}$) é chamada de matriz m por n (escreve-se $m \times n$), e os valores m e n são suas dimensões, tipo ou ordem. Uma matriz é representada por uma letra maiúscula e seus elementos por letras minúsculas, assim, uma matriz com m linhas e n colunas pode ser escrita

$$A_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]_{m\times n}\,.$$

Ou simplesmente $A = [a_{ij}]$, onde a_{ij} é o elemento (ou a entrada) da linha i e da coluna j. O símbolo $M_{(m\times n)}$ denota o conjunto das matrizes $m \times n$.

Dependendo dos valores de $m \times n$ as matrizes recebem nomes especiais, sendo que uma matriz $1 \times n$ é chamada de matriz linha enquanto que $m \times 1$ recebe o nome de matriz coluna e uma matriz $n \times n$ é chamada de matriz quadrada de ordem n.

Quando $A = [a_{ij}]$ é uma matriz quadrada de ordem n, os seus elementos a_{ii} formam a diagonal principal, sendo $a_{ij} = 0$ se $i \neq j$ recebe o nome de matriz diagonal ou

seja

$$\begin{bmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{bmatrix}$$

é uma matriz diagonal, e uma matriz diagonal que $a_{ii} = 1$

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

recebe o nome de matriz identidade, que normalmente é simbolizada pela letra I_n ou simplesmente por I.

Uma matriz triangular superior é uma matriz quadrada que seus elementos abaixo da diagonal principal é zero

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

portanto $a_{ij} = 0$, se $i > j$. Analogamente uma matriz é triangular inferior se os seus elementos acima da diagonal principal forem zero

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

portanto $a_{ij} = 0$, se $i < j$. Mas se $a_{ij} = 0$ para todos os elementos desta matriz

$$\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

ela é denominada matriz nula. Uma matriz quadrada é dita simétrica se $a_{ij} = a_{ji}$, um exemplo é

$$\begin{bmatrix} a & b & c & d \\ b & e & f & g \\ c & f & i & h \\ d & g & h & k \end{bmatrix}.$$

Observe que, no caso de uma matriz simétrica, a parte superior é uma reflexão da parte inferior, em relação à diagonal principal.

1.2 Operações matriciais

Nesta seção definiremos algumas propriedades das operações com matrizes. Veremos que muitas das regras básicas dos números reais também valem para matrizes.

1.2.1 Matriz transposta

Sendo $A = [a_{ij}]_{m\times n}$, podemos obter uma matriz $A^t = [b_{ij}]_{n\times m}$, cujas linhas são as colunas de A, isto é, $b_{ij} = a_{ji}$. A^t é denominada transposta de A.

Exemplo 1.2.1 *A transposta de* $\begin{pmatrix} 1 & 3 & 5 \\ 2 & 7 & 8 \\ 1 & 5 & 9 \end{pmatrix}$ *é* $\begin{pmatrix} 1 & 2 & 1 \\ 3 & 7 & 5 \\ 5 & 8 & 9 \end{pmatrix}$.

1.2.2 Soma e subtração de matrizes

Definição 1.2.1 *Seja* $A = [a_{ij}], B = [b_{ij}] \in M_{(m\times n)}$ *a soma de* A *e* B*, denotada por* $A + B$ *é a matriz* $S = [s_{ij}] \in M_{(m\times n)}$*, tal que:*

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

para cada i *de* $1 \leqslant i \leqslant m$ *e cada* j *de* $1 \leqslant j \leqslant n$.

Lembrando que a subtração entre matrizes é definida de modo análogo a soma de matrizes, ou seja:

$$A - B = [a_{ij}] + (-[b_{ij}]) = [a_{ij} - b_{ij}] = [s_{ij}] = S$$

Exemplo 1.2.2 *Note que a soma das matrizes abaixo obedece as regras da adição dos números reais.*

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

Propriedade 1.2.1 *Dadas as matrizes* A, B *e* C *de mesma ordem* $m \times n$*, temos:*
i) $A + B = B + A$ *(Lei da comutatividade para a adição);*
ii) $A + (B + C) = (A + B) + C$ *(Lei da Associatividade da adição);*
iii) $A + 0 = A$ *onde 0 é denota a matriz nula* $m \times n$.

Demostrações: Seja $A = [a_{ij}]$, $B = [b_{ij}]$, e $C = [c_{ij}]$ então:
i) $A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = B + A$;

ii) $A+(B+C) = [a_{ij}]+[b_{ij}+c_{ij}] = [a_{ij}+(b_{ij}+c_{ij})] = [(a_{ij}+b_{ij})+c_{ij}] = [a_{ij}+b_{ij}]+[c_{ij}] = (A+B)+C$;
iii) $A + 0 = [a_{ij}] + [0_{ij}] = [a_{ij} + 0_{ij}] = [a_{ij}] = A$. ■

1.2.3 Multiplicação por um escalar

Definição 1.2.2 *Se A é uma matriz e c é um escalar, então o produto $c \cdot A$ é a matriz obtida pela multiplicação de cada elemento de A por c.*

Exemplo 1.2.3 *Seja k um número real e $A = \begin{bmatrix} 1 & 5 \\ 2 & 10 \end{bmatrix}$, então $k \cdot A$ será:*

$$k \cdot \begin{bmatrix} 1 & 5 \\ 2 & 10 \end{bmatrix} = \begin{bmatrix} k \cdot 2 & k \cdot 3 \\ k \cdot 3 & k \cdot (-10) \end{bmatrix} = \begin{bmatrix} 2k & 3k \\ 3k & -10k \end{bmatrix}.$$

Propriedade 1.2.2 *Dadas as matrizes B e C de mesma ordem $m \times n$ e escalares a e b, temos:*
i) $a(B + C) = aB + aC$;
ii) $(a + b)C = aC + bC$;
iii) $a(bC) = (ab)C$.

Demonstrações: Sejam $B = [b_{ij}]$, $C = [c_{ij}]$ elementos de $M_{(m\times n)}$, a e b números reais, então:
i) $a(B + C) = a[b_{ij} + c_{ij}] = [ab_{ij} + ac_{ij}] = [ab_{ij}] + [ac_{ij}] = a[b_{ij}] + a[c_{ij}] = aB + aC$;
ii) $(a+b)C = (a+b)[c_{ij}] = [(a+b)c_{ij}] = [(ac_{ij})+(bc_{ij})] = [(ac_{ij})]+[(bc_{ij})] = a[c_{ij}]+b[c_{ij}] = aC + bC$;
iii) $a(bC) = a(b[c_{ij}]) = a[bc_{ij}] = [abc_{ij}] = [(ab)c_{ij}] = (ab)[c_{ij}] = (ab)C$.

■

Observação 1.2.1 *Quando discutimos matrizes, é usual chamar as quantidades numéricas de escalares, salvo menção explicita ao contrário, escalares são números reais.*

1.2.4 Produto entre matrizes

Definimos a multiplicação de uma matriz por um escalar, mas não a multiplicação de duas matrizes. Como matrizes são somadas somando os elementos correspondentes e subtraídas subtraindo elementos correspondentes, pareceria natural definir a multiplicação de matrizes multiplicando os elementos correspondentes. Ocorre que tal definição não seria muito útil na maioria dos problemas. A experiência levou os matemáticos à seguinte definição para multiplicação de matrizes.

Definição 1.2.3 *Se A é uma matriz $m \times r$ e B é uma matriz $r \times n$, então o produto AB é a matriz $m \times n$, sendo os elementos determinados como segue. Para obter o elemento da linha i e da coluna j de AB, multiplique a linha i de A com a coluna j de B, e adicione os produtos resultantes. Ou mais especificamente, sendo $A = [a_{ij}]_{m\times r}$, $B = [b_{ij}]_{r\times n}$ e $AB = [c_{ij}]_{m\times n}$, temos:*

$$c_{ij} = \sum_{p=1}^{r} a_{ip}b_{pj} = a_{i1}b_{1j} + a_{i2}b_{2j} \cdots + a_{ir}b_{rj}$$

para cada par i e j com $1 \leqslant i \leqslant m$ e $1 \leqslant j \leqslant n$.

Observe que a definição de multiplicação de matrizes exige que o número de colunas da matriz A seja igual ao número de linhas da matriz B, para que seja possível formar o produto AB.

Exemplo 1.2.4 *Suponhamos que a seguinte matriz forneça as quantidades das vitaminas A, B e C obtidas em cada unidade dos alimentos I e II.*

$$\begin{array}{c} \\ I \\ II \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \left[\begin{array}{ccc} 4 & 3 & 0 \\ 5 & 0 & 1 \end{array}\right] \end{array}$$

Se ingeridos 5 unidades do alimento I e 2 unidades do alimento II, quanto consumiremos de cada tipo de vitamina?

Podemos representar o consumo dos alimentos I e II pela matriz consumo:

$$\left[\begin{array}{cc} 5 & 2 \end{array}\right].$$

A operação que vai nos fornecer a quantidade ingerida de cada vitamina é o produto:

$$\left[\begin{array}{cc} 5 & 7 \end{array}\right] \cdot \left[\begin{array}{ccc} 4 & 3 & 0 \\ 5 & 0 & 1 \end{array}\right] = \left[\begin{array}{ccc} 5\cdot 4 + 2\cdot 5 & 5\cdot 3 + 2\cdot 0 & 5\cdot 0 + 2\cdot 1 \end{array}\right] = \left[\begin{array}{ccc} 30 & 15 & 2 \end{array}\right],$$

isto é, serão ingeridas 30 unidades de vitamina A, 15 de B e 2 de C.

Propriedade 1.2.3 *Quando é possível as operações, as seguintes propriedades são válidas:*
i) $AB \neq BA$ *(podendo mesmo um dos membros estar definido e o outro não);*
ii) $AI = IA = A$ *(Isto justifica o nome da matriz identidade);*
iii) $A(BC) = (AB)C$ *(Lei da Associatividade da multiplicação);*
iv) $A(B + C) = AB + AC$ *(Lei da distributividade à esquerda em relação a soma);*
v) $(A + B)C = AC + BC$ *(Lei da distributividade à direita em relação a soma);*
vi) $0 \cdot A = A \cdot 0 = 0$.

i) (Contra exemplo)
Sendo $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & 1 \\ 2 & 0 \end{bmatrix}$, temos:

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 1 \\ 8 & 2 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 3 & 1 \\ 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 7 \\ 2 & 4 \end{bmatrix}$$

logo, $AB \neq BA$.

Demonstração:

iii) Sejam $A_{m\times p}$, $B_{p\times q}$, $C_{q\times n}$ matrizes e representaremos o produto entre as três matrizes, respectivamente, como $[A(BC)]_{mn}$, logo:

$$[A(BC)]_{ij} = \sum_{k=1}^{p} a_{ik}(BC)_{kj} = \sum_{k=1}^{p} a_{ik}\left(\sum_{l=1}^{q} b_{kl}c_{lj}\right) = \sum_{k=1}^{p}\sum_{l=1}^{q} a_{ik}(b_{kl}c_{lj}) =$$

$$\sum_{k=1}^{p}\sum_{l=1}^{q}(a_{ik}b_{kl})c_{lj} = \sum_{l=1}^{q}\sum_{k=1}^{p}(a_{ik}b_{kl})c_{lj} = \sum_{l=1}^{q}\left(\sum_{k=1}^{p} a_{ik}b_{kl}\right)c_{lj} = \sum_{l=1}^{q}(AB)_{il}c_{lj} = [(AB)C]_{ij}.$$

■

As demonstrações dos itens *ii*) *iv*) *v*) e *vi*) podem ser encontradas em Iezzi e Hazzan (1977).

1.3 Determinantes

Quando nos referimos ao determinante, isto é, ao número associado a uma matriz quadrada $A = [a_{ij}]$, denotado por, *det* A ou $|A|$ ou *det* $[a_{ij}]$. Neste contexto, define-se os determinantes para a matriz de ordem $1, 2$ e 3 respectivamente, como:

$$|a| = a$$

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

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \tag{1.1}$$

Antes de fazermos uma definição explícita para o determinante de ordem n, será conveniente lembrarmos do significado de *Permutação*.

Dado n objetos distintos $a_1, a_2, \cdots, a_n$, uma permutação destes objetos consiste em dispô-los em uma determinada ordem. Já a quantidade de permutação de n objetos é dada por $n!$ (lê-se n fatorial), sendo $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1$, para $n > 0$. Se $n = 0$, define-se 0!=1.

Definição 1.3.1 *Dada uma permutação dos inteiros* $1, 2, \cdots, n$, *existe uma inversão quando um inteiro precede outro menor que ele.*

Exemplo 1.3.1 *Na Tabela (1.1), descrita abaixo, mostra-se as* 3! *permutações de* $1, 2, 3$ *e o número de inversões.*

Permutações	Número de inversões
(1 2 3)	0
(1 3 2)	1
(2 1 3)	1
(2 3 1)	2
(3 1 2)	2
(3 2 1)	3

Tabela 1.1: Permutações

.

Voltemos ao determinante da matriz de ordem 3, observe em sua definição, Equação (1.1), que aparecem todos os produtos $a_{1j_1}a_{2j_2}a_{3j_3}$, onde $(j_1j_2j_3)$ são as permutações de $1, 2$ e 3. Além disso, observa-se que o sinal do produto é negativo, quando temos um número ímpar de inversões, Veja na tabela (1.1) para verificar os sinais.

Como generalização, o determinante de uma matriz quadrada $[a_{ij}]_{n\times n}$ é dado pela definição a seguir.

Definição 1.3.2 *Seja* $A = [a_{ij}]$ *uma matriz quadrada, o determinante desta matriz será:*

$$det[a_{ij}] = \sum_{\rho}(-1)^j a_{1j_1}a_{2j_2}\ldots a_{nj_n},$$

onde j *é o número de inversões da permutações* $(j_1, j_2, \ldots j_n)$, *e* ρ *indica que a soma é estendida a todas as* $n!$ *permutações de* $(1\ 2\cdots n)$.

Exemplo 1.3.2 *O determinante da matriz* $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 3 & 2 & 3 \end{bmatrix}$, *é encontrado assim:*

$$\begin{vmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 3 & 2 & 3 \end{vmatrix} = 2\cdot 3\cdot 3 - 2\cdot 1\cdot 2 - 1\cdot 1\cdot 3 + 1\cdot 1\cdot 3 + 1\cdot 1\cdot 2 - 1\cdot 3\cdot 3 = 7.$$

Propriedade 1.3.1 *Algumas propriedades de determinantes;*
i) Se todos os elementos de uma fila (linha ou coluna) de uma matriz A *são nulos,* $det\ A\ = 0$*;*

ii) Se multiplicarmos uma fila da matriz por uma constante, o determinante fica multiplicado por esta constante;
iii) Uma vez trocada a posição de duas filas, o determinante troca de sinal;
iv) O determinante de uma matriz que tem duas filas iguais é zero;
v) O determinante não se altera se adicionar uma fila a outra fila multiplicada por uma constante;
vi) $det\ (A \cdot B) = det\ A \cdot\ det\ B$.

As informações omitidas podem ser encontradas em Boldrini et al. (1980).

A seguir mostraremos outro caminho para encontrar o determinante de uma matriz, através do denominado *desenvolvimento de Laplace*.

1.3.1 Matriz dos cofatores

No início da seção (1.3), definimos determinante de uma matriz $A = A_{3\times 3}$ como

$$det\ (A) = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \quad (1.2)$$

Rearranjando os termos da Equação (1.2) e fatorando, podemos reescrevê-los como:

$$det\ (A) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}). \quad (1.3)$$

As expressões que estão entre parenteses em (1.3) são os determinantes:

$$M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad M_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad M_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

As submatrizes de A que aparecem nestes determinantes têm um nome especial que explicitaremos na definição a seguir.

Definição 1.3.3 *Se A é uma matriz quadrada. Então, o determinante menor da entrada a_{ij}, denotado por M_{ij}, é definido como o determinante da submatriz que sobra quando suprimimos a i-ésima linha e a j-ésima coluna de A. O número $(-1)^{i+j}M_{ij}$ é denotado por C_{ij} e é chamado o cofator de a_{ij}.*

Tendo em vista a definição acima, a expressão (1.3) pode ser escrita em termos de menores e cofatores como

$$det\ A = a_{11}M_{11} + a_{12}(-M_{12}) + a_{13}M_{13} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \quad (1.4)$$

ou seja, o determinante de A pode ser calculado multiplicando as entradas da primeira linha de A pelos cofatores correspondentes e somando os produtos que resultam.

Exemplo 1.3.3 *Seja* $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 4 & -2 \\ 0 & 6 & 1 \end{bmatrix}$. *Vamos calcular o determinante de A como expansão em cofatores ao longo da primeira linha de A:*
Por (1.4) temos que,

$$\begin{vmatrix} 1 & 2 & 3 \\ -1 & 4 & -2 \\ 0 & 6 & 1 \end{vmatrix} = 1\cdot\begin{vmatrix} 4 & -2 \\ 6 & 1 \end{vmatrix} - 2\cdot\begin{vmatrix} -1 & -2 \\ 0 & 1 \end{vmatrix} + 3\cdot\begin{vmatrix} -1 & 4 \\ 0 & 6 \end{vmatrix} = 1\cdot 16 - 2\cdot(-1) + 3\cdot(-6) = 0.$$

Note que a exemplo da equação (1.4) podemos rearranjar os termos em relação a outras linha ou coluna, como segue:

$$\begin{aligned} det\ A &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\ &= a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} \\ &= a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} \\ &= a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32} \\ &= a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} \\ &= a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33}. \end{aligned} \tag{1.5}$$

Estas equações são chamadas de Expansões em cofatores de *det A*.

Teorema 1.3.1 *O determinante de uma matriz A de tamanho* $n \times n$ *pode ser calculado multiplicando as entradas de qualquer linha (ou colunas) pelos seus cofatores e somando os produtos resultantes, ou seja, para quaisquer* $1 \leq i \leq n$ *e* $1 \leq j \leq n$,

$$det\ A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$
(expansão em cofatores ao longo da j-ésima coluna)

e

$$det\ A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.$$
(expansão em cofatores ao longo da i-ésima linha)

Exemplo 1.3.4 *Calcular o determinante da matriz A, do exemplo (1.3.3), mas agora utilizando uma coluna. Assim temos:*

$$\begin{vmatrix} 1 & 2 & 3 \\ -1 & 4 & -2 \\ 0 & 6 & 1 \end{vmatrix} = 1\cdot\begin{vmatrix} 4 & -2 \\ 6 & 1 \end{vmatrix} - (-1)\cdot\begin{vmatrix} 2 & 3 \\ 6 & 1 \end{vmatrix} + 0\cdot\begin{vmatrix} 2 & 3 \\ 6 & 1 \end{vmatrix} = 1\cdot 16 + 1\cdot(-16) + 0\cdot(-16) = 0.$$

Como esperado o resultado não foi alterado.

Por meio das matrizes dos cofatores de uma matriz A, a qual denotamos por $\bar{A}$, podemos obter a matriz adjunta de A que é a transposta da matriz dos cofatores e será denotada por $adj(A)$, ou seja:

$$adj(A) = \bar{A}^t.$$

Exemplo 1.3.5 *Seja* $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix}$, *encontre a matriz adjunta de A:*

Encontrando os cofatores,

$$M_{11} = +1 \cdot \begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} = 0, \quad M_{12} = -1 \cdot \begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} = 2, \quad M_{13} = +1 \cdot \begin{vmatrix} 1 & 0 \\ 3 & 0 \end{vmatrix} = 0,$$

$$M_{21} = -1 \cdot \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = -1, \quad M_{22} = +1 \cdot \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -5, \quad M_{23} = -1 \cdot \begin{vmatrix} 1 & 1 \\ 3 & 0 \end{vmatrix} = 3,$$

$$M_{31} = +1 \cdot \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = 1, \quad M_{32} = -1 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = 1 \quad e \quad M_{33} = +1 \cdot \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1.$$

Assim,

$$\bar{A} = \begin{bmatrix} 0 & 2 & 0 \\ -1 & -5 & 3 \\ 1 & 1 & -1 \end{bmatrix}.$$

Como o $adj(A) = \bar{A}^t$ *temos que,*

$$adj(A) = \begin{bmatrix} 0 & -1 & 1 \\ 2 & -5 & 1 \\ 0 & 3 & -1 \end{bmatrix},$$

como queríamos encontrar.

No exemplo anterior, temos $adj(A) = \begin{bmatrix} 0 & -1 & 1 \\ 2 & -5 & 1 \\ 0 & 3 & -1 \end{bmatrix}$, e fazendo $A \cdot adj(A)$ teremos:

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 & 1 \\ 2 & -5 & 1 \\ 0 & 3 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = 2I.$$

Utilizando o Teorema (1.3.1) podemos mostrar que $det\ A = 2$, ou seja, $A \cdot adj(A) = detA \cdot I$. Isto não acontece por sorte, este fato é válido para todas as matrizes.

