INTRODUÇÃO À ÁLGEBRA LINEAR

Reginaldo J. Santos
Departamento de Matemática-ICEx
Universidade Federal de Minas Gerais
http://www.mat.ufmg.br/~regi

Julho 2010

Introdução à Álgebra Linear

Editor, Coordenador de Revisão, Supervisor de Produção, Capa e Ilustrações:
Reginaldo J. Santos

ISBN 85-7470-018-5

Ficha Catalográfica

Santos, Reginaldo J.
S237i Introdução à Álgebra Linear / Reginaldo J. Santos
- Belo Horizonte: Imprensa Universitária da UFMG, 2013.

1. Álgebra Linear I. Título

CDD: 512.5

Sumário

Prefácio

Este texto cobre o material para um curso de um semestre de Introdução à Álgebra Linear ou de Álgebra Linear Matricial. O texto pode, mas **não** é necessário, ser acompanhado um programa como o MATLAB®*, SciLab ou o Maxima.

O conteúdo é dividido em sete capítulos. O Capítulo 1 trata das matrizes e sistemas lineares. Aqui todas as propriedades da álgebra matricial são demonstradas. A resolução de sistemas lineares é feita usando somente o método de Gauss-Jordan (transformando a matriz até que ela esteja na forma escalonada reduzida). Este método requer mais trabalho do que o método de Gauss (transformando a matriz, apenas, até que ela esteja na forma escalonada). Ele foi o escolhido, por que também é usado no estudo da inversão de matrizes no Capítulo 2. Neste Capítulo é também estudado o determinante, que é definido usando cofatores. As demonstrações dos resultados deste capítulo podem ser, a critério do leitor, feitas somente para matrizes 3×3.

O Capítulo 3 trata de vetores no plano, no espaço e no $\mathbb{R}^n$. Os vetores são definidos inicialmente de forma geométrica, assim como a soma e a multiplicação por escalar. São provadas algumas propriedades geometricamente. Depois são introduzidos sistemas de coordenadas de forma natural sem a necessidade da definição

*MATLAB® é marca registrada de The Mathworks, Inc.

de base. O produto escalar é definido também geometricamente. São estudados também retas e planos no espaço. Depois, o conceito de vetor é generalizado para o $\mathbb{R}^n$. O conceito de dependência e independência linear é introduzido de forma algébrica, acompanhado da interpretação geométrica para os casos de $\mathbb{R}^2$ e $\mathbb{R}^3$.

No Capítulo 4 são tratados os conceitos de subespaços e de base de subespaços. São estudados os espaços linha e coluna de uma matriz e o seu posto. Ao final do capítulo os Espaços Vetoriais Abstratos são definidos. No Capítulo 5 são abordados o produto escalar e bases ortonormais. Além de subespaços ortogonais e quadrados mínimos.

Transformações Lineares de $\mathbb{R}^n$ em $\mathbb{R}^m$ são estudadas no Capítulo 6. O Capítulo 7 traz um estudo da diagonalização de matrizes em geral e a diagonalização de matrizes simétricas através de uma matriz ortogonal. É feita uma aplicação ao estudo das seções cônicas.

Os exercícios estão agrupados em três classes. Os "Exercícios Numéricos", que contém exercícios que são resolvidos fazendo cálculos, que podem ser realizados sem a ajuda de um computador ou de uma máquina de calcular. Os "Exercícios Teóricos", que contém exercícios que requerem demonstrações. Alguns são simples, outros são mais complexos. Os mais difíceis complementam a teoria e geralmente são acompanhados de sugestões. Os "Exercícios usando o MATLAB®", que contém exercícios para serem resolvidos usando o MATLAB® ou outro software. Os comandos necessários a resolução destes exercícios são também fornecidos juntamente com uma explicação rápida do uso. Os exercícios numéricos são imprescindíveis, enquanto a resolução dos outros, depende do nível e dos objetivos pretendidos para o curso.

O MATLAB® é um software destinado a fazer cálculos com matrizes (MATLAB® = MATrix LABoratory). Os comandos do MATLAB® são muito próximos da forma como escrevemos expressões algébricas, tornando mais simples o seu uso. Podem ser incorporados às rotinas pré-definidas, pacotes para cálculos específicos. Um pacote chamado `gaal` com funções que são direcionadas para o estudo de Geometria Analítica e Álgebra Linear pode ser obtido na web na página do autor, assim como um texto com uma introdução ao MATLAB® e instruções de como instalar o pacote `gaal`. O MATLAB® não é um software gratuito, embora antes a versão estudante vinha grátis ao se comprar o guia do usuário. Atualmente o SciLab é uma alternativa gratuita, mas que não faz cálculo simbólico. O Maxima é um programa de computação algébrica gratuito. Ambos podem ser usados como ferramenta auxiliar na aprendizagem de Álgebra Linear. Na página do autor na web podem ser encontrados pacotes de funções para estes programas além de links para as páginas do SciLab e do Maxima e várias páginas interativas que podem auxiliar na aprendizagem.

No fim de cada capítulo temos um "Teste do Capítulo", onde o aluno pode avaliar os seus conhecimentos. Os Exercícios Numéricos e os Exercícios usando o MATLAB® estão resolvidos após o último capítulo utilizando o MATLAB®. Desta forma o leitor que não estiver interessado em usar o software pode obter apenas as respostas dos exercícios, enquanto aquele que tiver algum interesse, pode ficar sabendo como os exercícios poderiam ser resolvidos fazendo uso do MATLAB® e do pacote `gaal`.

Gostaria de agradecer aos professores que colaboraram apresentando correções, críticas e sugestões, entre eles Helder C. Rodrigues e Francisco Satuf, Joana Darc A. S. da Cruz e Lucia Brasil.

Sugestão de Cronograma para 60 Horas

Capítulo 1	Seções 1.1 e 1.2	8 aulas
Capítulo 2	Seções 2.1 e 2.2	8 aulas
Capítulo 3	Seções 3.1 a 3.3	12 aulas
Capítulo 4	Seções 4.1 e 4.2	8 aulas
Capítulo 5	Seções 5.1 a 5.3	12 aulas
Capítulo 7	Seções 7.1 a 7.3	12 aulas
	Total	60 aulas

Sugestão de Cronograma para 90 Horas

Capítulo 1	Seções 1.1 e 1.2	10 aulas
Capítulo 2	Seções 2.1 e 2.3	12 aulas
Capítulo 3	Seções 3.1 a 3.3	12 aulas
Capítulo 4	Seções 4.1 a 4.3	12 aulas
Capítulo 5	Seções 5.1 a 5.3	12 aulas
Capítulo 6	Seções 6.1 a 6.3	15 aulas
Capítulo 7	Seções 7.1 a 7.3	12 aulas
	Total	85 aulas

Histórico

Julho 2010 Algumas correções. O texto foi totalmente reformatado.

Julho 2009 Algumas correções. Várias figuras foram refeitas.

Março 2008 Foi acrescentada a subseção opcional 'Forma Canônica de Jordan'. Foram corrigidos alguns erros.

Julho 2007 Foi acrescentado o Exemplo 2.16 na seção de Determinantes. Foram acrescentados um exercício na seção de Determinantes, um na de Produto Escalar em $\mathbb{R}^n$, quatro na de Diagonalização e um na de Diagonalização de Matrizes Simétricas. Foram corrigidos alguns erros.

Março 2007 A Seção 1.1 de Matrizes e a Seção 2.2 de Determinantes foram reescritas. Na seção 1.2 o Teorema 1.4 voltou a ser que toda matriz é equivalente por linhas a uma única matriz na forma escalonada reduzida. As seções 4.1, 5.1 e 5.3 foram reescritas e acrescentada uma aplicação à computação gráfica. Foi acrescentada a sub-seção opcional 'Diagonalização de Operadores' à seção 7.1. Foram acrescentados dois exercícios na seção de Matrizes, um na de Inversão de Matrizes, um na de Base e Dimensão, três na de Mudança de Coordenadas. Foram corrigidos alguns erros.

Maio 2004 Foram acrescentadas aplicações à criptografia (Exemplo na página 93) e a cadeias de Markov (Exemplos 1.9 na página 14, 1.16 na página 49 e 7.8 na página 408). Foi acrescentado um exercício na seção 1.1. Foi incluída a demonstração de que toda matriz é equivalente por linhas a uma única matriz escalonada reduzida. Este resultado era o Teorema 1.4 na página 26 que passou para o Apêndice III da seção 5.2. O Teorema 1.4 agora contém as propriedades da relação "ser equivalente por linhas" com a demonstração. O que antes era Exemplo 1.14 passou para o lugar do Exemplo 1.10. O Exemplo 2.5 foi modificado. A seção 'Base e Dimensão' foi reescrita. Foi acrescentada a Proposição 4.9 na página 249 que é útil na obtenção de uma base para um subespaço. O Teorema 4.3 do Apêndice III passou para o texto obrigatório da seção 4.3 e é agora o Teorema 4.2 na página 236. Foram acrescentados alguns exemplos e alguns exercícios à seção 4.3. Os exemplos 7.4 na página 398 e 7.5 na página 405 foram modificados. A seção 'Diagonalização de Matrizes' ganhou mais um exercício teórico.

Julho 2003 Várias correções incluindo respostas de exercícios. A seção 'Base e Dimensão' foi reescrita. Foi acrescentada uma seção de Espaços Vetoriais Abstratos no Capítulo 4. A seção 'Diagonalização de Matrizes′ ganhou mais dois exercícios teóricos. A seção ′Diagonalização de Matrizes Simétricas′ ganhou um apêndice sobre ′Autovalores Complexos′.

Julho 2002 Criado a partir do texto ′Geometria Analítica e Álgebra Linear′ para ser usado numa disciplina de Introdução à Álgebra Linear ou Álgebra Matricial.

1

Matrizes e Sistemas Lineares

1.1 Matrizes

Uma **matriz** A, $m \times n$ (m por n), é uma tabela de mn números dispostos em m linhas e n colunas

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

A i**-ésima linha** de A é

$$\begin{bmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{bmatrix},$$

para $i = 1, \ldots, m$ e a j**-ésima coluna** de A é

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix},$$

para $j = 1, \ldots, n$. Usamos também a notação $A = (a_{ij})_{m \times n}$. Dizemos que a_{ij} ou $[A]_{ij}$ é o **elemento** ou a **entrada** de posição i, j da matriz A.

Se $m = n$, dizemos que A é uma **matriz quadrada de ordem** n e os elementos $a_{11}, a_{22}, \ldots, a_{nn}$ formam a **diagonal (principal)** de A.

Exemplo 1.1. Considere as seguintes matrizes:

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

$$D = \begin{bmatrix} 1 & 3 & -2 \end{bmatrix}, \quad E = \begin{bmatrix} 1 \\ 4 \\ -3 \end{bmatrix} \text{ e } F = \begin{bmatrix} 3 \end{bmatrix}.$$

As matrizes A e B são 2×2. A matriz C é 2×3, D é 1×3, E é 3×1 e F é 1×1. De acordo com a notação que introduzimos, exemplos de elementos de algumas das matrizes dadas acima são $a_{12} = 2$, $c_{23} = -2$, $e_{21} = 4$, $[A]_{22} = 4$, $[D]_{12} = 3$.

Uma matriz que só possui uma linha é chamada **matriz linha**, e uma matriz que só possui uma coluna é chamada **matriz coluna**, No Exemplo 1.1 a matriz D é uma matriz linha e a matriz E é uma matriz coluna. Matrizes linha e matrizes coluna são chamadas de **vetores**. O motivo ficará claro na Seção 3.3 na página 207.

Dizemos que duas matrizes são iguais se elas têm o mesmo tamanho e os elementos correspondentes são iguais, ou seja, $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{p\times q}$ são **iguais** se $m = p$, $n = q$ e $a_{ij} = b_{ij}$ para $i = 1, \ldots, m$ e $j = 1, \ldots, n$.

Vamos definir operações matriciais análogas às operações com números e provar propriedades que são válidas para essas operações. Veremos, mais tarde, que um sistema de equações lineares pode ser escrito em termos de uma única equação matricial.

Vamos, agora, introduzir as operações matriciais.

1.1.1 Operações com Matrizes

Definição 1.1. A **soma** de duas matrizes de **mesmo tamanho** $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{m\times n}$ é definida como sendo a matriz $m \times n$

$$C = A + B$$

obtida somando-se os elementos correspondentes de A e B, ou seja,

$$c_{ij} = a_{ij} + b_{ij}\,,$$

para $i = 1, \ldots, m$ e $j = 1, \ldots, n$. Escrevemos também $[A + B]_{ij} = a_{ij} + b_{ij}$.

Exemplo 1.2. Considere as matrizes:

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

Se chamamos de C a soma das duas matrizes A e B, então

$$C = A + B = \begin{bmatrix} 1+(-2) & 2+1 & -3+5 \\ 3+0 & 4+3 & 0+(-4) \end{bmatrix} = \begin{bmatrix} -1 & 3 & 2 \\ 3 & 7 & -4 \end{bmatrix}$$

Definição 1.2. A **multiplicação de uma matriz** $A = (a_{ij})_{m\times n}$ **por um escalar** (número) α é definida pela matriz $m \times n$

$$B = \alpha A$$

obtida multiplicando-se cada elemento da matriz A pelo escalar α, ou seja,

$$b_{ij} = \alpha\, a_{ij}\,,$$

para $i = 1, \ldots, m$ e $j = 1, \ldots, n$. Escrevemos também $[\alpha A]_{ij} = \alpha\ a_{ij}$. Dizemos que a matriz B é um **múltiplo escalar** da matriz A.

Exemplo 1.3. O produto da matriz $A = \begin{bmatrix} -2 & 1 \\ 0 & 3 \\ 5 & -4 \end{bmatrix}$ pelo escalar -3 é dado por

$$-3\,A = \begin{bmatrix} (-3)(-2) & (-3)\,1 \\ (-3)\,0 & (-3)\,3 \\ (-3)\,5 & (-3)(-4) \end{bmatrix} = \begin{bmatrix} 6 & -3 \\ 0 & -9 \\ -15 & 12 \end{bmatrix}.$$

Definição 1.3. O **produto** de duas matrizes, tais que **o número de colunas da primeira matriz é igual ao número de linhas da segunda**, $A = (a_{ij})_{m\times p}$ e $B = (b_{ij})_{p\times n}$ é definido pela matriz $m \times n$

$$C = AB$$

obtida da seguinte forma:

$$c_{ij} \quad = \quad a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj}, \tag{1.1}$$

para $i = 1, \ldots, m$ e $j = 1, \ldots, n$. Escrevemos também $[AB]_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj}$.

A equação (1.1) está dizendo que o elemento i, j do produto é igual à soma dos produtos dos elementos da i-ésima linha de A pelos elementos correspondentes da j-ésima coluna de B.

$$\begin{bmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & \boxed{c_{ij}} & \vdots \\ c_{m1} & \ldots & c_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ \vdots & & \ldots & \vdots \\ \boxed{a_{i1}} & \boxed{a_{i2}} & \boxed{\ldots} & \boxed{a_{ip}} \\ \vdots & & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \ldots & \boxed{b_{1j}} & \ldots & b_{1n} \\ b_{21} & \ldots & \boxed{b_{2j}} & \ldots & b_{2n} \\ \vdots & \ldots & \boxed{\vdots} & \ldots & \vdots \\ b_{p1} & \ldots & \boxed{b_{pj}} & \ldots & b_{pn} \end{bmatrix}$$

A equação (1.1) pode ser escrita de forma compacta usando a **notação de somatório**.

$$[AB]_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$$

e dizemos "somatório de k variando de 1 a p de $a_{ik}b_{kj}$". O símbolo $\sum\limits_{k=1}^{p}$ significa que estamos fazendo uma soma em que o índice k está variando de $k = 1$ até $k = p$. Algumas propriedades da notação de somatório estão explicadas no Apêndice I na página 28.

Exemplo 1.4. Considere as matrizes:

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

Se chamamos de C o produto das duas matrizes A e B, então

$$C = AB = \begin{bmatrix} 1\,(-2) + 2 \cdot 0 + (-3)\,5 & 1 \cdot 1 + 2 \cdot 3 + (-3)\,(-4) & 0 \\ 3\,(-2) + 4 \cdot 0 + 0 \cdot 5 & 3 \cdot 1 + 4 \cdot 3 + 0\,(-4) & 0 \end{bmatrix} = \begin{bmatrix} -17 & 19 & 0 \\ -6 & 15 & 0 \end{bmatrix}.$$

Observação. No exemplo anterior o produto BA não está definido (por que?). Entretanto, mesmo quando ele está definido, BA pode não ser igual a AB, ou seja, o produto de matrizes **não é comutativo**, como mostra o exemplo seguinte.

Exemplo 1.5. Sejam $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix}$. Então,

$$AB = \begin{bmatrix} -2 & 7 \\ -6 & 15 \end{bmatrix} \quad \text{e} \quad BA = \begin{bmatrix} 1 & 0 \\ 9 & 12 \end{bmatrix}.$$

Vamos ver no próximo exemplo como as matrizes podem ser usadas para descrever quantitativamente um processo de produção.

Exemplo 1.6. Uma indústria produz três produtos, X, Y e Z, utilizando dois tipos de insumo, A e B. Para a manufatura de cada kg de X são utilizados 1 grama do insumo A e 2 gramas do insumo B; para cada kg de Y, 1 grama de insumo A e 1 grama de insumo B e, para cada kg de Z, 1 grama de A e 4 gramas de B. Usando matrizes podemos determinar quantos gramas dos insumos A e B são necessários na produção de x kg do produto X, y kg do produto Y e z kg do produto Z.

$$\begin{array}{r c} & \begin{array}{ccc} \text{X} & \text{Y} & \text{Z} \end{array} \\ \begin{array}{r} \text{gramas de A/kg} \\ \text{gramas de B/kg} \end{array} & \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 4 \end{bmatrix} \end{array} = A \qquad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \begin{array}{l} \text{kg de X produzidos} \\ \text{kg de Y produzidos} \\ \text{kg de Z produzidos} \end{array}$$

$$AX = \begin{bmatrix} x + y + z \\ 2x + y + 4z \end{bmatrix} \begin{array}{l} \text{gramas de A usados} \\ \text{gramas de B usados} \end{array}$$

Definição 1.4. A **transposta** de uma matriz $A = (a_{ij})_{m\times n}$ é definida pela matriz $n \times m$

$$B = A^t$$

obtida trocando-se as linhas com as colunas, ou seja,

$$b_{ij} = a_{ji}\,,$$

para $i = 1, \ldots, n$ e $j = 1, \ldots, m$. Escrevemos também $[A^t]_{ij} = a_{ji}$.

Exemplo 1.7. As transpostas das matrizes

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

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

A seguir, mostraremos as propriedades que são válidas para a álgebra matricial. Várias propriedades são semelhantes àquelas que são válidas para os números reais, mas deve-se tomar cuidado com as diferenças. Uma propriedade importante que é válida para os números reais, mas não é válida para as matrizes é a comutatividade do produto, como foi mostrado no Exemplo 1.5. Por ser compacta, usaremos a notação de somatório na demonstração de várias propriedades. Algumas propriedades desta notação estão explicadas no Apêndice I na página 28.

1.1.2 Propriedades da Álgebra Matricial

Teorema 1.1. *Sejam A, B e C matrizes com tamanhos apropriados, α e β escalares. São válidas as seguintes propriedades para as operações matriciais:*

(a) *(comutatividade)* $A + B = B + A$;

(b) *(associatividade)* $A + (B + C) = (A + B) + C$;

(c) *(elemento neutro) A matriz $\bar{0}$, $m \times n$, definida por $[\bar{0}]_{ij} = 0$, para $i = 1, \ldots, m$, $j = 1, \ldots, n$ é tal que*

$$A + \bar{0} = A,$$

para toda matriz A, $m \times n$. A matriz $\bar{0}$ é chamada ***matriz nula*** *$m \times n$.*

(d) *(elemento simétrico) Para cada matriz A, existe uma única matriz $-A$, definida por $[-A]_{ij} = -a_{ij}$ tal que*

$$A + (-A) = \bar{0}.$$

(e) *(associatividade)* $\alpha(\beta A) = (\alpha\beta)A$;

(f) *(distributividade)* $(\alpha + \beta)A = \alpha A + \beta A$;

(g) *(distributividade)* $\alpha(A + B) = \alpha A + \alpha B$;

(h) *(associatividade)* $A(BC) = (AB)C$;

(i) *(elemento neutro) Para cada inteiro positivo p a matriz, $p \times p$,*

$$I_p = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix},$$

*chamada **matriz identidade** é tal que*

$$A\,I_n = I_m A = A, \quad \textit{para toda matriz } A = (a_{ij})_{m\times n}.$$

(j) *(distributividade)* $A(B+C) = AB + AC$ *e* $(B+C)A = BA + CA$;

(k) $\alpha(AB) = (\alpha A)B = A(\alpha B)$;

(l) $(A^t)^t = A$;

(m) $(A+B)^t = A^t + B^t$;

(n) $(\alpha A)^t = \alpha\, A^t$;

(o) $(AB)^t = B^t A^t$;

Demonstração. Para provar as igualdades acima, devemos mostrar que os elementos da matriz do lado esquerdo são iguais aos elementos correspondentes da matriz do lado direito. Serão usadas várias propriedades dos números sem citá-las explicitamente.

(a) $[A+B]_{ij} = a_{ij} + b_{ij} = b_{ij} + a_{ij} = [B+A]_{ij}$;

(b) $[A+(B+C)]_{ij} = a_{ij} + [B+C]_{ij} = a_{ij} + (b_{ij} + c_{ij}) = (a_{ij} + b_{ij}) + c_{ij} =$
$= [A+B]_{ij} + c_{ij} = [(A+B)+C]_{ij}$;

(c) Seja X uma matriz $m \times n$ tal que

$$A + X = A \tag{1.2}$$

para qualquer matriz A, $m \times n$. Comparando os elementos correspondentes, temos que

$$a_{ij} + x_{ij} = a_{ij}\,,$$

ou seja, $x_{ij} = 0$, para $i = 1\ldots, m$ e $j = 1\ldots, n$. Portanto, a única matriz que satisfaz (1.2) é a matriz em que todos os seus elementos são iguais a zero. Denotamos a matriz X por $\bar{0}$.

(d) Dada uma matriz A, $m \times n$, seja X uma matriz $m \times n$, tal que

$$A + X = \bar{0}\,. \tag{1.3}$$

Comparando os elementos correspondentes, temos que

$$a_{ij} + x_{ij} = 0\,,$$

ou seja, $x_{ij} = -a_{ij}$, para $i = 1\ldots, m$ e $j = 1\ldots, n$. Portanto, a única matriz que satisfaz (1.3) é a matriz em que todos os seus elementos são iguais aos simétricos dos elementos de A. Denotamos a matriz X por $-A$.

(e) $[\alpha(\beta A)]_{ij} = \alpha[\beta A]_{ij} = \alpha(\beta a_{ij}) = (\alpha\beta)a_{ij} = [(\alpha\beta)A]_{ij}$.

(f) $[(\alpha+\beta)A]_{ij} = (\alpha+\beta)a_{ij} = (\alpha a_{ij}) + (\beta a_{ij}) = [\alpha A]_{ij} + [\beta A]_{ij} = [\alpha A + \beta A]_{ij}$.

(g)
$$\begin{aligned} [\alpha(A+B)]_{ij} &= \alpha[A+B]_{ij} = \alpha(a_{ij}+b_{ij}) = \alpha a_{ij} + \alpha b_{ij} = [\alpha A]_{ij} + [\alpha B]_{ij} \\ &= [\alpha A + \alpha B]_{ij}. \end{aligned}$$

(h) A demonstração deste item é a mais trabalhosa. Sejam A, B e C matrizes $m \times p$, $p \times q$ e $q \times n$ respectivamente. A notação de somatório aqui pode ser muito útil, pelo fato de ser compacta.

$$\begin{aligned} [A(BC)]_{ij} &= \sum_{k=1}^{p} a_{ik}[BC]_{kj} = \sum_{k=1}^{p} a_{ik}(\sum_{l=1}^{q} b_{kl}c_{lj}) = \sum_{k=1}^{p}\sum_{l=1}^{q} a_{ik}(b_{kl}c_{lj}) = \\ &= \sum_{k=1}^{p}\sum_{l=1}^{q}(a_{ik}b_{kl})c_{lj} = \sum_{l=1}^{q}\sum_{k=1}^{p}(a_{ik}b_{kl})c_{lj} = \sum_{l=1}^{q}(\sum_{k=1}^{p} a_{ik}b_{kl})c_{lj} = \\ &= \sum_{l=1}^{q}[AB]_{il}c_{lj} = [(AB)C]_{ij}\,. \end{aligned}$$

(i) Podemos escrever a matriz identidade em termos do delta de Kronecker que é definido por

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

como $[I_n]_{ij} = \delta_{ij}$. Assim,

$$[AI_n]_{ij} = \sum_{k=1}^{n} a_{ik}[I_n]_{kj} = \sum_{k=1}^{n} a_{ik}\delta_{kj} = a_{ij}.$$

A outra igualdade é análoga.

(j)

$$\begin{aligned}[A(B+C)]_{ij} &= \sum_{k=1}^{p} a_{ik}[B+C]_{kj} = \sum_{k=1}^{p} a_{ik}(b_{kj}+c_{kj}) = \sum_{k=1}^{p}(a_{ik}b_{kj}+a_{ik}c_{kj}) = \\ &= \sum_{k=1}^{p} a_{ik}b_{kj} + \sum_{k=1}^{p} a_{ik}c_{kj} = [AB]_{ij}+[AC]_{ij} = [AB+AC]_{ij}\,.\end{aligned}$$

A outra igualdade é inteiramente análoga a anterior e deixamos como exercício.

(k) $[\alpha(AB)]_{ij} = \alpha\sum_{k=1}^{p} a_{ik}b_{kj} = \sum_{k=1}^{p}(\alpha a_{ik})b_{kj} = [(\alpha A)B]_{ij}$ e

$[\alpha(AB)]_{ij} = \alpha\sum_{k=1}^{p} a_{ik}b_{kj} = \sum_{k=1}^{p} a_{ik}(\alpha b_{kj}) = [A(\alpha B)]_{ij}$.

(l) $[(A^t)^t]_{ij} = [A^t]_{ji} = a_{ij}$.

(m) $[(A+B)^t]_{ij} = [A+B]_{ji} = a_{ji}+b_{ji} = [A^t]_{ij}+[B^t]_{ij}$.

(n) $[(\alpha A)^t]_{ij} = [\alpha A]_{ji} = \alpha a_{ji} = \alpha[A^t]_{ij} = [\alpha A^t]_{ij}$.

(o) $[(AB)^t]_{ij} = [AB]_{ji} = \sum_{k=1}^{p} a_{jk}b_{ki} = \sum_{k=1}^{p}[A^t]_{kj}[B^t]_{ik} = \sum_{k=1}^{p}[B^t]_{ik}[A^t]_{kj} = [B^tA^t]_{ij}$.

■

A **diferença** entre duas matrizes de mesmo tamanho A e B é definida por

$$A - B = A + (-B),$$

ou seja, é a soma da matriz A com a simétrica da matriz B.

Sejam A uma matriz $n\times n$ e p um inteiro positivo. Definimos a **potência** p de A, por $A^p = \underbrace{A\ldots A}_{p \text{ vezes}}$. E para $p = 0$, definimos $A^0 = I_n$.

Exemplo 1.8. Vamos verificar se para matrizes A e B, quadradas, vale a igualdade

$$(A+B)(A-B) = A^2 - B^2. \tag{1.4}$$

Usando a propriedade (i) do teorema anterior obtemos

$$\begin{aligned}(A+B)(A-B) &= (A+B)A + (A+B)(-B)\\ &= AA + BA - AB - BB = A^2 + BA - AB - B^2\end{aligned}$$

Assim, $(A+B)(A-B) = A^2 - B^2$ se, e somente se, $BA - AB = 0$, ou seja, se, e somente se, $AB = BA$. Como o produto de matrizes não é comutativo, a conclusão é que a igualdade (1.4), **não** vale para matrizes em geral. Como contra-exemplo basta tomarmos duas matrizes que não comutem entre si. Sejam

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Para estas matrizes

$$A+B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad A-B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A^2 = A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \quad B^2 = B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Assim,

$$(A+B)(A-B) = \begin{bmatrix} -1 & 0 \\ -2 & 1 \end{bmatrix} \neq \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = A^2 - B^2.$$

1.1.3 Aplicação: Cadeias de Markov

Exemplo 1.9. Vamos supor que uma população é dividida em três estados (por exemplo: ricos, classe média e pobres) e que em cada unidade de tempo a probabilidade de mudança de um estado para outro seja constante no tempo, só dependa dos estados. Este processo é chamado **cadeia de Markov**.

Seja t_{ij} a probabilidade de mudança do estado j para o estado i em uma unidade de tempo (geração). Cuidado com a ordem dos índices. A matriz

$$T = \begin{array}{cc} \begin{array}{ccc} 1 & 2 & 3 \end{array} & \\ \left[\begin{array}{ccc} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{array}\right] & \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \end{array}$$

é chamada **matriz de transição**. Como exemplo vamos considerar a matriz de transição

$$T = \begin{array}{cc} \begin{array}{ccc} 1 & 2 & 3 \end{array} & \\ \left[\begin{array}{ccc} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{array}\right] & \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \end{array} \tag{1.5}$$

A distribuição da população inicial entre os três estados pode ser descrita pela seguinte matriz:

$$P_0 = \left[\begin{array}{c} p_1 \\ p_2 \\ p_3 \end{array}\right] \begin{array}{l} \text{está no estado 1} \\ \text{está no estado 2} \\ \text{está no estado 3} \end{array}$$

A matriz P_0 caracteriza a distribuição inicial da população entre os três estados e é chamada **vetor de estado**. Por exemplo,

$$P_0 = \left[\begin{array}{c} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{array}\right] \begin{array}{l} \text{está no estado 1} \\ \text{está no estado 2} \\ \text{está no estado 3} \end{array} \tag{1.6}$$

representa uma população dividida de forma que $1/3$ da população está em cada estado. Após uma unidade de tempo a população estará dividida entre os três estados da seguinte forma

$$P_1 = \begin{bmatrix} t_{11}p_1 + t_{12}p_2 + t_{13}p_3 \\ t_{21}p_1 + t_{22}p_2 + t_{23}p_3 \\ t_{31}p_1 + t_{32}p_2 + t_{33}p_3 \end{bmatrix} \begin{array}{l} \text{estará no estado 1} \\ \text{estará no estado 2} \\ \text{estará no estado 3} \end{array}$$

Lembre-se que t_{ij} é a probabilidade de mudança do estado j para o estado i. Assim o vetor de estado após uma unidade de tempo é dada pelo produto de matrizes:

$$P_1 = TP_0.$$

Por exemplo se a matriz de transição, T, é a matriz dada por (1.5) e a matriz de estado inicial, P_0, é a matriz dada por (1.6), então após uma unidade de tempo a matriz de estado será dada por

$$P_1 = TP_0 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}$$

Como estamos assumindo que em cada unidade de tempo a matriz de transição é a mesma, então após k unidades de tempo a população estará dividida entre os três estados segundo a matriz de estado

$$P_k = TP_{k-1} = T^2P_{k-2} = \cdots = T^kP_0$$

Assim a matriz T^k dá a transição entre k unidades de tempo.

