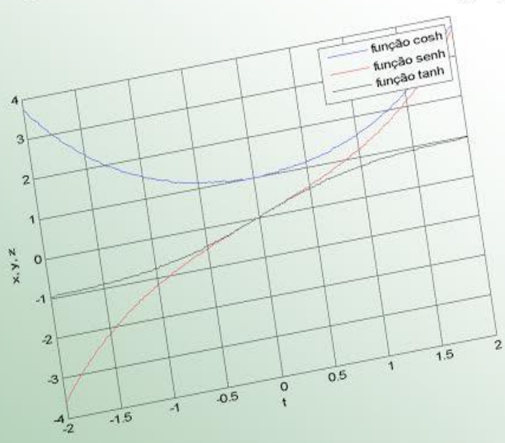
coleção Euler

# álgebra linear com OCTAVE

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & c_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & c_m \end{pmatrix}$$

$$f(\lambda) = |A - \lambda I| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}$$





**Marcelo H. Stoppa**

**Romes A. Borges**

# álgebra linear
# com OCTAVE

# álgebra linear com OCTAVE

M. H. Stoppa
R. A. Borges

coleção Euler

# APRESENTAÇÃO

**A ÁLGEBRA LINEAR** é considerada essencial ao tratamento de problemas ligados às mais variadas ciências como Matemática, Física, Engenharias, dentre outras. Com o objetivo de conceber enfoque especial e diferenciado para o estudo da Álgebra Linear, este livro contempla também, em todos os capítulos, discussões acerca de aplicações computacionais, tendo como ferramenta o OCTAVE, que é um software livre e de fácil utilização. Relacionando conteúdos essenciais de Álgebra Linear por meio de uma abordagem computacional, ele facilita a compreensão dos aspectos teóricos abordados em cada seção. Assim, este livro se caracteriza como uma importante ferramenta no tratamento de problemas de Álgebra Linear em estudos ligados às diversas áreas do conhecimento.

A adoção do software OCTAVE se dá, entre outros fatores, pela versatilidade de sua aplicação, facilidade de manipulação da álgebra matricial, tamanho do arquivo de instalação e requisitos computacionais mínimos necessários, além, é claro, pelo fato de ser ele um software livre.

No primeiro capítulo, que traz um tratamento especial ao estudo de matrizes, são abordados os principais tipos de matrizes, bem como as definições e operações básicas envolvendo este conteúdo. O estudo prossegue no segundo capítulo com a análise de sistemas de equações lineares, em que são abordados conceitos fundamentais como o de determinantes, inversa de uma matriz, dentre outros temas de importância para se compreender outros tópicos que serão tratados neste volume.

O estudo segue nos capítulos subsequentes abordando tópicos fundamentais de Álgebra Linear sempre com a preocupação de trabalhar a teoria utilizando recursos de computação. Assim, no terceiro capítulo, são discutidas as chamadas transformações no plano, em que são estudados alguns tipos de problemas envolvendo as formas em que estas transformações podem ocorrer. No capítulo seguinte, é apresentado um tratamento detalhado dos espaços e subespaços vetoriais, enunciando definições, teoremas e principais propriedades do conteúdo. O quinto capítulo aborda os fundamentos das transformações lineares; e o sexto exibe uma introdução ao estudo de autovalores e autovetores com seus principais fundamentos e aplicações.

Servindo como texto de referência, finalmente, o sétimo capítulo é dedicado ao OCTAVE; nele são mostrados os procedimentos básicos de utilização desse software, tais como seus comandos principais, entradas de matrizes, vetores e operadores.

Como este último capítulo servirá como fonte de consulta ao leitor durante o acompanhamento de todo o conteúdo do livro, recomenda-se ao estudante que ele seja lido previamente, de modo a facilitar a compreensão dos exemplos computacionais apresentados no decorrer do livro.

Para facilitar a leitura e a identificação dos diversos itens do texto, são utilizados recursos gráficos, ressaltando e diferenciando as definições ou propriedades de cada um: **Definição 1.1:** ; exemplos: **Exemplo 1.1:** ; exemplos resolvidos: **ER 1.1:** ; teoremas: **Teorema 1.1:** ; exemplos apresentados com a ajuda do OCTAVE: 💻 OCTAVE 1.1. No caso dos exemplos resolvidos, tanto as soluções

**Solução**: quanto as provas dos teoremas **Prova**: também são diretamente delimitadas, permitindo a fácil identificação do seu início e do seu fim.

Os exemplos resolvidos (indicados por **ER** no texto) são acompanhados, na maioria das vezes, da resolução do mesmo problema por meio de recursos do OCTAVE, e isto é indicado pelo ícone 💻, seguido da palavra OCTAVE. Neste caso, a fonte dentro da caixa é alterada para cronos pro, e todos os comandos que devem ser inseridos no OCTAVE são exibidos na fonte `courier new`. Este recurso de alteração de fonte também é utilizado para facilitar a visualização e contextualização do ambiente OCTAVE.

Para finalizar, deixamos um agradecimento especial a todos os alunos, professores e demais colaboradores, que enviaram sugestões, críticas e que, de alguma maneira, estimularam o desenvolvimento deste livro.

M. H. Stoppa

R. A. Borges

# SUmÁRiO

# 1. mATRiZES

**É mUItO COmUm** em várias áreas, como física, química, biologia, estatística, organizar um determinado conjunto de números em uma tabela retangular. Na realidade, no cotidiano é conveniente e até necessário usar conjuntos de números arranjados em linhas e colunas, de modo que suas posições sejam guardadas para uma posterior comparação e para uma variedade de outras razões. Neste capítulo, será apresentada a teoria de matrizes, exibindo definições para vários tipos de matrizes e suas principais propriedades.

## 1.1 DEFINIÇÕES

Considere uma empresa que fabrica três modelos de churrasqueira: um modelo elétrico, um modelo padrão, e um modelo a gás. Se a empresa deseja comparar as unidades da linha de produção e o trabalho envolvido em um mês de produção de cada um desses modelos, uma *tabela* pode ser usada para apresentar os dados:

$$
\begin{array}{l} \\ \\ \text{unidades de material} \\ \text{unidades de trabalho} \end{array}
\begin{array}{ccc} \text{modelo} & \text{modelo} & \text{modelo} \\ \text{elétrico} & \text{padrão} & \text{a gás} \\ \left( \begin{array}{c} 30 \\ 5 \end{array} \right. & \begin{array}{c} 26 \\ 7 \end{array} & \left. \begin{array}{c} 21 \\ 4 \end{array} \right) \end{array}
$$

Os números acima servem apenas para ilustrar, de modo muito simplificado, a aplicação de uma tabela de números reais. As unidades de material (que representam a quantidade de churrasqueiras produzidas) para os três modelos estão na primeira linha da tabela, as unidades de trabalho (que representam a quantidade de postos de trabalho utilizados) na segunda linha, e as unidades de produção para cada modelo (material e trabalho) nas colunas da tabela. Se o padrão no qual as unidades estão sendo guardadas está claramente definido, esta tabela pode ser representada simplesmente como:

$$
\begin{pmatrix} 30 & 26 & 21 \\ 5 & 7 & 4 \end{pmatrix} \tag{1.1}
$$

Outro exemplo do uso de tabelas retangulares de números reais é o que pode ser utilizado por um técnico de basquete que deseja manter um registro do número de participações em jogos e do desempenho de pontuação de seus três principais jogadores. Considere a seguinte tabela:

$$
\begin{array}{l} \\ \text{Jogador A} \\ \text{Jogador B} \\ \text{Jogador C} \end{array}
\begin{array}{ccc} \text{jogos} & \text{pontos} & \text{lances livres} \\ \left( \begin{array}{c} 18 \\ 15 \\ 12 \end{array} \right. & \begin{array}{c} 123 \\ 95 \\ 78 \end{array} & \left. \begin{array}{c} 52 \\ 37 \\ 62 \end{array} \right) \end{array}
$$

As linhas guardam os dados referentes a cada um dos três jogadores, e as colunas guardam da esquerda para a direita, respectivamente, o número de jogos participados, o número total de pontos e os pontos marcados por arremessos de lances livres, de cada jogador.

Uma vez que a representação das linhas e colunas está determinada, é possível guardar apenas os dados, ou simplesmente,

$$\begin{pmatrix} 18 & 123 & 52 \\ 15 & 95 & 37 \\ 12 & 78 & 62 \end{pmatrix} \tag{1.2}$$

Após esta breve introdução, podemos definir o que é uma matriz:

**Definição 1.1: Matrizes** são tabelas retangulares de elementos $a_{ij}$ dispostos da seguinte maneira:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{ou} \quad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{1.3}$$

onde cada elemento $a_{ij}$ tem dois índices: o **índice linha** e o **índice coluna,** respectivamente representados por i e j, que indicam a sua posição na tabela. Uma matriz é descrita entre parênteses ou colchetes, sem distinção.

Os elementos $a_{i1}$, $a_{i2}$, ..., $a_{in}$ são os elementos da i-ésima linha.

Os elementos $a_{1j}$, $a_{2j}$, ..., $a_{mj}$ são os elementos da j-ésima coluna.

O elemento $a_{ij}$ é o elemento que está na i-ésima linha e j-ésima coluna.

$$\left(\begin{array}{ccc:c:cc} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \hdashline a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \hdashline \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{m_j} & \cdots & a_{mn} \end{array}\right)$$

**Exemplo 1.1:** O elemento $a_{21}$ da matriz (1.1) é igual a 5; isto é, $a_{21}$ é o elemento que está na 2ª linha e 1ª coluna.

**Definição 1.2:** A matriz de m linhas e n colunas é chamada matriz de **ordem** m por n.

Assim, a ordem da matriz indica o "tamanho" da matriz, ou seja, quantas linhas e quantas colunas a matriz possui. A ordem de uma matriz, por exemplo, 3 por 2 é comumente indicada por 3×2.

**Exemplo 1.2:** A matriz (1.1) é de ordem 2 por 3 (2×3), enquanto a matriz (1.2) é de ordem 3 por 3 (3×3).

**Definição 1.3:** Quando o número de linhas é igual ao número de colunas, a matriz é chamada **quadrada**.

Uma matriz quadrada de ordem n por n é chamada simplesmente de ordem n. A matriz (1.2) é um exemplo de matriz quadrada de ordem 3.

**Definição 1.4:** Uma matriz é **real**, se cada um dos elementos da matriz é um número real.

As matrizes serão denotadas simbolicamente por letras maiúsculas A, B, C,..., ou por $(a_{ij})$, $(b_{ij})$, $(c_{ij})$,..., onde $a_{ij}$, $b_{ij}$, $c_{ij}$,..., representam, respectivamente, os elementos gerais das matrizes.

**ER 1.1:** Construa uma matriz quadrada $(a_{ij})$ de ordem 3, onde $a_{ij} = 4i - j^3$.

**Solução:** Se $a_{ij} = 4i - j^3$, então

$$a_{11} = 4(1) - (1)^3 = 3, \quad a_{12} = 4(1) - (2)^3 = -4, \quad a_{13} = 4(1) - (3)^3 = -23,$$

$$a_{21} = 4(2) - (1)^3 = 7, \quad a_{22} = 4(2) - (2)^3 = 0, \quad a_{23} = 4(2) - (3)^3 = -19,$$

$$a_{31} = 4(3) - (1)^3 = 11, \quad a_{32} = 4(3) - (2)^3 = 4, \quad a_{33} = 4(3) - (3)^3 = -15.$$

Portanto, a matriz desejada é

$$\begin{pmatrix} 3 & -4 & -23 \\ 7 & 0 & -19 \\ 11 & 4 & -15 \end{pmatrix}.$$

**Definição 1.5:** Duas matrizes $(a_{ij})$ e $(b_{ij})$ são **iguais** se e somente se elas têm mesma ordem, e $a_{ij} = b_{ij}$ para todos os pares (i,j).

Podemos entender a Definição 1.5 da seguinte maneira: duas matrizes são iguais se TODOS os elementos que ocupam posições correspondentes nas duas matrizes são iguais.

**OCTAVE 1.1**

Para definir a matriz $a_{ij} = 4i-j^3$ devemos utilizar a seguinte sequência de comandos:

```
> for ii = 1:3
>     for jj = 1:3
>         A(ii,jj) = 4*ii-jj^3;
>     end
> end
> A
> A =
     3    -4    -23
     7     0    -19
     11    4    -15
```

Aqui são utilizadas duas rotinas do tipo `for` para gerar os elementos das linhas (`ii`) e os elementos das colunas (`jj`). A matriz final fica armazenada na variável `A`. São utilizadas as variáveis `ii` e `jj` para não haver confusão com as unidades imaginárias `i` e `j` que são nativas do OCTAVE.

**ER 1.2:** Determine se cada par de matrizes são iguais ou não:

a) $\begin{pmatrix} 41 & 13 & 0.2 \\ -82 & 31 & -5 \end{pmatrix}$ e $\begin{pmatrix} 1.2 & 0.22 \\ -13 & 1.5 \end{pmatrix}$; b) $\begin{pmatrix} 55 & 2.1 \\ 11 & 48 \end{pmatrix}$ e $\begin{pmatrix} 55 & 2.1 \\ 11 & 48 \end{pmatrix}$;

c) $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ e $\begin{pmatrix} 5x \\ 5y \\ 5z \end{pmatrix}$; d) $\begin{pmatrix} 7 & -2.5 \\ 11 & 0 \\ 3.2 & 0 \end{pmatrix}$ e $\begin{pmatrix} 0 & -2.5 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$.

**Solução:** As matrizes em (a) não podem ser iguais, uma vez que elas não possuem a mesma ordem. As matrizes em (b) são obviamente iguais. As matrizes em (c) são iguais, se e somente se, x = y = z = 0. Embora as matrizes em (d) tenham a mesma ordem, elas não são iguais, uma vez que nem todos os elementos correspondentes são iguais (repare a primeira coluna de cada matriz).

Considere a matriz (1.2), que representa o desempenho de pontuação de três jogadores de basquete numa temporada. Suponha que a matriz que representa o desempenho das pontuações destes jogadores numa temporada seguinte de jogos seja

$$\begin{pmatrix} 17 & 141 & 91 \\ 17 & 82 & 30 \\ 16 & 102 & 59 \end{pmatrix}.$$

A matriz que representa a pontuação combinada de cada um dos três jogadores durante as duas temporadas pode ser obtida pela adição dos valores que ocupam posições correspondentes das duas matrizes:

$$\begin{pmatrix} 18 & 123 & 52 \\ 15 & 95 & 37 \\ 12 & 78 & 62 \end{pmatrix} + \begin{pmatrix} 17 & 141 & 91 \\ 17 & 82 & 30 \\ 16 & 102 & 59 \end{pmatrix} = \begin{pmatrix} 35 & 264 & 143 \\ 32 & 177 & 67 \\ 28 & 180 & 121 \end{pmatrix}.$$

Isto significa que, em duas temporadas, os jogadores A, B e C participaram de 35, 32 e 28 jogos, fizeram 264, 177 e 180 pontos e acertaram 143, 67 e 121 lances livres, respectivamente.

A soma de duas matrizes $(a_{ij})$ e $(b_{ij})$ é definida se, e somente se, as matrizes são de mesma ordem.

**Definição 1.6:** Suponha que $(a_{ij})$ e $(b_{ij})$ são matrizes de mesma ordem, então a **soma** $(a_{ij}) + (b_{ij})$ é definida como uma 3ª matriz $(c_{ij})$ de mesma ordem, onde cada $c_{ij}$ satisfaz a condição $c_{ij} = a_{ij} + b_{ij}$, para todos os pares (i,j).

**Propriedade 1.7:** Considere quaisquer duas matrizes de ordem m por n: $A = (a_{ij})$ e $B = (b_{ij})$. Como a adição de números reais é comutativa, então $a_{ij} + b_{ij} = b_{ij} + a_{ij}$ para quaisquer pares (i,j) e

$$A + B = B + A. \tag{1.4}$$

**Propriedade 1.8:** Considere quaisquer três matrizes de ordem m por n: $A = (a_{ij})$, $B = (b_{ij})$ e $C = (c_{ij})$. Como a adição de números reais é associativa, então para quaisquer pares (i,j), em elementos correspondentes, têm-se $(a_{ij} + b_{ij}) + c_{ij} = a_{ij} + (b_{ij} + c_{ij})$ e

$$(A + B) + C = A + (B + C). \tag{1.5}$$

**Assim, a *adição de matrizes reais é comutativa e associativa*.**

**Definição 1.9:** Uma **matriz nula** ou **matriz zero**, denotada por **0**, é uma matriz onde todos os seus elementos são zero. Para cada matriz A de ordem m por n existe uma matriz nula de ordem m por n tal que A + **0** = **0** + A = A. Esta matriz nula de ordem m por n é o ***elemento neutro aditivo*** para o conjunto de todas as matrizes de ordem m por n.

**ER 1.3:** Encontre a soma das matrizes A e B apresentadas a seguir:

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

**Solução:**

$$A + B = \begin{pmatrix} 1+0 & 2+(-4) & 4+1 \\ 1+(-1) & -5+1 & -1+8 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 5 \\ 0 & -4 & 7 \end{pmatrix}.$$

**OCTAVE 1.3**

Primeiramente é necessário definir as matrizes A e B, para depois executar a soma de maneira trivial:

```
> A = [1,2,4;1,-5,-1];
> B = [0 -4 1;-1 1 8];
> A + B
> ans =
        1    -2     5
        0    -4     7
```

Note que as matrizes A e B foram definidas, a primeira com os elementos das linhas separadas por vírgulas (,) e a segunda com os elementos das linhas separados por espaços. O resultado é o mesmo. O ponto e vírgula (;) ao final da linha de comando impede a exibição de A e B. A variável auxiliar `ans` guarda o resultado da soma A+B.

Considere novamente a matriz (1.1), que representa as unidades de material e o trabalho envolvido em um mês de produção de três modelos de churrasqueira. Suponha que a fábrica deseja dobrar sua produção de cada modelo. Então a matriz

$$\begin{pmatrix} 60 & 52 & 42 \\ 10 & 14 & 8 \end{pmatrix}$$

representa as unidades de material e trabalho envolvidas num simples mês de produção dos três modelos de churrasqueira. É conveniente representar a duplicação das entradas na matriz (1.1) como o produto da matriz e o número real 2, ou seja,

$$2\begin{pmatrix} 30 & 26 & 21 \\ 5 & 7 & 4 \end{pmatrix} = \begin{pmatrix} 60 & 52 & 42 \\ 10 & 14 & 8 \end{pmatrix}.$$

Note que

$$2\begin{pmatrix} 30 & 26 & 21 \\ 5 & & \end{pmatrix} = \begin{pmatrix} 30 & 26 & 21 \\ & & \end{pmatrix} + \begin{pmatrix} 30 & 26 & 21 \\ & & \end{pmatrix} = \begin{pmatrix} 60 & 52 & 42 \\ & & \end{pmatrix}.$$

**Definição 1.10:** O produto de um número real (escalar) k e uma matriz $(a_{ij})$, denotado por $k(a_{ij})$ ou $(a_{ij})k$ é chamado **múltiplo escalar** da matriz $(a_{ij})$ por k. O múltiplo escalar $k(a_{ij})$ é definido como uma matriz na qual os elementos são produtos de k pelos elementos correspondentes de $(a_{ij})$, e como a multiplicação de números reais é comutativa,

$$k(a_{ij}) = (a_{ij})k = (ka_{ij}) \tag{1.6}$$

Note que, para qualquer número real k e qualquer matriz real $A = (a_{ij})$, as matrizes A e kA são de mesma ordem. Em particular, se $k = -1$, então

$$A + (-1)A = (a_{ij}) + (-a_{ij}) = 0.$$

**Definição 1.11:** A matriz (-1)A é chamada **inversa aditiva** de A.

**Definição 1.12:** Se B e A são quaisquer duas matrizes de mesma ordem, a **diferença** B – A é definida pela relação

$$B - A = B + (-1)A. \tag{1.7}$$

Com relação aos coeficientes escalares, qualquer soma algébrica de múltiplos escalares de matrizes de mesma ordem satisfaz determinadas propriedades.

Por exemplo, se k e q são escalares e A e B são matrizes de mesma ordem, então

$$kA + qA = (k + q)A, \tag{1.8}$$

$$kqA = k(qA) = q(kA) = (kq)A, \tag{1.9}$$

$$k(A + B) = kA + kB. \tag{1.10}$$

Além disso, se $kA = \mathbf{0}$, então ou $k = 0$ ou A é uma matriz nula.

**ER 1.4:** Encontre 5A –3B onde $A = \begin{pmatrix} 2 & -1 \\ -5 & 4 \end{pmatrix}$ e $B = \begin{pmatrix} -3 & 5 \\ 1 & -2 \end{pmatrix}$.

**Solução:**

$$5A - 3B = 5\begin{pmatrix} 2 & -1 \\ -5 & 4 \end{pmatrix} - 3\begin{pmatrix} -3 & 5 \\ 1 & -2 \end{pmatrix} = 5\begin{pmatrix} 2 & -1 \\ -5 & 4 \end{pmatrix} + (-3)\begin{pmatrix} -3 & 5 \\ 1 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} 5(2) & 5(-1) \\ 5(-5) & 5(4) \end{pmatrix} + \begin{pmatrix} (-3)(-3) & (-3)(5) \\ (-3)(1) & (-3)(-2) \end{pmatrix}$$

$$= \begin{pmatrix} 10 & -5 \\ -25 & 20 \end{pmatrix} + \begin{pmatrix} 9 & -15 \\ -3 & 6 \end{pmatrix} = \begin{pmatrix} 19 & -20 \\ -28 & 26 \end{pmatrix}$$

**OCTAVE 1.4**

Definimos as matrizes A e B e executamos diretamente a operação 5A–3B:

```
> A = [2,-1;-5,4];
> B = [-3,5;1,-2];
> 5*A - 3*B
> ans =
      19   -20
     -28    26
```

## 1.2 MULTIPLICAÇÃO DE MATRIzES

Como preparação para a definição da multiplicação de matrizes, vamos considerar um sistema de equações lineares, com 3 equações e 3 variáveis da forma:

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases} \tag{1.11}$$

Este sistema pode ser representado por uma equação matricial simples:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}. \tag{1.12}$$

Os coeficientes de x, y e z podem ser obtidos de (1.11) ou (1.12).

**Definição 1.13:** Considerando o sistema (1.11) e sua forma matricial equivalente (1.12), este sistema pode ser escrito como AX = B onde A é a **matriz dos coeficientes**, X é a **matriz das variáveis** e B é a **matriz dos termos independentes**, dados respectivamente por

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

Os coeficientes de cada variável estão posicionados numa coluna, e os coeficientes das variáveis de cada equação estão numa linha.

O uso da matriz dos coeficientes como um operador em (1.12) exige a introdução da **multiplicação de matrizes**. Note que, do lado esquerdo de (1.12) o elemento $a_{11}x + a_{12}y + a_{13}z$ pode ser obtido das matrizes

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad e \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

somando-se os produtos dos elementos da linha 1 da matriz dos coeficientes e os elementos correspondentes da matriz das variáveis, de forma ordenada, ou seja,

$$(a_{11})(x) + (a_{12})(y) + (a_{13})(z).$$

Analogamente o elemento $(a_{21})(x) + (a_{22})(y) + (a_{23})(z)$ pode ser obtido somando-se os produtos dos elementos da linha 2 da matriz dos coeficientes e os elementos correspondentes da matriz das variáveis, ordenadamente.

De maneira semelhante, usando os elementos da linha 3 da matriz dos coeficientes, obtém-se $(a_{31})(x) + (a_{32})(y) + (a_{33})(z)$.

Agora, com esse raciocínio, pode-se definir o conceito de multiplicação de matrizes.

**Definição 1.14:** O **produto** AB de duas matrizes A e B é definido como uma nova matriz C, tal que o elemento $c_{ij}$ de C é obtido da soma de produtos de elementos da i-ésima linha de A e os elementos correspondentes na j-ésima coluna de B, ordenadamente. Isto exige que o número de colunas de A seja igual ao número de linhas de B. Se $A = (a_{ij})$ é uma matriz de ordem m por n e $B = (b_{ij})$ é uma matriz de ordem n por r, então $C = (c_{ij})$ é uma matriz de ordem m por r onde

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + ... + a_{in}b_{nj}. \tag{1.13}$$

Pode-se usar a notação de somatório e escrever (1.13) como

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}. \tag{1.14}$$

**Definição 1.15:** As matrizes A e B são ditas **conformáveis** para o produto AB, se e somente se, o número de colunas de A for igual ao número de linhas de B.

Consequentemente, pela Definição 1.14, o produto AB existe e duas matrizes podem ser multiplicadas apenas quando elas são conformáveis. No produto AB, costuma-se dizer que B está **pré-multiplicada** por A, e A está sendo **pós-multiplicada** por B. Mesmo que o produto AB exista, o produto BA pode não existir, uma vez que as matrizes A e B podem não ser conformáveis para o produto BA. Isto ilustra uma propriedade importante da multiplicação de matrizes, normalmente chamada propriedade não comutativa.

**ER 1.5:** Encontre os produtos AB e BA, se eles existirem, onde

$$A = \begin{pmatrix} 3 & 1 \\ 2 & -3 \end{pmatrix} \text{ e } B = \begin{pmatrix} 2 & -1 & 1 \\ 2 & 1 & -2 \end{pmatrix}.$$

**Solução:** As matrizes A e B são conformáveis para o produto AB, pois o número de colunas de A é igual ao número de linhas de B. Portanto, AB existe. Além disso, AB é de ordem 2 por 3, pois A é de ordem 2 por 2 e B é de ordem 2 por 3. Por definição, o elemento geral de AB é dado por $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j}$. Então

$$c_{11} = (3)(2) + (1)(2), \quad c_{12} = (3)(-1) + (1)(1), \quad c_{13} = (3)(1) + (1)(-2),$$
$$c_{21} = (2)(2) + (-3)(2), \; c_{22} = (2)(-1) + (-3)(1), \; c_{23} = (2)(1) + (-3)(-2).$$

Portanto,

$$AB = \begin{pmatrix} 8 & -2 & 1 \\ -2 & -5 & -4 \end{pmatrix}$$

E o produto BA não existe, pois as matrizes B e A não são conformáveis para o produto BA.

**OCTAVE 1.5**

Definimos as matrizes A e B e executamos os produtos:

```
> A = [3,1;2,-3];
> B = [2 -1 1;2 1 -2];
> A*B
> ans =
         8     -2      1
        -2     -5     -4
> B*A
error: operator *: nonconformant arguments (op1 is 2x3, op2 is 2x2)
error: evaluating binary operator '*' near line 16, column 2
>
```

Note a mensagem de erro produzida para o produto BA, informando que a ordem das matrizes não é compatível com a definição de multiplicação.

Se os elementos de qualquer linha ou coluna de uma matriz são considerados como componentes de um vetor, para cada par de valores (i,j), o elemento $c_{ij}$ de (1.13) é o escalar (algumas vezes chamado de ponto), produto da i-ésima linha de A pela j-ésima coluna de B.

**Definição 1.16:** Uma matriz com uma simples linha é também chamada de **matriz linha** ou **vetor linha**; uma matriz com uma simples coluna é chamada **matriz coluna** ou **vetor coluna**.

**ER 1.6:** Encontre as matrizes produto AB e BA do vetor linha A = (1 1 -1) e o vetor coluna

$$B = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}.$$

**Solução:** Como a matriz A é de ordem 1 por 3 e B é de ordem 3 por 1, as matrizes são conformáveis, independentemente da ordem considerada. Portanto, os produtos AB e BA existem:

$$AB = \left((1)(-2) + (1)(0) + (-1)(2)\right) = (-4);$$

$$BA = \begin{pmatrix} (-2)(1) & (-2)(1) & (-2)(-1) \\ (0)(1) & (0)(1) & (0)(-1) \\ (2)(1) & (2)(1) & (2)(-1) \end{pmatrix} = \begin{pmatrix} -2 & -2 & 2 \\ 0 & 0 & 0 \\ 2 & 2 & -2 \end{pmatrix}.$$

Note que o produto AB pode ser considerado uma matriz com um único elemento que representa o produto escalar de dois vetores cujas componentes são os elementos de A e B, respectivamente.

**OCTAVE 1.6**

Definimos os vetores linha A e coluna B e executamos os produtos normalmente:

```
> A = [1 1 -1];
> B = [-2;0;2];
> A*B
> ans = -4
> B*A
      -2     -2      2
       0      0      0
       2      2     -2
>
```

**ER 1.7:** Verifique que C(A + B) = CA + CB, onde

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix},\ B = \begin{pmatrix} 2 & -2 \\ 4 & 1 \end{pmatrix} \ \text{ e } \ C = \begin{pmatrix} 3 & 1 \\ 1 & -3 \\ 0 & 2 \end{pmatrix}.$$

**Solução:**

$$C(A+B) = \begin{pmatrix} 3 & 1 \\ 1 & -3 \\ 0 & 2 \end{pmatrix}\left[\begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} + \begin{pmatrix} 2 & -2 \\ 4 & 1 \end{pmatrix}\right] = \begin{pmatrix} 3 & 1 \\ 1 & -3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 6 & 1 \end{pmatrix} = \begin{pmatrix} 15 & -2 \\ -15 & -4 \\ 12 & 2 \end{pmatrix};$$

$$CA + CB = \begin{pmatrix} 3 & 1 \\ 1 & -3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} + \begin{pmatrix} 3 & 1 \\ 1 & -3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & -2 \\ 4 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 5 & 3 \\ -5 & 1 \\ 4 & 0 \end{pmatrix} + \begin{pmatrix} 10 & -5 \\ -10 & -5 \\ 8 & 2 \end{pmatrix} = \begin{pmatrix} 15 & -2 \\ -15 & -4 \\ 12 & 2 \end{pmatrix};$$

portanto, C(A+B) = CA + CB.

**OCTAVE 1.7**

Uma forma de verificar se a expressão C(A+B) = CA+CB é verdadeira, é mostrar que a diferença entre elas se anula, ou seja, que C(A+B) – (CA+CB) é uma matriz nula. Assim, definimos as matrizes A, B e C e calculamos esta expressão:

```
> A = [1 1;2 0];
> B = [2 -2;4 1];
> C = [3 1;1 -3;0 -2];
> C*(A+B) - (C*A+C*B)
> ans =
      0    0
      0    0
      0    0
>
```

Uma vez que a variável `ans` aponta uma matriz nula, está verificada a expressão.

**Teorema 1.1:** *A multiplicação de matrizes é distributiva com respeito à adição.*

**Prova:** Sejam A = $(a_{ij})$ e B = $(b_{ij})$ matrizes de ordem m por n, e seja C = $(c_{ij})$ uma matriz de ordem k por m. Então, A + B, CA, CB e C(A + B) existem. Os elementos da i-ésima linha de C são

$$c_{i1}, c_{12}, \ldots, c_{im},$$

e os elementos da j-ésima coluna de A + B são

$$a_{1j} + b_{1j}, a_{2j} + b_{2j}, \ldots, a_{mj} + b_{mj}.$$

Portanto, o ij-ésimo elemento de C(A + B) é

$$c_{i1}\left(a_{1j} + b_{1j}\right) + c_{i2}\left(a_{2j} + b_{2j}\right) + \ldots + c_{im}\left(a_{mj} + b_{mj}\right);$$

isto é,

$$\left(c_{i1}a_{1j} + c_{i2}a_{2j} + \ldots + c_{im}a_{mj}\right) + \left(c_{i1}b_{1j} + c_{i2}b_{2j} + \ldots + c_{im}b_{mj}\right),$$

a soma dos ij-ésimos elementos de CA e CB, respectivamente. Portanto,

$$C(A+B) = CA + CB. \tag{1.15}$$

A equação (1.15) representa a **propriedade distributiva à esquerda** da multiplicação de matrizes. A **propriedade distributiva à direita**

$$(A + B)C = AC + BC \tag{1.16}$$

também é válida desde que A + B, AC, BC e (A + B)C existam. Note que C(A + B) e (A + B)C geralmente não são iguais.

**ER 1.8:** Verifique que A(BC) = (AB)C onde

$$A = \begin{pmatrix} 2 & 2 \\ -1 & 5 \end{pmatrix},\ B = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 7 \end{pmatrix},\ e\ \ C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ -2 & -1 \end{pmatrix}.$$

**Solução:**

$$A(BC) = \begin{pmatrix} 2 & 2 \\ -1 & 5 \end{pmatrix}\left[\begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 7 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \\ -2 & -1 \end{pmatrix}\right] = \begin{pmatrix} 2 & 2 \\ -1 & 5 \end{pmatrix}\begin{pmatrix} 4 & 2 \\ -13 & -8 \end{pmatrix} =$$

$$= \begin{pmatrix} -18 & -12 \\ -69 & -42 \end{pmatrix}.$$

$$(AB)C = \left[\begin{pmatrix} 2 & 2 \\ -1 & 5 \end{pmatrix}\begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 7 \end{pmatrix}\right]\begin{pmatrix} 0 & 1 \\ 1 & 0 \\ -2 & -1 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 6 & 12 \\ -6 & 3 & 36 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} -18 & -12 \\ -69 & -42 \end{pmatrix};$$

Portanto, A(BC) = (AB)C.

**OCTAVE 1.8**

De forma análoga ao OCTAVE 1.8, definimos as matrizes A, B e C e calculamos a expressão A(BC) – (AB)C:

```
> A = [2 2;-1 5];
> B = [1 2 -1;-1 1 7];
> C = [0 1;1 0;-2 -1];
> A*(B*C) - (A*B)*C
> ans =
        0   0
        0   0
```

Novamente, `ans` aponta uma matriz nula e está verificada a expressão.

**Teorema 1.2:** *A multiplicação de matrizes é associativa.*

**Prova:** Sejam A = $(a_{ij})$, B = $(b_{ij})$ e C = $(c_{ij})$ matrizes de ordem k por m, m por n e n por p, respectivamente. Então, os produtos AB, BC, A(BC) e (AB)C existem. Os elementos da i-ésima linha de A são $a_{i1}$, $a_{i2}$, ..., $a_{im}$, e os elementos da j-ésima coluna de BC são

$$b_{11}c_{1j}+b_{12}c_{2j}+...+b_{1n}c_{nj}$$

$$b_{21}c_{1j}+b_{22}c_{2j}+...+b_{2n}c_{nj}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$b_{m1}c_{1j}+b_{m2}c_{2j}+...+b_{mn}c_{nj}.$$

Portanto, o ij-ésimo elemento de A(BC) é

$$a_{i1}\left(b_{11}c_{1j}+b_{12}c_{2j}+...+b_{1n}c_{nj}\right)+a_{i2}\left(b_{21}c_{1j}+b_{22}c_{2j}+...+b_{2n}c_{nj}\right)+...$$
$$...+a_{im}\left(b_{m1}c_{1j}+b_{m2}c_{2j}+...+b_{mn}c_{nj}\right);$$

isto é,

$$\left(a_{i1}b_{11}+a_{i2}b_{21}+...+a_{im}b_{m1}\right)c_{1j}+\left(a_{i1}b_{12}+a_{i2}b_{22}+...+a_{im}b_{m2}\right)c_{2j}+...$$
$$...+\left(a_{i1}b_{1n}+a_{i2}b_{2n}+...+a_{im}b_{mn}\right)c_{nj}$$

que representa o ij-ésimo elemento de (AB)C. Portanto,

$$A(BC)=(AB)C. \tag{1.17}$$

Considere as matrizes A = $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ e B = $\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$.

Note que AB = 0, mas nem A nem B são uma matriz nula. Na álgebra matricial, é possível ter **divisores nulos**. Existem elementos não nulos cujo produto é zero. Na álgebra de números reais, pode-se mostrar que divisores nulos não existem, de forma que, se a e b são números reais e ab = 0, então a = 0 ou b = 0.

## 1.3 MATRIZES DIAGONAIS

Um conjunto importante de matrizes é o daquelas denominadas **diagonais**. Estas matrizes têm um importante papel na resolução de diversos tipos de problemas, como, por exemplo, os problemas de autovalor, que serão estudados em mais detalhes no Capítulo 6.

**Definição 1.17:** Os elementos $a_{ij}$ onde i = j, de uma matriz $(a_{ij})$ são chamados **elementos diagonais** de $(a_{ii})$ e estão na chamada **diagonal principal**.

**Definição 1.18:** Uma matriz quadrada da forma

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{1.18}$$

é chamada **matriz diagonal**; isto é, a matriz diagonal é uma matriz quadrada $(a_{ij})$ onde $a_{ij} = 0$ se $i \neq j$ para todos os pares (i,j).

**Exemplo 1.3:** As matrizes abaixo são exemplos de matrizes diagonais e qualquer matriz nula de ordem n também é uma matriz diagonal:

$$\begin{pmatrix} 2 & 0 \\ 0 & -7 \end{pmatrix}, \begin{pmatrix} 7 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ e } \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}. \tag{1.19}$$

**Definição 1.19:** Se todos os $a_{ii}$ da matriz (1.18) são iguais, então a matriz diagonal é chamada **matriz escalar**.

A terceira matriz do Exemplo 1.3 é um exemplo de matriz escalar.

**ER 1.9:** Determine o efeito da pré-multiplicação e pós-multiplicação de qualquer matriz quadrada de ordem 2 por: (a) uma matriz diagonal conformável; (b) uma matriz escalar conformável.

**Solução:** Sejam $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ uma matriz quadrada qualquer de ordem 2, $D = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}$ uma matriz diagonal conformável, e $S = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ uma matriz escalar conformável.

a) Pela definição da multiplicação de matrizes, $DA = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} k_1 a & k_1 b \\ k_2 c & k_2 d \end{pmatrix}$; isto é, a pré-multiplicação da matriz A por D resulta numa matriz onde cada elemento da i-ésima linha é igual ao produto entre o elemento correspondente de A e o elemento diagonal na i-ésima linha de D. Analogamente, $AD = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} = \begin{pmatrix} k_1 a & k_2 b \\ k_1 c & k_2 d \end{pmatrix}$; isto é, a pós-multiplicação da matriz A por D resulta numa matriz onde cada elemento da i-ésima coluna é igual ao produto entre o elemento correspondente de A e o elemento diagonal na i-ésima coluna de D.

b) $SA = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ka & kb \\ kc & kd \end{pmatrix}$;

$AS = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} = \begin{pmatrix} ka & kb \\ kc & kd \end{pmatrix}$.

AS = AS = kA. Em geral, o produto entre quaisquer matrizes A e uma escalar conformável S com elementos diagonais k é equivalente ao múltiplo escalar kA.

**Definição 1.20:** Uma matriz escalar $(a_{ii})$, onde $a_{ii} = 1$ para todos os valores de i é chamada **matriz identidade** ou **matriz unidade**. Uma matriz identidade de ordem n será denotada por I e tem a propriedade de que, para toda matriz quadrada A de ordem n,

$$AI = IA = A \tag{1.20}$$

Assumimos que I é a única matriz com esta propriedade (1.20). Note que, se A é uma matriz de ordem m por n, a matriz identidade pré-multiplicativa é de ordem m por m, enquanto a matriz identidade pós-multiplicativa é de ordem n por n.

A matriz identidade é frequentemente denotada por $(\delta_{ij})$ onde o símbolo **delta de Kronecker** é definido por

$$\delta_{ij} = \begin{cases} 0, & (i \neq j) \\ 1, & (i = j) \end{cases} \tag{1.21}$$

**ER 1.10:** Mostre que AB = BA = I onde $A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 3 \end{pmatrix}$ e $B = \begin{pmatrix} 3 & 3 & -1 \\ 0 & 1 & 0 \\ -2 & -2 & 1 \end{pmatrix}$.

**Solução:**

$$AB = \begin{pmatrix} 3+0-2 & 3-1-2 & -1+0+1 \\ 0+0+0 & 0+1+0 & 0+0+0 \\ 6+0-6 & 6+0-6 & -2+0+3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$BA = \begin{pmatrix} 3+0-2 & -3+3+0 & -1+0+1 \\ 0+0+0 & 0+1+0 & 0+0+0 \\ -2+0+2 & 2-2+0 & -2+0+3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

portanto, AB = BA = I.

**OCTAVE 1.10**

Mais uma vez, basta apenas definir as matrizes A e B e efetuar os produtos AB e BA:

```
> A = [1 -1 1;0 1 0;2 0 3];
> B = [3 3 -1;0 1 0;-2 -2 1];
> A*B
> ans =
        1    0    0
        0    1    0
        0    0    1
> B*A
> ans =
        1    0    0
        0    1    0
        0    0    1
```

Diretamente, temos que AB = BA = I.

## 1.4 MATRIzES ESPECIAIS

Um tipo especial de matriz de grande utilidade é a matriz simétrica. As matrizes simétricas desempenham um papel muito importante em muitos ramos da matemática, bem como em outras ciências.

**Definição 1.21:** Uma matriz $(a_{ij})$ é chamada **matriz simétrica**, se, e somente se, $a_{ij} = a_{ji}$ para todos os pares (i,j).

**Exemplo 1.4:** A matriz

$$\begin{pmatrix} 5 & 3 & 0 \\ 3 & 4 & -6 \\ 0 & -6 & 5 \end{pmatrix}$$

é um exemplo de matriz simétrica de ordem 3.

**Definição 1.22:** Uma matriz $(a_{ij})$ é chamada **antissimétrica** se, e somente se, $a_{ij} = -a_{ji}$ para todos (i,j).

**Exemplo 1.5:** A matriz $\begin{pmatrix} 0 & -3 \\ 3 & 0 \end{pmatrix}$ é um exemplo de matriz antissimétrica de ordem 2.

Uma matriz deve ser necessariamente quadrada para que seja simétrica ou antissimétrica. Além disso, os elementos diagonais de uma matriz antissimétrica devem ser zero, uma vez que $a_{ii} = -a_{ii}$, se e somente se, $a_{ii} = 0$.

**ER 1.11:** Determine quais das seguintes matrizes são simétricas e quais são antissimétricas:

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 7 & -2 \\ 2 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0.5 \\ -0.5 & 0 \end{pmatrix}$$

$$E = \begin{pmatrix} 0 & 0 \\ 7 & 0 \end{pmatrix}, \quad F = (2.3), \quad G = \begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 5 \\ 1 & -5 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 3 \\ 2 & 1 \\ 1 & 0 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

**Solução:** As matrizes A, C, F, e J são simétricas, enquanto D, G, e J são antissimétricas. As matrizes B, E, e H não são nem simétricas nem antissimétricas. Note que a matriz nula J de ordem 2 pode ser considerada tanto simétrica quanto antissimétrica. As matrizes nulas de qualquer ordem n são as únicas que têm esta propriedade.

As matrizes simétricas e antissimétricas podem ser discutidas em termos de **transposição**, que é a operação de troca entre linhas e colunas de uma dada matriz.

**Definição 1.23:** Chama-se **transposta** de uma matriz A, matriz denotada por $A^T$, que é obtida pela transposição das linhas e colunas de A, ou seja, se $A = (a_{ij})$, então $A^T = (a_{ji})$.

Note que a transposta de um vetor coluna é um vetor linha. Em geral, a transposta de uma matriz de ordem m por n é uma matriz de ordem n por m. Veja a seguir a diferença entre matrizes simétrica e antissimétrica.

**Definição 1.24:** A é uma matriz **simétrica** se para todos os pares (i,j), $a_{ij} = a_{ji}$ e, então, $A^T = A$.

**Definição 1.25:** A é uma matriz **antissimétrica** se para todos os pares (i,j), $a_{ij} = -a_{ji}$ e, então, $A^T = -A$.

**OCTAVE 1.11**

A verificação de matrizes que são simétricas ou antissimétricas fica muito simples, utilizando-se a definição de matriz transposta e ainda o fato de que $A^T = A$ quando A é simétrica e $A^T = -A$ quando A é antissimétrica. Para tanto, considerando o ER 1.11, basta definir as matrizes e verificar (para M variando de A a J) se $M^T - M = 0$ ou $M^T + M = 0$, respectivamente, para M simétrica ou antissimétrica:

$$A=\begin{pmatrix}-1 & 0\\ 0 & 2\end{pmatrix},\quad B=\begin{pmatrix}7 & -2\\ 2 & 1\end{pmatrix},\quad C=\begin{pmatrix}1 & -2\\ -2 & 1\end{pmatrix},\quad D=\begin{pmatrix}0 & 0.5\\ -0.5 & 0\end{pmatrix}$$

$$E=\begin{pmatrix}0 & 0\\ 7 & 0\end{pmatrix},\quad F=\begin{pmatrix}2.3\end{pmatrix},\quad G=\begin{pmatrix}0 & 2 & -1\\ -2 & 0 & 5\\ 1 & -5 & 0\end{pmatrix},\quad H=\begin{pmatrix}1 & 3\\ 2 & 1\\ 1 & 0\end{pmatrix},\quad J=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}.$$

```
> A = [-1,0;0,2];
> B = [7,-2;2,1];
> C = [1,-2;-2,1];
> D = [0,0.5;-0.5,0];
> E = [0 0;7 0];
> F = [2.3];
> G = [0,2,-1;-2,0,5;1,-5,0];
> H = [1,3;2,1;1,0];
> J = [0,0;0,0];
> A' - A
> ans =
      0   0
```

```
        0   0
> A' + A
> ans =
       -2  0
        0  4
```

A operação de transposição no OCTAVE é realizada colocando um apóstrofo (') à direita da matriz a ser transposta. Observamos neste caso que, como A'–A = 0 e A'+A ≠ 0, onde 0 é uma matriz nula, A é simétrica. Continuamos o mesmo processo para as outras matrizes:

```
> B' - B
> ans =
        0    4
       -4    0
> B' + B
> ans =
       14  0
        0  2
```

Observamos que a matriz B não é simétrica nem antissimétrica. Continuando a avaliação com as outras matrizes, concluímos que C e F são simétricas, enquanto D e G são antissimétricas. A matriz J (nula) pode ser considerada simétrica e antissimétrica. As matrizes E e H não são simétricas nem antissimétricas.

Os seguintes teoremas são, em geral, verdadeiros para matrizes quaisquer, considerando que a adição e a multiplicação são aplicáveis para estas nos casos apontados.

**Teorema 1.3:** *A operação de transposição é **reflexiva**; isto é, $(A^T)^T = A$.*

**Prova:** Sejam $A = (a_{ij})$ e $A^T = (b_{ij})$. Então $b_{ij} = a_{ji}$ para todos os pares (i,j). A transposta de $A^T$ é $(b_{ji})$ e então $(b_{ji}) = (a_{ij})$. Portanto, $(A^T)^T = A$.