Teorema 1.3.2 *O produto de uma matriz quadrada A por sua adjunta é igual ao produto do determinante de A com a identidade. Ou seja:*

$$A \cdot \bar{A}^t = A \cdot adj(A) = det\ A \cdot I.$$

Uma demostração para este Teorema pode ser encontrado em Boldrini et al. (1980).

1.4 Matriz inversível

Definição 1.4.1 *Dada uma matriz quadrada A de ordem n, chamamos de inversa de A a uma matriz quadrada denotada por A^{-1} de ordem n tal que $A \cdot A^{-1} = A^{-1} \cdot A = I$.*

Exemplo 1.4.1 *Seja* $A = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$ *e* $B = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}$ *temos que:*

$$A \cdot B = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \cdot \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \qquad \text{e}$$

$$B \cdot A = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2,$$

logo podemos afirmar que $B = A^{-1}$.

Propriedade 1.4.1 *Se uma matriz possui uma inversa, então esta inversa é unica.*

Demostração: Suponha que $A \cdot B_1 = I \quad e \quad A \cdot B_2 = I$, assim temos que

$$B_1 = B_1 \cdot I = B_1 \cdot (A \cdot B_2) = (B_1 \cdot A) \cdot B_2 = I \cdot B_2 = B_2 \quad \Rightarrow \quad B_1 = B_2.$$

■

Uma consequência imediata do Teorema (1.3.2) é que temos um método para encontrar A^{-1}, ou seja,

$$A \cdot adj(A) = det\ A \cdot I \quad \Rightarrow \quad A \cdot \left(\frac{1}{det(A)} \cdot adj(A) \right) = I \quad \Rightarrow \quad \frac{1}{det(A)} \cdot adj(A) = A^{-1}.$$

Exemplo 1.4.2 *Encontre a inversa da matriz* $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$, *sabendo que* $\begin{vmatrix} 2 & 1 \\ 0 & 2 \end{vmatrix} = 4$.

Sendo a matriz dos cofatores $\begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}$, *logo a matriz adjunta é* $\begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$ *assim a inversa será:*

$$A^{-1} = \frac{1}{4} \cdot \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Teorema 1.4.1 *Uma matriz quadrada A admite uma inversa se, e somente se, $detA \neq 0$.*

Demonstração: Suponhamos que $A_{n\times n}$ tenha inversa, isto é , existe A^{-1} tal que $A \cdot A^{-1} = I$. Usando o determinante temos (Proposição 1.3.1(vi)),

$$1 = det\ I = det\ (A \cdot A^{-1}) = det\ A \cdot det\ A^{-1}.$$

Logo, se A tem inversa, então $det\ A \neq 0$ e $det\ A^{-1} = \dfrac{1}{det\ A}$. Portanto $det\ A \neq 0$ é uma condição necessária para que A tenha inversa.

Reciprocamente pelo Teorema (1.3.2) $A \cdot adj(A) = detA \cdot I$. Se $detA \neq 0$, então $A \cdot \dfrac{1}{detA} \cdot adj(A) = I$, e como a inversa é única, então $A^{-1} = \dfrac{1}{detA} \cdot adj(A)$. ■

1.5 Eliminação de linhas e formas escalonadas

O método de eliminação em sistemas lineares de equação lineares, veja Seção (1.6.3), consiste em efetuar repetidamente operações elementares sobre esse sistema, obtendo sistemas equivalentes. Até reduzi-lo a um de fácil resolução. Nesta seção reinterpretamos o método nas matrizes, explicitando seu caráter algorítmico.

1.5.1 Operações elementares

Seja A uma matriz $m \times n$. Para cada $1 \leq i \leq m$, denotemos por L_i a i-ésima linha de A. Então define-se as operações nas linhas da matriz A como se segue:

i) Permutação das linhas L_i e L_j, indicada por $L_i \longleftrightarrow L_j$.

ii) Substituição de uma linha L_i pela adição desta mesma linha com c vezes uma outra linha L_j, indicada por $L_i \longrightarrow L_i + cL_j$.

iii) Multiplicação de uma linha L_i por um número real c não nulo, indicada por $L_i \longrightarrow cL_i$.

Propriedade 1.5.1 *Toda operação elementar e nas linhas de matrizes em $M_{(m\times n)}$ é reversível, ou seja, existe uma operação elementar è tal que $è(e(A)) = A$ e $e(è(A)) = A$, para todo $A \in M_{(m\times n)}$.*

Demonstração: Se e é uma operação elementar do tipo $L_i \longleftrightarrow L_j$, então tome $è = e$. Se e é uma operação elementar do tipo $L_i \longrightarrow cL_i$, então tome è como a operação do tipo $L_i \longrightarrow \dfrac{1}{c}L_i$. Finalmente, se e é uma operação do tipo $L_i \longrightarrow L_i + cL_j$, então tome è como a operação $L_i \longrightarrow L_i - cL_j$. ■

Sejam A e B matrizes de ordem $m \times n$. A matriz A é dita *equivalente por linhas* à matriz B, denotada por $A \sim B$ ou $A \longrightarrow B$, se B pode ser obtida de A pela aplicação sucessiva de um número finito de operações elementares sobre linhas.

1.5.2 Forma escalonada de Gauss

Definição 1.5.1 *Uma matriz retangular está na forma escalonada se ela satisfaz as três seguintes propriedades:*

i) Todas as linhas não nulas estão acima de qualquer linha só de zeros.

ii) O elemento líder[1] de cada linha está numa coluna a direita do elemento líder da linha acima.

iii) Todos os elementos de uma coluna abaixo de um elemento líder são zeros.

Se uma matriz em forma escalonada satisfaz as seguintes condições adicionais então ela está na forma escalonada reduzida.

iv) O elemento líder de cada linha não nula é 1.

v) Cada elemento é o único elemento não-nula em sua coluna.

Exemplo 1.5.1 *A matriz* $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ *esta na forma escalonada pois obedece as três primeiras condições supracita; já a matriz* $\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{bmatrix}$ *esta está escalonada na forma reduzida. Pois obedece todas as condições.*

Vejamos agora um algoritmo que reduz por linhas uma matriz dada não nula qualquer a uma matriz na forma escalonada. O termo reduzir por linhas significa transformar uma matriz usando as operações elementares sobre linhas. Este processo é chamado de escalonamento de matrizes.
1) Seja K_1 a primeira coluna da matriz dada 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 $a_{1k_1} \neq 0$.
2) Para cada $i > 1$, realize a operação

$$L_i \longrightarrow L_i - \frac{a_{ik_1}}{a_{1k_1}} L_1.$$

Repita os passos 1 e 2 na matriz assim obtida, ignorando a primeira linha. Novamente, repita os Passos 1 e 2 nessa nova matriz, ignorando as duas primeiras linhas etc., até alcançar a última linha não nula.

[1] Um elemento líder de uma linha se refere ao primeiro elemento não-nulo considerado da esquerda para a direita (numa linha não-nula).

3) Se $L_1, \cdots, L_p$ são as linhas não nulas da matriz obtida após terminar o processo acima e se K_i é a coluna na qual aparece o primeiro elemento não nulo a_{ik_i} da linha L_i, aplique as operações

$$L_i \longrightarrow \frac{1}{a_{ik_i}} L_i, \quad \text{para todo } 1 \leqslant i \leqslant p.$$

4) Realize na matriz obtida até então as operações

$$L_l \longrightarrow L_l - a_{lk_i} L_i, \quad l = 1, \cdots, i-1,$$

para $i = 2$. Depois para $i = 3$, e assim por diante, até $i = p$. Dessa forma, obtemos uma matriz na forma escalonada reduzida. Contudo se o objetivo for só escalonar a matriz devemos parar o algoritmo no passo 2.

Exemplo 1.5.2 *Para escalonar a matriz* $\begin{bmatrix} 2 & 1 & 1 & 10 \\ 4 & 6 & 1 & 0 \\ 6 & 2 & 0 & 1 \end{bmatrix}$*, devemos anular os elementos situados em* a_{21}, a_{31} *e* a_{32}*, logo devemos fazer as operações elementares* $L_2 \to L_2 - \frac{4}{2} \cdot L_1$, $L_3 \to L_3 - \frac{6}{2} \cdot L_1$ *e* $L_3 \to L_3 - \left(\frac{-4}{2}\right) \cdot L_2$*, e assim obtemos as equivalências:*

$$\begin{bmatrix} 2 & 2 & 1 & 10 \\ 4 & 6 & 1 & 18 \\ 6 & 2 & 0 & 14 \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & 1 & 10 \\ 0 & 2 & -1 & -2 \\ 6 & 2 & 0 & 14 \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & 1 & 10 \\ 0 & 2 & -1 & -2 \\ 0 & -4 & -3 & -16 \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & 1 & 10 \\ 0 & 2 & -1 & -2 \\ 0 & 0 & -5 & -20 \end{bmatrix}.$$

Para anular os elementos a_{12} *e transformar os elementos* a_{11}, a_{22} *e* a_{33} *em* 1*, basta seguir com as operações elementares,* $L_1 \to L_1 - \left(\frac{-2}{2}\right) \cdot L_2$, $\quad L_1 \to \frac{1}{2} \cdot L_1$, $\quad e \quad L_2 \to \frac{1}{2} \cdot L_2$*. Especificamente, temos:*

$$\begin{bmatrix} 2 & 2 & 1 & 10 \\ 0 & 2 & -1 & -2 \\ 0 & 0 & -5 & -20 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 2 & 12 \\ 0 & 2 & -1 & -2 \\ 0 & 0 & -5 & -20 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 6 \\ 0 & 2 & -1 & -2 \\ 0 & 0 & -5 & -20 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 6 \\ 0 & 1 & -\frac{1}{2} & -1 \\ 0 & 0 & -5 & -20 \end{bmatrix}.$$

com as operações, $L_3 \to -\frac{1}{5} \cdot L_3$, $\quad L_1 \to L_1 - 1 \cdot L_3 \quad e \quad L_2 \to L_2 + \frac{1}{2} \cdot L_3$*. Segue que,*

$$\begin{bmatrix} 1 & 0 & 1 & 6 \\ 0 & 1 & -\frac{1}{2} & -1 \\ 0 & 0 & -5 & -20 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 6 \\ 0 & 1 & -\frac{1}{2} & -1 \\ 0 & 0 & 1 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & -\frac{1}{2} & -1 \\ 0 & 0 & 1 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}.$$

encontrando a matriz escalonada reduzida.

1.5.3 Matrizes elementares

O conceito de matriz elementar, que introduziremos aqui, será utilizado para demonstrar vários resultados em seções posteriores.

Uma matriz elementar $n \times n$, que denotaremos por E, é uma matriz que resulta da aplicação de uma operação elementar na matriz identidade I. Vejamos alguns exemplos no caso 4×4:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ k & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

As matrizes acima se obtém a partir de I_4 mediante as operações, $L_1 \leftrightarrow L_4$, $L_2 \to k \cdot L_2$ e $L_3 \to L_3 + k \cdot L_1$, respectivamente.

Exemplo 1.5.3 *Ao fazermos a operação elementar* $3 \cdot L_1$ *em* $\begin{bmatrix} 4 & 10 \\ 2 & 5 \end{bmatrix}$, *obtemos* $\begin{bmatrix} 12 & 30 \\ 2 & 5 \end{bmatrix}$ *que é exatamente o produto da matriz elementar* $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$ *por* $\begin{bmatrix} 4 & 10 \\ 2 & 5 \end{bmatrix}$.

Teorema 1.5.1 *Uma matriz elementar E é inversível e sua inversa é a matriz elementar E', que corresponde à operação com linhas inversa da operação efetuada por E.*

Demonstração: Como as operações elementares são inversíveis, como foi mostrado na Proposição (1.5.1), as matrizes elementares são inversíveis, pois se E for obtida através de uma operação elementar em I, então existe outra operação elementar de mesmo tipo que transforma E de volta para I. ■

Exemplo 1.5.4 *A inversa das matrizes elementares* $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ *e* $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}$ *são respectivamente:*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad e \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix}.$$

1.6 Matrizes de um sistema linear

Segundo Anton e Busby (2006) o primeiro exemplo conhecido do uso de matriz aumentada para descrever sistemas lineares aparece no livro chinês "Nove Capítulos de

Arte Matemática"publicado entre 200 a.C. e 100 a.C. durante a dinastia de Han. O problema proposto pelo manuscrito é:
Existem três tipos de milho, dos quais três montes do primeiro, dois do segundo e um do terceiro totalizam 39 medidas. Dois montes do primeiro, três do segundo e um do terceiro totalizam 34 medidas. Finalmente, um monte do primeiro, dois do segundo e três do terceiro totalizam 26 medidas. Quantas medidas de milho estão contidas em um monte de cada um dos tipos?

O Problema nos leva a um sistema linear de três equações e três incógnitas, que o autor escreve como:

$$\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 1 & 1 \\ 26 & 34 & 39 \end{array} .$$

1.6.1 Equação Linear

Definição 1.6.1 *Equação lineares nas variáveis $x_1, x_2, \ldots, x_n$ é toda equação do tipo $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ em que $a_1, a_2, \ldots, a_n$ e b são constantes reais e $a_1, a_2, \ldots, a_n$ não são ambas nulas, b é chamado de coeficiente (ou termo) independente da equação.*

Exemplo 1.6.1 *A equação $2x_1 + 4x_2 - 5x_3 = 8$ apresenta 3 variáveis, x_1, x_2 e x_3 com coeficiente $2, 4, -5$ e o 8 é o termo independente.*

1.6.1.1 Solução de uma equação linear

Dizemos que uma sequência $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, com $\alpha_i \in \mathbb{R}$ para $1 \leqslant i \leqslant n$, é a solução da equação $a_1x_1+a_2x_2+\ldots+a_nx_n = b$ quando a sentença $\alpha_1x_1+\alpha_2x_2+\ldots+\alpha_nx_n = b$ for verdadeira, com $\alpha_i \in \mathbb{R}, i = 1, 2, 3, \cdots, n$.

Exemplo 1.6.2 *A sequência $(2, -2)$, denominado par ordenado, é uma solução da equação $3x - 2y = 10$, pois, $3 \cdot 2 - 2 \cdot (-2) = 10$.*

1.6.2 Sistema linear

Definição 1.6.2 *Um conjunto finito de equações lineares nas variáveis $x_1, x_2, \ldots, x_n$ é chamado de sistema de equações lineares ou simplesmente de sistema linear.*

Um sistema arbitrário de m equações lineares em n variáveis (incógnitas) pode ser escrito como

$$\left\{\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & \cdots & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & \cdots & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & \cdots & a_{mn}x_n & = & b_m \end{array}\right. \quad a_{ij}, b_i \in \mathbb{R},\ com\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n \quad (1.6)$$

as variáveis $x_1, x_2, \ldots, x_n$ denotam constantes. Uma sequência de números $(\alpha_1,\ \alpha_2, \ldots,\ \alpha_n)$ é chamado de solução do sistema se $\alpha_1 = x_1,\ \alpha_2 = x_2,\ \ldots,\ \alpha_n = x_n$ é uma solução de cada equação do sistema.

Um sistema de equações lineares pode ter nenhuma solução, exatamente uma, ou então uma infinidade de soluções; os que não têm solução são chamados inconsistentes; os que tem dizemos consistentes.

Exemplo 1.6.3 *O sistema linear* $\begin{cases} 2x + 3y - 2z = 6 \\ -3x + 2y - z = 0 \end{cases}$ *tem como solução* $x = 1,\ y = 2$ *e* $z = 1$*, pois esses valores satisfazem ambas as equações.*

No entanto, $x = 2,\ y = 2$ *e* $z = 5$ *não é solução do sistema pois estes valores satisfazem apenas a equação* $2x + 3y - 2z = 6$*.*

1.6.2.1 Representação Matricial

Todo sistema de equações pode ser representado na forma matricial, os coeficientes a_{ij} formarão a matriz dos coeficientes, as incógnitas x_i serão os elementos da matriz das incógnitas x, e os termos independentes b_i formará a matriz dos termos independentes. Especificamente, o sistema (1.6) pode ser representado da forma

$$A \cdot X = B \quad \text{ou} \quad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

sendo $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ a matriz dos coeficientes, $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ e $\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ a matriz das incógnitas e a matriz dos termos independentes, respectivamente.

Exemplo 1.6.4 *Represente o sistema* $\begin{cases} 2x_1 + x_2 - x_3 = 4 \\ 3x_1 - 2x_2 - 2x_3 = 2 \\ x_1 + x_2 - x_3 = 2 \end{cases}$ *, na forma matricial.*

O sistema pode ser reescrito como, $\begin{bmatrix} 2 & 1 & -1 \\ 3 & -2 & -2 \\ 1 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}.$

1.6.3 Resolução com o método de Gauss

O método de Gauss consiste em realizar operações elementares sobre linhas no sistema $AX = B$, transformando-o em um sistema escalonado equivalente. A eliminação

de Gauss pode ser descrito nos seguintes termos:
Multiplicarmos os dois lados do sistema sucessivamente por matrizes elementares

$$E_k \dots E_2 \cdot E_1 \cdot A \cdot X = E_k \dots E_2 \cdot E_1 \cdot B \tag{1.7}$$

até obter a matriz escalonada, ou seja, triangular superior $U = E_k \dots E_2 \cdot E_1 \cdot A$, chamando $E_k \dots E_2 \cdot E_1$ de E. Assim o sistema $AX = B$ se transforma no sistema equivalente escalonado:

$$UX = EB. \tag{1.8}$$

Exemplo 1.6.5 *Escalonar o sistema* $\begin{cases} 1x_1 + 1x_2 - 2x_3 = 0 \\ 2x_1 + 3x_2 - 3x_3 = 3 \\ 3x_1 - 1x_2 + 2x_3 = 12 \end{cases}$ *utilizando o método de Gauss.*

O sistema reescrito em notação matricial ficará:

$$\begin{bmatrix} 1 & 1 & -2 \\ 2 & 3 & -3 \\ 3 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}. \tag{1.9}$$

Inicialmente, anularemos o segundo elemento da primeira coluna , ou seja:
1) Multiplicaremos (1.9) por $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$*, ou seja,*

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & -2 \\ 2 & 3 & -3 \\ 3 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}$$

logo,

$$\begin{bmatrix} 1 & 1 & -2 \\ -2+3 & -2+2 & 4-3 \\ 3 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}.$$

Encontrando, assim, o sistema equivalente:

$$\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 3 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}. \tag{1.10}$$

Prosseguimos com o escalonamento, agora vamos anular o terceiro elemento da primeira coluna de (1.10). Especificamente,

2) Vamos multiplicar (1.10) por $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$. *Portanto,*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 3 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}.$$

Obtendo,

$$\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ -3+3 & -3-1 & 6+2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}.$$

Encontrando, assim, o sistema equivalente

$$\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & -4 & 8 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix} \tag{1.11}$$

Nesta terceira etapa, queremos anular o terceiro elemento da segunda coluna de (1.11). Então,

3) Multiplicaremos (1.11) por $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}$, *ou seja:*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & -4 & 8 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}.$$

Portanto,

$$\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & 4-4 & 4+8 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -8-3 & 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}$$

encontrando, assim, o sistema (1.11) escalonado,

$$\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 12 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -11 & 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}$$

isto é,

$$\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 12 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 24 \end{bmatrix} \tag{1.12}$$

como queríamos.

Para encontrar a solução do sistema, faz se necessário o algoritmo da substituição reversa, que será explicitado a seguir.

1.6.3.1 Algoritmo da substituição reversa

Este algoritmo resolve o sistema $RX = B$ pelo método da substituição reversa, onde R é quadrada, inversível e triangular superior. Isto significa que

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_1 m \\ & r_{22} & \cdots & r_2 m \\ 0 & & \ddots & \vdots \\ & & & r_m m \end{bmatrix}$$

com $r_{ii} \neq 0$, para $i = 1; \cdots, m$. Para resolver o sistema $Rx = b$, iniciamos explicitando x_m na última equação e, retornando até a primeira, explicitando as variáveis principais de cada equação em função das variáveis determinadas nas etapas anteriores. Assim,

$$x_m = \frac{b_m}{r_{mm}}, \quad x_{m-1} = \frac{b_{m-1} - r_{m-1,m}x_m}{r_{m-1,m-1}}, \quad x_{m-2} = \frac{b_{m-2} - r_{m-2,m-1}x_{m-1} - r_{m-2,m}x_m}{r_{m-2,m-2}}, \cdots$$

e assim por diante. Agora estamos prontos para encontrar as soluções dos sistemas escalonados. As contas são completamente análogas para matriz dos coeficiente triangular inferior.

Exemplo 1.6.6 *Encontre os valores x_1, x_2 e x_3 do Exemplo 1.6.5.*
Descobriremos por substituição reversa:
Por (1.12) temos que $12x_3 = 24$, portanto $x_3 = 2$;
Substituindo x_3 na equação $x_2 + x_3 = 3$, segue que $x_2 = 1$;
Por fim, substituindo x_3 e x_2 em $x_1 + x_2 - 2x_3 = 1$, obtemos $x_1 = 3$.