Veremos na Seção 7.1 na página 390 como calcular rapidamente potências k de matrizes e assim como determinar a distribuição da população após k unidades de tempo para k um inteiro positivo qualquer.

Exercícios Numéricos (respostas na página 461)

1.1.1. Considere as seguintes matrizes

$$A = \begin{bmatrix} 2 & 0 \\ 6 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 4 \\ 2 & -8 \end{bmatrix}, \quad C = \begin{bmatrix} -6 & 9 & -7 \\ 7 & -3 & -2 \end{bmatrix}$$

$$D = \begin{bmatrix} -6 & 4 & 0 \\ 1 & 1 & 4 \\ -6 & 0 & 6 \end{bmatrix}, \quad E = \begin{bmatrix} 6 & 9 & -9 \\ -1 & 0 & -4 \\ -6 & 0 & -1 \end{bmatrix}$$

Se for possível calcule:

(a) $AB - BA$,

(b) $2C - D$,

(c) $(2D^t - 3E^t)^t$,

(d) $D^2 - DE$.

1.1.2. Conhecendo-se somente os produtos AB e AC, como podemos calcular $A(B + C)$, $B^t A^t$, $C^t A^t$ e $(ABA)C$?

1.1.3. Considere as seguintes matrizes

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

$$C = \begin{bmatrix} -2 & 1 & -1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}$$

$$E_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Verifique que:

(a) AB é diferente de BA.

(b) AE_j é a j-ésima coluna de A, para $j = 1, 2, 3$ e $E_i^t B$ é a i-ésima linha de B, para $i = 1, 2, 3$ (o caso geral está no Exercício 1.1.16 na página 23).

(c) $CD = [\ d_1C_1\ d_2C_2\ d_3C_3\ ]$, em que $C_1 = \begin{bmatrix} -2 \\ 0 \\ -1 \end{bmatrix}$, $C_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ e $C_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, são as colunas de C (o caso geral está no Exercício 1.1.17 (a) na página 24).

(d) $DC = \begin{bmatrix} d_1C_1 \\ d_2C_2 \\ d_3C_3 \end{bmatrix}$, em que $C_1 = \begin{bmatrix} -2 & 1 & -1 \end{bmatrix}$, $C_2 = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$ e $C_3 = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$ são as linhas de C (o caso geral está no Exercício 1.1.17 (b) na página 24).

(e) Escrevendo B em termos das suas colunas, $B = [\ B_1\ B_2\ ]$, em que $B_1 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$ e $B_2 = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}$, o produto AB pode ser escrito como $AB = A [\ B_1\ B_2\ ] = [\ AB_1\ AB_2\ ]$ (o caso geral está no Exercício 1.1.18 (a) na página 25).

(f) escrevendo A em termos das suas linhas, $A_1 = \begin{bmatrix} -3 & 2 & 1 \end{bmatrix}$ e $A_2 = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix}$, o produto AB pode ser escrito como $AB = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} B = \begin{bmatrix} A_1B \\ A_2B \end{bmatrix}$ (o caso geral está no Exercício 1.1.18 (b) na página 25).

1.1.4. Sejam

$$A = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 4 & -2 \end{bmatrix} \quad \text{e} \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Verifique que $xA_1 + yA_2 + zA_3 = AX$, em que A_j é a j-ésima coluna de A, para $j = 1, 2, 3$ (o caso geral está no Exercício 1.1.19 na página 25).

1.1.5. Encontre um valor de x tal que $AB^t = 0$, em que

$$A = \begin{bmatrix} x & 4 & -2 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 2 & -3 & 5 \end{bmatrix}.$$

1.1.6. Mostre que as matrizes $A = \begin{bmatrix} 1 & \frac{1}{y} \\ y & 1 \end{bmatrix}$, em que y é uma número real não nulo, verificam a equação $X^2 = 2X$.

1.1.7. Mostre que se A e B são matrizes que comutam com a matriz $M = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, então $AB = BA$.

1.1.8. (a) Determine todas as matrizes A, 2×2, **diagonais** (os elementos que estão fora da diagonal são iguais a zero) que comutam com toda matriz B, 2×2, ou seja, tais que $AB = BA$, para toda matriz B, 2×2.

(b) Determine todas as matrizes A, 2×2, que comutam com toda matriz B, 2×2, ou seja, tais que $AB = BA$, para toda matriz B, 2×2.

1.1.9. Verifique que $A^3 = \bar{0}$, para

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

O caso geral está no Exercício 1.1.29 na página 27.

Exercícios usando o MATLAB®

Uma vez inicializado o MATLAB®, aparecerá na janela de comandos um prompt >> ou `EDU>>`. O prompt significa que o MATLAB® está esperando um comando. Todo comando deve ser finalizado teclando-se **Enter**. Comandos que foram dados anteriormente podem ser obtidos novamente usando as teclas ↑ e ↓. Enquanto se estiver escrevendo um comando, este pode ser corrigido usando as teclas ←, →, **Delete** e **Backspace**. O MATLAB® faz diferença entre letras maiúsculas e minúsculas.

No MATLAB®, pode-se obter ajuda sobre qualquer comando ou função. O comando
`>> help`
(sem o prompt >>) mostra uma listagem de todos os pacotes disponíveis. Ajuda sobre um pacote específico ou sobre um comando ou função específica pode ser obtida com o comando
`>> help nome`,
(sem a vírgula e sem o prompt >>) em que `nome` pode ser o nome de um pacote ou o nome de um comando ou função.

Além dos comandos e funções pré-definidas, escrevemos um pacote chamado `gaal` com funções específicas para a aprendizagem de Geometria Analítica e Álgebra Linear. Este pacote pode ser obtido gratuitamente através da internet no endereço `http://www.mat.ufmg.br/~regi`, assim como um texto com uma introdução ao MATLAB® e instruções de como instalar o pacote `gaal`. Depois deste pacote ser devidamente instalado, o comando `help gaal` no prompt do MATLAB® dá informações sobre este pacote.

Mais informações sobre as capacidades do MATLAB® podem ser obtidas em [3, 24].

Vamos descrever aqui alguns comandos que podem ser usados para a manipulação de matrizes. Outros comandos serão introduzidos a medida que forem necessários.

`>> syms x y z` diz ao MATLAB® que as variáveis `x y` e `z` são simbólicas.

`>> A=[a11,a12,...,a1n;a21,a22,...; ...,amn]` cria uma matriz, m por n, usando os elementos `a11, a12, ..., amn` e a armazena numa variável de nome `A`. Por exemplo, `>> A=[1,2,3;4,5,6]` cria a matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$;

`>> I=eye(n)` cria a matriz identidade n por n e a armazena numa variável `I`;

`>> O=zeros(n)` ou `>> O=zeros(m,n)` cria a matriz nula n por n ou m por n, respectivamente, e a armazena numa variável `O`;

`>> A+B` é a soma de `A` e `B`,
`>> A*B` é o produto de `A` por `B`,
`>> A.'` é a transposta de `A`,
`>> A-B` é a diferença `A` menos `B`,
`>> num*A` é o produto do escalar `num` por `A`,
`>> A^k` é a potência `A` elevado a k.

`>> A(:,j)` é a coluna j da matriz `A`, `>> A(i,:)` é a linha i da matriz `A`.

`>> diag([d1,...,dn])` cria uma matriz diagonal, cujos elementos da diagonal são iguais aos elementos da matriz `[d1,...,dn]`, ou seja, são `d1,...,dn`.

`>> A=sym(A)` converte a matriz `A` numa matriz em que os elementos são armazenados no formato simbólico. A função `numeric` faz o processo inverso.

`>> solve(expr)` determina a solução da equação `expr=0`. Por exemplo, `>> solve(x^2-4)` determina as soluções da equação $x^2 - 4 = 0$;

Comando do pacote GAAL:

`>> A=randi(n)` ou `>> A=randi(m,n)` cria uma matriz `n` por `n` ou `m` por `n`, respectivamente, com elementos inteiros aleatórios entre -5 e 5.

1.1.10. Use o MATLAB® para calcular alguns membros da sequência A, $A^2, \ldots, A^k, \ldots$, para

(a) $A = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{3} \end{bmatrix}$;

(b) $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} \\ 0 & -\frac{1}{5} \end{bmatrix}$.

A sequência parece estar convergindo para alguma matriz? Se estiver, para qual?

1.1.11. Calcule as potências das matrizes dadas a seguir e encontre experimentalmente (por tentativa!) o menor inteiro $k > 1$ tal que (use o comando `>> A=sym(A)` depois de armazenar a matriz na variável `A`):

(a) $A^k = I_3$, em que
$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix};$$

(b) $A^k = I_4$, em que
$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix};$$

(c) $A^k = \bar{0}$, em que

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

1.1.12. Vamos fazer um experimento no MATLAB® para tentar ter uma ideia do quão comum é encontrar matrizes cujo produto comuta. No prompt do MATLAB® digite a seguinte linha:

```
>> c=0; for n=1:1000,A=randi(3);B=randi(3);if(A*B==B*A),c=c+1;end,end,c
```

(não esqueça das vírgulas e pontos e vírgulas!). O que esta linha está mandando o MATLAB® fazer é o seguinte:

- Criar um contador `c` e atribuir a ele o valor zero.
- Atribuir às variáveis `A` e `B`, 1000 matrizes 3×3 com entradas inteiras e aleatórias entre -5 e 5.
- Se `AB=BA`, ou seja, `A` e `B` comutarem, então o contador `c` é acrescido de 1.
- No final o valor existente na variável `c` é escrito.

Qual a conclusão que você tira do valor obtido na variável `c`?

1.1.13. Faça um experimento semelhante ao anterior, mas para o caso em que cada uma das matrizes é **diagonal**, isto é, os elementos que estão fora da diagonal são iguais a zero. Use a seta para cima $\uparrow$ para obter novamente a linha digitada e edite a linha no prompt do MATLAB® de forma a obter algo semelhante à linha:

```
>> c=0; for n=1:1000,A=diag(randi(1,3));B=diag(randi(1,3));if( ....
```

Qual a conclusão que você tira do valor obtido na variável `c`?

1.1.14. Faça um experimento semelhante ao anterior, mas para o caso em que uma das matrizes é diagonal. Use a seta para cima $\uparrow$ para obter novamente a linha digitada e edite a linha no prompt do MATLAB® de forma a obter a seguinte linha:

```
>> c=0; for n=1:1000,A=diag(randi(1,3));B=randi(3);if(A*B==B*A),c=c+1;A,B,end,end,c
```

Aqui são impressas as matrizes `A` e `B` quando elas comutarem. Qual a conclusão que você tira deste experimento? Qual a probabilidade de um tal par de matrizes comutarem?

1.1.15. Use o MATLAB® para resolver os **Exercícios Numéricos**.

Exercícios Teóricos

1.1.16. Sejam $E_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, E_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, E_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$ matrizes $n \times 1$.

(a) Mostre que se

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

é uma matriz $m \times n$, então AE_j é igual à coluna j da matriz A.

(b) Mostre que se

$$B = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1m} \\ b_{21} & b_{22} & \ldots & b_{2m} \\ \vdots & & \ldots & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nm} \end{bmatrix},$$

é uma matriz $n \times m$ então $E_i^t B$ é igual à linha i da matriz B.

1.1.17. Seja

$$D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}$$

uma **matriz diagonal** $n \times n$, isto é, os elementos que estão fora da diagonal são iguais a zero. Seja

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

(a) Mostre que o produto AD é obtido da matriz A multiplicando-se cada coluna j por λ_j, ou seja, se $A = [\, A_1 \; A_2 \; \ldots \; A_n \,]$, em que $A_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix}$ é a coluna j de A, então

$$AD = [\, \lambda_1 A_1 \; \lambda_2 A_2 \; \ldots \; \lambda_n A_n \,].$$

(b) Mostre que o produto DA é obtido da matriz A multiplicando-se cada linha i por λ_i, ou seja, se $A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}$, em que $A_i = [\, a_{i1} \; \ldots \; a_{in} \,]$ é a linha i de A, então

$$DA = \begin{bmatrix} \lambda_1 A_1 \\ \lambda_2 A_2 \\ \vdots \\ \lambda_n A_n \end{bmatrix}.$$

1.1.18. Sejam A e B matrizes $m \times p$ e $p \times n$, respectivamente.

(a) Mostre que a j-ésima coluna do produto AB é igual ao produto AB_j, em que $B_j = \begin{bmatrix} b_{1j} \\ \vdots \\ b_{pj} \end{bmatrix}$ é a j-ésima coluna de B, ou seja, se $B = [\; B_1 \; \ldots \; B_n \;]$, então

$$AB = A[\; B_1 \; \ldots \; B_n \;] = [\; AB_1 \; \ldots \; AB_n \;];$$

(b) Mostre que a i-ésima linha do produto AB é igual ao produto A_iB, em que $A_i = [\; a_{i1} \; \ldots \; a_{ip} \;]$ é a i-ésima linha de A, ou seja, se $A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}$, então

$$AB = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} B = \begin{bmatrix} A_1B \\ A_2B \\ \vdots \\ A_mB \end{bmatrix}.$$

1.1.19. Seja A uma matriz $m \times n$ e $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ uma matriz $n \times 1$. Prove que $AX = \sum_{j=1}^{n} x_j A_j$, em que A_j é a j-ésima coluna de A. (Sugestão: Desenvolva o lado direito e chegue ao lado esquerdo.)

1.1.20. (a) Mostre que se A é uma matriz $m \times n$ tal que $AX = \bar{0}$, para toda matriz X, $n \times 1$, então $A = \bar{0}$. (Sugestão: use o Exercício 16 na página 23.)

(b) Sejam B e C matrizes $m \times n$, tais $BX = CX$, para todo X, $n \times 1$. Mostre que $B = C$. (Sugestão: use o item anterior.)

1.1.21. Mostre que a matriz identidade I_n é a única matriz tal que $A\,I_n = I_n A = A$ para qualquer matriz A, $n \times n$. (Sugestão: Seja J_n uma matriz tal que $A\,J_n = J_n\,A = A$. Mostre que $J_n = I_n$.)

1.1.22. Se $AB = BA$ e p é um inteiro positivo, mostre que $(AB)^p = A^p B^p$.

1.1.23. Sejam A, B e C matrizes $n \times n$.

(a) $(A + B)^2 = A^2 + 2AB + B^2$? E se $AB = BA$? Justifique.

(b) $(AB)C = C(AB)$? E se $AC = CA$ e $BC = CB$? Justifique.

(Sugestão: Veja o Exemplo 1.8 na página 14.)

1.1.24. (a) Se A e B são duas matrizes tais que $AB = \bar{0}$, então $A = \bar{0}$ ou $B = \bar{0}$? Justifique.

(b) Se $AB = \bar{0}$, então $BA = \bar{0}$? Justifique.

(c) Se A é uma matriz tal que $A^2 = \bar{0}$, então $A = \bar{0}$? Justifique.

1.1.25. Dizemos que uma matriz A, $n \times n$, é **simétrica** se $A^t = A$ e é **anti-simétrica** se $A^t = -A$.

(a) Mostre que se A é simétrica, então $a_{ij} = a_{ji}$, para $i, j = 1, \ldots n$ e que se A é anti-simétrica, então $a_{ij} = -a_{ji}$, para $i, j = 1, \ldots n$. Portanto, os elementos da diagonal principal de uma matriz anti-simétrica são iguais a zero.

(b) Mostre que se A e B são simétricas, então $A + B$ e αA são simétricas, para todo escalar α.

(c) Mostre que se A e B são simétricas, então AB é simétrica se, e somente se, $AB = BA$.

(d) Mostre que se A e B são anti-simétricas, então $A + B$ e αA são anti-simétricas, para todo escalar α.

(e) Mostre que para toda matriz A, $n \times n$, $A + A^t$ é simétrica e $A - A^t$ é anti-simétrica.

(f) Mostre que toda matriz quadrada A pode ser escrita como a soma de uma matriz simétrica e uma anti-simétrica. (Sugestão: Observe o resultado da soma de $A + A^t$ com $A - A^t$.)

1.1.26. Para matrizes quadradas $A = (a_{ij})_{n\times n}$ definimos o **traço** de A como sendo a soma dos elementos da diagonal (principal) de A, ou seja, $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$.

(a) Mostre que $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.

(b) Mostre que $\mathrm{tr}(\alpha A) = \alpha\mathrm{tr}(A)$.

(c) Mostre que $\mathrm{tr}(A^t) = \mathrm{tr}(A)$.

(d) Mostre que $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. (Sugestão: Prove inicialmente para matrizes 2×2.)

1.1.27. Seja A uma matriz $n\times n$. Mostre que se $AA^t = \bar{0}$, então $A = \bar{0}$. (Sugestão: use o traço.) E se a matriz A for $m\times n$, com $m \neq n$?

1.1.28. Já vimos que o produto de matrizes não é comutativo. Entretanto, certos conjuntos de matrizes são comutativos. Mostre que:

(a) Se D_1 e D_2 são matrizes diagonais $n\times n$, então $D_1D_2 = D_2D_1$.

(b) Se A é uma matriz $n\times n$ e

$$B = a_0 I_n + a_1 A + a_2 A^2 + \ldots + a_k A^k,$$

em que $a_0, \ldots, a_k$ são escalares, então $AB = BA$.

1.1.29. Uma matriz A é chamada **nilpotente** se $A^k = \bar{0}$, para algum inteiro positivo k. Verifique que a matriz

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

é nilpotente.

Apêndice I: Notação de Somatório

São válidas algumas propriedades para a notação de somatório:

(a) O índice do somatório é uma variável muda que pode ser substituída por qualquer letra:

$$\sum_{i=1}^{n} f_i = \sum_{j=1}^{n} f_j.$$

(b) O somatório de uma soma pode ser escrito como uma soma de dois somatórios:

$$\sum_{i=1}^{n}(f_i + g_i) = \sum_{i=1}^{n} f_i + \sum_{i=1}^{n} g_i.$$

Pois,

$$\sum_{i=1}^{n}(f_i + g_i) = (f_1 + g_1) + \ldots + (f_n + g_n) =$$

$$= (f_1 + \ldots + f_n) + (g_1 + \ldots + g_n) = \sum_{i=1}^{n} f_i + \sum_{i=1}^{n} g_i.$$

Aqui foram aplicadas as propriedades associativa e comutativa da soma de números.

(c) Se no termo geral do somatório aparece um produto, em que um fator não depende do índice do somatório, então este fator pode "sair" do somatório:

$$\sum_{i=1}^{n} f_i\, g_k = g_k \sum_{i=1}^{n} f_i.$$

Pois,

$$\sum_{i=1}^{n} f_i\, g_k = f_1 g_k + \ldots + f_n g_k = g_k(f_1 + \ldots + f_n) = g_k \sum_{i=1}^{n} f_i.$$

Aqui foram aplicadas as propriedades distributiva e comutativa do produto em relação a soma de números.

(d) Num somatório duplo, a ordem dos somatórios pode ser trocada:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} = \sum_{j=1}^{m} \sum_{i=1}^{n} f_{ij}.$$

Pois,

$$\begin{aligned}
\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} &= \sum_{i=1}^{n} (f_{i1} + \ldots + f_{im}) = \\
&= (f_{11} + \ldots + f_{1m}) + \ldots + (f_{n1} + \ldots + f_{nm}) = \\
&= (f_{11} + \ldots + f_{n1}) + \ldots + (f_{1m} + \ldots + f_{nm}) = \\
&= \sum_{j=1}^{m} (f_{1j} + \ldots + f_{nj}) = \sum_{j=1}^{m} \sum_{i=1}^{n} f_{ij}.
\end{aligned}$$

Aqui foram aplicadas as propriedades comutativa e associativa da soma de números.

1.2 Sistemas de Equações Lineares

Muitos problemas em várias áreas da Ciência recaem na solução de sistemas lineares. Vamos ver como a álgebra matricial pode simplificar o estudo dos sistemas lineares.

Uma **equação linear** em n variáveis $x_1, x_2, \ldots, x_n$ é uma equação da forma

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b\,,$$

em que $a_1, a_2, \ldots, a_n$ e b são constantes reais;

Um **sistema de equações lineares** ou simplesmente **sistema linear** é um conjunto de equações lineares, ou seja, é um conjunto de equações da forma

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & \vdots & = & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}\right.$$

em que a_{ij} e b_k são constantes reais, para $i, k = 1, \ldots, m$ e $j = 1, \ldots, n$.

Usando o produto de matrizes que definimos na seção anterior, o sistema linear acima pode ser escrito como uma equação matricial

$$\boxed{A\,X = B,}$$

em que

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

Uma **solução** de um sistema linear é uma matriz $S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$ tal que as equações do sistema são satisfeitas quando substituímos $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$. O conjunto de todas as soluções do sistema é chamado **conjunto solução** ou **solução geral** do sistema. A matriz A é chamada **matriz do sistema linear.**

Exemplo 1.10. O sistema linear de duas equações e duas incógnitas

$$\left\{ \begin{array}{rcrcl} x & + & 2y & = & 1 \\ 2x & + & y & = & 0 \end{array} \right.$$

pode ser escrito como

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

A solução (geral) do sistema acima é $x = -1/3$ e $y = 2/3$ (verifique!) ou

$$X = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}.$$

Uma forma de resolver um sistema linear é substituir o sistema inicial por outro que tenha o mesmo conjunto solução do primeiro, mas que seja mais fácil de resolver. O outro sistema é obtido depois de aplicar sucessivamente uma série de operações, que não alteram a solução do sistema, sobre as equações. As operações que são usadas são:

- Trocar a posição de duas equações do sistema;

- Multiplicar uma equação por um escalar diferente de zero;
- Somar a uma equação outra equação multiplicada por um escalar.

Estas operações são chamadas de **operações elementares**. Quando aplicamos operações elementares sobre as equações de um sistema linear somente os coeficientes do sistema são alterados, assim podemos aplicar as operações sobre a matriz de coeficientes do sistema, que chamamos de **matriz aumentada**, ou seja, a matriz

$$[A \mid B] = \left[\begin{array}{cccc:c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & & \ldots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right].$$

Definição 1.5. Uma **operação elementar sobre as linhas** de uma matriz é uma das seguintes operações:

(a) Trocar a posição de duas linhas da matriz;

(b) Multiplicar uma linha da matriz por um escalar diferente de zero;

(c) Somar a uma linha da matriz um múltiplo escalar de outra linha.

O próximo teorema garante que ao aplicarmos operações elementares às equações de um sistema o conjunto solução não é alterado.

Teorema 1.2. *Se dois sistemas lineares $AX = B$ e $CX = D$, são tais que a matriz aumentada $[C \mid D]$ é obtida de $[A \mid B]$ aplicando-se uma operação elementar, então os dois sistemas possuem as mesmas soluções.*

Demonstração. A demonstração deste teorema segue-se de duas observações:

(a) Se X é solução de um sistema, então X também é solução do sistema obtido aplicando-se uma operação elementar sobre suas equações (verifique!).

(b) Se o sistema $CX = D$, é obtido de $AX = B$ aplicando-se uma operação elementar às suas equações (ou equivalentemente às linhas da sua matriz aumentada), então o sistema $AX = B$ também pode ser obtido de $CX = D$ aplicando-se uma operação elementar às suas equações, pois cada operação elementar possui uma operação elementar inversa do mesmo tipo, que desfaz o que a anterior fez (verifique!).

Pela observação (b), $AX = B$ e $CX = D$ podem ser obtidos um do outro aplicando-se uma operação elementar sobre as suas equações. E pela observação (a), os dois possuem as mesmas soluções. ■

Dois sistemas que possuem o mesmo conjunto solução são chamados **sistemas equivalentes**. Portanto, segue-se do Teorema 1.2 que aplicando-se operações elementares às equações de um sistema linear obtemos sistemas equivalentes.

1.2.1 Método de Gauss-Jordan

O método que vamos usar para resolver sistemas lineares consiste na aplicação de operações elementares às linhas da matriz aumentada do sistema até que obtenhamos uma matriz numa forma em que o sistema associado a esta matriz seja de fácil resolução.

Vamos procurar obter uma matriz numa forma em que todas as linhas não nulas possuam como primeiro elemento não nulo (chamado **pivô**) o número 1 . Além disso, se uma coluna contém um pivô, então todos os seus outros elementos terão que ser iguais a zero. Vamos ver no exemplo seguinte como conseguimos isso. Neste exemplo veremos como a partir do faturamento e do gasto com insumos podemos determinar quanto foi produzido de cada produto manufaturado em uma indústria.

Exemplo 1.11. Uma indústria produz três produtos, X, Y e Z, utilizando dois tipos de insumo, A e B. Para a manufatura de cada kg de X são utilizados 1 grama do insumo A e 2 gramas do insumo B; para cada kg de Y, 1 grama de insumo A e 1 grama de insumo B e, para cada kg de Z, 1 grama de A e 4 gramas de B. O preço de venda do kg de cada um dos produtos X, Y e Z é R$ 2,00, R$ 3,00 e R$ 5,00, respectivamente. Com a venda de toda a produção de X, Y e Z manufaturada com 1 kg de A e 2 kg de B, essa indústria arrecadou R$ 2500,00. Vamos determinar quantos kg de cada um dos produtos X, Y e Z foram vendidos. Como vimos no Exemplo 1.6 na página 7, usando matrizes o esquema de produção pode ser descrito da seguinte forma:

$$\begin{array}{r c} & \begin{array}{ccc} \text{X} & \text{Y} & \text{Z} \end{array} \\ \begin{array}{r} \text{gramas de A/kg} \\ \text{gramas de B/kg} \\ \text{preço/kg} \end{array} & \left[\begin{array}{ccc} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 5 \end{array}\right] \end{array} = A \qquad X = \left[\begin{array}{c} x \\ y \\ z \end{array}\right] \begin{array}{l} \text{kg de X produzidos} \\ \text{kg de Y produzidos} \\ \text{kg de Z produzidos} \end{array}$$

$$AX = \left[\begin{array}{c} x+y+z \\ 2x+y+4z \\ 2x+3y+5z \end{array}\right] = \left[\begin{array}{c} 1000 \\ 2000 \\ 2500 \end{array}\right] \begin{array}{l} \text{gramas de A usados} \\ \text{gramas de B usados} \\ \text{arrecadação} \end{array}$$

Assim precisamos resolver o sistema linear

$$\left\{\begin{array}{rcrcrcl} x & + & y & + & z & = & 1000 \\ 2x & + & y & + & 4z & = & 2000 \\ 2x & + & 3y & + & 5z & = & 2500 \end{array}\right.$$

cuja matriz aumentada é

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1000 \\ 2 & 1 & 4 & 2000 \\ 2 & 3 & 5 & 2500 \end{array}\right]$$

1ª eliminação:

Vamos procurar para pivô da 1ª linha um elemento não nulo da primeira coluna não nula (se for o caso, podemos usar a troca de linhas para "trazê-lo" para a primeira linha). Como o primeiro elemento da primeira coluna é igual a 1 ele será o primeiro pivô. Agora, precisamos "zerar" os outros elementos da 1ª coluna, que é a coluna do pivô, para isto, adicionamos à 2ª linha, -2 vezes a 1ª linha e adicionamos à 3ª linha, também, -2 vezes a 1ª linha.

$$\boxed{\begin{array}{l} -2\times 1^{\underline{a}}\text{ linha} + 2^{\underline{a}}\text{ linha} \longrightarrow 2^{\underline{a}}\text{ linha} \\ -2\times 1^{\underline{a}}\text{ linha} + 3^{\underline{a}}\text{ linha} \longrightarrow 3^{\underline{a}}\text{ linha} \end{array}} \qquad \left[\begin{array}{ccc:c} 1 & 1 & 1 & 1000 \\ 0 & -1 & 2 & 0 \\ 0 & 1 & 3 & 500 \end{array}\right]$$

2ª eliminação:

Olhamos para a sub-matriz obtida eliminando-se a 1ª linha. Escolhemos para pivô um elemento diferente de zero na 1ª coluna não nula desta sub-matriz. Vamos escolher o elemento de posição 2,2. Como temos que "fazer" o pivô igual a um, vamos multiplicar a 2ª linha por -1.

$$\boxed{-1\times 2^{\underline{a}}\text{ linha} \longrightarrow 2^{\underline{a}}\text{ linha}} \qquad \left[\begin{array}{ccr:r} 1 & 1 & 1 & 1000 \\ 0 & 1 & -2 & 0 \\ 0 & 1 & 3 & 500 \end{array}\right]$$

Agora, precisamos "zerar" os outros elementos da 2ª coluna, que é a coluna do pivô, para isto, somamos à 1ª linha, -1 vezes a 2ª e somamos à 3ª linha, também, -1 vezes a 2ª.

$$\boxed{\begin{array}{l} -1\times 2^{\underline{a}}\text{ linha} + 1^{\underline{a}}\text{ linha} \longrightarrow 1^{\underline{a}}\text{ linha} \\ -1\times 2^{\underline{a}}\text{ linha} + 3^{\underline{a}}\text{ linha} \longrightarrow 3^{\underline{a}}\text{ linha} \end{array}} \qquad \left[\begin{array}{ccr:r} 1 & 0 & 3 & 1000 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 5 & 500 \end{array}\right]$$

3ª eliminação:

Olhamos para a sub-matriz obtida eliminando-se a 1ª e a 2ª linha. Escolhemos para pivô um elemento diferente de zero na 1ª coluna não nula desta sub-matriz. Temos de escolher o elemento de posição 3,3 e como temos de "fazer" o pivô igual a 1, vamos multiplicar a 3ª linha por 1/5.

$$\boxed{\tfrac{1}{5}\times 3^{\underline{a}} \text{ linha} \longrightarrow 3^{\underline{a}} \text{ linha}} \qquad \left[\begin{array}{ccc:c} 1 & 0 & 3 & 1000 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 100 \end{array}\right]$$

Agora, precisamos "zerar" os outros elementos da 3ª coluna, que é a coluna do pivô, para isto, somamos à 1ª linha, −3 vezes a 3ª e somamos à 2ª linha, 2 vezes a 2ª.

$$\boxed{\begin{array}{l} -3\times 3^{\underline{a}} \text{ linha} + 1^{\underline{a}} \text{ linha} \longrightarrow 1^{\underline{a}} \text{ linha} \\ \;\; 2\times 3^{\underline{a}} \text{ linha} + 2^{\underline{a}} \text{ linha} \longrightarrow 2^{\underline{a}} \text{ linha} \end{array}} \qquad \left[\begin{array}{ccc:c} 1 & 0 & 0 & 700 \\ 0 & 1 & 0 & 200 \\ 0 & 0 & 1 & 100 \end{array}\right]$$

Portanto o sistema dado é equivalente ao sistema

$$\left\{\begin{array}{ccccc} x & & & = & 700 \\ & y & & = & 200 \\ & & z & = & 100 \end{array}\right.$$

que possui solução geral dada por

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 700 \\ 200 \\ 100 \end{bmatrix}.$$

Portanto, foram vendidos 700 kg do produto X, 200 kg do produto Y e 100 kg do produto Z.

A última matriz que obtivemos no exemplo anterior está na forma que chamamos de **escalonada reduzida**.