**Teorema 1.4:** *A transposta da soma (diferença) de duas matrizes é igual à soma (diferença) de suas transpostas; isto é,* $(A + B)^T = A^T + B^T$ *e* $(A - B)^T = A^T - B^T$*.*

**Prova:** Sejam $A = (a_{ij})$ e $B = (b_{ij})$ quaisquer duas matrizes de mesma ordem. Então, $A + B = (c_{ij})$ onde $c_{ij} = a_{ij} + b_{ij}$ para todos os pares (i,j);

$$A^T = (a_{ji}), \quad B^T = (b_{ji}),$$

$$(A+B)^T = (c_{ji}) = (a_{ji} + b_{ji}) = A^T + B^T.$$

De maneira similar, pode-se mostrar que $(A - B)^T = A^T - B^T$.

**Teorema 1.5:** *A transposta do produto de duas matrizes é igual ao produto de suas transpostas em ordem reversa; isto é,* $(AB)^T = B^TA^T$*.*

**Prova:** Sejam $A = (a_{ij})$ e $B = (b_{ij})$ matrizes de ordem k por m e m por n, respectivamente. Seja $AB = (c_{ij})$. Então, $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + ... + a_{im}b_{mj}$ é o elemento da j-ésima linha e i-ésima coluna de $(c_{ij})^T$; isto é, $(AB)^T$. Os elementos $b_{1j}, b_{2j}, ..., b_{mj}$ da j-ésima coluna de B são os elementos da j-ésima linha de $B^T$. Os elementos $a_{i1}, a_{i2}, ..., a_{im}$ da i-ésima linha de A são os elementos da i-ésima coluna de $A^T$. O elemento da j-ésima linha e i-ésima coluna de $B^TA^T$ é $b_{1j}a_{i1} + b_{2j}a_{i2} + ... + b_{mj}a_{im}$; isto é, $a_{i1}b_{1j} + a_{i2}b_{2j} + ... + a_{im}b_{mj}$. Portanto, $(AB)^T = B^TA^T$.

**Teorema 1.6:** *O produto de qualquer matriz e sua transposta é uma matriz simétrica; isto é:*

$$(AA^T)^T = AA^T.$$

**Prova:**

$$\begin{aligned} (AA^T)^T &= (A^T)^TA^T && \text{pelo Teorema 1.5} \\ &= AA^T && \text{pelo Teorema 1.3.} \end{aligned}$$

Portanto, $AA^T$ é uma matriz simétrica. Note que, se A é uma matriz de ordem m por n, então $AA^T$ é uma matriz simétrica de ordem m; $A^TA$ é uma matriz simétrica de ordem n.

Os seguintes teoremas são verdadeiros para matrizes quadradas.

**Teorema 1.7:** *A soma de qualquer matriz e sua transposta é uma matriz simétrica; isto é,* $\left(A + A^T\right)^T = A + A^T$.

**Prova:**

$\left(A + A^T\right)^T = A^T + \left(A^T\right)^T$ pelo Teorema 1.4

$= A^T + A$ pelo Teorema 1.3

$= A + A^T$ pois a adição de matrizes é comutativa.

Portanto, $A + A^T$ é uma matriz simétrica.

**Teorema 1.8:** *A diferença de qualquer matriz e sua transposta é uma matriz antissimétrica; isto é,* $(A - A^T)^T = -(A - A^T)$.

**Prova:**

$\left(A - A^T\right)^T = A^T - \left(A^T\right)^T$ pelo Teorema 1.4

$= A^T - A$ pelo Teorema 1.3

$= -(A - A^T)$.

Portanto, $A - A^T$ é uma matriz antissimétrica.

**Teorema 1.9:** *Se A e B são matrizes simétricas, então AB é matriz simétrica, se, e somente se, AB = BA.*

**Prova:** Sejam A e B matrizes simétricas de mesma ordem, tais que AB é uma matriz simétrica. Então,

$AB = (AB)^T$ pois AB é uma matriz simétrica

$= B^T A^T$ pelo Teorema 1.5

$= BA$ pois A e B são matrizes simétricas.

Portanto, AB = BA. Se AB = BA, onde A e B são matrizes simétricas, então

$$(AB)^T = (BA)^T$$

$= A^TB^T$ pelo Teorema 1.5

$= AB$, pois A e B são matrizes simétricas.

Portanto, AB é uma matriz simétrica.

**ER 1.12:** Verifique o Teorema 1.5 onde $A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 7 \\ 3 & 0 \end{pmatrix}$.

**Solução:**

$$AB = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 7 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 7 & 7 \\ -3 & 0 \end{pmatrix}; \quad (AB)^T = \begin{pmatrix} 7 & -3 \\ 7 & 0 \end{pmatrix};$$

$$B^TA^T = \begin{pmatrix} 1 & 3 \\ 7 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 7 & -3 \\ 7 & 0 \end{pmatrix}; \ (AB)^T = B^TA^T.$$

**OCTAVE 1.12**

Para verificar o Teorema 1.5, devemos avaliar se $(AB)^T = B^TA^T$, ou equivalentemente se $(AB)^T - B^TA^T = 0$. Para tanto, definimos as matrizes A e B, e calculamos o valor desta última expressão:

```
> A = [1 2;0 -1];
> B = [1 7;3 0];
> (A*B)'-B'*A'
> ans =
        0   0
        0   0
```

Assim, temos $(AB)^T - B^TA^T = 0$ (matriz nula) e, consequentemente, $(AB)^T = B^TA^T$, o que verifica o teorema para estas matrizes particulares.

## 1.5 MATRIzES COMPLEXAS

Aqui se trabalha principalmente com matrizes reais, entretanto, algumas matrizes especiais complexas, que serão úteis posteriormente, na prova de alguns teoremas sobre matrizes simétricas reais, são consideradas nesta seção.

**Definição 1.26:** Uma **matriz complexa** é uma matriz cujos elementos são números complexos.

***Como todo número real é complexo, toda matriz real é um tipo particular de matriz complexa.***

**Definição 1.27:** Se A é uma matriz complexa, então $\overline{A}$ denota a matriz obtida de A pela troca de cada elemento z = a + bi pelo seu conjugado $\overline{z}$ = a – bi. A matriz $\overline{A}$ é chamada **conjugada** da matriz A.

**Exemplo 1.6:** As matrizes seguintes são conjugadas uma da outra:

$$\begin{pmatrix} 3+i & 1 \\ i & 6-4i \end{pmatrix} \text{ e } \begin{pmatrix} 3-i & 1 \\ -i & 6+4i \end{pmatrix}$$

***Uma dada matriz A é real, se, e somente se,*** **A = $\overline{A}$.**

**Definição 1.28:** A transposta da conjugada de uma matriz A é denotada por $A^*$, isto é, $A^* = \left(\overline{A}\right)^T$.

**Exemplo 1.7:** Se A = $(a_{ij})$, então, $A^T = (a_{ji})$, $\overline{A} = \left(\overline{a_{ij}}\right)$ e $\left(\overline{A}\right)^T = \left(\overline{a_{ji}}\right) = \left(\overline{A^T}\right)$. Portanto, a transposta da conjugada de uma matriz é igual à conjugada da transposta da matriz.

**Definição 1.29:** Uma matriz A tal que A = $A^*$ é chamada **matriz Hermitiana**; isto é, uma matriz A = $(a_{ij})$ é uma matriz Hermitiana, se, e somente se, $a_{ij} = \overline{a_{ji}}$ para todos os pares (i,j).

Uma vez que $a_{ii} = \overline{a_{ii}}$, apenas se $a_{ii}$ é um número real, os elementos diagonais de uma matriz Hermitiana são números reais. Se A é uma matriz simétrica real, então $a_{ij} = a_{ji}$, e $a_{ij} = \overline{a_{ji}}$ para todos os pares (i,j), pois $a_{ji} = \overline{a_{ji}}$. Portanto, toda matriz simétrica real é uma matriz Hermitiana.

**Definição 1.30:** Uma matriz A tal que A = – $A^*$ é chamada **matriz anti-Hermitiana**; isto é, uma matriz A = $(a_{ij})$ é uma matriz anti-Hermitiana, se, e somente se, $a_{ij} = -\overline{a_{ji}}$ para todos os pares (i,j).

***Toda matriz antissimétrica real é anti-Hermitiana.***

**ER 1.13:** Verifique que A é uma matriz anti-Hermitiana onde

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

**Solução:**

$$\overline{A} = \begin{pmatrix} -2i & 3 & -i \\ -3 & 0 & -2+i \\ -i & 2+i & -i \end{pmatrix} \quad A^* = \left(\overline{A}\right)^T = \begin{pmatrix} -2i & 3 & -i \\ 3 & 0 & 2+i \\ -i & -2+i & -i \end{pmatrix}$$

$$-A^* = \begin{pmatrix} 2i & 3 & i \\ -3 & 0 & -2-i \\ i & 2-i & i \end{pmatrix} = A.$$

Portanto, A é uma matriz anti-Hermitiana.

**OCTAVE 1.13**

No OCTAVE, a unidade imaginária é representada pela letra `i` ou pela letra `j` (são caracteres especiais reservados), de modo que um número complexo z = 2 + 3i tem a seguinte forma:

```
> z = 2 + 3*i;
>
```

Uma matriz complexa tem a entrada de seus valores de forma natural, considerando cada componente da matriz como um número complexo. A diferença é que **a operação de transposição de uma matriz complexa no OCTAVE resulta numa transposta conjugada**. Desta forma, verificar se uma determinada matriz complexa é Hermitiana ou anti-Hermitiana, equivale a avaliar $A - A^* = 0$ ou $A + A^* = 0$, respectivamente. Para o caso de matrizes complexas, no OCTAVE, A – A' = 0 ou A + A' = 0. No caso do ER 1.13, basta definir a matriz A e avaliar a expressão A + A':

```
> A = [2*i,3,i;-3,0,-2-i;i;2-i;i]
> A =
   0 + 2i   3 + 0i   0 + 1i
```

```
   0 + 2i    3 + 0i    0 + 1i
  -3 + 0i    0 + 0i   -2 - 1i
   0 + 1i    2 - 1i    0 + 1i
> A+A'
> ans =
        0     0     0
        0     0     0
        0     0     0
>
```

E, portanto, a matriz A é uma matriz anti-Hermitiana.

**ER 1.14:** Verifique que $\overline{AB} = \overline{A}.\overline{B}$.

**Solução:** Sejam A = ($a_{ij}$) e B = ($b_{ij}$) matrizes de ordem m por n e n por p, respectivamente. Então o j-ésimo elemento de AB é dado por $\sum_{k=1}^{n} a_{ik}b_{kj}$ e o j-ésimo elemento de $\overline{AB}$ é dado por $\sum_{k\ 1} \overline{a_{ik}b_{kj}}$, pois o conjugado da soma de números complexos é igual à soma de seus conjugados. Portanto, como o conjugado do produto de dois números complexos é igual ao produto de seus conjugados,

$$\sum_{k=1}^{n} \overline{a_{ik}b_{kj}} \quad \sum_{k=1}^{n} \overline{a_{ik}.b_{kj}}\ .$$

Portanto, $\overline{AB} = \overline{A}.\overline{B}$; isto é, o conjugado do produto de duas matrizes é igual ao produto de seus conjugados.

**ER 1.15:** Verifique que $(AB)^* = A^*B^*$.

**Solução:**

$$(AB)^* = \left(\overline{AB}\right)^T \quad \text{por definição}$$

$$= (\overline{A}.\overline{B})^T \quad \text{pelos resultados do Exemplo 2}$$

$$= \left(\overline{A}\right)^T \left(\overline{B}\right)^T \quad \text{pelo Teorema 1.5}$$

$$= B^*A^* \quad \text{por definição.}$$

# 2. SiSTEmAS DE EQUAÇÕES LiNEARES

**Um dOS pROBlemAS** mais comuns em matemática consiste na resolução de sistemas de equações lineares. Grande parte dos problemas matemáticos oriundos de aplicações científicas e industriais envolve de alguma forma, durante a busca pela solução, sistemas lineares. De modo geral, é um caminho aceitável procurar reduzir problemas mais complexos em conjuntos de sistemas de equações lineares. Algumas das técnicas de resolução de sistemas lineares carecem do uso de um número real associado a uma dada matriz. Estamos nos referindo aos determinantes, que são números associados a matrizes quadradas e permitem determinar se uma dada matriz admite inversa. Estes conceitos serão abordados com mais detalhes neste capítulo, que começamos com a definição de determinantes.

## 2.1 DETERMINANTE

**Definição 2.1:** Associada a cada matriz quadrada A = ($a_{ij}$) de ordem 2, existe uma função matricial chamada **determinante**, denotada por detA ou por

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

A função determinante associa a cada matriz quadrada A de ordem 2 um único número real $a_{11}a_{22} - a_{21}a_{12}$ chamado **valor do determinante**; isto é, detA = $a_{11}a_{22} - a_{21}a_{12}$.

O valor do determinante de uma matriz de ordem 2 pode ser relembrado pelo seguinte esquema:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}^{-}_{+} = a_{11}a_{22} - a_{21}a_{12}. \quad (2.1)$$

Note que o valor do determinante é a diferença entre o produto dos elementos da diagonal principal e o produto dos outros dois restantes.

**ER 2.1:** Encontre o valor do determinante da matriz A = $\begin{bmatrix} 3 & -1 \\ 5 & 4 \end{bmatrix}$.

**Solução:**

$$\det A = \begin{vmatrix} 3 & -1 \\ 5 & 4 \end{vmatrix} = (3)(4) - (5)(-1) = 17.$$

**OCTAVE 2.1**

No OCTAVE, existe uma função que calcula diretamente o determinante de uma matriz, então definimos a matriz e usamos o comando `det`:

```
> A = [3,-1;5,4];
> det(A)
> ans =  17.000
```

**Definição 2.2:** Um **determinante** de uma matriz A = $(a_{ij})$ de ordem 3 é uma função matricial e é denotado por detA ou por

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

O valor do determinante de A é definido como

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$
$$-a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}\,.$$

O valor do determinante de uma matriz de ordem 3 pode ser relembrado pelo uso de um esquema particular, semelhante ao usado para determinantes de matrizes de ordem dois (2.1):

$$\begin{array}{c} -a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12} \\ \left| \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array} \right| \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{array} \\ +a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \end{array} \tag{2.2}$$

Em (2.2), as duas primeiras colunas do determinante são repetidas à direita. Os produtos dos três elementos ao longo das setas apontando para baixo e para a direita são contabilizados positivamente, enquanto os outros três elementos ao longo das setas apontando para cima e para a direita são considerados negativos. A soma algébrica desses seis produtos é o valor do determinante.

**ER 2.2:** Encontre o valor do determinante da matriz A = $\begin{pmatrix} 0 & 5 & 2 \\ 2 & -1 & -2 \\ 5 & 0 & 3 \end{pmatrix}$.

**Solução:** Usando o esquema de (2.2),

$$-(-10)-(0)-(30)$$

$$\left|\begin{array}{ccc|cc} 0 & 5 & 2 & 0 & 5 \\ 2 & -1 & -2 & 2 & -1 \\ 5 & 0 & 3 & 5 & 0 \end{array}\right.$$

$$+(0)+(-50)+(0)$$

Portanto, o valor do determinante da matriz A é (0) + (–50) + (0) – (–10) – (0) – (30), isto é, detA = –70.

**OCTAVE 2.2**

Diretamente, definimos a matriz e usamos o comando `det`:

```
> A = [0,5,2;2,-1,-2;5,0,3];
> det(A)
> ans =  -70.000
```

Note que o método das diagonais empregado para determinação de determinantes de matrizes de ordem 3 não é válido para determinantes de matrizes de ordem maior que 3.

Em geral, um **determinante** de uma matriz A = $(a_{ij})$ de ordem n é uma função da matriz quadrada A e é denotado por detA ou por

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

**Definição 2.3:** O valor do determinante de uma matriz de ordem n é definido como a soma dos n! termos da forma

$$(-1)^k a_{1i_1} a_{2i_2} \dots a_{ni_n}.$$

onde cada termo contém um e apenas um elemento de cada linha e um e apenas um elemento de cada coluna; isto é, os índices $i_1, i_2, \dots, i_n$ são iguais a 1, 2, ..., n tomados em alguma ordem. O expoente k representa o número de mudanças necessárias, entre dois elementos, para que os índices $i_1, i_2, ..., i_n$ sejam colocados na ordem 1, 2, ..., n.

**Exemplo 2.1:** Considere o termo contendo $a_{13}a_{21}a_{34}a_{42}$ na avaliação do determinante de uma matriz de ordem quatro. O valor k é 3, pois são necessárias três mudanças entre dois elementos, de modo que os índices (3,1,4,2) sejam colocados na ordem (1,2,3,4):

$$(3,1,4,2) \to (1,3,4,2) \to (1,2,4,3) \to (1,2,3,4)$$

$$a_{13}a_{21}a_{34}a_{42} = a_{21}a_{13}a_{34}a_{42} = a_{21}a_{42}a_{34}a_{13} = a_{21}a_{42}a_{13}a_{34}.$$

Portanto, o termo contendo o fator $a_{13}a_{21}a_{34}a_{42}$ tem o fator adicional $(-1)^3$; isto é, $-1$.

Embora os seguintes teoremas sejam válidos para determinantes de matrizes de ordem n, suas provas serão consideradas apenas para determinantes de matrizes de ordem 3. Estes teoremas são extremamente úteis na avaliação de determinantes de matrizes de ordem $n \geq 4$, pois os valores destes determinantes são raramente calculados por meio da definição. Além disso, vários destes teoremas serão usados para discutir certos conceitos em álgebra matricial.

**Teorema 2.1:** *O valor de um determinante permanece inalterado se linhas e colunas correspondentes são trocados; isto é, $\det A = \det A^T$.*

**Prova:** Considere os determinantes $\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ e $\det A^T = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$.

Note que os elementos de cada linha ou coluna de detA correspondem aos elementos da mesma coluna ou linha, respectivamente de $\det A^T$. Por definição,

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

e

$$\det A^T = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21}.$$

Por um simples rearranjo dos fatores e termos, nota-se que detA = det$A^T$.

A importância do Teorema 2.1 é que, para todo teorema provado, com respeito às linhas de um determinante, existe um teorema correspondente, com respeito às colunas.

**Teorema 2.2:** *Se quaisquer duas linhas (colunas) de um determinante são trocadas, então o sinal do valor do determinante é trocado.*

**Exemplo 2.2:**

$$a)\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = -\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{vmatrix};$$

$$b)\begin{vmatrix} 1 & 2 & -1 \\ 2 & 0 & 7 \\ -3 & 1 & 4 \end{vmatrix} = -\begin{vmatrix} 2 & 1 & -1 \\ 0 & 2 & 7 \\ 1 & -3 & 4 \end{vmatrix}$$

**Teorema 2.3:** *Se todo elemento de uma linha de um determinante é zero, então o valor do determinante é zero.*

**Prova:** Considere a definição do valor do determinante de uma matriz de ordem 3. Todo termo desta expressão contém um, e apenas um, elemento de cada linha como fator. Contudo, todo termo deve conter um fator que é um elemento da linha de zeros. Portanto, o valor do determinante é zero.

**Teorema 2.4:** *Se todo elemento de uma linha de um determinante é multiplicado por um fator k, então o valor do determinante é multiplicado por k.*

**Prova:** Todo termo na definição do valor do determinante de uma matriz de ordem 3 contém um, e apenas um, elemento de cada linha como fator. Assim, todo termo deve conter o fator k uma única vez; isto é, o valor do determinante será multiplicado por k.

**Exemplo 2.3:** $\begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$

**Teorema 2.5:** *O valor de um determinante permanece inalterado se cada elemento de uma linha é aumentado por um múltiplo escalar do elemento correspondente de outra linha.*

**Exemplo 2.4:** $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21}+ka_{31} & a_{22}+ka_{32} & a_{23}+ka_{33} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$

Note que os elementos da terceira linha são deixados intactos enquanto um múltiplo escalar destes elementos é adicionado a cada um dos elementos correspondentes na segunda linha. A prova do Teorema 2.5 pode ser obtida pelo uso da definição do valor do determinante de uma matriz de ordem 3.

O valor do determinante de uma matriz de ordem 3 foi definido pela equação

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \tag{2.3}$$

$$- a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.$$

Pelo rearranjo dos termos, o membro do lado direito de (2.3) pode ser escrito como

$$a_{11}\left(a_{22}a_{33} - a_{32}a_{23}\right) - a_{12}\left(a_{21}a_{33} - a_{31}a_{23}\right) + a_{13}\left(a_{21}a_{32} - a_{31}a_{22}\right),$$

e

$$a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{2.4}$$

Os três determinantes em (2.4) podem ser obtidos do determinante original pela eliminação de certas linhas e colunas. A regra mostrada por (2.4), para expressar o valor do determinante de uma matriz de ordem 3, pode ser generalizada para determinantes de matrizes de ordem mais alta.

**Definição 2.4:** Considere uma matriz quadrada $(a_{ij})$ de ordem n. O determinante obtido da matriz formada pela eliminação dos elementos da i-ésima linha e da j-ésima coluna é chamado **menor complementar** do elemento $a_{ij}$ e é denotado por $M_{ij}$.

**Definição 2.5:** O **cofator** do elemento $a_{ij}$, denotado pelo símbolo $A_{ij}$, é dado pelo menor complementar $M_{ij}$, multiplicado por um sinal positivo se i+j é par, ou um sinal negativo se i+j é ímpar, ou seja,

$$A_{ij} = (-1)^{i+j} M_{ij}. \tag{2.5}$$

**Exemplo 2.5:** Por exemplo, o menor complementar $a_{12}$ em $\det A = \begin{vmatrix} 1 & 0 & 1 \\ 5 & -1 & 3 \\ -1 & 2 & 7 \end{vmatrix}$ é

$$M_{12} = \begin{vmatrix} 5 & 3 \\ -1 & 7 \end{vmatrix} = (5)(7) - (-1)(7) = 35 + 7 = 42$$

e o cofator é

$$A_{12} = (-1)^{1+2} M_{12} = (-1)(42) = -42.$$

**Teorema 2.6:** *O valor de um determinante é igual à soma dos produtos de cada elemento pelo seu cofator de qualquer linha (ou coluna).*

O Teorema 2.6 é um teorema-chave para avaliação de determinantes de matrizes de ordem n. A prova deste teorema para matrizes de ordem n pode ser encontrado na maioria dos textos de álgebra linear.

**ER 2.3:** Avalie $\begin{vmatrix} 1 & 0 & 1 \\ 5 & -1 & 3 \\ -1 & 2 & 7 \end{vmatrix}$ usando os cofatores dos elementos da segunda coluna.

**Solução:** Consideremos a segunda coluna. Então, os cofatores de 0, –1 e 2, denotados por $A_{12}$, $A_{22}$ e $A_{32}$, respectivamente são dados por:

$$A_{12} = (-1)^{1+2} \begin{vmatrix} 5 & 3 \\ -1 & 7 \end{vmatrix} = -42; \quad A_{22} = (-1)^{2+2} \begin{vmatrix} 1 & 1 \\ -1 & 7 \end{vmatrix} = 8; \quad A_{32} = (-1)^{3+2} \begin{vmatrix} 1 & 1 \\ 5 & 3 \end{vmatrix} = 2.$$

Usando o Teorema 2.6

$$\begin{vmatrix} 1 & 0 & 1 \\ 5 & -1 & 3 \\ -1 & 2 & 7 \end{vmatrix} = (0)\,A_{12} + (-1)\,A_{22} + (2)\,A_{32} = (0)(-42) + (-1)(8) + (2)(2) = -4.$$

O símbolo |A| também é usado para denotar o determinante de A (neste caso, |A| não significa valor absoluto de A).

## 2.2 INVERSA DE UMA MATRIz

Esta seção lida com o problema de se encontrar uma inversa multiplicativa, caso exista, para uma dada matriz quadrada.

**Definição 2.6:** Uma **inversa multiplicativa à esquerda** de uma matriz A é uma matriz B tal que BA = I; uma **inversa multiplicativa à direita** de uma matriz A é uma matriz C tal que AC = I. Se as inversas à esquerda e à direita de uma matriz A são iguais, então, esta é chamada simplesmente **inversa multiplicativa** de A e é denotada por $A^{-1}$.

**Teorema 2.7:** *Uma inversa multiplicativa à esquerda de uma matriz quadrada A não nula é uma inversa multiplicativa de A.*

**Prova:** Suponha que BA = I, então

$A(BA) = AI$ pela pré-multiplicação por A,

$(AB)A = A$ pois a multiplicação de matrizes é associativa e $AI = A$

$(AB)A - A = 0 \Rightarrow (AB - I)A = 0$, 0 é matriz nula, mesma ordem de A,

$AB - I = 0$ pois A é não nula no produto anterior (hipótese),

$AB = I$.

Além disso, B é também uma inversa multiplicativa à direita de A. Portanto, B é uma inversa multiplicativa de A; isto é, $B = A^{-1}$.

Analogamente, pode-se mostrar que o Teorema 2.8 é verdadeiro.

**Teorema 2.8:** *Uma inversa multiplicativa à direita de uma matriz quadrada A é uma inversa multiplicativa de A.*

**Teorema 2.9:** *A inversa multiplicativa de uma matriz quadrada A, caso exista, é única.*

**Prova:** Sejam $A^{-1}$ e B quaisquer duas inversas multiplicativas da matriz quadrada A. Uma vez que $A^{-1}A = I$ e $BA = I$, então,

$$A^{-1}A = BA$$

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

$$A^{-1}\left(AA^{-1}\right) = B\left(AA^{-1}\right)$$

$$A^{-1}I = BI$$

$$A^{-1} = B.$$

Portanto, quaisquer duas inversas multiplicativas de uma matriz A são identicamente iguais; isto é, a inversa multiplicativa, caso exista, de uma matriz quadrada A é única.

**Exemplo 2.5:** Nem toda matriz quadrada tem inversa multiplicativa. A matriz $\begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}$ não tem inversa multiplicativa. Para verificar isto, considere que $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ é a inversa multiplicativa de $\begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}$, deve-se observar, então,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} a & 3a \\ c & 3c \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \qquad (2.6)$$

Isto exige que 4 equações sejam satisfeitas: $a = 1$, $3a = 0$, $c = 0$ e $3c = 1$. Entretanto, se $a = 1$, então $3a \neq 0$. Portanto, os valores de a, b, c e d que satisfazem (2.6) não existem, e a matriz $\begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}$ não tem inversa multiplicativa.

**ER 2.4:** Encontre a inversa multiplicativa, caso exista, de $\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$.

**Solução:**

Se $\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$ tem uma inversa multiplicativa $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

então $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. ou $\begin{pmatrix} a+4b & 3a+2b \\ c+4d & 3c+2d \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Portanto, $\begin{cases} a+4b=1 \\ 3a+2b=0 \end{cases}$ e $\begin{cases} c+4d=0 \\ 3c+2d=1 \end{cases}$

Resolvendo estes pares de equações, simultaneamente, encontra-se a = –1/5, b = 3/10, c = 4/10 e d = –1/10. Portanto, a inversa multiplicativa de

$\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$ é $\begin{pmatrix} -\frac{1}{5} & \frac{3}{10} \\ \frac{4}{10} & -\frac{1}{10} \end{pmatrix}$.

**OCTAVE 2.4**

No OCTAVE, existe uma função que calcula a inversa multiplicativa de uma matriz, caso exista. Então definimos a matriz e usamos o comando `inv`:

```
> A = [1,3;4,2];
> inv(A)
> ans =
  -0.20000   0.30000
   0.40000  -0.10000
```

Observe que a saída da matriz apresenta os resultados das frações, quando comparados com o ER 2.4, com 5 casas decimais.

**ER 2.5:** Encontre a forma da inversa multiplicativa, caso exista, da matriz quadrada genérica $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

**Solução:** Se $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ tem uma inversa multiplicativa $\begin{pmatrix} w & x \\ y & z \end{pmatrix}$, então

$$\begin{pmatrix} w & x \\ y & z \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} aw + cx & bw + dx \\ ay + cz & by + dz \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Além disso,

$$\begin{cases} aw + cx = 1 \\ bw + dx = 0 \end{cases} \text{ e } \begin{cases} ay + cz = 0 \\ by + dz = 1 \end{cases}$$

Se ad – bc ≠ 0, as soluções destes pares de equações são dadas por

$$w = \frac{d}{ad - bc}, \quad x = \frac{-b}{ad - bc}, \quad y = \frac{-c}{ad - bc}, \quad z = \frac{a}{ad - bc}.$$

Portanto, a inversa multiplicativa de $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ é $\begin{pmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{pmatrix}$ desde que ad–bc ≠ 0. Note que a inversa multiplicativa de uma matriz quadrada de ordem 2 existe se $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0.$

Os ER 2.4 e ER 2.5 ilustram o uso de um método um tanto quanto trabalhoso para se encontrar as inversas multiplicativas de matrizes quadradas de ordem maior que 2. Agora, será apresentado um método direto para se encontrar a inversa multiplicativa, caso exista, de uma matriz quadrada de qualquer ordem.

Se A = $(a_{ij})$, então pelo Teorema 2.6

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij} \quad \text{para i = 1,2,...,n.} \tag{2.7}$$

Além disso, é possível provar que, para qualquer determinante de ordem n, a soma dos produtos dos elementos de qualquer linha e os cofatores dos elementos correspondentes de outra linha é zero, ou seja,

$$\sum_{j=1}^{n} a_{hj} A_{ij} = 0 \quad \text{para h, i = 1,2,...,n onde h} \neq \text{i.} \tag{2.8}$$

As Equações (2.7) e (2.8) podem ser expressas como uma única equação

$$\sum_{j=1}^{n} a_{hj} A_{ij} = \delta_{ij} \det A \text{, para h,i = 1,2,...,n.} \tag{2.9}$$

As $n^2$ equações representadas por (2.9) podem ser escritas numa matriz da forma

$$\left(a_{ij}\right)\left(A_{ij}\right)^{T} = \left(\delta_{ij}\right)\det A \ . \tag{2.10}$$

Por exemplo, se n = 3 em (2.9), então, a equação (2.10) representa

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \det A.$$

Uma vez que $(a_{ij}) = A$ e $(\delta_{ij}) = I$, a equação (2.10) pode ser escrita como

$$A\left(A_{ij}\right)^T = I \det A. \tag{2.11}$$

Se $\det A \neq 0$, então

$$A \frac{\left(A_{ij}\right)^T}{\det A} = I. \tag{2.12}$$

Portanto, a matriz $\frac{\left(A_{ij}\right)^T}{\det A}$ é a inversa multiplicativa de A; isto é,

$$A^{-1} = \frac{\left(A_{ij}\right)^T}{\det A} = \frac{1}{\det A}\left(A_{ij}\right)^T. \tag{2.13}$$

***A inversa multiplicativa, caso exista, de uma matriz quadrada é o produto entre o inverso do determinante desta matriz e a transposta da matriz dos cofatores.***

O fato de que $\frac{\left(A_{ij}\right)^T}{\det A}$ não é apenas a inversa à direita de A, decorre do Teorema 2.8. Note que uma condição necessária e suficiente para a inversa multiplicativa da matriz A existir, é que $\det A \neq 0$.

**Definição 2.6:** Uma matriz quadrada A é dita **não singular** se detA ≠ 0, e **singular** se detA = 0.

Se A não é uma matriz quadrada, então é possível para A ter uma inversa multiplicativa à esquerda ou à direita, mas não ambas.

**Exemplo 2.6:** Considere $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

Então, toda matriz da forma $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ r & s \end{pmatrix}$, onde r e s são escalares arbitrários, é uma inversa multiplicativa à direita de A. Uma inversa multiplicativa à esquerda $\begin{pmatrix} m & n \\ w & x \\ y & z \end{pmatrix}$ não existe, pois

$$\begin{pmatrix} m & n \\ w & x \\ y & z \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} m & n & 0 \\ w & x & 0 \\ y & z & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

para todos os valores de m, n, w, x, y e z.

**ER 2.6:** Encontre a inversa multiplicativa, caso exista, de $A = \begin{pmatrix} -2 & 1 & 1 \\ 0 & 2 & -1 \\ 4 & 3 & 4 \end{pmatrix}$.

**Solução:** Usando o Teorema 2.6 e os cofatores dos elementos da primeira linha, obtém-se

$$\det A = -2\begin{vmatrix} 2 & -1 \\ 3 & 4 \end{vmatrix} - 1\begin{vmatrix} 0 & -1 \\ 4 & 4 \end{vmatrix} + 1\begin{vmatrix} 0 & 2 \\ 4 & 3 \end{vmatrix} = -2(11) - 1(4) + 1(-8) = -34.$$

Como $\det A \neq 0$, então $A^{-1}$ existe. Os cofatores dos elementos de A são dados por

$$A_{11} = \begin{vmatrix} 2 & -1 \\ 3 & 4 \end{vmatrix} = 11, \quad A_{12} = -\begin{vmatrix} 0 & -1 \\ 4 & 4 \end{vmatrix} = -4, \quad A_{13} = \begin{vmatrix} 0 & 2 \\ 4 & 3 \end{vmatrix} = -8,$$

$$A_{21} = -\begin{vmatrix} 1 & 1 \\ 3 & 4 \end{vmatrix} = -1, \quad A_{22} = \begin{vmatrix} -2 & 1 \\ 4 & 4 \end{vmatrix} = -12, \quad A_{23} = -\begin{vmatrix} -2 & 1 \\ 4 & 3 \end{vmatrix} = 10,$$

$$A_{31} = \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} = -3, \quad A_{32} = -\begin{vmatrix} -2 & 1 \\ 0 & -1 \end{vmatrix} = -2, \quad A_{33} = \begin{vmatrix} -2 & 1 \\ 0 & 2 \end{vmatrix} = -4.$$

Trocando cada elemento de A pelo seu cofator, obtemos a matriz

$$\left(A_{ij}\right) = \begin{pmatrix} 11 & -4 & -8 \\ -1 & -12 & 10 \\ -3 & -2 & -4 \end{pmatrix}.$$

Portanto,

$$A^{-1} = \frac{\left(A_{ij}\right)^T}{\det A} = \frac{1}{-34}\begin{pmatrix} 11 & -1 & -3 \\ -4 & -12 & -2 \\ -8 & 10 & -4 \end{pmatrix} = \begin{pmatrix} -\frac{11}{34} & \frac{1}{34} & \frac{3}{34} \\ \frac{4}{34} & \frac{12}{34} & \frac{2}{34} \\ \frac{8}{34} & -\frac{10}{34} & \frac{4}{34} \end{pmatrix}.$$

**OCTAVE 2.6**

Simplesmente definimos a matriz e usamos o comando `inv`:

```
> A = [-2,1,1;0,2,-1;4,3,4];
> inv(A)
> ans =
 -0.323529  0.029412   0.088235
  0.117647  0.352941   0.058824
  0.235294 -0.294118   0.117647
```

**ER 2.7:** Resolva o sistema de equações lineares $\begin{cases} 3x - 2y = 11 \\ x + 3y = 3 \end{cases}$

**Solução:** Este sistema de equações lineares pode ser expresso em forma matricial como

$$\begin{pmatrix} 3 & -2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 11 \\ 3 \end{pmatrix}.$$

A inversa multiplicativa de $\begin{pmatrix} 3 & -2 \\ 1 & 3 \end{pmatrix}$ é $\begin{pmatrix} \frac{3}{11} & \frac{2}{11} \\ -\frac{1}{11} & \frac{3}{11} \end{pmatrix}$. Pré-multiplicando ambos os lados da equação matricial que representa o sistema de equações lineares pela inversa multiplicativa, obtém-se

$$\begin{pmatrix} \frac{3}{11} & \frac{2}{11} \\ -\frac{1}{11} & \frac{3}{11} \end{pmatrix}\begin{pmatrix} 3 & -2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{3}{11} & \frac{2}{11} \\ -\frac{1}{11} & \frac{3}{11} \end{pmatrix}\begin{pmatrix} 11 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{39}{11} \\ \frac{-2}{11} \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{39}{11} \\ \frac{-2}{11} \end{pmatrix}.$$

Portanto, x = 39/11 e y = –2/11.

**OCTAVE 2.7**

Considerando a equação matricial AX = B, equivalente ao sistema de equações, onde A é matriz dos coeficientes, X é a matriz das variáveis e B é matriz dos termos independentes, definimos as matrizes A e B e calculamos $A^{-1}B$:

```
> A = [3,-2;1,3];
> B = [11;3];
> inv(A)*B
> ans =
  3.54545
 -0.18182
```

Assim, a saída da operação executada mostra a solução do sistema: x = 3.54545 e y = –0.18182.

## 2.3 ANÉIS E ISOMORFISMOS DE MATRIzES

O conjunto das matrizes quadradas de uma dada ordem, com as operações de adição e multiplicação de matrizes, forma um modelo de um sistema matemático abstrato particular chamado **anel**.

**Definição 2.7:** Um **anel** é um sistema matemático consistindo de um conjunto R com elementos a, b, c, d, e, f, g, ..., uma relação de equivalência denotada por = e duas operações binárias bem definidas + e ×, chamadas "adição" e "multiplicação", respectivamente, as quais satisfazem as seguintes propriedades:

1. *fechamento*: se a, b ∈ �, então $a + b \in$ � e $a \times b \in$ �;
2. *comutativa*: se a, b ∈ �, então $a + b = b + a$;
3. *associativa*: se a, b ∈ �, então $a + (b + c) = (a + b) + c$ e $a \times (b \times c) = (a \times b) \times c$;
4. *identidade aditiva*: existe um elemento z ∈ � tal que, para todo a ∈ �, $a + z = z + a = a$;
5. *inverso aditivo:* para todo a ∈ �, existe um elemento $(-a) \in$ �, tal que $a + (-a) = (-a) + a = z$;
6. *distributiva:* para todo a, b, c ∈ �, $a \times (b + c) = (a \times b) + (a \times c)$ e $(b + c) \times a = (b \times a) + (c \times a)$.

**ER 2.8:** Mostre que o conjunto S de matrizes quadradas reais de ordem n forma um anel.

**Solução:** Pelas definições da soma e do produto de duas matrizes, o conjunto S é fechado sob a adição e multiplicação de matrizes (propriedade 1). A adição de duas matrizes, de S é comutativa (propriedade 2) e associativa (propriedade 3), pois a adição de números reais é comutativa e associativa. A propriedade 4 é satisfeita porque a matriz nula de ordem n é o elemento neutro aditivo para o conjunto S. Uma vez que, para cada $(a_{ij}) \in S$ existe uma matriz $(-a_{ij})$ tal que $(a_{ii}) + (-a_{ii}) = 0$, a propriedade 5 é satisfeita. A propriedade 6 e a segunda parte da propriedade 3 foram provadas na seção 1.2. Portanto, o conjunto S de matrizes quadradas reais de ordem n forma um anel.

**Definição 2.8:** Uma correspondência biunívoca entre os elementos de um anel R e os elementos de um anel S é chamada **isomorfismo** se a soma e o produto de elementos correspondentes também são correspondentes; isto é, se quaisquer dois elementos $a,b \in R$ e quaisquer dois elementos $a', b' \in S$ são tais que a corresponde com a' $(a \leftrightarrow a')$ e b corresponde com b' $(b \leftrightarrow b')$, e a correspondência é um isomorfismo se $a + b \leftrightarrow a' + b'$ e $ab \leftrightarrow a'b'$.

Sempre que existe um isomorfismo entre os elementos de dois anéis R e S, os anéis são abstratamente idênticos e são ditos **isomórficos**. Apenas as notações usadas para representar os elementos e as operações dos dois anéis diferem.

Existem dois subconjuntos importantes do conjunto de matrizes reais quadradas de ordem dois. O primeiro é o conjunto de matrizes escalares de ordem dois $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$.

Este conjunto de matrizes é isomórfico ao conjunto dos números reais. Tanto o conjunto de matrizes escalares quanto o conjunto de números reais, com respeito às operações de adição e multiplicação definidas nos conjuntos, são exemplos de anéis.

**Exemplo 2.7:** Considere a correspondência biunívoca entre as matrizes $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ e os escalares k; isto é, $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \leftrightarrow k$.

De modo a estabelecer que os dois sistemas são isomórficos, é necessário mostrar que as somas e os produtos de elementos correspondentes nos dois sistemas estão em mesma correspondência um pra um. Ou seja, que

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \leftrightarrow a \quad e \quad \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} \leftrightarrow b.$$

Então,

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a+b & 0 \\ 0 & a+b \end{pmatrix} \leftrightarrow a+b \quad e$$

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} ab & 0 \\ & \end{pmatrix} \leftrightarrow ab.$$

Portanto, o anel de matrizes da forma $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ é isomórfico ao anel de números reais k.

O segundo subconjunto importante de matrizes reais quadradas de ordem dois é o conjunto de matrizes da forma $\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$.

Este conjunto de matrizes é isomórfico ao conjunto de números complexos x + yi. Tanto o conjunto de matrizes desta forma quanto o conjunto de números complexos, com respeito às operações de adição e multiplicação definidas nos conjuntos, são exemplos de anéis.

**Exemplo 2.8:** Considere a correspondência biunívoca entre as matrizes $\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$ e o conjunto dos números complexos $x + yi$; isto é, $\begin{pmatrix} x & y \\ -y & x \end{pmatrix} \leftrightarrow x + yi$.

Sejam $\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \leftrightarrow a + bi$ e $\begin{pmatrix} c & d \\ -d & c \end{pmatrix} \leftrightarrow c + di$.

Então,

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ -(b+d) & a+c \end{pmatrix}$$

e $(a+bi)+(c+di)=(a+c)+(b+d)i$ . Uma vez que

$$\begin{pmatrix} a+c & b+d \\ -(b+d) & a+c \end{pmatrix} \leftrightarrow (a+c)+(b+d)i\,,$$

então

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} c & d \\ & \end{pmatrix} \leftrightarrow (a+bi)+(c+di)\,,$$

o que demonstra que a correspondência é preservada mediante a adição. Para a multiplicação,

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix}$$

e $(a+bi)\times(c+di)=(ac-bd)+(ad+bc)i$ . Uma vez que

$$\begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix} \leftrightarrow (ac-bd)+(ad+bc)i\,,$$

Então,

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} c & d \\ & \end{pmatrix} \leftrightarrow (a+bi)\times(c+di)\,,$$

o que demonstra que a correspondência é preservada mediante a multiplicação.

**ER 2.9:** Verifique que a correspondência biunívoca entre as matrizes $\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$ e os números complexos x + yi é preservada mediante a multiplicação, para os números complexos 3 + i e 1 – 5i.

**Solução:**

Sejam 3 + i $\leftrightarrow \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}$ e 1 – 5i $\leftrightarrow \begin{pmatrix} 1 & -5 \\ 5 & 1 \end{pmatrix}$. Então,

$$(3+i)(1-5i)=\left[(3)(1)-(1)(-5)\right]+\left[(3)(-5)+(1)(1)\right]i=8-14i$$

e

$$\begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 1 & -5 \\ 5 & 1 \end{pmatrix}=\begin{pmatrix} 8 & -14 \\ 14 & 8 \end{pmatrix}$$

Uma vez que $8-14i \leftrightarrow \begin{pmatrix} 8 & -14 \\ 14 & 8 \end{pmatrix}$, então,

$$(3+i)\times(1-5i)\leftrightarrow\begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 1 & -5 \\ 5 & 1 \end{pmatrix}.$$

**OCTAVE 2.9**

Dado um número complexo z = a+bi, as partes real e imaginária, no OCTAVE, são avaliadas pelas funções `real(z)` e `imag(z)`, respectivamente. Desta forma, podemos relacionar o número complexo z com uma matriz da forma $\begin{pmatrix} \texttt{real(z)} & \texttt{imag(z)} \\ -\texttt{imag(z)} & \texttt{real(z)} \end{pmatrix}$. Assim, definimos os números complexos a considerar, definimos as matrizes associadas e avaliamos a multiplicação como segue:

```
> z1 = 3 + i;
> z2 = 1 - 1*i;
> z3 = z1*z2
z3 = 8 - 14i
> A1 = [real(z1),imag(z1);-imag(z1),real(z1)];
> A2 = [real(z2),imag(z2);-imag(z2),real(z2)];
> A3 = [real(z3),imag(z3);-imag(z3),real(z3)]
> A3 =
     8   -14
     14    8
```

```
> A1*A2
> ans =
      8    -14
      14    8
```

Observe que multiplicamos os complexos z1 e z2 e guardamos o resultado na variável z3. Na sequência, definimos as matrizes relacionadas a z1, z2 e z3. O resultado obtido pela definição para z3 é o mesmo obtido do produto das matrizes A1 e A2.

Uma razão para a importância do estudo da álgebra de matrizes é que muitos sistemas matemáticos são isomórficos a conjuntos de matrizes. Considere agora um sistema matemático que tem grande importância histórica no desenvolvimento da álgebra vetorial.

Para as duas definições seguintes, sejam i, j e k, tais que $i^2 = j^2 = k^2 = -1$, $ij = -ji = = k$, $jk = -kj = i$, e $ki = -ik = j$.

**Definição 2.9:** Um **quatérnio** é um elemento da forma $a + bi + cj + dk$, onde a, b, c e d são números reais.

**Definição 2.10:** Dois quatérnios $a + bi + cj + dk$ e $a' + b'i + c'j + d'k$ são **iguais** se, e somente se, $a = a'$, $b = b'$, $c = c'$, e $d = d'$.

Dois quatérnios são adicionados da mesma maneira que duas funções lineares em i, j e k. O produto de dois quatérnios quaisquer pode ser obtido, considerando a propriedade distributiva da multiplicação com respeito à adição. Sob estas definições de adição e multiplicação de quatérnios, mostra-se que o conjunto de quatérnios forma um anel.

Além disso, pode ser mostrado que o anel dos quatérnios $a + bi + cj + dk$ é isomórfico a dois diferentes conjuntos de matrizes: o conjunto das matrizes reais quadradas de ordem quatro da forma

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

e o conjunto das matrizes complexas quadradas de ordem dois da forma

$$\begin{pmatrix} a+bi & c+di \\ -c+di & a-bi \end{pmatrix}.$$

**ER 2.10:** Verifique que a correspondência biunívoca entre as matrizes $\begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}$ e os quatérnios a + bi + cj + dk é preservada sob a multiplicação para os quatérnios 2 + 3i + j + 4k e 1 – i + 3j – 2k.