Capítulo 2

Algumas matrizes no ensino superior

Neste capítulo faremos uma abordagem sobre transformações lineares definidas em espaços vetoriais com dimensões finitas, mostraremos que essas transformações podem ser representadas por matrizes e apresentaremos algumas decomposições matriciais bem como acharemos soluções de sistemas lineares com estas fatorações. Para isso, usaremos como referências Hefez (2011); Anton et al. (2001); Boldrini et al. (1980); Lima (2009).

2.1 Matriz de uma transformação linear

Segundo Anton et al. (2001) em meados do século dezessete foi materializada a ideia de utilizar pares de números para situar pontos no plano e ternos de números para situar pontos no espaço tridimensional. Na segunda metade do século dezoito, os matemáticos e físicos começaram a perceber que não havia necessidade de parar com ternos, pois quádruplos (a_1, a_2, a_3, a_4) de números poderiam ser considerados pontos de um espaço de dimensão quatro, quíntuplo $(a_1, a_2, a_3, a_4, a_5)$ de números como pontos num espaço de dimensão cinco e assim por diante, uma n-upla de números sendo pontos de um "espaço n-dimensional."

2.1.1 Espaços vetoriais

Definição 2.1.1 *: Um espaço vetorial V é um conjunto, cujos elementos são chamados vetores, onde esta definida duas operações: a adição, que a cada par de vetores u e $v \in V$ faz corresponder um novo vetor $u + v \in V$, chamado soma de u e v; e a multiplicação por um número real, que a cada número $a_1 \in \mathbb{R}$ e a cada vetor $v \in V$ faz corresponder um vetor $a_1 \cdot v$, chamado o produto de a_1 por v. Essas operações devem satisfazer, para quaisquer $a_1, a_2 \in \mathbb{R}$ e $u, v, w \in V$, as seguintes condições.*
Comutatividade: $u + v = v + u$;
Associatividade: $(u + v) + w = u + (v + w)$ e $(a_1 a_2)v = a_1(a_2 v)$;
Vertor nulo: *existe um vetor* $0 \in V$*, chamado vetor nulo, ou vetor zero, tal que* $v + 0 =$

$0 + v = v$ *para todo* $v \in V$*;*
Inverso aditivo: *para cada vetor* $v \in V$ *existe um vetor* $-v \in V$*, chamado de inverso aditivo, ou simétrico de* v*, tal que* $-v + v = v + (-v) = 0$*;*
Distributividade: $(a_1 + a_2)v = a_1v + a_2v$ *e* $a_1(u + v) = a_1u + a_1v$*;*
Multiplicação por 1: $1 \cdot v = v$*.*

Exemplo 2.1.1 *Para todo número natural* n*, o símbolo* $\mathbb{R}^n$ *representa o espaço vetorial euclidiano* n*-dimensional. Os elementos de* $\mathbb{R}^n$ *são as listas ordenadas* $u = (a_1, \cdots, a_n)$*,* $v = (a_1, \cdots, a_n)$ *de números reais.*

Observação 2.1.1 *Essas regras serão usadas implicitamente, quando trabalharmos com vetores.*

2.1.2 Combinação linear

Definição 2.1.2 *Sejam* V *um espaço vetorial sobre* $\mathbb{R}$*, e* $v_1, v_2, \ldots, v_n \in V$ *e* $a_1, a_2, \ldots, a_n$ *números pertencentes a* $\mathbb{R}$*. Então o vetor* $v = a_1v_1 + a_2v_2 + \ldots + a_nv_n$ *é um elemento de* V *ao que chamamos combinação linear de* $v_1, v_2, \ldots, v_n$*.*

Exemplo 2.1.2 *O vetor* $(9, 5)$ *pode ser escrito como combinação linear do conjunto* $\{(1, 0),\ (0, 1)\}$*, ou seja,* $(9, 5) = 9 \cdot (1, 0) + 5 \cdot (0, 1)$*.*

Definição 2.1.3 *Um conjunto* $\{v_1, v_2, \ldots, v_n\}$ *de vetores é definido linearmente dependente (LD), se e somente se, pelo menos um deles pode ser escrito como combinação linear dos outros vetores de seu conjunto. Esse conjunto é dito como linearmente independente (LI) se nenhum deles pode ser escrito como combinação linear dos outros vetores do conjunto.*

Exemplo 2.1.3 *Seja* $v_1 = (1, 1, 1),\ v_2 = (0, 2, 4),\ v_3 = (-1, -2, 3)$ *e* $v_4 = (0, 1, 8)$*; temos que o conjunto* $\{v_1, v_2, v_3, v_4\}$ *é LD pois,* $v_4 = v_1 + v_2 + v_3 = (1, 1, 1) + (0, 2, 4) + (-1, -2, 3) = (0, 1, 8)$*, mas o conjunto* $\{v_1, v_2, v_3\}$ *é LI? Vamos verificar.*
Suponhamos que exista $a, b \in \mathbb{R}$ *tal que* $av_1 + bv_2 = v_3$ *ou seja,*

$$a \cdot (1, 1, 1) + b \cdot (0, 2, 4) = (-1, -2, 3)$$

Temos que

$$(a, a + 2b, a + 4b) = (-1, -2, 3)$$

logo $a = -1$*, substituindo na equação* $a+2b = -2$ *encontramos* $b = -\dfrac{1}{2}$*, o que é impróprio para* $a + 4b = 3$ *pois* $a + 4b = -1 + 4 \cdot \left(-\dfrac{1}{2}\right) = -3 \neq 3$*. Portanto* $\{v_1, v_2, v_3\}$ *é LI.*

2.1.3 Base de um espaço vetorial

Estamos interessados em encontrar, em um espaço vetorial V, um subconjunto β tal que qualquer vetor de V seja uma combinação linear de elementos de β. Em outras palavras, queremos determinar um conjunto de vetores que gere V, tal que todos os elementos sejam necessários para gerar V, se pudermos encontrar tais vetores, teremos os alicerces de nosso espaço, este conjunto de vetores é denominado de base.

Definição 2.1.4 *Seja V um espaço vetorial finitamente gerado. Uma base de V é um subconjunto finito $\beta \subset V$ satisfazendo:*
i) β gera V, ou seja, o espaço gerado por β é igual a V.
ii) β é LI.

Observação 2.1.2 *Um espaço vetorial pode ter várias bases.*

Exemplo 2.1.4 *O conjunto $\{(1,0),(0,1)\}$ é chamado de base canônica do $\mathbb{R}^2$, pois qualquer elemento (x,y) pertencente a $\mathbb{R}^2$ é combinação desta base. Além disso, $\{(1,1),(0,1)\}$ também é uma outra base do $\mathbb{R}^2$.*

Teorema 2.1.1 *Qualquer base de um espaço vetorial tem sempre o mesmo número de elementos. Este número é chamado dimensão de V, e denotado dim V.*

Demonstração: Sejam $\{v_1, v_2, \cdots, v_n\}$ e $\{w_1, w_2, \cdots, w_m\}$ duas bases de V. Como $v_1, v_2, \cdots, v_n$ geram V e $w_1, w_2, \cdots, w_m$ são LI, isto implica que $m \leq n$.

Por outro lado, como $\{w_1, w_2, \cdots, w_m\}$ geram V e $v_1, v_2, \cdots, v_n$ são LI, isto implica que $n \leq m$. Portanto, $n = m$. ■

2.1.3.1 Matriz mudança de base

Em geral, aparecem matrizes complicadas (relacionadas a bases) na resolução de problemas e devemos simplificar as mesmas. Se usarmos bases onde as matrizes são diagonalizáveis simplificaremos muito tais problemas.

Consideremos duas bases ordenadas α e $\beta \in \mathbb{R}^n$ sendo $\alpha = \{v_1, v_2, \ldots, v_n\}$ e $\beta = \{w_1, w_2, \ldots, w_n\}$ e dado um vetor $v \in \mathbb{R}^n$, podemos escrevê-lo como

$$v = x_1v_1 + x_2v_2 + \ldots + x_nv_n \quad \text{e} \quad v = y_1w_1 + y_2w_2 + \ldots + y_nw_n$$

as coordenadas do vetor v em relação a base α é $[v]_\alpha = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ e as coordenadas do

vetor v em relação a base β é $[v]_\beta = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ já que $\{v_1, v_2, \ldots, v_n\}$ é base de V, podemos escrever os vetores $w_1, w_2, \ldots, w_n$ como combinação linear dos vetores $v_1, v_2, \ldots, v_n$, ou seja

$$\begin{cases} w_1 = a_{11}v_1 + a_{21}v_2 \ldots a_{n1}v_n \\ w_2 = a_{12}v_1 + a_{22}v_2 \ldots a_{n2}v_n \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad\quad \vdots \\ w_n = a_{1n}v_1 + a_{2n}v_2 \ldots a_{nn}v_n \end{cases}$$

substituindo temos:

$v = y_1w_1 + y_2w_2 + \cdots + y_nw_n = y_1(a_{11}v_1 + a_{21}v_2 \ldots a_{n1}v_n) + y_2(a_{12}v_1 + a_{22}v_2 \ldots a_{n2}v_n) + \ldots + y_n(a_{1n}v_1 + a_{2n}v_2 \ldots a_{nn}v_n) = (a_{11}y_1 + a_{12}y_2 + \ldots + a_{1n}y_n)v_1 + (a_{21}y_1 + a_{22}y_2 + \ldots + a_{2n}y_n)v_2 + \ldots + (a_{n1}y_1 + a_{n2}y_2 + \ldots + a_{nn}y_n)v_n.$

Como

$v = x_1v_1 + x_2v_2 + \ldots + x_nv_n,$

então:

$x_1v_1 + x_2v_2 + \ldots + x_nv_n = (a_{11}y_1 + a_{12}y_2 + \ldots + a_{1n}y_n)v_1 + (a_{21}y_1 + a_{22}y_2 + \ldots + a_{2n}y_n)v_2 + \ldots + (a_{n1}y_1 + a_{n2}y_2 + \ldots + a_{nn}y_n)v_n$

e como as coordenadas em relação a uma base são únicas, temos

$$\begin{cases} x_1 = a_{11}y_1 + a_{12}y_2 + \ldots + a_{1n}y_n \\ x_2 = a_{21}y_1 + a_{22}y_2 + \ldots + a_{2n}y_n \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad\quad\quad\quad \vdots \\ x_n = a_{n1}y_1 + a_{n2}y_2 + \ldots + a_{nn}y_n \end{cases}.$$

Em notação matricial

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

isto é, $\quad [v]_\alpha = [M]_\alpha^\beta [v]_\beta \quad$ onde

$$[M]_\alpha^\beta = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

é chamada matriz de mudança da base β para a base α. Observe que uma vez obtida $[M]_{\alpha}^{\beta}$ podemos encontrar as coordenadas de qualquer vetor v em relação a base α, multiplicando a matriz de mudança da base pelas coordenadas de v na base β (que conhecemos).

Exemplo 2.1.5 *Sejam* $\beta = \{v_1 = (1,1),\ v_2 = (1,0)\}$ *e* $\theta = \{u_1 = (1,2),\ u_2 = (-4-3)\}$ *duas bases do espaço vetorial real* $\mathbb{R}^2$*. Determine a matriz* P *de mudança de base* β *para a base* α*.*

Para determinar $P = [M]_{\alpha}^{\beta}$ *escrevem-se os vetores da base* α *como combinação linear dos vetores da base* β*, isto é:*

$$a(1,1) + b(1,0) = (1,2) \text{ ou seja } (a+b, a) = (1,2).$$

Assim temos que $a = 2$ *e* $b = -1$*. Portanto o vetor* u_1 *da base* α *se escreve:*

$$(1,2) = 2(1,1) - (1,0),$$

isto é, suas coordenadas em relação à base β *são,*

$$[u_1]_{\beta} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Considerando-se, agora, o vetor u_2 *tem-se:*

$$c(1,1) + d(1,0) = (-4,-3) \text{ ou seja } (c+d, c) = (-4,-3)$$

De onde se obtém $c = -3$ *e* $d = -1$*. Logo, o vetor* u_2 *da base* α *se escreve:*

$$(-4,-3) = -3(1,1) - (1,0),$$

ou seja, suas coordenadas em relação à base β *são:*

$$[u_2]_{\beta} = \begin{pmatrix} -3 \\ -1 \end{pmatrix}.$$

O sistema linear S*, é então:*

$$S : \begin{cases} u_1 = 2v_1 - v_2 \\ u_2 = -3v_1 - v_2 \end{cases}$$

e, portanto, a matriz dos coeficientes é:

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

A matriz P é a transposta dessa matriz; suas colunas são formadas pelas coordenadas de u_1 e u_2 em relação à base β:

$$P = [M]_{\alpha}^{\beta} = \begin{pmatrix} 2 & -3 \\ -1 & -1 \end{pmatrix}.$$

Para definirmos os conceitos de ortogonalidade entre vetores faz se necessário definirmos o produto interno.

2.1.4 Produto interno

Nesta seção apresentaremos os conceitos e propriedades elementares de um espaço vetorial com produto interno Euclidiano.

Definição 2.1.5 *Seja V um espaço vetorial. Um produto interno em V é uma função que a cada par de vetores u e v em V associa um número real, denotado por $\langle u, v\rangle$, que satisfaz as seguintes condições:*

Para quaisquer vetores u, v e w de V e qualquer número real λ,

i) $\langle u, u\rangle \geq 0$;

ii) $\langle u, u\rangle = 0$ *se, e somente se,* $u = 0$;

iii) $\langle u, v\rangle = \langle v, u\rangle$;

iv) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$;

v) $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$.

Exemplo 2.1.6 *Dados $u = (x_1, x_2, \cdots, x_n)$ e $v = (y_1, y_2, \cdots y_n)$, vetores em $\mathbb{R}^n$.*

Definimos

$$\langle u, v\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n. \tag{2.1}$$

Note que

$$\langle u, u\rangle = x_1^2 + x_2^2 + \cdots + x_n^2 \geq 0,$$

e que

$$\langle u, v\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n = y_1x_1 + y_2x_2 + \cdots + y_nx_n = \langle v, u\rangle,$$

mostrando que as condições *i)* e *iii)* da definição de produto interno são satisfeitas. A condição *ii)* também é satisfeita visto que

$$\langle u, u\rangle = x_1^2 + x_2^2 + \cdots + x_n^2 = 0 \Longleftrightarrow x_1 = \cdots = x_n = 0 \Longleftrightarrow u = 0.$$

Se $w = (z_1, z_2, \cdots, z_n)$, então

$$\langle u + v, w\rangle = (x_1 + y_1)z_1 + (x_2 + y_2)z_2 + \cdots + (x_n + y_n)z_n$$

$$= (x_1z_1 + x_2z_2 + \cdots + x_nz_n) + (y_1z_1 + y_2z_2 + \cdots + y_nz_n) = \langle u, w\rangle + \langle v, w\rangle$$

mostrando que condição iv) é satisfeita. A condição v também é satisfeita, pois se $\lambda \in \mathbb{R}$, então

$$\langle \lambda u, v\rangle = (\lambda x_1)y_1 + (\lambda x_2)y_2 + \cdots + (\lambda x_n)y_n = \lambda(x_1y_1 + x_2y_2 + \cdots + x_ny_n) = \lambda\langle u, v\rangle.$$

Definição 2.1.6 *Seja V um espaço com produto interno. Definimos a norma do vetor v de V, ou comprimento de v, denotado por $||v||$, como o número real*

$$||v|| = \langle v, v\rangle^{\frac{1}{2}}.$$

Se $||v|| = 1$, dizemos que v é um vetor unitário. A distância $d(u, v)$ entre dois vetores u e v de V é definida como

$$d(u, v) = ||u - v|| = \sqrt{\langle u - v, u - v\rangle}.$$

Exemplo 2.1.7 *Se $u = (x_1, x_2, \cdots, x_n)$ e $v = (y_1, y_2, \cdots y_n)$, vetores em $\mathbb{R}^n$ com produto interno usual, então*

$$||u|| = \langle u, u\rangle^{\frac{1}{2}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

e

$$d(u, v) = ||u - v|| = \langle u - v, u - v\rangle^{\frac{1}{2}} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

2.1.5 Base ortogonal

Seja V um espaço vetorial. Um conjunto $\beta = \{v_1, v_2, \ldots, v_n\}$ é uma base ortogonal de V se somente se $\langle v_i, v_j\rangle = 0$ com $i \neq j$. Se além disso cada elemento da base β tiver norma 1, ou seja $||v_i|| = 1$, para cada $i \in \{1, 2, \cdots, n\}$, então a base será chamada de ortonormal.

2.1.5.1 Matriz ortogonal

Uma matriz quadrada A é ortogonal se, e somente se, $A \cdot A^t = I = A^t \cdot A$ ou seja $A^{-1} = A^t$. Decorre desta definição que toda matriz ortogonal é formada por vetores colunas (ou linha) ortonormais.

Exemplo 2.1.8 *A matriz* $A = \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ \frac{-6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & \frac{-3}{7} \end{bmatrix}$ *é ortogonal, pois:*

$$A^T \cdot A = \begin{bmatrix} \frac{3}{7} & \frac{-6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & \frac{-3}{7} \end{bmatrix} \cdot \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ \frac{-6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & \frac{-3}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Nem todas as matrizes do tipo $M_{(n\times n)}$ são ortogonais. No entanto, podemos produzir uma matriz ortogonal a partir de uma matriz formada por coluna ou linhas LI. Para isto, devemos, em um primeiro momento ortogonalizar as linhas ou colunas por meio do processo que segue.

2.1.6 Processo de ortogonalização de Gram-Schmidt

Seja $\alpha = \{u_1, u_2, \ldots, u_n\}$ uma base não ortogonal do espaço vetorial euclidiano V, é possível, a partir desta base, determinar uma nova base ortogonal $\beta = \{w_1, w_2, \ldots, w_n\}$ para V. Para isto, Considere-se $u_1 = w_1$, temos que encontrar um valor a_{12} de modo que $w_2 = u_2 - a_{12} \cdot w_1$ seja ortogonal a w_1, teremos:

$$\langle u_2 - a_{12} \cdot w_1, w_1 \rangle = 0$$

$$\langle u_2, w_1 \rangle - \langle a_{12} \cdot w_1, w_1 \rangle = 0$$

$$a_{12} \cdot \langle w_1, w_1 \rangle = \langle u_2, w_1 \rangle$$

$$a_{12} = \frac{\langle w_1, u_2 \rangle}{\langle w_1, w_1 \rangle}$$

isto é,

$$w_2 = u_2 - \frac{\langle w_1, u_2 \rangle}{\langle w_1, w_1 \rangle} \cdot w_1$$

Assim, os vetores w_2 e w_1 são ortogonais.

Na sequência, considere-se o vetor $w_3 = u_3 - a_{13}w_1 - a_{23}w_2$ e encontre os valores de a_{23} e a_{13} de maneira que o vetor w_3 seja ortogonal aos vetores w_1 e w_2:

$$\begin{cases} \langle u_3 - a_{13} \cdot w_2 - a_{23} \cdot w_1, w_1 \rangle = \langle u_3, w_1 \rangle - a_{13} \cdot \langle w_2, w_1 \rangle - a_{23} \cdot \langle w_1, w_1 \rangle = 0 \\ \langle u_3 - a_{13} \cdot w_2 - a_{23} \cdot w_1, w_2 \rangle = \langle u_3, w_2 \rangle - a_{13} \cdot \langle w_2, w_2 \rangle - a_{23} \cdot \langle w_1, w_2 \rangle = 0 \end{cases}$$

Tendo em vista que $\langle w_1, w_2 \rangle = 0$, vem:

$$\begin{cases} \langle u_3, w_1 \rangle - a_{23} \cdot \langle w_1, w_1 \rangle = 0 \\ \langle u_3, w_2 \rangle - a_{13} \cdot \langle w_2, w_2 \rangle = 0 \end{cases}$$

logo,

$$a_{23} = \frac{\langle w_1, u_3\rangle}{\langle w_1, w_1\rangle}, \qquad a_{13} = \frac{\langle w_2, u_3\rangle}{\langle w_2, w_2\rangle}$$

isto é

$$w_3 = u_3 - \frac{\langle w_2, u_3\rangle}{\langle w_2, w_2\rangle} \cdot w_2 - \frac{\langle w_1, u_3\rangle}{\langle w_1, w_1\rangle} \cdot w_1$$

assim, os vetores w_1, w_2 *e* w_3 são ortogonais.