Definição 1.6. Uma matriz $A = (a_{ij})_{m\times n}$ está na forma **escalonada reduzida** quando satisfaz as seguintes condições:

(a) Todas as linhas nulas (formadas inteiramente por zeros) ocorrem abaixo das linhas não nulas;

(b) O **pivô** (1º elemento não nulo de uma linha) de cada linha não nula é igual a 1;

(c) O pivô de cada linha não nula ocorre à direita do pivô da linha anterior.

(d) Se uma coluna contém um pivô, então todos os seus outros elementos são iguais a zero.

Se uma matriz satisfaz as propriedades (a) e (c), mas não necessariamente (b) e (d), dizemos que ela está na forma **escalonada**.

Exemplo 1.12. As matrizes

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

são escalonadas reduzidas, enquanto

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 1 & 3 & -1 & 5 \\ 0 & 0 & -5 & 15 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

são escalonadas, mas **não** são escalonadas reduzidas.

Este método de resolução de sistemas, que consiste em aplicar operações elementares às linhas da matriz aumentada até que a matriz do sistema esteja na forma escalonada reduzida, é conhecido como **método de Gauss-Jordan**.

Exemplo 1.13. Considere o seguinte sistema

$$\left\{\begin{array}{rcrcrcr} x & + & 3y & + & 13z & = & 9 \\ & & y & + & 5z & = & 2 \\ & & -2y & - & 10z & = & -8 \end{array}\right.$$

A sua matriz aumentada é

$$\left[\begin{array}{rrr|r} 1 & 3 & 13 & 9 \\ 0 & 1 & 5 & 2 \\ 0 & -2 & -10 & -8 \end{array}\right]$$

1ª eliminação:

Como o pivô da 1ª linha é igual a 1 e os outros elementos da 1ª coluna são iguais a zero, não há nada o que fazer na 1ª eliminação.

$$\left[\begin{array}{rrr|r} 1 & 3 & 13 & 9 \\ 0 & 1 & 5 & 2 \\ 0 & -2 & -10 & -8 \end{array}\right]$$

2ª eliminação:

Olhamos para submatriz obtida eliminando-se a 1ª linha. Escolhemos para pivô um elemento não nulo da 1ª coluna não nula da submatriz. Escolhemos o elemento de posição 2,2. Como ele é igual a 1, precisamos, agora, "zerar" os outros elementos da coluna do pivô. Para isto somamos à 1ª linha, -3 vezes a 2ª e somamos à 3ª linha, 2 vezes a 2ª.

$$\boxed{\begin{array}{r} -3\times 2^{\underline{a}} \text{ linha} + 1^{\underline{a}} \text{ linha} \longrightarrow 1^{\underline{a}} \text{ linha} \\ 2\times 2^{\underline{a}} \text{ linha} + 3^{\underline{a}} \text{ linha} \longrightarrow 3^{\underline{a}} \text{ linha} \end{array}} \qquad \left[\begin{array}{ccc:c} 1 & 0 & -2 & 3 \\ 0 & 1 & 5 & 2 \\ 0 & 0 & 0 & -4 \end{array}\right]$$

Portanto o sistema dado é equivalente ao sistema

$$\left\{\begin{array}{rcrcrcr} x & & & - & 2z & = & 3 \\ & & y & + & 5z & = & 2 \\ & & & & 0 & = & -4 \end{array}\right.$$

que **não** possui solução.

Em geral, um sistema linear não tem solução se, e somente se, a última linha não nula da forma escalonada reduzida da sua matriz aumentada for da forma $[\,0\ \ldots\ 0\,|\,b'_m\,]$, com $b'_m \neq 0$.

Exemplo 1.14. Considere o seguinte sistema

$$\left\{\begin{array}{rcrcrcrcr} & & & & 3z & - & 9w & = & 6 \\ 5x & + & 15y & - & 10z & + & 40w & = & -45 \\ x & + & 3y & - & z & + & 5w & = & -7 \end{array}\right.$$

A sua matriz aumentada é

$$\left[\begin{array}{cccc:c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right]$$

1ª eliminação:
Como temos que "fazer" o pivô igual a um, escolhemos para pivô o elemento de posição 3,1. Precisamos "colocá-lo" na primeira linha, para isto, trocamos a 3ª linha com a 1ª .

$$\boxed{1^{\underline{a}} \text{ linha} \longleftrightarrow 4^{\underline{a}} \text{ linha}} \qquad \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 5 & 15 & -10 & 40 & -45 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right]$$

Agora, precisamos "zerar" os outros elementos da 1ª coluna, que é a coluna do pivô, para isto, adicionamos à 2ª linha, −5 vezes a 1ª.

$$\boxed{-5\times 1^{\underline{a}} \text{ linha} + 2^{\underline{a}} \text{ linha} \longrightarrow 2^{\underline{a}} \text{ linha}} \qquad \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -5 & 15 & -10 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right]$$

2ª eliminação:
Olhamos para a sub-matriz obtida eliminando-se a 1ª linha. Escolhemos para pivô um elemento diferente de zero na 1ª coluna não nula desta sub-matriz. Escolhemos o elemento de posição 2,3. Como temos que fazer o pivô igual a 1, multiplicamos a 2ª linha por $-1/5$.

$$\boxed{-(1/5)\times 2^{\underline{a}} \text{ linha} \longrightarrow 2^{\underline{a}} \text{ linha}} \qquad \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & 1 & -3 & 2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right]$$

Agora, precisamos "zerar" os outros elementos da 2ª coluna, que é a coluna do pivô, para isto, adicionamos à 1ª linha a 2ª e à 3ª linha, −3 vezes a 2ª.

$$\boxed{\begin{array}{c} 2^{\underline{a}} \text{ linha} + 1^{\underline{a}} \text{ linha} \longrightarrow 1^{\underline{a}} \text{ linha} \\ -3\times 2^{\underline{a}} \text{ linha} + 3^{\underline{a}} \text{ linha} \longrightarrow 3^{\underline{a}} \text{ linha} \end{array}} \qquad \left[\begin{array}{cccc|c} 1 & 3 & 0 & 2 & -5 \\ 0 & 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Esta matriz é escalonada reduzida. Portanto o sistema dado é equivalente ao sistema seguinte

$$\left\{ \begin{array}{rcrcrcrcr} x & + & 3y & & & + & 2w & = & -5 \\ & & & & z & - & 3w & = & 2. \end{array} \right.$$

A matriz deste sistema possui duas colunas sem pivôs. As variáveis que não estão associadas a pivôs podem ser consideradas **variáveis livres**, isto é, podem assumir valores arbitrários. Neste exemplo as variáveis y e w não estão associadas a pivôs e podem ser consideradas variáveis livres. Sejam $w = \alpha$ e $y = \beta$. As variáveis associadas aos pivôs terão os seus valores dependentes das variáveis livres,

$$z = 2 + 3\alpha, \quad x = -5 - 2\alpha - 3\beta.$$

Assim, a solução geral do sistema é

$$X = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -5 - 2\alpha - 3\beta \\ \beta \\ 2 + 3\alpha \\ \alpha \end{bmatrix} \quad \text{para todos os valores de } \alpha \text{ e } \beta \text{ reais.}$$

Em geral, se o sistema linear tiver solução e a forma escalonada reduzida da matriz aumentada possuir colunas sem pivôs, as variáveis que **não** estão associadas a pivôs podem ser consideradas **variáveis livres**, isto é, podem assumir valores arbitrários. As variáveis associadas aos pivôs terão os seus valores dependentes das variáveis livres.

Lembramos que o sistema linear não tem solução se a última linha não nula da forma escalonada reduzida da matriz aumentada do sistema for da forma $[\,0\ \ldots\ 0\,|\,b'_m\,]$, com $b'_m \neq 0$, como no Exemplo 1.13 na página 39.

Observação. Para se encontrar a solução de um sistema linear não é necessário transformar a matriz aumentada do sistema na sua forma escalonada reduzida, mas se a matriz está nesta forma, o sistema associado é o mais simples possível. Um outro método de resolver sistemas lineares consiste em, através da aplicação de operações elementares à matriz aumentada do sistema, se chegar a uma matriz que é somente **escalonada** (isto é, uma matriz que satisfaz as condições (a) e (c), mas não necessariamente (b) e (d) da Definição 1.6). Este método é conhecido como **método de Gauss**.

O próximo resultado mostra que um sistema linear que tenha mais de uma solução não pode ter um número finito de soluções.

Proposição 1.3. *Sejam A uma matriz $m \times n$ e B uma matriz $m \times 1$. Se o sistema linear $A\,X = B$ possui duas soluções distintas $X_0 \neq X_1$, então ele tem infinitas soluções.*

Demonstração. Seja

$$X_\lambda = (1-\lambda)X_0 + \lambda X_1, \quad \text{para } \lambda \in \mathbb{R}.$$

Vamos mostrar que X_λ é solução do sistema $A\,X = B$, para qualquer $\lambda \in \mathbb{R}$. Para isto vamos mostrar que $A\,X_\lambda = B$.

Aplicando as propriedades (i), (j) das operações matriciais (Teorema 1.1 na página 9) obtemos

$$A\,X_\lambda = A[(1-\lambda)X_0 + \lambda X_1] = A(1-\lambda)X_0 + A\lambda X_1 = (1-\lambda)A\,X_0 + \lambda A\,X_1$$

Como X_0 e X_1 são soluções de $A\,X = B$, então $A\,X_0 = B$ e $A\,X_1 = B$, portanto

$$A\,X_\lambda = (1-\lambda)B + \lambda B = [(1-\lambda) + \lambda]B = B,$$

pela propriedade (f) do Teorema 1.1.

Assim o sistema $A\,X = B$ tem infinitas soluções, pois para todo valor de $\lambda \in \mathbb{R}$, X_λ é solução e $X_\lambda - X_{\lambda'} = (\lambda - \lambda')(X_1 - X_0)$, ou seja, $X_\lambda \neq X_{\lambda'}$, para $\lambda \neq \lambda'$. Observe que para $\lambda = 0$, $X_\lambda = X_0$, para $\lambda = 1$, $X_\lambda = X_1$, para $\lambda = 1/2$, $X_\lambda = \frac{1}{2}X_0 + \frac{1}{2}X_1$, para $\lambda = 3$, $X_\lambda = -2X_0 + 3X_1$ e para $\lambda = -2$, $X_\lambda = 3X_0 - 2X_1$.

No Exemplo 3.4 na página 164 temos uma interpretação geométrica desta demonstração. ■

Para resolver sistemas lineares vimos aplicando operações elementares à matriz aumentada do sistema linear. Isto pode ser feito com quaisquer matrizes.

1.2.2 Matrizes Equivalentes por Linhas

Definição 1.7. Uma matriz $A = (a_{ij})_{m\times n}$ é **equivalente por linhas** a uma matriz $B = (b_{ij})_{m\times n}$, se B pode ser obtida de A aplicando-se uma sequência de operações elementares sobre as suas linhas.

Exemplo 1.15. Observando os Exemplos 1.11, 1.14 e 1.13, vemos que as matrizes

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 5 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 3 & -9 \\ 5 & 15 & -10 & 40 \\ 1 & 3 & -1 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 & 13 \\ 0 & 1 & 5 \\ 0 & -2 & -10 \end{bmatrix}$$

são equivalentes por linhas às matrizes

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

respectivamente. Matrizes estas que são escalonadas reduzidas.

Cuidado: elas são equivalentes por linhas, **não** são iguais!

A relação "ser equivalente por linhas" satisfaz as seguintes propriedades, cuja verificação deixamos como exercício para o leitor:

- Toda matriz é equivalente por linhas a ela mesma (reflexividade);
- Se A é equivalente por linhas a B, então B é equivalente por linhas a A (simetria);
- Se A é equivalente por linhas a B e B é equivalente por linhas a C, então A é equivalente por linhas a C (transitividade).

Toda matriz é equivalente por linhas a uma matriz na forma escalonada reduzida e a demonstração, que omitiremos, pode ser feita da mesma maneira que fizemos no caso particular das matrizes aumentadas dos Exemplos 1.11, 1.14 e 1.13. No Teorema 4.10 na página 251 mostramos que essa matriz escalonada reduzida é a única matriz na forma escalonada reduzida equivalente a A.

Teorema 1.4. *Toda matriz $A = (a_{ij})_{m \times n}$ é equivalente por linhas a uma única matriz escalonada reduzida*

$$R = (r_{ij})_{m \times n}.$$

O próximo resultado será usado para provar alguns resultados no capítulo de inversão de matrizes.

Proposição 1.5. *Seja R uma matriz $n \times n$, na forma escalonada reduzida. Se $R \neq I_n$, então R tem uma linha nula.*

Demonstração. Observe que o pivô de uma linha i está sempre numa coluna j com $j \geq i$. Portanto, ou a última linha de R é nula ou o pivô da linha n está na posição n, n. Mas, neste caso todas as linhas anteriores são não nulas e os pivôs de cada linha i está na coluna i, ou seja, $R = I_n$. ■

1.2.3 Sistemas Lineares Homogêneos

Um sistema linear da forma

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & 0 \\ \vdots & & & & & & \vdots & = & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & 0 \end{array} \right. \tag{1.7}$$

é chamado **sistema homogêneo**. O sistema (1.7) pode ser escrito como $A\,X = \bar{0}$. Todo sistema homogêneo admite pelo menos a solução $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ chamada de **solução trivial**. Portanto, todo sistema homogêneo tem solução. Além disso ou tem somente a solução trivial ou tem infinitas soluções

Observação. Para resolver um sistema linear homogêneo $A\,X = \bar{0}$, basta escalonarmos a matriz A do sistema, já que sob a ação de uma operação elementar a coluna de zeros não é alterada. Mas, é preciso ficar atento quando se escreve o sistema linear associado à matriz resultante das operações elementares, para se levar em consideração esta coluna de zeros que não vimos escrevendo.

Teorema 1.6. *Se $A = (a_{ij})_{m \times n}$, é tal que $m < n$, então o sistema homogêneo $AX = \bar{0}$ tem solução diferente da solução trivial, ou seja, todo sistema homogêneo com menos equações do que incógnitas tem infinitas soluções.*

Demonstração. Como o sistema tem menos equações do que incógnitas ($m < n$), o número de linhas não nulas r da forma escalonada reduzida da matriz aumentada do sistema também é tal que $r < n$. Assim, temos r pivôs e $n - r$ variáveis (incógnitas) livres, que podem assumir todos os valores reais. Logo, o sistema admite solução não trivial e portanto infinitas soluções. ■

O conjunto solução de um sistema linear homogêneo satisfaz duas propriedades interessantes. Estas propriedades terão um papel decisivo no estudo de subespaços de $\mathbb{R}^n$ na Seção 4.1 na página 227.

Proposição 1.7. *Seja $A = (a_{ij})_{m \times n}$.*

(a) *Se X e Y são soluções do sistema homogêneo, $AX = \bar{0}$, então $X + Y$ também o é.*

(b) *Se X é solução do sistema homogêneo, $AX = \bar{0}$, então αX também o é.*

Demonstração. (a) Se X e Y são soluções do sistema homogêneo $AX = \bar{0}$, então $AX = \bar{0}$ e $AY = \bar{0}$ e portanto $X + Y$ também é solução, pois
$A(X + Y) = AX + AY = \bar{0} + \bar{0} = \bar{0}$;

(b) Se X é solução do sistema homogêneo $AX = \bar{0}$, então αX também o é, pois $A(\alpha X) = \alpha AX = \alpha\bar{0} = \bar{0}$.

■

Estas propriedades não são válidas para sistemas lineares em geral. Por exemplo, considere o sistema linear $A\,X = B$, em que $A = [1]$ e $B = [1]$. A solução deste sistema é $X = [1]$. Mas, $X + X = 2\,X = 2$, não é solução do sistema.

Exemplo 1.16. Vamos retomar a cadeia de Markov do Exemplo 1.9 na página 14. Vamos supor que uma população é dividida em três estados (por exemplo: ricos, classe média e pobres) e que em cada unidade de tempo a probabilidade de mudança de um estado para outro seja constante no tempo, só dependa dos estados.

Seja t_{ij} a probabilidade de mudança do estado j para o estado i em uma unidade de tempo (geração). A matriz de transição é dada por

$$T = \begin{array}{c c} \begin{array}{ccc} ① & ② & ③ \end{array} & \\ \left[\begin{array}{ccc} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{array}\right] & \begin{array}{c} ① \\ ② \\ ③ \end{array} \end{array}$$

Por exemplo, a matriz de transição pode ser dada por

$$
\begin{array}{cc}
 & \begin{array}{ccc} ① & ② & ③ \end{array} \\
T \;=\; & \left[\begin{array}{ccc} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{array}\right] \begin{array}{c} ① \\ ② \\ ③ \end{array}
\end{array}
$$

Vamos descobrir qual distribuição inicial da população entre os três estados permanece inalterada, geração após geração. Ou seja, vamos determinar P tal que

$$TP = P \quad \text{ou} \quad TP = I_3P \quad \text{ou} \quad (T - I_3)P = \bar{0}.$$

Assim precisamos resolver o sistema linear homogêneo

$$
\left\{\begin{array}{rcrcrcl}
-\frac{1}{2}x & + & \frac{1}{4}y & & & = & 0 \\
\frac{1}{2}x & - & \frac{1}{2}y & + & \frac{1}{2}z & = & 0 \\
 & & \frac{1}{4}y & - & \frac{1}{2}z & = & 0
\end{array}\right.
$$

cuja matriz aumentada é

$$
\left[\begin{array}{rrr:r}
-\frac{1}{2} & \frac{1}{4} & 0 & 0 \\
\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 \\
0 & \frac{1}{4} & -\frac{1}{2} & 0
\end{array}\right]
$$

1ª eliminação:

$$
\boxed{-2\times 1^{\underline{a}} \text{ linha} \longrightarrow 2^{\underline{a}} \text{ linha}} \qquad
\left[\begin{array}{rrr:r}
1 & -\frac{1}{2} & 0 & 0 \\
\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 \\
0 & \frac{1}{4} & -\frac{1}{2} & 0
\end{array}\right]
$$

$-\frac{1}{2}\times 1^{\underline{a}}$ linha $+\ 2^{\underline{a}}$ linha $\longrightarrow 2^{\underline{a}}$ linha

$$\left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & -\frac{1}{2} & 0 \end{array}\right]$$

2ª eliminação:

$-4\times 2^{\underline{a}}$ linha $\longrightarrow 2^{\underline{a}}$ linha

$$\left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & \frac{1}{4} & -\frac{1}{2} & 0 \end{array}\right]$$

$\frac{1}{2}\times 2^{\underline{a}}$ linha $+\ 1^{\underline{a}}$ linha $\longrightarrow 1^{\underline{a}}$ linha
$-\frac{1}{4}\times 2^{\underline{a}}$ linha $+\ 3^{\underline{a}}$ linha $\longrightarrow 3^{\underline{a}}$ linha

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Portanto o sistema dado é equivalente ao sistema seguinte

$$\left\{\begin{array}{rcrcrcl} x & & & - & z & = & 0 \\ & & y & - & 2z & = & 0 \end{array}\right.$$

Seja $z = \alpha$. Então $y = 2\alpha$ e $x = \alpha$. Assim, a solução geral do sistema é

$$X = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \text{para todo } \alpha \in \mathbb{R}.$$

Tomando a solução tal que $p_1 + p_2 + p_3 = 1$ obtemos que se a população inicial for distribuída de forma que $p_1 = 1/4$ da população esteja no estado 1, $p_2 = 1/2$ da população esteja no estado 2 e $p_3 = 1/4$, esteja no estado 3, então esta distribuição permanecerá constante geração após geração.

1.2.4 Matrizes Elementares (opcional)

Definição 1.8. Uma **matriz elementar** $n \times n$ é uma matriz obtida da matriz identidade I_n aplicando-se uma, e somente uma, operação elementar.

Vamos denotar por E_{ij} a matriz elementar obtida trocando-se a linha i com a linha j da matriz I_n, $E_i(\alpha)$ a matriz elementar obtida multiplicando-se a linha i da matriz I_n pelo escalar $\alpha \neq 0$ e $E_{i,j}(\alpha)$ a matriz elementar obtida da matriz I_n, somando-se à linha j, α vezes a linha i.

$$E_{i,j} = \left[\begin{array}{ccccccccc} 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \ddots & & & & & & & \cdot \\ \cdot & & 1 & & & & & & \cdot \\ \cdot & & & 0 & \ldots & 1 & & & \cdot \\ \cdot & & & \vdots & \ddots & \vdots & & & \cdot \\ \cdot & & & 1 & \ldots & 0 & & & \cdot \\ \cdot & & & & & & 1 & & \cdot \\ \cdot & & & & & & & \ddots & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{array}\right] \begin{array}{l} \\ \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \\ \end{array}, E_i(\alpha) = \left[\begin{array}{ccccccc} 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \ddots & & & & & \cdot \\ \cdot & & 1 & & & & \cdot \\ \cdot & & & \alpha & & & \cdot \\ \cdot & & & & 1 & & \cdot \\ \cdot & & & & & \ddots & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{array}\right] \begin{array}{l} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{array}$$

$$\text{e} \quad E_{i,j}(\alpha) = \left[\begin{array}{ccccccc} 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \ddots & & & & & \cdot \\ \cdot & & 1 & & & & \cdot \\ \cdot & & \vdots & \ddots & & & \cdot \\ \cdot & & \alpha & \ldots & 1 & & \cdot \\ \cdot & & & & & \ddots & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{array}\right] \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array}$$

Exemplo 1.17. As matrizes seguintes são as matrizes elementares 2×2:

$$E_{1,2} = E_{2,1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E_1(\alpha) = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}, E_2(\alpha) = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix}, \text{ com } \alpha \neq 0,$$

$$E_{1,2}(\alpha) = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \quad \text{e} \quad E_{2,1}(\alpha) = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}.$$

Sejam $E_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, E_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, E_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$ matrizes $m \times 1$.

As matrizes elementares podem ser escritas em termos das matrizes E_i como

$$E_{i,j} = \begin{bmatrix} E_1^t \\ \vdots \\ E_j^t \\ \vdots \\ E_i^t \\ \vdots \\ E_m^t \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array}, \quad E_i(\alpha) = \begin{bmatrix} E_1^t \\ \vdots \\ \alpha E_i^t \\ \vdots \\ E_m^t \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \\ \end{array} \quad \text{e} \quad E_{i,j}(\alpha) = \begin{bmatrix} E_1^t \\ \vdots \\ E_i^t \\ \vdots \\ E_j^t + \alpha E_i^t \\ \vdots \\ E_m^t \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array}$$

Aplicar uma operação elementar em uma matriz, corresponde a multiplicar a matriz à esquerda por uma matriz elementar, como mostra o resultado a seguir.

Teorema 1.8. *Sejam E uma matriz elementar $m \times m$ e A uma matriz qualquer $m \times n$. Então, EA é igual à matriz obtida aplicando-se na matriz A a mesma operação elementar que originou E.*

Demonstração. Como a i-ésima linha de um produto de matrizes BA é igual a B_iA, em que B_i é a i-ésima linha da matriz B (Exercício 1.1.18 (b) na página 25) e $E_i^tA = A_i$, em que A_i é a linha i da matriz A (Exercício 16 (b) na página 23), então:

$$E_{i,j}A = \begin{array}{c} \\ \\ i \rightarrow \\ \\ j \rightarrow \\ \\ \\ \end{array} \begin{bmatrix} E_1^t \\ \vdots \\ E_j^t \\ \vdots \\ E_i^t \\ \vdots \\ E_m^t \end{bmatrix} A = \begin{bmatrix} E_1^tA \\ \vdots \\ E_j^tA \\ \vdots \\ E_i^tA \\ \vdots \\ E_m^tA \end{bmatrix} \begin{array}{c} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array} = \begin{bmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \\ A_m \end{bmatrix} \begin{array}{c} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array}$$

$$E_i(\alpha)A = \begin{array}{c} \\ \\ i \rightarrow \\ \\ \\ \end{array} \begin{bmatrix} E_1^t \\ \vdots \\ \alpha E_i^t \\ \vdots \\ E_m^t \end{bmatrix} A = \begin{bmatrix} E_1^tA \\ \vdots \\ \alpha E_i^tA \\ \vdots \\ E_m^tA \end{bmatrix} \begin{array}{c} \\ \\ \leftarrow i \\ \\ \\ \end{array} = \begin{bmatrix} A_1 \\ \vdots \\ \alpha A_i \\ \vdots \\ A_m \end{bmatrix} \begin{array}{c} \\ \\ \leftarrow i \\ \\ \\ \end{array}$$

$$E_{i,j}(\alpha)A = \begin{array}{r} \\ \\ i\rightarrow \\ \\ j\rightarrow \\ \\ \\ \end{array} \begin{bmatrix} E_1^t \\ \vdots \\ E_i^t \\ \vdots \\ E_j^t+\alpha E_i^t \\ \vdots \\ E_m^t \end{bmatrix} A = \begin{bmatrix} E_1^t A \\ \vdots \\ E_i^t A \\ \vdots \\ E_j^t A+\alpha E_i^t A \\ \vdots \\ E_m^t A \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array} = \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j+\alpha A_i \\ \vdots \\ A_m \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array}$$

■

Assim, aplicar uma sequência de operações elementares em uma matriz, corresponde a multiplicar a matriz à esquerda por um produto de matrizes elementares.

Exemplo 1.18. Quando usamos o método de Gauss-Jordan para resolver o sistema do Exemplo 1.11 na página 34, aplicamos uma sequência de operações elementares na matriz aumentada do sistema. Isto corresponde a multiplicar a matriz aumentada

$$[\,A\,|\,B\,] = \left[\begin{array}{ccc:c} 1 & 1 & 1 & 1000 \\ 2 & 1 & 4 & 2000 \\ 2 & 3 & 5 & 2500 \end{array}\right]$$

à esquerda pelas matrizes elementares

$$E_{1,2}(-2) = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_{1,3}(-2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix},$$

$$E_2(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_{2,1}(-1) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_{2,3}(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

$$E_3(\tfrac{1}{5}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix}, \quad E_{3,1}(-3) = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_{3,2}(2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix},$$

ou seja,

$$E_{3,2}(2)\, E_{3,1}(-3)\, E_3(\tfrac{1}{5})\, E_{2,3}(-1)\, E_{2,1}(-1)\, E_2(-1)\, E_{1,3}(-2)\, E_{1,2}(-2)\, [\, A \mid B\,] = \left[\begin{array}{ccc:c} 1 & 0 & 0 & 700 \\ 0 & 1 & 0 & 200 \\ 0 & 0 & 1 & 100 \end{array}\right].$$

Exercícios Numéricos (respostas na página 470)

1.2.1. Quais das seguintes matrizes estão na forma escalonada reduzida:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & -4 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix},$$

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

1.2.2. Em cada item suponha que a matriz aumentada de um sistema foi transformada usando operações elementares na matriz escalonada reduzida dada. Resolva o sistema correspondente.

(a) $\begin{bmatrix} 1 & 0 & 0 & -7 & 8 \\ 0 & 1 & 0 & 3 & 2 \\ 0 & 0 & 1 & 1 & -5 \end{bmatrix}$;

(c) $\begin{bmatrix} 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{bmatrix}$;

(b) $\begin{bmatrix} 1 & -6 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 0 & 4 & 7 \\ 0 & 0 & 0 & 1 & 5 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$;

(d) $\begin{bmatrix} 1 & 7 & 0 & 0 & -8 & -3 \\ 0 & 0 & 1 & 0 & 6 & 5 \\ 0 & 0 & 0 & 1 & 3 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

1.2.3. Resolva, usando o método de Gauss-Jordan, os seguintes sistemas:

(a) $\left\{ \begin{array}{rcrcrcr} x_1 & + & x_2 & + & 2x_3 & = & 8 \\ -x_1 & - & 2x_2 & + & 3x_3 & = & 1 \\ 3x_1 & - & 7x_2 & + & 4x_3 & = & 10 \end{array} \right.$;

(b) $\left\{ \begin{array}{rcrcrcr} 2x_1 & + & 2x_2 & + & 2x_3 & = & 0 \\ -2x_1 & + & 5x_2 & + & 2x_3 & = & 1 \\ 8x_1 & + & x_2 & + & 4x_3 & = & -1 \end{array} \right.$;

(c) $\left\{\begin{array}{rcrcrcr} & - & 2x_2 & + & 3x_3 & = & 1 \\ 3x_1 & + & 6x_2 & - & 3x_3 & = & -2 \\ 6x_1 & + & 6x_2 & + & 3x_3 & = & 5 \end{array}\right.$.

1.2.4. Os sistemas lineares seguintes possuem a mesma matriz A. Resolva-os usando o método de Gauss-Jordan. Observe que os dois sistemas podem ser resolvidos ao mesmo tempo escalonando a matriz aumentada $[\,A \,|\, B_1 \,|\, B_2\,]$.

(a) $\left\{\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 1 \\ 2x_1 & - & 5x_2 & + & x_3 & = & -2 \\ 3x_1 & - & 7x_2 & + & 2x_3 & = & -1 \end{array}\right.$; (b) $\left\{\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 2 \\ 2x_1 & - & 5x_2 & + & x_3 & = & -1 \\ 3x_1 & - & 7x_2 & + & 2x_3 & = & 2 \end{array}\right.$.

1.2.5. Seja $A = \begin{bmatrix} 1 & 0 & 5 \\ 1 & 1 & 1 \\ 0 & 1 & -4 \end{bmatrix}$.

(a) Encontre a solução geral do sistema $(A + 4I_3)X = \bar{0}$;

(b) Encontre a solução geral do sistema $(A - 2I_3)X = \bar{0}$.