**Solução:**

Sejam

$$2+3i+j+4k \leftrightarrow \begin{pmatrix} 2 & 3 & 1 & 4 \\ -3 & 2 & -4 & 1 \\ -1 & 4 & 2 & -3 \\ -4 & -1 & 3 & 2 \end{pmatrix} \text{ e}$$

$$1-i+3j-2k \leftrightarrow \begin{pmatrix} 1 & -1 & 3 & -2 \\ 1 & 1 & 2 & 3 \\ -3 & -2 & 1 & 1 \\ 2 & -3 & -1 & 1 \end{pmatrix}.$$

Então,

$$
\begin{aligned}
(2+3i+j+4k)\times(1-i+3j-2k) = & \left[(2)(1)-(3)(-1)-(1)(3)-(4)(-2)\right]+\\
& +\left[(2)(-1)+(3)(1)+(1)(-2)-(4)(3)\right]j+\\
& +\left[(2)(3)-(3)(-2)+(1)(1)+(4)(-1)\right]j+\\
& +\left[(2)(-2)+(3)(3)-(1)(-1)+(4)(1)\right]k=\\
& =10-13i+9j+10k,
\end{aligned}
$$

e

$$
\begin{pmatrix} 2 & 3 & 1 & 4 \\ -3 & 2 & -4 & 1 \\ -1 & 4 & 2 & -3 \\ -4 & -1 & 3 & 2 \end{pmatrix}
\begin{pmatrix} 1 & -1 & 3 & -2 \\ 1 & 1 & 2 & 3 \\ -3 & -2 & 1 & 1 \\ 2 & -3 & -1 & 1 \end{pmatrix}
=
\begin{pmatrix} 10 & -13 & 9 & 10 \\ 13 & 10 & -10 & 9 \\ -9 & 10 & 10 & 13 \\ -10 & -9 & -13 & 10 \end{pmatrix}.
$$

Uma vez que $10-13i+9j+10k \leftrightarrow \begin{pmatrix} 10 & -13 & 9 & 10 \\ 13 & 10 & -10 & 9 \\ -9 & 10 & 10 & 13 \\ -10 & -9 & -13 & 10 \end{pmatrix}$, então a correspondência biunívoca é preservada mediante a multiplicação.

**OCTAVE 2.10**

O OCTAVE trabalha de forma direta com os quatérnios, onde um quatérnio Q = a+bi+cj+dk é definido por meio da função `quaternion`. A diferença, aqui, é que o quatérnio no OCTAVE é um vetor cujos elementos se ordenam como Q = [b c d a], ou seja, os coeficientes do quatérnio podem ser identificados da seguinte maneira: a = Q(4), b = Q(1), c = Q(2) e d = Q(3). Relacionamos o quatérnio Q com a matriz da forma $\begin{pmatrix} Q(4) & Q(1) & Q(2) & Q(3) \\ -Q(1) & Q(4) & -Q(3) & Q(2) \\ -Q(2) & Q(3) & Q(4) & -Q(1) \\ -Q(3) & -Q(2) & Q(1) & Q(4) \end{pmatrix}$. A multiplicação entre quatérnios é efetuada pela função `qmult`. Assim, definimos os quatérnios, as matrizes associadas e avaliamos a multiplicação como segue:

```
> Q1 = quaternion(2,3,1,4);
> Q2 = quaternion(1,-1,3,-2);
> Q3 = qmult(Q1,Q2)
Q3 =
  10  -15    5  -10
>    A1 = [  Q1(4), Q1(1), Q1(2), Q1(3)
            -Q1(1), Q1(4),-Q1(3), Q1(2)
            -Q1(2), Q1(3), Q1(4),-Q1(1)
            -Q1(3),-Q1(2), Q1(1), Q1(4)]
>A1 =
     4   2   3   1
    -2   4  -1   3
    -3   1   4  -2
    -1  -3   2   4
> A2 = [  Q2(4), Q2(1), Q2(2), Q2(3)
         -Q2(1), Q2(4),-Q2(3), Q2(2)
         -Q2(2), Q2(3), Q2(4),-Q2(1)
         -Q2(3),-Q2(2), Q2(1), Q2(4)]
> A2 =
  -2   1  -1   3
  -1  -2  -3  -1
   1   3  -2  -1
  -3   1   1  -2
>    A3 = [  Q3(4), Q3(1), Q3(2), Q3(3)
            -Q3(1), Q3(4),-Q3(3), Q3(2)
            -Q3(2), Q3(3), Q3(4),-Q3(1)
            -Q3(3),-Q3(2), Q3(1), Q3(4)]
> A3 =
  -10   10  -15    5
  -10  -10   -5  -15
```

```
   -15     5  -10  -10
    -5    15   10  -10
 > A1*A2
 > ans =
   -10    10  -15    5
   -10   -10   -5  -15
    15     5  -10  -10
    -5    15   10  -10
```

Observe que multiplicamos os quatérnios Q1 e Q2 e guardamos o resultado na variável Q3. As matrizes relacionadas a estes quatérnios são, respectivamente, A1, A2 e A3. O resultado obtido pela definição para A3 é o mesmo obtido do produto das matrizes A1 e A2. As matrizes que aparecem aqui diferem das exibidas no ER 2.10, uma vez que, no OCTAVE, o termo independente do quatérnio é colocado na quarta posição do vetor.

## 2.4 POSTO DE UMA MATRIz

Antes da discussão a respeito do conceito de posto de uma matriz, é necessário definir o conceito de dependência linear.

**Definição 2.11:** Considere um conjunto de funções, vetores ou elementos $e_1$, $e_2$, ..., $e_n$. Estas funções, vetores ou elementos são ditos **linearmente dependentes** se existe um conjunto de escalares $k_1$, $k_2$, ..., $k_n$, nem todos nulos, tal que $k_1e_1 + k_2e_2 + ... + k_ne_n = 0$. Se eles não são linearmente dependentes, então se diz que eles são **linearmente independentes**.

**ER 2.11:** Mostre que as funções f(x,y,z), g(x,y,z) e h(x,y,z) são linearmente dependentes, onde $f(x,y,z) = x - y + z$, $g(x,y,z) = x + y + 2z$ e $h(x,y,z) = 3x + y + 5z$.

**Solução:** Se f(x,y,z), g(x,y,z) e h(x,y,z) são funções linearmente dependentes, então, existem escalares $k_1$, $k_2$ e $k_3$, não todos nulos, tais que
$k_1f(x,y,z) + k_2g(x,y,z) + k_3h(x,y,z) = 0$.
Considere $k_1(x - y + z) + k_2(x + y + 2z) + k_3(3x + y + 5z) = 0$. Então,

$(k_1 + k_2 + 3k_3)x + (-k_1 + k_2 + k_3)y + (k_1 + 2k_2 + 5k_3)z = 0$, e

$$\begin{cases} k_1 + k_2 + 3k_3 = 0 \\ -k_1 + k_2 + k_3 = 0 \\ k_1 + 2k_2 + 5k_3 = 0 \end{cases}$$

Resolvendo simultaneamente, adicionando a segunda equação à primeira e à terceira equações, obtém-se

$$\begin{cases} 2k_2 + 4k_3 = 0 \\ 3k_2 + 6k_3 = 0 \end{cases}$$

$k_2 = -2k_3$ e $k_1 = -k_3$. Seja $k_3 = -1$. Então, $k_2 = 2$, $k_1 = 1$, e

$1(x - y + z) + 2(x + y + 2z) - 1(3x + y + 5z) = 0$.

Portanto, f(x,y,z), g(x,y,z) e h(x,y,z) são funções linearmente dependentes.

**Definição 2.12:** Dada uma matriz A de ordem m por n, selecionando os elementos em quaisquer r linhas e r colunas de A, obtém-se **submatrizes quadradas** de ordem r. Os determinantes destas submatrizes quadradas de ordem r são chamados **r-menores** de A.

**Definição 2.13:** A ordem da maior submatriz quadrada cujo determinante tem um valor não nulo é chamada **posto** da matriz.

O posto de qualquer matriz diferente da matriz nula não pode ser zero; o posto de uma matriz m por n não pode exceder m ou n; e o posto de uma matriz nula é definido como zero.

***Em geral, o posto de uma matriz é igual ao maior número de vetores linha ou coluna linearmente independentes, uma vez que o valor do determinante de uma matriz quadrada é zero se os vetores linha ou coluna são linearmente dependentes.***

**ER 2.12:** Determine o posto da matriz A = $\begin{pmatrix} 4 & 2 & -1 & 3 \\ 0 & 5 & -1 & 2 \\ 12 & -4 & -1 & 5 \end{pmatrix}$.

**Solução:** Uma vez que a matriz A é de ordem 3 por 4, o posto de A não pode exceder 3. Cada um dos 3–menores de A, entretanto, tem valor zero; isto é,

$$\begin{vmatrix} 4 & 2 & -1 \\ 0 & 5 & -1 \\ 12 & -4 & -1 \end{vmatrix} = \begin{vmatrix} 4 & 2 & 3 \\ 0 & 5 & 2 \\ 12 & -4 & 5 \end{vmatrix} = \begin{vmatrix} 4 & -1 & 3 \\ 0 & -1 & 2 \\ 12 & -1 & 5 \end{vmatrix} = \begin{vmatrix} 2 & -1 & 3 \\ 5 & -1 & 2 \\ -4 & -1 & 5 \end{vmatrix} = 0.$$

Portanto, o posto de A não pode ser três. Uma vez que o 2–menor

$$\begin{vmatrix} 4 & 2 \\ 0 & 5 \end{vmatrix}$$

não é nulo, o posto de A é dois.

**OCTAVE 2.12**

O posto de uma matriz é calculado pelo comando `rank`:

```
> A = [4,2,-1,3;0,5,-1,2;12,-4,-1,5];
> rank(A)
> ans = 2
```

A dificuldade na determinação do posto de uma matriz de ordem mais alta pode ser reduzida considerando-se três **transformações elementares sobre linhas** na matriz:

(i) troca de duas linhas;

(ii) multiplicação dos elementos de qualquer linha por um escalar não nulo;

(iii) adição de uma linha multiplicada por um escalar não nulo à outra linha.

Dos Teoremas 2.2, 2.4 e 2.5, fica evidente que o posto de uma matriz não é alterado pela aplicação de cada uma das transformações elementares sobre linha, visto que

- por (i) apenas os sinais dos valores de alguns r–menores são alterados;
- por (ii) os valores de alguns r–menores são multiplicados por um escalar não nulo;
- por (iii) o valor de cada r–menor é inalterado.

Estas transformações elementares sobre linhas podem ser aplicadas em qualquer sequência.

**Definição 2.14:** Uma matriz é considerada na **forma linha escalonada** se, por meio de transformações elementares sobre linhas, o primeiro elemento não nulo em cada linha for igual a 1 e estiver posicionado à direita do primeiro elemento não nulo da linha antecedente.

Se uma matriz estiver na forma linha escalonada equivalente, seu posto pode facilmente ser determinado.

**ER 2.13:** Determine o posto da matriz $A = \begin{pmatrix} 0 & 2 & 4 & 6 \\ 3 & -1 & 4 & -2 \\ 6 & -1 & 10 & -1 \end{pmatrix}$.

**Solução:** Trocando a primeira com a segunda linha,

$$\begin{pmatrix} 3 & -1 & 4 & -2 \\ 0 & 2 & 4 & 6 \\ 6 & -1 & 10 & -1 \end{pmatrix};$$

multiplicando os elementos da primeira linha por (–2) e adicionando estes produtos aos elementos correspondentes da terceira linha,

$$\begin{pmatrix} 3 & -1 & 4 & -2 \\ 0 & 2 & 4 & 6 \\ 0 & 1 & 2 & 3 \end{pmatrix};$$

multiplicando os elementos da segunda linha por (–1/2) e adicionando estes produtos aos elementos correspondentes da terceira linha,

$$\begin{pmatrix} 3 & -1 & 4 & -2 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

multiplicando os elementos da primeira e da terceira linha por (1/3) e (1/2), respectivamente:

$$\begin{pmatrix} 3 & -1 & 4 & -2 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

A matriz A está agora, na forma linha escalonada, e pode-se verificar que todo 3–menor tem valor zero. Assim, o posto de A é dois, pois

$$\begin{vmatrix} 1 & -\frac{1}{3} \\ 0 & 1 \end{vmatrix} \neq 0.$$

Cada uma das transformações elementares sobre linhas numa matriz pode ser vista como uma pré-multiplicação da matriz por uma matriz adequada, obtida pelo desenvolvimento da transformação elementar sobre linha, sobre a matriz identidade.

**Exemplo 2.9:** Se a matriz cujo posto a ser determinado é de ordem 3 x n, então a primeira transformação elementar sobre linha (i) pode ser realizada multiplicando-se a dada matriz por uma das matrizes

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ ou } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \qquad (2.14)$$

A segunda transformação elementar (ii) pode ser realizada pela pré-multiplicação da matriz dada por uma das matrizes

$$\begin{pmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ ou } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{pmatrix}. \qquad (2.15)$$

A terceira operação elementar (iii) pode ser realizada pela pré-multiplicação da matriz dada por uma das matrizes

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

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{pmatrix} \text{ ou } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{pmatrix}. \tag{2.16}$$

**Definição 2.15:** As matrizes que permitem o desenvolvimento das transformações elementares são chamadas **matrizes de transformação elementar sobre linhas**.

**Exemplo 2.10:** As matrizes de (2.14), (2.15) e (2.16) são matrizes de transformação elementar sobre linhas, de ordem 3. Note que toda matriz de transformação elementar sobre linha é não singular.

É possível mostrar que toda matriz não singular é produto de matrizes de transformação elementar sobre linha. Então, toda matriz não singular A tem uma inversa não singular $A^{-1}$, que é igual ao produto de matrizes de transformação elementar sobre linhas. O produto das matrizes de transformação elementar sobre linha que transformam A em I é $A^{-1}$. Note que as matrizes usadas de transformação elementar sobre linha devem ser multiplicadas em ordem reversa de sua aplicação.

**ER 2.14:** Determine a inversa de $A = \begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix}$.

**Solução:** Uma vez que A é não singular, serão aplicadas transformações elementares sobre linhas de A, até que se obtenha a matriz I. O produto das matrizes que representam as transformações deve ser igual a $A^{-1}$. Por exemplo, multiplicando os elementos da primeira linha por (–4/3) e adicionando os produtos aos elementos correspondentes da segunda linha, obtém-se

$$\begin{pmatrix} 1 & 0 \\ -\frac{4}{3} & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix}=\begin{pmatrix} 3 & 1 \\ 0 & \frac{2}{3} \end{pmatrix};$$

multiplicando os elementos da primeira linha por (1/3) e os elementos da segunda linha por (3/2),

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

multiplicando os elementos da segunda linha por (–1/3) e adicionando os produtos aos elementos correspondentes da primeira linha,

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

Portanto,

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

## 2.5 SISTEMAS DE EQUAÇÕES LINEARES

Considere um sistema de m equações lineares em n variáveis $x_1, x_2, \ldots, x_n$:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = c_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = c_m \end{cases}. \qquad (2.17)$$

O sistema (2.17) pode ser escrito em forma matricial como

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}. \tag{2.18}$$

**Definição 2.16:** A matriz $(a_{ij})$ de ordem m por n é chamada **matriz dos coeficientes**.

**Definição 2.17:** A matriz composta dos mn elementos $a_{ij}$ mais uma coluna adicional cujos elementos são as constantes $c_i$ é chamada **matriz aumentada** do sistema.

**Exemplo 2.10:** A matriz

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & c_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & c_m \end{pmatrix} \tag{2.19}$$

é a matriz aumentada de (2.17).

***O posto da matriz aumentada é maior ou igual ao posto da matriz dos coeficientes.***

A solução, caso exista, para um sistema de equações lineares (2.17) pode ser determinada pela transformação da matriz aumentada numa forma escalonada por meio de transformações elementares sobre linha. O sistema de equações lineares, representado pela matriz aumentada transformada, será um sistema equivalente a (2.17), onde “equivalente” significa que o conjunto solução não é alterado.

**Definição 2.18:** Se o sistema de equações lineares (2.17) é tal que as equações são todas satisfeitas simultaneamente por, no mínimo, um conjunto de valores das variáveis, então ele é dito **consistente;** e o sistema é dito **inconsistente** se as equações não são satisfeitas simultaneamente por qualquer conjunto de valores das variáveis.

**ER 1.15:** Resolva o sistema de equações lineares por meio de transformações elementares sobre linhas na matriz aumentada:

$$\begin{cases} 2x - y + z = 1 \\ x + y - z = 2 \\ 3x - y + z = 0 \end{cases}$$

**Solução:** A matriz aumentada é

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

Trocando a primeira e a segunda linha, obtém-se:

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

Multiplicando os elementos da primeira linha por (−2) e adicionando os produtos aos elementos correspondentes da segunda linha, obtém-se:

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

Multiplicando os elementos da primeira linha por (−3) e adicionando os produtos aos elementos correspondentes da terceira linha,

$$\begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & -3 & 3 & -3 \\ 0 & -4 & 4 & -6 \end{pmatrix}.$$

Multiplicando os elementos da linha dois por (–1/3),

$$\begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & -4 & 4 & -6 \end{pmatrix}.$$

Multiplicando os elementos da segunda linha por (4) e adicionando os produtos aos elementos correspondentes da terceira linha, obtém-se:

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

Multiplicando os elementos da terceira linha três (–1/2):

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

A matriz aumentada transformada está na forma escalonada e representa o sistema de equações lineares a seguir:

$$\begin{cases} x + y - z = 2 \\ \quad y - z = 1 \\ \qquad 0 = 1 \end{cases}$$

Uma vez que nenhum dos valores x, y e z existem tais que 0 = 1, o sistema é inconsistente. Note que o posto da matriz aumentada é três enquanto o posto da matriz dos coeficientes é dois.

**OCTAVE 2.15**

O OCTAVE tem um comando especial (`rref`) que determina a forma escalonada de uma matriz, porém com um nível posterior, onde executa operações extras procurando zerar os outros elementos da coluna onde se encontram os pivôs de cada linha ( $\boxed{1}$ ). Observe o efeito do comando `rref` sobre a matriz do ER 2.15:

```
> A = [2,-1,1,1;1,1,-1,2;3,-1,1,0];
> rref(A)
> ans =
```

```
1     0     0     0
0     1    −1     0
0     0     0     1
```

Sobre a matriz $\begin{pmatrix} \boxed{1} & 1 & -1 & 2 \\ 0 & \boxed{1} & -1 & 1 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$ no ER 2.15 são executadas as seguintes operações:

a) multiplicamos a 2ª linha por –1 e adicionamos à primeira linha, obtendo a matriz $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

b) multiplicamos a 3ª linha por –1 e adicionamos às primeira e segunda linhas, obtendo a matriz $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, que é a matriz exibida pela variável temporária `ans`.

Neste caso o sistema de equações lineares representado por esta matriz é $\begin{cases} x = 0 \\ y - z = 0 \\ 0 = 1 \end{cases}$ (inconsistente).

**ER 2.16:** Resolva o sistema de equações lineares seguinte, por meio de transformações elementares sobre linhas, na matriz aumentada:

$$\begin{cases} x + 3y + z = 6 \\ 3x - 2y - 8z = 7 \\ 4x + 5y - 3z = 17 \end{cases}$$

**Solução:** A matriz aumentada é

$$\begin{pmatrix} 1 & 3 & 1 & 6 \\ 3 & -2 & -8 & 7 \\ 4 & 5 & -3 & 17 \end{pmatrix}$$

Multiplicando os elementos da primeira linha por (−3) e adicionando os produtos aos elementos correspondentes da segunda linha, obtém-se:

$$\begin{pmatrix} 1 & 3 & 1 & 6 \\ 3 & -11 & -11 & -11 \\ 4 & 5 & -3 & 17 \end{pmatrix}.$$

Multiplicando os elementos da primeira linha por (−4) e adicionando os produtos aos elementos correspondentes da terceira linha,

$$\begin{pmatrix} 1 & 3 & 1 & 6 \\ 3 & -11 & -11 & -11 \\ 0 & -7 & -7 & -7 \end{pmatrix}.$$

Multiplicando os elementos da segunda linha por (−1/11):

$$\begin{pmatrix} 1 & 3 & 1 & 6 \\ 3 & 1 & 1 & 1 \\ 0 & -7 & -7 & -7 \end{pmatrix}.$$

Multiplicando os elementos da segunda linha por (7) e adicionando os produtos aos elementos correspondentes da terceira linha, obtém-se:

$$\begin{pmatrix} 1 & 3 & 1 & 6 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

A matriz aumentada transformada está na forma escalonada e representa o sistema de equações lineares a seguir:

$$\begin{cases} x + 3y + z = 6 \\ \quad y + z = 1 \\ \quad\quad 0 = 0 \end{cases}$$

Portanto, existe um número infinito de soluções da forma x = 3 + 2z e y = 1 – z. O valor da variável z pode ser atribuído arbitrariamente, encontrando-se, assim, os valores de x e y.

**OCTAVE 2.16**

Utilizamos novamente o comando `rref` sobre a matriz aumentada A:

```
> A = [1,3,1,6;3,-2,-8,7;4,5,-3,17];
> rref(A)
> ans =
  1   0  -2   3
  0   1   1   1
  0   0   0   0
```

Sobre a matriz $\begin{pmatrix} 1 & 3 & 1 & 6 \\ 0 & \boxed{1} & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ no ER 2.16 o OCTAVE efetua uma operação extra no intuito de zerar os demais elementos da coluna do pivô da segunda linha,

multiplicando a 2º linha por –3 e adicionando à primeira linha, obtendo a matriz $\begin{pmatrix} 1 & 0 & -2 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Neste caso o sistema de equações lineares representado por esta matriz é $\begin{cases} x \quad -2z=3 \\ \quad y+z=1 \end{cases}$ que é equivalente ao sistema do ER 2.16.

Note que, neste exemplo, o posto da matriz aumentada e o posto da matriz dos coeficientes são ambos iguais a dois. Em geral, uma condição necessária e suficiente para um sistema de equações lineares ser consistente é que a matriz aumentada e a matriz dos coeficientes tenham o mesmo posto.

Seguem algumas proposições gerais a respeito de sistemas lineares de m equações em n variáveis, que serão aceitas sem demonstração.

- Se o sistema (2.17) é consistente no caso de M > n (mais equações que variáveis), as equações são necessariamente linearmente dependentes.

- Se o posto da matriz aumentada é n, existe uma única solução.
- Se o posto r da matriz aumentada é menor que n, a solução pode ser obtida pela escolha de valores arbitrários para n – r das variáveis e, a partir destes, podem ser determinados os valores das variáveis restantes; isto é, existe um número infinito de soluções.
- Quando m < n, os valores para no mínimo n – m variáveis necessitam ser escolhidos arbitrariamente, se o sistema é consistente.

**Definição 2.19:** Um sistema de equações lineares homogêneas é da forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = 0 \\ ...................................... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = 0 \end{cases} \qquad (2.20)$$

O sistema (2.20) pode ser expresso em forma matricial como AX = 0, onde A é a matriz dos coeficientes, X é a matriz coluna das variáveis $x_1$, $x_2$, ..., $x_n$, e 0 é a matriz coluna nula de ordem m por 1.

Tal sistema é necessariamente consistente, uma vez que ele tem a **solução trivial**

$$x_1 = x_2 = ... = x_n = 0.$$

Se existe uma solução **não trivial**, um número infinito de soluções existe; isto é, se $t_1$, $t_2$, ..., $t_n$ representa uma solução não trivial, então $kt_1$, $kt_2$, ..., $kt_n$ para qualquer número real k é também uma solução.

Esta propriedade de um sistema de equações homogêneas pode ser provada por substituição direta em (2.20).

# 3. TRANSFORMAÇÕES DO PLANO

**AS OPERAÇÕES REALIZADAS** no plano cartesiano ortogonal são de grande interesse, principalmente quando relacionadas aos recursos gráficos computacionais. São muito comuns os problemas que envolvem o movimento planar, manipulação, rotação e transformação de figuras. Na animação gráfica, as transformações mais interessantes são as que correspondem ao que denominamos movimento rígido, ou seja, as que preservam, de alguma maneira, a forma e o tamanho dos objetos.

## 3.1 APLICAÇÕES

De modo simplificado, em matemática, uma função é considerada como uma regra que associa cada elemento de um conjunto A a um elemento de um conjunto B, onde A e B são usualmente subconjuntos do conjunto de números reais.

**Exemplo 3.1:** Se A é constituído do conjunto de números reais x, e B é composto por um conjunto de números reais não negativos y, então $y = x^2$ representa uma função; isto é, para cada $x \in A$, a regra $y = x^2$ associa um elemento $y \in B$ onde y é o quadrado de x.

Nesta seção, será introduzida uma terminologia adicional para o conceito de função. Esta terminologia é baseada em conceitos geométricos e favorece a discussão dos problemas tratados neste capítulo.

**Definição 3.1:** Uma **aplicação de valor singular** T de um conjunto A *em* um conjunto B, é uma regra que associa cada elemento $a \in A$ a um único elemento $b \in B$. O elemento b é chamado **imagem** de a sob a aplicação T e é denotado por $b = T(a)$.

**Exemplo 3.2:** Considere dois conjuntos, A = {m,n} e B = {x,y,z}. Seja T a regra que associa m e n com x e z, respectivamente; então, T é uma aplicação de valor singular de A em B tal que $x = T(m)$ e $z = T(n)$. Um diagrama desta aplicação é exibido na Figura 3.1.

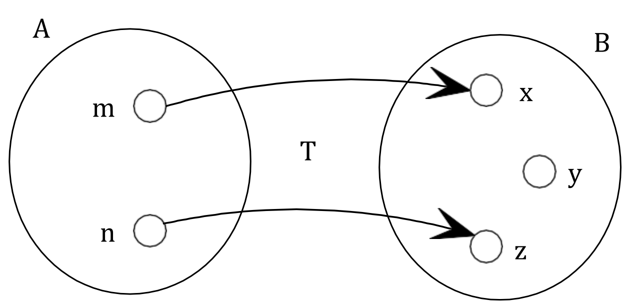



FIGURA 3.1 - Aplicação de valor singular

A Figura 3.2 ilustra uma situação que *não* pode ser descrita como uma aplicação de valor singular de A em B, pois nenhum elemento de A pode ter mais que uma imagem.

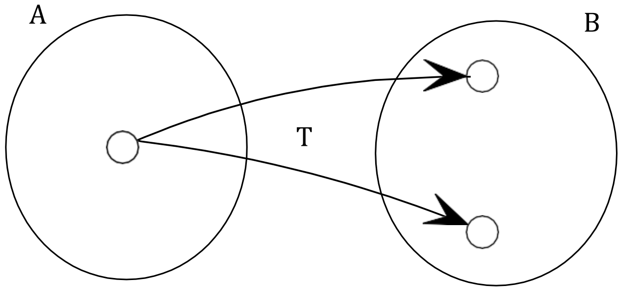



FIGURA 3.2 - Contraexemplo de aplicação de valor singular

**Definição 3.2:** Se T é uma aplicação de valor singular de A em B, tal que todo elemento de B é a imagem de algum elemento de A, então, T é uma aplicação de valor singular *sobrejetora* de A em B.

**Exemplo 3.3:** Considere dois conjuntos, A = {1,2,3} e B = {a,b,c}. Seja T a regra que associa a, b e c com 2, 1 e 3, respectivamente (Figura 3.3), então T é uma aplicação de valor singular sobrejetora de A em B.

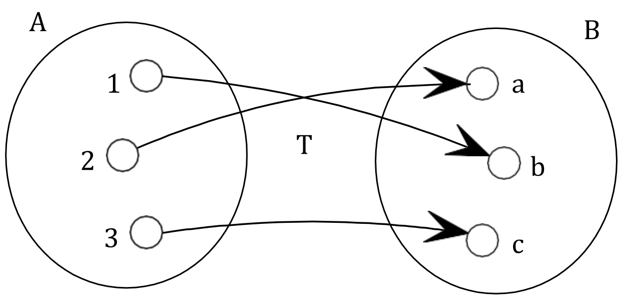



FIGURA 3.3 - Aplicação sobrejetora

**Exemplo 3.4:** Seja A = {1,2,3}, B = {m,n}, e T uma regra tal que T(1) = m, T(2) = m, e T(3) = n (Figura 3.4). Note que, enquanto cada elemento de B é imagem de algum elemento de A, m é imagem de 1 e 2 ao mesmo tempo. A aplicação T é, apesar disso, uma aplicação de valor singular sobrejetora de A em B.

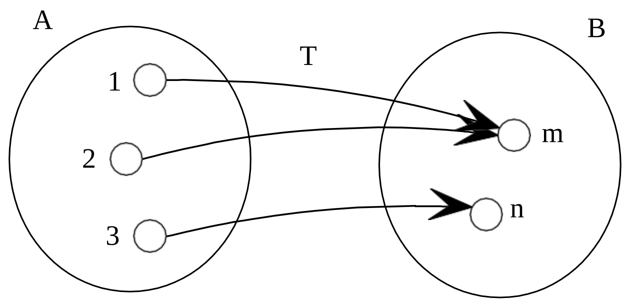



FIGURA 3.4 - Aplicação sobrejetora onde T(1)=T(2)

Entretanto, existe uma distinção entre as duas aplicações de valor singular sobrejetoras ilustradas nas Figuras 3.3 e 3.4. A aplicação da Figura 3.3 é chamada **aplicação um a um** de A sobre B e é um tipo especial de aplicação. Toda aplicação um a um de A sobre B é uma aplicação de valor singular sobrejetora de A em B, mas nem toda aplicação de valor singular sobrejetora de A em B é uma aplicação um a um de A sobre B. A aplicação na Figura 3.4 *não* é uma aplicação um a um de A sobre B.

Se T é uma aplicação um a um de A sobre B, então cada elemento de B é a imagem de exatamente um elemento de A. Portanto, é possível definir uma aplicação *inversa* de B sobre A.

**Definição 3.3:** Seja T uma aplicação um a um de A em B. A **aplicação inversa de B em A**, denotada por $T^{-1}$, associa a cada elemento $b \in B$ o elemento $a \in A$, que tem b como sua imagem sob T, isto é, $T^{-1}(T(a)) = a$ para todo $a \in A$.

**ER 3.1:** Sejam N o conjunto de inteiros, $n \in N$ e T uma aplicação de valor singular de N em N. Determine se T é uma aplicação sobrejetora de N em N, onde,

(a) $T(n) = n + 1$

(b) $T(n) = 2n + 1$

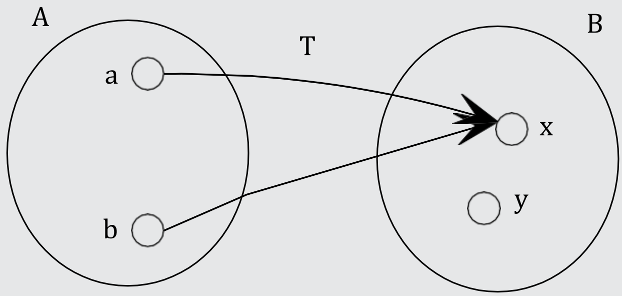
**Solução:**

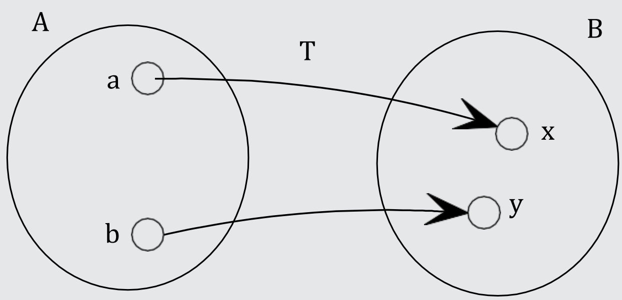
(a) T é uma aplicação de N sobre N, uma vez que cada inteiro é a única imagem de outro inteiro, uma unidade a menos. Além disso, T é uma aplicação um a um de N sobre N, pois cada inteiro é a imagem de um, e apenas um inteiro.

(b) T não é uma aplicação de N sobre N, pois nenhum inteiro par é a imagem de outro inteiro sob a aplicação; isto é, 2n + 1, não pode ser par para qualquer inteiro n.

**ER 3.2:** Seja T uma aplicação de valor singular de A em B onde A = {a,b} e B = {x,y}. Defina as possíveis aplicações.

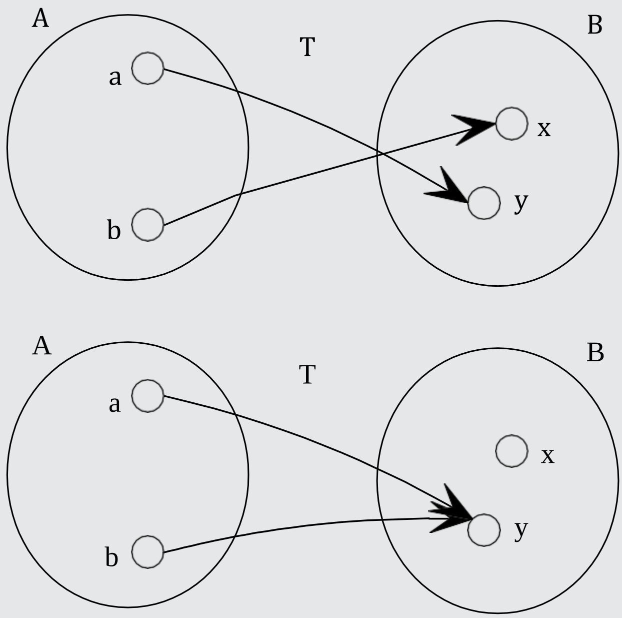
**Solução:** Existem quatro possíveis aplicações T de A em B:




(i) T(a) = x e T(b) = x

(ii) T(a) = x e T(b) = y




(iii) $T(a) = y$ e $T(b) = x$

(iv) $T(a) = y$ e $T(b) = y$

Note que as aplicações T, descritas em (ii) e (iii), são aplicações um a um sobrejetoras de A em B.

Algumas aplicações importantes do conjunto de pontos do plano serão ilustradas nas seções seguintes deste capítulo.

## 3.2 ROTAÇÕES

Algumas vezes, é necessário ou desejável, considerar uma aplicação do conjunto dos pontos de um plano nele mesmo, de modo a simplificar a análise de um dado problema. Tal aplicação é chamada **transformação ponto** ou **transformação do plano**, o que pode ser formalizado conforme a definição seguinte.

**Definição 3.4:** Uma **transformação do plano** é uma regra segundo a qual cada ponto P no plano é transformado ou *aplicado* sobre um ponto P' no mesmo plano.

Uma transformação do plano pode ser considerada como um movimento de pontos no plano com respeito à posição de seus respectivos pontos de imagem. Cada transformação do plano pode ser descrita algebricamente como um conjunto de equações que relacionam as coordenadas cartesianas retangulares de cada ponto no plano às coordenadas de sua imagem. Existem vários tipos de transformações do plano, tais como rotações, reflexões, dilatações, expansões, cisalhamentos, projeções e inversões, dentre outras.

Nesta seção e nas seções seguintes, algumas das mais importantes transformações do plano são consideradas e representadas em forma matricial.

**Exemplo 3.5:** Considere uma **rotação do plano** sobre a origem sob um ângulo θ, mantendo o mesmo sistema de referência de coordenadas cartesianas retangulares. Seja P (x,y) qualquer ponto no plano, então, $(x,y) = x\vec{i} + y\vec{j}$, onde $\vec{i} = (1,0)$ e $\vec{j} = (0,1)$ são vetores unitários. Sob uma rotação do plano sobre a origem, cada ponto P é aplicado sobre o ponto P': (x',y') como na Figura 3.5. Note que $|\overrightarrow{OP}| = |\overrightarrow{OP'}|$. Então,

$$\overrightarrow{OP} = x\vec{i} + y\vec{j} = |\overrightarrow{OP}|\cos\phi\vec{i} + |\overrightarrow{OP}|\,\text{sen}\phi\vec{j} = |\overrightarrow{OP'}|\cos\phi\vec{i} + |\overrightarrow{OP'}|\,\text{sen}\phi\vec{j},$$

e

$$\begin{aligned}\overrightarrow{OP'} &= x'\vec{i} + y'\vec{j} = |\overrightarrow{OP'}|\cos(\theta+\phi)\,\vec{i} + |\overrightarrow{OP'}|\,\text{sen}(\theta+\phi)\,\vec{j}\\ &= |\overrightarrow{OP'}|\,(\cos\theta\cos\phi - \text{sen}\theta\,\text{sen}\phi)\,\vec{i} + |\overrightarrow{OP'}|\,(\text{sen}\theta\cos\phi + \cos\theta\,\text{sen}\phi)\,\vec{j}\\ &= (x\cos\theta - y\,\text{sen}\theta)\,\vec{i} + (x\,\text{sen}\theta + y\cos\theta)\,\vec{j}.\end{aligned}$$

Portanto,

$$\begin{cases} x' = x\cos\theta - y\,\text{sen}\theta \\ y' = x\,\text{sen}\theta + y\cos\theta \end{cases} \tag{3.1}$$

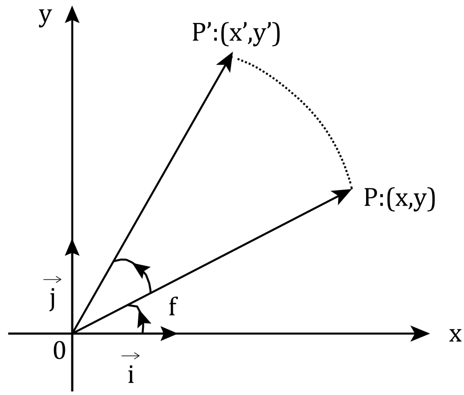



FIGURA 3.5 - Rotação do plano

A relação entre as coordenadas de cada ponto P e sua imagem P' pode ser representada na forma matricial como

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \qquad (3.2)$$

**Definição 3.5:** A matriz $\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$ da transformação rotação definida por (3.2) é chamada **matriz de rotação**. (3.3)

A matriz de rotação (3.3) pode ser considerada como um *operador* que leva cada ponto (x,y) no plano sobre sua imagem

$$\left(x\cos\theta - y\operatorname{sen}\theta, x\operatorname{sen}\theta + y\cos\theta\right)$$

onde o plano é rotacionado sobre a origem sob um ângulo θ. Quando considerado como um operador, a matriz da forma (3.3) pode ser usada para representar, descrever, ou caracterizar uma rotação do plano sobre a origem sob um ângulo θ. Note que o determinante de toda matriz de rotação é igual a um. Além disso, tanto os vetores linha quanto os vetores coluna de uma matriz de rotação podem ser considerados como um par de vetores unitários ortogonais.

**ER 3.3:** Determine as coordenadas da imagem do ponto P(5, $\sqrt{3}$) após uma rotação do plano sobre a origem sob um ângulo de 30º (Figura 3.6).

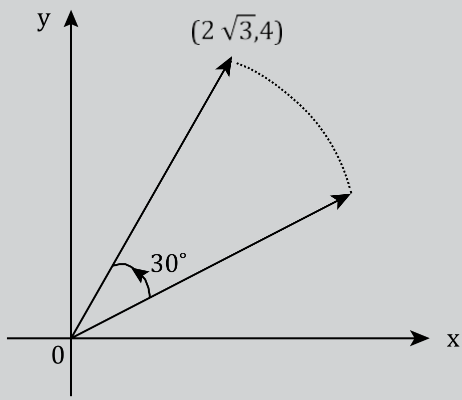



FIGURA 3.6 - Rotação do plano em 30º

**Solução:** O ângulo de rotação é θ = 30º, cosθ = $\frac{\sqrt{3}}{2}$ e senθ = $\frac{1}{2}$. Portanto, por (3.3), a matriz de rotação para a transformação do plano é

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

Assim, pela transformação (3.2), a imagem (x',y') do ponto P (5, $\sqrt{3}$) é dada por:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 5 \\ \sqrt{3} \end{pmatrix} = \begin{pmatrix} 2\sqrt{3} \\ 4 \end{pmatrix};$$

isto é, as coordenadas da imagem de P (5, $\sqrt{3}$) sob uma rotação do plano sobre a origem sob um ângulo de 30º são (2$\sqrt{3}$,4).

**OCTAVE 3.3**

Utilizando a matriz de rotação indicada no ER 3.3, basta multiplicá-la pelo vetor de coordenadas de interesse:

```
> R = [sqrt(3)/2,-1/2;1/2,sqrt(3)/2];
> P = [5;sqrt(3)];
> R*P
ans =

   3.4641
   4.0000
```

Observe que $\sqrt{3}$ é calculado por `sqrt(3)`. O resultado exibido são as coordenadas de interesse, com 4 casas decimais.

**ER 3.4:** Determine a equação satisfeita pelo conjunto dos pontos obtidos pela rotação do plano sobre a origem sob um ângulo de 45º, da equação $x^2 - 2xy + y^2 - \sqrt{2}x - \sqrt{2}y = 0$ (Figura 3.7).

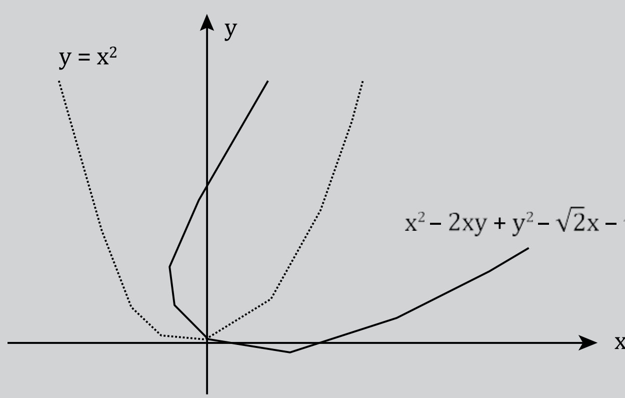



FIGURA 3.7 - Rotação de parábola em 45°

**Solução:** Uma vez que o ângulo de rotação é θ = 45º, cosθ = $\frac{\sqrt{2}}{2}$ e senθ = $\frac{\sqrt{2}}{2}$. O ponto imagem (x',y') de cada ponto (x,y) é dado pela equação matricial:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Então,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Trocando x por $\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y'$ e y por $-\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y'$ em

$x^2 - 2xy + y^2 - \sqrt{2}x - \sqrt{2}y = 0$, a equação satisfeita pelo conjunto de pontos imagens torna-se

$$\left(\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y'\right)^2 - 2\left(\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y'\right)\left(-\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y'\right) + \left(-\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y'\right)^2 -$$
$$-\sqrt{2}\left(\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y'\right) - \sqrt{2}\left(-\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y'\right) = 0$$

Simplificando, y' = $(x')^2$; isto é, y = $x^2$, pois (x',y') são as coordenadas de cada ponto imagem com referência ao eixo–xy.

**Definição 3.6:** Um ponto é chamado **ponto fixo** sob uma transformação do plano, se ele é levado nele mesmo, ou seja, T(x,y) = (x,y).

Note que, quando $\theta \neq k360^{o}$, para qualquer inteiro k, a origem é o único ponto fixo sob uma rotação do plano em torno da origem sob ângulo $\theta$, $R(\theta) \neq I$. Quando $\theta = k.360^{o}$ para algum inteiro k, todo ponto do plano é um ponto fixo, e a matriz de rotação (3.3) é igual à matriz identidade de ordem 2.

**Definição 3.7:** Uma rotação do plano sobre a origem representada por uma matriz identidade é chamada **transformação identidade**.

Ocasionalmente é necessário aplicar uma transformação do plano sobre outra.

**Definição 3.8:** Se $T_1$ e $T_2$ são matrizes que representam transformações, então a **transformação produto**, denotada por $T_2T_1$, é a aplicação de $T_1$ seguida pela aplicação de $T_2$, e é definida como

$$T_2T_1\begin{pmatrix} x \\ y \end{pmatrix} = T_2\left[T_1\begin{pmatrix} x \\ y \end{pmatrix}\right]. \tag{3.4}$$

A rotação do plano em torno da origem sob um ângulo $\theta$ é uma aplicação um-pra-um do conjunto dos pontos do plano nele mesmo. Portanto, existe uma transformação inversa para cada rotação. Considere uma rotação do plano em torno da origem sob um ângulo $(-\theta)$. Conforme (3.3), a matriz de rotação desta transformação é

$$\begin{pmatrix} \cos(-\theta) & -\operatorname{sen}(-\theta) \\ \operatorname{sen}(-\theta) & \cos(-\theta) \end{pmatrix};$$

isto é,

$$\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \tag{3.5}$$

uma vez que $\cos(-\theta) = \cos\theta$ e $\operatorname{sen}(-\theta) = -\operatorname{sen}\theta$. Então, como

$$\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

uma rotação do plano em torno da origem sob um ângulo $(-\theta)$ é a transformação inversa de uma rotação do plano em torno da origem sob um ângulo $\theta$.

## 3.3 REFLEXÕES, DILATAÇÕES E EXPANSÕES

Nesta seção, alguns tipos adicionais de transformações do plano são apresentados. Considere as transformações do plano representadas pelas matrizes

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \text{ e } \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.6}$$

Estas matrizes levam cada ponto (x,y) do plano nos pontos (–x,y) e (x,–y), respectivamente.

**Definição 3.9:** As aplicações um a um do plano nele mesmo, representadas pelas matrizes (3.6), são chamadas **reflexões do plano**.

Cada ponto do plano é levado em seu "simétrico" com respeito a um dos eixos coordenados pelas **matrizes de reflexão** (3.6). Note que sob a reflexão descrita por

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

os pontos do eixo–y são pontos fixos, e sob a reflexão descrita por

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

os pontos do eixo–x são fixos.

Outro par de matrizes que representam reflexões do plano é dado por

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ e } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.7}$$

A matriz $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ leva cada ponto (x,y) sobre (y,x); isto é, cada ponto do plano é levado em seu simétrico com respeito à reta y = x. A matriz $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ leva cada ponto (x,y) sobre (–y, –x); isto é, cada ponto do plano é levado em seu simétrico com respeito à reta y = –x.

Quanto às matrizes de reflexão dadas por (3.6) e (3.7), têm-se:[1]

- o produto entre quaisquer duas destas matrizes é uma matriz de rotação;
- o determinante de cada uma destas matrizes é igual a –1.

Além disso, o produto entre quaisquer duas reflexões do plano, no que se refere à interseção de retas que passam pela origem, é uma rotação do plano sobre a origem.

As matrizes de reflexão de (3.7) podem ser consideradas como produto das matrizes de reflexão de (3.6) e das matrizes de rotação da forma (3.3).