Continue o processo até que tenham sido obtido um conjunto ortogonal $\{w_1, w_2, \ldots, w_n\}$ de vetor que é base de V. Observe que os vetores são definidos recursivamente por

$$u_1 = w_1$$

$$w_k = u_k - a_{1k} \cdot w_1 - \ldots - a_{(k-1,k)} \cdot w_{(k-1)}$$

Onde $k = 2, 3, ..., n$, e

$$a_{ik} = \frac{\langle w_i, u_k\rangle}{\langle w_i, w_i\rangle}.$$

A partir da base ortogonal $\{w_1, w_2, \ldots, w_n\}$ pode-se determinar uma base ortonormal $\{q_1, q_2, \ldots, q_n\}$, onde $q_i = \dfrac{w_i}{||w_i||}$.

Esta base ortonormal pode ser obtida ao mesmo tempo em que se obtém a base ortogonal. Comece com

$$u_1 = w_1 \qquad \text{com} \qquad q_1 = \frac{w_1}{||w_1||}$$

continue com o processo de ortogonização, tomando

$$w_2 = u_2 - r_{12}q_1 \quad \text{com} \quad q_2 = \frac{w_2}{||w_2||}$$

$$w_3 = u_3 - r_{13} \cdot q_1 - r_{23} \cdot q_2 \qquad \text{e considerando} \qquad q_3 = \frac{w_3}{||w_3||}$$

e assim por diante, até que, num passo genérico k,

$$w_k = u_k - r_{1k} \cdot q_1 - \ldots - r_{k-1,k}q_{k-1} \qquad \text{sendo que} \qquad q_k = \frac{w_k}{||w_k||},$$

onde $r_{ik} = \dfrac{\langle q_i, u_k\rangle}{\langle q_i, q_i\rangle}$, sendo $\langle q_i, q_i\rangle = 1$ temos $r_{ik} = \langle q_i, u_k\rangle$ para $k = 2, \ldots, n$ e $i = 1, \ldots, k-1$.

Exemplo 2.1.9 *Seja* $\alpha = \{v_1 = (5, 2); v_2 = (1, 0)\}$ *uma base do* $\mathbb{R}^2$. *Vamos obter a partir de* α *uma base* $\beta = \{u_1, u_2\}$ *ortogonal em relação ao produto interno usual.*
Iniciaremos fazendo,

$$u_1 = v_1 = (5, 2),$$

continuando o processo

$$u_2 = v_2 - a_1 u_1$$

Sendo $a_1 = \dfrac{\langle u_1, v_2\rangle}{\langle u_1, u_1\rangle}$, *teremos* $u_2 = v_2 - \dfrac{\langle u_1, v_2\rangle}{\langle u_1, u_1\rangle} \cdot u_1$, *ou seja,*

$$u_2 = (1,0) - \frac{\langle (5,2),(1,0)\rangle}{\langle (5,2),(5,2)\rangle} \cdot (5,2) = (\tfrac{4}{29}, \tfrac{-10}{29}).$$

Portanto a base β *ortogonal será* $\{(5,2), (\frac{4}{29}, \frac{-10}{29})\}$.

Já para encontrar a base ortonormal $\{q_1, q_2\}$ *devemos fazer,* $q_1 = \frac{u_1}{||u_1||}$ *e* $q_2 = \frac{u_2}{||u_2||}$. *Sendo* $||u_1|| = \sqrt{5^2 + 2^2} = \sqrt{29}$ *e* $||u_2|| = \sqrt{\left(\frac{4}{29}\right)^2 + \left(\frac{-10}{29}\right)^2} = \frac{2\sqrt{29}}{29}$, *temos:*

$$q_1 = \tfrac{(5,2)}{\sqrt{29}} = \left(\tfrac{5}{\sqrt{29}}, \tfrac{2}{\sqrt{29}}\right) \quad e \quad q_2 = \frac{\left(\frac{4}{29}, \frac{-10}{29}\right)}{\frac{2\sqrt{29}}{29}} = \left(\tfrac{2}{\sqrt{29}}, -\tfrac{5}{\sqrt{29}}\right)$$

2.1.7 Transformações lineares

Sejam U, V espaços vetoriais. Uma transformação linear $T : U \longrightarrow V$ é uma correspondência que associa a cada vetor $u \in U$ um vetor $T(u) \in V$ de modo que, para quaisquer $u, v \in U$ e $\alpha \in \mathbb{R}$, temos que:

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

o vetor $T(u)$ chama-se a imagem (ou o transformado) de u pela transformação T.

Exemplo 2.1.10 *A função* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, *dada por* $T(x, y, z) = (x - y, y - z)$, *é uma transformação linear.*

De fato, se $v_1 = (x_1, y_1, z_1) \in \mathbb{R}^3$, $v_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$ *e* $a \in \mathbb{R}$, *então*

$$\begin{aligned} T(v_1 + av_2) &= T(x_1 + ax_2, y_1 + ay_2, z_1 + az_2) \\ &= (x_1 + ax_2 - (y_1 + ay_2), y_1 + ay_2 - (z_1 + az_2)) \\ &= ((x_1 - y_1) + a(x_2 - y_2), (y_1 - z_1) + a(y_2 + z_2)) \\ &= (x_1 - y_1, y_1 - z_1) + a(x_2 - y_2, y_2 - z_2) \\ &= T(v_1) + aT(v_2), \end{aligned}$$

e, portanto, T *é uma transformação linear de* $\mathbb{R}^3$ *em* $\mathbb{R}^2$.

2.1.7.1 Matriz de uma transformação linear

Nesta subseção, veremos que se U e V são espaços vetoriais de dimensão finita, com bases fixadas, então uma transformação linear $T : U \to V$ pode ser representada por uma matriz. A vantagem de uma tal representação é que muitos problemas associados

às transformações lineares, entre espaços de dimensão finita, podem ser resolvidos com a teoria das matrizes.

Sejam U e V espaços vetoriais de dimensão n e m, respectivamente, sobre $\mathbb{R}$. Consideremos uma transformação linear $T : U \to V$. Dadas as bases $\alpha = \{u_1, \ldots, u_n\}$ de U e $\beta = \{v_1, \ldots, v_m\}$ de V. Como β é uma base de V, podemos determinar de modo único, números reais a_{ij}, com $1 \leq i \leq n,\ 1 \leq j \leq m$, tais que

$$T(u_i) = a_{1i}v_1 + \cdots + a_{ji}v_j + \ldots + a_{mi}v_m. \tag{2.2}$$

Tomemos agora u em U. Temos que $u = k_1u_1 + \cdots + k_nv_n$, em que $k_i \in \mathbb{R}$ para $1 \leq i \leq n$. Pela linearidade de T e por (2.2), segue que

$$\begin{aligned} T(u) &= k_1T(u_1) + \cdots + k_nT(u_n) \\ &= k_1(a_{11}v_1 + \cdots + a_{m1}v_m) + \cdots + k_n(a_{1n}v_1 + \cdots + a_{mn}v_m) \\ &= (a_{11}k_1 + \cdots + a_{1n}k_n)v_1 + \cdots + (a_{m1}k_1 + \cdots + a_{mn}k_n)v_m. \end{aligned}$$

Logo,

$$[T(u)]_\beta = \begin{bmatrix} a_{11}k_1 & +\cdots+ & a_{1n}k_n \\ & \vdots & \\ a_{m1}k_1 & +\cdots+ & a_{mn}k_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} k_1 \\ \vdots \\ k_n \end{bmatrix} = [T]^\alpha_\beta \cdot [u]_\alpha \tag{2.3}$$

onde definimos

$$[T]^\alpha_\beta = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

A matriz $[T]^\alpha_\beta$, que representa T em relação às bases α e β, é chamada a matriz de T nas bases α e β.

Exemplo 2.1.11 *Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $T(x, y, z) = (2x + y - z, 3x - 2y + 3z)$. Sejam $\beta = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ e $\alpha = \{(1, 3), (1, 4)\}$ bases de $\mathbb{R}^3$ e $\mathbb{R}^2$ respectivamente. Procuremos $[T]^\beta_\alpha$.*
Calculando T nos elementos da base β, temos:

$$\begin{aligned} T(1,1,1) &= (2,4) &&= 4(1,3) - 2(1,4) \\ T(1,1,0) &= (3,1) &&= 11(1,3) - 8(1,4) \\ T(1,0,0) &= (2,3) &&= 5(1,3) - 3(1,4) \end{aligned}$$

Então

$$T]^\beta_\alpha = \begin{bmatrix} 4 & 11 & 5 \\ -2 & -8 & -3 \end{bmatrix}$$

Quando a transformação linear for de um espaço V à ele mesmo, este será chamado de operador linear sobre V.

2.2 Decomposições matriciais

Nesta seção estudaremos as decomposições LU, Cholesky e QR; sendo a decomposição LU um novo olhar sobre o método de eliminação de Gauss, e a fatoração de Cholesky uma especialização da decomposição LU, já a fatoração QR é uma interpretação matricial do processo de Gram-Schmidt.

2.2.1 Decomposição LU

A decomposição LU é um método que fatora a matriz A em uma matriz triangular inferior L (que vem do inglês lower, inferior) e uma matriz triangular superior U (que vem do inglês upper, superior), utilizando-se do método de eliminação de Gauss. Uma condição suficiente (mas não necessária) para que uma matriz A possa ser escalonada sem transposições de linhas, portanto admite uma decomposição A = LU, é que suas submatrizes principais sejam inversíveis.

Seja A uma matriz inversível a ser escalonado e aplicando matrizes elementares $E_n \dots E_3 \cdot E_2 \cdot E_1 = E$ temos

$$E \cdot A = U \tag{2.4}$$

aplicando a inversa E^{-1} em ambos os lados de (2.4) teremos

$$E^{-1} \cdot E \cdot A = E^{-1} \cdot U$$

e chamando E^{-1} de L, segue que:

$$A = L \cdot U \tag{2.5}$$

encontramos a fatoração de A.

Exemplo 2.2.1 *Note que ao continuarmos o Exemplo (1.6.5) da eliminação de Gauss, ou seja, que pede para escalonar o sistema* $\begin{cases} 1x_1 + 1x_2 - 2x_3 = 0 \\ 2x_1 + 3x_2 - 3x_3 = 3 \\ 3x_1 - 1x_2 + 2x_3 = 12 \end{cases}$ *e teve como resultado o sistema*

$$\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 12 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -11 & 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}. \tag{2.6}$$

De fato:

Sendo $E_3 \cdot E_2 \cdot E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -11 & 4 & 1 \end{bmatrix} = E$, *sabendo que* $E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} \cdot E_3 \cdot E_2 \cdot E_1 = I$, *assim devemos calcular o produto* $E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} = E^{-1}$, *ou seja:*

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -4 & 1 \end{bmatrix} = E^{-1}. \tag{2.7}$$

Multiplicando (2.7) em (2.6), obtemos:

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 12 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -11 & 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}.$$

Portanto,

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 12 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 12 \end{bmatrix}.$$

Transformando a matriz dos coeficientes A em um produto de uma triangular inferior L e uma matriz triangular superior U, ou seja:

$$\begin{bmatrix} 1 & 1 & -2 \\ 2 & 3 & -3 \\ 3 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 12 \end{bmatrix}.$$

Um fato notável acontece no produto das inversas das matrizes elementares observe em (2.7), sendo k_i os elementos diferentes de zero das matrizes elementares que situam abaixo da diagonal principal, vale a igualdade:

$$L = E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} \cdots E_i^{-1} = \begin{bmatrix} 1 & & & \\ k_1 & 1 & & \\ k_2 & k_3 & \ddots & \\ \vdots & \vdots & K_i & 1 \end{bmatrix}.$$

2.2.1.1 Solução do sistema $AX = B$ por fatoração LU

A vantagem do método LU mora no fato de trabalhamos somente com a matriz dos coeficientes do sistema linear. Podendo assim resolver vários sistemas lineares que tenha a mesma matriz dos coeficientes e termos independentes diferentes, com uma única fatoração.

Encontrada a fatoração da matriz dos coeficiente A, o sistema $AX = B$ pode ser reescrito como $LUX = B$ então defina $UX = Y$, resolva $LY = B$ e, em seguida, $UX = B$ para obter a solução do sistema original.

Exemplo 2.2.2 *Resolver o sistema* $\begin{cases} 3x_1 + x_2 + x_3 = 7 \\ x_1 + 2x_2 + 3x_3 = 11 \\ x_1 + x_2 + 2x_3 = 7 \end{cases}$ *com o método LU.*

Neste exemplo para efeito didático utilizaremos as operações elementares para escalonar a matriz dos coeficientes.

Iniciaremos o processo reescrevendo o sistema na forma matricial

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 11 \\ 7 \end{bmatrix}$$

Para encontrar as matrizes equivalentes foram utilizadas as seguintes operações elementares respectivamente; $L_2 \to L_2 - \frac{1}{3} \cdot L_1$, $L_3 \to L_3 - \frac{1}{3} \cdot L_1$ *e* $L_3 \to L_3 - \frac{2}{5} \cdot L_1$,

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 1 \\ 0 & \frac{5}{3} & \frac{8}{3} \\ 1 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 1 \\ 0 & \frac{5}{3} & \frac{8}{3} \\ 0 & \frac{2}{3} & \frac{5}{3} \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 1 \\ 0 & \frac{5}{3} & \frac{8}{3} \\ 0 & 0 & \frac{3}{5} \end{bmatrix}$$

Assim, encontramos a matriz triangular superior $U = \begin{bmatrix} 3 & 1 & 1 \\ 0 & \frac{5}{3} & \frac{8}{3} \\ 0 & 0 & \frac{3}{5} \end{bmatrix}$ *e a matriz triangular inferior* $L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{1}{3} & \frac{2}{5} & 1 \end{bmatrix}$, *que é formada pelos valores inverso das operações elementares e 1's na sua diagonal principal.*

Encontrada a fatoração da matriz dos coeficiente podemos reescrever o $A \cdot X = B$ *como* $L \cdot (U \cdot X) = B$ *e chamando* $(U \cdot X)$ *de* Y, *Temos um novo sistema* $L \cdot Y = B$. *Encontraremos então os valores* Y *por substituição reversa, ou seja:*

$$\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{1}{3} & \frac{2}{5} & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 11 \\ 7 \end{bmatrix}$$

logo temos:

$y_1 = 7$;

$\frac{y_1}{3} + y_2 = 11$ *substituindo* y_1 *na equação, temos que* $y_2 = \frac{26}{3}$;

$\frac{y_1}{3} + \frac{2y_2}{5} + y_3 = 7$ *substituindo* y_1 *e* y_2 *na equação, teremos* $y_3 = \frac{6}{5}$.

Repetindo o processo para encontrar a solução do sistema $U \cdot x = y$, ou seja:

$$\begin{bmatrix} 3 & 1 & 1 \\ 0 & \frac{5}{3} & \frac{8}{3} \\ 0 & 0 & \frac{3}{5} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ \frac{26}{3} \\ \frac{6}{5} \end{bmatrix}$$

$\frac{3x_3}{5} = \frac{6}{5}$ *logo,* $x_3 = 2$*;*
$\frac{5x_2}{3} + \frac{8x_3}{3} = \frac{26}{3}$ *substituindo* x_3*, teremos* $x_2 = 2$*;*
$3x_1 + x_2 + x_3 = 7$ *substituindo os valores de* x_2 *e* x_3 *na equação, teremos* $x_1 = 1$.
Encontrando assim a solução $(1, 2, 2)$ *que procurávamos.*

2.2.2 Fatoração PLU

Nem sempre uma matriz A possui uma decomposição LU e um exemplo clássico é $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. Entretanto, a matriz $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ obtida de A pela permutação das linhas possui uma decomposição LU

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

A priori, não se sabe quais são (nem mesmo se vão ser necessárias) as transposições de linhas durante o processo de escalonamento de uma matriz A. Entretanto, depois de efetuado o processo, dispomos da relação de todas as transposições feitas. Efetuando, na mesma ordem em que foram feitas, todas essas transposições nas linhas da matriz identidade, obtemos uma matriz $P \in M_{(m \times m)}$, que se chama uma matriz de permutação. O produto PA corresponde a efetuar sobre a matriz A, antecipadamente, todas as transposições de linhas que seriam necessárias durante o escalonamento. Portanto a matriz PA pode ser escalonada usando apenas operações elementares do tipo $L_i - kL_j$. Assim, tem-se a decomposição $PA = LU$.

2.2.3 Fatoração de Cholesky

Sabendo que o produto de uma matriz por sua transposta é uma matriz simétrica com elementos positivos na sua diagonal ($l_i = c_i$ isto implica que $l_i \cdot c_i = l_i^2$), o que provoca a seguinte definição:

Se A é uma matriz simétrica positiva definida[1], então existe uma única matriz triangular inferior L com diagonal estritamente positiva, tal que $A = L \cdot L^t$. A dedução das fórmulas são feitas pela observação da multiplicação das duas matrizes genéricas, supondo que já tenha a matriz dos coeficientes que atenda os pré-requisitos estabelecidos

[1](critério de Sylvester) Todas as submatrizes possuem determinante positivo

anteriormente, calculemos:

$$L \cdot L^T = \begin{bmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & \cdots & 0 \\ l_{31} & l_{32} & l_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{nn} \end{bmatrix} \cdot \begin{bmatrix} l_{11} & l_{21} & l_{31} & \cdots & l_{n1} \\ 0 & l_{22} & l_{32} & \cdots & l_{n2} \\ 0 & 0 & l_{33} & \cdots & l_{n3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & l_{nn} \end{bmatrix} =$$

$$\begin{bmatrix} l_{11}^2 & l_{11} \cdot l_{21} & l_{11} \cdot l_{31} & \cdots & l_{11} \cdot l_{n1} \\ l_{21} \cdot l_{11} & l_{21}^2 + l_{22}^2 & l_{21} \cdot l_{31} + l_{22} \cdot l_{32} & \cdots & l_{21} \cdot l_{1n} + l_{22} \cdot l_{2n} \\ l_{31} \cdot l_{11} & l_{31} \cdot l_{21} + l_{32} \cdot l_{22} & l_{31}^2 + l_{32}^2 + l_{33}^2 & \cdots & l_{31} \cdot l_{n1} + \cdots + l_{33} l_{n3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} \cdot l_{11} & l_{n1} \cdot l_{21} + l_{n2} \cdot l_{22} & l_{n1} \cdot l_{31} + l_{n2} \cdot l_{22} + l_{n3} \cdot l_{33} & \cdots & l_{n1}^2 + \ldots + l_{nn}^2 \end{bmatrix} =$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} = A \tag{2.8}$$

Da igualdade acima, segue que:

$$\begin{bmatrix} l_{11}^2 \\ l_{21} \cdot l_{11} \\ l_{31} \cdot l_{11} \\ \vdots \\ l_{n1} \cdot l_{11} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{bmatrix} \tag{2.9}$$

E portanto de (2.9) tem-se que:
$l_{11}^2 = a_{11} \Rightarrow l_{11} = \sqrt{a_{11}}$; $\quad l_{11} \cdot l_{21} = a_{21} \Rightarrow l_{21} = \dfrac{a_{21}}{\sqrt{a_{11}}}$; $\quad l_{11} \cdot l_{31} = a_{31} \Rightarrow l_{31} = \dfrac{a_{31}}{\sqrt{a_{11}}}$, $\cdots$
$l_{i1} = \dfrac{a_{i1}}{\sqrt{a_{11}}}$, $\quad$ para $\quad i = 2, \ldots, n$.

Continuando o processo, temos a seguintes igualdade:

$$\begin{bmatrix} l_{11} \cdot l_{21} \\ l_{21}^2 + l_{22}^2 \\ l_{31} \cdot l_{21} + l_{32} \cdot l_{22} \\ \vdots \\ l_{n1} \cdot l_{21} + l_{n2} \cdot l_{22} \cdot l_{11} \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{n2} \end{bmatrix} \tag{2.10}$$

Logo, de (2.10), segue:

$$l_{11}^2 + l_{22}^2 = a_{22} \quad \Rightarrow \quad l_{22} = \sqrt{a_{22} - l_{21}^2}, \quad l_{31} \cdot l_{21} + l_{32} \cdot l_{22} = a_{32} \;\Rightarrow\; l_{32} = \frac{a_{32} - l_{31} \cdot l_{21}}{l_{22}}, \cdots$$

$$\begin{bmatrix} l_{11} \cdot l_{n1} \\ l_{12} \cdot l_{1n} + l_{22} \cdot l_{2n} \\ l_{31} \cdot l_{n1} + l_{32} \cdots l_{n2} + l_{33} l_{n3} \\ \vdots \\ l_{n1}^2 + l_{n2}^2 \ldots + l_{n3}^2 + l_{nn}^2 \end{bmatrix} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \vdots \\ a_{nn} \end{bmatrix} \tag{2.11}$$

Fazendo estas igualdades deduziremos as seguintes fórmulas gerais:

$$l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2} \qquad ou \qquad l_{ii} = \sqrt{a_{ii} - l_{i1}^2 + l_{i2}^2 + \ldots + l_{ii-1}^2} \tag{2.12}$$

Além disso,

$$l_{ij} = \frac{1}{l_{jj}}(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}) \qquad ou \qquad l_{ij} = \frac{1}{l_{jj}} \left[a_{ij} - (l_{i1} \cdot l_{j1} + l_{i2} \cdot l_{j2} + \ldots + l_{i(j-1)} \cdot l_{j(j-1)})\right] \tag{2.13}$$

Para $i > j, \quad e \quad j > 1$.