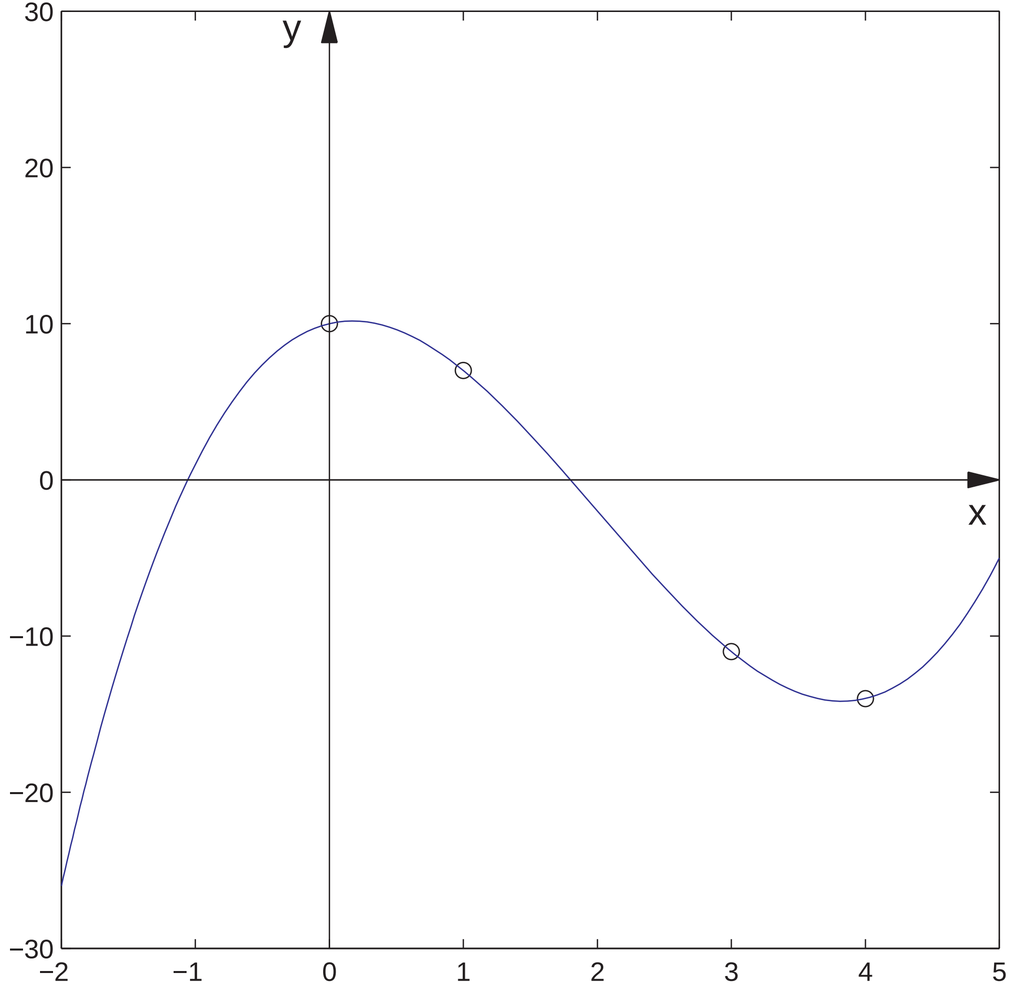
1.2.6. Para cada sistema linear dado, encontre todos os valores de a para os quais o sistema não tem solução, tem solução única e tem infinitas soluções:

(a) $\left\{\begin{array}{rcrcrcl} x & + & 2y & - & 3z & = & 4 \\ 3x & - & y & + & 5z & = & 2 \\ 4x & + & y & + & (a^2-14)z & = & a+2 \end{array}\right.$;

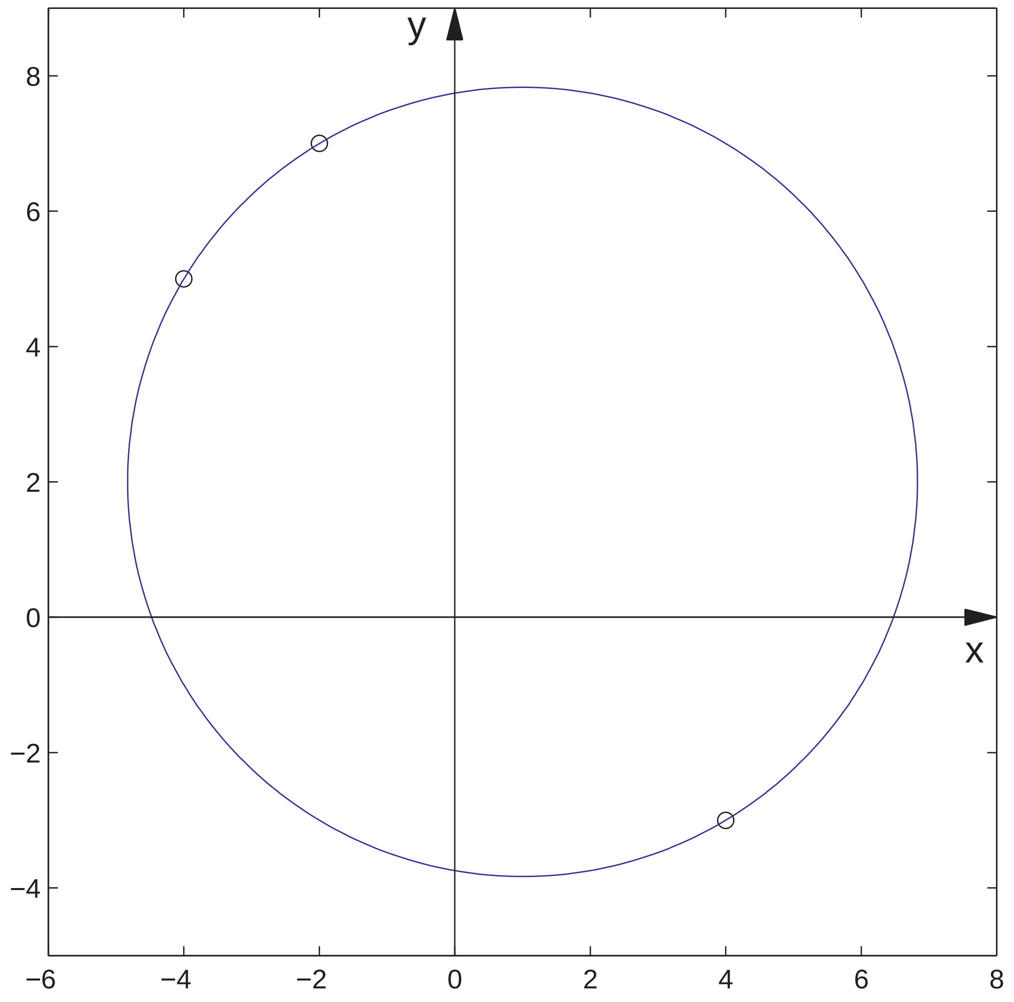
(b) $\left\{\begin{array}{rcrcrcl} x & + & y & + & z & = & 2 \\ 2x & + & 3y & + & 2z & = & 5 \\ 2x & + & 3y & + & (a^2-1)z & = & a+1 \end{array}\right.$.

1.2.7. Uma indústria produz três produtos, X, Y e Z, utilizando dois tipos de insumo, A e B. Para a manufatura de cada kg de X são utilizados 2 gramas do insumo A e 1 grama do insumo B; para cada kg de Y, 1 grama de insumo A e 3 gramas de insumo B e, para cada kg de Z, 3 gramas de A e 5 gramas de B. O preço de venda do kg de cada um dos produtos X, Y e Z é R$ 3,00, R$ 2,00 e R$ 4,00, respectivamente. Com a venda de toda a produção de X, Y e Z manufaturada com 1,9 kg de A e 2,4 kg de B, essa indústria arrecadou R$ 2900,00. Determine quantos kg de cada um dos produtos X, Y e Z foram vendidos. (Sugestão: veja o Exemplo 1.11 na página 34.)

1.2.8. Determine os coeficientes a, b, c e d da função polinomial $p(x) = ax^3 + bx^2 + cx + d$, cujo gráfico passa pelos pontos $P_1 = (0, 10), P_2 = (1, 7), P_3 = (3, -11)$ e $P_4 = (4, -14)$.


30
20
10
0
−10
−20
−30
−2
−1
0
1
2
3
4
5
y
x

1.2.9. Determine coeficientes a, b e c da equação do círculo, $x^2 + y^2 + ax + by + c = 0$, que passa pelos pontos $P_1 = (-2, 7), P_2 = (-4, 5)$ e $P_3 = (4, -3)$.


y
x
8
6
4
2
0
−2
−4
−6
−4
−2
0
2
4
6
8

1.2.10. Encontre condições sobre os b_i's para que cada um dos sistemas seja **consistente** (isto é, tenha solução):

(a) $\left\{\begin{array}{rcrcrcl} x_1 & - & 2x_2 & + & 5x_3 & = & b_1 \\ 4x_1 & - & 5x_2 & + & 8x_3 & = & b_2 \\ -3x_1 & + & 3x_2 & - & 3x_3 & = & b_3 \end{array}\right.$; (b) $\left\{\begin{array}{rcrcrcl} x_1 & - & 2x_2 & - & x_3 & = & b_1 \\ -4x_1 & + & 5x_2 & + & 2x_3 & = & b_2 \\ -4x_1 & + & 7x_2 & + & 4x_3 & = & b_3 \end{array}\right.$.

1.2.11. (Relativo à sub-seção 1.2.4) Considere a matriz

$$A = \begin{bmatrix} 0 & 1 & 7 & 8 \\ 1 & 3 & 3 & 8 \\ -2 & -5 & 1 & -8 \end{bmatrix}.$$

Encontre matrizes elementares E, F, G e H tais que $R = EFGHA$ é uma matriz escalonada reduzida. (Sugestão: veja o Exemplo 1.18 na página 55.)

1.2.12. Resolva, usando o método de Gauss-Jordan, os seguintes sistemas:

(a) $\left\{\begin{array}{rcrcrcrcrcrcl} x_1 & + & 2x_2 & & & - & 3x_4 & + & x_5 & & & = & 2 \\ x_1 & + & 2x_2 & + & x_3 & - & 3x_4 & + & x_5 & + & 2x_6 & = & 3 \\ x_1 & + & 2x_2 & & & - & 3x_4 & + & 2x_5 & + & x_6 & = & 4 \\ 3x_1 & + & 6x_2 & + & x_3 & - & 9x_4 & + & 4x_5 & + & 3x_6 & = & 9 \end{array}\right.$;

(b) $\left\{\begin{array}{rcrcrcrcrcrcr} x_1 & + & 3x_2 & - & 2x_3 & & & + & 2x_5 & & & = & 0 \\ 2x_1 & + & 6x_2 & - & 5x_3 & - & 2x_4 & + & 4x_5 & - & 3x_6 & = & -1 \\ & & & & 5x_3 & + & 10x_4 & & & + & 15x_6 & = & 5 \\ 2x_1 & + & 6x_2 & & & + & 8x_4 & + & 4x_5 & + & 18x_6 & = & 6 \end{array}\right.$;

1.2.13. Considere a matriz $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & -2 & a \\ 2 & 2\,a-2 & -a-2 & 3\,a-1 \\ 3 & a+2 & -3 & 2\,a+1 \end{bmatrix}$. Determine o conjunto solução do sistema $AX = B$, em que $B = [\ 4\ 3\ 1\ 6\ ]^t$, para todos os valores de a.

1.2.14. Resolva os sistemas lineares cujas matrizes aumentadas são:

(a) $\begin{bmatrix} 1 & 2 & 3 & 1 & 8 \\ 1 & 3 & 0 & 1 & 7 \\ 1 & 0 & 2 & 1 & 3 \end{bmatrix}$;

(b) $\begin{bmatrix} 1 & 1 & 3 & -3 & 0 \\ 0 & 2 & 1 & -3 & 3 \\ 1 & 0 & 2 & -1 & -1 \end{bmatrix}$;

(c) $\begin{bmatrix} 1 & 2 & 3 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 1 & 3 & 3 & 0 \end{bmatrix}$;

Exercícios usando o MATLAB®

Comandos do MATLAB®:

`>> A=[A1,...,An]` cria uma matriz `A` formada pelas matrizes, definidas anteriormente, `A1, ..., An` colocadas uma ao lado da outra;

`>> expr=subs(expr,x,num)` substitui na expressão `expr` a variável `x` por `num`.

`>> p=poly2sym([an,...,a0],x)` armazena na variável `p` o polinômio $a_n x^n + \ldots + a_0$.

`>> clf` limpa a figura ativa.

Comandos do pacote GAAL:

`>> B=opel(alpha,i,A)` ou `>> oe(alpha,i,A)`faz a operação elementar
`alpha×linha i ==> linha i` da matriz `A` e armazena a matriz resultante em `B`.

`>> B=opel(alpha,i,j,A)` ou `>> oe(alpha,i,j,A)` faz a operação elementar
`alpha×linha i + linha j ==> linha j` da matriz `A` e armazena em `B`.

`>> B=opel(A,i,j)` ou `>> oe(A,i,j)` faz a troca da linha i com a linha j da matriz `A` e armazena a matriz resultante em `B`.

`>> B=escalona(A)` calcula passo a passo a forma escalonada reduzida da matriz `A` e armazena a matriz resultante na variável `B`.

`>> matvand(P,k)` obtém a matriz de Vandermonde de ordem `k`, se `P=[x1;...;xn]` e a matriz de Vandermonde generalizada no caso em que `P=[x1,y1;...;xn,yn]`.

`>> po([x1,y1;x2,y2;...xk,yk])` desenha os pontos `(x1,y1),...,(xk,yk)`.

`>> plotf1(f,[a,b])` desenha o gráfico da função dada pela expressão simbólica `f` no intervalo `[a,b]`.

`>> plotci(f,[a,b],[c,d])` desenha o gráfico da curva dada implicitamente pela expressão `f(x,y)=0` na região do plano `[a,b]x[c,d]`.

`>> p=poly2sym2([a,b,c,d,e,f],x,y)` armazena na variável `p` o polinômio em duas variáveis $ax^2 + bxy + cy^2 + dx + ey + f$.

`>> eixos` desenha os eixos coordenados.

1.2.15. (a) Use o comando `P=randi(4,2)`, para gerar 4 pontos com entradas inteiras e aleatórias entre -5 e 5. Os pontos estão armazenados nas linhas da matriz `P`.

(b) Use o MATLAB® para *tentar* encontrar os coeficientes a, b, c e d da função polinomial $p(x) = ax^3 + bx^2 + cx + d$ cujo gráfico passa pelos pontos dados pelas linhas da matriz `P`. A matriz `A=matvand(P(:,1),3)` pode ser útil na solução deste problema, assim como a matriz `B=P(:,2)`. Se não conseguiu, repita o passo anterior. Por que pode não ser possível?

(c) Desenhe os pontos e o gráfico do polinômio com os comandos `clf`, `po(P)`, `syms x`, `p=poly2sym(R(:,5),x)`, `plotf1(p,[-5,5])`, em que `R` é forma escalonada reduzida da matriz `[A,B]`.

(d) Desenhe os eixos coordenados com o comando `eixos`.

1.2.16. (a) Use o comando `P=randi(5,2)`, para gerar 5 pontos com entradas inteiras e aleatórias entre -5 e 5. Os pontos estão armazenados nas linhas da matriz `P`.

(b) Use o MATLAB® para *tentar* encontrar os coeficientes a, b, c, d, e e f da cônica, curva de equação $ax^2 + bxy + cy^2 + dx + ey + f = 0$, cujo gráfico passa pelos pontos cujas coordenadas são dadas pelas linhas da matriz `P`. A matriz `A=matvand(P,2)` pode ser útil na solução deste problema. Se não conseguiu, repita o passo anterior. Por que pode não ser possível?

(c) Desenhe os pontos e a cônica com os comandos
`clf, po(P), syms x y, p=poly2sym2([-R(:,6);1],x,y), plotci(p,[-5,5],[-5,5])`, em que `R` é a forma escalonada reduzida da matriz `A`.

(d) Desenhe os eixos coordenados com o comando `eixos`.

1.2.17. Use o MATLAB® e resolva os **Exercícios Numéricos a partir do Exercício 1.2.3.**

Exercícios Teóricos

1.2.18. Mostre que toda operação elementar possui inversa, do mesmo tipo, ou seja, para cada operação elementar existe uma outra operação elementar do mesmo tipo que desfaz o que a operação anterior fez.

1.2.19. Prove que:

(a) Toda matriz é equivalente por linhas a ela mesma (reflexividade);

(b) Se A é equivalente por linhas a B, então B é equivalente por linhas a A (simetria);

(c) Se A é equivalente por linhas a B e B é equivalente por linhas a C, então A é equivalente por linhas a C (transitividade).

1.2.20. (a) Sejam X_1 e X_2 soluções do sistema homogêneo $A\,X = \bar{0}$. Mostre que $\alpha X_1 + \beta X_2$ é solução, para quaisquer escalares α e β. (Sugestão: veja o Exemplo 1.7.)

(b) Sejam X_1 e X_2 soluções do sistema $A\,X = B$. Mostre que se $\alpha X_1 + \beta X_2$ é solução, para quaisquer escalares α e β, então $B = \bar{0}$. (Sugestão: faça $\alpha = \beta = 0$.)

1.2.21. Sejam A uma matriz $m \times n$ e $B \neq \bar{0}$ uma matriz $m \times 1$.

(a) Mostre que se X_1 é uma solução do sistema $AX = B$ e Y_1 é uma solução do sistema homogêneo associado $AX = \bar{0}$, então $X_1 + Y_1$ é solução de $AX = B$.

(b) Seja X_0 solução particular do sistema $AX = B$. Mostre que toda solução X do sistema $AX = B$, pode ser escrita como $X = X_0 + Y$, em que Y é uma solução do sistema homogêneo associado, $AX = \bar{0}$. Assim, a solução geral do sistema $AX = B$ é a soma de uma solução particular de $AX = B$ com a solução geral do sistema homogêneo associado $AX = \bar{0}$. (Sugestão: Escreva $X = X_0 + (X - X_0)$ e mostre que $X - X_0$ é solução do sistema homogêneo $AX = \bar{0}$.)

Teste do Capítulo

1. Para o sistema linear dado, encontre todos os valores de a para os quais o sistema não tem solução, tem solução única e tem infinitas soluções:

$$\left\{\begin{array}{ccccccc} x & + & 2y & + & z & = & 3 \\ x & + & y & - & z & = & 2 \\ x & + & y & + & (a^2-5)z & = & a \end{array}\right.$$

2. Se possível, encontre os valores de x, y e z tais que:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}\begin{bmatrix} -40 & 16 & x \\ 13 & -5 & y \\ 5 & -2 & z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3. Sejam

$$D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad \text{e} \quad P = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

Sabendo-se que $A = P^t DP$, calcule D^2, PP^t e A^2.

4. Responda **Verdadeiro** ou **Falso**, justificando:

(a) Se $A^2 = -2A^4$, então $(I_n + A^2)(I_n - 2A^2) = I_n$;

(b) Se $A = P^t DP$, onde D é uma matriz diagonal, então $A^t = A$;

(c) Se D é uma matriz diagonal, então $DA = AD$, para toda matriz A, $n \times n$;

(d) Se $B = AA^t$, então $B = B^t$.

(e) Se B e A são tais que $A = A^t$ e $B = B^t$, então $C = AB$, é tal que $C^t = C$.

2

Inversão de Matrizes e Determinantes

2.1 Matriz Inversa

Todo número real a, não nulo, possui um inverso (multiplicativo), ou seja, existe um número b, tal que $a\,b = b\,a = 1$. Este número é único e o denotamos por a^{-1}. Apesar da álgebra matricial ser semelhante à álgebra dos números reais, nem todas as matrizes A *não nulas* possuem inversa, ou seja, nem sempre existe uma matriz B tal que $A\,B = B\,A = I_n$. De início, para que os produtos AB e BA estejam definidos e sejam iguais é preciso que as matrizes A e B sejam quadradas. Portanto, somente as matrizes quadradas podem ter inversa, o que já diferencia do caso dos números reais, pois todo número não nulo tem inverso. Mesmo entre as matrizes quadradas, muitas não possuem inversa, apesar do conjunto das que não tem inversa ser bem menor do que o conjunto das que tem (Exercício 2.2.9 na página 132).

Definição 2.1. Uma matriz quadrada $A = (a_{ij})_{n\times n}$ é **invertível** ou **não singular**, se existe uma matriz $B = (b_{ij})_{n\times n}$ tal que

$$A\,B = B\,A = I_n\,, \tag{2.1}$$

em que I_n é a matriz identidade. A matriz B é chamada de **inversa** de A. Se A não tem inversa, dizemos que A é **não invertível** ou **singular**.

Exemplo 2.1. Considere as matrizes

$$A = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} -1/2 & 1/6 \\ 0 & 1/3 \end{bmatrix}.$$

A matriz B é a inversa da matriz A, pois $A\,B = B\,A = I_2$.

Teorema 2.1. *Se uma matriz $A = (a_{ij})_{n\times n}$ possui inversa, então a inversa é única.*

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

$$AB = BA = I_n = AC = CA$$

e assim,

$$B = B\,I_n = B(AC) = (BA)C = I_nC = C\,.$$

■

Denotamos a inversa de A, quando ela existe, por A^{-1}. Devemos chamar atenção para o fato de que o índice superior -1, aqui, não significa uma potência, tão pouco uma divisão. Assim como no caso da transposta, em que A^t significa a transposta de A, aqui, A^{-1} significa a inversa de A.

2.1.1 Propriedades da Inversa

Teorema 2.2. *(a) Se A é invertível, então A^{-1} também o é e*

$$(A^{-1})^{-1} = A\,;$$

(b) Se $A = (a_{ij})_{n\times n}$ e $B = (b_{ij})_{n\times n}$ são matrizes invertíveis, então AB é invertível e

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

(c) Se $A = (a_{ij})_{n\times n}$ é invertível, então A^t também é invertível e

$$(A^t)^{-1} = (A^{-1})^t\,.$$

Demonstração. Se queremos mostrar que uma matriz é a inversa de uma outra, temos que mostrar que os produtos das duas matrizes são iguais à matriz identidade.

(a) Uma matriz B é a inversa de A^{-1} se

$$A^{-1}B = BA^{-1} = I_n\,.$$

Mas, como A^{-1} é a inversa de A, então

$$AA^{-1} = A^{-1}A = I_n\,.$$

Como a inversa é única, então $B = A$ é a inversa de A^{-1}, ou seja, $(A^{-1})^{-1} = A$.

(b) Temos que mostrar que a inversa de AB é $B^{-1}A^{-1}$, ou seja, mostrar que os produtos $(AB)(B^{-1}A^{-1})$ e $(B^{-1}A^{-1})(AB)$ são iguais à matriz identidade. Mas, pelas propriedades (h) e (i) do Teorema 1.1 na página 9:

$$\begin{aligned} (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} &= I_n, \\ (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B &= I_n. \end{aligned}$$

(c) Queremos mostrar que a inversa de A^t é $(A^{-1})^t$. Pela propriedade (o) do Teorema 1.1 na página 9:

$$\begin{aligned} A^t(A^{-1})^t = (A^{-1}A)^t = I_n^t &= I_n, \\ (A^{-1})^tA^t = (AA^{-1})^t = I_n^t &= I_n. \end{aligned}$$

■

O teorema seguinte, cuja demonstração será omitida no momento (Subseção 2.1.2), garante que basta verificarmos uma das duas igualdades em (2.1) para sabermos se uma matriz é a inversa de outra.

Teorema 2.3. *Sejam A e B matrizes $n \times n$.*

(a) *Se $BA = I_n$, então $AB = I_n$;*

(b) *Se $AB = I_n$, então $BA = I_n$;*

Assim, para verificar que uma matriz A é invertível, quando temos uma matriz B que é candidata a inversa de A, basta fazer um dos produtos AB ou BA e verificar se um deles é igual a I_n. O próximo exemplo ilustra este fato.

Exemplo 2.2. Seja $A = (a_{ij})_{n\times n}$ uma matriz tal que $A^3 = \bar{0}$ (A pode não ser a matriz nula!). Vamos mostrar que a inversa de $I_n - A$ é $I_n + A + A^2$. Para provar isto, devemos multiplicar a matriz $I_n - A$, pela matriz que possivelmente seja a inversa dela, aqui $I + A + A^2$, e verificar se o produto das duas é igual a matriz identidade I_n.

$$(I_n - A)(I_n + A + A^2) = I_n(I_n + A + A^2) - A(I_n + A + A^2) = I_n + A + A^2 - A - A^2 - A^3 = I_n.$$

Aqui foram usadas as propriedades (i) e (j) do Teorema 1.1 na página 9.

2.1.2 Matrizes Elementares e Inversão (opcional)

As matrizes elementares têm um papel importante no estudo da inversão de matrizes e da solução de sistemas lineares.

Proposição 2.4. *Toda matriz elementar é invertível e sua inversa é também uma matriz elementar. Usando a notação introduzida na página 52, temos:*

(a) $E_{i,j}^{-1} = E_{j,i} = E_{i,j}$;

(b) $E_i(\alpha)^{-1} = E_i(1/\alpha)$, *para* $\alpha \neq 0$;

(c) $E_{i,j}(\alpha)^{-1} = E_{i,j}(-\alpha)$.

Demonstração. Seja E uma matriz elementar. Esta matriz é obtida de I_n aplicando-se uma operação elementar. Seja F a matriz elementar correspondente a operação que transforma E de volta em I_n. Agora, pelo Teorema 1.8 na página 54, temos que $F\,E = E\,F = I_n$. Portanto, F é a inversa de E. ■

Teorema 2.5. *Seja A uma matriz $n \times n$. As seguintes afirmações são equivalentes:*

(a) *Existe uma matriz B, $n \times n$, tal que $BA = I_n$.*

(b) *A matriz A é equivalente por linhas à matriz identidade I_n.*

(c) *A matriz A é invertível.*

Demonstração. **(a)⇒(b)** Se $BA = I_n$, então o sistema $A\,X = \bar{0}$ tem somente a solução trivial, pois $X = I_n X = BAX = B\,\bar{0} = \bar{0}$. Isto implica que a matriz A é equivalente por linhas à matriz identidade I_n, pois caso contrário a forma escalonada reduzida de A teria uma linha nula (Proposição 1.5 na página 46).

(b)⇒(c) A matriz A ser equivalente por linhas à I_n significa, pelo Teorema 1.8 na página 54, que existem matrizes elementares $E_1, \ldots, E_k$, tais que

$$
\begin{aligned}
E_k \ldots E_1 A &= I_n && (2.2)\\
(E_1^{-1} \ldots E_k^{-1}) E_k \ldots E_1 A &= E_1^{-1} \ldots E_k^{-1} \\
A &= E_1^{-1} \ldots E_k^{-1}. && (2.3)
\end{aligned}
$$

Aqui, usamos o fato de que as matrizes elementares são invertíveis (Proposição 2.4). Portanto, A é invertível como o produto de matrizes invertíveis.

(c)⇒(a) Claramente.

■

Se A é invertível, então multiplicando-se ambos os membros de (2.2) à direita por A^{-1} obtemos

$$E_k \ldots E_1 I_n = A^{-1}.$$

Assim, a mesma sequência de operações elementares que transforma a matriz A na matriz identidade I_n transforma também I_n em A^{-1}.

A demonstração do Teorema 2.3 na página 74, agora, é uma simples consequência do Teorema anterior.

Demonstração do Teorema 2.3. (a) Vamos mostrar que se $BA = I_n$, então A é invertível e $B = A^{-1}$. Se $BA = I_n$, então pelo Teorema 2.5, A é invertível e $B = BI_n = BAA^{-1} = I_nA^{-1} = A^{-1}$. Logo, $AB = BA = I_n$.

(b) Se $AB = I_n$, então pelo item anterior B é invertível e $B^{-1} = A$. Portanto

$$BA = AB = I_n.$$

■

Segue da demonstração, do Teorema 2.5 (equação (2.3)) o resultado seguinte.

Teorema 2.6. *Uma matriz A é invertível se, e somente se, ela é um produto de matrizes elementares.*

Exemplo 2.3. Vamos escrever a matriz A do Exemplo 2.5 na página 82 como o produto de matrizes elementares. Quando encontramos a inversa da matriz A, aplicamos uma sequência de operações elementares em $[\,A\,|\,I_3\,]$ até que encontramos a matriz $[\,I_3\,|\,A^{-1}\,]$. Como as operações são por linha, esta mesma sequência de operações elementares transforma A em I_n. Isto corresponde a multiplicar a matriz $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 5 \end{bmatrix}$ **à esquerda** pelas matrizes elementares

$$E_{1,2}(-2) = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_{1,3}(-2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix},$$

$$E_2(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_{2,1}(-1) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_{2,3}(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

$$E_3(\tfrac{1}{5}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix}, \quad E_{3,1}(-3) = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_{3,2}(2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix},$$

ou seja,

$$E_{3,2}(2)\ E_{3,1}(-3)\ E_3(\tfrac{1}{5})\ E_{2,3}(-1)\ E_{2,1}(-1)\ E_2(-1)\ E_{1,3}(-2)\ E_{1,2}(-2)\ A = I_3.$$

Multiplicando à esquerda pelas inversas das matrizes elementares correspondentes obtemos

$$A = E_{1,2}(2)\ E_{1,3}(2)\ E_2(-1)\ E_{2,1}(1)\ E_{2,3}(1)\ E_3(5)\ E_{3,1}(3)\ E_{1,2}(-2).$$

2.1.3 Método para Inversão de Matrizes

O exemplo seguinte mostra, para matrizes 2×2, não somente uma forma de descobrir se uma matriz A tem inversa mas também, como encontrar a inversa, no caso em que ela exista. Ou seja, escalonamos a matriz $[A \mid I_2]$ e encontramos a sua forma escalonada reduzida $[R \mid S]$. Se $R = I_2$, então a matriz A é invertível e a inversa $A^{-1} = S$. Caso contrário, a matriz A não é invertível.

Exemplo 2.4. Seja $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Devemos procurar uma matriz $B = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ tal que $AB = I_2$, ou seja,

$$\left\{ \begin{array}{ccccccccc} ax & + & bz & & & & & = & 1 \\ cx & + & dz & & & & & = & 0 \\ & & & & ay & + & bw & = & 0 \\ & & & & cy & + & dw & = & 1 \end{array} \right.$$

Este sistema pode ser desacoplado em dois sistemas independentes que possuem a mesma matriz, que é a matriz A. Podemos resolvê-los simultaneamente. Para isto, basta escalonarmos a matriz aumentada

$$\left[\begin{array}{cc:c:c} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right] = [\, A \,|\, I_2 \,].$$

Os dois sistemas têm solução única se, e somente se, a forma escalonada reduzida da matriz $[\, A \,|\, I_2 \,]$ for da forma $[\, I_2 \,|\, S \,] = \left[\begin{array}{cc:c:c} 1 & 0 & s & t \\ 0 & 1 & u & v \end{array} \right]$ (verifique, observando o que acontece se a forma escalonada reduzida da matriz A não for igual a I_2). Neste caso, $x = s, z = u$ e $y = t, w = v$, ou seja, a matriz A possuirá inversa,

$$A^{-1} = B = S = \begin{bmatrix} s & t \\ u & v \end{bmatrix}.$$

Para os leitores da Subseção 2.1.2 o próximo teorema é uma simples consequência do Teorema 2.5 na página 76. Entretanto a demonstração que daremos a seguir fornece um método para encontrar a inversa de uma matriz, se ela existir.

Teorema 2.7. *Uma matriz A, $n \times n$, é invertível se, e somente se, A é equivalente por linhas à matriz identidade I_n.*

Demonstração. Pelo Teorema 2.3 na página 74, para verificarmos se uma matriz A, $n \times n$, é invertível, basta verificarmos se existe uma matriz B, tal que

$$A\,B = I_n\,. \tag{2.4}$$

Vamos denotar as colunas de B por $X_1, X_2, \ldots, X_n$, ou seja, $B = [\, X_1 \ \ldots \ X_n \,]$, em que

$$X_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix},\ X_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix},\ \ldots,\ X_n = \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{bmatrix}$$

e as colunas da matriz identidade I_n, por $E_1, E_2, \ldots, E_n$, ou seja, $I_n = [\, E_1 \ \ldots \ E_n \,]$, em que

$$E_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ E_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix},\ \ldots,\ E_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Assim a equação (2.4) pode ser escrita como

$$A\,[\,X_1\;\ldots\;X_n\,] = [\,AX_1\;\ldots\;AX_n\,] = [\,E_1\;\ldots\;E_n\,],$$

pois a j-ésima coluna do produto AB é igual a A vezes a j-ésima coluna da matriz B (Exercício 18 na página 25). Analisando coluna a coluna a equação anterior vemos que encontrar B é equivalente a resolver n sistemas lineares

$$A\,X_j = E_j \quad \text{para } j = 1\ldots,n.$$

Cada um dos sistemas pode ser resolvido usando o método de Gauss-Jordan. Para isso, formaríamos as matrizes aumentadas $[A \mid E_1]$, $[A \mid E_2], \ldots, [A \mid E_n]$. Entretanto, como as matrizes dos sistemas são todas iguais à A, podemos resolver todos os sistemas simultaneamente formando a matriz $n \times 2n$

$$[\,A \mid E_1\,E_2\ldots E_n\,] = [\,A \mid I_n\,].$$

Transformando $[\,A \mid I_n\,]$ na sua forma escalonada reduzida, que vamos denotar por $[\,R \mid S\,]$, vamos chegar a duas situações possíveis: ou a matriz R é a matriz identidade, ou não é.