**Exemplo 3.6:**

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \tag{3.8}$$

$$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \tag{3.9}$$

Existem matrizes de reflexão diferentes daquelas de (3.6) e (3.7), as quais, levando cada ponto do plano em seu simétrico, com respeito a uma reta r, são iguais a

1 Sugestão: verifique isto como exercício.

$$T^{-1}RT, \tag{3.10}$$

onde T é uma matriz de transformação que leva a reta r sobre o eixo–x, e R é uma matriz que representa a reflexão do plano com respeito ao eixo–x (Figura 3.8).

Em (3.8) e (3.9), a transformação (3.10) foi utilizada para expressar as matrizes que representam reflexões do plano com respeito às retas y = x e y = –x, respectivamente. Em (3.8), T é uma matriz que representa uma rotação do plano sobre a origem sob um ângulo $\theta = -45º$, e em (3.9), T é uma matriz que representa uma rotação do plano sobre a origem sob um ângulo $\theta = 45º$.

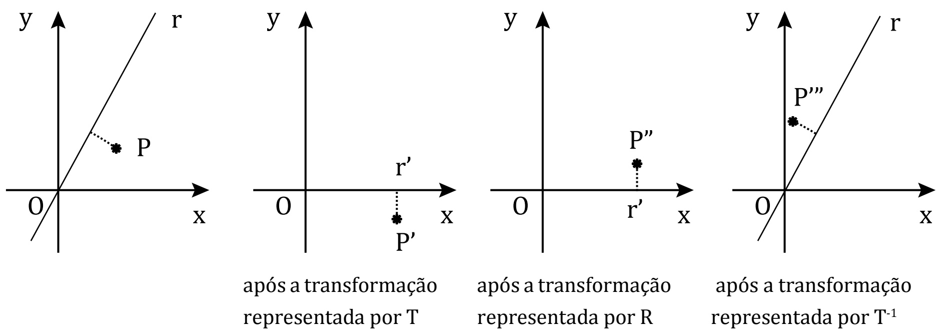



FIGURA 3.8 - Sequência de transformações no plano

Se T é uma matriz de transformação que leva a reta r sobre o eixo–y, então, em (3.10), R é a matriz que representa uma reflexão do plano com respeito ao eixo–y. Outra aplicação de (3.10) é considerada a seguir.

**ER 3.5:** Determine a matriz de reflexão F que leva cada ponto (x,y) do plano sobre seu simétrico no que se refere à reta $y = \sqrt{3}\,x$. Determine a imagem do ponto P $(\sqrt{3},1)$ sob a reflexão do plano representado por F (Figura 3.9).

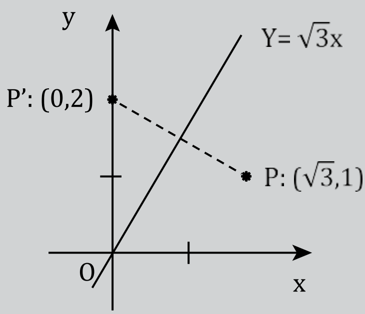



FIGURA 3.9 - Reflexão de ponto

**Solução:** Uma rotação do plano sobre a origem sob um ângulo θ = 30º leva à reta y = $\sqrt{3}x$ sobre o eixo–y. Tal rotação do plano pode ser representada pela matriz

$$T = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}, \text{ cuja inversa é } T^{-1} = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}.$$

A matriz R que representa uma reflexão do plano com respeito ao eixo–y é

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

Portanto, conforme (3.10), a matriz de reflexão F que leva cada ponto (x,y) do plano em seu simétrico com respeito à reta y = $\sqrt{3}$ x é dada por

$$F = T^{-1}RT = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix};$$

isto é,

$$F = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}.$$

A imagem do ponto P, $(\sqrt{3},1)$ sob a reflexão do plano representado por F, é P'(0,2), pois

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

💻 OCTAVE 3.5

Definimos a matriz T, calculamos a sua inversa $T^{-1}$, e entramos com a matriz de reflexão, R, em relação ao eixo–y. Calculamos a matriz de transformação $F = T^{-1}RT$, e finalmente aplicamos à coordenada de interesse:

```
> T = [sqrt(3)/2,-1/2;1/2,sqrt(3)/2];
> T_inv = inv(T);
> R = [-1,0;0,1];
> F = T_inv*R*T;
> F*[sqrt(3);1]
ans =

   3.3307e-016
   2.0000e+000
```

Observe que, por consequência da aproximação dos cálculos efetuados no computador, o resultado é dado em notação científica, e as coordenadas parecem estranhas. Porém, considerando o resultado `3.3307e-016 = 0.000000000003307` como zero, e `2.0000e+000` como 2, temos exatamente as coordenadas (0,2) encontradas previamente.

Outra classe importante de transformações do plano pode ser representada pelas matrizes escalares de ordem dois $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$, onde k > 0. O efeito desta transformação é levar cada ponto P do plano sobre um ponto P’ no plano, tal que $OP' = kOP'$.

Se k > 1, a matriz representa um “alongamento uniforme” do plano, e se 0 < k < 1, a matriz representa uma “compressão uniforme” do plano.

**Definição 3.10:** Uma transformação do plano representada por uma matriz $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ é chamada **dilatação do plano** (k > 0).

As transformações representadas pelas matrizes diagonais de ordem dois $\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$, onde $d_1 > 0$ e $d_2 > 0$, estão intimamente relacionadas com as dilatações do plano. Muitas figuras geométricas planas são distorcidas sob estas transformações.

**Definição 3.11:** Uma transformação do plano representada por uma matriz do tipo $\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$ é chamada **expansão do plano** ($d_1 > 0$ e $d_2 > 0$).

O conjunto de dilatações do plano é um subconjunto do conjunto de expansões do plano. Note que uma expansão do plano é uma aplicação um a um do conjunto de pontos do plano nele mesmo.

**ER 3.6:** Determine o efeito da dilatação do plano representada por $\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$ sobre a reta y = 3x + 2.

**Solução:** Sob a dilatação do plano, cada ponto (x,y) no plano é levado a (x’,y’), onde

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Portanto,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Isto é, $x = \frac{1}{4}x'$ e $y = \frac{1}{4}y'$. Portanto, sob a dada dilatação do plano, a reta $y = 3x + 2$ é levada sobre a reta $y' = 3x' + 8$; isto é, $y = 3x + 8$. Note que a imagem da reta dada é uma reta com a mesma inclinação.

**ER 3.7:** Determine o efeito da expansão do plano representada por $\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ sobre a elipse $\frac{x^2}{9} + \frac{y^2}{1} = 1$ (Figura 3.10).

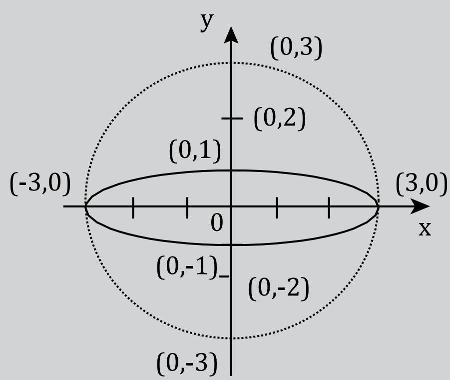



FIGURA 3.10 - Expansão sobre elipse

**Solução:** Sob a expansão do plano, cada ponto (x,y) do plano é levado sobre o ponto (x’,y’) onde

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Portanto,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix};$$

isto é, x = x' e y = $\frac{1}{3}$y'. Assim, sob a expansão do plano representada por $\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$, a elipse $\frac{x^2}{9} + \frac{y^2}{1} = 1$ é observada no círculo $\frac{(x')^2}{9} + \frac{(y')^2}{9} = 1$; isto é, $x^2 + y^2 = 9$.

## 3.4 OUTRAS TRANSFORMAÇÕES

As transformações do plano discutidas até aqui são exemplos de aplicações um a um do plano nele mesmo. Os exemplos seguintes são apresentados para considerar algumas transformações interessantes do plano, duas das quais não são aplicações um a um.

**ER 3.8:** A transformação do plano representada por uma matriz da forma $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ é chamada **cisalhamento paralelo ao eixo–x**. Determine o efeito do cisalhamento (Figura 3.11) paralelo ao eixo–x representado por $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ sobre o retângulo com vértices em (0,0), (2,0), (2,1) e (0,1).

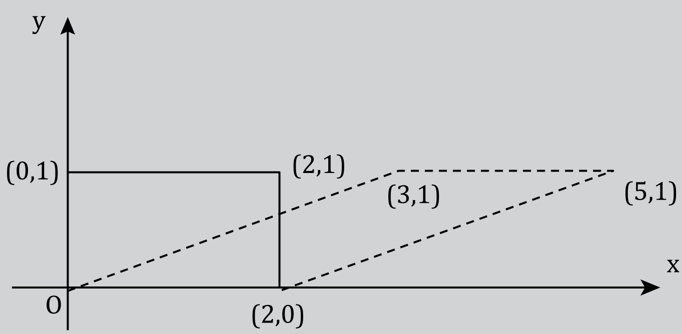



FIGURA 3.11 - Cisalhamento paralelo ao eixo-x

**Solução:** Sob o cisalhamento paralelo ao eixo–x, representado por $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, cada ponto (x,0) é um ponto fixo; isto é, cada ponto do eixo–x é levado nele mesmo. Portanto, o segmento de reta com extremidades (0,0) e (2,0) é levado nele mesmo. Cada ponto (x,1) é levado num ponto de coordenadas (x+3,1); isto é, cada ponto na reta y = 1 é levado num ponto da reta posicionada 3 unidades à direita. Portanto, o segmento de reta com extremidades (0,1) e (2,1) é levado sobre o segmento com extremidades (3,1) e (5,1). Cada ponto (0,y) é levado no ponto de coordenadas (3y,y); isto é, cada ponto no eixo–y é levado sobre a reta 3y = x. Como resultado, o segmento com extremidades (0,0) e (0,1) é levado no segmento de reta com extremidades (0,0) e (3,1). Cada ponto (2,y) é levado sobre o ponto de coordenadas (2+3y,y); isto é, cada ponto na reta x = 2 é levado sobre um ponto da reta 2 + 3y = x. O resultado é que o segmento com extremidades (2,0) e (2,1) é levado sobre o segmento de reta com extremidades (2,0) e (5,1). Portanto, o retângulo é levado no paralelogramo de vértices (0,0), (2,0), (5,1) e (3,1).

**ER 3.9:** Determine o efeito da transformação do plano representada por $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

**Solução:** Sob esta transformação, cada ponto (x,y) é levado sobre um ponto de coordenadas (x,0). A matriz representa uma projeção vertical dos pontos do plano sobre o eixo–x. A transformação representada por $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ é um exemplo de uma aplicação sobre pontos do plano, mas não sobre ele todo.

**ER 3.10:** Determine o efeito da transformação do plano representada por $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$.

**Solução:** Sob esta transformação, cada ponto (x,y) é levado num ponto de coordenadas (x,x). A matriz representa uma projeção do plano sobre a reta y = x. A transformação do plano representada por $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ é um exemplo de aplicação do conjunto de pontos do plano, mas não nele todo.

As transformações do plano representado pelas matrizes do ER 3.9 e do 3.10 são chamadas **projeções do plano**. Existem outras projeções do plano.

## 3.5 TRANSFORMAÇÕES LINEARES HOMOGÊNEAS

As transformações do plano que foram discutidas até aqui são exemplos de uma classe geral de transformações chamadas **transformações do plano lineares homogêneas**.

**Definição 3.12:** As transformações do plano lineares homogêneas são aquelas expressas na forma matricial:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{3.11}$$

Assim, sob uma transformação linear homogênea do plano, cada ponto (x,y) é levado sobre sua imagem (x',y') onde

$$\begin{cases} x' = ax + by \\ y' = cx + dy \end{cases} \tag{3.12}$$

**Definição 3.13:** Se a matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ da transformação é não singular, a transformação definida por (3.11) é chamada **transformação homogênea linear não singular do plano**.

Por exemplo, rotações do plano sobre a origem, reflexões do plano com respeito a uma reta pela origem, dilatações do plano, expansões do plano, e cisalhamentos paralelos aos eixos coordenados são transformações homogêneas lineares não singulares do plano. As projeções do plano consideradas na Seção 3.4 são transformações homogêneas do plano que são não singulares.

**Teorema 3.1:** *Toda transformação homogênea linear não singular do plano é uma aplicação um a um do conjunto de pontos do plano nele mesmo.*

**Prova:** Seja T uma matriz não singular de ordem 2. Considere $\begin{pmatrix} x' \\ y' \end{pmatrix} = T\begin{pmatrix} x \\ y \end{pmatrix}$.

Então, uma vez que $T^{-1}$ existe,

$$T^{-1}\begin{pmatrix} x' \\ y' \end{pmatrix} = T^{-1}T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

isto é, dado qualquer ponto P’: (x’,y’), é possível encontrar o ponto P (x,y) tal que P’ é a imagem de P sob a transformação representada por T. Portanto, cada ponto do plano é imagem de algum ponto do plano sob a transformação representada por T, e T representa uma aplicação do conjunto de pontos do plano sobre ele mesmo.

Sejam A: $(x_1,y_1)$ e B: $(x_2,y_2)$ quaisquer dois pontos do plano que têm a mesma imagem sob a transformação representada por T; isto é, sejam

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \text{ e } \begin{pmatrix} x' \\ y' \end{pmatrix} = T\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

Então,

$$T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \qquad T\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

$$T^{-1}T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \qquad T^{-1}T\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \qquad \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

Uma vez que A e B são pontos idênticos se eles têm a mesma imagem, T representa uma aplicação um a um do conjunto de pontos do plano nele mesmo.

Um dos propósitos da discussão de reflexões do plano com respeito a uma reta pela origem, expansões do plano, e cisalhamentos paralelos aos eixos coordenados pode ser notado no seguinte importante e interessante teorema.

**Teorema 3.2:** *Toda transformação homogênea linear não singular do plano é o produto de uma reflexão do plano com respeito à reta y = x, uma expansão do plano, e um cisalhamento paralelo a um eixo de coordenadas.*

**Prova:** Do estudo da Seção 2.4 sobre transformações elementares sobre linhas, decorre que toda matriz não singular $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ pode ser expressa como um produto de matrizes de transformações elementares sobre linhas; isto é, $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ é um produto de matrizes da forma

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \text{ e } \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix} \tag{3.13}$$

A primeira matriz de (3.13) representa reflexões do plano com respeito à reta y = x; a segunda e a terceira matrizes representam expansões do plano; e a quarta e a quinta matrizes representam cisalhamentos paralelos aos eixos coordenados.

## 3.6 MATRIzES ORTOGONAIS

Considere uma matriz quadrada real A para a qual $AA^T = I$. Então, $A^T$ é a inversa à direita de uma matriz ortogonal A e, uma vez que A é uma matriz quadrada, suas matrizes inversas à esquerda e à direita são iguais. Portanto, $A^TA = I$. Além disso, $A^T = A^{-1}$.

**Definição 3.14:** Uma matriz é **ortogonal** se, e somente se, sua transposta e sua inversa são iguais, ou seja, $A^T = A^{-1}$.

A partir desta definição, podem ser facilmente verificadas as proposições seguintes:

**Proposição 3.1:** Se A é uma matriz ortogonal, sua inversa (transposta) é uma matriz ortogonal.

**Prova:** $(A^{-1})(A^{-1})^T = (A^T)(A^T)^T = A^TA = I$.

**Proposição 3.2:** Se duas matrizes A e B são matrizes ortogonais, seu produto é uma matriz ortogonal.

**Prova:** $(AB)^T = B^TA^T = B^{-1}A^{-1} = (AB)^{-1}$.

**Proposição 3.3:** Se A é uma matriz ortogonal, o determinante de A é igual a 1 ou –1.

**Prova:** Têm-se $\det AA^T = \det A \det A^T = \det A \det A = 1$, e então $\det A = \pm 1$.

**Definição 3.18:** Se $\det A = 1$, a matriz ortogonal A é chamada **matriz ortogonal própria**, e se $\det A = -1$, a matriz ortogonal A é chamada **matriz ortogonal imprópria**.

As matrizes de transformação que representam rotações do plano sobre a origem são exemplos de matrizes ortogonais próprias e as reflexões do plano com respeito a uma reta pela origem são exemplos de matrizes ortogonais impróprias.

**ER 3.11:** Verifique que a matriz A é uma matriz ortogonal própria onde

$$A = \begin{pmatrix} \frac{12}{13} & \frac{5}{13} \\ -\frac{5}{13} & \frac{12}{13} \end{pmatrix}.$$

**Solução:**

$$AA^T = \begin{pmatrix} \frac{12}{13} & \frac{5}{13} \\ -\frac{5}{13} & \frac{12}{13} \end{pmatrix}\begin{pmatrix} \frac{12}{13} & -\frac{5}{13} \\ \frac{5}{13} & \frac{12}{13} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Portanto, como $AA^T = I$, então A é uma matriz ortogonal. Além disso, detA = 1. Portanto, A é uma matriz ortogonal própria.

**OCTAVE 3.11**

Definimos a matriz A, calculamos a sua transposta (TA) e guardamos o produto destas na variável P. Na sequência, calculamos o determinante de P:

```
> A = [12/13,5/13;-5/13,12/13];
> TA = A';
> P = A*TA
P =
   1.00000   0.00000
   0.00000   1.00000
> det(P)
ans = 1.0000
```

Desta forma, ao encontrar o valor do determinante igual a 1, concluímos que A é ortogonal própria.

Uma vez que as matrizes que representam rotações do plano sobre a origem e reflexões do plano com respeito a uma reta pela origem, são matrizes ortogonais, é possível apresentar uma prova consistente de que o produto escalar de dois vetores no plano é um escalar invariante sob estas transformações. Isto é, se $a' = a_1'i + a_2'j$ e $b' = b_1'i + b_2'j$ são as imagens dos vetores $a = a_1 i + a_2 j$ e $b = b_1 i + b_2 j$, respectivamente, sob uma rotação do plano sobre a origem ou sob uma reflexão do plano com respeito a uma reta pela origem, então $a_1b_1 + a_2b_2 = a_1'b_1' + a_2'b_2'$. Seja R uma matriz de rotação ou uma matriz de reflexão.

Então,

$$\begin{pmatrix} a_1' & a_2' \end{pmatrix}^T = R\begin{pmatrix} a_1 & a_2 \end{pmatrix}^T \text{ e } \begin{pmatrix} a_1 & a_2 \end{pmatrix} R^T = \begin{pmatrix} a_1' & a_2' \end{pmatrix}.$$

Analogamente,

$$R\begin{pmatrix} b_1 & b_2 \end{pmatrix}^T = \begin{pmatrix} b_1' & b_2' \end{pmatrix}^T.$$

Então, multiplicando termo a termo, obtém-se:

$$\begin{pmatrix} a_1 & a_2 \end{pmatrix} R^T R \begin{pmatrix} b_1 & b_2 \end{pmatrix}^T = \begin{pmatrix} a_1' & a_2' \end{pmatrix}\begin{pmatrix} b_1' & b_2' \end{pmatrix}^T.$$

Assim, como R é uma matriz ortogonal, $R^TR = I$. Portanto,

$$\begin{pmatrix} a_1 & a_2 \end{pmatrix}\begin{pmatrix} b_1 & b_2 \end{pmatrix}^T = \begin{pmatrix} a_1' & a_2' \end{pmatrix}\begin{pmatrix} b_1' & b_2' \end{pmatrix}^T$$

e

$$a_1b_1 + a_2b_2 = a_1'b_1' + a_2'b_2'.$$

Já que os valores dos segmentos de reta, distâncias, e a medida de ângulos podem ser expressos em termos de produtos escalares de dois vetores, segue imediatamente que estas propriedades são invariantes sob rotações do plano sobre a origem e sob reflexões do plano com respeito a retas pela origem.

**ER 3.12:** Verifique que o produto escalar dos vetores $a = 2i + j$ e $b = 3i - j$ é um escalar invariante sob uma rotação do plano sobre a origem representada por

$$\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}.$$

**Solução:** Sob a rotação dada, $a' = a_1'i + a_2'j$ onde

$$\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Analogamente, $b' = b_1'i + b_2'j$ onde

$$\begin{pmatrix} b_1' \\ b_2' \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

Então, $a.b = (2)(3) + (1)(-1) = 5$ e $a'.b' = (2)(1) + (-1)(-3) = 5$. Portanto, o produto escalar a.b é um escalar invariante sob a rotação do plano sobre a origem, representado por

$$\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

**OCTAVE 3.12**

O produto escalar entre dois vetores é determinado pelo comando `dot`. Os vetores no OCTAVE são considerados matrizes coluna. Assim, definindo os vetores a e b, o produto escalar destes é guardado na variável c, como segue:

```
> a = [2;1];
> b = [3;-1];
> c = dot(a,b)
c = 5
```

Agora, definida a matriz de rotação R, e os vetores a e b, uma vez rotacionados, são guardados, respectivamente, em ar e br. Finalmente é verificado o produto escalar entre ar e br:

```
> R = [3/5,4/5;-4/5,3/5];
> ar = R*a;
> br = R*b;
```

```
> cr = dot(ar,br)
cr = 5
```

Note que mesmo após a rotação, o produto vetorial continua o mesmo (c=cr=5).

## 3.7 TRANSLAÇÕES

O conjunto de transformações lineares homogêneas do plano discutidas na seção 3.5 é um subconjunto do conjunto de **transformações lineares gerais do plano**. Toda transformação linear geral do plano pode ser descrita pelas equações

$$\begin{cases} x' = ax + by + e \\ y' = cx + dy + f \end{cases} \tag{3.14}$$

onde (x',y') é a imagem do ponto (x,y). Não é possível expressar uma transformação linear geral do plano descrita por (3.14) na forma de matriz em que a matriz da transformação seja uma matriz quadrada de ordem 2. Para expressar (3.14) numa forma matricial, será necessário considerar o uso de coordenadas cartesianas retangulares homogêneas.

**Definição 3.18:** Num plano Euclidiano coordenado, as **coordenadas homogêneas** de um ponto com **coordenadas não homogêneas** (x,y) são quaisquer escalares ordenados $(x_1,x_2,x_3)$ onde $x_3 \neq 0$ e para os quais $x = x_1/x_3$, e $y = x_2/x_3$.

**Exemplo 3.7:** O ponto do plano com coordenadas não homogêneas (1,–3) pode ser representado por um conjunto infinito de coordenadas homogêneas da forma (k,–3k,k) onde k é um número real não nulo. Um conjunto de coordenadas homogêneas para o ponto (x,y) é sempre da forma (x,y,1); todos os outros conjuntos de coordenadas homogêneas são da forma (kx,ky,k) onde $k \neq 0$.

Em geral, é o contexto de uma discussão que indicará se coordenadas homogêneas ou não homogêneas estão sendo consideradas.

Agora, uma transformação linear geral do plano, descrita por (3.14), pode ser expressa em forma matricial como

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & e \\ c & d & f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \tag{3.15}$$

onde (x’,y’,1) são as coordenadas homogêneas da imagem do ponto com coordenadas homogêneas (x,y,1).

Note que toda transformação linear homogênea do plano pode ser expressa na forma matricial

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{3.16}$$

Por exemplo, uma rotação do plano sobre a origem sob um ângulo θ pode ser expressa como

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\mathrm{sen}\theta & 0 \\ \mathrm{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

A equação matricial (3.16) representa as duas equações (3.12), que definem uma transformação linear homogênea do plano, e uma terceira equação trivial, 1 = 1, que não impõe quaisquer condições na relação entre as coordenadas de um ponto e sua imagem.

**Definição 3.19:** Uma transformação do plano onde cada ponto do plano é levado sobre um ponto fixado a uma distância $x_0$ paralela ao eixo–x e a uma distância $y_0$ paralela ao eixo–y (Figura 3.12) é chamada **translação do plano** e é definida pelas equações:

$$\begin{cases} x' = x + x_0 \\ y' = y + y_0 \end{cases} \tag{3.17}$$

onde (x’,y’) é a imagem do ponto (x,y).

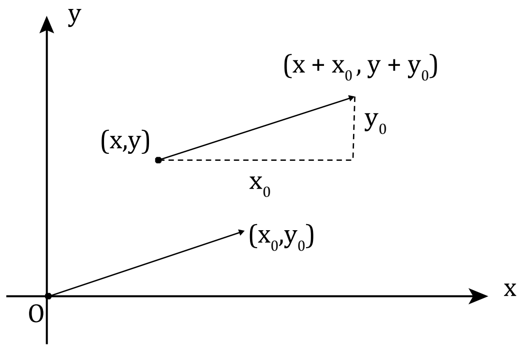



FIGURA 3.12 - Translação do plano

Uma translação do plano pode ser expressa na forma matricial como

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{3.18}$$

onde (x',y',1) são as coordenadas homogêneas da imagem do ponto com coordenadas homogêneas (x,y,1).

**Definição 3.20:** A matriz

$$\begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.19}$$

da transformação de translação definida por (3.18) é chamada **matriz de translação**.

Toda matriz que representa uma translação do plano é da forma (3.19). Note que o determinante de toda matriz de translação é igual a 1.

Uma translação do plano é uma aplicação um a um do conjunto de pontos do plano nele mesmo. Além disso, não existem pontos fixados sob a translação (3.18) quando $x_0$ e $y_0$ não são ambos nulos; isto é, cada ponto é levado sobre outro ponto. Quando $x_0 = y_0 = 0$, todo ponto do plano é um ponto fixado.

**ER 3.13:** Determine as coordenadas da imagem de P (3,2) sob uma translação do plano que leva a origem sobre (4,–1).

**Solução:** A translação dada do plano pode ser descrita pela equação matricial

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

uma vez que $\begin{pmatrix} 4 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Portanto, a imagem de P (3,2) é o ponto P’: (x’,y’) onde

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 1 \\ 1 \end{pmatrix};$$

isto é, P’: (7,1) é a imagem do ponto P (3,2) sob a translação do plano que leva a origem sobre (4,–1).

**OCTAVE 3.13**

Uma vez definida a matriz de translação, multiplica-se diretamente pelo vetor de interesse e encontram-se as novas coordenadas:

```
> T = [1,0,4;0,1,-1;0,0,1];
> v = [3;2;1];
> T*v
ans =
   7
   1
   1
```

Então, o vetor (3,2) é levado no vetor (7,1), por meio da translação considerada.

**ER 3.14:** Determine a equação satisfeita pelo conjunto das imagens de $x^2 + y^2 + 4x + 6y + 9 = 0$ sob a translação do plano representado por

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

**Solução:** Sob a translação dada, cada ponto (x,y) é levado sobre o ponto (x + 2, y + 3). Portanto, se x é trocado por x – 2 e y por y–3 em $x^2 + y^2 + 4x + 6y + 9 = 0$, a equação satisfeita pelo conjunto dos pontos imagem torna-se

$$\begin{aligned} &(x-2)^2 + (y-3)^2 + 4(x-2) + 6(y-3) + 9 = 0 \\ &x^2 - 4x + 4 + y^2 - 6y + 9 + 4x - 8 + 6y - 18 + 9 = 0 \\ &x^2 + y^2 = 4; \end{aligned}$$

isto é, um círculo com centro em (–2,–3) e raio 2 unidades é levado sobre um círculo com centro na origem e raio 2 unidades (Figura 3.13).

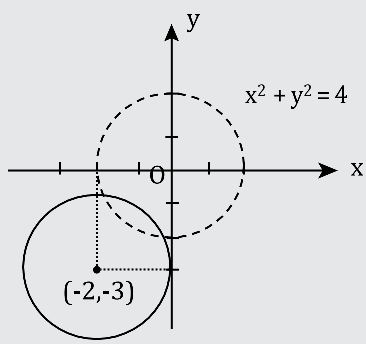



FIGURA 3.13 - Translação de circunferência

**ER 3.15:** Mostre que a imagem do ponto P (x,y) sob uma rotação do plano sobre a origem, seguida de uma translação do plano, geralmente não é a imagem do ponto sob as mesmas transformações consideradas em ordem reversa.

**Solução**: As matrizes que representam uma rotação do plano sobre a origem e uma translação do plano são, respectivamente, das formas

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ e } \begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix}$$

se elas são consideradas operar sobre o ponto P com coordenadas expressas na forma homogênea. Então, as coordenadas da imagem de P sob uma rotação seguida por uma translação são expressas por

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\operatorname{sen}\theta + x_0 \\ x\operatorname{sen}\theta - y\cos\theta + y_0 \\ 1 \end{pmatrix};$$

isto é, (x,y) é levado sobre $\left(x\cos\theta - y\operatorname{sen}\theta + x_0, x\operatorname{sen}\theta - y\cos\theta + y_0\right)$.

As coordenadas da imagem de P, sob as mesmas transformações consideradas em ordem reversa, são expressas por

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\operatorname{sen}\theta + x_0\cos\theta - y_0\operatorname{sen}\theta \\ x\operatorname{sen}\theta + y\cos\theta + x_0\operatorname{sen}\theta + y_0\cos\theta \\ 1 \end{pmatrix};$$

isto é, (x,y) é levado sobre

$\left(x\cos\theta - y\operatorname{sen}\theta + x_0\cos\theta - y_0\operatorname{sen}\theta, x\operatorname{sen}\theta + y\cos\theta + x_0\operatorname{sen}\theta + y_0\cos\theta\right)$

Em geral, os pontos $\left(x\cos\theta - y\operatorname{sen}\theta + x_0, x\operatorname{sen}\theta - y\cos\theta + y_0\right)$ e $\left(x\cos\theta - y\operatorname{sen}\theta + x_0\cos\theta - y_0\operatorname{sen}\theta, x\operatorname{sen}\theta + y\cos\theta + x_0\operatorname{sen}\theta + y_0\cos\theta\right)$ não são os mesmos.

**ER 3.16:** Use os resultados do ER 15 para determinar a imagem do ponto P (–5,2) sob uma rotação do plano sobre a origem sob um ângulo de 90º, seguido por uma translação do plano que leva a origem sobre (2,3).

**Solução:** Sob transformações tomadas nesta ordem, a imagem de cada ponto (x,y) é

$$\left(x\cos\theta - y\operatorname{sen}\theta + x_0, x\operatorname{sen}\theta - y\cos\theta + y_0\right).$$

Uma vez que $\theta = 90º$, $x_0 = 2$, $y_0 = 3$, $x = -5$ e $y = 2$, a imagem de P (–5,2) é (0,–2).

## 3.8 TRANSFORMAÇÕES DE MOVIMENTO RÍGIDO

Nas seções 3.2 e 3.3, foram consideradas rotações do plano sobre a origem e reflexões do plano com respeito a uma reta pela origem. Por meio da translação e de uma generalização do conceito expresso pela expressão (3.10) da Seção 3.3, é possível, agora, considerar rotações do plano sobre qualquer ponto do plano e reflexões do plano com respeito a qualquer reta no plano. Os três exemplos seguintes ilustram o procedimento para se determinar matrizes que representam tais transformações.

**ER 3.17:** Determine a matriz $\Re$ que representa uma rotação do plano sobre o ponto (1,1,1) sob um ângulo de 45º. Determine a imagem da origem sob a rotação do plano (Figura 3.14).

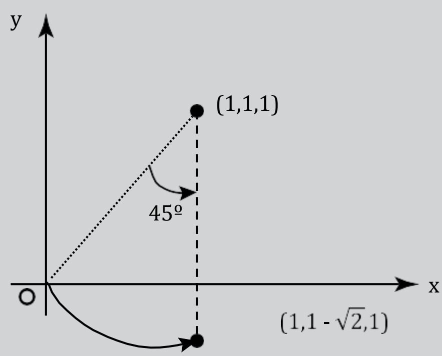



FIGURA 3.14 - Rotação do plano sobre ponto

**Solução:** A translação do plano que leva o ponto (1,1,1) sobre a origem é representada pela matriz

$$T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \text{ cuja inversa é } T^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

A matriz R = $\begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ representa uma rotação do plano sobre a origem sob um ângulo θ = 45º. Portanto, pela generalização do conceito expresso pela expressão (3.10) da Seção 3.3, a matriz $\mathfrak{R}$ que representa uma rotação do plano sobre o ponto (1,1,1) sob um ângulo de 45º é dada por

$$\mathfrak{R} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix};$$

isto é,

$$\mathfrak{R} = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & 1 \\ \sqrt{2}/2 & \sqrt{2}/2 & 1-\sqrt{2} \\ 0 & 0 & 1 \end{pmatrix}$$

A imagem da origem sob a rotação representada por $\mathfrak{R}$ é $\left(1, 1-\sqrt{2}, 1\right)$, pois

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

**OCTAVE 3.17**

Aqui, o software auxilia na multiplicação das matrizes. Basta defini-las, efetuar o produto diretamente e aplicar a matriz resultante às coordenadas consideradas:

```
> r2 = sqrt(2);
> R= [1,0,1;0,1,1;0,0,1]*[r2/2,-r2/2,0; r/2,r2/2,0;0,0,1]*[1,0,-1;
0,1,-1;0,0,1];
> origem = [0;0;1];
> R*origem
```

```
ans =
    1.00000
   -0.41421
    1.00000
```

Para simplificar a entrada da matriz da transformação, inicialmente definiu-se r2 como a raiz quadrada de 2 (r2=sqrt(2)). Note que a ordenada do vetor resultante é dada diretamente como um número natural (–0,41421) correspondente a 1 – sqrt(2):

```
> 1 - sqrt(2)
ans = -0.41421
```

**ER 3.18:** Determine a reflexão da matriz F que leva cada ponto (x,y,1) do plano sobre sua imagem com respeito à reta y = 2.

**Solução:** A translação do plano que leva a reta y = 2 sobre o eixo–x é representado pela matriz $T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$, cuja inversa é $T^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. A matriz R = $= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ representa uma reflexão com respeito ao eixo–x. Portanto, por (3.10), a matriz de reflexão F que leva cada ponto (x,y,1) do plano sobre sua imagem espelho, com respeito à reta y = 2, é dada como

$$F = T^{-1}RT = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

que é $F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$.

💻 **OCTAVE 3.18**

A determinação da matriz F é realizada, definindo-se as matrizes T e R, calculando a inversa de T e definindo F = inv(T)*R*T:

```
> T = [1,0,0;0,1,-2;0,0,1];
> R = [1,0,0;0,-1,0;0,0,1];
> F = inv(T)*R*T

> F =
   1   0   0
   0  -1   4
   0   0   1
```

**ER 3.19:** Determine a matriz de reflexão F que leva cada ponto (x,y,1) do plano sobre sua imagem espelho com respeito à reta x – y + 2 = 0.

**Solução:** Uma translação do plano representada pela matriz $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ leva a reta x – y + 2 = 0 sobre a reta x – y = 0; uma rotação sobre a origem sob um ângulo de (–45º), representada pela matriz $\begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, leva a reta x – y = 0 sobre o eixo–x. Portanto, a transformação produto que leva a reta x – y + 2 = 0 sobre o eixo–x é representada pela matriz

$$T = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 & \sqrt{2} \\ -\sqrt{2}/2 & \sqrt{2}/2 & -\sqrt{2} \\ 0 & 0 & 1 \end{pmatrix},$$

cuja inversa é

$$T^{-1} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & -2 \\ \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

A matriz $R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ representa a reflexão do plano com respeito ao eixo–x. Portanto, por (3.10), a matriz de reflexão F, que leva cada ponto (x,y,1) sobre sua imagem espelho com respeito à reta x – y + 2 = 0, é dada como

$$F = T^{-1}RT = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & -2 \\ \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 & \sqrt{2} \\ -\sqrt{2}/2 & \sqrt{2}/2 & -\sqrt{2} \\ 0 & 0 & 1 \end{pmatrix};$$

que é $F = \begin{pmatrix} 0 & 1 & -2 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}$.

Note que sob a transformação representada por F, todo ponto da reta x – y + 2 = 0 é um ponto fixo; isto é,

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

**OCTAVE 3.19**

De forma análoga, definem-se as matrizes T e R. Daí, F é calculada diretamente por F = inv(T)*R*T:

```
> r2 = sqrt(2);
> T = [r2/2 r2/2 0;-r2/2,r2/2,0;0,0,1]*[1 0 2;0 1 0;0 0 1];
> R = [1,0,0;0,-1,0;0,0,1];
> F = inv(T)*R*T
> F =

    0   1  -2
    1   0   2
    0   0   1
```

Mais uma vez, a definição r2 = sqrt(2) simplifica a entrada dos elementos da matriz, quando se tem raiz quadrada de 2.

**Definição 3.21:** O conjunto de rotações do plano sobre um ponto, reflexões do plano com respeito a uma reta, e translações do plano é chamado conjunto de **transformações de movimento rígido** porque a distância entre quaisquer dois pontos do plano é um escalar invariante sob estas transformações.

A geometria Euclidiana é algumas vezes caracterizada como um estudo de propriedades das figuras geométricas que permanecem invariantes sob as transformações de movimento rígido. O conjunto de transformações de movimento rígido é um subconjunto do conjunto geral de transformações lineares do plano. Como indicado na Figura 3.15, o conjunto R de rotações do plano sobre a origem e reflexões do plano com respeito a uma reta pela origem é a interseção do conjunto M de transformações de movimento rígido e o conjunto H das transformações lineares homogêneas.

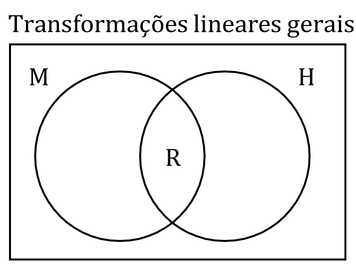



FIGURA 3.15 - Conjuntos de transformações lineares

O conjunto de produtos ordenados de transformações de movimento rígido pode ser representado por uma matriz da forma

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix}, \text{ onde } \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 1 \qquad (3.20)$$

Além disso,

$$\sum_{i=1}^{2} a_{ij} = 1 \text{, para j = 1,2;} \tag{3.21}$$

e

$$\sum_{i=1}^{2} a_{i1}a_{i2} = 0 \, . \tag{3.22}$$

Agora, serão provados dois teoremas que afirmam que uma transformação linear geral do plano é uma transformação de movimento rígido se, e somente se, a matriz que representa a transformação satisfaz as condições (3.21) e (3.22).

---

**Teorema 3.3:** *Seja* T = $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix}$ *representando uma transformação linear do plano sob a qual a distância é um escalar invariante. Então,*

$$a_{11}^2 + a_{21}^2 = 1,$$
$$a_{12}^2 + a_{22}^2 = 1,$$

*e*

$$a_{11}a_{12} + a_{21}a_{22} = 0.$$

**Prova:** Considere os pontos O (0,0,1) e P (1,0,1). Sob a transformação representada por T, as imagens de O e P são O’ ($a_{13}$, $a_{23}$, 1) e P’ ($a_{11}$ + $a_{13}$, $a_{21}$ + $a_{23}$, 1), respectivamente. Como $|\overrightarrow{O'P'}| = |\overrightarrow{OP}|$ sob a transformação representada por T,

$$\sqrt{[(a_{11}+a_{13})-a_{13}]^2+[(a_{21}+a_{23})-a_{23}]^2} = \sqrt{(1-0)^2+(0-0)^2}$$
$$\sqrt{a_{11}^2+a_{21}^2} = \sqrt{1^2}$$
$$a_{11}^2+a_{21}^2 = 1$$

De maneira análoga, escolhendo O: (0,0,1) e P (0,1,1), pode ser mostrado que $a_{12}^2 + a_{22}^2 = 1.$

Agora, considere os pontos O (0,0,1) e P (1,1,1). Sob a transformação representada por T, os pontos imagens de O e P são O' ($a_{13}$, $a_{23}$, 1) e P' ($a_{11}$ + $a_{12}$ + + $a_{13}$, $a_{21}$ + $a_{22}$ + $a_{23}$, 1), respectivamente. Novamente, como $\left|\overrightarrow{O'P'}\right| = \left|\overrightarrow{OP}\right|$ sob a transformação representada por T,

$$\sqrt{\left(a_{11}+a_{12}\right)^2+\left(a_{21}+a_{22}\right)^2} = \sqrt{1^2+1^2}$$

$$\left(a_{11}+a_{12}\right)^2+\left(a_{21}+a_{22}\right)^2=2$$

$$\left(a_{11}^2+a_{21}^2\right)+\left(a_{12}^2+a_{22}^2\right)+2\left(a_{11}a_{12}+a_{21}a_{22}\right)=2$$

$$1+1+2\left(a_{11}a_{12}+a_{21}a_{22}\right)=2$$

$$a_{11}a_{12}+a_{21}a_{22}=0$$

**Teorema 3.4:** *Seja* T = $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix}$ *representando uma transformação linear geral do plano tal que* $a_{11}^2+a_{21}^2=1$, $a_{12}^2+a_{22}^2=1$ *e* $a_{11}a_{12}+a_{21}a_{22}=0$. *Então, a distância é um escalar invariável sob a transformação representada por* T.

**Prova:** Considere quaisquer dois pontos $P_1$: ($x_1$,$y_1$,1) e $P_2$: ($x_2$,$y_2$,1) no plano. Sob a transformação representada por T, a imagem dos pontos $P_1$ e $P_2$ são $P_1$': ($a_{11}x_1$ + $a_{12}y_1$ + $a_{13}$, $a_{21}x_1$ + $a_{22}y_1$ + $a_{23}$, 1) e $P_2$': ($a_{11}x_2$ + $a_{12}y_2$ + $a_{13}$, $a_{21}x_2$ + + $a_{22}y_2$ + $a_{23}$,1), respectivamente. Agora,

$$\left|\overrightarrow{P_1'P_2'}\right| = \sqrt{\left[a_{11}\left(x_2-x_1\right)+a_{12}\left(y_2-y_1\right)\right]^2+\left[a_{21}\left(x_2-x_1\right)+a_{22}\left(y_2-y_1\right)\right]^2}$$

$$=\sqrt{\left(a_{11}^2+a_{21}^2\right)\left(x_2-x_1\right)^2+2\left(a_{11}a_{12}+a_{21}a_{22}\right)\left(x_2-x_1\right)\left(y_2-y_1\right)+\left(a_{12}^2+a_{22}^2\right)\left(y_2-y_1\right)^2}$$

$$=\sqrt{\left(x_2-x_1\right)^2+\left(y_2-y_1\right)^2}$$

$$=\left|\overrightarrow{P_1P_2}\right|;$$

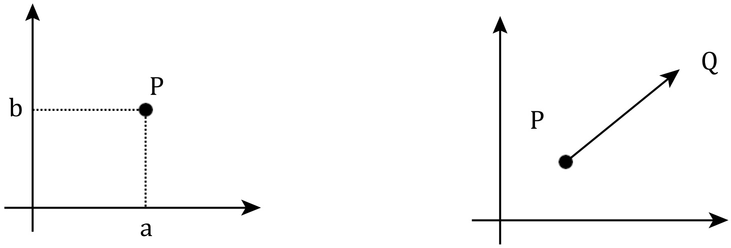
isto é, a distância entre as imagens de $P_1$ e $P_2$ é igual à distância entre $P_1$' e $P_2$'. Portanto, a distância é um escalar invariante sob a transformação representada por T.

# 4. ESPAÇO VETORIAL

**É NOTÓRIA A** importância dos vetores na teoria física clássica; por exemplo, a velocidade, a aceleração e as forças que atuam sobre um determinado objeto são descritas por vetores. A noção geométrica dos vetores é muito importante neste caso, pois auxilia na determinação das forças resultantes. Para tanto, é imprescindível a teoria das operações algébricas entre vetores. Neste capítulo trataremos destas operações entre vetores, preparando a base teórica necessária para o bom entendimento dos espaços vetoriais.

## 4.1 VETORES NO PLANO

No plano cartesiano ortogonal, um ponto P do plano é identificado pelo par (a,b) de números reais, onde as quantidades a e b são as coordenadas do ponto P (Figura 4.1(a)).




(a) (b)

FIGURA 4.1 - Ponto e vetor no plano

Considerados os pontos P e Q do plano, podemos definir dois segmentos de reta orientados $\overrightarrow{PQ}$ e , $\overrightarrow{QP}$ $P$ o primeiro com ponto inicial em P e ponto final em Q (Figura 4.1(b)) e o segundo, com ponto inicial em Q e ponto final em P. Embora o conjunto de pontos $\overrightarrow{PQ}$ e $\overrightarrow{QP}$ sejam iguais, considerada a orientação, eles são distintos. Diremos aqui que eles são segmentos opostos.

**Definição 4.1:** Dois segmentos orientados são ditos **equivalentes** se tiverem o mesmo comprimento, a mesma direção e o mesmo sentido. A direção é determinada pela reta que suporta o segmento.

**Exemplo 4.1:** Observe a Figura 4.2 abaixo

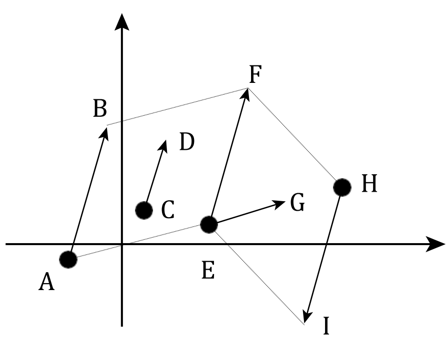



FIGURA 4.2 - Conjunto de vetores no plano

$\overrightarrow{AB}, \overrightarrow{CD}$ e $\overrightarrow{EF}$ têm a mesma direção; $\overrightarrow{EG}$ e $\overrightarrow{CD}$ têm o mesmo comprimento; $\overrightarrow{AB}, \overrightarrow{HI}$ e $\overrightarrow{EF}$ têm o mesmo comprimento, mas os únicos com orientações equivalentes são $\overrightarrow{AB}$ e $\overrightarrow{EF}$.

**Definição 4.2:** Para qualquer segmento orientado no plano existe outro equivalente a este, cujo ponto inicial é a origem.

**Exemplo 4.2:** Observe o segmento orientado $\overrightarrow{PQ}$ na Figura 4.3 abaixo. O segmento $\overrightarrow{OA}$ tem a mesma orientação de $\overrightarrow{PQ}$, e uma vez que eles têm a mesma direção e o mesmo sentido, eles são equivalentes.

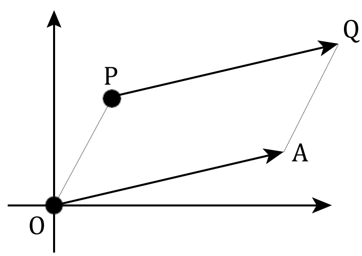



FIGURA 4.3 - Vetores equivalentes

**Definição 4.3:** Os segmentos orientados com ponto inicial na origem são denominados **vetores no plano** e são determinados exclusivamente pelo seu ponto final, já que o ponto inicial é fixo na origem.