Exemplo 2.2.3 *Resolver sistema* $\begin{cases} x_1 + x_2 = 3 \\ x_1 + 5x_2 - 2x_3 = 5 \\ -2x_2 + 2x_3 = 2 \end{cases}$ *pelo método de Cholesky.*

Reescrevendo o sistema na forma matricial, obtemos:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 5 & -2 \\ 0 & -2 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix} \tag{2.14}$$

Observa-se que a matriz dos coeficientes é simétrica e que os determinantes de suas sub matrizes são positivos

$$detA_1 = |1| = 1, \qquad detA_2 = \begin{vmatrix} 1 & 1 \\ 1 & 5 \end{vmatrix} = 4 \quad e \quad detA_3 = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 5 & -2 \\ 0 & -2 & 2 \end{vmatrix} = 4,$$

logo este sistema pode ser resolvido pelo método de Cholesky.

Assim podemos utilizar as fórmulas (2.12) e (2.13) para encontrar os valores l_{ij}*:*

$$l_{11} = \sqrt{a_{11}} = \sqrt{1} = 1, \qquad l_{12} = \frac{a_{12}}{l_{11}} = \frac{1}{1} = 1, \qquad l_{13} = \frac{a_{13}}{l_{11}} = \frac{0}{1} = 0,$$

$$l_{22} = \sqrt{a_{22} - l_{12}^2} = \sqrt{5 - 1^2} = 2, \qquad l_{23} = \frac{1}{l_{22}} \left[a_{23} - (l_{12} \cdot l_{13})\right] = \frac{1}{2} \cdot [-2 - (1 \cdot 0)] = -1,$$

$$l_{33} = \sqrt{a_{33} - (l_{13}^2 + l_{23}^2)} = \sqrt{2 - (0^2 + (-1)^2)} = 1.$$

Portanto, a matriz L *e* L^t *seguem:*

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad e \qquad L^t = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

Resolvendo o sistema $L.Y = B$*, por substituição reversa,*

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix}$$

Temos:
$y_1 = 3$;
$y_1 + 2y_2 = 5$*, substituindo* y_1 *na equação, encontramos* $y_2 = 1$;
$-y_2 + y_3 = 2$*, substituindo* y_2 *na equação, encontramos* $y_3 = 3$.

Resolvendo o sistema $L^t \cdot X = Y$*,*

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$$

Logo,
$x_3 = 3$;
$2x_2 - x_3 = 1$*, substituindo* x_3 *na equação, encontramos* $x_2 = 2$;
$x_1 + x_2 = 3$*, substituindo* x_2 *na equação, encontramos* $x_1 = 1$.
Assim a solução do sistema é (1,2,3).

2.2.4 Decomposição QR

Esta é uma interpretação matricial do processo de Gram-Schmidt. Na qual toda matriz inversível $A = [a_{ij}] \in M_{n\times n}$ admite uma decomposição do tipo $A = QR$, onde Q é ortogonal e R é triangular superior, com elementos positivos na diagonal.

Nos últimos anos a fatoração QR tem assumido importância crescente como fundamento matemático de uma grande variedade de algoritmos numéricos práticos, incluindo algoritmos largamente usados para computar autovalores de matrizes grandes.

Para ver como surge a fatoração QR, considere $u_1, u_2, \ldots, u_n$ vetores- coluna de A e $q_1, q_2, \ldots, q_n$ vetores ortonormais obtidos pela aplicação do Processo de Gram-Schmidt

à matriz A com normalizações; assim

$$A = \left[\begin{array}{ccccccc} u_1 & | & u_2 & | & ... & | & u_n \end{array} \right] \; e \; Q = \left[\begin{array}{ccccccc} q_1 & | & q_2 & | & ... & | & q_n \end{array} \right]$$

Os vetores $u_1, u_2, \ldots, u_n$ podem ser escritos em termos dos vetores $q_1, q_2, \ldots, q_n$ como

$$\begin{array}{ccccccc} u_1 & = & \langle q_1, u_1 \rangle q_1 & + & \langle q_2, u_1 \rangle q_2 & + \cdots + & \langle q_n, u_1 \rangle q_n \\ u_2 & = & \langle q_1, u_2 \rangle q_1 & + & \langle q_2, u_2 \rangle q_2 & + \cdots + & \langle q_n, u_2 \rangle q_n \\ \vdots & & \vdots & & \vdots & & \vdots \\ u_n & = & \langle q_1, u_n \rangle q_1 & + & \langle q_2, u_n \rangle q_2 & + \cdots + & \langle q_n, u_n \rangle q_n \end{array}$$

na forma matricial, tem-se:

$$\left[\begin{array}{c} u_1 | u_2 | \cdots | u_n \end{array} \right] = \left[\begin{array}{c} q_1 | q_2 | \cdots | q_n \end{array} \right] \cdot \left[\begin{array}{cccc} \langle q_1, u_1 \rangle & \langle q_1, u_2 \rangle & \cdots & \langle q_1, u_n \rangle \\ \langle q_2, u_1 \rangle & \langle q_2, u_2 \rangle & \cdots & \langle q_2, u_n \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle q_n, u_1 \rangle & \langle q_n, u_2 \rangle & \cdots & \langle q_n, u_n \rangle \end{array} \right]$$

ou, mais concisamente, por

$$A = QR$$

No entanto, é uma propriedade de processo de Gram-Schmidt que, para $j \geq 2$, o vetor q_j ortogonal a $u_1, u_2, \ldots, u_{j-1}$; assim, todas as entradas abaixo da diagonal principal de R são nulas,

$$R = \left[\begin{array}{cccc} \langle q_1, u_1 \rangle & \langle q_1, u_2 \rangle & \cdots & \langle q_1, u_n \rangle \\ 0 & \langle q_2, u_2 \rangle & \cdots & \langle q_2, u_n \rangle \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \langle q_n, u_n \rangle \end{array} \right]$$

Teorema 2.2.1 *Se A é uma matriz $m \times n$ ($m \geq n$) com colunas linearmente independentes, então A pode ser fatorada como $A = QR$ onde Q é uma matriz $m \times n$ com colunas ortonormais e R é uma matriz $n \times n$ triangular superior inversível.*

Exemplo 2.2.4 *Seja $A = \{u_1, u_2, u_3\}$ uma base qualquer não ortogonal, aplicaremos o processo de ortogonização de Gram-Schmidt e encontraremos uma base Ortogornal $\{w_1, w_2, w_3\}$ em seguida, encontraremos uma base ortonormal $\{q_1, q_2, q_3\}$, onde $r_{ik} = \dfrac{w_i}{||w_i||}$ quando $i = k$ ou ainda, $r_{ik} = \langle q_i, w_k \rangle$ quando $i \neq k$:*

$w_1 = u_1$, *normalizando e reorganizando* $u_1 = r_{11} \cdot q_1$;

$w_2 = u_2 - r_{12} \cdot q_1$, *normalizando e reorganizando* $u_2 = r_{12} \cdot q_1 + r_{22} \cdot q_2$;

$w_3 = u_3 - r_{13} \cdot q_1 - r_{23} \cdot q_2$, *normalizando e reorganizando* $u_3 = r_{13} \cdot q_1 + r_{23} \cdot q_2 + r_{33} \cdot q_3$.

Em forma matricial:

$$A = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} = Q \cdot R.$$

2.2.5 Solução de $AX = B$ usando a decomposição QR

Quando A é uma matriz quadrada de ordem m, cujas colunas são linearmente independentes, o sistema $AX = B$ possui uma única solução. Para resolver este sistema usando a decomposição QR, procedemos do seguinte modo:

1. Calcule a decomposição $A = QR$.
2. Determine $y = Q^t \cdot B$.
3. Resolva o sistema $RX = Y$ por substituição inversa.

Exemplo 2.2.5 *Encontre a solução do sistema linear* $\begin{cases} x + y + z = 2 \\ x + y = 3 \\ x + z = 1 \end{cases}$ *, com a utilização da decomposição QR.*

Colocando o sistema em forma matricial teremos

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}.$$

Observe que os vetores coluna da matriz dos coeficiente, ou seja, $v_1 = (1, 1, 1)$, $v_2 = (1, 1, 0)$ e $v_3 = (1, 0, 1)$ não são ortogonais. Desta feita, usaremos o processo de ortogonalização de Gram-Schmidt para ortonaliza-los e em sequência normaliza-los:
Seja,

$$w_1 = v_1.$$

Para ortonormalizar o vetor w_1 temos que encontrar o $||w_1||$,

$$||w_1|| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3},$$

fazendo $q_1 = \dfrac{w_1}{||w_1||}$, teremos o vetor unitário

$$q_1 = \left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\right)$$

para encontrar w_2 temos que encontrar r_{12}, ou seja:

$$r_{12} = \frac{\langle q_1, v_2 \rangle}{\langle q_1, q_1 \rangle} \quad se \quad \langle q_1, q_1 \rangle = 1 \quad temos\ que \quad r_{12} = \langle q_1, v_2 \rangle$$

fazendo o cálculo temos

$$r_{12} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \cdot (1,0,1) = \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + 0 = \frac{2}{\sqrt{3}},$$

calculando w_2

$$w_2 = v_2 - r_{12} \cdot q_1 \qquad w_2 = (1,1,0) - \frac{2}{\sqrt{3}} \cdot \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right).$$

O módulo de $||w_2||$ *é*

$$||w_2|| = \sqrt{\left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2} = \frac{\sqrt{6}}{3}$$

Normalizando o vetor w_2 *para encontrar* q_2*,*

$$q_2 = \frac{w_2}{||w_2||} \qquad q_2 = \frac{\left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)}{\frac{\sqrt{6}}{3}} = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right).$$

Para encontrar w_3 *teremos que encontrar* r_{13} *e* r_{23}

$$r_{13} = \langle q_1, v_3 \rangle = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \cdot (1,0,1) = \frac{2}{\sqrt{3}},$$

$$r_{23} = \langle q_2, v_3 \rangle = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right) \cdot (1,0,1) = -\frac{1}{\sqrt{6}}$$

encontraremos agora o w_3

$$w_3 = (1,0,1) - \frac{2}{\sqrt{3}} \cdot \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) - \left(-\frac{1}{\sqrt{6}}\right) \cdot \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right) = \left(\frac{1}{2}, -\frac{1}{2}, 0\right)$$

o módulo de w_3 *é*

$$w_3 = \sqrt{\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + 0^2} = \frac{\sqrt{2}}{2}$$

normalizando para encontrar w_3

$$q_3 = \frac{w_3}{||w_3||} \qquad q_3 = \frac{\left(\frac{1}{2}, -\frac{1}{2}, 0\right)}{\frac{\sqrt{2}}{2}} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right).$$

Agora vamos reescrever o sistema com a matriz dos coeficientes em forma QR

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ 0 & \frac{\sqrt{6}}{3} & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

lembrando que a matriz Q *é ortogonal e a inversa de uma matriz ortogonal é ela transposta, e multiplicando* Q^{-1} *no lado esquerdo do sistema temos*

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ 0 & \frac{\sqrt{6}}{3} & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} =$$

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix},$$

fazendo as operações encontraremos o sistema equivalente

$$\begin{bmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ 0 & \frac{\sqrt{6}}{3} & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{6}{\sqrt{3}} \\ \frac{3}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Resolvendo por substituição reversa:

$\frac{\sqrt{2}}{2}z = -\frac{1}{\sqrt{2}}$ *logo,* $z = -1$*;*

$\frac{\sqrt{6}}{3}y - \frac{1}{\sqrt{6}}z = \frac{3}{\sqrt{6}}$ *substituindo* z *na equação, teremos* $y = 1$;

$\sqrt{3}x + \frac{2}{\sqrt{3}}y + \frac{2}{\sqrt{3}}z = \frac{6}{\sqrt{3}}$ *substituindo* z *e* y *na equação, teremos* $x = 2$.

Assim a solução do sistema é (2,1,-1).

Capítulo 3

Algumas aplicações de matrizes

A Criptografia, a sequência de Fibonacci, e a cadeia de Markov estão relacionados com matrizes. Os citados tópicos serão abordados a seguir com a finalidade de propor o ensino de matrizes agregadas à algumas aplicações, o que torna a aprendizado mais significativo. Este capítulo esta alicerçado em Boldrini et al. (1980); Hefez (2011); Anton et al. (2001); Taha et al. (2008).

Na próxima seção são apresentados, conceitos de teoria dos números[1] e matrizes, aplicados em um sistema criptográfico.

3.1 O uso de matrizes na Criptografia

A necessidade de guardar mensagens secretas, consideradas importantes, de forma que somente pessoas certas possam decifrá-las, vem acompanhando a sociedade a milênios. A Criptografia do grego kryptós - significa secreto, oculto, e gráphein - escrever, é um conjunto de técnicas para escrever mensagens em código, de modo que somente o remetente e o destinatário sejam capazes de interpretá-la. Sendo cifrar o ato de transformar um texto normal em texto secreto e decodificar a operação inversa, ou seja consiste em transformar um texto cifrado em texto normal.

Existem diversos métodos para transformar dados normal (texto puro) em texto cifrado. O que será abordado nesta seção será a Cifras de Hill o qual se utiliza de álgebra linear para cifrar e decifrar mensagens.

3.1.1 A evolução da Criptografia

Cerca de 1900 a.C. acontece o primeiro relato da história da Criptografia. Numa vila egípcia perto do rio Nilo chamada Menet Khufu, Khnumhotep II era um arquiteto do faraó Amenemhet II. Ele construiu alguns monumentos para o faraó, os quais precisavam

[1] A teoria dos números é o estudo dos números naturais ou inteiros positivos 1, 2, 3, 4,... e suas propriedades.

ser documentados. Nem é preciso dizer que estas informações, escritas em tabletes de argila, não eram para cair no domínio público. O escriba de Khnumhotep II teve a ideia de substituir algumas palavras ou trechos de texto destes tabletes. Caso o documento fosse roubado, o ladrão não encontraria o caminho que o levaria ao tesouro, morreria de fome perdido nas catacumbas da pirâmide. Esse pode ser considerado o primeiro exemplo documentado da escrita cifrada.

Atualmente, a Criptografia é utilizada em transações eletrônicas, como movimentações bancárias executados na internet e entre outras situações da vida cotidiana, os quais necessitam de uma comunicação confidencial para o tráfego de dados.

Historicamente a Criptografia aconteceu em três fases distintas: a Criptografia manual, a Criptografia por máquinas e a Criptografia em rede.

Para uma melhor compreensão dos processos de cifragem e decifragem utilizando Cifras de Hill, faremos uma breve introdução de aritmética modular.

3.1.2 Aritmética modular

Definição 3.1.1 *Sejam a e b dois números inteiros e m um número inteiro positivo maior que 1. Dizemos a e b são congruentes módulo m se o m divide a diferença $a-b$. Quando os inteiros a e b são congruentes módulo m, escrevemos que*

$$a \equiv b \ mod \ m$$

Pelo algoritmo da divisão, dado um inteiro positivo m, qualquer inteiro a é congruo módulo m a exatamente um dos inteiros $0, 1, 2, \cdots, m-1$. Este inteiro é chamado resíduo de a módulo m.

Assim pode-se dividir o conjunto $\mathbb{Z}$, dos números inteiros em subconjuntos, sendo cada subconjunto formado por todos os inteiros que tem o mesmo resto quando divididos por m. Ou seja

$$\begin{aligned} \bar{0} &= \{x \in \mathbb{Z};\ x \equiv 0 \ \ mod\ m\} \\ \bar{1} &= \{x \in \mathbb{Z};\ x \equiv 1 \ \ mod\ m\} \\ \vdots & \qquad\qquad \vdots \\ \overline{m-1} &= \{x \in \mathbb{Z};\ x \equiv m-1 \ \ mod\ m\} \end{aligned}$$

O conjunto $\bar{a} = \{x \ \in \mathbb{Z}; x \equiv a \ \ mod\ m\}$ é chamado de classe residual módulo m do elemento $a \in \mathbb{Z}$. O conjunto de todas as classes residuais módulo m é representada por $\mathbb{Z}_m$. Portanto, pode ser mostrado (Hefez (2011)) que em $\mathbb{Z}_m$ existe m classes distintas, ou seja,

$$\mathbb{Z}_m = \{\bar{0}, \bar{1}, \cdots, \overline{m-1}\}.$$

Observação 3.1.1 *Se a é um inteiro não negativo, então seu resíduo módulo m é simplesmente o resto da divisão de a por m.*

Teorema 3.1.1 *Dados um número inteiro positivo m e um número inteiro a, seja R o resto da divisão de $|a|$ por m. Então o resíduo r de a é dado por*

$$r = \begin{cases} R, & se \quad a \geqslant 0, \\ m - R, & se \quad a < 0 \quad e \quad R \neq 0 \\ 0, & se \quad a < 0 \quad e \quad R = 0 \end{cases} .$$

A definição de soma e Adição e Multiplicação em $\mathbf{Z}_m$ é respectivamente:

$$\bar{a} + \bar{b} = \overline{a + b} \quad e \quad \bar{a} \cdot \bar{b} = \overline{a \cdot b}.$$

Definição 3.1.2 *Um elemento $\bar{a} \in \mathbb{Z}_m$ é inversível, quando existir $\bar{b} \in \mathbb{Z}_m$ tal que $\bar{a} \cdot \bar{b} = 1$. Neste caso, diremos que $\bar{b}$ é o inverso de $\bar{a}$.*

Exemplo 3.1.1 *A tabela abaixo, chamada de tábua da multiplicação módulo 5, mostra o produto de todos elementos de Z_5. Na referida tabela iremos omitir as barras dos elementos de $\mathbb{Z}_5$, para não sobrecarregá-la.*

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Tabela 3.1: Tábua de multiplicação módulo 5

Observa-se que todos os elementos de $\mathbb{Z}_5^*$ possuem inverso multiplicativo, porém nem sempre isto acontece.

Exemplo 3.1.2 *Um exemplo é a tabela a seguir.*

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Tabela 3.2: Tábua de multiplicação módulo 4

Na tabela 3.2, veja que o elemento $\bar{2}$ não tem inverso.

Os elementos inversíveis são caracterizados pela proposição que segue:

Proposição 3.1.1 *Um número $\bar{a} \in \mathbb{Z}_m$ é inversível se, e somente se, a e m não tem fatores primos comuns, isto é, $mdc(a, m) = 1$.*

Demonstração: Suponha que $\bar{a}$ seja inversível, então existe $\bar{b} \in \mathbb{Z}_m$ tal que $\bar{1} = \bar{a} \cdot \bar{b} = \overline{a \cdot b}$. Logo, $ab \equiv 1 \mod m$. Consequentemente, $mdc(a, m) = 1$.

Reciprocamente, se $mdc(a, m) = 1$, existem naturais b e t tais que $ab - mt = 1$, e assim, $\overline{1 + mt} = \overline{ab}$. Logo,

$$\bar{1} = \bar{1} + \overline{mt} = \overline{1 + mt} = \overline{a \cdot b} = \bar{a} \cdot \bar{b}.$$

Portanto $\bar{a}$ é inversível. ■

Exemplo 3.1.3 *O número 5 tem um inverso módulo 26, pois* $mdc(5, 26) = 1$*, mas* 8 *não tem inverso módulo 26 pois* $mdc(8, 26) = 2$*.*

Já no caso de uma matriz A inversível módulo m, ou seja $AB \equiv BA \equiv I \mod m$ onde I é a matriz identidade; é possível mostrar que a matriz B é única e denotaremos B por A^{-1}.

Teorema 3.1.2 *Uma matriz* 2×2 *com entradas em* $\mathbb{Z}_m$ *é inversível módulo* m *se, e somente se, o resíduo de* $det(A)$ *módulo* m *tem um inverso multiplicativo módulo* m*.*

Demonstração: Seja $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, com a, b, c e $d \in \mathbb{Z}_m$ logo $det(A) = D = ad - bc \in \mathbb{Z}_m$. Suponhamos que a matriz A possua uma inversa multiplicativa módulo m, isto é, existe uma matriz A^{-1} com entrada em $\mathbb{Z}_m$ tal que,

$$A \cdot A^{-1} = A^{-1} \cdot A = I. \tag{3.1}$$

Tomando determinantes, em (3.1):

$$det(A) \cdot det(A^{-1}) = det(A \cdot A^{-1}) = det(I) = 1 \mod m.$$

Logo $det(A^{-1})$ é o inverso multiplicativo módulo m de $det(A)$.