- Se $R = I_n$, então a forma escalonada reduzida da matriz $[\,A \mid I_n\,]$ é da forma $[\,I_n \mid S\,]$. Se escrevemos a matriz S em termos das suas colunas $S = [\,S_1\,S_2\ldots S_n\,]$, então as soluções dos sistemas $A\,X_j = E_j$ são $X_j = S_j$ e assim $B = S$ é tal que $A\,B = I_n$ e pelo Teorema 2.3 na página 74 A é invertível.

- Se $R \neq I_n$, então a matriz A não é equivalente por linhas à matriz identidade I_n. Então, pela Proposição 1.5 na página 46 a matriz R tem uma linha nula. O que implica que cada um dos sistemas $A\,X_j = E_j$ ou não tem solução única ou não tem solução. Isto implica que a matriz A não tem inversa, pois as colunas da (única) inversa seriam X_j, para $j = 1,\ldots n$. ■

Observação. Da demonstração do Teorema 2.7 obtemos não somente uma forma de descobrir se uma matriz A tem inversa mas também, como encontrar a inversa, no caso em que ela exista. Ou seja, escalonamos a matriz $[A \mid I_n]$ e encontramos a sua forma escalonada reduzida $[R \mid S]$. Se $R = I_n$, então a matriz A é invertível e a inversa $A^{-1} = S$. Caso contrário, a matriz A não é invertível. Vejamos os exemplos seguintes.

Exemplo 2.5. Vamos encontrar, se existir, a inversa de

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 5 \end{bmatrix}$$

1ª eliminação:

$\boxed{\begin{array}{l} -2\times 1^{\underline{a}} \text{ linha} + 2^{\underline{a}} \text{ linha} \longrightarrow 2^{\underline{a}} \text{ linha} \\ -2\times 1^{\underline{a}} \text{ linha} + 3^{\underline{a}} \text{ linha} \longrightarrow 3^{\underline{a}} \text{ linha} \end{array}}$

$$\left[\begin{array}{rrr:rrr} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & -2 & 1 & 0 \\ 0 & 1 & 3 & -2 & 0 & 1 \end{array}\right]$$

2ª eliminação:

$\boxed{-1\times 2^{\underline{a}} \text{ linha} \longrightarrow 2^{\underline{a}} \text{ linha}}$

$$\left[\begin{array}{rrr:rrr} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 2 & -1 & 0 \\ 0 & 1 & 3 & -2 & 0 & 1 \end{array}\right]$$

$\boxed{\begin{array}{l} -1\times 2^{\underline{a}} \text{ linha} + 1^{\underline{a}} \text{ linha} \longrightarrow 1^{\underline{a}} \text{ linha} \\ -1\times 2^{\underline{a}} \text{ linha} + 3^{\underline{a}} \text{ linha} \longrightarrow 3^{\underline{a}} \text{ linha} \end{array}}$

$$\left[\begin{array}{rrr:rrr} 1 & 0 & 3 & -1 & 1 & 0 \\ 0 & 1 & -2 & 2 & -1 & 0 \\ 0 & 0 & 5 & -4 & 1 & 1 \end{array}\right]$$

3ª eliminação:

$$\boxed{\tfrac{1}{5}\times 3^{\underline{a}}\text{ linha} \longrightarrow 3^{\underline{a}}\text{ linha}} \qquad \left[\begin{array}{ccc:ccc} 1 & 0 & 3 & -1 & 1 & 0 \\ 0 & 1 & -2 & 2 & -1 & 0 \\ 0 & 0 & 1 & -\frac{4}{5} & \frac{1}{5} & \frac{1}{5} \end{array}\right]$$

$$\boxed{\begin{array}{r} -3\times 3^{\underline{a}}\text{ linha} + 1^{\underline{a}}\text{ linha} \longrightarrow 1^{\underline{a}}\text{ linha} \\ 2\times 3^{\underline{a}}\text{ linha} + 2^{\underline{a}}\text{ linha} \longrightarrow 2^{\underline{a}}\text{ linha} \end{array}} \qquad \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & \frac{7}{5} & \frac{2}{5} & -\frac{3}{5} \\ 0 & 1 & 0 & \frac{2}{5} & -\frac{3}{5} & \frac{2}{5} \\ 0 & 0 & 1 & -\frac{4}{5} & \frac{1}{5} & \frac{1}{5} \end{array}\right]$$

Assim, a matriz $[A \mid I_3]$ é equivalente por linhas à matriz acima, que é da forma $[I_3 \mid S]$, portanto a matriz A é invertível e a sua inversa é a matriz S, ou seja,

$$A^{-1} = \begin{bmatrix} \frac{7}{5} & \frac{2}{5} & -\frac{3}{5} \\ \frac{2}{5} & -\frac{3}{5} & \frac{2}{5} \\ -\frac{4}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix}.$$

Exemplo 2.6. Vamos determinar, se existir, a inversa da matriz

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

Para isso devemos escalonar a matriz aumentada

$$[A \mid I_3] = \left[\begin{array}{ccc:ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

1ª eliminação:

$\boxed{-1\times 1^{\underline{a}} \text{ linha} + 2^{\underline{a}} \text{ linha} \longrightarrow 2^{\underline{a}} \text{ linha}}$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

2ª eliminação:

$\boxed{-1\times 2^{\underline{a}} \text{ linha} \longrightarrow 2^{\underline{a}} \text{ linha}}$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$\boxed{\begin{array}{l} -2\times 2^{\underline{a}} \text{ linha} + 1^{\underline{a}} \text{ linha} \longrightarrow 1^{\underline{a}} \text{ linha} \\ -1\times 2^{\underline{a}} \text{ linha} + 3^{\underline{a}} \text{ linha} \longrightarrow 3^{\underline{a}} \text{ linha} \end{array}}$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & -1 & 2 & 0 \\ 0 & 1 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right]$$

Assim, a matriz $[A \mid I_3]$ é equivalente por linhas à matriz acima, que é da forma $[R \mid S]$, com $R \neq I_3$. Assim, a matriz A não é equivalente por linhas à matriz identidade e portanto **não** é invertível.

Se um sistema linear $A\,X = B$ tem o **número de equações igual ao número de incógnitas**, então o conhecimento da inversa da matriz do sistema A^{-1}, reduz o problema de resolver o sistema a simplesmente fazer um produto de matrizes, como está enunciado no próximo teorema.

Teorema 2.8. *Seja A uma matriz $n \times n$.*

(a) *O sistema associado* $AX = B$ *tem solução única se, e somente se,* A *é invertível. Neste caso a solução é* $X = A^{-1}B$*;*

(b) *O sistema homogêneo* $A\,X = \bar{0}$ *tem solução não trivial se, e somente se,* A *é singular (não invertível).*

Demonstração. (a) Se a matriz A é invertível, então multiplicando $A\,X = B$ por A^{-1} à esquerda em ambos os membros obtemos

$$\begin{aligned}
A^{-1}(A\,X) &= A^{-1}B \\
(A^{-1}A)X &= A^{-1}B \\
I_n X &= A^{-1}B \\
X &= A^{-1}B.
\end{aligned}$$

Aqui foram usadas as propriedades (h) e (i) do Teorema 1.1 na página 9. Portanto, $X = A^{-1}B$ é a única solução do sistema $A\,X = B$. Por outro lado, se o sistema $A\,X = B$ possui solução única, então a forma escalonada reduzida da matriz aumentada do sistema $[A \mid B]$ é da forma $[R \mid S]$, em que $R = I_n$. Pois a matriz A é quadrada e caso R fosse diferente da identidade possuiria uma linha de zeros (Proposição 1.5 na página 46) o que levaria a que o sistema $A\,X = B$ ou não tivesse solução ou tivesse infinitas soluções. Logo, a matriz A é equivalente por linhas à matriz identidade o que pelo Teorema 2.7 na página 80 implica que A é invertível.

(b) Todo sistema homogêneo possui pelo menos a solução trivial. Pelo item anterior, esta será a única solução se, e somente se, A é invertível. ■

Vamos ver no próximo exemplo que se conhecemos a inversa de uma matriz, então a produção de uma indústria em vários períodos pode ser obtida apenas multiplicando-se a inversa por matrizes colunas que contenham a arrecadação e as quantidades dos insumos utilizados em cada período.

Exemplo 2.7. Uma indústria produz três produtos, X, Y e Z, utilizando dois tipos de insumo, A e B. Para a manufatura de cada kg de X são utilizados 1 grama do insumo A e 2 gramas do insumo B; para cada kg de Y, 1 grama de insumo A e 1 grama de insumo B e, para cada kg de Z, 1 grama de A e 4 gramas de B. O preço de venda do kg de cada um dos produtos X, Y e Z é R$ 2,00, R$ 3,00 e R$ 5,00, respectivamente. Como vimos no Exemplo 1.6 na página 7, usando matrizes o esquema de produção pode ser descrito da seguinte forma:

$$\begin{array}{r c}
 & \begin{array}{ccc} \text{X} & \text{Y} & \text{Z} \end{array} \\
\begin{array}{r} \text{gramas de A/kg} \\ \text{gramas de B/kg} \\ \text{preço/kg} \end{array} & \left[\begin{array}{ccc} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 5 \end{array}\right]
\end{array} = A \qquad X = \left[\begin{array}{c} x \\ y \\ z \end{array}\right] \begin{array}{l} \text{kg de X produzidos} \\ \text{kg de Y produzidos} \\ \text{kg de Z produzidos} \end{array}$$

$$AX = \left[\begin{array}{c} x+y+z \\ 2x+y+4z \\ 2x+3y+5z \end{array}\right] \begin{array}{l} \text{gramas de A usados} \\ \text{gramas de B usados} \\ \text{arrecadação} \end{array}$$

No Exemplo 2.5 na página 82 determinamos a inversa da matriz

$$A = \left[\begin{array}{ccc} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 5 \end{array}\right]$$

que é

$$A^{-1} = \left[\begin{array}{rrr} \frac{7}{5} & \frac{2}{5} & -\frac{3}{5} \\ \frac{2}{5} & -\frac{3}{5} & \frac{2}{5} \\ -\frac{4}{5} & \frac{1}{5} & \frac{1}{5} \end{array}\right].$$

Sabendo-se a inversa da matriz A podemos saber a produção da indústria sempre que soubermos quanto foi gasto do insumo A, do insumo B e a arrecadação.

(a) Se em um período com a venda de toda a produção de X, Y e Z manufaturada com 1 kg de A e 2 kg de B, essa indústria arrecadou R$ $2500,00$, então para determinar quantos kg de cada um dos produtos X, Y e Z foram vendidos simplesmente multiplicamos A^{-1} pela matriz

$$B = \begin{bmatrix} 1000 \\ 2000 \\ 2500 \end{bmatrix} \begin{array}{l} \text{gramas de A usados} \\ \text{gramas de B usados} \\ \text{arrecadação} \end{array}$$

ou seja,

$$\begin{array}{l} \text{kg de X produzidos} \\ \text{kg de Y produzidos} \\ \text{kg de Z produzidos} \end{array} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = X = A^{-1}B = \begin{bmatrix} \frac{7}{5} & \frac{2}{5} & -\frac{3}{5} \\ \frac{2}{5} & -\frac{3}{5} & \frac{2}{5} \\ -\frac{4}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 1000 \\ 2000 \\ 2500 \end{bmatrix} = \begin{bmatrix} 700 \\ 200 \\ 100 \end{bmatrix}$$

Portanto, foram produzidos 700 kg do produto X, 200 kg de Y e 100 kg de Z.

(b) Se em outro período com a venda de toda a produção de X, Y e Z manufaturada com 1 kg de A e $2,1$ kg de B, essa indústria arrecadou R\$ $2900,00$, então para determinar quantos kg de cada um dos produtos X, Y e Z foram vendidos simplesmente multiplicamos A^{-1} pela matriz

$$B = \begin{bmatrix} 1000 \\ 2100 \\ 2900 \end{bmatrix} \begin{array}{l} \text{gramas de A usados} \\ \text{gramas de B usados} \\ \text{arrecadação} \end{array}$$

ou seja,

$$\begin{array}{l} \text{kg de X produzidos} \\ \text{kg de Y produzidos} \\ \text{kg de Z produzidos} \end{array} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = X = A^{-1}B = \begin{bmatrix} \frac{7}{5} & \frac{2}{5} & -\frac{3}{5} \\ \frac{2}{5} & -\frac{3}{5} & \frac{2}{5} \\ -\frac{4}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 1000 \\ 2100 \\ 2900 \end{bmatrix} = \begin{bmatrix} 500 \\ 300 \\ 200 \end{bmatrix}$$

Portanto, foram produzidos 500 kg do produto X, 300 kg de Y e 200 kg de Z.

Vamos mostrar a recíproca do item (b) do Teorema 2.2 na página 72. Este resultado será útil na demonstração de que o determinante do produto de matrizes é o produto dos determinantes (Subseção 2.2.2 na página 119).

Proposição 2.9. *Se A e B são matrizes $n \times n$, com AB invertível, então A e B são invertíveis.*

Demonstração. Considere o sistema $(AB)X = \bar{0}$. Se B **não** fosse invertível, então existiria $X \neq \bar{0}$, tal que $B\,X = \bar{0}$ (Teorema 2.8 na página 84). Multiplicando-se por A, teríamos $AB\,X = \bar{0}$, o que, novamente pelo Teorema 2.8 na página 84, contradiz o fato de AB ser invertível. Portanto, B é invertível. Agora, se B e AB são invertíveis, então A também é invertível, pois $A = (AB)B^{-1}$, que é o produto de duas matrizes invertíveis. ■

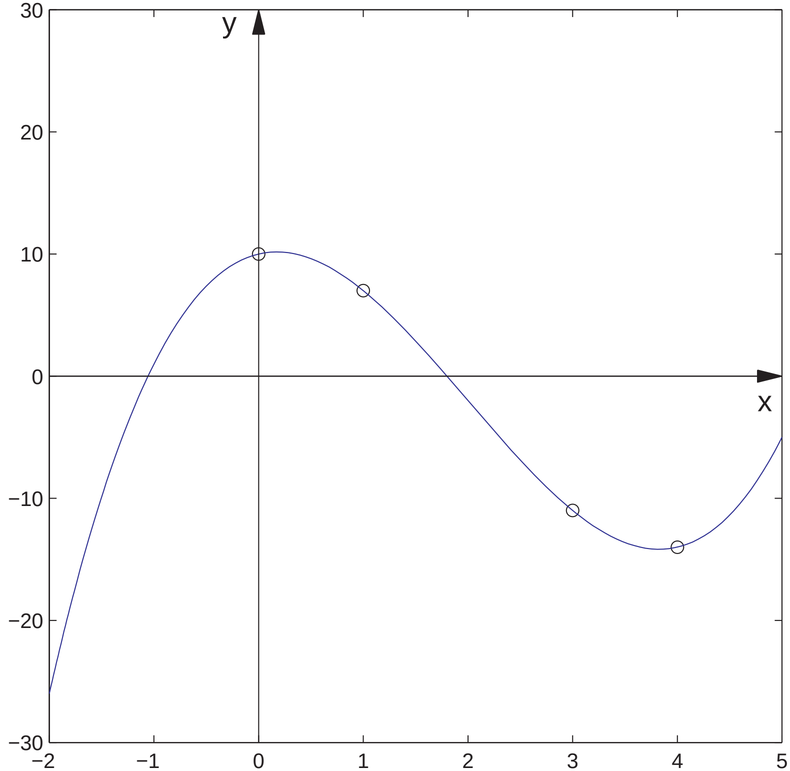
2.1.4 Aplicação: Interpolação Polinomial

Sejam $P_1 = (x_1, y_1), \ldots, P_n = (x_n, y_n)$, com $x_1, \ldots, x_n$ números distintos. Considere o problema de encontrar um polinômio de grau $n-1$

$$p(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0,$$

que *interpola* os dados, no sentido de que $p(x_i) = y_i$, para $i = 1, \ldots, n$.

Por exemplo se os pontos são $P_1 = (0, 10), P_2 = (1, 7), P_3 = (3, -11), P_4 = (4, -14)$ então o problema consiste em encontrar um polinômio de grau 3 que interpola os pontos dados (veja o Exercício 1.2.8 na página 59).



Vamos mostrar que existe, um e somente um, polinômio de grau no máximo igual a $n-1$, que interpola n pontos, com abscissas distintas. Substituindo os pontos no polinômio $p(x)$, obtemos um sistema linear $AX = B$, em que

$$X = \begin{bmatrix} a_{n-1} \\ a_{n-2} \\ \vdots \\ a_0 \end{bmatrix}, \quad B = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \text{e} \quad A = \begin{bmatrix} x_1^{n-1} & x_1^{n-2} & \ldots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \ldots & x_2 & 1 \\ \vdots & \vdots & & \vdots & \\ x_n^{n-1} & x_n^{n-2} & \ldots & x_n & 1 \end{bmatrix}.$$

A matriz A é chamada **matriz de Vandermonde**.

Vamos mostrar que $AX = B$ tem somente uma solução. Pelo Teorema 2.8 na página 84, um sistema de n equações e n incógnitas $AX = B$ tem solução única se, e somente se, o sistema homogêneo associado, $AX = \bar{0}$, tem somente a solução trivial. $X = [\, a_{n-1} \ \cdots \ a_0 \,]$ é solução do sistema homogêneo se, e somente se, o polinômio de grau $n-1$, $p(x) = a_{n-1}x^{n-1} + \cdots + a_0$, se anula em n pontos distintos. O que implica que o polinômio $p(x)$ é o polinômio com todos os seus coeficientes iguais a zero. Portanto, o sistema homogêneo $A\,X = \bar{0}$ tem somente a solução trivial. Isto prova que existe, um e somente um, polinômio de grau no máximo igual a $n-1$, que interpola n pontos, com abscissas distintas.

Assim a solução do sistema linear é $X = A^{-1}B$. Como a matriz A depende apenas das abscissas dos pontos, tendo calculado a matriz A^{-1} podemos determinar rapidamente os polinômios que interpolam vários conjuntos de pontos, desde que os pontos de todos os conjuntos tenham as mesmas abscissas dos pontos do conjunto inicial.

2.1.5 Aplicação: Criptografia

	a	b	c	d	e	f	g	h	i	j	k	l	m	n
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
o	p	q	r	s	t	u	v	w	x	y	z	à	á	â
15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
ã	ç	é	ê	í	ó	ô	õ	ú	ü	A	B	C	D	E
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
U	V	W	X	Y	Z	À	Á	Â	Ã	Ç	É	Ê	Í	Ó
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74
Ô	Õ	Ú	Ü	0	1	2	3	4	5	6	7	8	9	:
75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
;	<	=	>	?	@	!	"	#	$	%	&	'	(	)
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104
*	+	,	-	.	/	[	\	]	_	{	|	}		
105	106	107	108	109	110	111	112	113	114	115	116	117		

Tabela 2.1. Tabela de conversão de caracteres em números

Vamos transformar uma mensagem em uma matriz da seguinte forma. Vamos quebrar a mensagem em pedaços de tamanho 3 e cada pedaço será convertido em uma matriz coluna usando a Tabela 2.1 de conversão entre caracteres e números.

Considere a seguinte mensagem criptografada

$$\texttt{1ydobbr,?} \tag{2.5}$$

Quebrando a mensagem criptografada em pedaços de tamanho 3 e convertendo cada pedaço para uma coluna de números usando a Tabela 2.1 obtemos a matriz

$$Y = \begin{bmatrix} 80 & 15 & 18 \\ 25 & 2 & 107 \\ 4 & 2 & 94 \end{bmatrix}$$

Sabendo-se que esta mensagem foi criptografada fazendo o produto da mensagem inicial pela matriz

$$M = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

então

$$X = M^{-1}Y$$

será a mensagem inicial convertida para números, ou seja,

$$X = M^{-1}Y = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 80 & 15 & 18 \\ 25 & 2 & 107 \\ 4 & 2 & 94 \end{bmatrix} = \begin{bmatrix} 59 & 15 & 5 \\ 21 & 0 & 13 \\ 4 & 2 & 94 \end{bmatrix}$$

Convertendo para texto usando novamente a Tabela 2.1 obtemos que a mensagem que foi criptografada é

`Tudo bem?` (2.6)

Exercícios Numéricos (respostas na página 492)

2.1.1. Seja A uma matriz 3×3. Suponha que $X = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ é solução do sistema homogêneo $A\,X = \bar{0}$. A matriz A é singular ou não? Justifique.

2.1.2. Se possível, encontre as inversas das seguintes matrizes:

(a) $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$;

(b) $\begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 1 \\ 1 & 3 & 2 \end{bmatrix}$;

(c) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 3 & 3 & 2 \end{bmatrix}$;

(d) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix}$;

(e) $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$;

(f) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ 5 & 9 & 1 & 6 \end{bmatrix}$;

2.1.3. Encontre todos os valores de a para os quais a matriz $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & a \end{bmatrix}$ tem inversa.

2.1.4. Se

$$A^{-1} = \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix} \quad \text{e} \quad B^{-1} = \begin{bmatrix} 2 & 5 \\ 3 & -2 \end{bmatrix},$$

encontre $(A\,B)^{-1}$.

2.1.5. Resolva o sistema $A\,X = B$, se $A^{-1} = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$.

2.1.6. Seja
$$A = \begin{bmatrix} 1 & -1 \\ -4 & 1 \end{bmatrix}.$$
mostraremos no Exemplo 7.6 na página 406 que
$$P = \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$$
são tais que
$$A = PDP^{-1}.$$
Determine A^k, para $k = 1, 2, 3, \ldots$

2.1.7. (Relativo à Subseção 2.1.2) Encontre matrizes elementares $E_1, \ldots, E_k$ tais que $A = E_1 \ldots E_k$, para
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}.$$

Exercícios usando o MATLAB®

Comandos do MATLAB®:

`>> M=[A,B]` atribui à matriz `M` a matriz obtida colocando lado a lado as matrizes `A` e `B`.

`>> A=[A1,...,An]` cria uma matriz `A` formada pelas matrizes, definidas anteriormente, `A1, ..., An` colocadas uma ao lado da outra;

`>> M=A(:,k:l)` atribui à matriz `M` a submatriz da matriz `A` obtida da coluna `l` à coluna `k` da matriz `A`.

Comandos do pacote GAAL:

`>> B=opel(alpha,i,A)` ou `B=oe(alpha,i,A)` faz a operação elementar `alpha*linha i ==> linha i` da matriz `A` e armazena a matriz resultante em `B`.

`>> B=opel(alpha,i,j,A)` ou `B=oe(alpha,i,j,A)` faz a operação elementar `alpha*linha i + linha j ==> linha j` da matriz `A` e armazena a matriz resultante na variável `B`.

`>> B=opel(A,i,j)` ou `B=oe(A,i,j)` faz a troca da linha i com a linha j da matriz `A` e armazena a matriz resultante na variável `B`.

`>> B=escalona(A)` calcula passo a passo a forma escalonada reduzida da matriz `A` e armazena a matriz resultante na variável `B`.

2.1.8. O pacote `GAAL` contém alguns arquivos com mensagens criptografadas e uma chave para decifrá-las. Use os comandos a seguir para ler dos arquivos e atribuir às variáveis correspondentes, uma mensagem criptografada e a uma chave para decifrá-la.

```
>> menc=lerarq('c:/matlab/toolbox/gaal/menc1.txt')
>> key=lerarq('c:/matlab/toolbox/gaal/key.txt')
```

Com estes comandos foram lidos os arquivos `menc1.txt` e `key.txt` e atribuídos os resultados às variáveis `menc` e `key` respectivamente. Para converter a mensagem criptografada e a chave para matrizes numéricas use os comandos do pacote `gaal`:

```
>> y=char2num(menc), M=char2num(key)
```

Sabendo-se que a mensagem criptografada (convertida para números), `y`, foi originalmente obtida multiplicando-se a matriz `M` pela mensagem original (convertida para números), `x`, determine `x`. Descubra a mensagem usando o comando do pacote `gaal`, `num2char(x)`, que converte a matriz para texto. Decifre as mensagens que estão nos arquivos `menc2.txt` e `menc3.txt`. Como deve ser a matriz `M` para que ela possa ser uma matriz chave na criptografia?

2.1.9. Resolva os **Exercícios Numéricos a partir do Exercício 2.1.2** usando o MATLAB®.

Exercícios Teóricos

2.1.10. (a) Mostre que a matriz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ é invertível se, e somente se, $ad - bc \neq 0$ e neste caso a inversa é dada por

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

(Sugestão: encontre a forma escalonada reduzida da matriz $[\,A \mid I_2\,]$, para $a \neq 0$ e para $a = 0$.)

(b) Mostre que se $ad - bc \neq 0$, então o sistema linear

$$\left\{\begin{array}{ccccc} ax & + & by & = & g \\ cx & + & dy & = & h \end{array}\right.$$

tem como solução

$$x = \frac{gd - bh}{ad - bc}, \quad y = \frac{ah - gc}{ad - bc}$$

Sugestão para os próximos 4 exercícios: Para verificar que uma matriz B é a inversa de uma matriz A, basta fazer um dos produtos AB ou BA e verificar que é igual a I_n.

2.1.11. Se A é uma matriz $n \times n$ e $A^k = \bar{0}$, para k um inteiro positivo, mostre que

$$(I_n - A)^{-1} = I_n + A + A^2 + \ldots + A^{k-1}.$$

2.1.12. Seja A uma **matriz diagonal**, isto é, os elementos que estão fora da diagonal são iguais a zero ($a_{ij} = 0$, para $i \neq j$). Se $a_{ii} \neq 0$, para $i = 1, \ldots, n$, mostre que A é invertível e a sua inversa é também uma matriz diagonal com elementos na diagonal dados por $1/a_{11}, 1/a_{22}, \ldots, 1/a_{nn}$.

2.1.13. Sejam A e B matrizes quadradas. Mostre que se $A + B$ e A forem invertíveis, então

$$(A + B)^{-1} = A^{-1}(I_n + BA^{-1})^{-1}.$$

2.1.14. Seja J_n a matriz $n \times n$, cujas entradas são iguais a 1. Mostre que se $n > 1$, então

$$(I_n - J_n)^{-1} = I_n - \frac{1}{n-1} J_n.$$

(Sugestão: observe que $J_n^2 = nJ_n$.)

2.1.15. Mostre que se B é uma matriz invertível, então $AB^{-1} = B^{-1}A$ se, e somente se, $AB = BA$. (Sugestão: multiplique a equação $AB = BA$ por B^{-1}.)

2.1.16. Mostre que se A é uma matriz invertível, então $A + B$ e $I_n + BA^{-1}$ são ambas invertíveis ou ambas não invertíveis. (Sugestão: multiplique $A + B$ por A^{-1}.)

2.1.17. Sejam A e B matrizes $n \times n$. Mostre que se B não é invertível, então AB também não o é.

2.1.18. Mostre que se A e B são matrizes $n \times n$, invertíveis, então A e B são equivalentes por linhas.

2.1.19. Sejam A uma matriz $m \times n$ e B uma matriz $n \times m$, com $n < m$. Mostre que AB não é invertível. (Sugestão: Mostre que o sistema $(AB)X = \bar{0}$ tem solução não trivial.)

2.2 Determinantes

Vamos inicialmente definir o determinante de matrizes 1×1. Para cada matriz $A = [a]$ definimos o **determinante** de A, indicado por $\det(A)$, por $\det(A) = a$. Vamos, agora, definir o determinante de matrizes 2×2 e a partir daí definir para matrizes de ordem maior. A cada matriz A, 2×2, associamos um número real, denominado **determinante** de A, por:

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

Para definir o determinante de matrizes quadradas maiores, precisamos definir o que são os menores de uma matriz. Dada uma matriz $A = (a_{ij})_{n \times n}$, o **menor** do elemento a_{ij}, denotado por $\tilde{A}_{ij}$, é a submatriz $(n-1) \times (n-1)$ de A obtida eliminando-se a i-ésima linha e a j-ésima coluna de A, que tem o seguinte aspecto:

$$\tilde{A}_{ij} = \begin{array}{cc} \begin{array}{ccccc} \phantom{a_{11}} & & j & & \phantom{a_{1n}} \end{array} & \\ \left[\begin{array}{ccccc} a_{11} & \ldots & & \ldots & a_{1n} \\ \vdots & & & & \vdots \\ & & a_{ij} & & \\ \vdots & & & & \vdots \\ a_{n1} & \ldots & & \ldots & a_{nn} \end{array}\right] & \begin{array}{c} \\ \\ i \\ \\ \\ \end{array} \end{array}$$

Exemplo 2.8. Para uma matriz $A = (a_{ij})_{3\times 3}$,

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

Agora, vamos definir os cofatores de uma matriz quadrada $A = (a_{ij})_{3\times 3}$. O **cofator** do elemento a_{ij}, denotado por $\tilde{a}_{ij}$, é definido por

$$\tilde{a}_{ij} = (-1)^{i+j} \det(\tilde{A}_{ij}),$$

ou seja, o cofator $\tilde{a}_{ij}$, do elemento a_{ij} é igual a mais ou menos o determinante do menor $\tilde{A}_{ij}$, sendo o mais e o menos determinados pela seguinte disposição:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

Exemplo 2.9. Para uma matriz $A = (a_{ij})_{3\times 3}$,

$$\tilde{a}_{23} = (-1)^{2+3} \det(\tilde{A}_{23}) = -\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = -\det \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} = a_{31}a_{12} - a_{11}a_{32}$$

Vamos, agora, definir o determinante de uma matriz 3×3. Se

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

então, o determinante de A é igual à soma dos produtos dos elementos da 1ª linha pelos seus cofatores.

$$\begin{aligned}\det(A) &= a_{11}\tilde{a}_{11} + a_{12}\tilde{a}_{12} + a_{13}\tilde{a}_{13}\\ &= a_{11}\det\begin{bmatrix} a_{22} & a_{23}\\ a_{32} & a_{33}\end{bmatrix} - a_{12}\det\begin{bmatrix} a_{21} & a_{23}\\ a_{31} & a_{33}\end{bmatrix} + a_{13}\det\begin{bmatrix} a_{21} & a_{22}\\ a_{31} & a_{32}\end{bmatrix}\\ &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}).\end{aligned}$$

Da mesma forma que a partir do determinante de matrizes 2×2, definimos o determinante de matrizes 3×3, podemos definir o determinante de matrizes quadradas de ordem maior. Supondo que sabemos como calcular o determinante de matrizes $(n-1) \times (n-1)$ vamos definir o determinante de matrizes $n \times n$.