A cada ponto no plano P(a,b) é associado um único vetor **v** = $\overrightarrow{OP}$ e, reciprocamente, dado um vetor, este fica associado a um único ponto do plano, que é seu ponto final. Consequentemente, a correspondência entre pontos do plano e vetores é biunívoca.

Desta forma, um vetor **v** = $\overrightarrow{OP}$ é representado, simplesmente, pelas coordenadas do seu ponto final P(a,b). Usamos a notação **v** = $\begin{bmatrix} a \\ b \end{bmatrix}$ ou **v** = (a,b) para identificar um vetor cujo ponto final é (a,b). No contexto deste livro, os vetores serão graficamente marcados por negrito, ou seja, v é um escalar e **v** é um vetor.

**Definição 4.4:** A origem do plano fica associada a um vetor denominado **vetor nulo**, que tem os pontos inicial e final coincidentes com a origem. O vetor nulo é representado por **0** = (0,0).

**Definição 4.5:** O **oposto** de um vetor **v** = $\overrightarrow{OP}$ é o vetor **w** = $\overrightarrow{OQ}$ que tem o mesmo comprimento, mesma direção e sentido oposto. Em termos de coordenadas, se **v** = (a,b), então **w** = (–a,–b) e denotamos **w** = –**v**.

## 4.2 OPERAÇÕES COM VETORES NO PLANO

Na álgebra vetorial, é importante que estejam bem definidas as operações entre vetores. E, dentre as mais importantes, estão a multiplicação por escalar e a adição, operações que darão, posteriormente, os fundamentos para descrever um subespaço vetorial. Estas operações são definidas a seguir.

a) Multiplicação de um vetor **v** por um escalar k.

- Se k > 0, o vetor **w** = k**v** possui a mesma direção de **v**, o mesmo sentido e comprimento k vezes o comprimento de **v**.
- Se k < 0, o vetor **w** = k**v** será igual ao oposto do vetor |k|**v**, onde |k| representa o valor absoluto do escalar k.
- Se k = 0, **w** = k**v** será o vetor nulo.

A multiplicação de um vetor por um escalar corresponde à multiplicação de cada coordenada do vetor por esse escalar. Assim, se **v** = (a,b) e **w** = k**v**, então **w** = (ka,kb).

**Exemplo 4.3:** Para **v** = (3,1), **w1** = 2**v** = (6,2) e **w2** = –2**v** = (–6,–2) (Figura 4.4).

b) Adição de dois vetores.

- Se **v** = (a,b) e **w** = (c,d), então o vetor soma será **v** + **w** = (a + c, b + d).
- A soma de um vetor **v** = (a,b) com seu oposto **w** = –**v** = (–a,–b) é o vetor nulo. Isto é, **v** + **w** = **v** + (–**v**) = (a – a, b – b) = (0,0).

- A **diferença** entre dois vetores **v** e **w** é a soma do primeiro com o oposto do segundo vetor:

**v** – **w** = **v** + (-**w**)

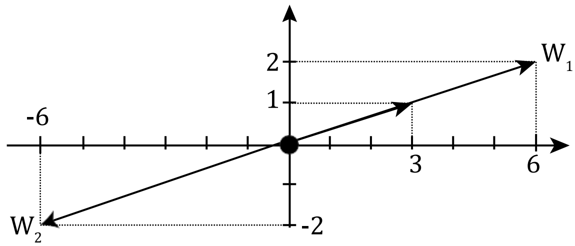



FIGURA 4.4 - Diferença entre vetores

## 4.3 VETORES NO ESPAÇO

Analogamente ao caso de vetores no plano, podemos estender a ideia de vetor ao espaço tridimensional, onde um ponto P do espaço é identificado com uma terna de números reais (x,y,z), onde x, y e z são as coordenadas do ponto P (Figura 4.5).

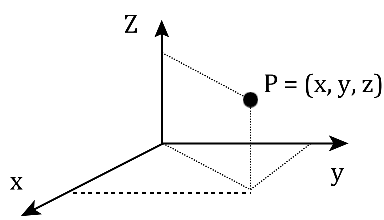



FIGURA 4.5 - Ponto no espaço 3D

Agora cada ponto do espaço P(a,b,c) é associado a um único vetor **v** = $OP$ e, reciprocamente, dado um vetor, este fica associado a um único ponto do espaço, que é seu ponto final.

Assim, um vetor **v** = $\overrightarrow{OP}$ é representado pelas coordenadas do seu ponto final P(a,b,c). Denotamos **v** $= \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ ou **v** = (a,b,c) para identificar um vetor cujo ponto final é (a,b,c).

A origem do espaço representa o vetor nulo **0** = (0,0,0). Se V é o conjunto de vetores no espaço, então V = {($x_1$,$x_2$,$x_3$); $x_i \in$ R} = R ´ R ´ R = $R^3$.

De forma análoga às operações com vetores no plano, o processo de multiplicação por escalar e adição de vetores pode ser estendido a vetores no espaço (três coordenadas) conforme a Definição 4.6, a seguir. Na realidade estas operações podem ser generalizadas para vetores do espaço n-dimensional, ou seja, vetores com n coordenadas; porém, nos casos de dimensão maior que 3, não é possível a visualização gráfica.

**Definição 4.6:** Sejam **u** = ($x_1$,$x_2$,$x_3$) e **v** = ($y_1$,$y_2$,$y_3$) vetores no espaço. A soma de dois vetores e o produto de um vetor por um escalar são denotados respectivamente por **u** + **v** e k**u**, e são definidos da seguinte forma:
**u** + **v** = ($x_1 + y_1$, $x_2 + y_2$, $x_3 + y_3$) e k**u** = ($kx_1$, $kx_2$, $kx_3$).

Antes de definir um espaço vetorial é importante discutir uma outra estrutura denominada *corpo*. Considere-se um conjunto *C*, munido de duas operações, chamadas de *adição* (+) e *multiplicação* (.), que obedecem às seguintes propriedades, para quaisquer elementos $x, y, z \in C$:

- $(x + y) + z = x + (y + z)$ e $(x.y).z = x.(y.z)$ (propriedade associativa);
- $x + y = y + x$ e $x.y = y.x$ (propriedade comutativa);
- $\exists\ \alpha, \beta \in C$ tal que, $x + \alpha = \alpha + x = x$ e $x.\beta = \beta.x = x$ (existência do elemento neutro);
- $\exists\ -x, x^{-1} \in C$ tal que, $x + (-x) = \alpha$ e $x.x^{-1} = \beta$ (existência do elemento inverso);
- $x.(y + z) = x.y + x.z$ (distributiva da multiplicação em relação à adição);
- Se $\alpha$ e $\beta$ representam os elementos neutros de *C*, respectivamente às operações de adição e multiplicação, então *C* é um **corpo comutativo**.

## 4.4 ESPAÇOS VETORIAIS

Um dos objetivos da teoria de Espaços Vetoriais é fundamentar a álgebra de determinados conjuntos que se comportam de maneira semelhante quando observadas combinações lineares dos seus elementos. De forma geral, é interessante lidar com as propriedades comuns de sistemas algébricos constituídos por um conjunto de operações (no caso, a multiplicação por escalar e a adição), que permitam uma definição natural de combinação linear de seus elementos.

**Definição 4.7:** Seja V um conjunto não vazio, onde estão definidas as operações:

i) $\forall$ $\mathbf{u}$, $v \in V \rightarrow \mathbf{u} + v \in V$

ii) $\forall\ \alpha \in \mathrm{R}$; $\forall\ \mathbf{u} \in V \rightarrow \alpha\, \mathbf{u} \in V$

Este conjunto V é um **espaço vetorial real**, se forem verificadas as propriedades:

1. $\forall\ \mathbf{u}, \mathbf{v}, w \in V$: $(\mathbf{u} + \mathbf{v}) + w = \mathbf{u} + (\mathbf{v} + w)$;
2. $\forall\ \mathbf{u}, \mathbf{v} \in V$, segue que: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
3. Existe $0 \in V$ (elemento nulo) tal que para todo $\mathbf{u} \in V$: $0 + \mathbf{u} = \mathbf{u}$;
4. Para cada $\mathbf{u} \in V$, existe $-\mathbf{u} \in V$ (elemento oposto) tal que: $\mathbf{u} + (-\mathbf{u}) = 0$;
5. Para quaisquer $\alpha, \beta \in \mathrm{R}$ e qualquer $\mathbf{u} \in V$: $(\alpha.\ \beta).\mathbf{u} = \alpha\,(\beta.\ \mathbf{u})$;
6. Para quaisquer $\alpha, \beta \in \mathrm{R}$ e todo $\mathbf{u} \in V$: $(\alpha + \beta).\ \mathbf{u} = \alpha.\mathbf{u} + \beta.\mathbf{u}$;
7. Para todo escalar $\alpha \in \mathrm{R}$ e quaisquer $\mathbf{u}, \mathbf{v} \in V$: $\alpha.(\mathbf{u} + \mathbf{v}) = \alpha.\mathbf{u} + \alpha.\mathbf{v}$;
8. Para qualquer $\mathbf{u} \in V$ tem-se que: $1.\mathbf{u} = \mathbf{u}$.

Na Definição 4.7, se em vez de escalares reais tivermos escalares complexos, V será um **espaço vetorial complexo**.

Neste contexto, a palavra **vetor** é utilizada de maneira mais geral, ou seja, designa um elemento de um espaço vetorial. Por exemplo, considerando o espaço vetorial das matrizes quadradas reais de ordem 2, denotado por V = M(2,2), definidas a multiplicação por escalar e a adição usuais, um vetor deste espaço será na realidade uma matriz real 2x2.

**Exemplo 4.4:** Considere o espaço n-dimensional real, $V = R^n = \{(x_1,x_2,...,x_n); x_i \in R\}$. Sejam **u** e **v** dois vetores de V dados por $\mathbf{u} = (x_1,x_2,...,x_n)$, $\mathbf{v} = (y_1,y_2,...,y_n)$. A soma entre **u** e **v** é definida por $\mathbf{u} + \mathbf{v} = (x_1 + y_1,x_2 + y_2,...,x_n + y_n)$ e o produto de um vetor por um escalar é definido como $a\mathbf{u} = (ax_1,ax_2,...,ax_n)$ $(a \in R)$. Nestas condições é simples verificar que V é um espaço vetorial real, ou seja, as propriedades 1 a 8 são satisfeitas. Neste caso, o vetor nulo é $\mathbf{0} = (0, 0, 0, ..., 0)$ e $-\mathbf{u} = (-x_1,-x_2,...,-x_n)$.

**Exemplo 4.5:** Considere, agora, V como o conjunto das matrizes de ordem 2, cujos elementos são números complexos, ou seja, $V = \left\{\begin{bmatrix} z_1 & z_2 \\ z_3 & z_4 \end{bmatrix}\right\}$, onde os elementos da matriz são números complexos. Assim, se $\mathbf{u}, \mathbf{v} \in V$ são dados por

$$\mathbf{u} = \begin{bmatrix} z_1 & z_2 \\ z_3 & z_4 \end{bmatrix} \text{ e } \mathbf{v} = \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix},$$

a soma é $\mathbf{u} + \mathbf{v} = \begin{bmatrix} z_1 & z_2 \\ z_3 & z_4 \end{bmatrix} + \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = \begin{bmatrix} z_1 + w_1 & z_2 + w_2 \\ z_3 + w_3 & z_4 + w_4 \end{bmatrix}$ e o produto por escalar é definido por $a\mathbf{u} = a\begin{bmatrix} z_1 & z_2 \\ z_3 & z_4 \end{bmatrix} = \begin{bmatrix} az_1 & az_2 \\ az_3 & az_4 \end{bmatrix}$, onde $z_i + w_i$ $(i = 1,2,3,4)$ é uma soma de números complexos e $az_i$ $(i = 1,2,3,4)$ é produto de um escalar por um número complexo.

O Exemplo 4.4 é um espaço vetorial real, enquanto o Exemplo 4.5 é um exemplo de espaço vetorial complexo.

## 4.5 SUBESPAÇOS VETORIAIS

Um subespaço vetorial de V é um subconjunto $S \subset V$, que é também espaço vetorial segundo o par de operações que caracterizam V. Fica claro que todo elemento de S já possui satisfeitos todos os axiomas operacionais citados acima, ou seja, para concluir que S é um subespaço vetorial de V, basta verificar o fechamento.

Num posterior estudo sobre subespaços vetoriais, o leitor chegará à conclusão de que se um conjunto S é um subespaço, então S possui o vetor nulo (denotado por **0**). A definição terá uma melhor "precisão" ao se incluir esse item, pois, se no conjunto S não estiver presente o vetor nulo, já se pode concluir de imediato que não se trata de um subespaço vetorial.

**Definição 4.8:** Dado um espaço vetorial real V, um subconjunto $W \neq$ Æ, será um **subespaço vetorial** de V se:

i) $\forall\ \mathbf{u},\mathbf{v} \in W \Rightarrow \mathbf{u} + \mathbf{v} \in W$

ii) $\forall\ a \in \mathrm{R}, \mathbf{u} \in W \Rightarrow a\mathbf{u} \in W$

---

**Exemplo 4.6:** Considere os subconjuntos $L = \left\{ (x,y) \in R^2 ;\ y = x + 1 \right\}$ e $S = \left\{ \left(x,y\right) \in R^2\ ;\ y = x \right\}$ do $R^2$. Sabemos que S é um subespaço de $R^2$ e L não o é, pois, ao se somarem vetores de S, o vetor resultante pertence à reta definida pelo conjunto S. O mesmo não ocorre com L.

Observe que S possui o vetor nulo, ou seja, em S a reta passa pela origem, enquanto o mesmo não ocorre com L. Pode-se afirmar que toda reta que passa pela origem do plano cartesiano representa um subespaço vetorial de $R^2$; da mesma forma, todas as retas que não passam pela origem não representam subespaços vetoriais de $R^2$. Pode-se estender este conceito para $R^3$, pois todos os planos e todas as retas que passam pela origem representam subespaços vetoriais do $R^3$ e, consequentemente, se um plano ou uma reta não passam pela origem não representam um subespaço vetorial.

---

Tanto o conjunto dos números reais quanto o conjunto dos números complexos gozam da propriedade de serem, ao mesmo tempo, espaço vetorial e corpo. Não necessariamente um espaço vetorial precisa ser definido sobre o corpo dos reais; pode ser definido sobre um corpo específico. Por exemplo, o conjunto dos números complexos é um espaço vetorial sobre o corpo dos reais, e é também um espaço vetorial sobre o corpo dos complexos. É interessante a observação de que o conjunto dos números reais não é um espaço vetorial sobre o corpo dos complexos. Basta observar que um número complexo, multiplicado por um número real, nem sempre gera um número real, o que significa que a operação de *multiplicação por escalar*, quando o corpo é o conjunto dos números complexos e o espaço vetorial é o conjunto dos números reais, não é fechada.

Todo espaço vetorial admite pelo menos dois subespaços, denominados subespaços triviais, o conjunto formado somente pelo vetor nulo e o próprio espaço vetorial (**0** e V).

Agora veremos um importante subespaço que aparece quando se resolvem sistemas lineares homogêneos.

---

**Exemplo 4.7:** Considere o sistema linear homogêneo $\begin{cases} 3x+5y+z=0 \\ 2x+2y+3z=0 \\ 2x+4y-z=0 \end{cases}$ que tem a seguinte forma matricial $\begin{bmatrix} 3 & 5 & 1 \\ 2 & 2 & 3 \\ 2 & 4 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (*). Assim, estamos procurando no espaço vetorial M(3,1), que são as matrizes de ordem 3x1 reais, os vetores que satisfazem a equação matricial (*). O conjunto dos vetores-solução é um subespaço de M(3,1)? Para responder a esta pergunta, basta considerar dois vetores-solução $\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ e $\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$ e verificar que sua soma $\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}+\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$ ainda é um vetor-solução:

$$\begin{bmatrix} 3 & 5 & 1 \\ 2 & 2 & 3 \\ 2 & 4 & -1 \end{bmatrix}\left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}+\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}\right)=\begin{bmatrix} 3 & 5 & 1 \\ 2 & 2 & 3 \\ 2 & 4 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}+\begin{bmatrix} 3 & 5 & 1 \\ 2 & 2 & 3 \\ 2 & 4 & -1 \end{bmatrix}\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}+\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Isto é, a soma é solução. Além disso, multiplicando $\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ por uma constante k, teremos:

---

$$\begin{bmatrix} 3 & 5 & 1 \\ 2 & 2 & 3 \\ 2 & 4 & -1 \end{bmatrix}\left(k.\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}\right) = k.\left(\begin{bmatrix} 3 & 5 & 1 \\ 2 & 2 & 3 \\ 2 & 4 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}\right) = k.\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Isto é, o produto de uma constante por uma solução continua sendo uma solução. Portanto, o conjunto W dos vetores-solução de (*) é um subespaço de M(3,1).

A seguir, veremos um exemplo onde um subconjunto de um espaço vetorial não é subespaço vetorial.

**Exemplo 4.8:** Sejam V = $\mathbb{R}^2$ e W = $\{(x,x^3);\ x \in \mathbb{R}\}$. Considerando **u** = (1,1) e **v** = (2,8) em W, temos **u** + **v** = (1,1) + (2,8) = (3,9) $\notin$ W. Portanto, W não é subespaço vetorial de V, pois a primeira condição da Definição 4.8 deveria ser satisfeita para quaisquer **u** e **v** $\in$ W e isto não acontece no caso dos vetores considerados.

**Teorema 4.1 (Interseção de subespaços):** *Dados $W_1$ e $W_2$, subespaços de um espaço vetorial V, a interseção $W_1 \cap W_2$ ainda é um subespaço de V.*

Uma das características mais importantes quando se trabalha com espaços vetoriais é a capacidade de geração de novos vetores a partir de um conjunto de vetores dados. Para tanto, é definido o conceito de combinação linear de elementos de um espaço vetorial.

**Definição 4.9:** O subespaço construído recebe o nome de *subespaço vetorial de* V *gerado por* B (ou apenas "*o gerado por B*"), e cada elemento de [B] é escrito como *combinação linear* dos elementos de B.

**Definição 4.10:** Um espaço vetorial $V$ é finitamente gerado se existir um conjunto finito $S \subset V$ de tal forma que $[S]=V$. Salvo menção contrária, consideraremos apenas espaços vetoriais finitamente gerados.

NOTAS:

1. Estende-se a definição acima para o caso $B = \phi$ mediante a seguinte convenção: $[\phi] = \{0\}$.
2. No caso de $B \subset V$ ser um conjunto com infinitos elementos, define-se $[B]$ como: $u \in [B] \Leftrightarrow \exists\ u_1,...,u_n \in B$ e $\exists\ \alpha_1,...,\alpha_n \in C$ tais que $u = \sum_{i=1}^{n} \alpha_i u_i$.

Pode-se afirmar que qualquer conjunto finito de vetores de $V$ gera um subespaço vetorial de $V$ incluindo o conjunto $\phi$ que por definição gera o subespaço nulo $\{0\}$.

Seja $V$ um espaço vetorial finitamente gerado sobre um corpo $C$. Decorre da forma como foram definidos os conjuntos $B$ e $[B]$, as seguintes propriedades:

**Propriedade 4.1:** $B \subset [B]$ (Todo subconjunto de $V$ está contido no subespaço gerado por ele).

**Propriedade 4.2:** $B_1 \subset B_2 \subset V \Rightarrow [B_1] \subset [B_2]$ (Se $B_1$ é subconjunto de $B_2$ que por sua vez é subconjunto de um espaço vetorial $V$, então, o subespaço gerado por $B_1$ é subconjunto do gerado por $B_2$).

**Propriedade 4.3:** $[B] = [[B]]$ (o subespaço gerado por $B$ é igual ao gerado pelo gerado por $B$).

**Propriedade 4.4:** Se $S_1$ e $S_2$ são subconjuntos de $V$, então: $[S_1 \cup S_2] = [S_1] + [S_2]$.

**Propriedade 4.5:** Se $[B] = S$ e $[S] = K$, então $[B] = K$ (Se um conjunto $B$ gera um subespaço $S$, que por sua vez gera um outro subespaço $K$, então, por transitividade, o conjunto $B$ gera o subespaço $K$).

**Exemplo 4.9:** Seja o espaço vetorial das matrizes quadradas reais de ordem n, denotado por V = M(n,n). Se considerarmos o conjunto das matrizes triangulares superiores e das matrizes triangulares inferiores aqui denotados, respectivamente, por $W_1$ e $W_2$, subconjuntos de V, é fácil verificar que $W_1$ e $W_2$ são subespaços de V. Então, $W_1 \cap W_2$ é o conjunto das matrizes diagonais, e de acordo com o Teorema 4.1, também é um subespaço de V.

**Exemplo 4.10:** Considere o espaço vetorial tridimensional real $V = \mathbb{R}^3$, e $W_1$ e $W_2$ representando retas não coincidentes que passam pela origem. Então $W_1 \cap W_2 = \{\mathbf{0}\}$ e $W_1 \cup W_2$ é o "feixe" formado pelas duas retas, que não é subespaço vetorial de $\mathbb{R}^3$. De fato, se somarmos dois vetores **u** e **v** pertencentes a $W_1 \cup W_2$, vemos que **u** + **v** está no plano que contém $W_1$ e $W_2$, mas $\mathbf{u} + \mathbf{v} \notin W_1 \cup W_2$. Assim, $W_1 \cup W_2$ não é subespaço de V. Entretanto, podemos construir um conjunto W, que contém $W_1$ e $W_2$ e é subespaço de V. W que será formado por todos os vetores de V que forem a soma de vetores de $W_1$ com vetores de $W_2$. Assim, $W = W_1 + W_2$ e será chamado **soma de $W_1$ com $W_2$**.

**Teorema 4.2 (Soma de subespaços):** *Sejam $W_1$ e $W_2$ subespaços de um espaço vetorial V. Então, o conjunto: $W_1 + W_2 = \{\mathbf{v} \in V: \mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2, \mathbf{w}_1 \in W_1 \text{ e } \mathbf{w}_2 \in W_2\}$ é subespaço de V.*

**Exemplo 4.11:** No Exemplo 4.10, a soma de subespaços $W = W_1 + W_2$ é o plano que contém as duas retas e, de acordo com o Teorema 4.2, também é um subespaço de V.

**Exemplo 4.13:** Sejam $W_1$ e $W_2$ os subespaços do espaço das matrizes quadradas reais de ordem 2 (M(2,2)) gerados da seguinte forma: $W_1 = \left\{\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}\right\}$ e $W_2 = \left\{\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}\right\}$, onde $a,b,c,d \in \mathbb{R}$. Então, $W_1 + W_2 = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right\} = M(2,2)$.

**Definição 4.9:** Considerando dois subespaços quaisquer $W_1$ e $W_2$ de um espaço vetorial V, quando $W_1 \cap W_2 = \{\mathbf{0}\}$, então $W_1 + W_2$ é chamado **soma direta** de $W_1$ com $W_2$ e denotamos esta soma direta por $W_1 \oplus W_2$.

**Exemplo 4.16:** Considere o espaço vetorial (plano) $V = \mathbb{R}^2$, e os vetores unitários $\mathbf{v_1} = (1,0)$, $\mathbf{v_2} = (0,1)$. Logo, $\mathbf{v_1}$ e $\mathbf{v_2}$ geram o plano $\mathbb{R}^2$ ($V = [\mathbf{v_1},\mathbf{v_2}]$), pois, dado que $\mathbf{v} = (x,y) \in V$, temos que $(x,y) = x(1,0) + y(0,1)$, ou seja, todo vetor pode ser escrito como uma combinação linear de $\mathbf{v_1}$ e $\mathbf{v_2}$ ($\mathbf{v} = x\mathbf{v_1} + y\mathbf{v_2}$).

**Exemplo 4.17:** Considere o espaço vetorial das matrizes quadradas reais de ordem 2 (M(2,2)). Dados os vetores $\mathbf{v_1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{v_2} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, o espaço gerado por $\mathbf{v_1}$ e $\mathbf{v_2}$ é dado por $[\mathbf{v_1},\mathbf{v_2}] = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} : a,b \text{ reais} \right\}$.

## 4.6 COMBINAÇÃO LINEAR

No contexto de espaços vetoriais é importante saber quando um dado vetor é uma combinação de um dado conjunto de outros vetores. Além disso, cabe saber se realmente todos os vetores deste conjunto são necessários para escrever, em termos de combinação linear, o vetor dado. O que acontece é que nem sempre todos os vetores são indispensáveis, aparecendo no conjunto vetores excedentes que não contribuem com a caracterização do vetor de interesse. Desta forma é interessante o conceito de dependência linear que caracteriza estes vetores excedentes.

**Definição 4.10:** Consideremos um espaço vetorial real V (ou complexo), $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ e $a_1, a_2, \ldots, a_n$ números reais (ou complexos). Chamamos **combinação linear de** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ o vetor $\mathbf{v} \in V$, definido por

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n.$$

Vamos, agora, entender o importante conceito de subespaço gerado. Vamos primeiramente fixar os vetores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$. Desta forma, o conjunto W de todos os vetores de V que são combinação linear destes é um subespaço vetorial.

**Definição 4.11:** Fixados os vetores $\mathbf{v}_{1,} \mathbf{v}_{2, ...,} \mathbf{v}_n \in V$, o conjunto W de todos os vetores que são combinação linear dos $\mathbf{v}_i$ (i = 1, 2, ..., n) é chamado **subespaço gerado** pelos $\mathbf{v}_i$ e usamos a notação: $W = [\mathbf{v}_{1,} \mathbf{v}_{2, ...,} \mathbf{v}_n]$.

Podemos representar o espaço gerado pelos $\mathbf{v}_i$, da seguinte forma:

$$W = [\mathbf{v}_{1,} \mathbf{v}_{2, ...,} \mathbf{v}_n] = \{\mathbf{v} \in V: \mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + _{...} + a_n\mathbf{v}_{n,} a_i \in \mathbb{R}, 1 \leq i \leq n\}$$

onde W é o "menor" subespaço de V que contém o conjunto de vetores $\{\mathbf{v}_{1,} \mathbf{v}_{2, ...,} \mathbf{v}_n\}$. O termo "menor" significa que qualquer outro subespaço W' de V que contenha $\{\mathbf{v}_{1,} \mathbf{v}_{2, ...,} \mathbf{v}_n\}$ deverá satisfazer W' É W.

**Exemplo 4.14:** Considere o espaço vetorial $V = \mathbb{R}^3$ e $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$. Então $[\mathbf{v}] = \{a\mathbf{v}: a \in \mathbb{R}\}$, isto é, $[\mathbf{v}]$ é a reta que contém o vetor $\mathbf{v}$ (Figura 4.6).

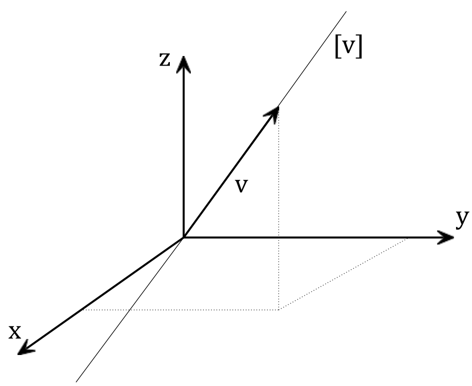



FIGURA 4.6 - Reta que contém vetor

**Exemplo 4.15:** Se $\mathbf{v_1},\mathbf{v_2} \in \mathbb{R}^3$ são tais que $\alpha\mathbf{v_1} \neq \mathbf{v_2}$ para todo $\alpha \in \mathbb{R}$, então $[\mathbf{v_1},\mathbf{v_2}]$ será o plano que passa pela origem e contém $\mathbf{v_1}$ e $\mathbf{v_2}$ (Figura 4.7). Observe que se $\mathbf{v_3} \in [\mathbf{v_1},\mathbf{v_2}]$, então, $[\mathbf{v_1},\mathbf{v_2},\mathbf{v_3}] = [\mathbf{v_1},\mathbf{v_2}]$, pois todo vetor que pode ser escrito como combinação linear de $\mathbf{v_1}$, $\mathbf{v_2}$ e $\mathbf{v_3}$ é uma combinação linear de $\mathbf{v_1}$ e $\mathbf{v_2}$ (pois $\mathbf{v_3}$ é combinação linear de $\mathbf{v_1}$ e $\mathbf{v_2}$).

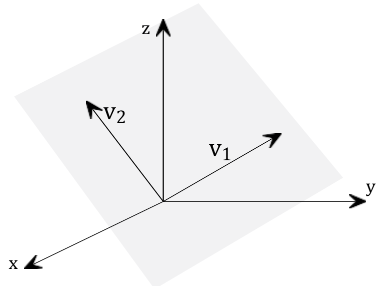



FIGURA 4.7 - Vetores contidos num plano do espaço

**Exemplo 4.16:** Considere o espaço vetorial (plano) V = $\mathbb{R}^2$, e os vetores unitários $\mathbf{v_1}$ = (1,0), $\mathbf{v_2}$ = (0,1). Logo $\mathbf{v_1}$ e $\mathbf{v_2}$ geram o plano $\mathbb{R}^2$ (V = [$\mathbf{v_1}$,$\mathbf{v_2}$]), pois, dado $\mathbf{v}$ = (x,y) $\in$ V, temos que (x,y) = x(1,0) + y(0,1), ou seja, todo vetor pode ser escrito como uma combinação linear de $\mathbf{v_1}$ e $\mathbf{v_2}$ ($\mathbf{v} = x\mathbf{v_1} + y\mathbf{v_2}$).

**Exemplo 4.17:** Considere o espaço vetorial das matrizes quadradas reais de ordem 2 (M(2,2)). Dados os vetores $\mathbf{v_1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{v_2} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, o espaço gerado por $\mathbf{v_1}$ e $\mathbf{v_2}$ é dado por $[\mathbf{v_1},\mathbf{v_2}] = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} : a,b \text{ reais} \right\}$.

## 4.7 DEPENDÊNCIA E INDEPENDÊNCIA LINEAR

Uma vez entendidos os conceitos de combinação linear e de subespaços gerados, torna-se imprescindível o entendimento dos conceitos de dependência e independência linear, que permitem avaliar quando um determinado vetor pode ser descrito como uma combinação linear de outros.

**Definição 4.12:** Um conjunto $M = \{v_1,...,v_n\} \subset V$ é *linearmente independente* (LI) se, e somente se, ao se tomar qualquer combinação linear do vetor nulo com os elementos de $M$, obtivermos somente a solução $\alpha_i = 0$, $i$=1, 2,..., $n$. Ou seja, $\sum_{i=1}^{n} \alpha_i v_i = 0 \Rightarrow \alpha_1 = \alpha_2 = \cdots = 0$.

Se existir algum $\alpha_i \neq 0$, que satisfaça a combinação linear dos vetores de $M$, dizemos que $\{v_1, ..., v_n\}$ é *linearmente dependente* (LD), ou que os vetores $v_1, ..., v_n$ são LD.

**Exemplo 4.18:** No espaço vetorial $V = IR^3$, os vetores $v_1 = (2, -1, 3)$, $v_2 = (-1, 0, -2)$ e $v_3 = (2, -3, 1)$ formam um conjunto linearmente dependente, pois $3v_1 + 4v_2 - v_3 =$ $= 0$, ou seja:

$$3(2, -1, 3) + 4\,(-1, 0, -2) - (2, -3, 1) = (0, 0, 0).$$

**Propriedade 4.6:** Se um conjunto finito $L \subset C$ contém o vetor nulo, então esse conjunto é LD.

**Propriedade 4.7:** Se $S = \{u\} \subset V$ e $u \neq 0$, então $S$ é LI.

**Propriedade 4.8:** $S = \{u_1, u_2, ..., u_n\} \subset V$ é LD se, e somente se, um dos seus vetores é combinação linear dos outros.

**Propriedade 4.9:** Se $S_1$ e $S_2$ são subconjuntos finitos e não vazios de $V$ com $S_1 \subset S_2$ e $S_1$ LD, então $S_2$ também é LD.

**Propriedade 4.10:** Se $S_1$ e $S_2$ são subconjuntos finitos e não vazios de $V$ com $S_1 \subset S_2$ e $S_2$ LI, então $S_1$ também é LI.

É importante observar que, se um conjunto com $n \geq 2$ elementos é LI, ao se retirar um vetor desse conjunto, o novo conjunto continua sendo LI, o que não ocorre necessariamente no caso de um conjunto LD. Pode-se retirar um vetor de um conjunto LD, e o novo conjunto continuar sendo LD ou se tornar LI.

**Propriedade 4.11:** Se $S = \{v_1, ..., v_n\}$ é LI, e para um certo $v \in V$ tivermos $S \cup \{v\} = \{v_1, ..., v_n, v\}$ LD, então, o vetor $v$ é combinação linear dos vetores de $S$, isto é, $v \in [S]$.

A propriedade diz que, ao se acrescentar um vetor de *V* no conjunto *S*, e se por consequência o novo conjunto for *linearmente dependente*, então *v* pertence ao subespaço gerado por *S*, ou seja, *v* é combinação linear dos vetores de *S*.

**Propriedade 4.12:** Se $S = \{u_1, u_2, ..., u_j, ..., u_n\}$ e $u_j \in [S - \{u_j\}]$ (isto é, $u_j$ é combinação linear dos vetores de *S*, então $[S] = [S - \{u_j\}]$.

**Exemplo 4.19:** Sejam V = $\mathbb{R}^3$ e $\mathbf{v_1},\mathbf{v_2},\mathbf{v_3} \in$ V. Então o conjunto $\{\mathbf{v_1},\mathbf{v_2},\mathbf{v_3}\}$ é LD se estes três vetores estiverem no mesmo plano que passa pela origem. (A verificação deste fato é deixada a cargo do leitor!).

**Exemplo 4.20:** Sejam V = $\mathbb{R}^2$, $e_1 = (1,0)$ e $e_2 = (0,1)$. Então temos que $e_1$ e $e_2$ são LI, pois $a_1e_1 + a_2e_2 = \mathbf{0} \Rightarrow a_1(1,0) + a_2(0,1) = (0,0) \Rightarrow (a_1,a_2) = (0,0) \Rightarrow a_1 = a_2 = 0$.

**Exemplo 4.21:** Analogamente, para V= $\mathbb{R}^3$ e $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ e $e_3 = (0,0,1)$ temos que $e_1$, $e_2$ e $e_3$ são LI. (A verificação deste fato é deixada a cargo do leitor!).

**Exemplo 4.22:** Sejam V = $\mathbb{R}^2$ e o conjunto P = {(1,–1),(1,0),(1,1)}. Então, o conjunto P é LD, pois $\frac{1}{2}(1,-1) - 1(1,0) + \frac{1}{2}(1,1) = (0,0)$.

## 4.8 BASE DE UM ESPAÇO VETORIAL

Outra característica interessante a ser estudada é a busca por um conjunto finito de vetores, de modo que qualquer outro vetor do espaço possa ser escrito como uma combinação linear destes. Isto nos leva à determinação de um conjunto finito de vetores que gera todo o espaço vetorial, o que é muito interessante, pois esses vetores do conjunto (finito) são suficientes para descrever todos os outros elementos do espaço. Daí a noção de **base** de um espaço vetorial, definida a seguir.

**Definição 4.13:** Um conjunto $\{\mathbf{v_1}, ..., \mathbf{v_n}\}$ de vetores de um espaço vetorial V será uma **base** de V, se este conjunto for LI e se ele gerar o espaço, ou seja:

(i) $\{\mathbf{v_1}, ..., \mathbf{v_n}\}$ é LI

(ii) $[\mathbf{v_1}, ..., \mathbf{v_n}] = V$

Não se tem a garantia de que todo espaço vetorial é gerado por um número mínimo de vetores, isto é, nada garante que todo espaço vetorial possui uma *base*. Pode-se afirmar que todo espaço vetorial possui uma base logo após a demonstração do seguinte teorema:

**Teorema 4.3:** *Todo espaço vetorial finitamente gerado possui uma base.*

**Prova:** Se considerarmos $V = \{0\}$, devido à convenção feita a respeito desse assunto, o conjunto $\phi$ é uma base de $V$. Caso contrário, existe um subconjunto finito e não vazio$S \subset V$, de maneira que $[S] = V$. Como $S \neq \{0\}$, existem subconjuntos não vazios de $S$ que são LI. Considere um dos subconjuntos LI de $S$ com o maior número possível de elementos. Indicando por $B$ esse subconjunto, $B$ é uma base de $V$, pois, devido à maneira como foi tomado, ao se considerar um vetor $u \in S - B$ , tem-se que $B \cup \{u\}$ é LD e, portanto, $u \in [B]$ donde se conclui que $[B] = S$ . Se $[B] = S$ e $[S] = V$, então$[B] = V$.

A seguir será enunciado um resultado importante, que diz respeito ao número de vetores de uma ou mais bases de um espaço vetorial finitamente gerado.

**Teorema 4.4 (Invariância):** *Seja V um espaço vetorial finitamente gerado. Então duas bases quaisquer de V têm o mesmo número de vetores.*

**Definição 4.14:** Seja $V$ um espaço finitamente gerado sobre um corpo $C$. Denomina-se *dimensão de V* (dim $V$) o número de vetores de qualquer uma de suas bases. Diz-se também, neste caso, que $V$ é um *espaço de dimensão finita*.

Sem o teorema da invariância, a definição de dimensão fica inconsistente. Outro resultado utilizado na demonstração do teorema da dimensão é o chamado **teorema do completamento**. Esse teorema propõe que, se um conjunto finito LI de um espaço vetorial não for base, este pode ser completado até se tornar uma.

**Exemplo 4.23:** Considere o espaço bidimensional V = $\mathbb{R}^2$. Os conjuntos {(1,0),(0,1)} e {(1,1),(0,1)} são bases de V. Então, dimV = 2.

**Teorema 4.5 (Completamento):** *Seja V um espaço vetorial de dimensão $n \geq 1$ sobre um corpo de escalares. Se $D = \{u_1, ..., u_r\} \subset V$ é um conjunto LI com $r < n$ vetores, então existem $n - r$ vetores de V, de maneira que $B = \{u_1, ..., u_r, u_{r+1}, ..., u_n\}$ é uma base de V.*

**Exemplo 4.24:** Considere o plano V = $\mathbb{R}^2$, e os vetores $\mathbf{e_1}$ = (1,0) e $\mathbf{e_2}$ = (0,1). Então, o conjunto {$\mathbf{e_1}$,$\mathbf{e_2}$} é base de V, conhecida como **base canônica** de $\mathbb{R}^2$, e os vetores $\mathbf{e_1}$ e $\mathbf{e_2}$ são chamados de **vetores canônicos do plano**. O conjunto {(1,1),(0,1)} também é uma base de V = $\mathbb{R}^2$.

De fato, se (0,0) = a(1,1) + b(0,1) = (a,a +b), então, a = b = 0.

Isto é, {(1,1),(0,1)} é LI. Temos ainda que estes vetores geram o plano, ou seja, [(1,1),(0,1)] = V, pois dado $\mathbf{v}$ = (x,y) $\in$ V, temos (x,y) = x(1,1) + (y–x)(0,1), ou seja, todo vetor de $\mathbb{R}^2$ é uma combinação linear dos vetores (1,1) e (0,1).

**Exemplo 4.25:** O conjunto de vetores {(0,3),(0,2)} não é base de $\mathbb{R}^2$, pois é um conjunto LD. De fato, se (0,0) = a(0,3) + b(0,2), temos que 3a = −2b e a e b não são necessariamente zero, a saber, a = 2 e b = −3.

**Exemplo 4.26:** Seja V = $\mathbb{R}^3$ o espaço tridimensional. Então, o conjunto {(1,0,0),(0,1,0),(0,0,1)} é uma base de $\mathbb{R}^3$. Esta é a **base canônica** de $\mathbb{R}^3$. É fácil verificar que $\{e_1,e_2,e_3\}$ é LI e $(x,y,z) = xe_1 + ye_2 + ze_3$, ou seja, $V = \mathbb{R}^3 = [e_1,e_2,e_3]$.

**Exemplo 4.27:** O conjunto $\{(1,0,0),(0,1,0)\}$ não é base de $\mathbb{R}^3$, pois, apesar de ser LI, não gera todo $\mathbb{R}^3$, isto é, $[(1,0,0),(0,1,0)] \neq \mathbb{R}^3$; por exemplo, não conseguimos escrever o vetor $(0,0,3)$ como combinação linear de $(1,0,0)$ e $(0,1,0)$, ou seja, não encontramos a e b que satisfaçam a equação $a(1,0,0) + b(0,1,0) = (0,0,3)$.

**Exemplo 4.28:** Seja o espaço das matrizes quadradas de ordem 2, $V = M(2,2)$. Podemos verificar facilmente que o conjunto formado pelas matrizes $\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\ & \end{bmatrix}, \begin{bmatrix}0 & 0\\ & \end{bmatrix}, \begin{bmatrix}0 & 0\\ & 0\end{bmatrix}$ é uma base de V.

Quando trabalhamos com espaços de funções, podemos encontrar espaços que não possuem base finita, em que é necessário um conjunto infinito de vetores para gerar o espaço. Isto não significa lidar com combinações lineares infinitas, mas, sim, que cada vetor do espaço é uma combinação linear finita daquela "base infinita". Desta forma, escrevemos cada vetor com uma quantidade finita de vetores da "base infinita". Porém, aqui, consideraremos sempre espaços vetoriais que tenham base finita.

A seguir são apresentados três importantes teoremas, cuja demonstração será deixada a cargo do leitor.

**Teorema 4.6:** *Seja um espaço vetorial V gerado por um conjunto finito de vetores* $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v}_n$. *Então, qualquer conjunto com mais de n vetores é necessariamente LD.*

**Exemplo 4.29:** Uma vez considerada a base canônica de $\mathbb{R}^3$, é direto que, $\dim\mathbb{R}^3 = 3$.

**Exemplo 4.30:** Seja $V = M(2,2)$ o espaço das matrizes quadradas de ordem 2; então, de acordo com o Exemplo 27 da Seção 4.9, sua base tem 4 elementos e, portanto, $\dim V = 4$.

**Definição 4.15:** Quando um espaço vetorial V admite uma base finita, dizemos que V é um **espaço vetorial de dimensão finita**.

**Exemplo 4.31:** Suponha que a dimensão de um determinado espaço V é dimV = 3. Se encontramos um conjunto de 3 vetores {**v**, **w, z**}, onde **v**, **w** e **z** são LI, podemos afirmar que eles formam uma base e, portanto, V = [**v**, **w**, **z**], ou seja, V é gerado por **v**, **w** e **z**.

**Teorema 4.7:** *Se U e W são subespaços de um espaço vetorial V que tem dimensão finita, então dimU ≤ dimV e dimW ≤ dimV. Além disso, dim(U + W) = dimU + + dimW – dim(UÇW).*

**Teorema 4.8:** *Dada uma base β = {$\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$} de V, cada vetor de V é escrito de maneira única como combinação linear de $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$.*

**Definição 4.16:** Sejam β = {$\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$} base de V e **v** ∈ V, onde $\mathbf{v} = a_1\mathbf{v_1} + \ldots + a_n\mathbf{v_n}$. Chamamos os números $a_1, \ldots, a_n$ de **coordenadas** de **v** em relação à base β e denotamos por

$$[v]_\beta = \begin{bmatrix} a_1 \\ M \\ a_n \end{bmatrix}$$

**Exemplo 4.32:** Seja o espaço vetorial V = $\mathbb{R}^2$. Se β = {(1,0),(0,1)} é a base canônica de V, então o vetor (3,1) pode ser escrito como (3,1) = 3(1,0) + (1)(0,1). De maneira mais geral, um vetor genérico (x,y) é escrito como (x,y) = x(1,0) + y(0,1). Portanto $[(3,1)]_\beta = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ e $[(x,y)]_\beta = \begin{bmatrix} x \\ y \end{bmatrix}$. Se β' = {(1,1),(0,1)}, então (3,1) = x(1,1) + + y(0,1), resultando x = 3 e y = –2. Então, (3,1) = 3(1,1) – 2(0,1), donde $[(3,1)]_{\beta'} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

É relevante ressaltar que a ordem dos elementos de uma base influi diretamente na matriz das coordenadas de um vetor em relação a esta base. Por exemplo, quando temos:

$\beta_1 = \{(1,0),(0,1)\}$ e $\beta_2 = \{(0,1),(1,0)\}$, então $\left[(x,y)\right]_{\beta_1} = \begin{bmatrix} x \\ y \end{bmatrix}$ mas $\left[(x,y)\right]_{\beta_2} = \begin{bmatrix} y \\ x \end{bmatrix}$.

Em virtude disto, ao considerarmos uma base $\beta = \{\mathbf{v_1},\mathbf{v_2}, ..., \mathbf{v_n}\}$, estaremos sempre considerando que a base seja ordenada, isto é, que os vetores estão ordenados na ordem em que aparecem.

**Exemplo 4.33:** Considerando $V = \{(x,y,z): x + y - z = 0\}$ e $W = \{(x,y,z): x = y\}$ vamos determinar V + W. Observe que $V = [(1,0,1),(0,1,1)]$ e $W = [(1,1,0),(0,0,1)]$, então, $V + W = [(1,0,1),(0,1,1),(1,1,0),(0,0,1)]$. Dado $(x,y,z) \in \mathbb{R}^3$ podemos escrever:

$$(x,y,z) = \alpha(1,0,1) + \beta(0,1,1) + \gamma(1,1,0) + \delta(0,0,1).$$

Considere que $\alpha = x$, $\beta = y$, $\gamma = 0$ e $\delta = z - x - y$ (resolva o sistema para obter estes valores). Observe que a solução deste sistema não é única, uma vez que 4 vetores em $\mathbb{R}^3$ são necessariamente LD. Portanto, $V + W = \mathbb{R}^3$. De acordo com o Teorema 4.7 temos:

dim$\mathbb{R}^3$ = dimV + dimW – dim(VÇW) $\Rightarrow$ dim(VÇW) = 1.

E assim, VÇW = $\{(x,y,z): x + y - z = 0$ e $x = y\} = \{(x,y,z): x = y = z/2\} = [(1,1,1/2)]$.