Reciprocamente, suponhamos que $mdc(m, D) = 1$. Então, existe $D^{-1} \in \mathbb{Z}_m$, tal que, $D \cdot D^{-1} = 1 \mod m$. É fácil verificar que

$$A^{-1} = \begin{bmatrix} D^{-1}d & -D^{-1}b \\ -D^{-1}c & D^{-1}a \end{bmatrix},$$

é a matriz inversa de A.

■

Uma consequência imediata da Proposição 3.1.1 e do Teorema 3.1.2 é o seguinte corolário:

Corolário 3.1.1 *Uma matriz quadrada* A *com entradas em* $\mathbb{Z}_m$ *é inversível módulo* m *se, e somente se,* m *e o resíduo de* $det(A)$ *módulo* m *não tem fatores primos comuns.*

Exemplo 3.1.4 *Uma matriz $A_{n\times n}$ é inversível módulo* 26 *se, e somente* $mdc(det(A), 26) = 1$ *ou seja* $det(A)$ *não pode ser múltiplo de* 2 *ou* 13.

3.1.3 Cifras de Hill

Criado por Lester S. Hill em 1929, este método de Criptografia, envolve-se duas matizes, uma para cifrar e uma para decifrar, sendo um tipo de cifra de substituição, em bloco, baseada em álgebra linear e aritmética modular, que não é vulnerável à análise de frequência das letras do alfabeto, porém pode ser quebrada utilizando recursos da própria álgebra linear.

Inicialmente vamos supor que cada letra de texto comum e de texto cifrado, excetuando o Z, tem um valor numérico que específica sua posição no alfabeto padrão. A Z será atribuído o valor zero, pois estaremos interessados em trabalhar com aritmética módulo 26, gerando a seguinte tabela:

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

Tabela 3.3: Relação de letras e números

Neste método a mensagem original é dividida em blocos de n letras, cuja a mensagem numérica equivalente $M = (m_1, m_2, \cdots, m_n)$, usando a tabela (3.3), é transformada pela chave $A = [a_{ij}]_n,\ a_{ij} \in \mathbb{Z}_{26}$, na mensagem cifrada $C = (c_1, c_2, \cdots, c_n)$, em $\mathbb{Z}_{26}$, ou seja, c_i é uma combinação linear de $m_i, \forall i$ tal que $1 \leqslant i \leqslant n$. Isto é,

$$\begin{array}{ccccccc} c_1 & = & a_{11}m_1 & + & \cdots & + & a_{1n}m_n\ (mod\ 26) \\ c_2 & = & a_{21}m_1 & + & \cdots & + & a_{2n}m_n\ (mod\ 26) \\ \vdots & & \vdots & & \vdots & & \vdots \\ c_1 & = & a_{n1}m_1 & + & \cdots & + & a_{nn}m_n\ (mod\ 26). \end{array}$$

Em notação matriciais temos:

$$C = A \cdot M,$$

onde

$$C = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},\quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \quad e \quad M = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix}$$

são matrizes com entradas em $\mathbb{Z}_{26}$.

Observação 3.1.2 *O número n é a ordem da matriz M chave de codificação e ainda, caso o texto a ser codificado não seja multiplo de n devemos adicionar letras fictícia para completar o último bloco do texto.*

Já para decodificar uma mensagem o receptor deve ter em sua posse a matriz inversa A^{-1}, o que obriga a matriz A ser inversível módulo 26. Sendo assim pode-se fazer:

$$C = A \cdot M \quad \Rightarrow \quad A^{-1} \cdot C = A^{-1} \cdot A \cdot M \quad \Rightarrow \quad A^{-1} \cdot C = M$$

e com uso da tabela (3.3) terá a mensagem original de volta.

Para facilitar as contas é bom termos em mãos a tabela de inverso módulo 26 que explicitarei abaixo.

$\bar{a}$	1	3	5	7	9	11	15	17	19	21	23	25
$\bar{a}^{-1}$	1	9	21	15	3	19	7	23	11	5	17	25

Tabela 3.4: Inversos multiplicativos módulo 26

A seguir faremos um roteiro detalhando passo a passo desse algoritmo de codificação e decodificação através de um exemplo.

Exemplo 3.1.5 *Mostre se possível, como cifrar e decifrar a mensagem,* **A MATEMÁTICA GOVERNA O MUNDO**, *utilizando como chave a matriz,* $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 3 & 2 & 3 \end{bmatrix}$.

Sendo o $det(A) = 7$, temos que existe A^{-1} módulo 26 (pois $mdc(7, 26) = 1$), logo é possível enviar a mensagem, pois o receptor terá como decodificá-la.

Primeiramente vamos escrever a citada mensagem em forma matricial e fazer conversão da mesma em números de acordo com a tabela 3.3

$$\begin{bmatrix} A & M & A & T & E & M & A & T \\ I & C & A & G & O & V & E & R \\ N & A & O & M & U & N & D & O \end{bmatrix} = \begin{bmatrix} 1 & 13 & 1 & 20 & 5 & 13 & 1 & 20 \\ 9 & 3 & 1 & 22 & 15 & 22 & 5 & 18 \\ 14 & 1 & 15 & 13 & 21 & 14 & 4 & 15 \end{bmatrix} = M$$

assim podemos fazer a multiplicação de A por M módulo 26,

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 3 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 13 & 1 & 20 & 5 & 13 & 1 & 20 \\ 9 & 3 & 1 & 22 & 15 & 22 & 5 & 18 \\ 14 & 1 & 15 & 13 & 21 & 14 & 4 & 15 \end{bmatrix} \equiv \begin{bmatrix} 25 & 4 & 18 & 23 & 20 & 1 & 11 & 21 \\ 16 & 23 & 19 & 21 & 19 & 15 & 20 & 21 \\ 11 & 22 & 24 & 13 & 4 & 21 & 25 & 11 \end{bmatrix}$$

e através da tabela 3.3 fazer a conversão de números para letras,

$$\begin{bmatrix} 25 & 4 & 18 & 23 & 20 & 1 & 11 & 21 \\ 16 & 23 & 19 & 21 & 19 & 15 & 20 & 21 \\ 11 & 22 & 24 & 13 & 4 & 21 & 25 & 11 \end{bmatrix} = \begin{bmatrix} Y & D & R & W & T & A & K & U \\ P & V & S & U & S & O & T & U \\ k & V & X & M & D & U & Y & K \end{bmatrix}$$

assim a mensagem a ser enviada será:
Y D R W T A K U P V S U S O T U K V X M D U Y K.

De posse da tabela (3.3), da matriz A^{-1} e tendo recebido a mensagem **Y D R W T A K U P V S U S O T U K V X M D U Y K** o receptor deve fazer os seguinte procedimentos.

Com auxílio da tabela (3.3) para transformar letras em números

$$\begin{bmatrix} Y & D & R & W & T & A & K & U \\ P & V & S & U & S & O & T & U \\ K & V & X & M & D & U & Y & K \end{bmatrix} = \begin{bmatrix} 25 & 4 & 18 & 23 & 20 & 1 & 11 & 21 \\ 16 & 23 & 19 & 21 & 19 & 15 & 20 & 21 \\ 11 & 22 & 24 & 13 & 4 & 21 & 25 & 11 \end{bmatrix}.$$

fazer o produto de $A^{-1} \cdot C \equiv M \mod 26$, ou seja,

$$\begin{bmatrix} 1 & 11 & 22 \\ 0 & 19 & 11 \\ 25 & 11 & 23 \end{bmatrix} \cdot \begin{bmatrix} 25 & 4 & 18 & 23 & 20 & 1 & 11 & 21 \\ 16 & 23 & 19 & 21 & 19 & 15 & 20 & 21 \\ 11 & 22 & 24 & 13 & 4 & 21 & 25 & 11 \end{bmatrix} \equiv \begin{bmatrix} 1 & 13 & 1 & 20 & 5 & 13 & 1 & 20 \\ 9 & 3 & 1 & 22 & 15 & 22 & 5 & 18 \\ 14 & 1 & 15 & 13 & 21 & 14 & 4 & 15 \end{bmatrix}.$$

relacionar a matriz encontrada com as letras da tabela 3.3,

$$\begin{bmatrix} 1 & 13 & 1 & 20 & 5 & 13 & 1 & 20 \\ 9 & 3 & 1 & 22 & 15 & 22 & 5 & 18 \\ 14 & 1 & 15 & 13 & 21 & 14 & 4 & 15 \end{bmatrix} = \begin{bmatrix} A & M & A & T & E & M & A & T \\ I & C & A & G & O & V & E & R \\ N & A & O & M & U & N & D & O \end{bmatrix}$$

encontrando assim a mensagem que lhe foi enviada, “A MATEMÁTICA GOVERNA O MUNDO.”

3.2 Matrizes e sequência de Fibonacci

É possível observar sequência numérica em diversas situações do nosso cotidiano. Dentre estas, a sequência de Fibonacci merece um destaque especial por conta de sua aplicabilidade, propriedades e de suas curiosidades. Sendo esta uma proposta para ensinar alguns conceitos de matrizes aos alunos da educação básica, para que haja a fixação de conteúdos e novas descobertas.

3.2.1 Um breve histórico sobre a origem da sequência de Fibonacci

Segundo Mol (2013), "Leonardo de Pisa (c. 1170-1250), também conhecido como Fibonacci, é considerado o mais importante matemático da Europa Medieval". Além disso, Eves (2011) também afirma que "Fibonacci foi um matemático invulgarmente capaz, sem rivais nos nove séculos da Idade Média".

O nome Fibonacci quer dizer "filho de Bonnacci". Seu pai Guiliermo Bonnacci era ligado aos negócios mercantis e passou a trabalhar em Benjaia, na África. Devido a isso, Leonardo teve uma parte de sua educação no continente africano, onde teve contato com a álgebra árabe e com o sistema de numeração indo-arábico. Além disso, conheceu a cultura matemática de diferentes povos (Egito, Sicília, Grécia, Síria e Provença).

Em 1202, Fibonacci(assim chamaremos Leonardo de Pisa), voltando para a cidade de Pisa, escreveu o livro Liber Abaci (Livro do ábaco ou Livro do Cálculo); já em 1220, escreveu a obra Practica Geometriae e, por volta de 1225, escreveu os livros Liber Quadratorum e Flos. No Liber Abaci, Fibonacci introduziu o sistema de numeração indo-arábico na Europa, caracterizando-o com nove símbolos e o zero, apresentando aplicações à matemática comercial, conversões de pesos e medidas, cálculos de taxas de juros e de câmbio, médias, entre outras. Apesar das notórias contribuições de Fibonacci para a matemática nas mais diversos áreas, segundo Mol (2013) "Fibonacci é hoje, no entanto, mais conhecido pela chamada sequência de Fibonacci, apresentada no Liber Abaci como resposta para um problema envolvendo o crescimento de uma população de coelhos".

Vejamos o problema proposto por Fibonacci.

Quantos casais de coelhos teriam ao final de 1 ano se:

- No primeiro mês temos um coelho macho e um coelho fêmea. Estes dois coelhos acabaram de nascer.
- Um coelho só atinge a maturidade sexual ao fim de um mês.
- O período de gestação de um coelho dura um mês.
- Ao atingirem a maturidade sexual, a fêmea irá dar à luz todos os meses.
- A mãe irá dar a luz todos os meses um coelho macho e um coelho fêmea.
- Os coelhos nunca morrem.

De acordo com Hefez (2011) a solução que Fibonacci propôs foi:

mês	número de casais do mês anterior	número de casais recém-nascidos	total
1^{o}	0	1	1
2^{o}	1	0	1
3^{o}	1	1	2
4^{o}	2	1	3
5^{o}	3	2	5
6^{o}	5	3	8
7^{o}	8	5	13
8^{o}	13	8	21
9^{o}	21	13	34
10^{o}	34	21	55
11^{o}	55	34	89
12^{o}	89	55	144

Tabela 3.5: Solução do problema dos coelhos de Fibonacci

3.2.2 Definição da sequência de Fibonacci

Assim a solução para o problema anterior é

$$(1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144).$$

Fibonacci em suas observações percebeu que cada termo da sequência a partir do terceiro termo é obtido através da soma dos dois termos antecessores. Surge então uma sequência fabulosa, que possui diversas propriedades interessantes em vários ramos da Matemática, além de ter inúmeras aparições em fenômenos da natureza.

A sequência supracitada que será nosso objeto de estudo nesta seção é a sequência

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...$$

que chamaremos de sequência de Fibonacci, onde os termos são chamados de números de Fibonacci.

Aqui devemos nos perguntar, existe alguma relação matemática que define essa sequência?

A resposta é afirmativa, existe sim uma relação que define a sequência de Fibonacci. Vejamos:

Definição 3.2.1 *Chama-se sequência de Fibonacci a sequência* (f_n) *definida por*

$$f_{n+1} = f_n + f_{n-1}, \forall n \geq 2, \tag{3.2}$$

onde $f_1 = f_2 = 1$, *"a sequência de Fibonacci é, por definição, uma sequência recursiva" Contador (2012).*

3.2.3 Aparições na natureza

Vamos apresentar algumas situações na natureza onde a sequência de Fibonacci se manisfesta nas mais variadas formas.

3.2.3.1 Aparições nas flores

Inicialmente, mostraremos algumas aparições da sequência de Fibonacci nas flores, mas vejamos bem, são aparições, nem todas as flores apresentam relações com esta sequência.

O nosso primeiro exemplo é da margarida, que de acordo com Zahn (2011) "As margaridas geralmente têm 13, 21, 34, 55 ou 89 pétalas". Note que estes são números consecutivos da sequência de Fibonacci.

Figura 3.1: Margarida de 13 pétalas

Um outro exemplo onde aparecem os números de Fibonacci são os lírios, em que a quantidade de sépalas (que tem a função de proteger a estrutura) e as pétalas (cuja função é atrair polinizadores) são um número de Fibonacci. De acordo com Ferri (1983), "os lírios possuem 3 pétalas e 3 sépalas". Vejamos:

Figura 3.2: Lírio vermelho

Podemos ainda citar as quaresmeiras, flores que segundo Ferri (1983) “possuem 5 pétalas”. Logo temos mais um exemplo de flor em que o número de pétalas é um número de fibonacci. Observemos:

Figura 3.3: Quaresmeira. Foto: Jardineiro.net

Embora existam vários outros exemplos dentro da natureza em que aparecem os números de Fibonacci ou a própria sequência, finalizaremos essa seção apresentando duas manifestações interessantes da sequência citada. A primeira delas é o “girassol”, onde suas sementes preenchem o miolo dispostas em dois conjuntos de espirais, sendo 21 no sentido horário e 34 no anti-horário. A segunda é a “pinha”, onde as sementes crescem e se organizam em dois conjuntos de espirais, sendo 08 espirais no horário e 13 no sentido anti-horário. Vejamos na imagem abaixo:

Figura 3.4: Sementes de girassol e pinha

3.2.4 Sequência de Fibonacci e matrizes

A sequência de Fibonacci, também aparece no contexto das matrizes. Existe uma matriz que a medida que vamos elevando-a a expoentes inteiros positivos maiores que um obtemos como resposta, os números de Fibonacci. Dada a matriz

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

observemos o seu comportamento. Se fizermos A^2, teremos:

$$A^2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} f_3 & f_2 \\ f_2 & f_1 \end{bmatrix},$$

onde todos os elementos da matriz são números de Fibonacci. Fazendo agora A^3, obtemos:

$$A^3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^3 = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} f_4 & f_3 \\ f_3 & f_2 \end{bmatrix},$$

onde também todos os elementos são números de Fibonacci. Continuando com raciocínio análogo, calculando A^4, encontramos:

$$A^4 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^4 = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} f_5 & f_4 \\ f_4 & f_3 \end{bmatrix}.$$

Calculando agora A^5, temos:

$$A^5 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^5 = \begin{bmatrix} 8 & 5 \\ 5 & 3 \end{bmatrix} = \begin{bmatrix} f_6 & f_5 \\ f_5 & f_4 \end{bmatrix}.$$

Novamente temos todos os elementos da matriz sendo números de Fibonacci. Seguindo esse raciocínio, é natural conjecturarmos[2] que

$$A^n = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n = \begin{bmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{bmatrix}.$$

De fato, a conjectura é verdadeira e vamos prová-la por indução[3].

[2]Conjecturar: Ato ou efeito de inferir ou deduzir que algo é provável, com base em presunções, evidências incompletas, pressentimentos; conjetura, hipótese, presunção, suposição.

[3]Ver Apêndice.

Demonstração: Note que para $n = 2$, a nossa conjectura é verdadeira, pois

$$A^2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} f_3 & f_2 \\ f_2 & f_1 \end{bmatrix}.$$

Suponhamos que seja verdadeira a conjectura para $n = k \in \mathbb{N}$, isto é,

$$A^k = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k = \begin{bmatrix} f_{k+1} & f_k \\ f_k & f_{k-1} \end{bmatrix}.$$

E queremos mostrar que também é válida para $n = k + 1 \in \mathbb{N}$, ou seja,

$$A^{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{k+1} = \begin{bmatrix} f_{k+2} & f_{k+1} \\ f_{k+1} & f_k \end{bmatrix},$$

com efeito,

$$A^{k+1} = A^k.A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k . \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} f_{k+1} & f_k \\ f_k & f_{k-1} \end{bmatrix} . \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow$$

$$A^{k+1} = \begin{bmatrix} f_{k+1} + f_k & f_{k+1} \\ f_k + f_{k-1} & f_k \end{bmatrix} = \begin{bmatrix} f_{k+2} & f_{k+1} \\ f_{k+1} & f_k \end{bmatrix}.$$

Provando que a validez de $n = k \in \mathbb{N}$ implica a validade para $n = k + 1 \in \mathbb{N}$ e desta forma, por indução temos que o resultado é verdadeiro para todo $n \in \mathbb{N}$. ■

3.2.4.1 Aplicação em determinantes de ordem maior que 2

Como dito anteriormente, a sequência de Fibonacci e os seus números estão presentes em algumas matrizes. Agora mostraremos outra situação desta aparição, desta vez, no contexto dos determinantes.

Inicialmente, vamos calcular o determinante da matriz

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 5 & 8 \\ 13 & 21 & 34 \end{pmatrix}.$$

Fazendo alguns cálculos, encontramos a resposta $det\ A = 0$.

Agora calculemos $det\ B$, onde

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 8 & 13 \\ 21 & 34 & 55 \end{pmatrix}.$$

Realizando os cálculos necessários, a resposta encontrada é $det\ B = 0$.

Calculando $det\ C$, onde

$$C = \begin{pmatrix} 2 & 3 & 5 \\ 8 & 13 & 21 \\ 34 & 55 & 89 \end{pmatrix}$$

obtemos $det\ C = 0$, calculando ainda $det\ D$, onde

$$D = \begin{pmatrix} 8 & 13 & 21 \\ 34 & 55 & 89 \\ 144 & 233 & 377 \end{pmatrix},$$

temos como resultado $det\ D = 0$.

Agora podemos fazer o seguinte questionamento, o que essas quatro matrizes têm em comum, além de terem seus respectivos determinantes iguais a 0? O leitor mais atento, já deve ter percebido que todos os elementos das quatro matrizes são números de Fibonacci e também, são números consecutivos desta sequência.

Podemos ainda fazer outra pergunta: todas as matrizes de ordem 3, cujos elementos são números consecutivos de Fibonacci, têm sempre seu determinante nulo?

A resposta é sim, e a sua justificativa é bastante simples. Como a sequência de Fibonacci é definida como a soma dos dois termos anteriores, então a terceira coluna das matrizes de ordem 3 será a soma das duas primeiras colunas. Sendo assim, a $3^{\underline{a}}$ coluna é combinação linear das duas primeiras, e segundo Iezzi e Hazzan (1977) "se uma matriz quadrada $M = [a_{ij}]$, de ordem n tem uma linha (ou coluna) que é combinação linear de outras linhas (ou colunas), então det M = 0", o que completa a nossa justificativa. O leitor pode encontrar a demonstração da afirmação no próprio livro.

De posse das afirmações do parágrafo anterior podemos fazer a seguinte generalização:

Dada uma matriz $A = [a_{ij}]$, de ordem n, tal que $n \geq 3$ cujos termos são números consecutivos de Fibonacci, então $det\ A = 0$. O fato se dá pela mesma justificativa anterior.

3.3 Cadeia de Markov

Esta seção apresenta as cadeias de Markov com uma linguagem que possa ser aplicado no ensino médio. Faz se uma introdução da teoria de probabilidade, possibilitando oferecer ao aluno a oportunidade de ter uma visão aplicada de matrizes em outras áreas do conhecimento. Tornando as aulas mais dinâmicas e atraentes, sobretudo com relação ao tópico matrizes.

3.3.1 Introdução

Em 1907 o matemático russo Andrei Andreyevich Markov começou o estudo deste importante tipo de processo, que apenas o resultado de uma dada experiência atual pode afetar o resultado da experiência seguinte, ou seja as experiências anteriores não influenciam as experiências futuras. Esta propriedade é conhecida como "perda de memória"ou Propriedade de Markov, e é o que caracteriza uma Cadeia de Markov (também conhecida como Processo Markoviano).