Vamos definir, agora, os cofatores de uma matriz quadrada $A = (a_{ij})_{n\times n}$. O **cofator** do elemento a_{ij}, denotado por $\tilde{a}_{ij}$, é definido por

$$\tilde{a}_{ij} = (-1)^{i+j}\det(\tilde{A}_{ij}),$$

ou seja, o cofator $\tilde{a}_{ij}$, do elemento a_{ij} é igual a mais ou menos o determinante do menor $\tilde{A}_{ij}$, sendo o mais e o menos determinados pela seguinte disposição:

$$\begin{bmatrix} + & - & + & - & \ldots\\ - & + & - & + & \ldots\\ + & - & + & - & \ldots\\ \vdots & \vdots & \vdots & \ddots & \ddots\end{bmatrix}$$

Definição 2.2. Seja $A = (a_{ij})_{n\times n}$. O **determinante** de A, denotado por $\det(A)$, é definido por

$$\det(A) = a_{11}\tilde{a}_{11} + a_{12}\tilde{a}_{12} + \ldots + a_{1n}\tilde{a}_{1n} = \sum_{j=1}^{n} a_{1j}\tilde{a}_{1j}, \tag{2.7}$$

em que $\tilde{a}_{1j} = (-1)^{1+j}\det(\tilde{A}_{1j})$ é o cofator do elemento a_{1j}. A expressão (2.8) é chamada **desenvolvimento em cofatores do determinante de** A em termos da 1ª linha.

Exemplo 2.10. Seja

$$A = \begin{bmatrix} \boxed{\begin{matrix} 0 & 0 & 0 & -3 \end{matrix}} \\ \begin{matrix} 1 & 2 & 3 & 4 \\ -1 & 3 & 2 & 5 \\ 2 & 1 & -2 & 0 \end{matrix} \end{bmatrix}.$$

Desenvolvendo-se o determinante de A em cofatores, obtemos

$$\det(A) = 0\tilde{a}_{11} + 0\tilde{a}_{12} + 0\tilde{a}_{13} + (-3)(-1)^{1+4}\det(B), \quad \text{em que} \quad B = \begin{bmatrix} \boxed{\begin{matrix} 1 & 2 & 3 \end{matrix}} \\ \begin{matrix} -1 & 3 & 2 \\ 2 & 1 & -2 \end{matrix} \end{bmatrix}.$$

Mas o $\det(B)$ também pode ser calculado usando cofatores,

$$\begin{aligned}
\det(B) &= 1B_{11} + 2B_{12} + 3B_{13} \\
&= 1(-1)^{1+1}\det(\tilde{B}_{11}) + 2(-1)^{1+2}\det(\tilde{B}_{12}) + 3(-1)^{1+3}\det(\tilde{B}_{13}) \\
&= \det\begin{bmatrix} 3 & 2 \\ 1 & -2 \end{bmatrix} - 2\det\begin{bmatrix} -1 & 2 \\ 2 & -2 \end{bmatrix} + 3\det\begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix} \\
&= -8 - 2\,(-2) + 3\,(-7) \\
&= -25
\end{aligned}$$

Portanto,

$$\det(A) = 3\det(B) = -75.$$

Exemplo 2.11. Usando a definição de determinante, vamos mostrar que o determinante de uma matriz **triangular inferior** (isto é, os elementos situados acima da diagonal principal são iguais a zero) é o produto dos elementos da diagonal principal. Vamos mostrar inicialmente para matrizes 3×3. Seja

$$A = \begin{bmatrix} \boxed{a_{11} \quad 0 \quad 0} \\ \begin{matrix} a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{matrix} \end{bmatrix}$$

Desenvolvendo-se o determinante de A em cofatores, obtemos

$$\det(A) = a_{11} \det \begin{bmatrix} a_{22} & 0 \\ a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33}.$$

Vamos supor termos provado que para qualquer matriz $(n-1) \times (n-1)$ triangular inferior, o determinante é o produto dos elementos da diagonal principal. Então vamos provar que isto também vale para matrizes $n \times n$. Seja

$$A = \begin{bmatrix} \boxed{a_{11} \quad 0 \quad \ldots \quad \ldots \quad 0} \\ \begin{matrix} a_{21} & a_{22} & 0 & & \vdots \\ \vdots & & \ddots & & 0 \\ a_{n1} & & \ldots & & a_{nn} \end{matrix} \end{bmatrix}$$

Desenvolvendo-se o determinante de A em cofatores, obtemos

$$\det(A) = a_{11} \det \begin{bmatrix} a_{22} & 0 & \ldots & \ldots & 0 \\ a_{32} & a_{33} & 0 & & \vdots \\ \vdots & & \ddots & & 0 \\ a_{n2} & & \ldots & & a_{nn} \end{bmatrix} = a_{11}a_{22}\ldots a_{nn},$$

pois o determinante acima é de uma matriz $(n-1) \times (n-1)$ triangular inferior. Em particular, para a matriz identidade, I_n,

$$\det(I_n) = 1.$$

2.2.1 Propriedades do Determinante

Vamos provar uma propriedade importante do determinante. Para isso vamos escrever a matriz $A = (a_{ij})_{n\times n}$ em termos das suas linhas

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_{k-1} \\ A_k \\ A_{k+1} \\ \vdots \\ A_n \end{bmatrix},$$

em que A_i é a linha i da matriz A, ou seja, $A_i = [\, a_{i1}\; a_{i2}\; \ldots\; a_{in} \,]$. Se a linha A_k é escrita na forma $A_k = \alpha X + \beta Y$, em que X e Y são matrizes linha $1 \times n$, então o determinante pode ser decomposto como mostra o resultado seguinte.

Teorema 2.10. *Seja $A = (a_{ij})_{n\times n}$ escrita em termos das suas linhas, denotadas por A_i, ou seja, $A_i = [\, a_{i1}\, a_{i2} \ldots a_{in}\,]$. Se para algum k, a linha $A_k = \alpha X + \beta Y$, em que $X = [\, x_1 \ldots x_n\,]$, $Y = [\, y_1 \ldots y_n\,]$ e α e β são escalares, então:*

$$\det\begin{bmatrix} A_1 \\ \vdots \\ A_{k-1} \\ \alpha X + \beta Y \\ A_{k+1} \\ \vdots \\ A_n \end{bmatrix} = \alpha \det\begin{bmatrix} A_1 \\ \vdots \\ A_{k-1} \\ X \\ A_{k+1} \\ \vdots \\ A_n \end{bmatrix} + \beta \det\begin{bmatrix} A_1 \\ \vdots \\ A_{k-1} \\ Y \\ A_{k+1} \\ \vdots \\ A_n \end{bmatrix}.$$

Aqui, $A_k = \alpha X + \beta Y = [\, \alpha x_1 + \beta y_1 \,\ldots\, \alpha x_n + \beta y_n\,]$.

Demonstração. Vamos provar aqui somente para $k = 1$. Para $k > 1$ é demonstrado no Apêndice II na página 135. Se $A_1 = \alpha X + \beta Y$, em que $X = [\, x_1 \ldots x_n\,]$,

$Y = [\, y_1 \ldots y_n \,]$ e α e β são escalares, então:

$$\begin{aligned}
\det \begin{bmatrix} \alpha X + \beta Y \\ A_2 \\ \vdots \\ A_n \end{bmatrix} &= \sum_{j=1}^{n} (-1)^{1+j} (\alpha x_j + \beta y_j) \det(\tilde{A}_{1j}) \\
&= \alpha \sum_{j=1}^{n} x_j \det(\tilde{A}_{1j}) + \beta \sum_{j=1}^{n} y_j \det(\tilde{A}_{1j}) \\
&= \alpha \det \begin{bmatrix} X \\ A_2 \\ \vdots \\ A_n \end{bmatrix} + \beta \det \begin{bmatrix} Y \\ A_2 \\ \vdots \\ A_n \end{bmatrix}
\end{aligned}$$

■

Exemplo 2.12. O cálculo do determinante da matriz a seguir pode ser feito da seguinte forma:

$$\det \begin{bmatrix} \cos t & \operatorname{sen} t \\ 2\cos t - 3\operatorname{sen} t & 2\operatorname{sen} t + 3\cos t \end{bmatrix} = 2\det \begin{bmatrix} \cos t & \operatorname{sen} t \\ \cos t & \operatorname{sen} t \end{bmatrix} + 3\det \begin{bmatrix} \cos t & \operatorname{sen} t \\ -\operatorname{sen} t & \cos t \end{bmatrix} = 3$$

Pela definição de determinante, o determinante deve ser calculado fazendo-se o desenvolvimento em cofatores segundo a 1ª linha. O próximo resultado, que não vamos provar neste momento (Apêndice II na página 135), afirma que o determinante pode ser calculado fazendo-se o desenvolvimento em cofatores segundo *qualquer linha* ou *qualquer coluna*.

Teorema 2.11. *Seja A uma matriz $n \times n$. O determinante de A pode ser calculado fazendo-se o desenvolvimento em cofatores segundo* ***qualquer linha*** *ou* ***qualquer coluna****.*

$$\begin{aligned}
\det(A) &= a_{i1}\tilde{a}_{i1} + a_{i2}\tilde{a}_{i2} + \ldots + a_{in}\tilde{a}_{in} = \sum_{j=1}^{n} a_{ij}\tilde{a}_{ij}, \quad \textit{para } i = 1, \ldots, n, && (2.8)\\
&= a_{1j}\tilde{a}_{1j} + a_{2j}\tilde{a}_{2j} + \ldots + a_{nj}\tilde{a}_{nj} = \sum_{i=1}^{n} a_{ij}\tilde{a}_{ij}, \quad \textit{para } j = 1, \ldots, n, && (2.9)
\end{aligned}$$

em que $\tilde{a}_{ij} = (-1)^{i+j}\det(\tilde{A}_{ij})$ é o cofator do elemento a_{ij}. A expressão (2.8) é chamada ***desenvolvimento em cofatores do determinante de A em termos da i-ésima linha*** *e (2.9) é chamada* ***desenvolvimento em cofatores do determinante de A em termos da j-ésima coluna****.*

Temos a seguinte consequência deste resultado.

Corolário 2.12. *Seja A uma matriz $n \times n$. Se A possui duas linhas iguais, então $\det(A) = 0$.*

Demonstração. O resultado é claramente verdadeiro para matrizes 2×2. Supondo que o resultado seja verdadeiro para matrizes $(n-1) \times (n-1)$, vamos provar que ele é verdadeiro para matrizes $n \times n$. Suponhamos que as linhas k e l sejam iguais, para $k \neq l$. Desenvolvendo o determinante de A em termos de uma linha i, com $i \neq k, l$, obtemos

$$\det(A) = \sum_{j=1}^{n} a_{ij}\tilde{a}_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\tilde{A}_{ij}).$$

Mas, cada $\tilde{A}_{ij}$ é uma matriz $(n-1) \times (n-1)$ com duas linhas iguais. Como estamos supondo que o resultado seja verdadeiro para estas matrizes, então $\det(\tilde{A}_{ij}) = 0$. Isto implica que $\det(A) = 0$. ■

No próximo resultado mostramos como varia o determinante de uma matriz quando aplicamos operações elementares sobre suas linhas.

Teorema 2.13. *Sejam A e B matrizes $n \times n$.*

(a) *Se B é obtida de A multiplicando-se uma linha por um escalar α, então*

$$\det(B) = \alpha \det(A)\,;$$

(b) *Se B resulta de A pela troca da posição de duas linhas $k \neq l$, então*

$$\det(B) = -\det(A)\,;$$

(c) *Se B é obtida de A substituindo a linha l por ela somada a um múltiplo escalar de uma linha k, $k \neq l$, então*

$$\det(B) = \det(A)\,.$$

Demonstração. (a) Segue diretamente do Teorema 2.10 na página 106.

(b) Sejam

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_l \\ \vdots \\ A_n \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} A_1 \\ \vdots \\ A_l \\ \vdots \\ A_k \\ \vdots \\ A_n \end{bmatrix}.$$

Agora, pelo Teorema 2.10 na página 106 e o Corolário 2.12, temos que

$$
\begin{aligned}
0 &= \det\begin{bmatrix} A_1 \\ \vdots \\ A_k + A_l \\ \vdots \\ A_k + A_l \\ \vdots \\ A_n \end{bmatrix} = \det\begin{bmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_k \\ \vdots \\ A_n \end{bmatrix} + \det\begin{bmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_l \\ \vdots \\ A_n \end{bmatrix} + \det\begin{bmatrix} A_1 \\ \vdots \\ A_l \\ \vdots \\ A_k \\ \vdots \\ A_n \end{bmatrix} + \det\begin{bmatrix} A_1 \\ \vdots \\ A_l \\ \vdots \\ A_l \\ \vdots \\ A_n \end{bmatrix} \\
&= 0 + \det(A) + \det(B) + 0.
\end{aligned}
$$

Portanto, $\det(A) = -\det(B)$.

(c) Novamente, pelo Teorema 2.10 na página 106, temos que

$$
\det\begin{bmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_l + \alpha A_k \\ \vdots \\ A_n \end{bmatrix} = \det\begin{bmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_l \\ \vdots \\ A_n \end{bmatrix} + \alpha \det\begin{bmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_k \\ \vdots \\ A_n \end{bmatrix} = \det\begin{bmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_l \\ \vdots \\ A_n \end{bmatrix}.
$$

■

Exemplo 2.13. Vamos calcular o determinante da matriz

$$
A = \begin{bmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{bmatrix}
$$

usando operações elementares para transformá-la numa matriz triangular superior e aplicando o Teorema 2.13.

$$
\begin{aligned}
\det(A) &= -\det\begin{bmatrix} 3 & -6 & 9 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{bmatrix} && \boxed{1^{\underline{a}} \text{ linha} \longleftrightarrow 2^{\underline{a}} \text{ linha}} \\
&= -3\det\begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{bmatrix} && \boxed{1/3\times 1^{\underline{a}} \text{ linha} \longrightarrow 1^{\underline{a}} \text{ linha}} \\
&= -3\det\begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 10 & -5 \end{bmatrix} && \boxed{-2\times 1^{\underline{a}} \text{ linha}+3^{\underline{a}} \text{ linha} \longrightarrow 3^{\underline{a}} \text{ linha}} \\
&= -3\det\begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & -55 \end{bmatrix} && \boxed{-10\times 2^{\underline{a}} \text{ linha}+3^{\underline{a}} \text{ linha} \longrightarrow 3^{\underline{a}} \text{ linha}} \\
&= (-3)(-55) = 165
\end{aligned}
$$

Quando multiplicamos uma linha de uma matriz por um escalar α o determinante da nova matriz é igual a α multiplicado pelo determinante da matriz antiga. Mas o que estamos calculando aqui é o determinante da matriz antiga, por isso ele é igual a $1/\alpha$ multiplicado pelo determinante da matriz nova.

Para se calcular o determinante de uma matriz $n \times n$ pela expansão em cofatores, precisamos fazer n produtos e calcular n determinantes de matrizes $(n-1)\times(n-1)$, que por sua vez vai precisar de $n-1$ produtos e assim por diante. Portanto, ao todo são necessários da ordem de $n!$ produtos. Para se calcular o determinante de uma matriz 20×20, é necessário se realizar $20! \approx 10^{18}$ produtos. Os computadores pessoais realizam da ordem de 10^8 produtos por segundo. Portanto, um computador pessoal precisaria de cerca de 10^{10} segundos ou 10^3 anos para calcular o determinante de uma matriz 20×20 usando a expansão em cofatores. Entretanto usando o método apresentado no exemplo anterior para o cálculo do determinante, é necessário apenas da ordem de n^3 produtos. Ou seja, para calcular o determinante de uma matriz 20×20 usando o método apresentado no exemplo anterior um computador pessoal gasta muito menos de um segundo.

A seguir estabelecemos duas propriedades do determinante que serão demonstradas somente na Subseção 2.2.2 na página 119.

Teorema 2.14. *Sejam A e B matrizes $n \times n$.*

(a) *O determinante do produto de A por B é igual ao produto dos seus determinantes,*

$$\det(AB) = \det(A)\det(B)\,.$$

(b) *Os determinantes de A e de sua transposta A^t são iguais,*

$$\det(A) = \det(A^t)\,;$$

Observação. Como o determinante de uma matriz é igual ao determinante da sua transposta (Teorema 2.14 (b)), segue-se que todas as propriedades que se referem a linhas são válidas com relação às colunas.

Exemplo 2.14. Seja $A = (a_{ij})_{n\times n}$. Vamos mostrar que se A é invertível, então

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

Como $A\,A^{-1} = I_n$, aplicando-se o determinante a ambos os membros desta igualdade e usando o Teorema 2.14, obtemos

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

Mas, $\det(I_n) = 1$ (Exemplo 2.11 na página 104, a matriz identidade também é triangular inferior!). Logo, $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

Exemplo 2.15. Se uma matriz quadrada é tal que $A^2 = A^{-1}$, então vamos mostrar que $\det(A) = 1$. Aplicando-se o determinante a ambos os membros da igualdade acima, e usando novamente o Teorema 2.14 e o resultado do exemplo anterior, obtemos

$$(\det(A))^2 = \frac{1}{\det(A)}.$$

Logo, $(\det(A))^3 = 1$. Portanto, $\det(A) = 1$.

O resultado seguinte caracteriza em termos do determinante as matrizes invertíveis e os sistemas lineares homogêneos que possuem solução não trivial.

Teorema 2.15. *Seja A uma matriz $n \times n$.*

(a) *A matriz A é invertível se, e somente se,* $\det(A) \neq 0$.

(b) *O sistema homogêneo $AX = \bar{0}$ tem solução não trivial se, e somente se,* $\det(A) = 0$.

Demonstração. (a) Seja R a forma escalonada reduzida da matriz A.

A demonstração deste item segue-se de três observações:

- Pelo Teorema 2.13 na página 110, $\det(A) \neq 0$ se, e somente se, $\det(R) \neq 0$.
- Pela Proposição 1.5 da página 46, ou $R = I_n$ ou a matriz R tem uma linha nula. Assim, $\det(A) \neq 0$ se, e somente se, $R = I_n$.
- Pelo Teorema 2.7 na página 80, $R = I_n$ se, e somente se, A é invertível.

(b) Pelo Teorema 2.8 na página 84, o sistema homogêneo $AX = \bar{0}$ tem solução não trivial se, e somente se, a matriz A não é invertível. E pelo item anterior, a matriz A é não invertível se, e somente se, $\det(A) = 0$.

■

Exemplo 2.16. Considere a matriz

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

(a) Determinar os valores de $\lambda \in \mathbb{R}$ tais que existe $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \neq \bar{0}$ que satisfaz $AX = \lambda X$.

(b) Para cada um dos valores de λ encontrados no item anterior determinar todos $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ tais que $AX = \lambda X$.

Solução:

(a) Como a matriz identidade I_3 é o elemento neutro do produto, então

$$AX = \lambda X \quad \Leftrightarrow \quad AX = \lambda I_3 X.$$

Subtraindo-se $\lambda I_3 X$ obtemos

$$AX - \lambda I_3 X = \bar{0} \quad \Leftrightarrow \quad (A - \lambda I_3)X = \bar{0}.$$

Agora, este sistema homogêneo tem solução não trivial ($X \neq \bar{0}$) se, e somente se,

$$\det(A - \lambda I_3) = 0.$$

Mas

$$\det \begin{bmatrix} 2-\lambda & 2 & 2 \\ 0 & 2-\lambda & 0 \\ 0 & 1 & 3-\lambda \end{bmatrix} = -(\lambda-2)^2(\lambda-3) = 0$$

se, e somente se, $\lambda = 2$ ou $\lambda = 3$. Assim, somente para $\lambda = 2$ e $\lambda = 3$ existem vetores $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \neq \bar{0}$ tais que $AX = \lambda X$.

(b) Para $\lambda = 2$:

$$(A-2I_3)X=\bar{0} \quad \Leftrightarrow \quad \begin{bmatrix} 0 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \left\{ \begin{array}{rcrcl} 2y & + & 2z & = & 0 \\ y & + & z & = & 0 \end{array} \right.$$

que tem solução o conjunto dos $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \beta \\ -\alpha \\ \alpha \end{bmatrix}$, para todos os valores de $\alpha, \beta \in \mathbb{R}$.

Para $\lambda = 3$:

$$(A-3I_3)X=\bar{0} \quad \Leftrightarrow \quad \begin{bmatrix} -1 & 2 & 2 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \left\{ \begin{array}{rcrcrcl} -x & + & 2y & + & 2z & = & 0 \\ & & -y & & & = & 0 \\ & & y & & & = & 0 \end{array} \right.$$

que tem solução o conjunto dos $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2\alpha \\ 0 \\ \alpha \end{bmatrix}$, para todos os valores de $\alpha \in \mathbb{R}$.

Exemplo 2.17. A matriz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ é invertível se, e somente se,

$$\det(A) = ad - bc \neq 0.$$

Neste caso a inversa de A é dada por

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

como pode ser verificado multiplicando-se a candidata a inversa pela matriz A.

Observe que este exemplo fornece uma regra para se encontrar a inversa de uma matriz 2×2: troca-se a posição dos elementos da diagonal principal, troca-se o sinal dos outros elementos e divide-se todos os elementos pelo determinante de A.

Exemplo 2.18. Considere o sistema linear de 2 equações e 2 incógnitas

$$\left\{ \begin{array}{ccccc} ax & + & by & = & g \\ cx & + & dy & = & h \end{array} \right.$$

A matriz deste sistema é

$$A = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right].$$

Se $\det(A) \neq 0$, então a solução do sistema é

$$X = A^{-1}B = \frac{1}{\det(A)} \left[\begin{array}{rr} d & -b \\ -c & a \end{array} \right] \left[\begin{array}{c} g \\ h \end{array} \right] = \frac{1}{\det(A)} \left[\begin{array}{c} dg - bh \\ -cg + ah \end{array} \right] = \frac{1}{\det(A)} \left[\begin{array}{c} \det \left[\begin{array}{cc} g & b \\ h & d \end{array} \right] \\ \det \left[\begin{array}{cc} a & g \\ c & h \end{array} \right] \end{array} \right]$$

ou seja,

$$x = \frac{\det \left[\begin{array}{cc} g & b \\ h & d \end{array} \right]}{\det \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]}, \quad y = \frac{\det \left[\begin{array}{cc} a & g \\ c & h \end{array} \right]}{\det \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]}$$

esta é a chamada **Regra de Cramer** para sistemas de 2 equações e 2 incógnitas. A Regra de Cramer para sistemas de n equações e n incógnitas será apresentada na Subseção 2.2.3.

2.2.2 Matrizes Elementares e o Determinante (opcional)

Relembramos que uma matriz elementar é uma matriz que se obtém aplicando-se uma operação elementar na matriz identidade. Assim, aplicando-se o Teorema 2.13 na página 110 obtemos o resultado seguinte.

Proposição 2.16. *(a) Se $E_{i,j}$ é a matriz elementar obtida trocando-se as linhas i e j da matriz identidade, então* $\det(E_{i,j}) = -1$.

(b) Se $E_i(\alpha)$ é a matriz elementar obtida da matriz identidade, multiplicando-se a linha i por α, então $\det(E_i(\alpha)) = \alpha$.

(c) Se $E_{i,j}(\alpha)$ é a matriz elementar obtida da matriz identidade, somando-se à linha j, α vezes a linha i, então $\det(E_{i,j}(\alpha)) = 1$.

Lembramos também que uma matriz é invertível se, e somente se, ela é o produto de matrizes elementares (Teorema 2.6 na página 77). Além disso, o resultado da aplicação de uma operação elementar em uma matriz é o mesmo que multiplicar a matriz à esquerda pela matriz elementar correspondente.

Usando matrizes elementares podemos provar o Teorema 2.14 na página 113.

Demonstração do Teorema 2.14.

(a) Queremos provar que $\det(AB) = \det(A)\det(B)$. Vamos dividir a demonstração deste item em três casos:

Caso 1: Se $A = E$ é uma matriz elementar. Este caso segue-se diretamente da proposição anterior e do Teorema 2.13 na página 110.

Caso 2: Se A é invertível, então pelo Teorema 2.6 na página 77 ela é o produto de matrizes elementares, $A = E_1 \ldots E_k$. Aplicando-se o caso anterior sucessivas vezes, obtemos

$$\det(AB) = \det(E_1)\ldots\det(E_k)\det(B) = \det(E_1 \ldots E_k)\det(B) = \det(A)\det(B).$$

Caso 3: Se A é singular, pela Proposição 2.9 na página 89, AB também é singular. Logo,

$$\det(AB) = 0 = 0\det(B) = \det(A)\det(B).$$

(b) Queremos provar que $\det(A) = \det(A^t)$. Vamos dividir a demonstração deste item em dois casos.

Caso 1: Se A é uma matriz invertível, pelo Teorema 2.6 na página 77 ela é o produto de matrizes elementares, $A = E_1 \dots E_k$. É fácil ver que se E é uma matriz elementar, então $\det(E) = \det(E^t)$ (verifique!). Assim,

$$\det(A^t) = \det(E_k^t) \dots \det(E_1^t) = \det(E_k) \dots \det(E_1) = \det(E_1 \dots E_k) = \det(A).$$

Caso 2: Se A não é invertível, então A^t também não o é, pois caso contrário, pelo Teorema 2.2 na página 72, também $A = (A^t)^t$ seria invertível. Assim neste caso, $\det(A^t) = 0 = \det(A)$. ■

2.2.3 Matriz Adjunta e Inversão (opcional)

Vamos definir a adjunta de uma matriz quadrada e em seguida enunciar e provar um teorema sobre a adjunta que permite provar vários resultados sobre matrizes, entre eles um que fornece uma fórmula para a inversa de uma matriz e também a regra de Cramer. Tanto a adjunta quanto os resultados que vem a seguir são de importância teórica.

Definição 2.3. Seja A uma matriz $n \times n$. Definimos a matriz **adjunta (clássica)** de A, denotada por $\operatorname{adj}(A)$, como a transposta da matriz formada pelos cofatores de A, ou seja,

$$\operatorname{adj}(A) = \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \ldots & \tilde{a}_{1n} \\ \tilde{a}_{21} & \tilde{a}_{22} & \ldots & \tilde{a}_{2n} \\ \vdots & & \ldots & \vdots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \ldots & \tilde{a}_{nn} \end{bmatrix}^t = \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{21} & \ldots & \tilde{a}_{n1} \\ \tilde{a}_{12} & \tilde{a}_{22} & \ldots & \tilde{a}_{n2} \\ \vdots & & \ldots & \vdots \\ \tilde{a}_{1n} & \tilde{a}_{2n} & \ldots & \tilde{a}_{nn} \end{bmatrix},$$

em que, $\tilde{a}_{ij} = (-1)^{i+j} \det(\tilde{A}_{ij})$ é o cofator do elemento a_{ij}, para $i, j = 1, \ldots, n$.

Exemplo 2.19. Seja

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

Vamos calcular a adjunta de B.

$$\tilde{b}_{11} = (-1)^{1+1} \det \begin{bmatrix} 3 & 2 \\ 0 & -2 \end{bmatrix} = -6, \quad \tilde{b}_{12} = (-1)^{1+2} \det \begin{bmatrix} 0 & 2 \\ 0 & -2 \end{bmatrix} = 0,$$
$$\tilde{b}_{13} = (-1)^{1+3} \det \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} = 0, \quad \tilde{b}_{21} = (-1)^{2+1} \det \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} = 4,$$
$$\tilde{b}_{22} = (-1)^{2+2} \det \begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix} = -2, \quad \tilde{b}_{23} = (-1)^{2+3} \det \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = 0,$$
$$\tilde{b}_{31} = (-1)^{3+1} \det \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} = -5, \quad \tilde{b}_{32} = (-1)^{3+2} \det \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix} = -2,$$
$$\tilde{b}_{33} = (-1)^{3+3} \det \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} = 3,$$

Assim, a adjunta de B é

$$\operatorname{adj}(B) = \begin{bmatrix} -6 & 0 & 0 \\ 4 & -2 & 0 \\ -5 & -2 & 3 \end{bmatrix}^t = \begin{bmatrix} -6 & 4 & -5 \\ 0 & -2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

Na definição do determinante são multiplicados os elementos de uma linha pelos cofatores da mesma linha. O teorema seguinte diz o que acontece se somamos os produtos dos elementos de uma linha com os cofatores de outra linha ou se somamos os produtos dos elementos de uma coluna com os cofatores de outra coluna.

Lema 2.17. *Se A é uma matriz $n \times n$, então*

$$a_{k1}\tilde{a}_{i1} + a_{k2}\tilde{a}_{i2} + \ldots + a_{kn}\tilde{a}_{in} = 0 \quad \textit{se } k \neq i; \tag{2.10}$$

$$a_{1k}\tilde{a}_{1j} + a_{2k}\tilde{a}_{2j} + \ldots + a_{nk}\tilde{a}_{nj} = 0 \quad \textit{se } k \neq j; \tag{2.11}$$

em que, $\tilde{a}_{ij} = (-1)^{i+j}\det(\tilde{A}_{ij})$ é o cofator do elemento a_{ij}, para $i, j = 1, \ldots, n$.

Demonstração. Para demonstrar a equação (2.10), definimos a matriz A^* como sendo a matriz obtida de A substituindo a i-ésima linha de A por sua k-ésima linha, ou seja,

$$A = \begin{array}{c l} \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_k \\ \vdots \\ A_n \end{bmatrix} & \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow k \\ \\ \\ \end{array} \end{array} \quad \text{e} \quad A^* = \begin{array}{c l} \begin{bmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_k \\ \vdots \\ A_n \end{bmatrix} & \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow k \\ \\ \\ \end{array} \end{array} .$$

Assim, A^* possui duas linhas iguais e pelo Corolário 2.12 na página 108, $\det(A^*) = 0$. Mas, o determinante de A^* desenvolvido segundo a sua i-ésima linha é exatamente a equação (2.10).
A demonstração de (2.11) é feita de forma análoga, mas usando o item (d) do Teorema 2.13, ou seja, que $\det(A) = \det(A^t)$. ■

Teorema 2.18. *Se A é uma matriz $n \times n$, então*

$$A(\operatorname{adj}(A)) = (\operatorname{adj}(A))A = \det(A)I_n$$

Demonstração. O produto da matriz A pela matriz adjunta de A é dada por

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & & \ldots & \vdots \\ \boxed{a_{i1} \quad a_{i2} \quad \ldots \quad a_{in}} \\ \vdots & & \ldots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{bmatrix} \begin{bmatrix} \tilde{a}_{11} & \ldots & \tilde{a}_{j1} & \ldots & \tilde{a}_{n1} \\ \tilde{a}_{12} & \ldots & \tilde{a}_{j2} & \ldots & \tilde{a}_{n2} \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ \tilde{a}_{1n} & \ldots & \tilde{a}_{jp} & \ldots & \tilde{a}_{nn} \end{bmatrix}$$

O elemento de posição i, j de $A\operatorname{adj}(A)$ é

$$(A\operatorname{adj}(A))_{ij} = \sum_{k=1}^{n} a_{ik}\tilde{a}_{jk} = a_{i1}\tilde{a}_{j1} + a_{i2}\tilde{a}_{j2} + \ldots a_{in}\tilde{a}_{jn}\,.$$

Pelo Lema 2.17, equação (2.10) e do Teorema 2.11 na página 108 segue-se que

$$(A\,\text{adj}(A))_{ij} = \begin{cases} \det(A) & \text{se } i = j \\ 0 & \text{se } i \neq j. \end{cases}$$

Assim,

$$A\,\text{adj}(A) = \begin{bmatrix} \det(A) & 0 & \dots & 0 \\ 0 & \det(A) & \dots & 0 \\ \vdots & & \dots & \vdots \\ 0 & 0 & \dots & \det(A) \end{bmatrix} = \det(A) I_n\,.$$

Analogamente, usando Lema 2.17, equação (2.11), se prova que $\text{adj}(A)\;A = \det(A)I_n$. ■

Exemplo 2.20. Vamos mostrar que se uma matriz A é singular, então $\text{adj}(A)$ também é singular. Vamos separar em dois casos.