## 4.9 MUDANÇA DE BASE

Como vimos anteriormente, um dado espaço vetorial pode ser representado por várias bases diferentes. A pergunta que vem neste momento é: como os vetores escritos nestas bases se relacionam? Assim, nesta seção veremos como relacionar os vetores de um espaço vetorial escritos em bases diferentes. Sejam

$\beta = \{\mathbf{u_1},...,\mathbf{u_n}\}$ e $\beta' = \{\mathbf{w_1},...,\mathbf{w_n}\}$ duas bases ordenadas de um mesmo espaço vetorial. Dado um vetor $v \in V$, podemos escrevê-lo como $\mathbf{v} = x_1\mathbf{u_1} + ... + x_n\mathbf{v_n}$ e $\mathbf{v} = y_1\mathbf{w_1} + ... + y_n\mathbf{w_n}$, respectivamente nas bases $\beta$ e $\beta'$. Como podemos relacionar as coordenadas de v em relação à base $\beta$ com as coordenadas do mesmo vetor v em relação à base $\beta'$? De modo que

$$[v]_\beta = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ e } [v]_{\beta'} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Como $\{\mathbf{u_1},...,\mathbf{u_n}\}$ é base de V, podemos escrever os vetores $\mathbf{w_i}$ como combinação linear dos $\mathbf{u_j}$, ou seja:

$$\begin{cases} w_1 = a_{11}u_1 + a_{21}u_2 + \ldots + a_{n1}u_n \\ w_2 = a_{12}u_1 + a_{22}u_2 + \ldots + a_{n2}u_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ w_n = a_{1n}u_1 + a_{2n}u_2 + \ldots + a_{nn}u_n \end{cases}$$

Substituindo estas últimas equações em $\mathbf{v} = y_1\mathbf{w_1} + ... + y_n\mathbf{w_n}$ temos:

$$\mathbf{v} = y_1\mathbf{w_1} + ... + y_n\mathbf{w_n} = y_1(a_{11}\mathbf{u_1} + ... + a_{n1}\mathbf{u_n}) + ... + y_n(a_{1n}\mathbf{u_1} + ... + a_{nn}\mathbf{u_n})$$
$$= (a_{11}y_1 + ... + a_{1n}y_n)\mathbf{u_1} + ... + (a_{n1}y_1 + ... + a_{nn}y_n)\mathbf{u_n}$$

Mas $\mathbf{v} = x_1\mathbf{u_1} + ... + x_n\mathbf{u_n}$ e, como as coordenadas em relação a uma base são únicas, temos:

$$x_1 = a_{11}y_1 + a_{12}y_2 + ... + a_{1n}y_n$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x_n = a_{n1}y_1 + a_{n2}y_2 + ... + a_{nn}y_n$$

Em forma matricial

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

onde, denotando que $[I]_{\beta}^{\beta'} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, temos $[v]_{\beta} = [I]_{\beta}^{\beta'}[v]_{\beta'}$.

**Definição 4.17:** A matriz $[I]_{\beta}^{\beta'}$, obtida conforme o procedimento anterior, é chamada **matriz de mudança da base β' para a base β**.

Comparando $[I]_{\beta}^{\beta'}$ com $\begin{cases} w_1 = a_{11}u_1 + a_{21}u_2 + \ldots + a_{n1}u_n \\ w_2 = a_{12}u_1 + a_{22}u_2 + \ldots + a_{n2}u_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ w_1 = a_{1n}u_1 + a_{2n}u_2 + \ldots + a_{nn}u_n \end{cases}$ observamos que esta matriz é obtida, colocando as coordenadas em relação a β dos $\mathbf{w_i}$ na i-ésima coluna. Note que, uma vez obtida $[I]_{\beta}^{\beta'}$, podemos encontrar as coordenadas de qualquer vetor **v** em relação à base β, multiplicando a matriz pelas coordenadas de **v** na base β' (supostamente conhecidas).

**Exemplo 4.34:** Sejam β = {(2,–1),(3,4)} e β' = {(1,0),(0,1)} bases de $\mathbb{R}^2$. Procuremos inicialmente $[I]_{\beta}^{\beta'}$. Como $\mathbf{w_1}$ = (1,0) = $a_{11}$(2, –1) + $a_{21}$(3,4) = (2$a_{11}$ + + 3$a_{21}$, – $a_{11}$ + 4$a_{21}$), obtemos o sistema:

$$\begin{cases} 2a_{11} + 3a_{21} = 1 \\ -a_{11} + 4a_{21} = 0 \end{cases}$$

donde $a_{11} = \dfrac{4}{11}$ e $a_{21} = \dfrac{1}{11}$.

Por outro lado, $\mathbf{w_2}$ = (0,1) = $a_{12}$(2,–1) + $a_{22}$(3,4) = (2$a_{12}$ + 3$a_{22}$, –$a_{12}$ + 4$a_{22}$), o que fornece o sistema:

$$\begin{cases} 2a_{12} + 3a_{22} = 0 \\ -a_{12} + 4a_{22} = 1 \end{cases}$$

Donde, resolvendo a operação, obtemos $a_{12} = \frac{-3}{11}$ e $a_{22} = \frac{2}{11}$. Portanto,

$$[I]^{\beta'}_{\beta} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \frac{4}{11} & \frac{-3}{11} \\ \frac{1}{11} & \frac{2}{11} \end{bmatrix}.$$

Podemos usar esta matriz para encontrar, por exemplo, $[\mathbf{v}]_{\beta}$ para $\mathbf{v} = (5,-8)$:

$$\left[(5,-8)\right]_{\beta} = [I]^{\beta'}_{\beta}\left[(5,-8)\right]_{\beta} = \begin{bmatrix} \frac{4}{11} & \frac{-3}{11} \\ \frac{1}{11} & \frac{2}{11} \end{bmatrix} \begin{bmatrix} 5 \\ -8 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

Isto é, (5,–8) = 4(2,–1) – 1(3,4).

É claro que, se o problema fosse apenas encontrar as coordenadas de (5,–8) em relação à base β, poderíamos simplesmente resolver o sistema (5,–8) = a(2,–1) + + b(3,4); porém, o cálculo feito através da matriz de mudança de base é operacionalmente vantajoso quando trabalhamos com mais vetores, pois neste caso não teremos de resolver um sistema de equações para cada vetor.

Se no início desta seção tivéssemos começado escrevendo os $\mathbf{u}_i$ em função dos $\mathbf{w}_j$, chegaríamos à relação análoga: $[v]_{\beta'} = [I]^{\beta}_{\beta'}[v]_{\beta}$. É fácil ver que as matrizes $[I]^{\beta}_{\beta'}$ e $[I]^{\beta'}_{\beta}$ são inversíveis e $\left([I]^{\beta'}_{\beta}\right)^{-1} = [I]^{\beta}_{\beta'}$.

**Teorema 4.9:** *A matriz* $[I]^{\beta'}_{\beta}$*, obtida conforme o procedimento anterior, é a* **inversa da matriz de mudança da base** $[I]^{\beta}_{\beta'}$*, e vice-versa.*

**Exemplo 4.35:** No Exemplo 34, podemos obter $[I]_{\beta}^{\beta'}$ a partir de $[I]_{\beta'}^{\beta}$. Note que $[I]_{\beta'}^{\beta}$ é fácil de ser calculada, pois β’ é a base canônica. Assim, $(2,-1) = 2(1,0) - 1(0,1)$ e $(3,4) = 3(1,0) + 4(0,1)$ donde $[I]_{\beta'}^{\beta} = \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}$.

Então $[I]_{\beta}^{\beta'} = \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{4}{11} & \frac{-3}{11} \\ \frac{1}{11} & \frac{2}{11} \end{bmatrix}$.

**Exemplo 4.36:** Consideremos em $R^2$ a base $\beta = \{\mathbf{e_1},\mathbf{e_2}\}$ e a base $\beta' = \{\mathbf{f_1},\mathbf{f_2}\}$ obtida da base canônica β pela rotação de um ângulo θ (Figura 4.8). Dado um vetor $\mathbf{v} \in R^2$ de coordenadas $[v]_{\beta} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ em relação à base β, quais são as coordenadas $[v]_{\beta'} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ em relação à base β’?

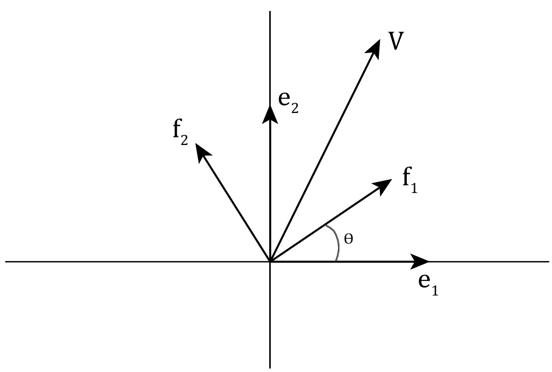
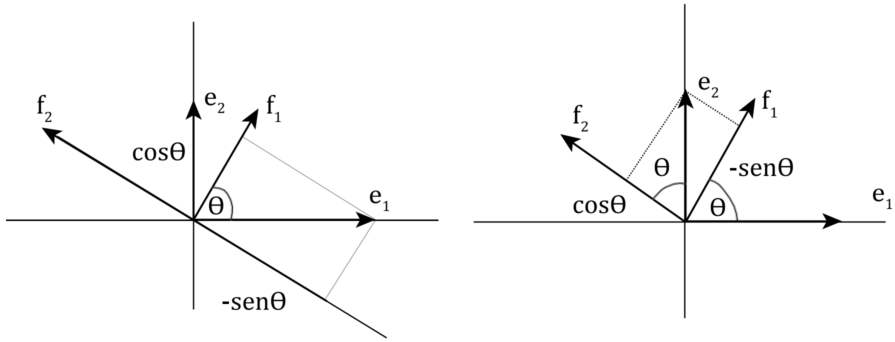



FIGURA 4.8 - Mudança de coordenadas de vetor no plano

Temos, então, $\mathbf{v} = x_1\mathbf{e_1} + x_2\mathbf{e_2} = y_1\mathbf{f_1} + y_2\mathbf{f_2}$ e queremos calcular $[v]_{\beta'} = [I]_{\beta'}^{\beta}[v]_{\beta}$, ou seja, temos que achar a matriz $[I]_{\beta'}^{\beta}$. Para isto, devemos escrever $\mathbf{e_1}$ e $\mathbf{e_2}$ em função de $\mathbf{f_1}$ e $\mathbf{f_2}$ (Figuras 4.9a e 4.9b):

$$\mathbf{e_1} = \cos\theta\mathbf{f_1} - \operatorname{sen}\theta\mathbf{f_2}$$

$$\mathbf{e_2} = \operatorname{sen}\theta\mathbf{f_1} + \cos\theta\mathbf{f_2}$$



(a) (b)

FIGURA 4.9 - Vetores representados em outra base no plano

Portanto, $[I]^{\beta}_{\beta'} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$ donde $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, ou seja:

$$y_1 = x_1 \cos\theta + x_2 \operatorname{sen}\theta$$
$$y_2 = -x_1 \operatorname{sen}\theta + x_2 \cos\theta.$$

No Exemplo 36, em particular para $\theta = \pi/3$, para $\mathbf{v} = (-2,3)$, isto é $\mathbf{v} = -2\mathbf{e_1} + 3\mathbf{e_2}$, temos $[v]_{\beta'} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ e queremos determinar as coordenadas de **v** na base $\beta' = \{\mathbf{f_1}, \mathbf{f_2}\}$ (Figura 4.10).

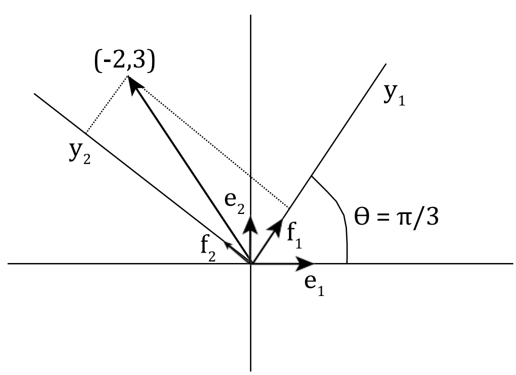


FIGURA 4.10 - Mudança de base

Como vimos, $[v]_{\beta'} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ onde:

$$y_1 = x_1\cos\theta + x_2\text{sen}\theta = -2\cos(\pi/3) + 3\text{sen}(\pi/3)$$
$$y_2 = -x_1\text{sen}\theta + x_2\cos\theta = 2\text{sen}(\pi/3) + 3\cos(\pi/3)$$

donde $[v]_{\beta'} = \begin{bmatrix} \dfrac{-2+3\sqrt{3}}{2} \\ \dfrac{3+2\sqrt{3}}{2} \end{bmatrix}$, ou seja $\mathbf{v} = \left(\dfrac{-2+3\sqrt{3}}{2}\right)f_1 + \left(\dfrac{3+2\sqrt{3}}{2}\right)f_2$.

# 5. TRANSFORMAÇÕES LINEARES

## 5.1 INTRODUÇÃO

As funções lineares descrevem o tipo mais simples de dependência entre variáveis e muitos problemas podem ser representados por estas funções.

Sejam dois conjuntos $U$ e $V$ , ambos não vazios (Figura 5.1). Uma aplicação de $U$ em $V$ é uma relação pela qual cada elemento de $U$ está associado a um único elemento de $V$. Se $F$ indica essa lei, e $u$ indica um elemento genérico de $U$, então o elemento associado a $u$ é representado por $F(u)$ (lê-se: $F$ aplicado em $u$) e se denomina *imagem* de $u$ por $F$.

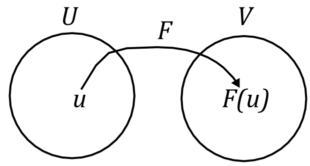



FIGURA 5.1 - Imagem de aplicação

O conjunto $U$ é o domínio e o conjunto $V$ é o contradomínio da aplicação $F$. Para indicar que $F$ é uma aplicação de $U$ em $V$, costuma-se escrever $F: U \to V$, ou, ainda, indicando por $u$ o elemento genérico de $U$, escrever, $u$ ▪ $F(u)$.

## 5.2 CONCEITOS E TEOREMAS

**Definição 5.1:** Duas aplicações $F: U \to V$ e $G: U \to V$ são iguais se, e somente se, $F(u) = G(u)$, $\forall\ u \in U$.

**Definição 5.2:** Dado $W \subset U$ denomina-se *imagem* de $W$ por $F$ o seguinte subconjunto de $V$: $F(W) = \{F(u): u \in W\}$. Se $U = W$, então $F(U)$ recebe o nome de imagem de $F$ e a notação será $\text{Im}(F)$ (Figura 5.2). Portanto, $\text{Im}(F) = \{F(u): u \in U\}$.

**Definição 5.3:** Uma aplicação $F: U \to V$ se diz *injetora* se, e somente se, $\forall\ u_1, u_2 \in U$, $F(u_1) = F(u_2) \Rightarrow u_1 = u_2$. Ou, em outra formulação, se, e somente se, $\forall\ u_1, u_2 \in U$, $u_1 \neq u_2 \Rightarrow F(u_1) \neq F(u_2)$.

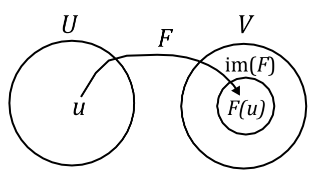



FIGURA 5.2 - Aplicação injetora

**Definição 5.4:** Uma aplicação $F: U \to V$ se diz *sobrejetora* se, e somente se, $\text{Im}(F) = V$ (Figura 5.3), ou seja, para todo $v \in V$, existe $u \in U$ tal que $v = F(u)$.

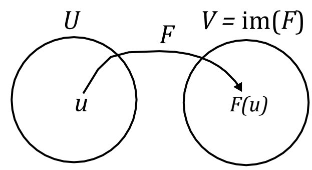



FIGURA 5.3 - Aplicação sobrejetora

**Definição 5.5:** Uma aplicação $F: U \to V$ se diz *bijetora* se, e somente se, $F$ é *injetora* e *sobrejetora* simultaneamente.

**Definição 5.6:** Sejam $U$ e $V$ espaços vetoriais sobre um corpo $C$. Uma aplicação $F: U \to V$ é chamada *transformação (ou aplicação) linear de U em V* se, e somente se:

$$\forall\, u,v \in \mathrm{V},\ \mathrm{F}(u+v) = \mathrm{F}(u) + \mathrm{F}(v)$$

$$\forall\ \mathrm{k} \in \mathbb{R} \text{ e } v \in \mathrm{V},\ \mathrm{F}(\mathrm{k}v) = \mathrm{k}\ \mathrm{F}(v)$$

ou simplesmente, $F(u+\alpha v) = F(u) + \alpha F(v)$ .

Sejam $U$ e $V$ espaços vetoriais sobre um corpo $C$ e uma transformação linear $F: U \to V$. Da definição de transformação linear, decorrem as seguintes propriedades:

**Propriedade 5.1:** Uma transformação linear $F$ transforma o vetor nulo de $U$ no vetor nulo de $V$, ou seja, $\mathrm{F}(\mathbf{0}) = \mathbf{0}$.

**Propriedade 5.2:** $F(-u) = -\mathrm{F}(u),\ \forall\ u \in U$.

**Propriedade 5.3:** $F(u_1 - u_2) = F(u_1) - F(u_2),\ \forall\ u_1, u_2 \in U$.

**Propriedade 5.4:** Se $W$ é um subespaço de $U$, então a imagem de $W$ por $F$ é um subespaço de $V$.

Tem-se aqui a importante afirmação de que uma *transformação linear* leva subespaço do domínio em subespaço do contradomínio, ou seja, uma *transformação linear* preserva a *estrutura de subespaço vetorial*.

**Propriedade 5.5:** Sendo *F: U* $\to$ *V* linear, então, $F\left(\sum_{i=1}^{n} \alpha_i u_i\right) = \sum_{i=1}^{n} \alpha_i F(u_i)$.

## 5.3 NÚCLEO E IMAGEM DE UMA TRANSFORMAÇÃO

**Definição 5.7:** Sejam *U* e *V* espaços vetoriais sobre um corpo *C* e *F: U* $\to$ *V* uma transformação linear. Indica-se por $Ker(F)$ (Figura 5.4) e denomina-se *núcleo* de *F* o seguinte subconjunto de *U*:

$$Ker(F) = \{u \in U : F(u) = 0\}$$

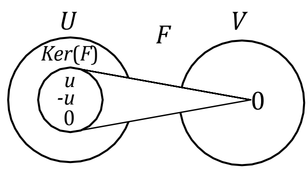



FIGURA 5.4 - Núcleo de uma transformação

O núcleo de uma transformação linear é um subconjunto do domínio que se caracteriza por possuir elementos que são levados sempre ao vetor nulo do contradomínio.

Decorre da definição que o núcleo de uma transformação é um conjunto não vazio, pois $F(0) = 0 \Rightarrow 0 \in Ker(F)$. Usando esse último argumento e verificando o fechamento, prova-se que o conjunto $Ker(F)$ constitui um subespaço vetorial de *U*.

**Teorema 5.1:** *Se U e V são espaços vetoriais sobre um corpo C e F: U* $\to$ *V é uma transformação linear, então* $Ker(F)$ *é um subespaço vetorial de U.*

**Teorema 5.2:** *Se U e V são espaços vetoriais sobre um corpo C e F: U* $\to$ *V é uma transformação linear, então F é injetora se, e somente se,* $Ker(F) = \{0\}$.

**Teorema 5.3: (Dimensão):** *Sejam U e V espaços vetoriais ambos de dimensão finita sobre C. Dada uma transformação F: U → V linear, então:*

$$dimU = dimKer(F) + dimIm(F).$$

**Prova:** Seja $B = \{u_1, ..., u_r\}$ uma base de $Ker(F)$. Pelo ***teorema do completamento*** pode-se considerar $C = \{u_1, ..., u_r, v_1, ..., v_s\}$ como sendo uma base de $U$. Basta mostrar agora que o conjunto $D = \{F(v_1), ..., F(v_s)\}$ é base de $\mathrm{Im}(F)$ e o teorema estará demonstrado.

Dados $\alpha_i, \beta_j \in C$, $i = (1,..., r)$, $j = (1,..., s)$ e definindo $u = \alpha_1 u_1 + ... + \alpha_r u_r + \beta_1 v_1 + ... + \beta_s v_s \in U$ tal que $F(u) = v \in \mathrm{Im}(F)$, tem-se o seguinte argumento:

$v = F(u) = F(\alpha_1 u_1 + ... + \alpha_r u_r + \beta_1 v_1 + ... + \beta_s v_s)$

$\Leftrightarrow v = \alpha_1 F(u_1) + ... + \alpha_r F(u_r) + \beta_1 F(v_1) + ... + \beta_s F(v_s)$.

Como o conjunto $B = \{u_1, ..., u_r\}$ é base de $Ker(F)$,

$v = \alpha_1 0 + ... + \alpha_r 0 + \beta_1 F(v_1) + ... + \beta_s F(v_s)$, $F(u_i) = 0$, $i = (1,..., r)$.

Com isso $v = \beta_1 F(v_1) + ... + \beta_s F(v_s)$, Logo, $[D] = \mathrm{Im}(F)$.

Para mostrar que $D$ é um conjunto linearmente independente, considere a igualdade $\beta_1 F(v_1) + ... + \beta_s F(v_s) = 0$.

$\beta_1 F(v_1) + ... + \beta_s F(v_s) = 0 \Leftrightarrow F(\beta_1 v_1 + ... + \beta_s v_s) = 0$, ou seja, $\beta_1 v_1 + ... + \beta_s v_s \in Ker(F)$.

Ora, se este vetor pertence ao conjunto mencionado, pode-se escrevê-lo como sendo uma combinação linear da base $B$.

$\beta_1 v_1 + ... + \beta_s v_s = \lambda_1 u_1 + ... + \lambda_r u_r \Leftrightarrow \beta_1 v_1 + ... + \beta_s v_s - \lambda_1 u_1 - \lambda_r u_r = 0$.

Como o conjunto $C$ é L.I., consequentemente, $\beta_1 = ... = \beta_s = \lambda_1 = ... = \lambda_r = 0$, o que prova a tese apresentada.

Observe que, de acordo com a estrutura que foi montada, $\dim U = r + s$, $\dim Ker(F) = r$ e $\dim \mathrm{Im}(F) = s$, portanto, $\dim U = \dim Ker(F) + \dim \mathrm{Im}(F)$..

**Exemplo 5.1:** As seguintes funções são aplicações lineares:

(i) Sejam V = W = R e a aplicação Q: R → R, dada por Q(x) = 0,2x. Então, Q é uma aplicação linear, pois Q(x + y) = 0,2(x + y) = 0,2x + + 0,2y = Q(x) + Q(y) e Q(kx) = 0,2(kx) = k(0,2x), para quaisquer x, y e k reais.

(ii) Sejam V = $R^4$, W = R e a aplicação Q: $R^4 \to$ R, dada por $Q\left(\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}\right) = \begin{bmatrix} 5 & -0{,}1 & 0{,}1 & 25 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$. Então, Q é uma aplicação linear, pois, para quaisquer $(x_1,y_1,z_1,w_1)$ e $(x_2,y_2,z_2,w_2)$ em $R^4$ e k em R: obtém-se que

$$Q\left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \\ w_2 \end{bmatrix}\right) = \begin{bmatrix} 5 & -0{,}1 & 0{,}1 & 25 \end{bmatrix}\left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \\ w_2 \end{bmatrix}\right) =$$

$$= \begin{bmatrix} 5 & -0{,}1 & 0{,}1 & 25 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{bmatrix} + \begin{bmatrix} 5 & -0{,}1 & 0{,}1 & 25 \end{bmatrix}\begin{bmatrix} x_2 \\ y_2 \\ z_2 \\ w_2 \end{bmatrix}$$

e

$$Q\left(k\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{bmatrix}\right) = \begin{bmatrix} 5 & -0{,}1 & 0{,}1 & 25 \end{bmatrix}\left(k\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{bmatrix}\right) = k\begin{bmatrix} 5 & -0{,}1 & 0{,}1 & 25 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{bmatrix}.$$

**Exemplo 5.2:** Sejam V = W = R, e a aplicação F: R $\to$ R definida por $u \mapsto \alpha u$ ou $F(u) = \alpha u$. Então F é uma aplicação linear. (Tente verificar!).

**Exemplo 5.3:** Se a aplicação F: R $\to$ R, é dada por $F(u) = u^2$, então F não é uma aplicação linear. De fato:

$$F(u+v) = (u+v)^2 = u^2 + 2uv + v^2 \neq u^2 + v^2 = F(u) + F(v).$$

**Exemplo 5.4:** Se V = $R^2$, W = $R^3$, e F: $R^2 \to R^3$, é dada por F(x,y) = (7x,0,x+y). Considerando **u** = $(x_1,y_1)$ e **v** = $(x_2,y_2)$ vetores quaisquer em V e k um escalar real, temos que F é uma transformação linear, visto que

$$\begin{aligned} F(u + v) &= F(x_1 + x_2, y_1 + y_2) = [7(x_1 + x_2), 0, (x_1,y_1) + (x_2,y_2)] = \\ &= (7x_1 + 7x_2, 0, x_1 + x_2 + y_1 + y_2) = \\ &= (7x_1, 0, x_1 + y_1) + (7x_2, 0, x_2 + y_2) = F(u) + F(v) \end{aligned}$$

e

$$F(ku) = F(kx_1,ky_1) = (7kx_1, 0, kx_1+ky_1) = k(7x_1, 0, x_1 + y_1) = kF(u).$$

**Obs.:** Decorre da definição que uma transformação linear T: V $\to$ W leva o vetor nulo de V no vetor nulo de W, isto é, se $\mathbf{0} \in V$, $T(\mathbf{0}) = \mathbf{0} \in W$. Isso ajuda a detectar transformações lineares. Se $T(\mathbf{0}) \neq \mathbf{0}$, T não é linear, porém, o fato de $T(\mathbf{0}) = \mathbf{0}$ não garante que T seja linear.

**Exemplo 5.5:** Sejam V = W = $P_n$ (polinômios de grau $\leq$ n) e D: $P_n \mapsto P_n$, (D(f) = f'), a aplicação derivada que a cada polinômio f associa sua derivada, a qual também é um polinômio (com 1 grau a menos). Assim, D é uma aplicação linear, pois, para quaisquer funções derivadas,

$$D(f+g) = D(f) + D(g) \text{ e } D(kf) = kD(f).$$

**Exemplo 5.6:** A aplicação nula F: V $\mapsto$ V ($F(u) = \mathbf{0}$) é linear. (Tente verificar!)

**Exemplo 5.7:** Sejam V = $\mathbb{R}^n$ e W = $\mathbb{R}^m$. Seja uma matriz $m \times n$. Definimos $L_A$: $\mathbb{R}^n \to \mathbb{R}^m$ por $L_A(v) = A.v$, onde $v$ é tomado como vetor coluna, $v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $L_A(v) = A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$. Das propriedades de operações de matrizes:

$L_A(u+v) = A(u+v) = Au + Av = L_A(u) + L_A(v)$ e $L_A(ku) = A(ku) = k(Au) = kL_A(u)$, e portanto $L_A$ é uma transformação linear.

Toda matriz $m \times n$ pode ser associada a uma transformação linear de $\mathbb{R}^n \to \mathbb{R}^m$. De modo geral, podemos dizer que uma matriz produz uma transformação linear, e também, uma transformação linear de $\mathbb{R}^n$ em $\mathbb{R}^m$ pode ser representada por uma matriz $m \times n$.

**Exemplo 5.8:** Consideremos a matriz $A = \begin{bmatrix} 7 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$, e a aplicação $L_A : \mathbb{R}^2 \to \mathbb{R}^3$, conforme o Exemplo 5.7. Assim, tem-se que

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} 7 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7x_1 \\ 0 \\ x_1 + x_2 \end{bmatrix}$$

E, portanto, que a aplicação $L_A(x_1,x_2) = (2x_1,0,x_1+x_2)$ é exatamente aquela do Exemplo 5.4.

**Teorema 5.4:** *Dados dois espaços vetoriais reais V e W e uma base de V, {$\mathbf{v_1}$,...,$\mathbf{v_n}$}, sejam $\mathbf{w_1}$,...,$\mathbf{w_n}$ elementos arbitrários de W, então existe uma única aplicação linear T: V $\longmapsto$ W tal que T($\mathbf{v_1}$) = $\mathbf{w_1}$, ..., T($\mathbf{v_n}$) = $\mathbf{w_n}$. Se $\mathbf{v}$ = $a_1\mathbf{v_1}$ + + $a_n\mathbf{v_n}$, esta aplicação é dada por*

$$T(\mathbf{v}) = a_1T(\mathbf{v_1}) + ..... + a_nT(\mathbf{v_n}) = a_1\mathbf{w_1} + ... + a_n\mathbf{w_n}.$$

Pode-se verificar que T, assim definida, é linear e que é a única que satisfaz as condições exigidas.

**ER 5.1:** Qual é a transformação linear T: $\mathbb{R}^2 \to \mathbb{R}^2$ tal que T(1,0) = (2,–1,0) e T(0,1) = (0,0,1)?

**Solução:** Temos neste caso $e_1$ = (1,0) e $e_2$ = (0,1) base de $\mathbb{R}^2$, $w_1$ = (2,–1,0) e $w_2$ = (0,0,1). Dado $v = (x_1,x_2)$ arbitrário, temos $v = x_1e_1 + x_2e_2$ e

$$T(v) = x_1T(e_1) + x_2T(e_2) = x_1(2,-1,0) + x_2(0,0,1) = (2x_1,-x_1,x_2).$$

**Exemplo 5.9:** Seja T: $\mathbb{R}^2 \to \mathbb{R}$, definida por T(x,y) = x+y. Neste caso, ker(T) = = {(x,y) $\in \mathbb{R}^2$; x + y = 0}, isto é, ker(T) é a reta y = –x. Podemos dizer ainda que

$$\ker(T) = \{(x,-x);\ x \in \mathbb{R}\} = \{x(1,-1);\ x \in \mathbb{R}\} = [(1,-1)] \text{ e } ImT = \mathbb{R},$$

pois dado $w \in \mathbb{R}$, $w = T(w,0)$.

**Exemplo 5.10:** Seja a transformação linear T: $\mathbb{R}^3 \to \mathbb{R}^3$ dada por T(x,y,z) = = (x,2y,0). Então, a imagem de T é dada por:

Im(T) = {(x,2y,0); x,y $\in \mathbb{R}$} = {x(1,0,0) +y(0,2,0); x,y $\in \mathbb{R}$} = {(1,0,0),(0,2,0)}.

Observe que dim(Im(T)) = 2. O núcleo de T é dado por:

$$\ker(T) = \{(x,y,z): T(x,y,z) = (0,0,0)\} = \{(x,y,z):$$
$$(x,2y,0) = (0,0,0)\} = \{(0,0,z): z \in \mathbb{R}\} =$$
$$= \{z(0,0,1): z \in \mathbb{R}\} = [(0,0,1)]$$

Observe que dim(ker(T)) = 1.

**Teorema 5.5:** *Seja T: V $\rightarrow$ W uma aplicação linear. Então, dim (kerT) + dim (ImT) = = dim V.*

**Teorema 5.6:** *Se dim V = dim W, então T linear é injetora se, e somente se, T é sobrejetora.*

**Teorema 5.7:** *Seja T: V $\rightarrow$ W uma aplicação linear injetora. Se dim V = dim W, então T leva base em base.*

**Definição 5.8:** Quando uma transformação linear T: V $\rightarrow$ W for injetora e sobrejetora, ao mesmo tempo, diz-se que ela é decorrente de **isomorfismo** e, neste caso, os dois espaços vetoriais são ditos **isomorfos**.

Em termos da álgebra linear, diz-se que dois espaços vetoriais isomorfos são idênticos, têm a mesma dimensão e levam base em base. Além disso, um isomorfismo T: V $\rightarrow$ W tem uma aplicação inversa $T^{-1}$: W $\rightarrow$ V que é linear e também é um isomorfismo.

**ER 5.2:** Seja T: $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por T(x,y,z) = (x – 2y,z,x + y). Verifique que T é um isomorfismo e calcule a sua inversa $T^{-1}$.

**Solução:** Se T for injetora, pelo Teorema 5.4, também é sobrejetora, donde pela Definição 5.6, é um isomorfismo. Isto equivale a mostrar que ker(T) = = {(0,0,0)}.

Mas, ker(T) = {(x,y,z): T(x,y,z) = (0,0,0)} e T(x,y,z) = (0,0,0) se, e somente se, (x – 2y, z, x + y) = (0,0,0). Resolvendo o sistema de equações lineares:

$$\begin{cases} x - 2y = 0 \\ z = 0 \\ x + y = 0 \end{cases}$$

encontramos x = y = z = 0 como única solução e, portanto, T é um isomorfismo.

Tomando a base canônica de $\mathbb{R}^3$, sua imagem pela T é:

$$T(\{\mathbf{e_1},\mathbf{e_2},\mathbf{e_3}\}) = \{T(1,0,0),T(0,1,0),T(0,0,1)\} = \{(1,0,1),(-2,0,1),(0,1,0)\}$$

que é ainda uma base de $\mathbb{R}^3$. Calculemos agora a aplicação inversa de T. Como

$$T(1,0,0) = (1,0,1),\ T(0,1,0) = (-2,0,1) \text{ e } T(0,0,1) = (0,1,0),$$

temos que

$$T^{-1}(1,0,1) = (1,0,0),\ T^{-1}(-2,0,1) = (0,1,0) \text{ e } T^{-1}(0,1,0) = (0,0,1).$$

Queremos calcular $T^{-1}(x,y,z)$. Para isto escrevemos $(x,y,z)$ em relação à base$\{(1,0,1),(-2,0,1),(0,1,0)\}$:

$$(x,y,z) = \frac{x+2z}{3}(1,0,1) + \frac{z-x}{3}(-2,0,1) + y(0,1,0)$$

Então, $T^{-1}(x,y,z) = \dfrac{x+2z}{3}T^{-1}(1,0,1) + \dfrac{z-x}{3}T^{-1}(-2,0,1) + yT^{-1}(0,1,0)$, ou seja:

$$T^{-1}(x,y,z) = \left(\frac{x+2z}{3}, \frac{z-x}{3}, y\right)$$

## 5.4 APLICAÇÕES LINEARES E MATRIzES

Veremos que, dados dois espaços vetoriais V e W com bases $\beta$ e $\beta'$ e uma matriz A, podemos encontrar uma transformação linear, associando $\beta$, $\beta'$ e A.

**ER 5.3:** Considerando $\mathbb{R}^2$, as bases $\beta = \{(1,0),(0,1)\}$ e $\beta' = \{(1,1),(-1,1)\}$ e a matriz $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, como se associa a esta matriz A uma aplicação linear que depende de A e das bases dadas $\beta$ e $\beta'$, isto é, $T_A: \mathbb{R}^2 \to \mathbb{R}^2$, definida por $\mathbf{v} \mapsto T_A(\mathbf{v})$?

**Solução:** Considere $\mathbf{v} = (x,y)$, um vetor qualquer do plano. Sejam $X = [\mathbf{v}]_\beta = \begin{bmatrix} x \\ y \end{bmatrix}$ e $AX = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ y \end{bmatrix} = [T_A(\mathbf{v})]_{\beta'}$. Então, $T_A(\mathbf{v}) = 2x(1,1) + y(-1,1) =$ $= (2x - y, 2x + y)$.

Note que, se $\mathbf{v} = (2,1)$, então $T_A(2,1) = (3,5)$, e se tivéssemos partido de $\beta = \beta' =$ $= \{(1,0),(0,1)\}$, teríamos obtido $T_A(\mathbf{v}) = (2x,y) = A\mathbf{v}$.

De modo geral, fixadas as bases $\beta = \{\mathbf{v}_1,...,\mathbf{v}_n\}$ e $\beta' = \{\mathbf{w}_1,...,\mathbf{w}_m\}$ à matriz

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}_{mxn}$$

podemos associar a aplicação $T_A: \mathbb{R}^n \to \mathbb{R}^m$

$\mathbf{v} \mapsto T_A(\mathbf{v})$ da seguinte forma:

$$\text{Seja } X = [\mathbf{v}]_\beta =,\ AX = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Então, $T_A(\mathbf{v}) = y_1\mathbf{w}_1 + ... + y_m\mathbf{w}_m$ onde $y_i = A_i X$ e $A_i$ é a i-ésima linha de A.

Em geral, dada uma matriz $A_{mxn}$, ela é encarada como uma aplicação linear $T_A: \mathbb{R}^n \to \mathbb{R}^m$ em relação às base canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$.

**ER 5.4:** Sejam $A = \begin{bmatrix} 1 & -3 & 5 \\ 2 & 4 & -1 \end{bmatrix}$, $\beta = \{(1,0),(0,1)\}$ e $\beta' = \{(1,0,0),(0,1,0),(0,0,1)\}$ e $T_A: \mathbb{R}^3 \to \mathbb{R}^2$. Encontremos esta transformação linear.

**Solução:** Seja $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $AX = \begin{bmatrix} 1 & -3 & 5 \\ 2 & 4 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - 3y + 5z \\ 2x + 4y - z \end{bmatrix}$. Então:

$T_A(x,y,z) = (x - 3y + 5z)(1,0) + (2x + 4y - z)(0,1) = (x - 3y + 5z, 2x + 4y - z)$.

**Exemplo 5.11:** Encontrar a matriz associada a uma transformação linear. Seja T: V $\to$ W linear, $\beta = \{\mathbf{v}_1,...,\mathbf{v}_n\}$ base de V e $\beta' = \{\mathbf{w}_1,...,\mathbf{w}_m\}$ base de W. Então $T(\mathbf{v}_1),..., T(\mathbf{v}_n)$ são vetores de W e, portanto,

$$T(\mathbf{v}_1) = a_{11}\mathbf{w}_1 + ... + a_{m1}\mathbf{w}_m$$

$$.......................................$$

$$T(\mathbf{v}_n) = a_{1n}\mathbf{w}_1 + ... + a_{mn}\mathbf{w}_m$$

A transposta da matriz dos coeficientes deste sistema, anotada por $[T]_{\beta'}^{\beta}$ é chamada matriz de T em relação às bases de β e β'.

$$[T]_{\beta'}^{\beta} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = A$$

Observe que T passa a ser a aplicação linear associada à matriz A e às bases β e β', isto é, $T = T_A$.

**Exemplo 5.12:** Sejam $T(x, y, z) = (2x + y - z, x + 2y)$ e as bases
$A = \{(1, 0, 0), (2, -1, 0), (0, 1, 1)\} \in IR^3$ e $B = \{(-1, 1), (0, 1)\} \in IR^2$

$$\begin{array}{c} \\ \\ [T]_B^A = \end{array} \begin{array}{ccc} T(v_1) & T(v_2) & T(v_3) \\ \downarrow & \downarrow & \downarrow \\ \left[\begin{array}{c} a_{11} \\ a_{21} \end{array}\right. & \begin{array}{c} a_{12} \\ a_{22} \end{array} & \left.\begin{array}{c} a_{13} \\ a_{23} \end{array}\right] \end{array}$$

$T(v_1) = T(1, 0, 0) = (2, 1) \rightarrow T(v_1) = (a_{11}w_1 + a_{21}w_2) \rightarrow a_{11}(-1, 1) + a_{21}(0, 1) = (2, 1)$

$$\begin{cases} -a_{11} = 2 & \to a = -2 \\ a_{12} + a_{21} = 1 & \to a_{21} = 3 \end{cases}$$

Analogamente, obtém-se: $a_{12} = -3$, $a_{22} = 3$, $a_{13} = 0$, $a_{23} = 2$. Assim,

$$[T]_B^A = \begin{bmatrix} -2 & -3 & 0 \\ 3 & 3 & 2 \end{bmatrix}.$$

**Exemplo 5.13:** Seja T a transformação linear do Exemplo 11 e sejam $\beta = \{(1,0,0),(0,1,0),(0,0,1)\}$ e $\beta' = \{(1,0),(0,1)\}$. Calculemos $[T]^{\beta}_{\beta'}$.

$$T(1,0,0) = (2,3) = 2(1,0) + 3(0,1)$$
$$T(0,1,0) = (1,-2) = 1(1,0) - 2(0,1)$$
$$T(0,0,1) = (-1,4) = -1(1,0) + 4(0,1)$$

Então, $[T]^{\beta}_{\beta'} = \begin{bmatrix} 2 & 1 & -1 \\ 3 & -2 & 4 \end{bmatrix}$.

É importante ressaltar, aqui, que usamos simplesmente [T] para denotar a matriz de uma transformação linear T: $\mathbb{R}^m \to \mathbb{R}^n$ em relação às bases canônicas. Assim, no Exemplo 5.12, $[T]^{\beta}_{\beta'}$ = [T]. Também é comum usar a notação simplificada T**v** = T(**v**).

**Exemplo 5.14:** Sejam T: V → V, onde T(**v**) = **v** (T é a identidade) e $\beta = \{\mathbf{v}_1,...,\mathbf{v}_n\}$ e $\beta' = \{\mathbf{v}'_1,...,\mathbf{v}'_n\}$ bases de V. Calculemos $[T]^{\beta}_{\beta'}$. Como

$$T\mathbf{v}_1 = \mathbf{v}_1 = a_{11}\mathbf{v}'_1 + ... + a_{n1}\mathbf{v}'_n$$

..........................................

$$T\mathbf{v}_n = \mathbf{v}_n = a_{1n}\mathbf{v}'_1 + ... + a_{nn}\mathbf{v}'_n$$

$[T]^{\beta}_{\beta'} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = [I]^{\beta}_{\beta'}$ é a matriz de mudança de base.

**Exemplo 5.14:** Dadas as bases $\beta = \{(1,1),(0,1)\}$ de $\mathbb{R}^2$ e $\beta' = \{(0,3,0),(-1,0,0),(0,1,1)\}$

de $\mathbb{R}^3$, encontremos a transformação linear T: $\mathbb{R}^2 \to \mathbb{R}^3$ cuja matriz é $[T]^{\beta}_{\beta'} = \begin{bmatrix} 0 & 2 \\ -1 & 0 \\ -1 & 3 \end{bmatrix}$, donde:

$$T(1,1) = 0(0,3,0) - 1(-1,0,0) - 1(0,1,1) = (1,-1,-1)$$
$$T(0,1) = 2(0,3,0) + 0(-1,0,0) + 3(0,1,1) = (0,9,3)$$

Para encontrar T(x,y), escrevemos (x,y) em relação à base β:

$$(x,y) = x(1,1) + (y - x)(0,1).$$

Aplicando T e usando a linearidade, temos:

$$T(x,y) = xT(1,1) + (y - x)T(0,1) = x(1,-1,-1) + (y - x)(0,9,3) = (x, 9y - 10x, 3y - 4x).$$

---

**Teorema 5.6:** *Sejam V e W espaços vetoriais, α base de V, β base de W e T: V → W uma aplicação linear. Então, para todo* **v** ∈ *V vale:*

$$\left[T(v)\right]_\beta = \left[T\right]_\beta^\alpha . \left[v\right]_\alpha$$

---

**ER 5.5:** Seja a transformação linear T: $\mathbb{R}^2 \to \mathbb{R}^3$ dada por $\left[T\right]_\beta^\alpha = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -2 & 3 \end{bmatrix}$

onde $\alpha = \{(1,0),(0,1)\}$ é base de $\mathbb{R}^2$, $\beta = \{(1,0,1),(-2,0,1),(0,1,0)\}$ é base de $\mathbb{R}^3$.

Qual é a imagem do vetor **v** = (2,–3) pela aplicação T?

**Solução:** Para isto achamos as coordenadas do vetor **v** em relação à base α, obtendo $[\mathbf{v}]_\alpha = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$; a seguir, usando o teorema temos:

$$\left[Tv\right]_\beta = \left[T\right]_\beta^\alpha \left[v\right]_\alpha = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ 13 \end{bmatrix},$$

ou seja, T**v** = 5(1,0,1) – 3(–2,0,1) –13(0,1,0) = (11,–13,2).

**Definição 5.7:** Chama-se **nulidade** de uma matriz A, a diferença entre o número de colunas de A e o posto de A.

**Teorema 5.7:** *Seja T: V → W uma aplicação linear e α e β bases de V e W respectivamente. Então:*

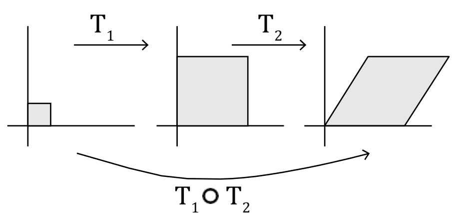
$$\dim Im(T) = \text{posto de } [T]_{\beta}^{\alpha}$$

$$\dim \ker(T) = \text{nulidade de } [T]_{\beta}^{\alpha} = \text{número de colunas} - \text{posto de } [T]_{\beta}^{\alpha}.$$

**Teorema 5.8:** *Sejam $T_1$: V → W e $T_2$: W → U transformações lineares e α, β e γ bases de V, W e U, respectivamente. Então, a composta de $T_1$ com $T_2$, $T_2 \circ T_1$: V → U, é linear e:*

$$[T_2 \circ T_1]_{\gamma}^{\alpha} = [T_2]_{\gamma}^{\beta} [T_1]_{\beta}^{\alpha}.$$

**Exemplo 5.15:** Considere uma expansão do plano $\mathbb{R}^2$ dada por $T_1(x,y) = 2(x,y)$, e um cisalhamento dado por $T_2(x,y) = (x + 2y, y)$. Ao efetuarmos primeiro a expansão e depois o cisalhamento, teremos a sequência:




As matrizes (em relação à base canônica ξ) das transformações são

$$[T_1]_{\xi}^{\xi} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \text{ e } [T_2]_{\xi}^{\xi} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Então, a matriz (em relação à base canônica de $\mathbb{R}^2$) da aplicação que expande e cisalha (que é justamente a composta $T_2 \circ T_1$) será

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

**ER 5.6:** Sejam as transformações lineares $T_1$: $\mathrm{R}^2 \to \mathrm{R}^3$ e $T_2$: $\mathrm{R}^3 \to \mathrm{R}^2$ cujas matrizes em relação às bases $\alpha$ = {(1,0),(0,2)}, $\beta$ = {(1/3,0,–3),(1,1,15),(2,0,5)} e $\gamma$ = {(2,0),(1,1)} são:

$$\left[T_1\right]_\beta^\alpha = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \text{ e } \left[T_2\right]_\gamma^\beta = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Qual é a transformação linear composta $T_2 \mathrm{d} T_1$: $\mathrm{R}^2 \to \mathrm{R}^2$, ou seja, $(T_2 \mathrm{d} T_1)$(x,y)?