Para formalizar o estudo das Cadeias de Markov faz-se necessário os conceitos básicos de probabilidade.

3.3.2 Conceitos básicos de probabilidade

Para uma melhor compreensão da definição de probabilidade, devemos ter bem claro o que é experimento aleatório, espaço amostral e evento.

3.3.2.1 Experimento aleatório

Chamamos de experimentos aleatórios aqueles que, repetidos em idênticas condições, produzem resultados diferentes. Embora não saibamos qual o resultado que irá ocorrer num experimento, em geral, conseguimos descrever o conjunto de todos os resultados possíveis que podem ocorrer. As variações de resultados, de experimento para experimento, são devidas a uma multiplicidade de causas que não podemos controlar, as quais denominamos acaso, Hazzan (2013).

3.3.2.2 Espaço amostral

Chamamos de espaço amostral, e normalmente é denotado pela letra grega Ω, o conjunto não-vazio de todas as possibilidades de resultados de um experimento.

Exemplo 3.3.1 *No lançamento simultâneo de duas moedas qual é o espaço amostral? Denotaremos k como cara e c como coroa. Assim segue que:*

$$\Omega = \{(c, c); (c, k); (k, c); (k, k)\}.$$

Observação 3.3.1 *Denotaremos por #(A) o número de elemento do conjunto A, logo diremos que o Ω é finito, se $\#\ (\Omega) = n$, com $n \in \mathbb{N}$; caso contrário diremos que Ω é infinito.*

3.3.2.3 Evento

Evento é um subconjunto do espaço amostral, normalmente denotado por uma letra maiúscula. Entre os eventos, salientamos o evento impossível, simbolizado por $\varnothing$ (conjunto vazio) e o próprio Ω chamado evento certo.

Se usarmos certas operações entre eventos (conjuntos), poderemos combinar eventos (conjuntos), para formar novos eventos (conjuntos). Considere A e B eventos de Ω, então:
i) O conjunto $A \cup B$ será também um evento que ocorrerá se, e somente se, A ou B (ou ambos) ocorrem. Dizemos que $A \cup B$ é a união entre o evento A e o evento B.
ii) Veja que $A \cap B$ será também um evento que ocorrerá se, e somente se, A e B ocorrerem simultaneamente. Dizemos que $A \cap B$ é a interseção entre o evento A e o evento B. Em particular, se $A \cap B = \varnothing$. A e B são chamados mutuamente exclusivos.
iii) O complementar de A, denotado por $A^{\subset}$ será também um evento que ocorrerá se, e somente se, A não ocorrer ou seja $A^{\subset} = \Omega - A$.

Se tivermos uma sequência $A_1, A_2, \ldots, A_n$ de eventos, então

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n, \tag{3.3}$$

que também é um evento e ocorrerá se, e somente se, ao menos um dos eventos A_i ocorrer. Dizemos que $A_1 \cup A_2 \cup \ldots \cup A_n$ é a união dos eventos $A_1, A_2, \ldots, A_n$.

Além disso, podemos obter o seguinte conjunto

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n, \tag{3.4}$$

que será também um evento que ocorrerá se, e somente se, todos os eventos A_i ocorrerem simultaneamente.

Exemplo 3.3.2 *Um dado é lançado e observado o número da face de cima.*
Logo, $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Sejam os eventos:
A: ocorrência de número par $= \{2, 4, 6\}$
B: ocorrência de número maior ou igual a $4 = \{4, 5, 6\}$
C: ocorrência de número ímpar $= \{1, 3, 5\}$.

Então teremos:
i) $A \cup B = \{2, 4, 5, 6\}$
ii) $A \cup C = \Omega$
iii) $A \cap B = \{4, 6\}$
iv) $B \cap C = \emptyset$

3.3.2.4 Definição de probabilidade

Segundo Morgado et al. (1991) "a definição de probabilidade como quociente do número de 'casos favoráveis' sobre o número de 'casos possíveis' foi a primeira definição

formal de probabilidade, e apareceu pela primeira vez em forma clara na obra Liber de Ludo Aleae de Jerônimo Cardano (1501-1576)."Assim podemos definir a probabilidade P de um evento A como

$$P(A) = \frac{\#(A)}{\#(\Omega)} \tag{3.5}$$

Uma consequência imediata desta definição são as seguintes propriedades:
i) Para todo evento A, $0 \leq P(A) \leq 1$;
ii) $P(\Omega) = 1$;
iii) $P(\varnothing) = 0$ (porque $\#(\varnothing) = 0$);
iv) Se $A \cap B = \varnothing$ então $P(A \cup B) = P(A) + P(B)$.
v) Se $A \cap B \neq \varnothing$ então $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Observação 3.3.2 *Uma demostração para as propriedades pode ser encontrada em Morgado et al. (1991).*

Exemplo 3.3.3 *Uma urna contém* 100 *bolinhas idênticas numeradas de* 1 *a* 100*. Uma bolinha é escolhida e observado seu número. Qual a probabilidade de observarmos um número múltiplo de* 6 *e de 8 simultaneamente?*
Veja que $\Omega = \{1, 2, 3, \cdots, 99, 100\}$*. Além disso, um múltiplo de* 6 *e* 8 *simultaneamente terá que ser múltiplo de* 24*, portanto o evento* A *será;* $A = \{24, 48, 72, 96\}$*. Portanto, segue que:*

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{4}{100} = \frac{1}{25}.$$

3.3.2.5 Variável aleatória

Em matemática, uma variável aleatória pode ser definida como uma função que associa a todo evento pertencente a uma partição do espaço amostral Ω a um único número real, isto é, $X : \Omega \to \mathbb{R}$. É comum a representação das variáveis aleatórias por letras maiúsculas e dos valores assumidos por letras minúsculas.

Quando a imagem (variação) de X é finita ou infinita contável, a variável aleatória é chamada de variável aleatória discreta.

Exemplo 3.3.4 *O lançamento de um dado de seis lados é um exemplo de variável aleatória discreta finita. O dado fornece um valor inteiro em todos os lançamentos, de modo que não existe a possibilidade de ele cair de lado e fornecer um valor fracionário como* $2,5555$*.*

Exemplo 3.3.5 *Já o número de carros que passam por um pedágio é um exemplo de variável aleatória discreta infinita. Passará uma infinidade de carros, porém nunca passará a metade de um carro por um pedágio (não haverá frações no número de carros que passarão por um pedágio).*

3.3.3 Cadeia de Markov

Um Processo de Markov é um processo estocástico[1] onde as distribuições de probabilidade para o seu desenvolvimento futuro dependem somente do estado presente, não levando em consideração como o processo chegou a tal estado.

Segundo Grigoletti (2010) os processos markovianos são modelados formalmente pelos modelos de Markov, que são sistemas de transições de estados, onde os estados são representados em termos de seus vetores probabilísticos, que podem variar no espaço temporal (discreto ou contínuo), e as transições entre estados são probabilísticas e dependem apenas do estado corrente. Se o espaço de estados é discreto (enumerável), então o modelo de Markov é denominado de cadeias de Markov, uma definição mais formal é dada a seguir .

Definição 3.3.1 *Uma sucessão de variáveis aleatórias* $\{X_t\}$ *é chamada de Cadeias de Markov se para* $t = 0, 1, 2, \cdots$ *e todos os estados* i_t

$$P(X_t = i_t | X_{t-1} = i_{t-1},\ X_{t-2} = i_{t-2},\ \cdots,\ X_1 = i_1,\ X_0 = i_0) = P(X_t = i_t | X_{t-1} = i_{t-1}). \tag{3.6}$$

Em uma Cadeia markoviana com n estados exaustivos e mutuamente exclusivos, as probabilidades em um ponto específico do tempo $t = 0, 1, 2, \cdots$ são habitualmente expressas por

$$p_{ij} = P(x_t = j | X_{t-1} = i),\ (i, j) = 1, 2, \cdots, n,\quad t = 0, 1, 2, \cdots, T. \tag{3.7}$$

Isso é conhecido como probabilidade de transição em uma etapa de passar do estado i em $t-1$ ao estado j em t. Por definição,

$$\sum_j p_{ij} = 1,\ i = 1, 2, \cdots, n \ \text{ e } \ p_{ij} \geq 0,\ (i, j) = 1, 2, \cdots, n. \tag{3.8}$$

Um modo conveniente de resumir as probabilidades de transição em uma etapa é usar a seguinte notação matricial:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nn} \end{pmatrix}.$$

[1]Sendo que o processo estocástico de tempo discreto é simplesmente uma descrição da relação entre as variáveis aleatórias $x_0, x_1, \cdots$, ou seja é um processo que evolui de forma aleatória ao longo do tempo tomando valores (estados) de um conjunto dado.

A matriz P define a denominada cadeia de Markov. A seguir temos um exemplo de cadeia de Markov.

Exemplo 3.3.6 *Todo ano, no inicio da estação de plantio de mudas (março a setembro), um jardineiro usa um teste químico para verificar a condição do solo. Dependendo do resultado do teste, a produtividade para a nova estação cai em um dos três estados: 1) bom: 2) razoável e 3) ruim. Ao longo dos anos, o jardineiro observou que a condição do solo no ano anterior causava um impacto sobre a produtividade no ano corrente e que a situação podia ser descrita pela seguinte cadeia de Markov:*

$$\begin{array}{c} \\ \text{Estado presente} \end{array} \begin{array}{c} \text{Estado futuro} \\ \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \left[\begin{array}{ccc} 0,2 & 0,5 & 0,3 \\ 0 & 0,5 & 0,5 \\ 0 & 0 & 1 \end{array}\right]. \end{array} \tag{3.9}$$

As probabilidades de transição mostram que a condição do solo pode se deteriorar ou se manter, mas nunca melhorar. Se a condição do solo neste ano for boa (estado 1), há 20% de chance de não mudar no ano seguinte, 50% de chance de se tornar razoável (estado 2) e 30% de chance de deteriorar até uma condição ruim (estado 3). Se a condição do solo neste ano for razoável (estado 2), a produtividade no ano seguinte pode permanecer razoável com probabilidade de 50% ou torna-se ruim (estado 3). Também com probabilidade de 50%. Por fim, uma condição ruim neste ano (estado 3) só pode resultar em igual condição no próximo ano (com probabilidade 1).

O jardineiro pode alterar as probabilidades de transição P usando fertilizante para melhorar a condição do solo. Nesse caso, a matriz de transição se torna:

$$\begin{array}{c} \\ \text{Estado presente} \end{array} \begin{array}{c} \text{Estado futuro} \\ \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \left[\begin{array}{ccc} 0,30 & 0,60 & 0,10 \\ 0,10 & 0,60 & 0,30 \\ 0,05 & 0,40 & 0,55 \end{array}\right]. \end{array} \tag{3.10}$$

Agora, a utilização do fertilizante permite melhorias na condição de deterioração. Há 10% de chance da condição do solo mudar de razoável para boa (estado 2 para estado 1), 5% de chance de mudar de ruim para boa (estado 3 para estado 1) e 40% de chance de uma condição ruim torna-se razoável (estado 3 para estado 2).

Cadeias de Markov são frequentemente descritas por um diagrama, onde as bordas do diagrama são rotulados pelas probabilidades de ir de um estado no tempo n para outros estados no tempo $n+1$. A mesma informação é representada pela matriz de transição do momento n para o tempo $n+1$.

Exemplo 3.3.7 *O diagrama abaixo, é usado para mostrar de forma clara as transições dos estado* $1, 2$ *e* 3.

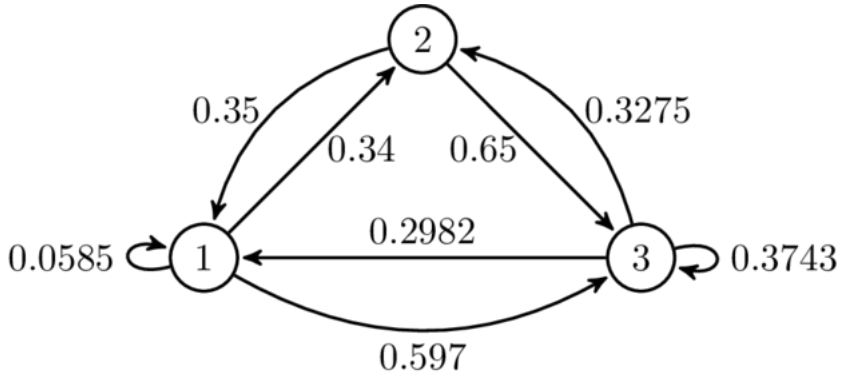

Figura 3.5: Diagrama de transição

As informações contida no diagrama (3.5), pode ser representado pela matriz de transição:

$$\begin{array}{c} \\ \text{Estado presente} \end{array} \begin{array}{c} \text{Estado futuro} \\ \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \begin{bmatrix} 0.0585 & 0.34 & 0.597 \\ 0.35 & 0 & 0.65 \\ 0.2982 & 0.3275 & 0.3743 \end{bmatrix} \end{array}. \tag{3.11}$$

3.3.3.1 Classificação dos estados

Os estados de uma cadeia de markov podem ser classificados com base na probabilidade de transição p_{ij} de P; ou seja, um estado é **absorvente** se retorna para ele mesmo, assim p_{ii} é igual a 1, mas se um estado puder alcançar outro estado e não pode voltar a ele mesmo será chamado de **transiente**, e um estado será denominado como **recorrente** se a probabilidade de voltar ao mesmo estado em que estava com base em outros estado for 1, e por último, um estado é **periódico** com período $t > 1$ se um retorno só for possível em $t, 2t, 3t, \cdots$ etapas.

Um conjunto de estados é dito um conjunto fechado se o processo ao entrar em um desses estados, este irá permanecer nos estados indefinidamente. Com isso, pode-se afirmar que o conjunto é formado por estados recorrentes.

Exemplo 3.3.8 *Suponha que a cadeia de Markov possui a seguinte matriz de transição:*

$$\mathrm{P}= \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0,2 & 0,8 & 0 & 0 & 0 \\ 0,5 & 0,5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0,35 & 0,65 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

O estado 3 *é transiente porque se o processo está no estado* 3*, há uma probabilidade positiva que ele nunca irá retornar para este estado. O estado* 4 *também é um estado transiente porque se o processo começa neste estado, imediatamente o processo o deixa e nunca mais irá retornar a este estado.*

Os estados 0 *e* 1 *são recorrentes. Através de P percebe que se o processo começar a partir de um desses dois estados, este nunca deixará estes dois estados. Além disto, sempre quando o processo move-se a partir de um destes estados para o outro, este irá retornar para o estado original eventualmente.*

O estado 2 *é um estado absorvente, pois, uma vez que o processo entra no estado* 2*, este nunca mais o deixará.*

Os estados $0, 1,$ *e* 2 *formam um conjunto fechado, uma vez que se o processo entrar em um destes estados, nunca os deixará.*

3.3.3.2 Matriz de transição de N fases

Dadas as probabilidades iniciais $a^{(0)} = (a_j^{(0)})$ de iniciar no estado j e a matriz de transição P de uma cadeia de Markov, as probabilidades absolutas $a^{(n)} = (a_j^{(n)})$ de estar no estado j após n transições $(n > 0)$ são calculadas da seguinte maneira:

$$\begin{aligned} a^{(1)} &= a^{(0)}P \\ a^{(2)} &= a^{(1)}P = a^{(0)}PP = a^{(0)}P^2 \\ a^{(3)} &= a^{(2)}P = a^{(0)}P^2P = a^{(0)}P^3 \end{aligned}$$

Continuando da mesma maneira, obtemos

$$a^n = a^{(0)}P^n, \quad n = 1, 2, \cdots$$

A matriz P^n é como matriz de transição em n etapas. Por esses cálculos, podemos ver que

$$P^n = P^{n-1}P \quad \text{ou} \quad P^n = P^{n-m}P^m, \quad 0 < m < n$$

Essas equações são conhecidas como equações de Chapman-Kolomogorov.

Exemplo 3.3.9 *A seguinte matriz de transição se aplica ao problema do jardineiro, com fertilizante, Exemplo (3.3.6):*

$$\mathrm{P}= \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0,30 & 0,60 & 0,10 \\ 0,10 & 0,60 & 0,30 \\ 0,05 & 0,40 & 0,55 \end{bmatrix} \end{array}$$

A condição inicial do solo é boa, isto é $a^{(0)} = (1, 0, 0)$. *Determine as probabilidades absolutas dos três estados do sistema após* $1, 8$ *e* 16 *estações de plantio de mudas.*

$$P \cdot P = \begin{bmatrix} 0,30 & 0,60 & 0,10 \\ 0,10 & 0,60 & 0,30 \\ 0,05 & 0,40 & 0,55 \end{bmatrix} \cdot \begin{bmatrix} 0,30 & 0,60 & 0,10 \\ 0,10 & 0,60 & 0,30 \\ 0,05 & 0,40 & 0,55 \end{bmatrix} = \begin{bmatrix} 0,155 & 0,58 & 0,265 \\ 0,105 & 0,54 & 0,355 \\ 0,0825 & 0,49 & 0,4275 \end{bmatrix} = P^2$$

$$\begin{aligned} P^2 \cdot P^2 &= \begin{bmatrix} 0,155 & 0,58 & 0,265 \\ 0,105 & 0,54 & 0,355 \\ 0,0825 & 0,49 & 0,4275 \end{bmatrix} \cdot \begin{bmatrix} 0,155 & 0,58 & 0,265 \\ 0,105 & 0,54 & 0,355 \\ 0,0825 & 0,49 & 0,4275 \end{bmatrix} \\ &= \begin{bmatrix} 0,106787 & 0,53295 & 0,360262 \\ 0,102262 & 0,52645 & 0,371287 \\ 0,099506 & 0,521925 & 0,378568 \end{bmatrix} = P^4, \end{aligned}$$

$$\begin{aligned} P^4 \cdot P^4 &= \begin{bmatrix} 0,106787 & 0,53295 & 0,360262 \\ 0,102262 & 0,52645 & 0,371287 \\ 0,099506 & 0,521925 & 0,378568 \end{bmatrix} \cdot \begin{bmatrix} 0,106787 & 0,53295 & 0,360262 \\ 0,102262 & 0,52645 & 0,371287 \\ 0,099506 & 0,521925 & 0,378568 \end{bmatrix} \\ &= \begin{bmatrix} 0,101752 & 0,525513 & 0,372733 \\ 0,101701 & 0,525434 & 0,372863 \\ 0,101669 & 0,525383 & 0,372946 \end{bmatrix} = P^8, \end{aligned}$$

$$\begin{aligned} P^8 \cdot P^8 &= \begin{bmatrix} 0,10175274 & 0,525513931 & 0,372733329 \\ 0,101701 & 0,525434 & 0,372863 \\ 0,101669 & 0,525383 & 0,372946 \end{bmatrix} \cdot \begin{bmatrix} 0,101752 & 0,525513 & 0,372733 \\ 0,101701 & 0,525434 & 0,372863 \\ 0,101669 & 0,525383 & 0,372946 \end{bmatrix} \\ &= \begin{bmatrix} 0,101694 & 0,525423 & 0,372881 \\ 0,101694 & 0,525423 & 0,372881 \\ 0,101694 & 0,525423 & 0,372881 \end{bmatrix} = P^{16} \end{aligned}$$

Assim,

$$a^{(1)} = (1\ 0\ 0) \cdot \begin{bmatrix} 0,30 & 0,60 & 0,10 \\ 0,10 & 0,60 & 0,30 \\ 0,05 & 0,40 & 0,55 \end{bmatrix} = (0,30 \quad 0,60 \quad 0,10),$$

$$a^{(8)} = (1\ 0\ 0) \cdot \begin{bmatrix} 0,101752 & 0,525513 & 0,372733 \\ 0,101701 & 0,525434 & 0,372863 \\ 0,101669 & 0,525383 & 0,372946 \end{bmatrix} = (0,101752 \quad 0,525513 \quad 0,372733),$$

$$a^{(16)} = (1\ 0\ 0) \cdot \begin{bmatrix} 0,101694 & 0,525423 & 0,372881 \\ 0,101694 & 0,525423 & 0,372881 \\ 0,101694 & 0,525423 & 0,372881 \end{bmatrix} = (0,101694 \quad 0,525423 \quad 0,372881).$$

Observa-se que as linhas da matriz $P^{(8)}$ e o vetor de probabilidades absolutas são quase idênticos. O resultado é mais pronunciado para $P^{(16)}$. Ele demonstra que, à medida que aumenta o número de transições, as probabilidades absolutas são independentes da inicial $a^{(0)}$. Veremos na próxima subseção que qualquer cadeia de Markov regular possui um vetor estado fixo q tal que, para qualquer escolha $x^{(0)}$, o $x^{(0)}P^n$ converge a q quando n aumenta.