(a) Se $A = \bar{0}$, então $\text{adj}(A)$ também é a matriz nula, que é singular.

(b) Se $A \neq \bar{0}$, então pelo Teorema 2.18 na página 124, $\text{adj}(A)\,A = \bar{0}$. Mas, então, se $\text{adj}(A)$ fosse invertível, então A seria igual à matriz nula (por que?), que estamos assumindo não ser este o caso. Portanto, $\text{adj}(A)$ tem que ser singular.

Corolário 2.19. *Seja A uma matriz $n \times n$. Se $\det(A) \neq 0$, então*

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A)\,;$$

Demonstração. Se $\det(A) \neq 0$, então definindo $B = \dfrac{1}{\det(A)}\text{adj}(A)$, pelo Teorema 2.18 temos que

$$A\,B = A(\frac{1}{\det(A)}\text{adj}(A)) = \frac{1}{\det(A)}(A\ \text{adj}(A)) = \frac{1}{\det(A)}\det(A)I_n = I_n\,.$$

Aqui, usamos a propriedade (j) do Teorema 1.1 na página 9. Portanto, A é invertível e B é a inversa de A. ■

Exemplo 2.21. No Exemplo 2.17 na página 117 mostramos como obter rapidamente a inversa de ma matriz 2×2. Usando o Corolário 2.19 podemos também obter a inversa de uma matriz 2×2,

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

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A) = \frac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \quad \text{se } \det(A) \neq 0$$

Ou seja, a inversa de uma matriz 2×2 é facilmente obtida trocando-se a posição dos elementos da diagonal principal, trocando-se o sinal dos outros elementos e dividindo-se todos os elementos pelo determinante de A.

Exemplo 2.22. Vamos calcular a inversa da matriz

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

A sua adjunta foi calculada no Exemplo 2.19 na página 122. Assim,

$$B^{-1} = \frac{1}{\det(B)}\operatorname{adj}(B) = \frac{1}{-6}\begin{bmatrix} -6 & 4 & -5 \\ 0 & -2 & -2 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{2}{3} & \frac{5}{6} \\ 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}.$$

Corolário 2.20 (Regra de Cramer). *Se o sistema linear $AX = B$ é tal que a matriz A é $n \times n$ e invertível, então a solução do sistema é dada por*

$$x_1 = \frac{\det(A_1)}{\det(A)},\ x_2 = \frac{\det(A_2)}{\det(A)}, \ldots, x_n = \frac{\det(A_n)}{\det(A)},$$

em que A_j é a matriz que se obtém de A substituindo-se a sua j-ésima coluna por B, para $j = 1, \ldots, n$.

Demonstração. Como A é invertível, pelo Corolário 2.19

$$X = A^{-1}\,B = \frac{1}{\det(A)}\mathrm{adj}(A)B.$$

A entrada x_j é dada por

$$x_j = \frac{1}{\det(A)}(\tilde{a}_{1j}b_1 + \ldots + \tilde{a}_{nj}b_n) = \frac{\det(A_j)}{\det(A)},$$

em que A_j é a matriz que se obtém de A substituindo-se a sua j-ésima coluna por B, para $j = 1, \ldots, n$ e $\det(A_j)$ foi calculado fazendo o desenvolvimento em cofatores em relação a j-ésima coluna de A_j. ■

Se a matriz A não é invertível, então a regra de Cramer não pode ser aplicada. Pode ocorrer que $\det(A) = \det(A_j) = 0$, para $j = 1, \ldots, n$ e o sistema não tenha solução (verifique!). A regra de Cramer tem um valor teórico, por fornecer uma fórmula para a solução de um sistema linear, quando a matriz do sistema é quadrada e invertível.

Existem sistemas $AX = B$ de n equações e n incógnitas, com $n > 2$, em que

$$\det(A) = \det(A_1) = \cdots = \det(A_n) = 0$$

e o sistema não tem solução. Por exemplo o sistema

$$\left\{\begin{array}{rcrcrcr} x & + & 2y & + & 3z & = & 1 \\ 2x & + & 4y & + & 6z & = & 3 \\ 3x & + & 6y & + & 9z & = & 2 \end{array}\right.$$

é tal que

$$\det(A) = \det(A_1) = \det(A_2) = \det(A_3) = 0,$$

mas ele não tem solução (verifique!).

Exercícios Numéricos (respostas na página 496)

2.2.1. Se $\det(A) = -3$, encontre
(a) $\det(A^2)$; (b) $\det(A^3)$; (c) $\det(A^{-1})$; (d) $\det(A^t)$;

2.2.2. Se A e B são matrizes $n \times n$ tais que $\det(A) = -2$ e $\det(B) = 3$, calcule $\det(A^t B^{-1})$.

2.2.3. Seja $A = (a_{ij})_{3\times 3}$ tal que $\det(A) = 3$. Calcule o determinante das matrizes a seguir:

(a) $\begin{bmatrix} a_{11} & a_{12} & a_{13}+a_{12} \\ a_{21} & a_{22} & a_{23}+a_{22} \\ a_{31} & a_{32} & a_{33}+a_{32} \end{bmatrix}$ (b) $\begin{bmatrix} a_{11}+a_{12} & a_{11}-a_{12} & a_{13} \\ a_{21}+a_{22} & a_{21}-a_{22} & a_{23} \\ a_{31}+a_{32} & a_{31}-a_{32} & a_{33} \end{bmatrix}$

2.2.4. Calcule o determinante das matrizes a seguir:

(a) $\begin{bmatrix} e^{rt} & te^{rt} \\ re^{rt} & (1+rt)e^{rt} \end{bmatrix}$ (b) $\begin{bmatrix} \cos\beta t & \operatorname{sen}\beta t \\ \alpha\cos\beta t - \beta\operatorname{sen}\beta t & \alpha\operatorname{sen}\beta t + \beta\cos\beta t \end{bmatrix}$

2.2.5. Calcule o determinante de cada uma das matrizes seguintes usando operações elementares para transformá-las em matrizes triangulares superiores.

(a) $\begin{bmatrix} 1 & -2 & 3 & 1 \\ 5 & -9 & 6 & 3 \\ -1 & 2 & -6 & -2 \\ 2 & 8 & 6 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 1 & 3 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 \end{bmatrix}$.

2.2.6. Determine todos os valores de λ para os quais $\det(A - \lambda I_n) = 0$, em que

(a) $A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 3 & 0 \\ 3 & 2 & -2 \end{bmatrix}$

(c) $A = \begin{bmatrix} 2 & -2 & 3 \\ 0 & 3 & -2 \\ 0 & -1 & 2 \end{bmatrix}$ (d) $A = \begin{bmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{bmatrix}$

2.2.7. Determine os valores de $\lambda \in \mathbb{R}$ tais que existe $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \neq \bar{0}$ que satisfaz $AX = \lambda X$.

(a) $A = \begin{bmatrix} 2 & 0 & 0 \\ 3 & -1 & 0 \\ 0 & 4 & 3 \end{bmatrix}$;

(b) $A = \begin{bmatrix} 2 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$;

(c) $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}$;

(d) $A = \begin{bmatrix} 2 & 2 & 3 & 4 \\ 0 & 2 & 3 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

2.2.8. Para as matrizes do exercício anterior, e os valores de λ encontrados, encontre a solução geral do sistema $AX = \lambda X$, ou equivalentemente, do sistema homogêneo $(A - \lambda I_n)X = \bar{0}$.

Exercícios usando o MATLAB®

Comandos do MATLAB®:

`>> det(A)` calcula o determinante da matriz `A`.

Comando do pacote GAAL:

`>> detopelp(A)` calcula o determinante de `A` aplicando operações elementares até que a matriz esteja na forma triangular superior.

2.2.9. Vamos fazer um experimento no MATLAB® para tentar ter uma ideia do quão comum é encontrar matrizes invertíveis. No prompt do MATLAB® digite a seguinte linha:

```
>> c=0; for n=1:1000,A=randi(2);if(det(A)~=0),c=c+1;end,end,c
```

(não esqueça das vírgulas e pontos e vírgulas!). O que esta linha está mandando o MATLAB® fazer é o seguinte:

- Criar um contador `c` e atribuir a ele o valor zero.
- Atribuir à variável `A`, 1000 matrizes 2×2 com entradas inteiras aleatórias entre -5 e 5.
- Se $\det(\mathtt{A}) \neq 0$, então o contador `c` é acrescido de 1.
- No final o valor existente na variável `c` é escrito.

Qual a conclusão que você tira do valor obtido na variável `c`?

2.2.10. Resolva, com o MATLAB®, os **Exercícios Numéricos a partir do Exercício 4.**

Exercícios Teóricos

2.2.11. Mostre que se $\det(AB) = 0$, então ou A é singular ou B é singular.

2.2.12. O determinante de AB é igual ao determinante de BA? Justifique.

2.2.13. Mostre que se A é uma matriz não singular tal que $A^2 = A$, então $\det(A) = 1$.

2.2.14. Mostre que se $A^k = \bar{0}$, para algum k inteiro positivo, então A é singular.

2.2.15. Mostre que se $A^t = A^{-1}$, então $\det(A) = \pm 1$;

2.2.16. Mostre que se α é um escalar e A é uma matriz $n \times n$, então $\det(\alpha A) = \alpha^n \det(A)$.

2.2.17. Mostre que A, $n \times n$, é invertível se, e somente se, $A^t A$ é invertível.

2.2.18. Sejam A e P matrizes $n \times n$, sendo P invertível. Mostre que $\det(P^{-1}AP) = \det(A)$.

2.2.19. Mostre que se uma matriz $A = (a_{ij})_{n\times n}$ é **triangular superior**, (isto é, os elementos situados abaixo da diagonal são iguais a zero) então $\det(A) = a_{11}a_{22}\ldots a_{nn}$.

2.2.20. (a) Mostre que se $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, então $\det(A) = 0$ se, e somente se, uma linha é múltiplo escalar da outra. E se A for uma matriz $n \times n$?

(b) Mostre que se uma linha A_i de uma matriz $A = (a_{ij})_{n\times n}$, é tal que $A_i = \alpha A_k + \beta A_l$, para α e β escalares e $i \neq k, l$, então $\det(A) = 0$.

(c) Mostre que se uma linha A_i de uma matriz $A = (a_{ij})_{n\times n}$, é tal que $A_i = \sum_{k\neq i} \alpha_k A_k$, para $\alpha_1, \ldots, \alpha_k$ escalares, então $\det(A) = 0$.

2.2.21. Mostre que o **determinante de Vandermonde** é dado por

$$V_n = \det \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & & \vdots & \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{bmatrix} = \prod_{i>j}(x_i - x_j).$$

A expressão à direita significa o produto de todos os termos $x_i - x_j$ tais que $i > j$ e $i, j = 1, \ldots, n$. (Sugestão: Mostre primeiro que $V_3 = (x_3 - x_2)(x_2 - x_1)(x_3 - x_1)$. Suponha que o resultado é verdadeiro para matrizes de Vandermonde de ordem $n-1$, mostre que o resultado é verdadeiro para matrizes de Vandermonde de ordem n. Faça as seguintes operações nas colunas da matriz, $-x_1 C_{i-1} + C_i \to C_i$, para $i = n, \ldots, 2$. Obtenha $V_n = (x_n - x_1) \ldots (x_2 - x_1) V_{n-1}$.)

2.2.22. Sejam A, B e D matrizes $p \times p$, $p \times (n-p)$ e $(n-p) \times (n-p)$, respectivamente. Mostre que

$$\det \begin{bmatrix} A & B \\ \bar{0} & D \end{bmatrix} = \det(A) \det(D).$$

(Sugestão: O resultado é claramente verdadeiro para $n = 2$. Suponha que o resultado seja verdadeiro para matrizes de ordem $n-1$. Desenvolva o determinante da matriz em termos da 1ª coluna, escreva o resultado em termos de determinantes de ordem $n-1$ e mostre que o resultado é verdadeiro para matrizes de ordem n.)

2.2.23. Dado um polinômio

$$p(t) = (-1)^n (t^n + a_{n-1} t^{n-1} + \cdots + a_0)$$

Verifique que a matriz

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}_{n \times n},$$

é tal que $\det(A - tI_n) = p(t)$. Esta matriz é chamada **matriz companheira** do polinômio $p(t)$. (Sugestão: verifique para $n = 2$ e depois supondo que seja verdade para matrizes $(n-1) \times (n-1)$ mostre que é verdade para matrizes $n \times n$ expandindo em cofatores em relação a primeira coluna)

Apêndice II: Demonstração do Teorema 2.11 na página 108

Demonstração do Teorema 2.10 na página 106 para $k > 1$. Deixamos como exercício para o leitor a verificação de que para matrizes 2×2 o resultado é verdadeiro. Supondo que o resultado seja verdadeiro para matrizes $(n-1) \times (n-1)$, vamos provar para matrizes $n \times n$. Sejam

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_{k-1} \\ \alpha X + \beta Y \\ A_{k+1} \\ \vdots \\ A_n \end{bmatrix}, \quad B = \begin{bmatrix} A_1 \\ \vdots \\ A_{k-1} \\ X \\ A_{k+1} \\ \vdots \\ A_n \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} A_1 \\ \vdots \\ A_{k-1} \\ Y \\ A_{k+1} \\ \vdots \\ A_n \end{bmatrix}.$$

Suponha que $k = 2, \ldots, n$. As matrizes $\tilde{A}_{1j}$, $\tilde{B}_{1j}$ e $\tilde{C}_{1j}$ só diferem na $(k-1)$-ésima linha (lembre-se que a primeira linha é retirada!). Além disso, a $(k-1)$-ésima linha de $\tilde{A}_{1j}$ é igual a α vezes a linha correspondente de $\tilde{B}_{1j}$ mais β vezes a linha correspondente de $\tilde{C}_{1j}$ (esta é a relação que vale para a k-ésima linha de A). Como estamos supondo o resultado verdadeiro para matrizes $(n-1) \times (n-1)$, então

$\det(\tilde{A}_{1j}) = \alpha \det(\tilde{B}_{1j}) + \beta \det(\tilde{C}_{1j})$. Assim,

$$\begin{aligned}
\det(A) &= \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det(\tilde{A}_{1j}) \\
&= \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\left[\alpha\det(\tilde{B}_{1j}) + \beta\det(\tilde{C}_{1j})\right] \\
&= \alpha\sum_{j=1}^{n}(-1)^{1+j}b_{1j}\det(\tilde{B}_{1j}) + \beta\sum_{j=1}^{n}(-1)^{1+j}c_{1j}\det(\tilde{C}_{1j}) \\
&= \alpha\det(B) + \beta\det(C),
\end{aligned}$$

pois $a_{1j} = b_{1j} = c_{1j}$, para $j = 1, \ldots, n$. ■

Lema 2.21. *Sejam $E_1 = [\,1\,0\,\ldots\,0\,], E_2 = [\,0\,1\,0\,\ldots\,0\,], \ldots, E_n = [\,0\,\ldots\,0\,1\,]$. Se A é uma matriz $n \times n$, cuja i-ésima linha é igual a E_k, para algum k ($1 \leq k \leq n$), então*

$$\det(A) = (-1)^{i+k}\det(\tilde{A}_{ik}).$$

Demonstração. É fácil ver que para matrizes 2×2 o lema é verdadeiro. Suponha que ele seja verdadeiro para matrizes $(n-1) \times (n-1)$ e vamos provar que ele é verdadeiro para matrizes $n \times n$. Podemos supor que $1 < i \leq n$.

Seja B_j a matriz $(n-2) \times (n-2)$ obtida de A eliminando-se as linhas 1 e i e as colunas j e k, para $1 \leq j \leq n$.

Para $j < k$, a matriz $\tilde{A}_{1j}$ é uma matriz $(n-1) \times (n-1)$ cuja $(i-1)$-ésima linha é igual a E_{k-1}. Para $j > k$, a matriz $\tilde{A}_{1j}$ é uma matriz $(n-1) \times (n-1)$ cuja $(i-1)$-ésima linha é igual a E_k. Como estamos supondo o lema verdadeiro para estas matrizes e como pelo Teorema 2.10 na página 106 se uma matriz tem uma linha nula o seu determinante é igual a zero, então $\det(\tilde{A}_{1k}) = 0$, segue-se que

$$\det(\tilde{A}_{1j}) = \begin{cases} (-1)^{(i-1)+(k-1)}\det(B_j) & \text{se } j < k, \\ 0 & \text{se } j = k, \\ (-1)^{(i-1)+k}\det(B_j) & \text{se } j > k. \end{cases} \tag{2.12}$$

Usando (2.12), obtemos

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det(\tilde{A}_{ij}) \\ &= \sum_{j<k}^{n}(-1)^{1+j}a_{1j}(-1)^{(i-1)+(k-1)}\det(B_j) + \sum_{j>k}^{n}(-1)^{1+j}a_{1j}(-1)^{(i-1)+k}\det(B_j) \end{aligned}$$

Por outro lado, temos que

$$(-1)^{i+k}\det(\tilde{A}_{ik}) = (-1)^{i+k}\left[\sum_{j<k}^{n}(-1)^{1+j}a_{1j}\det(B_j) + \sum_{j>k}^{n}(-1)^{1+(j-1)}a_{1j}\det(B_j)\right]$$

É simples a verificação de que as duas expressões acima são iguais. ■

Demonstração do Teorema 2.11 na página 108.

Pelo Teorema 2.14 na página 113 basta provarmos o resultado para o desenvolvimento em termos das linhas de A. Sejam

$$E_1 = [1\,0\,\ldots\,0], E_2 = [0\,1\,0\,\ldots\,0], \ldots, E_n = [0\,\ldots\,0\,1].$$

Observe que a linha i de A pode ser escrita como $A_i = \sum_{j=1}^{n} a_{ij} E_j$. Seja B_j a matriz obtida de A substituindo-se a linha i por E_j. Pelo Teorema 2.10 na página 106 e o Lema 2.21 segue-se que

$$\det(A) = \sum_{j=1}^{n} a_{ij} \det(B_j) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\tilde{A}_{ij}).$$

■

2.3 Matrizes Particionadas em Blocos (opcional)

2.3.1 Operações Matriciais em Blocos

As matrizes podem ser subdivididas em blocos, por exemplo a matriz

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

pode ser dividida em quatro submatrizes, A_{11}, A_{12}, A_{21} e A_{22},

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{rrr:rr} 5 & 4 & 3 & 4 & 4 \\ -3 & 3 & -1 & 2 & 4 \\ \hdashline 1 & 0 & 1 & -3 & -1 \\ 0 & -5 & 3 & -1 & 4 \end{array}\right].$$

Dependendo das subdivisões feitas, as matrizes podem ser operadas em termos dos seus blocos. Com relação à soma, duas matrizes podem ser somadas por blocos se os blocos correspondentes nas matrizes forem do mesmo tamanho, ou seja, se os blocos correspondentes podem ser somados. Vamos analisar a seguir quando podemos fazer o produto de matrizes em termos dos seus blocos.

Seja A uma matriz $m \times p$ e B uma matriz $p \times n$. Podemos particionar A e B em blocos e expressar o produto em termos de submatrizes de A e B. Considere os seguintes casos:

Caso 1:

Se $B = [\, B_1 \;\; B_2 \,]$, em que B_1 é uma matriz $p \times t$ e B_2 é uma matriz $p \times (n-t)$, então

$$AB = A\,[\, B_1 \;\; B_2 \,] = [\, AB_1 \;\; AB_2].$$

Caso 2:

Se $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$, em que A_1 é uma matriz $t \times p$ e A_2 é uma matriz $(m-t) \times p$, então

$$AB = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} B = \begin{bmatrix} A_1 B \\ A_2 B \end{bmatrix}.$$

Caso 3:

Se $A = [\, A_1 \;\; A_2 \,]$, em que A_1 é uma matriz $m \times t$ e A_2 é uma matriz $m \times (p-t)$ e $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$, em que B_1 é uma matriz $t \times n$ e B_2 é uma matriz $(p-t) \times n$, então

$$[AB]_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} = \sum_{k=1}^{t} a_{ik} b_{kj} + \sum_{k=t+1}^{p} a_{ik} b_{kj} = [A_1 B_1]_{ij} + [A_2 B_2]_{ij}.$$

Portanto,

$$AB = [\, A_1 \;\; A_2 \,] \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = A_1 B_1 + A_2 B_2.$$

Caso 4:

Sejam as matrizes A e B particionadas em blocos como segue:

$$A = \begin{array}{cc} \left[\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right] & \begin{array}{c} r \\ m-r \end{array} \\ \begin{array}{cc} t & p-t \end{array} & \end{array}, \qquad B = \begin{array}{cc} \left[\begin{array}{cc} B_{11} & B_{12} \\ B_{21} & B_{22} \end{array}\right] & \begin{array}{c} t \\ p-t \end{array} \\ \begin{array}{cc} s & n-s \end{array} & \end{array}$$

Sejam

$$A_1 = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}, \quad A_2 = \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix}, \quad B_1 = [\, B_{11} \; B_{12} \,] \quad \text{e} \quad B_2 = [\, B_{21} \; B_{22} \,].$$

Segue do Caso 3 que

$$AB = [\, A_1 \; A_2 \,] \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = A_1B_1 + A_2B_2.$$

Agora, segue-se dos Casos 1 e 2, que

$$A_1B_1 = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} B_1 = \begin{bmatrix} A_{11}B_1 \\ A_{21}B_1 \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} & A_{11}B_{12} \\ A_{21}B_{11} & A_{21}B_{12} \end{bmatrix}$$

$$A_2B_2 = \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} B_2 = \begin{bmatrix} A_{12}B_2 \\ A_{22}B_2 \end{bmatrix} = \begin{bmatrix} A_{12}B_{21} & A_{12}B_{22} \\ A_{22}B_{21} & A_{22}B_{22} \end{bmatrix}.$$

Portanto,

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

Observe que para que seja possível fazer o produto por blocos é necessário que o número de colunas dos blocos da primeira matriz seja igual ao número de linhas dos blocos correspondentes da segunda matriz. O que fizemos acima pode ser generalizado para um número qualquer de blocos. Se os blocos possuem os tamanhos adequados, a multiplicação por blocos pode ser feita da mesma forma como é feita a multiplicação usual de matrizes.

Exemplo 2.23. Sejam

$$A = \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \quad \text{e} \quad B = \left[\begin{array}{r|r} -2 & -1 \\ -3 & 4 \\ \hline -5 & 0 \\ 2 & -1 \end{array}\right].$$

Usando o particionamento das matrizes em blocos é mais simples fazer o produto AB.

$$AB = \left[\begin{array}{c|c} \bar{0}_2 \begin{bmatrix} -2 \\ -3 \end{bmatrix} + I_2 \begin{bmatrix} -5 \\ 2 \end{bmatrix} & \bar{0}_2 \begin{bmatrix} -1 \\ 4 \end{bmatrix} + I_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} \\ \hline I_2 \begin{bmatrix} -2 \\ -3 \end{bmatrix} + \bar{0}_2 \begin{bmatrix} -5 \\ 2 \end{bmatrix} & I_2 \begin{bmatrix} -1 \\ 4 \end{bmatrix} + \bar{0}_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} \end{array}\right] = \left[\begin{array}{r|r} -5 & 0 \\ 2 & -1 \\ \hline -2 & -1 \\ -3 & 4 \end{array}\right].$$

2.3.2 Inversa de Matrizes em Blocos

Proposição 2.22. *Sejam A e D matrizes $p \times p$ e $(n-p) \times (n-p)$, respectivamente. A matriz*

$$M = \begin{bmatrix} A & \bar{0} \\ \bar{0} & D \end{bmatrix}$$

é invertível se, e somente se, A e D são invertíveis. No caso em que M é invertível, então

$$M^{-1} = \begin{bmatrix} A^{-1} & \bar{0} \\ \bar{0} & D^{-1} \end{bmatrix} \tag{2.13}$$

Demonstração. Se A e D são invertíveis é fácil verificar que a matriz dada em (2.13) é a inversa de M.

Reciprocamente, suponha que M é invertível. Seja $N = M^{-1}$. Vamos particionar a matriz N da mesma maneira que M, ou seja,

$$N = \begin{bmatrix} E & F \\ G & H \end{bmatrix}.$$

Como $M\,N = N\,M = I_n$, então

$$\begin{bmatrix} A & \bar{0} \\ \bar{0} & D \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} I_p & \bar{0} \\ \bar{0} & I_{n-p} \end{bmatrix} = \begin{bmatrix} E & F \\ G & H \end{bmatrix} \begin{bmatrix} A & \bar{0} \\ \bar{0} & D \end{bmatrix}.$$

De em que segue-se que, $DH = I_{n-p} = HD$ e assim D é invertível. Além disso, $AE = I_p = EA$ e portanto A é invertível. ■

2.3.3 Determinante de Matrizes em Blocos

Proposição 2.23. *Sejam A, B e D matrizes $p \times p$, $p \times (n-p)$ e $(n-p) \times (n-p)$, respectivamente. Seja $M = \begin{bmatrix} A & B \\ \bar{0} & D \end{bmatrix}$. Então,*

$$\det(M) = \det \begin{bmatrix} A & B \\ \bar{0} & D \end{bmatrix} = \det(A)\det(D).$$

Demonstração. O resultado é claramente verdadeiro para $n = 4$. Supondo que o resultado seja verdadeiro para matrizes $(n-1) \times (n-1)$, vamos provar para matrizes $n \times n$. Expandindo o determinante em termos da 1ª coluna da matriz

$$M = \left[\begin{array}{cc} A & B \\ \bar{0} & D \end{array} \right],$$

obtemos

$$\det(M) = \sum_{i=1}^{p} (-1)^{i+1} a_{i1} \det(\tilde{M}_{i1})$$

Como estamos supondo o resultado verdadeiro para matrizes $(n-1) \times (n-1)$, então

$$\begin{aligned} \det(M) &= \sum_{i=1}^{p} (-1)^{i+1} a_{i1} (\det(\tilde{A}_{i1}) \det(D)) \\ &= \left(\sum_{i=1}^{p} (-1)^{i+1} a_{i1} \det(\tilde{A}_{i1}) \right) \det(D) = \det(A) \det(D). \end{aligned}$$

■

Exercícios Numéricos

2.3.1. Sejam

$$E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \left[\begin{array}{cc:cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hdashline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right]$$

Realize os seguintes produtos em blocos:

(a) $\begin{bmatrix} \bar{0}_2 & I_2 \\ I_2 & \bar{0}_2 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$;

(b) $\begin{bmatrix} E & \bar{0}_2 \\ \bar{0}_2 & I_2 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$;

(c) $\begin{bmatrix} I_2 & F \\ \bar{0}_2 & I_2 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$;

(d) $\begin{bmatrix} I_2 & \bar{0}_2 \\ E & I_2 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$.

Exercícios Teóricos

2.3.2. Seja A uma matriz invertível $n \times n$. Faça os seguintes produtos:

(a) $A^{-1} \begin{bmatrix} A & I_n \end{bmatrix}$

(b) $\begin{bmatrix} A & I_n \end{bmatrix}^t \begin{bmatrix} A & I_n \end{bmatrix}$

(c) $\begin{bmatrix} A & I_n \end{bmatrix} \begin{bmatrix} A & I_n \end{bmatrix}^t$

(d) $\begin{bmatrix} A \\ I_n \end{bmatrix} A^{-1}$

(e) $\begin{bmatrix} A^{-1} \\ I_n \end{bmatrix} \begin{bmatrix} A & I_n \end{bmatrix}$.

2.3.3. Seja A uma matriz $m \times n$, e suponha que $A = XY^t$, em que X é uma matriz $m \times k$ e Y é uma matriz $n \times k$. Escreva $X = [\, X_1 \ldots X_k \,]$ e $Y = [\, Y_1 \ldots Y_k \,]$. Escreva A como uma soma de k matrizes definidas em termos das colunas de X e de Y.

2.3.4. Sejam A, B e D matrizes $p \times p$, $p \times (n-p)$ e $(n-p) \times (n-p)$, respectivamente. Mostre que a matriz

$$M = \begin{bmatrix} A & B \\ \bar{0} & D \end{bmatrix}$$

é invertível se, e somente se, A e D são invertíveis. No caso em que M é invertível, então

$$M^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ \bar{0} & D^{-1} \end{bmatrix}$$

(Sugestão: mostre que $\begin{bmatrix} A & B \\ \bar{0} & D \end{bmatrix} \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ \bar{0} & D^{-1} \end{bmatrix} = \begin{bmatrix} I_n & \bar{0} \\ \bar{0} & I_{n-p} \end{bmatrix}$.)

2.3.5. Sejam A, B e D matrizes $p \times p$, $p \times (n-p)$ e $(n-p) \times (n-p)$, respectivamente. Seja

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

em que A é invertível. Mostre que M pode ser fatorada no produto

$$M = \begin{bmatrix} I_p & \bar{0} \\ \tilde{C} & I_{n-p} \end{bmatrix} \begin{bmatrix} A & B \\ \bar{0} & \tilde{D} \end{bmatrix},$$

em que $\tilde{C} = CA^{-1}$ e $\tilde{D} = D - CA^{-1}B$.
(Sugestão: faça o produto e verifique que é igual a M.)

2.3.6. Sejam A, B, C e D matrizes $n \times n$, com A invertível. Seja

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

(a) Mostre que

$$\det(M) = \det(AD - ACA^{-1}B).$$

(Sugestão: use a Proposição 2.23 na página 143 e o exercício anterior.)

(b) Se $AC = CA$, mostre que

$$\det(M) = \det(AD - CB)$$

(c) Em [23] Exercício 7.77 na página 396 é pedido para provar que se $AC = CA$, então

$$\det(M) = \det(A)\det(D) - \det(B)\det(C).$$

Mostre que este resultado é **falso**. Sugestão: use o item anterior, tome $A = C$ e D, B tais que $\det(D - B) \neq \det(D) - \det(B)$.

Teste do Capítulo

---

1. Calcule o determinante da matriz seguinte usando operações elementares para transformá-la em uma matriz triangular superior.

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

---

2. Se possível, encontre a inversa da seguinte matriz:

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

---

3. Encontre todos os valores de λ para os quais a matriz $A - \lambda I_4$ tem inversa, em que

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

---

4. Responda **Verdadeiro** ou **Falso**, justificando:

(a) Se $A^2 = -2A^4$, então $(I + A^2)^{-1} = I - 2A^2$;

(b) Se $A^t = -A^2$ e A é não singular, então determinante de A é -1;

(c) Se $B = AA^tA^{-1}$, então $\det(A) = \det(B)$.