**Solução:** Para isto, usamos o teorema anterior para achar a matriz da composta:

$$\left[T_2 \mathrm{o} T_1\right]_\gamma^\alpha = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}$$

Agora, escrevemos as coordenadas do vetor (x,y) em relação à base $\alpha$:

$$\left[(x,y)\right]_\alpha = \begin{bmatrix} x \\ \frac{y}{2} \end{bmatrix}.$$

Então, usando o Teorema 5.8, temos $\left[(T_2 \circ T_1)(x,y)\right]_\gamma = \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \frac{y}{2} \end{bmatrix} = \begin{bmatrix} x-y \\ 0 \end{bmatrix}.$
Portanto,

$$(T_2 \mathrm{d} T_1)(x,y) = (x-y)(2,0) + 0(1,1) = (2x - 2y,0).$$

**Teorema 5.9:** *Se T: V $\to$ W é uma transformação linear inversível (T é um isomorfismo) e $\alpha$ e $\beta$ são as bases de V e W, então $T^{-1}$: W $\to$ V é um operador linear e* $\left[T^{-1}\right]_\alpha^\beta = \left(\left[T\right]_\beta^\alpha\right)^{-1}$.

**Teorema 5.10:** *Seja T: V $\to$ W uma transformação linear e $\alpha$ e $\beta$ bases de V e W. Então, T é inversível se, e somente se,* $\det\left[T\right]_\beta^\alpha \neq 0$.

**Exemplo 5.16:** Seja T: $\mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear dada por $[T]_\xi^\xi = \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}$ onde $\xi$ é a base canônica de $\mathbb{R}^2$. Como $\det[T]_\xi^\xi = 1$, o Teorema 5.10 afirma que T é inversível. Pelo Teorema 5.9, temos que

$$[T^{-1}]_\xi^\xi = \left([T]_\xi^\xi\right)^{-1} = \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -4 \\ -2 & 3 \end{bmatrix}.$$

Então, $\left[T^{-1}(x,y)\right]_\xi = [T^{-1}]_\xi^\xi \begin{bmatrix} x \\ y \end{bmatrix}_\xi = \begin{bmatrix} 3 & -4 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x-4y \\ -2x+3y \end{bmatrix}$, ou seja, $T^{-1}(x,y) = (3x - 4y, -2x + 3y)$.

Se T: V $\to$ W é uma aplicação linear, $\alpha$, $\alpha'$ bases de V e $\beta$, $\beta'$ bases de W, podemos relacionar as matrizes $[T]_\beta^\alpha$ e $[T]_{\beta'}^{\alpha'}$ do seguinte modo:

**Teorema 5.11:** *Dada a matriz de transformação linear em relação a certas bases α e β e as matrizes de mudança de base para novas bases α' e β', pode-se encontrar a matriz de mesma transformação linear, desta vez em relação às novas bases α' e β', ou seja,* $[T]_{\beta'}^{\alpha'} = [I \circ T \circ I]_{\beta'}^{\alpha'} = [I]_{\beta'}^{\beta}[T]_\beta^\alpha[I]_\alpha^{\alpha'}$.

**Definição 5.8:** Se $[T]_\beta^\beta = A[T]_\alpha^\alpha A^{-1}$, as matrizes $[T]_\alpha^\alpha$ e $[T]_\beta^\beta$ são ditas **semelhantes**.

# 6. AUTOVALORES E AUTOVETORES

## 6.1 FUNÇÕES CARACTERÍSTICAS

Associada a cada matriz quadrada A = ($a_{ij}$) de ordem n existe uma função chamada **função característica** de A, dada por

$$f(\lambda) = |A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} \tag{6.1}$$

A equação

$$f(\lambda) = |A - \lambda I| = 0 \tag{6.2}$$

pode ser expressa na forma polinomial

$$c_0 \lambda^n + c_1 \lambda^{n-1} + \ldots + c_{n-1} \lambda + c_n = 0 \tag{6.3}$$

**Definição 6.1:** A equação (6.3), determinada conforme as Equações (6.1) e (6.2), é chamada **equação característica** da matriz A.

**ER 6.1:** Encontre a equação característica da matriz A = $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 3 \end{pmatrix}$.

**Solução:** A equação característica de A é $\begin{vmatrix} 1-\lambda & 2 & 0 \\ 2 & 2-\lambda & 2 \\ 0 & 2 & 3-\lambda \end{vmatrix} = 0$; isto é, $\lambda^3 - 6\lambda^2 + 3\lambda + 10 = 0$.

Algumas vezes a tarefa de expressar a equação característica de uma matriz na forma polinomial pode ser consideravelmente simplificada pela introdução do conceito de traço de uma matriz.

**Definição 6.2:** A soma dos elementos da diagonal de uma matriz A é chamada **traço** de A e é denotada por tr(A).

**Exemplo 6.1:** O traço de uma matriz A em ER 6.1 é tr(A) = 1 + 2 + 3 = 6.

Sejam $t_1 = tr(A)$, $t_2 = tr(A^2)$,..., $t_n = tr(A^n)$. Pode ser mostrado que os coeficientes da equação característica são dados pelas equações:

$$\begin{aligned} c_0 &= 1, \\ c_1 &= -t_1, \\ c_2 &= -\frac{1}{2}\left(c_1 t_1 + t_2\right), \\ c_3 &= -\frac{1}{3}\left(c_2 t_1 + c_1 t_2 + t_3\right), \\ &\ldots\ldots\ldots\ldots \\ c_n &= -\frac{1}{n}\left(c_{n-1} t_1 + c_{n-2} t_2 + \ldots + c_1 t_{n-1} + t_n\right). \end{aligned} \tag{6.4}$$

As equações (6.4) tornam possível o cálculo dos coeficientes da equação característica de uma matriz A pela soma dos elementos da diagonal das matrizes da forma $A^n$. Este processo numérico pode ser facilmente programado para matrizes de alta ordem, ou para pequenos valores de n. Os cálculos podem ser feitos manualmente sem dificuldade.

**ER 6.2:** Encontre a equação característica da matriz A no Exemplo 6.1 com o uso das equações (6.4).

**Solução:** A equação característica de A é da forma $c_0\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3 = 0$.

Agora

$$t_1 = tr(A) = 1 + 2 + 3 = 6$$

e como

$$A^2 = \begin{pmatrix} 5 & 6 & 4 \\ 6 & 12 & 10 \\ 4 & 10 & 13 \end{pmatrix} \text{ e } A^3 = \begin{pmatrix} 17 & 30 & 24 \\ 30 & 56 & 54 \\ 24 & 54 & 59 \end{pmatrix},$$

$$t_2 = tr(A^2) = 5 + 12 + 13 = 30,$$

$$t_3 = tr(A^3) = 17 + 56 + 59 = 132.$$

Então, usando as equações (6.4),

$$c_0 = 1,$$

$$c_1 = -6,$$

$$c_2 = -\frac{1}{2}\left[(-6)(6) + 30\right] = 3,$$

$$c_3 = -\frac{1}{3}\left[(3)(6) + (-6)(30) + 132\right] = 10.$$

Portanto, $\lambda^3 - 6\lambda^2 + 3\lambda + 10 = 0$ é a equação característica de A.

**OCTAVE 6.2**

O traço de uma matriz é calculado pelo comando `trace`. Então definimos a matriz A, calculamos o traço de A e, na sequência, calculamos $A^2$ e $A^3$ e seus respectivos traços:

```
> A = [1,2,0;2,2,2;0,2,3];
> t1 = trace(A);
> A2 = A^2;
> A3 = A^3;
> t2 = trace(A2);
> t3 = trace(A3);
```

Podemos, então, determinar os coeficientes da equação característica, diretamente das Equações 6.4, de posse dos valores já calculados t1, t2 e t3, definindo ainda c0 =1 e calculando os demais:

```
> c0 = 1;
c0 = 1
> c1 = -t1
c1 = -6
> c2 = -1/2*(c1*t1+t2)
c2 = 3
> c3 = -1/3*(c2*t1+c1*t2+t3)
C3 = 10
```

**Definição 6.3:** As n raízes $\lambda_1, \lambda_2, ..., \lambda_n$ da equação característica (6.3), de uma matriz A, são chamados os **autovalores** de A.

**ER 6.3:** Determine os autovalores da matriz A, onde A = $\begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}$.

**Solução:** A equação característica de A é $\begin{vmatrix} 3-\lambda & 1 \\ 2 & 2-\lambda \end{vmatrix} = 0$. Na forma expandida, isto torna-se $\lambda^2 - 5\lambda + 4 = 0$. Os autovalores $\lambda_1 = 1$ e $\lambda_2 = 4$ são as raízes da equação característica.

💻 **OCTAVE 6.3**

O OCTAVE calcula diretamente os autovalores de uma matriz pelo comando `eig`. Então, definimos a matriz A e calculamos os autovalores:

```
> A = [3 1;2 2];
> eig(A)
ans =
    4
    1
```

**ER 6.4:** Prove que o traço de uma matriz A de ordem n é igual à soma dos n autovalores de A, isto é, $\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i$ .

**Solução:** Seja A = $(a_{ij})$. Por definição, $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$ . Considere a função característica de A expressa na forma fatorada:

$$|A-\lambda I| = (-1)^n (\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_n).$$

O coeficiente de $\lambda^{n-1}$ em $|A - \lambda I|$ é $(-1)^{n-1}\sum_{i=1}^{n} a_{ii}$ ; o coeficiente de $\lambda^{n-1}$ na forma fatorada da função característica é $(-1)^{n+1}\sum_{i=1}^{n}\lambda_i$ . Portanto,

$$(-1)^{n-1}\sum_{i=1}^{n} a_{ii} = (-1)^{n+1}\sum_{i=1}^{n}\lambda_i$$
$$\sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n}\lambda_i$$

isto é, $\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i$ .

Muitas aplicações da álgebra matricial em matemática, física e engenharia envolvem o conceito de um conjunto de vetores não nulos que são levados no vetor nulo por meio de uma matriz $A - \lambda_i I$, onde $\lambda_i$ é um autovalor da matriz A.

**Definição 6.4:** Todo vetor coluna não nulo, denotado por $X_i$, tal que

$$\left(A-\lambda_i I\right)X_i=0 \tag{6.5}$$

é chamado um **autovetor** da matriz A.

Isto garante que no mínimo um autovetor exista para cada $\lambda_i$, uma vez que a equação (6.5) representa um sistema de n equações lineares homogêneas que tem uma solução não trivial $X_i \neq 0$ se, e somente se, $|A - \lambda I| = 0$; isto é, se, e somente se, $\lambda_i$ é um autovalor de A. Além disso, note que todo escalar não nulo múltiplo de um autovetor associado com um autovalor é também um autovetor associado com aquele autovalor.

**ER 6.5:** Determine um conjunto de autovetores da matriz A do ER 3.

**Solução:** Associados com $\lambda_1 = 1$ estão os autovetores $(x_1\ x_2)^T$ para os quais

$$\left(A-I\right)\left(x_1 \quad x_2\right)^T=0$$

isto é,

$$\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Segue que $-2x_1 = x_2$. Se $x_1$ é escolhido como algum escalar arbitrário conveniente, digamos 1, $x_2$ torna-se $-2$. Portanto, $(1\ -2)^T$ é um autovetor associado com o autovalor 1.

Analogamente, associado com $\lambda_2 = 4$ estão os autovetores $(x_1\quad x_2)^T$ para os quais

$$\left(A-4I\right)\left(x_1 \quad x_2\right)^T=0$$

isto é,

$$\begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Portanto, $x_1 = x_2$, e $(1\ 1)^T$ é um autovetor associado com o autovalor 4. Além disso, um conjunto de autovetores de uma matriz A é $\{(1\ -2)^T, (1\ \ 1)^T\}$. Deve-se notar que $(k\ -2k)^T$ e $(k\ k)^T$, onde k é qualquer escalar não nulo, representam formas gerais dos autovetores de A.

OCTAVE 6.5

O comando `eig` pode ser utilizado de modo a fornecer, automaticamente, os autovalores e os autovetores associados. Uma vez definida a matriz A, guardamos os autovalores numa matriz AVL e os autovetores numa matriz AVT, da seguinte forma:

```
> A = [3 1;2 2];
> [AVT,AVL]=eig(A)
AVT =
     0.70711  -0.44721
     0.70711   0.89443
AVL =
     4   0
     0   1
```

Note que os autovalores estão na diagonal da matriz AVL. Os autovalores associados a 4 e 1 são, respectivamente, a 19 e 29 colunas de AVT. Uma vez que qualquer múltiplo destes autovetores é um autovetor, podemos simplificar estes valores, dividindo-os por uma de suas coordenadas, como segue:

```
> v1 = AVT(:,1)/AVT(1,1)
v1 =
     1
     1
> v2 = AVT(:,2)/AVT(1,2)
v2 =
     1
    -2
```

Os autovetores, agora, são exatamente os que aparecem em ER 6.5.

Os autovalores de uma matriz são também chamados **valores próprios**, **valores latentes** e **valores característicos** da matriz. Os autovetores de uma matriz também são chamados **vetores próprios**, **vetores latentes** e **vetores característicos** da matriz.

## 6.2 INTERPRETAÇÃO GEOMÉTRICA DOS AUTOVETORES

Considere uma expansão do plano representada pela matriz $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$. Os autovalores de A são $\lambda_1 = 3$ e $\lambda_2 = 2$. Cada autovetor associado com $\lambda_1$ é da forma $(k \ \ 0)^T$, onde k é qualquer escalar não nulo, pois

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

Além disso, o conjunto de vetores da forma $(k \ \ 0)^T$ é t. q. $A(k \ \ 0)^T = \lambda_1(k \ \ 0)^T$; isto é,

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

Portanto, o conjunto dos autovetores associados com $\lambda_1 = 3$ é levado sobre ele mesmo, mediante a transformação representada por A, e a imagem de cada autovetor é um múltiplo escalar fixado deste autovetor. O múltiplo escalar fixado é igual ao autovalor com o qual o conjunto de autovetores está associado.

Analogamente, todo autovetor associado com $\lambda_2$ é da forma $(0 \ \ k)^T$, onde k é qualquer escalar não nulo. O conjunto de vetores da forma $(0 \ \ k)^T$ satisfaz

$$A\left(0 \quad k\right)^T = \lambda_2\left(0 \quad k\right)^T;$$

isto é,

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

Portanto, o conjunto dos autovetores associados com $\lambda_2 = 2$ é levado nele mesmo sob a transformação representada por A, e a imagem de cada autovetor é um múltiplo escalar fixado do autovalor. O múltiplo escalar fixado é $\lambda_{2,}$ isto é, 2.

Note que os conjuntos de vetores das formas $(k \ \ 0)^T$ e $(0 \ \ k)^T$ estão sobre o eixo–x e o eixo–y, respectivamente (Figura 6.1). Sob a expansão do plano representada pela matriz $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$, os *espaços vetoriais unidimensionais* contendo os conjuntos de vetores das formas $(k\ 0)^T$ e $(0\ k)^T$ são levados neles mesmos, respectivamente, e são chamados **espaços vetoriais invariantes**. Os espaços vetoriais invariantes ajudam a caracterizar ou a descrever uma transformação particular do plano.

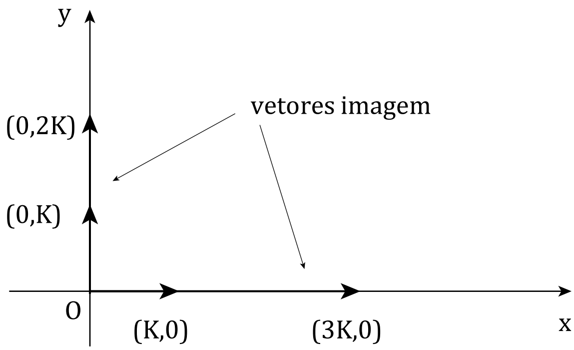



FIGURA 6.1 – Vetores imagem

**ER 6.6:** Determine os espaços vetoriais invariantes sob um cisalhamento paralelo ao eixo–x representado pela matriz $A = \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$.

**Solução:** Os autovalores de A são $\lambda_1 = 1$ e $\lambda_2 = 1$. Associados com cada autovalor, está o conjunto de autovetores da forma $(k\ \ 0)^T$, onde k é qualquer escalar não nulo. Então,

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

e o espaço vetorial unidimensional contendo o conjunto de vetores da forma $(k\ 0)^T$ é um espaço vetorial invariante. Além disso, uma vez que $\lambda_1 = \lambda_2 = 1$, cada vetor no espaço vetorial é sua própria imagem.

**ER 6.7:** Determine os espaços vetoriais invariantes sob uma projeção do plano representado pela matriz $A = \begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}$.

**Solução:** Os autovalores de A são $\lambda_1 = 1$ e $\lambda_2 = 0$. Associado com o autovalor $\lambda_1 = 1$, está o conjunto de autovetores da forma $(k\ \ k)^T$, onde k é qualquer escalar não nulo; associado com o autovalor $\lambda_2 = 0$, está o conjunto de autovetores da forma $(0\ \ k)^T$, onde k é qualquer escalar não nulo. Então,

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

Uma vez que os vetores da forma $(0\ k)^T$ são levados sobre o vetor nulo, o espaço vetorial unidimensional contendo esses vetores é levado nele mesmo, mas não é uma aplicação sobrejetora, o espaço não é considerado um espaço vetorial invariante. Entretanto, o espaço vetorial unidimensional contendo o conjunto de vetores da forma $(k\ \ k)^T$ é um espaço vetorial invariante. Note que estes vetores estão sobre a reta y = x e que o plano é levado sobre esta reta, por meio da projeção do plano representada pela matriz A.

## 6.3 TEOREMAS

Nesta seção, alguns teoremas, sobre autovalores e autovetores de matrizes em geral e de matrizes simétricas em particular, serão provados. Estes teoremas são importantes para um entendimento das seções seguintes deste livro.

Note que, no ER 6.5 da Seção 6.1, os autovetores associados com os autovalores distintos da matriz A são linearmente independentes; isto é,

$$k_1\begin{pmatrix} 1 \\ -2 \end{pmatrix} + k_2\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

implica $k_1 = k_2 = 0$. Esta não é uma coincidência. O seguinte teorema afirma uma condição suficiente para que autovetores associados com os autovalores de uma matriz sejam linearmente independentes.

**Teorema 6.1:** *Se os autovalores de uma matriz são distintos, então os autovetores associados são linearmente independentes.*

**Prova:** Seja A uma matriz quadrada de ordem n com autovalores distintos $\lambda_1$, $\lambda_2$, ..., $\lambda_n$ e autovetores associados $X_1$, $X_2$, ..., $X_n$, respectivamente. Suponha que o conjunto de vetores seja linearmente dependente. Então, existem escalares $k_1$, $k_2$, ..., $k_n$ não todos nulos, tais que

$$k_1X_1 + k_2X_2 + \ldots + k_nX_n = 0 \tag{6.6}$$

Considere a pré-multiplicação de ambos os lados de (6.6) por

$$(A - \lambda_2 I)(A - \lambda_3 I)\ldots(A - \lambda_n I).$$

Usando a equação (6.5), obtemos

$$k_1(A - \lambda_2 I)(A - \lambda_3 I)\ldots(A - \lambda_n I)X_1 = 0. \tag{6.7}$$

Uma vez que $(A - \lambda_1 I)X_1 = 0$, então $AX_1 = \lambda_1 X_1$. Portanto, a equação (6.7) pode ser escrita como

$$k_1(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)\ldots(\lambda_1 - \lambda_n)X_1 = 0,$$

o que implica $k_1 = 0$. Analogamente, pode ser provado que $k_2 = k_3 = \ldots = k_n = 0$, o que contraria a hipótese. Portanto, o conjunto dos autovetores são linearmente independentes.

Deve-se notar que, se os autovalores de uma matriz não são distintos, os autovetores associados podem ou não ser linearmente independentes.

**Exemplo 6.2:** Considere as matrizes $A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$ e $B = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$. Ambas as matrizes tem $\lambda_1 = \lambda_2 = 3$; isto é, um autovalor de multiplicidade 2. Todo vetor não nulo da forma $(x_1\ x_2)^T$ é um autovetor de A para $\lambda_1$ e $\lambda_2$. Portanto, é possível escolher quaisquer dois vetores linearmente independentes tais como $(1\quad 0)^T$ e $(0\quad 1)^T$ como autovetores de A associados com $\lambda_1$ e $\lambda_2$, respectivamente. Apenas um vetor da forma $(x_1\ 0)^T$ é, portanto, um autovetor de B para $\lambda_1$ e $\lambda_2$. Quaisquer dois vetores desta forma são linearmente dependentes; isto é, um é uma função linear do outro.

**Teorema 6.2:** *Se A é uma matriz Hermitiana, então os autovalores de A são reais.*

**Prova:** Seja A uma matriz Hermitiana, $\lambda_i$ qualquer autovalor de A, e $X_i$ um autovetor associado com $\lambda_i$. Então,

$$(A - \lambda_i I)X_i = 0$$
$$AX_i - \lambda_i X_i = 0$$
$$X_i^* AX_i - \lambda_i X_i^* X_i = 0$$

Uma vez que todo autovetor é um vetor não nulo, $X_i^*X_i$ é um número real não nulo e

$$\lambda_i = \frac{X_i^* AX_i}{X_i^* X_i}.$$

Além disso,

$$X_i^* AX_i = X_i^* A^* X_i = \left(X_i^* AX_i\right)^* = \overline{X_i^* AX_i}$$

isto é, $X_i^*AX_i$ é igual ao seu próprio conjugado e, portanto, é real. Além disso, $\lambda_i$ é igual ao quociente de dois números reais, e é real.

**Teorema 6.3:** *Se A é uma matriz simétrica real, então os autovalores de A são reais.*

**Prova:** Uma vez que toda matriz simétrica real é uma matriz Hermitiana, a prova segue do Teorema 6.2.

Antes de apresentar o próximo teorema, é necessário considerar a seguinte definição: dois autovetores complexos $X_1$ e $X_2$ são definidos como ortogonais se $X_1^*X_2 = 0$.

**Exemplo 6.3:** Se $X_1 = (-i\ 2)^T$ e $X_2 = (2i\ 1)^T$, então, $X_1^*X_2 = (-i\ 2)\ (2i\ 1)^T = 0$. Portanto, $X_1$ e $X_2$ são ortogonais.

**Teorema 6.4:** *Se A é uma matriz Hermitiana, então os autovetores de A associados com autovalores distintos são vetores ortogonais.*

**Prova:** Seja A uma matriz Hermitiana, e sejam $X_1$ e $X_2$ autovetores associados com quaisquer autovalores distintos $\lambda_1$ e $\lambda_2$, respectivamente. Então,

$$(A - \lambda_1 I)X_1 = 0 \quad \text{e} \quad (A - \lambda_2 I)X_2 = 0;$$

isto é,

$$AX_1 = \lambda_1 X_1 \quad \text{e} \quad AX_2 = \lambda_2 X_2. \tag{6.8}$$

Multiplicando ambos os lados da primeira equação de (6.8) por $X_2^*$, obtém-se

$$\lambda_1 X_2^* X_1 = X_2^* A X_1 = X_2^* A^* X_1 = (AX_2)^* X_1 = \lambda_2 X_2^* X_1.$$

Então,

$$X_2^* X_1 = \frac{\lambda_2}{\lambda_1} X_2^* X_1.$$

Uma vez que $\lambda_1$ e $\lambda_2$ são reais e distintos, $X_2^* X_1$ deve ser zero. Portanto, $X_1$ e $X_2$ são autovetores ortogonais.

**Teorema 6.5:** *Se A é uma matriz real simétrica, então os autovetores de A, associados com autovalores distintos, são vetores ortogonais.*

**Prova:** Uma vez que toda matriz real simétrica é uma matriz Hermitiana, a prova segue do Teorema 6.4.

## 6.4 DIAGONALIzAÇÃO DE MATRIzES

Nota-se que um autovetor $X_i$ tal que $(A - \lambda_i I)X_i = 0$, para $i = 1,2,\ldots,n$, pode ser associado com cada autovalor $\lambda_i$. Esta relação pode ser expressa de forma alternativa como

$$AX_i = \lambda_i X_i, \text{ para } i = 1,2,\ldots,n. \tag{6.9}$$

Se uma matriz quadrada de ordem n, cujas colunas são autovetores $X_i$ de A, é construída e denotada por X, então as equações de (6.9) podem ser expressas na forma

$$AX = X\Lambda, \tag{6.10}$$

onde $\Lambda$ é uma matriz diagonal cujos elementos são os autovalores de A; isto é,

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}. \tag{6.11}$$

Foi provado que os autovetores associados a autovalores distintos são linearmente independentes (Teorema 6.1). Portanto, a matriz X será não singular se os $\lambda_i$ são distintos. Se ambos os lados da equação (6.10) são multiplicados por $X^{-1}$, o resultado é

$$X^{-1}AX = \Lambda. \tag{6.12}$$

Assim, pelo uso de uma matriz de autovetores e sua inversa, é possível transformar qualquer matriz A com autovalores distintos numa matriz diagonal cujos elementos são os autovalores de A.

**Definição 6.5:** A transformação expressa por (6.12) é conhecida como **diagonalização** da matriz A.

Se os autovalores não são distintos, a diagonalização da matriz A pode não ser possível.

**Exemplo 6.4:** a matriz A = $\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$ não pode ser diagonalizada como em (6.12).

Uma matriz tal como a matriz A na equação (6.2) algumas vezes é mencionada como sendo **similar** à matriz diagonal.

**Definição 6.6:** Considere A e B duas matrizes quadradas de mesma ordem. Se existe uma matriz não singular C, tal que $C^{-1}AC = B$, então A e B são chamadas **matrizes similares**, e a transformação de A para B é chamada **transformação de similaridade**.

**Definição 6.7:** A matriz B da Definição 6.6 é uma matriz diagonal cujos elementos da diagonal são os autovalores de A, e é chamada **forma canônica clássica** da matriz A. Ela é a única matriz, exceto para a ordem na qual os autovalores aparecem ao longo da diagonal principal.

**Definição 6.8:** A matriz X de (6.12), cujas colunas são autovetores da matriz A, frequentemente é chamada **matriz modal** de A.

Lembre-se de que cada autovetor pode ser multiplicado por qualquer escalar não nulo. Portanto, uma matriz modal de A não é única.

**ER 6.8:** Determine se A = $\begin{pmatrix} 6 & 2 \\ -2 & 1 \end{pmatrix}$ e B = $\begin{pmatrix} 8 & 6 \\ -3 & -1 \end{pmatrix}$ são matrizes similares.

**Solução:** Se A e B são matrizes similares, existe uma matriz quadrada não singular C de ordem 2 tal que $C^{-1}AC = B$; isto é, $AC = CB$. Seja C = $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ onde $ad - bc \neq 0$. Então,

$$\begin{pmatrix} 6 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 8 & 6 \\ -3 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 6a+2c & 6b+2d \\ -2a+c & -2b+d \end{pmatrix} = \begin{pmatrix} 8a-3b & 6a-b \\ 8c-3d & 6c-d \end{pmatrix}$$

Esta equação matricial leva ao sistema de equações homogêneas

$$\begin{cases} 2a-3b-2c=0 \\ 2a+7c-3d=0 \\ 6a-7b-2d=0 \\ 2b+6c-2d=0 \end{cases}$$

O sistema de equações homogêneas tem um número infinito de soluções da forma a = 3t – 7s, b = 2t – 6s, c = 2s, e d = 2t, onde s e t são escalares arbitrários reais. Portanto, existe uma matriz não singular

$$C=\begin{pmatrix} 3t-7s & 2t-6s \\ 2s & 2t \end{pmatrix}$$

onde s e t são escalares arbitrários reais, com a garantia de que det C ≠ 0; isto é,

$$6t^2 - 18t + 12s^2 \neq 0,$$

$$(6t - 12s)(t - s) \neq 0,$$

$$t \neq 2s \ \text{ e } \ t \neq s.$$

Portanto, A e B são matrizes similares.

Como uma ilustração de que A e B são matrizes similares, sejam s = 0 e t = 1. Então,

$$C=\begin{pmatrix} 3 & 2 \\ 0 & 2 \end{pmatrix},\ C^{-1}=\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ 0 & \frac{1}{2} \end{pmatrix},$$

e

$$C^{-1}AC=\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 6 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ 0 & 2 \end{pmatrix}=\begin{pmatrix} 8 & 6 \\ -3 & -1 \end{pmatrix}=B.$$

**ER 6.9:** Prove que matrizes similares têm determinantes iguais e autovalores iguais.

**Solução:** Sejam A e B matrizes similares. Então, existe uma matriz quadrada não singular C de mesma ordem que A e B, tal que $C^{-1}AC = B$. Uma vez que o determinante do produto de duas matrizes é igual ao produto de seus determinantes, segue que

$$\begin{aligned} \det B &= \det C^{-1} \det A \det C \\ &= \det C^{-1} \det C \det A \\ &= \det\left(C^{-1}C\right)\det A \\ &= \det I \det A \\ &= \det A \end{aligned}$$

Então,

$$\begin{aligned} \det(A-\lambda I) &= \det\left[C^{-1}(A-\lambda I)C\right] \\ &= \det\left(C^{-1}AC - \lambda C^{-1}IC\right) \\ &= \det(B-\lambda I) \end{aligned}$$

isto é, A e B têm a mesma função característica. Uma consequência imediata disso é que A e B têm a mesma equação característica e os mesmos autovalores.

Deve-se tomar o cuidado de observar que o inverso da afirmação do Exemplo 6.2 não é necessariamente verdade. Por exemplo, as matrizes $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ têm os mesmos autovalores $\lambda_1 = \lambda_2 = 1$ e $\det A = \det B$, mas $C^{-1}AC = I$ para qualquer matriz quadrada não singular C de ordem 2 e $I \neq B$. Portanto, A e B não podem ser matrizes similares.

**ER 6.10:** Prove que $A^n X_i = \lambda_i^n X_i$ para todos os números naturais n.

**Solução:** Esta útil relação pode ser provada por indução matemática. Tem-se $AX_i = \lambda_i X_i$, por (6.9). Agora, suponha que $A^k X_i = \lambda_i^k X_i$, onde k é qualquer inteiro positivo. Então,

$$
\begin{aligned}
A^{k+1}X_i &= A\lambda_i^k X_i \\
&= \lambda_i^k A X_i \\
&= \lambda_i^{k+1} X_i \quad \text{por (4.9).}
\end{aligned}
$$

Portanto,

$$A^n X_i = \lambda_i^n X_i \tag{4.13}$$

para todos os números naturais n.

Foi mostrado que, se A é uma matriz real simétrica de ordem n com n autovalores distintos reais, os autovetores associados são mutuamente ortogonais (Teorema 6.5). Uma matriz de autovetores pode ser feita ortogonal própria se cada autovetor é normalizado por um múltiplo escalar apropriado.

**Definição 6.9:** Uma transformação de similaridade empregando uma matriz ortogonal modal é chamada **transformação ortogonal**; isto é, uma transformação ortogonal de uma matriz A é da forma $C^TAC$, onde C é uma matriz ortogonal.

Se uma matriz simétrica de ordem n tem autovalores múltiplos, é sempre possível determinar n autovetores unitários mutuamente ortogonais. É possível mostrar que r autovetores linearmente independentes que são ortogonais aos outros autovetores podem ser associados com um autovalor de multiplicidade r. Além disso, é sempre possível escolher estes vetores ortogonais a cada um dos outros.

**Teorema 6.6:** *Toda matriz simétrica real pode ser ortogonalmente transformada na forma canônica clássica.*

O Teorema 6.6 é algumas vezes chamado de **Teorema dos eixos principais**. Uma aplicação deste teorema à geometria analítica será considerada posteriormente.

**ER 6.11:** Determine uma matriz modal ortogonal própria que transforma a matriz A na forma canônica clássica, onde A = $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$.

**Solução:** A equação característica de A é $\lambda^2 - 6\lambda + 8 = 0$; então, os autovalores de A são $\lambda_1 = 2$ e $\lambda_2 = 4$. Associado com o autovalor $\lambda_1 = 2$, está o autovetor unitário $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}^T$.

Associado com o autovalor $\lambda_2 = 4$, está o autovetor unitário $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^T$.

Portanto, uma matriz modal ortogonal que transforma a matriz A na forma canônica é

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

isto é,

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

## 6.5 O TEOREMA DE CAyLEy-HAMILTON

Um importante teorema da teoria de matrizes é o **Teorema deCayley-Hamilton**:

**Teorema 6.7:** *Toda matriz quadrada A satisfaz sua própria equação característica* $|A-\lambda I| = 0$.

Mais precisamente, se $\lambda$ é trocado pela matriz A de ordem n e cada número real $c_n$ é trocado pelo múltiplo escalar $c_nI$, onde I é a matriz identidade de ordem n, então a equação característica da matriz A torna-se uma equação matricial válida; isto é,

$$c_0A^n + c_1A^{n-1} + \ldots + c_{n-1}A + c_nI = 0. \tag{6.14}$$

Um argumento heurístico pode ser usado para provar o Teorema de Cayley--Hamilton para uma matriz A com autovalores distintos.

**Prova:** Trocando a variável $\lambda$ pela matriz quadrada A e $c_n$ por $c_nI$ na expressão para a função característica de A obtém-se

$$f(A) = c_0A^n + c_1A^{n-1} + \ldots + c_{n-1}A + c_nI. \tag{6.15}$$

Pós-multiplicando ambos os lados da equação (6.15) por um autovetor $X_i$ de A, associado com $\lambda_i$, obtém-se

$$f(A)X_i = \left(c_0\lambda_i^n + c_1\lambda_i^{n-1} + \ldots + c_{n-1}\lambda_i + c_n\right)X_i$$

pois $A^kX_i = \lambda_i^kX_i$ (verifique isto como exercício). Uma vez que

$$c_0\lambda_i^n + c_1\lambda_i^{n-1} + \ldots + c_{n-1}\lambda_i + c_n = 0 \quad \text{para} \quad i = 1,2,\ldots,n,$$

então,

$$f(A)X_i = 0 \quad \text{para} \quad i = 1,2,\ldots,n.$$

Portanto,

$$f(A)X = 0, \tag{6.16}$$

onde X é uma matriz de autovetores. Como os autovetores são linearmente independentes pelo Teorema 6.1, a matriz de autovetores tem uma única inversa $X^{-1}$. Se ambos os lados da equação (6.16) são pós-multiplicados por $X^{-1}$, o resultado é f(A) = 0, o que conclui a prova do teorema.

**ER 6.12:** Mostre que a matriz A = $\begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}$ satisfaz sua própria equação característica.

**Solução:** A função característica de A é $f(\lambda) = |A - \lambda I| = \begin{vmatrix} 3-\lambda & -2 \\ 1 & 2-\lambda \end{vmatrix} = f(\lambda) = \lambda^2 -$ $- 5\lambda + 8$. Trocando $\lambda$ por A e 8 por 8I, onde I é a matriz identidade de ordem 2, obtém-se

$$f(A) = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}^2 - 5\begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix} + 8\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 7 & -10 \\ 5 & 2 \end{pmatrix} + \begin{pmatrix} -15 & 10 \\ -5 & -10 \end{pmatrix} + \begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Portanto, f(A) = 0 e o Teorema de Cayley-Hamilton foi verificado para a matriz A.

**OCTAVE 6.12**

Definindo a matriz A, de posse da equação característica, o cálculo é direto:

```
> A = [3 -2;1 2];
> A^2-5*A+8*eye(2)
ans =
    0    0
    0    0
```

A matriz identidade de ordem 2 é criada pelo comando `eye(2)`.

O Teorema de Cayley-Hamilton pode ser aplicado ao problema de determinação da inversa de uma matriz não singular A. Seja

$$c_0\lambda^n + c_1\lambda^{n-1} + \ldots + c_{n-1}\lambda + c_n = 0$$

a equação característica de A.

Note que, uma vez que A é uma matriz não singular, $\lambda_i \neq 0$; isto é, todo autovalor é não nulo, e $c_n \neq 0$. Pelo Teorema de Cayley-Hamilton,

$$c_0A^n + c_1A^{n-1} + \ldots + c_{n-1}A + c_nI = 0,$$

e

$$I = -\frac{1}{c_n}\left(c_0 A^n + c_1 A^{n-1} + \ldots + c_{n-1} A\right). \qquad (6.17)$$

Se ambos os lados de (6.17) são multiplicados por $A^{-1}$, o resultado é

$$A^{-1} = -\frac{1}{c_n}\left(c_0 A^{n-1} + c_1 A^{n-2} + \ldots + c_{n-1} I\right). \qquad (6.18)$$

Note que o cálculo de uma inversa pelo uso da equação (6.18) é facilmente adaptável a computadores digitais de alta velocidade e não é difícil calcular manualmente para pequenos valores de n. No cálculo das potências da matriz A, necessárias na equação (6.18), também são obtidas as informações a respeito do $tr(A^k)$ para o cálculo dos $c_i$.

**ER 6.13:** Use o Teorema de Cayley-Hamilton para encontrar a inversa de $A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & -3 \\ 2 & 2 & 4 \end{pmatrix}$.

**Solução:** A equação característica de A é $\begin{vmatrix} 1-\lambda & 0 & 1 \\ -1 & 1-\lambda & -3 \\ 2 & 2 & 4-\lambda \end{vmatrix} = 0$; ou $\lambda^3 - 6\lambda^2 + 13\lambda - 6 = 0$. Pelo Teorema de Cayley-Hamilton,

$$A^3 - 6A^2 + 13A - 6I = 0,$$

$$I = (A^3 - 6A^2 + 13A)/6,$$

e

$$A^{-1} = \frac{1}{6}\left(A^2 - 6A + 13I\right).$$

Portanto,

$$A^{-1}=\frac{1}{6}\left[\begin{pmatrix}1&0&1\\-1&1&-3\\2&2&4\end{pmatrix}^2-6\begin{pmatrix}1&0&1\\-1&1&-3\\2&2&4\end{pmatrix}+13\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\right]=\frac{1}{6}\left[\begin{pmatrix}3&2&5\\-8&-5&-16\\8&10&12\end{pmatrix}+\right.$$

$$\left.+\begin{pmatrix}-6&0&-6\\6&-6&18\\-12&-12&-24\end{pmatrix}+\begin{pmatrix}13&0&0\\0&13&0\\0&0&13\end{pmatrix}\right]=\begin{pmatrix}\frac{5}{3}&\frac{1}{3}&-\frac{1}{6}\\-\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\\-\frac{2}{3}&-\frac{1}{3}&\frac{1}{6}\end{pmatrix}$$

Assim, $A^{-k}$ onde k é um inteiro positivo é definido como igual a $(A^{-1})^k$. Pelo uso da equação (6.18), é possível expressar qualquer potência inteira negativa de uma matriz não singular A de ordem n em termos de uma função linear das primeiras (n – 1) potências de A.

**ER 6.14:** Encontre $A^{-3}$ para a matriz não singular $A=\begin{pmatrix}2&4\\1&1\end{pmatrix}$. Verifique que $A^{-3}$ é a inversa de $A^3$.

**Solução:** A equação característica de A é $\lambda^2-3\lambda-2=0$, e $c_0=1$, $c_1=-3$, e $c_2=-2$. Usando a equação (6.18),

$$A^{-1}=\frac{1}{2}(A-3I).\text{ Então,}$$

$$A^{-2}=A^{-1}A^{-1}=\frac{1}{2}\left(I-3A^{-1}\right)=\frac{11}{4}I-\frac{3}{4}A\text{ e }A^{-3}=A^{-1}A^{-2}=\frac{11}{4}A^{-1}-\frac{3}{4}I=\frac{11}{8}A-\frac{39}{8}I.$$

Portanto,

$$A^{-3}=\frac{11}{8}\begin{pmatrix}2&4\\1&1\end{pmatrix}-\frac{39}{8}\begin{pmatrix}1&0\\0&1\end{pmatrix}=\begin{pmatrix}-\frac{17}{8}&\frac{11}{2}\\\frac{11}{8}&-\frac{7}{2}\end{pmatrix}$$

Assim, como $A^3 = \begin{pmatrix} 28 & 44 \\ 11 & 17 \end{pmatrix}$, então $\begin{pmatrix} 28 & 44 \\ 11 & 17 \end{pmatrix} \begin{pmatrix} -\frac{17}{8} & \frac{11}{2} \\ \frac{11}{8} & -\frac{7}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

## 6.6 FORMAS QUADRÁTICAS

Antes de discutir a simplificação da equação quadrática geral em duas variáveis por meio de transformações de movimento rígido, é necessário introduzir o conceito de formas quadráticas reais.

**Definição 6.10:** Uma **forma quadrática** é um polinômio homogêneo de segundo grau em n variáveis $w_1, w_2, ..., w_n$; isto é, um polinômio da forma

$$\sum_{i,j=1}^{n} a_{ij} w_i w_j \tag{6.19}$$

Todas as formas quadráticas podem ser expressas como uma matriz produto $W^T A W$, onde $W = (w_1\ w_2 \ldots w_n)^T$ e $A = (a_{ij})$ é uma matriz simétrica real única. Portanto, toda forma quadrática real pode ser ortogonalmente transformada na forma $\lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2$, onde os $\lambda_i$ são os autovalores de A, por uma escolha apropriada de uma matriz ortogonal própria C tal que $C^T A C$ é uma matriz diagonal; isto é,

$$W^T A W = Z^T C^T A C Z = Z^T \Lambda Z, \tag{6.20}$$

onde $W = CZ$ e $Z = (z_1\ z_2 \ldots z_n)^T$. É claro, a matriz ortogonal apropriada C que transforma uma matriz simétrica real A numa matriz diagonal é uma matriz de autovetores unitários de A.

Se (6.19) é um polinômio homogêneo de segundo grau em duas variáveis, a matriz modal C da transformação ortogonal em (6.20) pode ser considerada uma matriz de rotação. Por exemplo, o problema de redução da forma quadrática $ax^2 + 2bxy + cy^2$ para a forma canônica $a'x'^2 + c'y'^2$ pode ser obtida por meio de uma rotação do plano sobre a origem. Quando a forma quadrática é escrita na forma matricial, o resultado é

$$\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Uma rotação do plano sobre a origem sob um ângulo θ pode ser expressa pela equação matricial

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Uma vez que a transposta de um produto de duas matrizes é igual ao produto das transpostas das duas matrizes em ordem reversa, então

$$\begin{pmatrix} x' & y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}$$

representa a forma quadrática após a rotação do plano sobre a origem sob um ângulo θ. Deseja-se que θ, o ângulo de rotação, seja tal que

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \quad \begin{pmatrix} a' & 0 \\ 0 & c' \end{pmatrix}$$

Isto constitui claramente um problema de transformação de uma matriz simétrica numa matriz diagonal. Primeiramente, mostra-se que, neste caso, os elementos diagonais da matriz diagonal são os autovalores da matriz simétrica. Portanto, para obter a rotação própria, é necessário apenas resolver a equação característica da forma quadrática:

$$\begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} = 0;$$

isto é, $\lambda^2 - (a+c)\lambda + (ac - b^2) = 0$.

Os autovalores $\lambda_1$ e $\lambda_2$ tornam-se os coeficientes de $x'^2$ e $y'^2$; isto é,

$$\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' & y' \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} = \left(\lambda_1 x'^2 + \lambda_2 y'^2\right).$$

O ângulo de rotação θ pode ser determinado, considerando-se a equação matricial

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix};$$

isto é,

$$\begin{cases} a\cos^2\theta - b\operatorname{sen}\theta\cos\theta - b\operatorname{sen}\theta\cos\theta + c\operatorname{sen}^2\theta = \lambda_1 \\ a\operatorname{sen}\theta\cos\theta + b\cos^2\theta - b\operatorname{sen}^2\theta - c\operatorname{sen}\theta\cos\theta = 0 \\ a\operatorname{sen}\theta\cos\theta - b\operatorname{sen}^2\theta + b\operatorname{sen}^2\theta - c\operatorname{sen}\theta\cos\theta = 0 \\ a\operatorname{sen}^2\theta + b\operatorname{sen}\theta\cos\theta + b\operatorname{sen}\theta\cos\theta + c\cos^2\theta = \lambda_2. \end{cases} \tag{6.21}$$

Se a segunda e a terceira equações de (6.21) são adicionadas, o resultado é

$$(a-c)\operatorname{sen}\theta\cos\theta + b\left(\cos^2\theta - \operatorname{sen}^2\theta\right) = 0.$$

Além disso, $\dfrac{2b}{c-a} = \dfrac{\operatorname{sen}2\theta}{\cos 2\theta} = \tan 2\theta.$

Portanto,

$$\theta = \frac{1}{2}\arctan\frac{2b}{c-a}. \tag{6.22}$$

Note que os vetores unitários ao longo dos eixos coordenados são as imagens dos vetores unitários $(\cos\theta\ -\operatorname{sen}\theta)^T$ e $(\operatorname{sen}\theta\ \cos\theta)^T$ associados com os autovalores da matriz simétrica da forma quadrática. Estes autovetores unitários estão ao longo dos eixos principais de $ax^2 + 2bxy + y^2 = 0$.

## 6.7 CLASSIFICAÇÃO DAS CÔNICAS

Uma equação de segundo grau mais geral, em duas variáveis x e y, pode ser escrita na forma

$$f(x,y) = ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0, \tag{6.23}$$

onde a, b, c, d, e, e f são números reais.

**Definição 6.11:** A equação (6.23) no plano Euclidiano coordenado é chamada **seção cônica**, ou simplesmente **cônica**.

Por meio de uma rotação do plano sobre a origem, uma translação do plano, ou ambas, é possível representar qualquer cônica de uma maneira simplificada padrão, ou seja, numa forma canônica. Estas figuras planas podem ser estudadas mais facilmente em suas formas canônicas. Existem nove classes de cônicas; isto é, a equação (6.23) representa um dos nove tipos de figuras planas chamadas cônicas. Duas destas figuras planas são imaginárias, de modo que não existem pontos reais que satisfaçam a equação (6.23).

Se a função geral do segundo grau f (x,y) da equação (6.23) é escrita na forma

$$f(x,y) = axx + bxy + dx + byx + cyy + ey + dx + ey + f, \qquad (6.24)$$

que é outra forma valiosa de f (x,y), esta função pode ser expressa como o produto de 3 matrizes:

$$\begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \qquad (6.25)$$

**Definição 6.12:** A matriz $\Delta$ dada por

$$\Delta = \begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix} \qquad (6.26)$$

tem importância central e define a cônica dada por (6.24); é uma matriz real simétrica e é chamada **matriz da seção cônica**.