3.3.3.3 Cadeias absorvente

Seja qual for o estado inicial da cadeia de Markov absorvente, após um número finito de transições ele estará em um dos estados absorventes e nesse estado ficará definitivamente.

Definição 3.3.2 *Diz-se que uma cadeia de Markov é absorvente se ela tem um estado absorvente e se de cada estado não absorvente é possível ir para algum estado absorvente. Essa última condição significa que, para cada estado não absorvente i, existe um estado absorvente j tal que, para algum $n, p_{ij}^{(n)} > 0$.*

Exemplo 3.3.10 *Um bêbado caminha na rua. Cada número de* 1 *a* 3 *representa um quarteirão, enquanto o número* 0 *representa a casa dele e o número* 4 *representa o bar.*

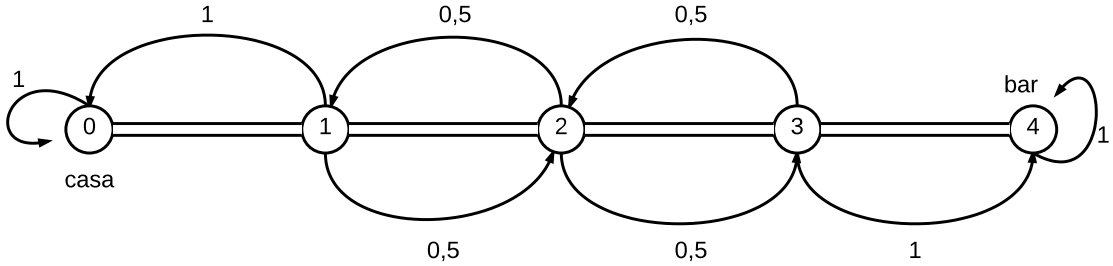

Figura 3.6: Andar do bêbado

A matriz P correspondente a este diagrama é:

$$\mathrm{P}= \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 0,5 & 0 & 0,5 & 0 & 0 \\ 0 & 0,5 & 0 & 0,5 & 0 \\ 0 & 0 & 0,5 & 0 & 0,5 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \end{array}$$

Este é um bom exemplo de uma cadeia absorvente.

A pergunta que surgem sobre esta sequência é: Na média, quantas vezes um dado estado transiente será visitado até que o processo seja absorvido?

A resposta desta questão dependem do estado inicial e da matriz de transição. Considere a matriz P com r estados absorventes e t estados transientes. A matriz P canônica é formada conforme abaixo:

i) I é uma matriz identidade r por r;

ii) O é uma matriz 0 r por t;

iii) R é uma matriz t por r;

iv) Q é uma matriz t por t.

Sendo a matriz canônica obtida reorganizando P colocando nas últimas linhas e colunas os estados absorventes.

$$P = \left[\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right] \tag{3.12}$$

Reorganizando a matriz P do Exemplo (3.3.10) na forma canônica teremos:

$$\mathrm{P}= \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 0 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 0 & 4 \end{array} \\ \left[\begin{array}{ccc|cc} 0 & 0,5 & 0 & 0,5 & 0 \\ 0,5 & 0 & 0,5 & 0 & 0 \\ 0 & 0,5 & 0 & 0 & 0,5 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \end{array}$$

Para uma cadeia de Markov absorvente, a matriz $N = (I - Q)^{-1}$ é chamada matriz fundamental para P. Sendo um elemento n_{ij} de N fornece o número esperado de vezes que o processo estará no estado transiente $a^{(j)}$ caso o estado inicial seja o estado $a^{(i)}$.

Continuando o Exemplo(3.3.10) temos $Q = \begin{bmatrix} 0 & 0,5 & 0 \\ 0,5 & 0 & 0,5 \\ 0 & 0,5 & 0 \end{bmatrix}$ assim:

$$I - Q = \begin{bmatrix} 1 & -0,5 & 0 \\ -0,5 & 1 & -0,5 \\ 0 & -0,5 & 1 \end{bmatrix}$$

Portanto,

$$N = (I-Q)^{-1} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1,5 & 1 & 0,5 \\ 1 & 2 & 1 \\ 0,5 & 1 & 1,5 \end{bmatrix} \end{array}$$

Iniciando-se no estado 2, o número médio de vezes em que o sistema permanece nos estados 1, 2 e 3 será, respectivamente, 1, 2 e 1.

3.3.3.4 Cadeia regulares

No Exemplo (3.3.9) nós vimos que os vetores estado convergem a um vetor fixo à medida que os período cresce. Uma pergunta pertinente seria se sempre os vetores estados convergem para um vetor fixo (estacionário) em uma cadeia de Markov. Um exemplo simples mostra que não

Exemplo 3.3.11 *Seja* $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ *e* $a^{(0)} = (1,0)$.
Então, como $P^2 = I$ *e* $P^3 = P$, *temos:*

$$a^{(0)} = a^{(2)} = a^{(4)} = \cdots = (1,0)$$

e

$$a^{(1)} = a^{(3)} = a^{(5)} = \cdots = (0,1).$$

Portanto este sistema oscila indefinidamente entre $(1,0)$ *e* $(0,1)$, *e portanto não converge a nenhum vetor fixo.*

Porém se impormos uma restrição à matriz de transição, pode-se mostrar que o sistema se aproxima do vetor fixo. Esta condição é descrita na definição abaixo.

Definição 3.3.3 *Uma matriz de transição é regular se uma potência positiva da matriz tem todas as entradas positivas.*

Esta é uma das mais importantes características exibidas por muitas cadeias de Markov, o comportamento de equilíbrio em longo prazo, isso significa que, as probabilidades de o sistema estar em cada um dos vários estados pouco ou nada variam.

Exemplo 3.3.12 *Vamos supor que o nível econômico de um homem é classificado em três categorias: rico (R), classe média (M), e pobre (P). Vamos supor que dos filhos de um homem rico 95% são ricos e 5% de classe média. No caso de um indivíduo de classe média, 10% são ricos, 70% de classe média e 20% pobres, e, finalmente, no caso de um homem pobre, 30% são de classe média e 70% são pobres. Supondo que cada homem tem um filho, podemos formar uma cadeia de Markov observando uma família através de gerações sucessivas. A matriz de probabilidade de transição é:*

$$\mathrm{P}= \begin{array}{c} \\ R \\ M \\ P \end{array} \begin{array}{c} \begin{array}{ccc} R & M & P \end{array} \\ \begin{bmatrix} 0,95 & 0,05 & 0 \\ 0,10 & 0,70 & 0,20 \\ 0 & 0,30 & 0,70 \end{bmatrix} \end{array} \tag{3.13}$$

Observe que existe probabilidades nulas na matriz de transição, porém

$$P^2 = \begin{bmatrix} 0,9075 & 0,0825 & 0,01 \\ 0,165 & 0,555 & 0,28 \\ 0,03 & 0,42 & 0,55 \end{bmatrix}$$

tem todas as entradas maior que zero, assim esta cadeia é regular logo tem um vetor equilíbrio a longo prazo.

Para descobrir um vetor de equilibro de uma cadeia regular não há a necessidade de fazer potência de matrizes, basta resolver a equação:

$$E = E \cdot P \tag{3.14}$$

Sendo E o vetor de equilíbrio.

Exemplo 3.3.13 *Encontre o vetor de equilíbrio do Exercício (3.3.12).*
Seja $E = (x\ y\ z)$, e pela equação (3.14), temos:

$$(x\ y\ z\ ) = (x\ y\ z\ ) \cdot \begin{bmatrix} 0,95 & 0,05 & 0 \\ 0,10 & 0,70 & 0,20 \\ 0 & 0,30 & 0,70 \end{bmatrix}$$

Que nos leva ao sistema

$$\begin{cases} 0,95x + 0,10y = x \\ 0,05x + 0,70y + 0,30z = y \\ 0,20y + 0,70z = z \end{cases},$$

sabendo que $x + y + z = 1$*, devemos substituir esta equação por uma equação do sistema, e ao resolvê-lo encontraremos,* $x = \frac{6}{11},\ y = \frac{3}{11}$ *e* $z = \frac{2}{11}$*. Sendo então* $E = (\frac{6}{11},\ \frac{3}{11},\ \frac{2}{11})$*.*

Conclusão

Motivado por um histórico de ensino de matrizes sem relação com a realidade, onde o professor é um transmissor e o aluno um receptor do conhecimento, sendo esta uma das críticas a escola dita tradicional, construímos este trabalho com o objetivo de ser um material de apoio, indicando um caminho para o ensino de matrizes. Para atingir este objetivo, recomendamos que o professor se apoie no terceiro capítulo para que o aluno entenda o que está exposto no primeiro de forma contextualizada. A citada sugestão de apoio deve-se ao fato de que os conteúdos abordados em cifras de Hill, sequência de Fibonacci e cadeia de Markov podem contribuir para uma aprendizagem mais significativa.

Cifras de Hill aborda desde as simples permutações com as letras do alfabeto, passando por números primos e divisibilidade. Descrevemos ainda a sequência de Fibonacci, como os seus termos se manifestam nas mais variadas situações, na natureza, em matrizes e nos determinantes, por fim, mostramos que muitas aplicações de matrizes podem ser obtidas e estudadas por meio das cadeias de Markov, em seus aspectos probabilísticos, sendo estes conteúdos abordados a nível de ensino médio, oferecendo ao leitor uma nova visão das aplicações de matrizes no nosso dia a dia.

Assim ao planejar uma aula sobre matrizes, escolhe-se qual a habilidade a ser atingida, e uma estratégia para atingi-las, e a partir, de aplicações na realidade do aluno ou que desperte curiosidade no mesmo.

Uma sugestão para o professor que deseja ensinar multiplicação de matrizes, matriz inversa e matriz identidade é utilizar como estratégia ensinar Criptografia através da cifras de Hill.

Mas se o objetivo for ensinar potências de matrizes, fazer a introdução às regras de determinantes e relacionar as matrizes com sequências, um opção interessante que provocará curiosidade nos alunos será a sequência de Fibonacci.

Se as habilidades a ser alcançada for produto de matrizes, probabilidade aplicada, vetores probabilísticos, previsões, regras de produto matriciais, linha e coluna de uma matriz ou treinar os alunos a uma leitura diferenciada de matrizes aplicadas em várias áreas do conhecimento, um ótimo caminho é a cadeia de Markov.

Desta maneira o professor desmistificará o conceito de ciência inerte no ensino de matrizes, que necessita apenas de memorização de procedimentos e fórmulas. Também, colocará os alunos como sujeito ativo do processo de ensino e aprendizagem.

Apêndice A

Indução matemática

Indução matemática é uma poderosa ferramenta matemática usada para provar a verdade de um número infinito de proposições e Teoremas.

É comum vermos tal princípio enunciado em dois passos:
(i) A base: mostrar que a afirmação é válida para $n = 1$.
(ii) O passo indutivo: mostrar que se a afirmação é válida para $n = k$, então também é válido para $n = k + 1$.

O Princípio da Boa Ordem, que será assumido como um postulado, será importante na compreensão da demonstração do Teorema intitulado Princípio da Indução Matemática.

Postulado 1 (Princípio Boa Ordem) *Todo subconjunto não vazio de números naturais contém um elemento mínimo.*

Teorema A.0.1 (Princípio de Indução Matemática) *Seja B um subconjunto dos números naturais. Se B possui as seguintes propriedades:*
(i) $1 \in B$
(ii) $k + 1 \in B$ *sempre que* $k \in B$*, então B contém todos os naturais.*

Demonstração: Desejamos provar que se B é um subconjunto dos naturais, possuindo as propriedades (i) e (ii), então B necessariamente contém todos os naturais. Vamos provar por redução ao absurdo. Vamos supor que, mesmo possuindo as propriedades (i) e (ii) B não contém todos os naturais. Seja A o conjunto dos naturais não contidos em B. Pelo Princípio da Boa Ordenação, A possui um menor elemento e este é maior que 1, pois $1 \in B$. Seja a_0 este elemento. Note que $a_0 - 1 \in B$ e como B satisfaz (ii), então o sucessor de $a_0 - 1$, que é a_0 também deve pertencer a B. O que é um absurdo, logo A é vazio, o que prova o resultado. ■

Percebamos que esse método funciona provando que a afirmação é válida para um valor inicial, isto é, o caso base, e provando que o processo utilizado para ir de um

valor para o próximo valor também é verdadeiro, então podemos obter qualquer valor repetindo esse processo.

A situação a seguir ilustra bem como funciona o Princípio de Indução Matemática e é conhecida como "efeito dominó".

Imagine que tenhamos uma fila de dominós em pé, se pudermos garantir que:
(i) o primeiro dominó cairá.
(ii) sempre que um dominó cair, seu próximo vizinho também cairá.
Então podemos concluir que todos os dominós cairão.

Vejamos agora algumas aplicações do Princípio de Indução Matemática.

Exemplo A.0.1 *Este exemplo ilustra o primeiro registro da utilização do Princípio de Indução Matemática feita por Francesco Maurolycus em 1575. Trata-se da determinação de uma fórmula exata em função de* $n \geq 1$ *para a soma dos n primeiros números ímpares, ou seja, busca-se uma fórmula para*

$$S_n = 1 + 3 + 5 + 7 + ... + (2n - 1). \tag{A.1}$$

Solução: Vamos calcular a soma para alguns valores particulares de n, isto é, para $n = 1, 2, 3, 4, 5$ e 6. Assim, usando (A.1), obtemos:

$$\begin{gathered} S_1 = 1; \\ S_2 = 1 + 3 = 4; \\ S_3 = 1 + 3 + 5 = 9; \\ S_4 = 1 + 3 + 5 + 7 = 16; \\ S_5 = 1 + 3 + 5 + 7 + 9 = 25; \\ S_6 = 1 + 3 + 5 + 7 + 9 + 11 = 36. \end{gathered}$$

Analisando os casos particulares, tudo nos leva a crer que uma fórmula para tal soma é $S_n = n^2$, mas como não provamos tal fato, a chamaremos de conjectura. Utilizaremos o Teorema (A.0.1) para verificarmos se a nossa conjectura é ou não verdadeira.
Demonstração: Seja $p(n) : S_n = n^2$.
$p(1) = 1^2 = 1$, logo $p(1)$ é verdadeiro.

Suponhamos que $p(n)$ seja verdadeiro para algum $n = k \in \mathbb{N}$, isto é,

$$p(k) : S_k = k^2. \tag{A.2}$$

Queremos mostrar que $p(k + 1)$ também é verdadeiro, ou seja,

$$p(k+1) : S_{k+1} = (k+1)^2.$$

Com efeito, somando o termo seguinte $(2k+1)$ em ambos os membros de (A.2), obtemos

$$S_k + (2k+1) = k^2 + (2k+1). \tag{A.3}$$

Como $S_k + (2k+1)$ é S_{k+1}, então $(A.3)$ torna-se

$$S_{k+1} = k^2 + 2k + 1 \Rightarrow S_{k+1} = (k+1)^2,$$

que é o $p(k+1)$. Assim, $p(k+1)$ é verdade, e portanto, pelo Teorema $(A.0.1)$, $p(n)$ é verdade para todo $n \in \mathbb{N}$. ■

Conta-se que certo dia, com a intenção de manter a turma em silêncio, um professor pediu aos alunos que somassem os números naturais de 1 a 100, ou seja, $(1 + 2 + 3 + ... + 100)$ e, assim que terminassem, colocassem a solução sobre sua mesa. Quase que imediatamente, um garoto de 10 anos chamado Carl Friedrich Gauss colocou sobre a mesa do professor a resposta encontrada. Ele olhou para o menino com pouco caso, enquanto os demais alunos trabalhavam arduamente. Quando conferiu os resultados, o professor verificou que a única resposta correta era a de Gauss, 5.050, mas sem fazê-la acompanhar de nenhum cálculo. Vejamos no exemplo abaixo a solução da questão resolvida por Gauss, mas estendida para um $n \in \mathbb{N}$.

Exemplo A.0.2 *Determine uma fórmula para a soma dos n primeiros números naturais. Seja*

$$S_n = 1 + 2 + 3 + ... + n.$$

Demonstração: Seja S_n a soma de tais números, isto é,

$$S_n = 1 + 2 + 3 + ... + n. \tag{A.4}$$

Notemos inicialmente que S_n pode ser escrita também como

$$S_n = n + (n-1) + ... + 3 + 2 + 1, \tag{A.5}$$

Somando $(A.4)$ com $(A.5)$ obtemos:

$$2S_n = (n+1) + (n+1) + ... + (n+1). \tag{A.6}$$

e assim, a equação $(A.6)$ reduz-se em

$$S_n = \frac{n(n+1)}{2} \tag{A.7}$$

Verificaremos, por indução sobre n, a validade da fórmula de $(A.7)$. Seja $p(n) : S_n = \frac{n(n+1)}{2}$. Notemos que

$$p(1) = \frac{1(1+1)}{2} = 1,$$

logo $p(1)$ é verdadeiro.

Suponhamos que $p(k)$ seja verdadeiro para algum $n = k \in \mathbb{N}$, isto é,

$$p(k) : S_k = \frac{k(k+1)}{2} \tag{A.8}$$

Queremos mostrar que $p(k+1)$ também é verdadeiro, ou seja,

$$P(k+1) : S_{k+1} = \frac{(k+1)(k+2)}{2}.$$

Com efeito, somando $(k+1)$ a ambos os membros de $(A.8)$, obtemos:

$$S_k + (k+1) = (k+1) + \frac{k(k+1)}{2}. \tag{A.9}$$

Como $S_k + (k+1)$ é S_{k+1}, então $(A.9)$ torna-se

$$S_{k+1} = \frac{k(k+1) + 2(k+1)}{2}.$$

Colocando $(k+1)$ em evidência, temos

$$S_{k+1} = \frac{(k+1)(k+2)}{2},$$

que é o $p(k+1)$. Logo $p(k+1)$ é verdadeiro, e portanto, pelo Teorema $(A.0.1)$ temos que $(A.7)$ é válida para todo $n \in \mathbb{N}$. ■

Referências Bibliográficas

Anton, H. e Busby, R. C. (2006). *Álgebra linear contemporânea*. Bookman Editora.

Anton, H., Rorres, C., e Doering, C. I. (2001). *Álgebra linear com aplicações*, volume 8. Bookman.

Boldrini, J. L., Costa, S. I., Figueredo, V., e Wetzler, H. G. (1980). *Álgebra linear*. Harper & Row.

Borges, F. P. (2015). Sequência de fibonacci e algumas aplicações. Dissertação de Mestrado, Universidade Federal de Mato Grosso.

Contador, P. R. M. (2012). *Matemática, uma breve história*. Livraria da Física, São Paulo, 4 edição.

Cruz, L., Chueiri, V., e Gonçalves, E. (2012). Introdução ao estudo da álgebra linear. *São Paulo*.

Eves, H. (2011). *Introdução à história da Matemática; Tradução Hygino H. Domingues*. Editora da Unicamp, Campinas, 5 edição.

Ferri, M. G. (1983). *Botânica: morfologia externa das plantas(organografia)*. Nobel, São Paulo, 15 edição.

Grigoletti, P. S. (2010). Cadeias de markov. *Escola de Informática da Universidade Catóolica de Pelotas, Pelotas*.

Hazzan, S. (2013). *Fundamentos de matemática elementar, 5: combinatória, probabilidade*. Atual.

Hefez, A. (2011). *Elementos de aritmética*. SBM, Rio de Janeiro.

Hefez, A. e Fernandez, C. d. S. (2012). Introdução à álgebra linear. *Coleção PROFMAT, SBM*.

Iezzi, G. e Hazzan, S. (1977). *Fundamentos de Matemática Elementar*, volume 4. Atual, São Paulo, 2 edição.

Lima, E. L. (2009). *Álgebra linear*. Matematica universitaria. IMPA, Rio de Janeiro.

Médio, P. C. N. D. E. (2006). Ciências da natureza, matemática e suas tecnologias. *Brasília: MEC/SEB*.

Mol, R. S. (2013). *Introdução à história da Matemática*. CAED-UFMG, Belo Horizonte.

Morgado, A. C. d. O., de Carvalho, J. P., Carvalho, P. P., e Fernandez, P. (1991). Análise combinatória e probabilidade. *Sociedade Brasileira de Matemática, Rio de Janeiro*.

Santos, J. L. d. (2013). A arte de cifrar, criptografar, esconder e salvaguardar como fontes motivadoras para atividades de matemática básica. Dissertação de Mestrado, Universidade Federal da Bahia.

Taha, H. A., Marques, A. S., e Scarpel, R. A. (2008). *Pesquisa operacional*. Pearson Education do Brasil.

Wikipédia (2017). Variável aleatória — wikipédia, a enciclopédia livre. [Online; accessed 14-março-2017].

Zahn, M. (2011). *Sequência de Fibonacci e o Número de Ouro*. Editora Ciência Moderna Ltda., Rio de Janeiro.