Em adição a $\Delta$, a matriz cujo determinante é o menor de f em $\Delta$, a matriz simétrica real

$$F = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \qquad (6.27)$$

é de fundamental importância na análise de f (x,y). O resultado de uma pré-multiplicação da matriz F por um vetor linha $(x \;\; y)$ e uma pós-multiplicação da matriz F pelo vetor coluna $(x\ y)^T$ representa a porção da forma quadrática da função geral de segundo grau; isto é,

$$(x \quad y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2.$$

Após uma rotação do plano sobre a origem sob um ângulo θ definida pela equação matricial

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta & 0 \\ -\operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$$

onde θ é $\frac{1}{2}\arctan\frac{2b}{c-a}$, a equação geral de uma cônica pode ser expressa na forma

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2\alpha x' + 2\beta y' + f = 0; \tag{6.28}$$

isto é,

$$(x' \quad y' \quad 1)\begin{pmatrix} \lambda_1 & 0 & \alpha \\ 0 & \lambda_2 & \beta \\ \alpha & \beta & f \end{pmatrix}\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = 0, \tag{6.29}$$

onde os $\lambda_i$ são os autovalores da matriz F e

$$\begin{cases} \alpha = d\cos\theta - e\operatorname{sen}\theta \\ \beta = d\operatorname{sen}\theta + e\cos\theta \end{cases} \tag{6.30}$$

Os termos lineares em x' e y' de (6.28) podem ser removidos por uma translação do plano definido pela equação matricial

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\frac{\alpha}{\lambda_1} \\ 0 & 1 & -\frac{\beta}{\lambda_2} \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x'' \\ y'' \\ 1 \end{pmatrix}. \tag{6.31}$$

**Definição 6.13:** Se pelo menos um dos $\lambda_i$ é igual a zero, não existe uma translação do plano que possa remover todos os termos lineares da equação (6.28) e, consequentemente, não existe um centro geométrico para a cônica que é chamada **cônica não central**.

Por outro lado, se existe algum $\lambda_i$ igual a zero, a matriz F é necessariamente singular.

**Definição 6.14:** Quando F é não singular, existe um centro geométrico para a cônica e a cônica é chamada **cônica central**.

Após uma translação do plano descrita por (6.31), a equação central da cônica pode ser expressa como

$$\lambda_1 x''^2 + \lambda_2 y''^2 + f' = 0; \tag{6.31}$$

isto é,

$$\begin{pmatrix} x'' & y'' & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & f' \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ 1 \end{pmatrix} = 0, \tag{6.32}$$

onde

$$f' = f - \frac{\alpha^2}{\lambda_1} - \frac{\beta^2}{\lambda_2}. \tag{6.33}$$

Se um e apenas um $\lambda_i$ é zero, então existe uma translação do plano que deve remover um dos termos lineares da equação (6.28).

**Exemplo 6.5:** Considere $\lambda_2 = 0$. A translação do plano definida pela equação matricial

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\frac{\alpha}{\lambda_1} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ 1 \end{pmatrix} \tag{6.34}$$

transforma a equação (6.28) na forma

$$\lambda_1 x''^2 + 2\beta y'' + f' = 0; \tag{6.35}$$

isto é,

$$\begin{pmatrix} x'' & y'' & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & \beta \\ 0 & \beta & f' \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ 1 \end{pmatrix} = 0, \tag{6.36}$$

onde

$$f' = f - \frac{\alpha^2}{\lambda_1}. \tag{6.37}$$

**ER 6.15:** Transforme a equação da cônica $5x^2 + 6xy + 5y^2 - 4x + 4y - 4 = 0$ numa forma canônica (Figura 6.2).

**Solução:** A equação da cônica pode ser escrita na forma

$$\begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} 5 & 3 & -2 \\ 3 & 5 & 2 \\ -2 & 2 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0.$$

A equação característica de $F = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ é $\lambda^2 - 10\lambda + 16 = 0$. Portanto, os autovalores de F são $\lambda_1 = 8$ e $\lambda_2 = 2$ com os autovetores unitários associados

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^T \text{ e } \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^T,$$

respectivamente. Portanto, a rotação do plano definida pela equação matricial

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$$

deve transformar a equação da cônica para a forma de (6.28):

$$\begin{pmatrix} x' & y' & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 & 3 & -2 \\ 3 & 5 & 2 \\ -2 & 2 & -4 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = 0,$$

$$\begin{pmatrix} x' & y' & 1 \end{pmatrix} \begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 2\sqrt{2} \\ 0 & 2\sqrt{2} & -4 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = 0;$$

isto é,

$$8x'^2 + 2y'^2 + 4\sqrt{2}y' - 4 = 0.$$

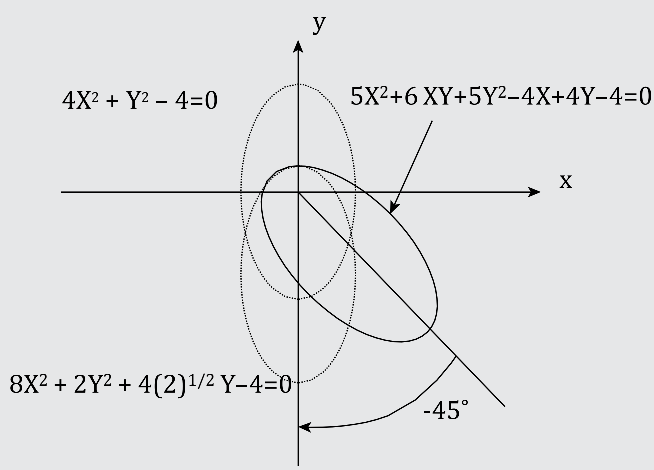



FIGURA 6.2 - Forma canônica para uma cônica

Uma translação do plano definida por (6.31), onde $\alpha = 0$, $\beta = 2\sqrt{2}$, $\lambda_1 = 8$, e $\lambda_2 = 8$ deve transformar a equação da cônica na forma de (6.32). Daí resulta que

$$\begin{pmatrix} x'' & y'' & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 2\sqrt{2} \\ 0 & 2\sqrt{2} & -4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\sqrt{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ 1 \end{pmatrix} = 0,$$

$$\begin{pmatrix} x'' & y'' & 1 \end{pmatrix} \begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -8 \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ 1 \end{pmatrix} = 0,$$

$$8x''^2 + 2y''^2 - 8 = 0,$$

$$4x''^2 + y''^2 - 4 = 0;$$

isto é,

$$4x^2 + y^2 - 4 = 0.$$

Portanto, a cônica é uma elipse real. Note que os vetores unitários são paralelos aos eixos principais da cônica original.

Fica evidente, da discussão, que a classe de cônicas representada pela equação (6.23) pode ser determinada por uma investigação dos valores algébricos dos autovalores da matriz F. As cônicas serão, agora, classificadas de acordo com os autovalores de F.

Se $\lambda_1$ e $\lambda_2$ são não nulos e têm os mesmos sinais algébricos, então a equação (6.32) representa uma elipse imaginária, uma elipse real, ou um ponto real dependente de f' ter o mesmo sinal dos $\lambda_i$, diferir em sinal dos $\lambda_i$, ou ser zero, respectivamente. Além disso, se os autovalores são iguais no caso especial da elipse real, a cônica é um círculo.

Se $\lambda_1$ e $\lambda_2$ são não nulos e têm sinais opostos, então a equação (6.32) representa uma hipérbole ou duas retas concorrentes, dependente de f' ser não nulo ou zero, respectivamente.

Para o caso em que apenas um autovalor é zero, considere $\lambda_2$ igual a zero; não existe perda de generalidade em se fazer isso. Então, a equação (6.23) pode ser transformada na forma (6.36). Se $\beta$ é não nulo, então a equação (6.36) representa uma parábola; se $\beta$ é igual a zero, então a equação (6.36) representa duas retas paralelas, se $\lambda_1$ e f' têm sinais opostos; duas retas paralelas imaginárias se $\lambda_1$ e f' têm sinais opostos e duas retas coincidentes se f' é igual a zero.

**ER 6.16:** Identifique a classe da cônica representada pela equação $x^2 - 4xy + 4y^2 - 4 = 0$.

**Solução:**

A matriz da seção cônica é $\begin{pmatrix} 1 & -2 & 0 \\ -2 & 4 & 0 \\ 0 & 0 & -4 \end{pmatrix}$

A equação característica da matriz F é $\begin{vmatrix} 1-\lambda & -2 \\ -2 & 4-\lambda \end{vmatrix} = 0;$

isto é, $\lambda^2 - 5\lambda = 0$. Portanto, os autovalores de F são $\lambda_1 = 5$ e $\lambda_2 = 0$ com autovetores associados

$$\begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix}^T \quad \text{e} \quad \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}^T.$$

Uma vez que a inversa da matriz de rotação e a matriz ortogonal própria de autovetores são idênticas,

$$\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix},$$

$\cos\theta = 1/\sqrt{5}$ e $\operatorname{sen}\theta = 2/\sqrt{5}$. Além disso, d = e = 0; f = –4. Então, $\alpha = \beta = 0$ por (6.30), e f' = –4 por (6.38). Uma vez que um autovalor $\lambda_2$ é zero e o autovalor restante $\lambda_1$ e f' tem sinais opostos, a equação da cônica representa duas retas paralelas.

Note que a equação da cônica pode ser expressa na forma fatorada como $(x - 2y + 2)(x - 2y - 2) = 0$. Portanto, as equações das duas retas paralelas são $x - 2y + 2 = 0$ e $x - 2y - 2 = 0$.

# 7. OCTAVE

## 7.1 INTRODUÇÃO

Existem vários softwares auxiliares na realização das mais diversas atividades de ensino e pesquisa em matemática, dentre estes, pode-se citar o MATHEMATICA®, o MAPLE®, o MATLAB®, o MAXIMA e o OCTAVE. Estes dois últimos são softwares livres, ou seja, são gratuitos e de código aberto, passíveis de contribuições e alterações, sob determinadas condições, ao seu código original. Esses programas contam em seu desenvolvimento com a participação de pessoas no mundo inteiro e tornam-se cada vez mais de interesse da comunidade científica e do público em geral. Muitos programas proprietários possuem similares gratuitos que podem auxiliar a reduzir os custos dessas ferramentas computacionais, ao mesmo tempo que dão acesso irrestrito a ferramentas robustas. Um desses programas é o GNU-OCTAVE, utilizado para cálculos numéricos. É devido à aplicabilidade e à facilidade de acesso que este livro trabalha com o OCTAVE.

O OCTAVE foi originalmente desenvolvido para ser um software auxiliar de um livro texto de graduação, utilizado no projeto de um reator químico, que estava sendo escrito por James B. Rawlings, da Universidade Wisconsin-Madison, e John G. Ekerdt, da Universidade do Texas (Eaton et al., 1997).

A partir daí, o OCTAVE se tornou muito mais que apenas um pacote computacional auxiliar destinado ao uso em sala de aula. Inicialmente, não existia a pretensão de se construir um software de ampla aplicação e o foco do projeto era apenas proporcionar uma ferramenta que auxiliasse os alunos na resolução de problemas reais e que pudesse, também, ser usada num contexto um pouco mais amplo que os problemas oriundos do projeto de reatores químicos. O OCTAVE é de simples utilização, e mesmo iniciantes são capazes de dominar os comandos básicos rapidamente, usando-os de maneira eficiente sem muita dificuldade.

Eaton et al. (1997) afirmam que as pessoas são levadas naturalmente a pensar que o nome OCTAVE busca fazer uma analogia com a música, mas, na realidade, é o nome do professor de um dos criadores do software.

O OCTAVE acabou sendo usado em diversas disciplinas da graduação do Departamento de Engenharia Química e do Departamento de Matemática da Universidade do Texas, para ensinar equações diferenciais e também álgebra linear.

De modo geral, o OCTAVE tem uma sintaxe semelhante à do MATLAB®, oferecendo recursos suficientes para a maioria dos usuários na resolução de problemas numéricos de álgebra linear, manipulação de polinômios, integração de equações diferenciais ordinárias e equações algébrico-diferenciais, além de vários outros recursos. Uma das grandes diferenças entre o OCTAVE e o MATLAB® é o ambiente gráfico que eles utilizam; enquanto o segundo utiliza um ambiente GUI (*graphics user interface*) o OCTAVE faz uso de um CLI (*command line interface*). O OCTAVE possui um método interativo, em que os comandos são digitados diretamente no prompt; todavia, também aceita o modo batch, como no MatLab®, que é representado pelos "m-files". A grande vantagem do OCTAVE é que ele também interpreta, na maioria dos casos, estes "m-files", bastando que estes sejam salvos no local adequado.

Mesmo para usos não específicos nota-se uma tendência crescente de valorização de softwares livres, como substituto do Windows®, do Linux. Entre os vários motivos da opção pelo software livre, que seria mais importante que a questão do preço, está a liberdade. Várias empresas e instituições acadêmicas estão optando pelo Linux, porque nele se alia, à liberdade, a eficiência que esse tipo de software oferece. Assim, o OCTAVE se apresenta como uma das melhores alternativas de software livre, quando comparado ao MATLAB®. Trata-se de um software que tem direitos reservados e algumas limitações de distribuição que buscam assegurar que todas as pessoas possam usufruir da mesma liberdade de uso e de redistribuição. Essas condições de uso e redistribuição são especificadas na licença geral pública de GNU que acompanha o software.

O OCTAVE está disponível na Internet em *ftp://ftp.che.wisc.edu/pub/octave* ou, ainda, no site oficial *http://www.gnu.org/software/octave/download.html*; informações adicionais estão disponíveis em *http://www.che.wisc.edu/octave*. Embora originalmente o OCTAVE tenha sido implementado para o ambiente Linux, existem hoje versões para Mac OS X® e Windows®, que podem ser encontradas em *http://octave.sourceforge.net/.*

Ao executar o OCTAVE é exibida uma janela reproduzida na Figura 7.1:

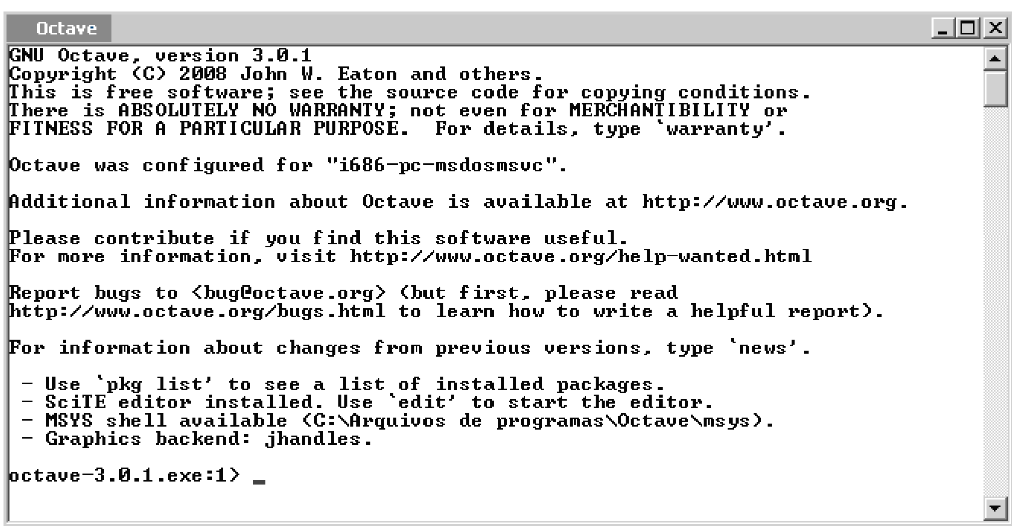



FIGURA 7.1 - Interface do OCTAVE

Na última linha aparece o prompt do OCTAVE, a partir do qual, com o uso do cursor, são digitados os comandos.

Alguns efeitos visuais básicos podem ser ajustados no OCTAVE. Clicando com o botão direito do mouse sobre a faixa azul e, em seguida, em "Propriedades", pode-se, por exemplo, alterar as cores da fonte e do fundo.

O OCTAVE possui uma função interna que verifica automaticamente erros de comandos inválidos. Um erro comum, denominado erro de sintaxe, acontece quando o OCTAVE não entende algo que foi digitado. Por exemplo, se ao digitar com erros um determinado comando, obtém-se:

```
>> functon y=cubica(x) y=x^3;endfunction
```

o OCTAVE responderá com uma mensagem do tipo:

```
parse error:
   syntax error
>> functon y=cubica(x) y=x^3;endfunction

            ^
```

Para a maioria dos erros de sintaxe, o OCTAVE usa um sinal (^), para marcar o ponto aproximado onde se localiza o erro. Neste exemplo, o OCTAVE acusou um erro de comando mal escrito, `functon` em vez de `function`. Estas mensagens facilitam muito a identificação e correção dos erros de compilação.

Outra forma de erro ocorre quando o OCTAVE não reconhece uma variável ou a estrutura de um comando, como no exemplo abaixo, em que a variável `t` não foi definida:

```
>> t
error: 't' undefined near line 1 column 1
```

No OCTAVE é possível obter comandos anteriores sem redigitá-los novamente bastando utilizar as teclas direcionais “para cima” e “para baixo” (as setas ↑ e ↓) do computador. Estes comandos são visualizados do último para o primeiro. É possível, ainda, utilizar as teclas “esquerda” e “direita” (as setas ¬ e ®) para movimentar o cursor durante a digitação de um comando, para corrigir, apagar (teclas *DELETE* ou *BACKSPACE*) ou inserir novos caracteres.

## 7.2 OPERADORES ARITMÉTICOS

Os principais operadores aritméticos do OCTAVE são apresentados a seguir.

QUADRO 1 - Operações e símbolos no OCTAVE

| operações | símbolos |
|---|---|
| adição | + |
| subtração | – |
| multiplicação | * |
| divisão | / |
| potenciação | ^ |

As operações elementares são apresentadas no OCTAVE em uma hierarquia usual, conforme a ordem apontada abaixo:

QUADRO 2 - Ordem de execução das operações

| ordem | operações |
|---|---|
| 1º | potenciação |
| 2º | multiplicação/divisão |
| 3º | adição/subtração |

Isso significa que, por exemplo, se entrarmos com a operação `2+3^4`, primeiramente, o OCTAVE efetua a potência (3^4 = 81) e depois adiciona 2 a este resultado, exibindo `83`:

```
>> 2+3^4
ans = 83
```

Observe, no exemplo anterior, que o OCTAVE atribui o resultado das operações realizadas a uma variável padrão denominada *ans*. Esta variável pode ser guardada em outra variável mais adequada conforme o desejo do usuário, por exemplo:

```
>> x = ans
x = 83
```

Uma vez armazenado o valor de uma variável, esta pode ser utilizada para gerar novos cálculos, como:

```
>> y = 2*x
y = 166
```

Observe que o sinal de "=" faz uma atribuição, de modo que o resultado da operação executada à direita desse sinal é armazenado na variável determinada pelo usuário; no exemplo acima, a variável em consideração é "`y`". Vírgulas podem ser utilizadas para separar comandos numa mesma linha como no exemplo a seguir:

```
>> x=34,y=56,t=x
x = 34
y = 56
t = 34
```

Uma vez que os valores das variáveis foram atribuídos (exemplo anterior), um ponto e vírgula (;) após um determinado comando atribui o resultado à variável designada e não exibe o resultado na tela:

```
>> TT=x+y+(t*y)/x;
```

Utiliza-se este procedimento em programas longos quando não interessam os resultados intermediários. Vale lembrar que a exibição na tela, do resultado de operações, aumenta o custo computacional. A qualquer momento o valor atribuído a uma determinada variável pode ser visualizado digitando o seu nome:

```
>> TT
TT = 146
```

Além das prioridades peculiares entre os operadores (`+`, `-`, `*` e `/`), os parênteses são úteis para alterar a ordem de uma operação; os parênteses mais internos são avaliados antes dos mais externos, numa ordem "de dentro para fora".

## 7.3 VARIÁVEIS

O OCTAVE tem certas regras para nomear as variáveis – elas devem ter nomes iniciados por letras e não podem conter espaços nem caracteres de pontuação. Além disso, o OCTAVE distingue letras maiúsculas de minúsculas; por exemplo, as variáveis `Aa`, `aA`, `aa` e `AA` são todas diferentes para o OCTAVE.

Algumas variáveis, denominadas *keywords* (*palavras-chave*), possuem um significado predefinido no OCTAVE. Estas *keywords* são reservadas e não podem ser utilizadas como variáveis quaisquer; do contrário, o OCTAVE pode incorrer em algum tipo de erro. O Quadro 3 abaixo mostra alguns exemplos de *keywords* reservadas:

Quadro 3 - Exemplos de *keywords* do Octave

| **Comandos** | | | | |
|---|---|---|---|---|
| break | for | varargout | diary | mlock |
| case | function | while | echo | more |
| catch | global | __end__ | edit_history | munlock |

| | | | | |
|---|---|---|---|---|
| continue | if | addpath | format | pkg |
| do | otherwise | autoload | help | rmdir |
| else | persistent | cd | history | rmpath |
| elseif | replot | chdir | iscommand | run_history |
| end | return | clear | iskeyword | save |
| endfor | static | dbclean | isvarname | type |
| endfunction | switch | dbstatus | load | warning |
| endif | try | dbstop | lookfor | which |
| endswitch | until | dbtype | mislocked | who |
| endwhile | varargin | dbwhere | mkdir | whos |

Para saber o que cada um destes comandos faz, basta digitar `help palavra`, onde `palavra` é qualquer um dos comandos exibidos no Quadro 3. Assim, o OCTAVE exibe informações detalhadas sobre as peculiaridades do comando digitado.

Com exceção das *keywords*, que são predefinidas e não podem ser consideradas variáveis, qualquer variável pode ser renomeada, ou pode receber atribuição de qualquer valor, em qualquer momento.

O usuário pode visualizar as variáveis declaradas, apagar qualquer uma delas, ou apagar todas elas, conforme os seguintes comandos:

- `who` - mostra as variáveis declaradas pelo usuário;
- `clear teste` - apaga a variável chamada `teste`;
- `clear` - apaga todas as variáveis declaradas pelo usuário.

## 7.4 FORMATOS NUMÉRICOS DO OCTAVE

O OCTAVE possui critérios específicos para a apresentação de resultados numéricos, que levam em conta o formato e o número de casas que estão sendo utilizados. A seguir são exibidos alguns dos formatos de exibição das variáveis no OCTAVE. A sintaxe é `format tipo`, onde `tipo` é uma das *keywords* a seguir:

QUADRO 4 - Tipos de formatos numéricos no Octave

| Tipos | | | |
|---|---|---|---|
| short | long | free | compact |
| short e | long e | bank | loose |
| short E | long E | plus | rat |
| short g | long g | native-hex | native-bit |
| short G | long G | hex | bit |

O formato padrão do OCTAVE é `format short`, em que os resultados numéricos são exibidos com 5 algarismos significativos, por exemplo:

```
>> 7/8
ans = 0.87500
```

O `format bank` exibe apenas 2 casas decimais com arredondamento automático:

```
>> format bank
>> 7/8
ans = 0.88
```

O `format long` exibe 15 algarismos significativos:

```
>> format long
>> 7/8
ans = 0.875000000000000
```

Mais uma vez, o comando `help format` exibe as características do comando `format`.

## 7.5 MATRIZES E VETORES

A criação de uma matriz no OCTAVE é feita digitando-se os elementos da matriz entre colchetes, e separando os elementos de uma linha por espaços ou vírgulas, e as linhas por ponto e vírgula; pode-se também teclar <*ENTER*> após a digitação de cada linha. A entrada de um vetor é considerada como um caso particular de matriz, apresentando apenas uma linha ou uma coluna, respectivamente, para o caso de vetor linha ou vetor coluna.

A matriz pode ser guardada em uma variável, para cálculos ou consultas posteriores:

```
>> A=[1,2,3;4,5,6;7,8,9]
A =
1 2 3
4 5 6
7 8 9
```

ou

```
>> A=[1 2 3
4 5 6
7 8 9]
A =
1 2 3
4 5 6
7 8 9
```

Pode-se criar uma matriz sem que ela seja mostrada na tela, de modo análogo à criação de qualquer variável, bastando colocar um ponto e vírgula no final da linha de comando:

```
>> A=[1,2,3;4,5,6;7,8,9];
```

Existem comandos para criação de matrizes específicas, como *identidade*, *nula* etc. Por exemplo, para gerar uma matriz onde todos os elementos são iguais a 1, utiliza-se o comando `ones`:

```
>> A=ones(3)
A =
1 1 1
1 1 1
1 1 1
```

Para gerar uma matriz nula, é utilizado o comando `zeros`:

```
>> Z=zeros(3)
Z =
0 0 0
0 0 0
0 0 0
```

Para gerar uma matriz com valores aleatórios (entre 0 e 1), utiliza-se o comando `rand`:

```
>> B=rand(3)
B =
0.22359 0.69924 0.44681
0.58534 0.75207 0.81055
0.42150 0.38117 0.68350
```

Nos três exemplos anteriores, são criadas matrizes quadradas, uma vez que o argumento entre parênteses é apenas um número, nestes casos, 3. Porém, podem-se criar matrizes de qualquer ordem considerando os argumentos entre parênteses como os números de linha e de coluna. Observe-se o exemplo:

```
>> Z=zeros(3,2)
Z =
0 0
0 0
0 0
```

Pode-se criar uma matriz identidade por meio do comando `eye`:

```
>> I=eye(3)
I =
```

```
1 0 0
0 1 0
0 0 1
```

É possível apagar qualquer linha ou coluna de uma matriz, por meio de um simples comando. Considere, por exemplo, a matriz A:

```
>> A=[1,2,3;4,5,6;7,8,9];
1 2 3
4 5 6
7 8 9
```

Para apagar na sequência, respectivamente, a terceira linha e a segunda coluna, faz-se o que segue:

```
>> A(3,:)=[]
A =
4 5 6
7 8 9
```

```
>> A(:,2)=[]
A =
4 6
7 9
```

De modo geral, para apagar a m-ésima linha de uma matriz A qualquer, executa-se o comando `A(m,:)=[]`, e, de modo análogo, para apagar a n-ésima coluna desta matriz A, o comando `A(:,n)=[]`.

O Octave obedece à terminologia usual para matrizes utilizando a notação de matrizes B(i,j), onde i representa a linha, e j a coluna. Por exemplo, considerando a matriz A anterior, o elemento da i-ésima linha e j-ésima coluna é exibido por meio do comando A(i,j). Quando se utiliza o símbolo ":" no lugar de uma linha ou coluna, ele significa "todos os elementos":

| | |
|---|---|
| >> A=[1,2,3;4,5,6;7,8,9];<br>1 2 3<br>4 5 6<br>7 8 9<br>>> A(1,2)<br>ans = 2<br>>> A(2,2)<br>ans = 5<br>>> A(1,:)<br>ans =<br>1 2 3 | >> A(:,2)<br>ans =<br>2<br>6<br>8<br>>> B(:,1:2)<br>ans =<br>1 2<br>5 6<br>7 8 |

O último comando à direita fornece os elementos das linhas e da coluna 1 até a 2.

No OCTAVE é possível acrescentar linhas ou colunas extras a uma dada matriz já existente. Vale lembrar que devem ser tomadas as devidas precauções com o número de elementos adequados a cada linha ou coluna da matriz original, pois qualquer discrepância faz com o que o OCTAVE apresente mensagem de erro. Veja os exemplos a seguir:

```
>> A=[1,2,3;4,5,6;7,8,9];
1 2 3
4 5 6
7 8 9
```

Para adicionar uma quarta coluna à matriz A, executa-se o seguinte comando:

```
>> A=[A,[11;12;13]]
A =
1 2 3 11
4 5 6 12
7 8 9 13
```

Para acrescentar uma quarta linha à matriz A, executa-se o seguinte comando:

```
>> A=[A;[11 12 13]]
A =
 1  2  3
 4  5  6
 7  8  9
11 12 13
```

## 7.6 OPERAÇÕES MATRICIAIS

O OCTAVE executa diretamente as operações elementares matriciais de forma semelhante às operações escalares, desde que sejam satisfeitas as exigências sobre a ordem das matrizes, de modo que as operações sejam consistentemente efetuadas. Veja os exemplos seguintes:

| | |
|---|---|
| `>> A=[1,2,3;4,5,6;7,8,9];`<br>`1 2 3`<br>`4 5 6`<br>`7 8 9` | `>> B=[11,22,33;44,55,66;77,88,99];`<br>`11 22 33`<br>`44 55 66`<br>`77 88 99` |
| `>> A + B`<br>`ans =`<br>`12  24   36`<br>`48  60   72`<br>`84  96  108` | `>> 5*A`<br>`ans =`<br>` 5  10  15`<br>`20  25  30`<br>`35  40  45` |
| `>> A - B`<br>`ans =`<br>`-10  -20  -30`<br>`-40  -50  -60`<br>`-70  -80  -90` | `>> A^2`<br>`ans =`<br>` 30   36   42`<br>` 66   81   96`<br>`102  126  150` |

Abaixo é apresentado um resumo das principais operações matriciais:

QUADRO 5 - Principais operações matriciais

| Símbolo | Operação | Exemplo |
|---|---|---|
| + | adição | A + B |
| – | subtração | A – B |
| * | multiplicação | A*B |
| ^ | potenciação ($A^k$) | A^k |
| .^ | potenciação elemento a elemento ($a_{ij}^k$) | A.^k |
| ` | transposição | A’ |

## 7.7 DECLARAÇÕES DE CONTROLE

As declarações de controle podem ser uma simples expressão, ou uma lista de rotinas mais complexas, ou declarações condicionais. As declarações de controle tais como `if`, `while`, `for`, dentre outras, controlam o fluxo de execução em códigos escritos no OCTAVE.

Todas estas declarações de controle são iniciadas por *keywords* especiais de modo a diferenciá-las de expressões comuns. Em geral, as declarações de controle contêm outras declarações que podem ou não ser executadas, dependendo do resultado do condicional considerado.

Toda declaração de controle tem uma declaração `end` correspondente que marca o fim da declaração de controle. Por exemplo, `endif` e `endwhile` determinam o final, respectivamente, das declarações `if` e `while`. As declarações contidas entre uma declaração de controle e a sua declaração `end` correspondente são denominadas *corpo* da declaração de controle.

- **A declaração** `if`

O `if` no OCTAVE é uma declaração de tomada de decisão. Existem três formas básicas de uma declaração `if`. A sua forma mais simples é dada por:

```
if (condição)
    then-corpo
endif
```

onde *condição* é uma expressão que controla o que o resto da declaração deve executar. E *then-corpo* é executada apenas se a *condição* é verdadeira.

A *condição* numa declaração `if` é verdadeira se é não nula, e falsa se seu valor é zero. Se o valor da expressão condicional é uma matriz ou vetor, ela é considerada verdadeira apenas se todos os seus valores são não nulos.

A segunda forma de uma declaração `if` é:

```
if (condição)
    then-corpo
else
    else-corpo
endif
```

Se a *condição* é verdadeira, *then-corpo* é executada, caso contrário *else-corpo* é executada. Considere o seguinte exemplo, onde `rem(x,2)==0` verifica se a variável `x` é divisível por 2, e `printf` é utilizado para exibir a mensagem entre aspas na tela.

```
if (rem(x,2)==0)
    printf("xiseven\n");
else
    printf("xisodd\n");
endif
```

A forma mais geral da declaração `if` permite múltiplas decisões, conforme o seguinte:

```
if (condição)
    then-corpo
elseif (condição)
    elseif-corpo
else
    else-corpo
endif
```

Qualquer número de `elseif` pode ser utilizado; neste caso cada uma das condições é testada na sequência, e se uma destas é verdadeira, seu *corpo* é executado. Por outro lado, se nenhuma das `elseif` é verdadeira e existe um `else`, então seu *corpo* é executado. Aparece, então, apenas um `else` e este deve ser a última parte da declaração.

No exemplo a seguir, se a primeira condição é verdadeira, ou seja, se a variável `x` é divisível por 2, então a primeira mensagem é exibida na tela. Se ela é falsa, então a segunda condição é avaliada. Se a segunda condição é verdadeira, ou seja, se a variável `x` é divisível por 3, então a segunda mensagem é exibida na tela. Caso contrário, a terceira mensagem é exibida.

```
if(rem(x,2)==0)
    printf("x é par");
elseif(rem(x,3)==0)
    printf("x é ímpar e divisível por 3");
else
    printf("x é ímpar");
endif
```

- **A declaração** `while`

No contexto de programação, um *laço* (loop) é uma parte do programa que pode ser executado duas ou mais vezes sucessivamente, ou seja, podem ser realizadas ações em repetição. A declaração `while` é o tipo de loop mais simples no OCTAVE. As ações são repetidas enquanto uma dada condição é verdadeira. De forma semelhante à declaração `if`, a condição num `while` é considerada verdadeira se seu valor é não nulo, e falsa se este valor é zero. E, também de forma semelhante ao `if`, se o valor da expressão condicional é um vetor ou uma matriz, esta condição será verdadeira apenas se todos os elementos forem não nulos.

Uma declaração `while` no OCTAVE se apresenta da seguinte forma:

```
while (condição)
      corpo
endwhile
```

onde *corpo* é uma declaração ou lista de declarações denominadas *corpo* do loop, e a *condição* é uma expressão que controla até quando o loop deve ser executado.

Uma das funções do `while` é o teste de condição. Se a condição é verdadeira, o `while` executa o corpo da declaração. Após esta execução a condição é testada novamente e, sendo verdadeira, mais uma vez o corpo é executado. Este processo se repete até que a condição seja falsa. Se a condição é inicialmente falsa, o corpo nunca é executado.

O exemplo seguinte cria uma variável `fibonacci`, que contém os primeiros elementos da sequência de Fibonacci:

```
fibonacci=ones(1,10);
i=3;
while(i<=10)  fibonacci(i)=fibonacci(i-
    1)+fibonacci(i-2); i=i+1;
endwhile
```

Nessa sequência o *corpo* é constituído de duas declarações e o loop trabalha de forma muito simples. Primeiramente, é atribuído o valor 3 para a variável `i` e é criado um vetor com dez elementos todos iguais a um (`fibonacci`). Então o `while` testa até quando esta variável é menor ou igual a 10. O primeiro elemento a ser calculado, para quando `i = 3` é `fibonacci(3) = fibonacci(2) + fibonacci(1)`, que resulta em 2. A segunda declaração do *corpo* é um contador e adiciona uma unidade à variável `i`, que passa a ser 4. E o processo se repete até que `i = 11`, o que faz a condição do `while` se tornar falsa e o processo parar.

- **A declaração** `for`

A declaração `for` é mais conveniente para iterações de um loop. A forma geral do `for` é dada por:

```
for var = expression
    corpo
endfor
```

O corpo, neste caso, representa qualquer declaração ou lista de declarações; expressão é qualquer expressão válida e `var` pode assumir várias formas. De maneira geral, `var` é uma variável simples ou indexada. Se o valor da expressão é uma estrutura, `var` pode ser uma lista.

A atribuição da expressão na declaração `for` trabalha um pouco diferente da atribuição natural das declarações no OCTAVE. O seguinte exemplo mostra outra maneira de criar um vetor que contém os 10 primeiros elementos da sequência de Fibonacci; diferentemente do exemplo anterior, agora utiliza a declaração `for`:

```
fibonacci=ones(1,10);
    for i=3:10
    fibonacci(i)=fibonacci(i-1)+fibonacci(i-2);
endfor
```

Este código produz uma lista de valores de 3 a 10 (números naturais). Dentro do loop, cada um destes valores é atribuído à variável `i` e o i-ésimo elemento da variável `fibonacci` é calculado pela expressão `fibonacci(i)= = fibonacci(i-1)+ fibonacci(i-2)`. Este processo é executado para todos os elementos da lista.

Mostraremos a seguir, algumas das funções mais comuns, exemplificando sua utilização no OCTAVE:

| | |
|---|---|
| valor absoluto ou magnitude de um número complexo: | `>> abs(3 + 4i)`<br>`ans= 5` |
| inverso do cosseno: | `>> acos(0.8)`<br>`ans =  0.64350` |
| inverso do cosseno hiperbólico: | `>> acosh(0.8)`<br>`ans = 0.00000 + 0.64350i` |
| ângulo de um número complexo em radianos: | `>> angle(4+i)`<br>`ans = 0.24498` |
| inverso do seno: | `>>asin(0.5)`<br>`ans = 0.52360` |
| inverso do seno hiperbólico: | `>> asinh(0.5)`<br>`ans = 0.48121` |
| inverso da tangente: | `>> 4*atan(1) % o resultado é o valor aproximado de pi`<br>`ans = 3.1416` |
| inverso da tangente em radianos nos quatro quadrantes do ângulo de um número complexo: | `>> s=atan2(3,2i)`<br>`s =     1.5708` |
| inverso da tangente hiperbólico: | `>> atanh(pi/4)`<br>`ans =  1.0593` |
| conjugado de um complexo: | `>>conj(1- 4i)`<br>`ans = 1 + 4i` |
| cosseno: | `>> cos(pi/3)`<br>`ans = 0.50000` |
| exponencial: | `>> exp(1)`<br>`ans = 2.7183` |

| | |
|---|---|
| parte imaginária de um número complexo: | `>> imag(5 + 2i)`<br>`ans = 2` |
| mínimo múltiplo comum dos inteiros x e y: | `>> lcm(18,81)`<br>`ans = 162` |
| logaritmo natural: | `>>h=log(2)`<br>`h = 0.69315` |
| logaritmo na base 10: | `>> log10(5)`<br>`ans = 0.69897` |
| parte real de um número complexo: | `>> real(5 + 2i)`<br>`ans = 5` |
| resto da divisão de x por y: | `R=rem(23,4)`<br>`R = 3` |
| arredondamento em direção ao inteiro mais próximo: | `round(-2.6)`<br>`ans = -3` |
| se x é menor que zero, a função `sign` retorna ao valor –1; se x é igual a zero, retorna ao valor 0; se x é positivo, retorna ao valor 1 | `>> sign(-2.6)`<br>`ans = -1` |
| cálculo do seno de um número: | `>> sin(0.5)`<br>`ans = 0.47943` |
| seno hiperbólico: | `>> sinh(0.5)`<br>`ans =    0.5211` |
| cosseno hiperbólico: | `>> cosh(0.5)`<br>`ans = 1.1276` |
| raiz quadrada: | `>>y=sqrt(3^2+4^2)`<br>`y = 5` |
| tangente: | `>>tan(sqrt(3))`<br>`ans = -6.1475` |

Também é possível gerar gráficos de várias funções. Como exemplo, apresentamos a seguir a implementação no OCTAVE para funções exponenciais.

```
% função exponencial
t=-5:0.1:5;
x=exp(t);
```

```
% com a variável t positiva
y=exp(-t);
% com a variável t negativa
plot(t,x,t,y,'r')
% plota dois gráficos positivos, um crescente
legend('função x', 'função y')
grid
axis([-5 5 -5 100])
%limita os eixos
title('função exponencial')
xlabel('eixo x')
ylabel('eixo y')
```

A Figura 7.2 mostra o gráfico resultado das funções apresentadas acima:

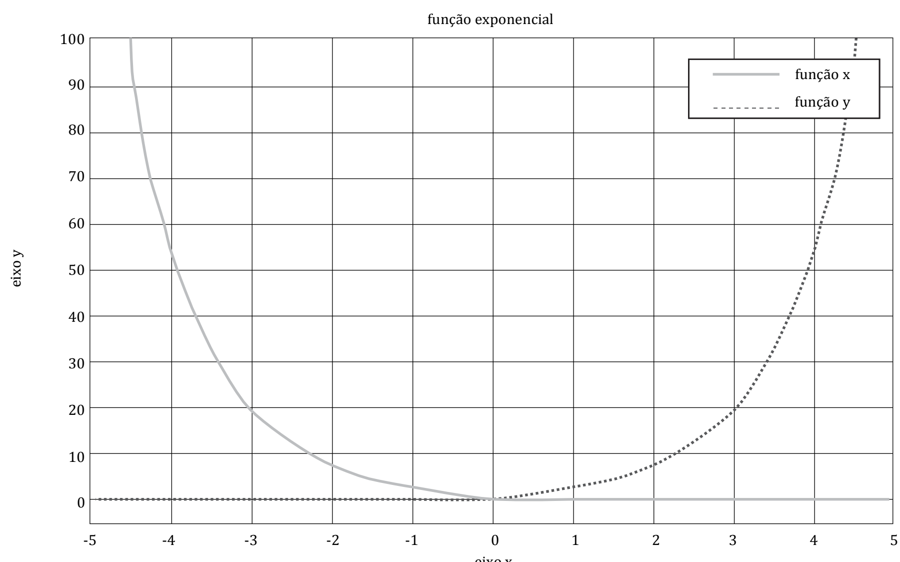



FIGURA 7.2 - Gráfico de funções exponenciais

No caso das funções hiperbólicas, são exibidos os gráficos de senhx, coshx e tanhx, cuja implementação no OCTAVE é a seguinte:

```
t=-2:0.01:2;
x=cosh(t);
y=sinh(t);
z=tanh(t);
plot(t,x,'b',t,y,'r',t,z,'k')
grid
>> legend('função cosh, função senh, função tanh')
>> xlabel('t')
>> ylabel('x, y, z')
>> title('Gráfico de Funções Hiperbólicas')
```

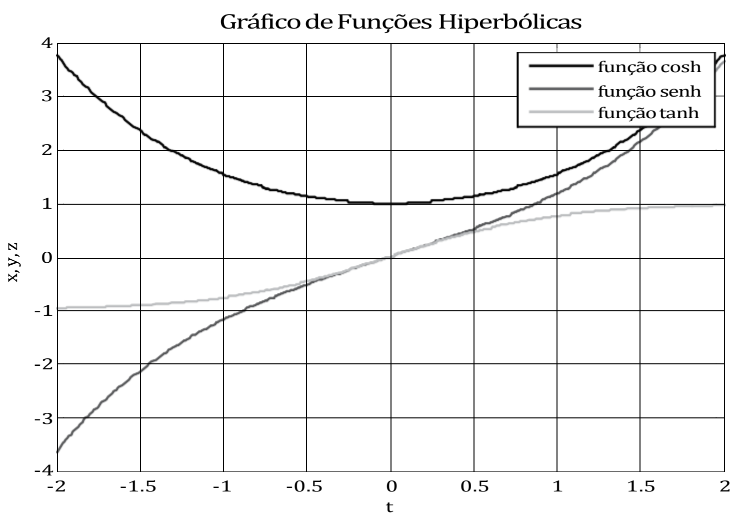



FIGURA 7.3 - Gráfico de funções hiperbólicas

# REFERÊNCIAS


ALENCAR FILHO, E. de. *Iniciação à lógica matemática*. São Paulo: Nobel, 2002.

BOLDRINI, J. L.; COSTA, S. I. R.; RIBEIRO, V. L.; WETZLER, H. G. *Álgebra linear*. 3. ed. São Paulo: Harper & Row do Brasil, 1980.

CAMARGO, I.; BOULOS, P. *Geometria analítica:* um tratamento vetorial. 3. ed. São Paulo: Pearson Prentice Hall, 1987.

BUENO, H. P. *Álgebra linear:* um segundo curso. São Paulo: SBM, 2006.

CALLIOLI, C. A.; DOMINGUES, H. H.; COSTA, R. C. F. *Álgebra linear e aplicações*. 6. ed. São Paulo: Atual, 1990.

CORRÊA, P. S. Q. *Álgebra linear e geometria analítica*. Rio de Janeiro: Interciência, 2006.

DAGHLIAN, J. *Lógica e álgebra de Boole*. 4. ed. São Paulo: Atlas, 1995.

EATON, J. W; BATEMAN, D.; HAUBERG, S.; WEHBRING, R. *GNU Octave*: a high-level interactive language for numerical computations, Edition 3, version, 3.8.0, 1997.

Disponível em: https://www.gnu.org/software/octave/octave.pdf. Acesso em: out. 2014.

GERSTING, J. L. *Fundamentos matemáticos para a Ciência da Computação*. 5. ed. Rio de Janeiro: LTC, 2004.

HERNSTEIN, I. N. *Tópicos de álgebra*. Tradução de Adalberto Bergamasco. São Paulo: Polígono, 1970.

HOFFMAN, K.; KUNZE, R. *Linear algebra*. 2nd Ed. New Jersey: Prentice Hall, 1971.

KOLMAN, B.; HILL, D. R. *Introdução à álgebra linear*: com aplicações. 8. ed. São Paulo: LTC, 2006.

LANG, S. *Algebra*. Reading, Mass: Addison-Wesley, 1971.

LAX, P. *Linear algebra*. New York: John Wiley, 1997.

LAY, D. C. *Álgebra linear e suas aplicações*. 2. ed. Rio de Janeiro: LTC, 1999.

LEITHOLD, L. *O cálculo com Geometria Analítica*. 3. ed. São Paulo: Harbra, 1994, v. 1.

LIMA, E. L. *Álgebra linear*. 6. ed. Rio de Janeiro: IMPA, 2003.

LIPSCHUTZ, S. *Álgebra linear*. 3. ed. São Paulo: Makron Books, 1994.

MATTHEWS, K. R. *Elementary linear algebra*. Queensland: Department of Mathematics, University of Queensland, 1991.

MENEZES, P. B.; Haeusler, E. H. *Matemática discreta para computação e informática*. Porto Alegre: Sagra Luzzato, 2004.

MONTEIRO, L. H. A. *Sistemas dinâmicos*. 3. ed. São Paulo: Livraria da Física, 2011.

NOBLE, B.; DANIEL, J. W. *Álgebra linear aplicada*. 2. ed. Rio de Janeiro: Prentice Hall do Brasil, 1986.

PETTOFREZZO, A. J. *Matrices and transformations*. New York: Dover Publications, Inc., 1966.

POOLE, D. *Álgebra linear*. São Paulo: Thomson Learning, 2006.

SCHEINERMAN, E. R, *Matemática discreta:* uma introdução. São Paulo: Thomson, 2003.

SILVA, V. V. *Álgebra linear*. Goiânia: Ed. UFG, 1999.

STEINBRUCH, A.; WINTERLE, P. *Álgebra linear*. 2. ed. Rio de Janeiro: McGraw-Hill, 2008.

STRANG, G. *Linear algebra and its applications*. 3. ed. San Diego: HBJ, 1988.

# ÍNDICE REMISSIVO

## A



## C

## V

*Revisão*
Cânone Editoração Ltda.

*Projeto gráfico e editoração eletrônica*
Alanna Oliva

*Capa*
M. H. Stoppa

Dados internacionais de catalogação-na-publicação (CIP)
GPT/BSCAC/UFG

Stoppa, M. H.
S883a Álgebra linear com octave / M. H. Stoppa, R. A. Borges. – Goiânia: Gráfica UFG, 2015.
242 p. – (Coleção Euler; 01)

ISBN: 978-85-68359-72-3.

1. Álgebra linear. 2. Matrizes. 3. Sistemas – equações lineares. I. Borges, R. A. II. Título.

CDU: 512.64