(d) $\det(A + B) = \det A + \det B$

3

Espaços $\mathbb{R}^n$

3.1 Vetores no Plano e no Espaço

Muitas grandezas físicas, como velocidade, força, deslocamento e impulso, para serem completamente identificadas, precisam, além da magnitude, da direção e do sentido. Estas grandezas são chamadas **grandezas vetoriais** ou simplesmente **vetores**.

Geometricamente, vetores são representados por **segmentos (de retas) orientados** (segmentos de retas com um sentido de percurso) no plano ou no espaço. A ponta da seta do segmento orientado é chamada **ponto final ou extremidade** e o outro

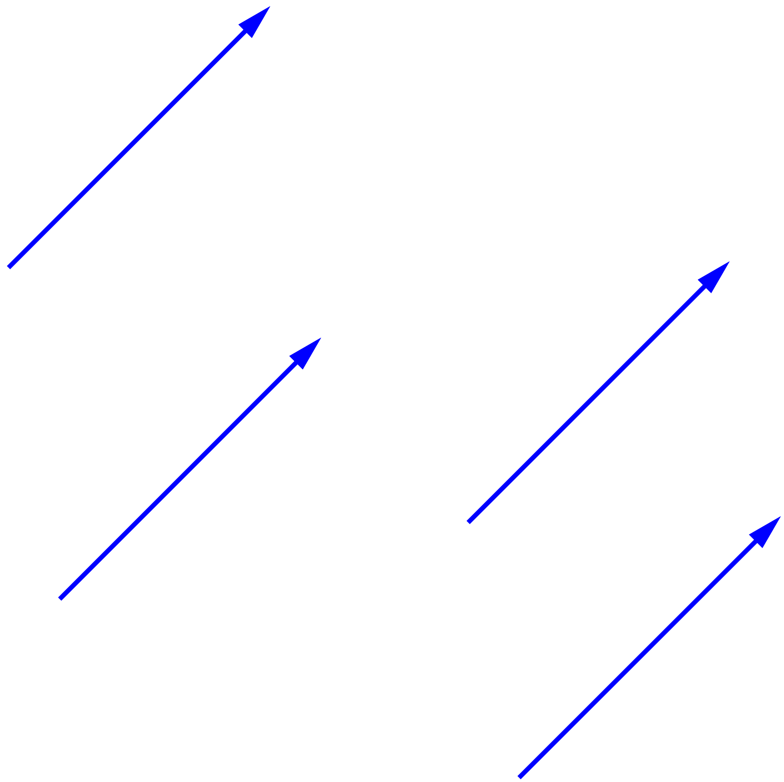


Figura 3.1. Segmentos orientados representando o mesmo vetor

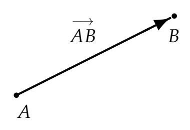
ponto extremo é chamado de **ponto inicial ou origem** do segmento orientado. A direção e o sentido do segmento orientado identifica a direção e o sentido do vetor. O comprimento do segmento orientado representa a magnitude do vetor.

Um vetor poder ser representado por vários segmentos orientados. Este fato é análogo ao que ocorre com os números racionais e as frações. Duas frações representam o mesmo número racional se o numerador e o denominador de cada uma delas estiverem na mesma proporção. Por exemplo, as frações $1/2$, $2/4$ e $3/6$ representam o mesmo número racional. De forma análoga, dizemos que dois segmentos orientados representam o mesmo vetor se possuem o mesmo comprimento, a mesma direção e o mesmo sentido. A definição de igualdade de vetores também é análoga a igualdade de números racionais. Dois números racionais a/b e c/d são iguais, quando $ad = bc$. Analogamente, dizemos que dois vetores são iguais se eles possuem o mesmo comprimento, a mesma direção e o mesmo sentido.

Na Figura 3.1 temos 4 segmentos orientados, com origens em pontos diferentes, que representam o mesmo vetor, ou seja, são considerados como vetores iguais, pois possuem a mesma direção, mesmo sentido e o mesmo comprimento.

Se o ponto inicial de um representante de um vetor V é A e o ponto final é B, então escrevemos

$$V = \overrightarrow{AB}$$



3.1.1 Soma de Vetores e Multiplicação por Escalar

A soma, $V + W$, de dois vetores V e W é determinada da seguinte forma:

- tome um segmento orientado que representa V;

- tome um segmento orientado que representa W, com origem na extremidade de V;
- o vetor $V + W$ é representado pelo segmento orientado que vai da origem de V até a extremidade de W.

Da Figura 3.2, deduzimos que a soma de vetores é comutativa, ou seja,

$$V + W = W + V, \tag{3.1}$$

para quaisquer vetores V e W. Observamos também que a soma $V + W$ está na diagonal do paralelogramo determinado por V e W, quando estão representados com a mesma origem.

Da Figura 3.3, deduzimos que a soma de vetores é associativa, ou seja,

$$V + (W + U) = (V + W) + U, \tag{3.2}$$

para quaisquer vetores V, W e U.

O vetor que tem a sua origem coincidindo com a sua extremidade é chamado **vetor nulo** e denotado por $\bar{0}$. Segue então, que

$$V + \bar{0} = \bar{0} + V = V, \tag{3.3}$$

para todo vetor V.

Para qualquer vetor V, o **simétrico** de V, denotado por $-V$, é o vetor que tem mesmo comprimento, mesma direção e sentido contrário ao de V. Segue então, que

$$V + (-V) = \bar{0}. \tag{3.4}$$

Definimos a **diferença** W **menos** V, por

$$W - V = W + (-V).$$

Segue desta definição, de (3.1), (3.2), (3.4) e de (3.3) que

$$W + (V - W) = (V - W) + W = V + (-W + W) = V + \bar{0} = V.$$

Assim, a diferença $V - W$ é um vetor que somado a W dá V, portanto ele vai da extremidade de W até a extremidade de V, desde que V e W estejam representados por segmentos orientados com a mesma origem.

A **multiplicação de um vetor** V **por um escalar** α, $\alpha\, V$, é determinada pelo vetor que possui as seguintes características:

(a) é o vetor nulo, se $\alpha = 0$ ou $V = \bar{0}$,

(b) caso contrário,

   i. tem comprimento $|\alpha|$ vezes o comprimento de V,

   ii. a direção é a mesma de V (neste caso, dizemos que eles são **paralelos**),

   iii. tem o mesmo sentido de V, se $\alpha > 0$ e
   tem o sentido contrário ao de V, se $\alpha < 0$.

As propriedades da multiplicação por escalar serão apresentadas mais a frente. Se $W = \alpha\, V$, dizemos que W é **um múltiplo escalar** de V. É fácil ver que dois vetores não nulos são paralelos (ou **colineares**) se, e somente se, um é um múltiplo escalar do outro.

As operações com vetores podem ser definidas utilizando um **sistema de coordenadas retangulares**. Em primeiro lugar, vamos considerar os vetores no plano.

Seja V um vetor no plano. Definimos as **componentes de** V como sendo as coordenadas (v_1, v_2) do ponto final do representante de V que tem ponto inicial na origem. Escrevemos simplesmente

$$\boxed{V = (v_1, v_2).}$$

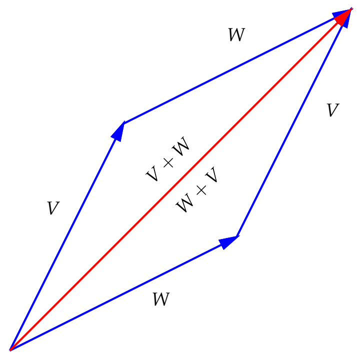


Figura 3.2. $V + W = W + V$

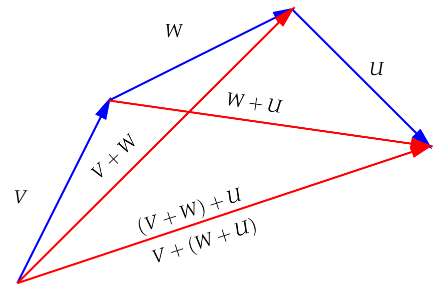


Figura 3.3. $V + (W + U) = (V + W) + U$

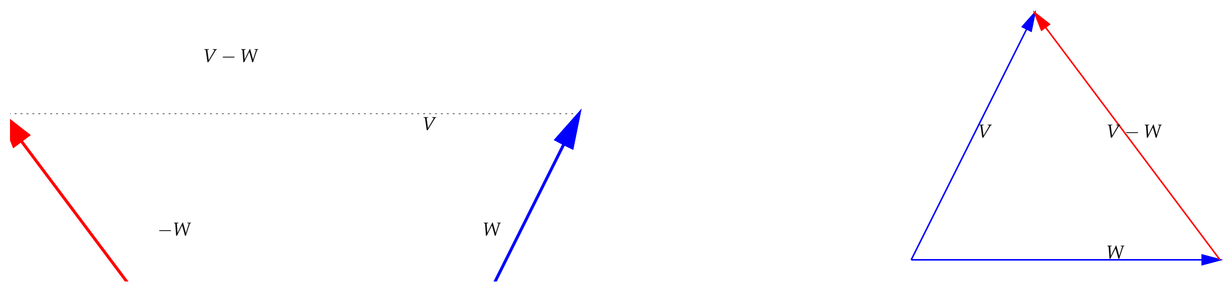


Figura 3.4. A diferença $V - W$

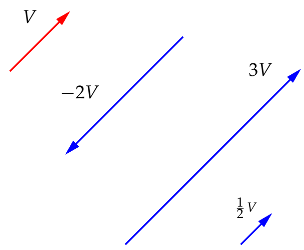


Figura 3.5. Multiplicação de vetor por escalar

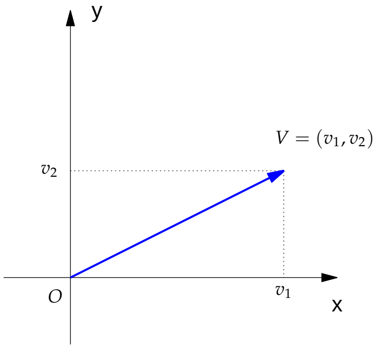


Figura 3.6. As componentes do vetor V no plano

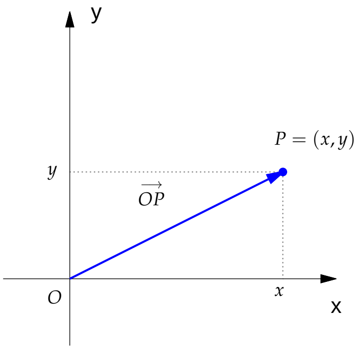


Figura 3.7. As coordenadas de P são iguais as componentes de $\overrightarrow{OP}$

Assim, as coordenadas de um ponto P são iguais as componentes do vetor $\overrightarrow{OP}$, que vai da origem do sistema de coordenadas ao ponto P. Em particular, o vetor nulo, $\bar{0} = (0, 0)$. Em termos das componentes, podemos realizar facilmente as operações: soma de vetores e multiplicação de vetor por escalar.

- Como ilustrado na Figura 3.8, a **soma** de dois vetores $V = (v_1, v_2)$ e $W = (w_1, w_2)$ é dada por

$$\boxed{V + W = (v_1 + w_1, v_2 + w_2);}$$

- Como ilustrado na Figura 3.9, a **multiplicação** de um vetor $V = (v_1, v_2)$ por um escalar α é dada por

$$\boxed{\alpha\, V = (\alpha\, v_1, \alpha\, v_2).}$$

Definimos as **componentes de um vetor no espaço** de forma análoga a que fizemos com vetores no plano. Vamos inicialmente introduzir um sistema de coordenadas retangulares no espaço. Para isto, escolhemos um ponto como origem O e como eixos coordenados, três retas orientadas (com sentido de percurso definido), passando pela origem, perpendiculares entre si. Estes serão os eixos x, y e z. O eixo z é o eixo vertical. Os eixos x e y são horizontais e satisfazem a seguinte propriedade. Suponha que giramos o eixo x pelo menor ângulo até que coincida com o eixo y. Se os dedos da mão direita apontam na direção do semi-eixo x positivo de forma que o semi-eixo y positivo esteja do lado da palma da mão, então o polegar aponta no sentido do semi-eixo z positivo. Cada par de eixos determina um plano chamado de **plano coordenado**. Portanto os três planos coordenados são: xy, yz e xz.

A cada ponto P no espaço associamos um terno de números (x, y, z), chamado de **coordenadas do ponto** P como segue.

- Passe três planos por P paralelos aos planos coordenados.

- A interseção do plano paralelo ao plano xy, passando por P, com o eixo z determina a coordenada z.
- A interseção do plano paralelo ao plano xz, passando por P, com o eixo y determina a coordenada y
- A interseção do plano paralelo ao plano yz, passando por P, com o eixo x determina a coordenada x.

Alternativamente, podemos encontrar as coordenadas de um ponto P como segue.

- Trace uma reta paralela ao eixo z, passando por P;
- A interseção da reta paralela ao eixo z, passando por P, com o plano xy é o ponto P'. As coordenadas de P', (x, y), no sistema de coordenadas xy são as duas primeiras coordenadas de P.
- A terceira coordenada é igual ao comprimento do segmento PP', se P estiver acima do plano xy e ao comprimento do segmento PP' com o sinal negativo, se P estiver abaixo do plano xy.

Agora, estamos prontos para utilizarmos um sistema de coordenadas cartesianas também nas operações de vetores no espaço. Seja V um vetor no espaço. Como no caso de vetores do plano, definimos as **componentes de** V como sendo as coordenadas (v_1, v_2, v_3) do ponto final do representante de V que tem ponto inicial na origem. Escrevemos simplesmente

$$\boxed{V = (v_1, v_2, v_3).}$$

Assim, as coordenadas de um ponto P são iguais as componentes do vetor $\overrightarrow{OP}$ que vai da origem do sistema de coordenadas ao ponto P. Em particular, o vetor nulo, $\bar{0} = (0, 0, 0)$. Assim como fizemos para vetores no plano, para vetores no espaço a soma de vetores e a multiplicação de vetor por escalar podem ser realizadas em termos das componentes.

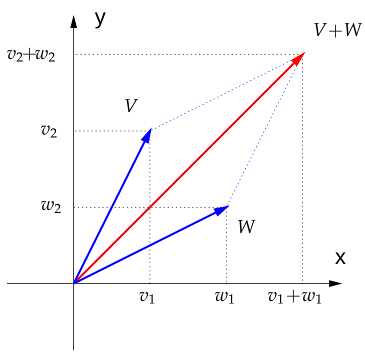


Figura 3.8. A soma de dois vetores no plano

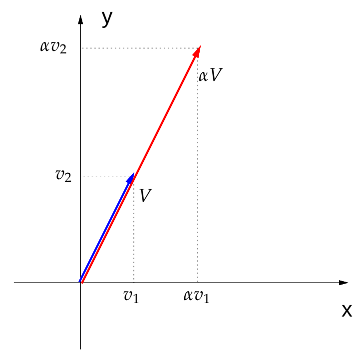


Figura 3.9. A multiplicação de vetor por escalar no plano

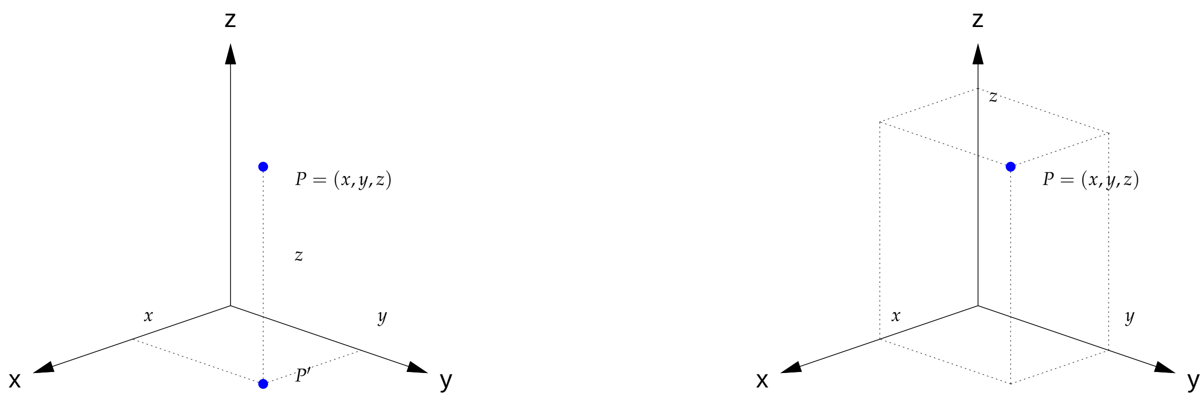


Figura 3.10. As coordenadas de um ponto no espaço

- Se $V = (v_1, v_2, v_3)$ e $W = (w_1, w_2, w_3)$, então a adição de V com W é dada por

$$\boxed{V + W = (v_1 + w_1, v_2 + w_2, v_3 + w_3);}$$

- Se $V = (v_1, v_2, v_3)$ e α é um escalar, então a multiplicação de V por α é dada por

$$\boxed{\alpha\, V = (\alpha\, v_1, \alpha\, v_2, \alpha\, v_3).}$$

Exemplo 3.1. Se $V = (1, -2, 3)$, $W = (2, 4, -1)$, então

$$V + W = (1 + 2, -2 + 4, 3 + (-1)) = (3, 2, 2), \quad 3V = (3 \cdot 1, 3\,(-2), 3 \cdot 3) = (3, -6, 9).$$

Quando um vetor V está representado por um segmento orientado com ponto inicial fora da origem (Figura 3.13), digamos em $P = (x_1, y_1, z_1)$, e ponto final em $Q = (x_2, y_2, z_2)$, então as componentes do vetor V são dadas por

$$\boxed{V = \overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = (x_2 - x_1, y_2 - y_1, z_2 - z_1).}$$

Portanto, as componentes de V são obtidas subtraindo-se as coordenadas do ponto Q (extremidade) das do ponto P (origem). O mesmo se aplica a vetores no plano.

Exemplo 3.2. As componentes do vetor V que tem um representante com ponto inicial $P = (5/2, 1, 2)$ e ponto final $Q = (0, 5/2, 5/2)$ são dadas por

$$V = \overrightarrow{PQ} = (0 - 5/2, 5/2 - 1, 5/2 - 2) = (-5/2, 3/2, 1/2).$$

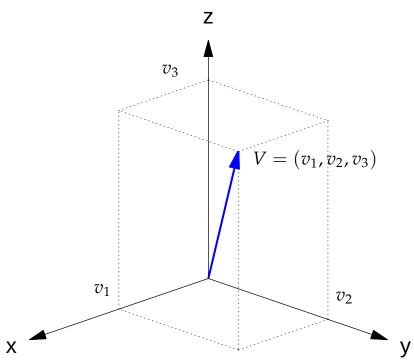


Figura 3.11. As componentes de um vetor no espaço

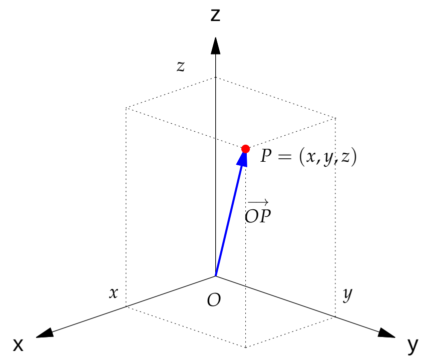


Figura 3.12. As coordenadas de P são iguais as componentes de $\overrightarrow{OP}$

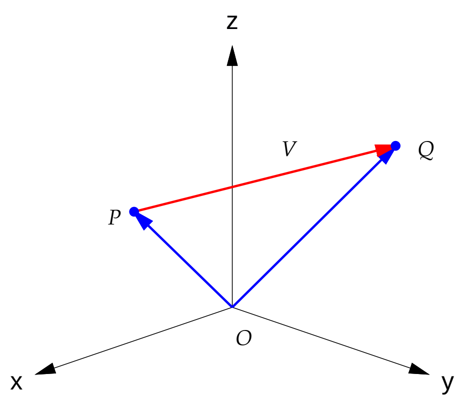


Figura 3.13. $V = \overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$

Observação. O vetor é "livre", ele não tem posição fixa, ao contrário do ponto e do segmento orientado. Por exemplo, o vetor $V = (-5/2, 3/2, 1/2)$, no exemplo acima, estava representado por um segmento orientado com a origem no ponto $P = (5/2, 1, 2)$. Mas, poderia ser representado por um segmento orientado cujo ponto inicial poderia estar em qualquer outro ponto.

Um vetor no espaço $V = (v_1, v_2, v_3)$ pode também ser escrito na notação matricial como uma **matriz linha** ou como uma **matriz coluna**:

$$V = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad \text{ou} \quad V = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}.$$

Estas notações podem ser justificadas pelo fato de que as operações matriciais

$$V + W = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{bmatrix}, \quad \alpha V = \alpha \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \\ \alpha v_3 \end{bmatrix}$$

ou

$$V + W = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} + \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} = \begin{bmatrix} v_1 + w_1 & v_2 + w_2 & v_3 + w_3 \end{bmatrix},$$

$$\alpha V = \alpha \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} \alpha v_1 & \alpha v_2 & \alpha v_3 \end{bmatrix}$$

produzem os mesmos resultados que as operações vetoriais

$$V + W = (v_1, v_2, v_3) + (w_1, w_2, w_3) = (v_1 + w_1, v_2 + w_2, v_3 + w_3),$$

$$\alpha V = \alpha(v_1, v_2, v_3) = (\alpha v_1, \alpha v_2, \alpha v_3).$$

O mesmo vale, naturalmente, para vetores no plano.

No teorema seguinte enunciamos as propriedades mais importantes da soma de vetores e multiplicação de vetores por escalar.

Teorema 3.1. *Sejam U, V e W vetores e α e β escalares. São válidas as seguintes propriedades:*

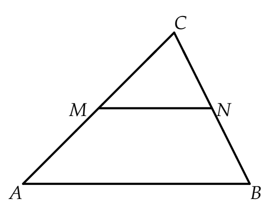
(a) $U + V = V + U$;
(b) $(U + V) + W = U + (V + W)$;
(c) $U + \bar{0} = U$;
(d) $U + (-U) = \bar{0}$;
(e) $\alpha(\beta U) = (\alpha\beta)U$;
(f) $\alpha(U + V) = \alpha U + \alpha V$;
(g) $(\alpha + \beta)U = \alpha U + \beta U$;
(h) $1U = U$.

Demonstração. Segue diretamente das propriedades da álgebra matricial (Teorema 1.1 na página 9). ■

Exemplo 3.3. Vamos usar vetores e as suas propriedades para provar um resultado conhecido de geometria plana. Seja um triângulo ABC e sejam M e N os pontos médios de AC e BC, respectivamente. Vamos provar que MN é paralelo a AB e tem comprimento igual a metade do comprimento de AB.

Devemos provar que

$$\overrightarrow{MN} = \frac{1}{2} \overrightarrow{AB} .$$



Agora, a partir da figura ao lado temos que

$$\overrightarrow{MN} = \overrightarrow{MC} + \overrightarrow{CN}\,.$$

Como M é ponto médio de AC e N é ponto médio de BC, então

$$\overrightarrow{MC} = \frac{1}{2}\,\overrightarrow{AC} \quad \text{e} \quad \overrightarrow{CN} = \frac{1}{2}\,\overrightarrow{CB}\,.$$

Logo,

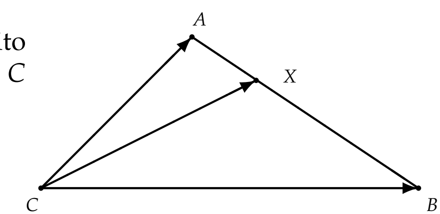
$$\overrightarrow{MN} = \frac{1}{2}\,\overrightarrow{AC} + \frac{1}{2}\,\overrightarrow{CB} = \frac{1}{2}(\overrightarrow{AC} + \overrightarrow{CB}) = \frac{1}{2}\,\overrightarrow{AB}\,.$$

Exemplo 3.4. Dados quatro pontos A, B, C e X tais que $\overrightarrow{AX} = \lambda\,\overrightarrow{AB}$, vamos escrever $\overrightarrow{CX}$ como uma soma de múltiplos escalares de $\overrightarrow{CA}$ e $\overrightarrow{CB}$, que é chamada **combinação linear** de $\overrightarrow{CA}$ e $\overrightarrow{CB}$.

Como $\overrightarrow{AX} = \lambda\,\overrightarrow{AB}$, então os vetores $\overrightarrow{AX}$ e $\overrightarrow{AB}$ são paralelos e portanto o ponto X só pode estar na reta definida por A e B. Vamos desenhá-lo entre A e B, mas isto não vai representar nenhuma restrição.

O vetor que vai de C para X, pode ser escrito como uma soma de um vetor que vai de C para A com um vetor que vai de A para X,

$$\overrightarrow{CX} = \overrightarrow{CA} + \overrightarrow{AX}\,.$$



Agora, por hipótese $\overrightarrow{AX} = \lambda\,\overrightarrow{AB}$, o que implica que $\overrightarrow{CX} = \overrightarrow{CA} + \lambda\,\overrightarrow{AB}$.
Mas, $\overrightarrow{AB} = \overrightarrow{CB} - \overrightarrow{CA}$, portanto $\overrightarrow{CX} = \overrightarrow{CA} + \lambda(\overrightarrow{CB} - \overrightarrow{CA})$. Logo,

$$\overrightarrow{CX} = (1-\lambda)\,\overrightarrow{CA} + \lambda\,\overrightarrow{CB}\,.$$

Observe que para $\lambda = 0$, $\overrightarrow{CX} = \overrightarrow{CA}$, para $\lambda = 1$, $\overrightarrow{CX} = \overrightarrow{CB}$, para $\lambda = 1/2$, $\overrightarrow{CX} = \frac{1}{2}\overrightarrow{CA} + \frac{1}{2}\overrightarrow{CB}$, para $\lambda = 1/3$, $\overrightarrow{CX} = \frac{2}{3}\overrightarrow{CA} + \frac{1}{3}\overrightarrow{CB}$.

Exemplo 3.5. Vamos mostrar, usando vetores, que o ponto médio de um segmento que une os pontos $A = (x_1, y_1, z_1)$ e $B = (x_2, y_2, z_2)$ é $M = \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}, \frac{z_1+z_2}{2}\right)$.

O ponto M é o ponto médio de AB se, e somente se, $\overrightarrow{AM} = \frac{1}{2}\overrightarrow{AB}$. Então, aplicando o exemplo anterior (com o ponto C sendo a origem O), $\overrightarrow{OM} = \frac{1}{2}\overrightarrow{OA} + \frac{1}{2}\overrightarrow{OB}$. Como as coordenadas de um ponto são iguais as componentes do vetor que vai da origem até aquele ponto, segue-se que $\overrightarrow{OM} = \frac{1}{2}(x_1, y_1, z_1) + \frac{1}{2}(x_2, y_2, z_2)$ e

$$M = \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}, \frac{z_1+z_2}{2}\right).$$

3.1.2 Norma e Produto Escalar

Já vimos que o **comprimento** de um vetor V é definido como sendo o comprimento de qualquer um dos segmentos orientados que o representam. O comprimento do vetor V também é chamado de **norma de** V e é denotado(a) por $||V||$. Segue do Teorema de Pitágoras que a norma de um vetor é dada por

$$\boxed{||V|| = \sqrt{v_1^2 + v_2^2},}$$

no caso em que $V = (v_1, v_2)$ é um vetor no plano, e por

$$\boxed{||V|| = \sqrt{v_1^2 + v_2^2 + v_3^2},}$$

no caso em que $V = (v_1, v_2, v_3)$ é um vetor no espaço (verifique usando as Figuras 3.14 e 3.15).

Um vetor de norma igual a 1 é chamado **vetor unitário**.

A **distância entre dois pontos** $P = (x_1, y_1, z_1)$ **e** $Q = (x_2, y_2, z_2)$ é igual à norma do vetor $\overrightarrow{PQ}$ (Figura 3.13 na página 161). Como

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = (x_2 - x_1, y_2 - y_1, z_2 - z_1),$$

então a distância de P a Q é dada por

$$\boxed{\text{dist}(P, Q) = ||\overrightarrow{PQ}|| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.}$$

Analogamente, a **distância entre dois pontos** $P = (x_1, y_1)$ **e** $Q = (x_2, y_2)$ no plano é igual à norma do vetor $\overrightarrow{PQ}$, que é dada por

$$\boxed{\text{dist}(P, Q) = ||\overrightarrow{PQ}|| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.}$$

Exemplo 3.6. A norma do vetor $V = (1, -2, 3)$ é

$$||V|| = \sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{14}.$$

A distância entre os pontos $P = (2, -3, 1)$ e $Q = (-1, 4, 5)$ é

$$\text{dist}(P, Q) = ||\overrightarrow{PQ}|| = ||(-1 - 2, 4 - (-3), 5 - 1)|| = ||(-3, 7, 4)|| = \sqrt{(-3)^2 + 7^2 + 4^2} = \sqrt{74}.$$

Se $V = (v_1, v_2, v_3)$ e α é um escalar, então da definição da multiplicação de vetor por escalar e da norma de um vetor segue-se que

$$||\alpha V|| = ||(\alpha v_1, \alpha v_2, \alpha v_3)|| = \sqrt{(\alpha v_1)^2 + (\alpha v_2)^2 + (\alpha v_3)^2} = \sqrt{\alpha^2(v_1^2 + v_2^2 + v_3^2)},$$

ou seja,

$$\boxed{||\alpha V|| = |\alpha|\, ||V||.} \tag{3.5}$$

Dado um vetor V **não nulo**, o vetor

$$\boxed{U = \left(\frac{1}{||V||}\right) V.}$$

é um **vetor unitário na direção de** V, pois por (3.5), temos que

$$||U|| = \left|\frac{1}{||V||}\right| ||V|| = 1.$$

Exemplo 3.7. Um vetor unitário na direção do vetor $V = (1, -2, 3)$ é o vetor

$$U = \left(\frac{1}{||V||}\right) V = \left(\frac{1}{\sqrt{14}}\right)(1, -2, 3) = (\frac{1}{\sqrt{14}}, \frac{-2}{\sqrt{14}}, \frac{3}{\sqrt{14}}).$$

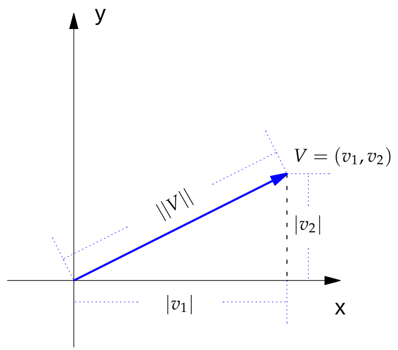


Figura 3.14. A norma de um vetor V no plano

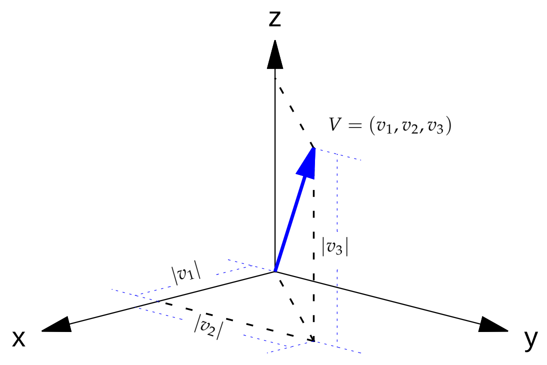


Figura 3.15. A norma de um vetor V no espaço

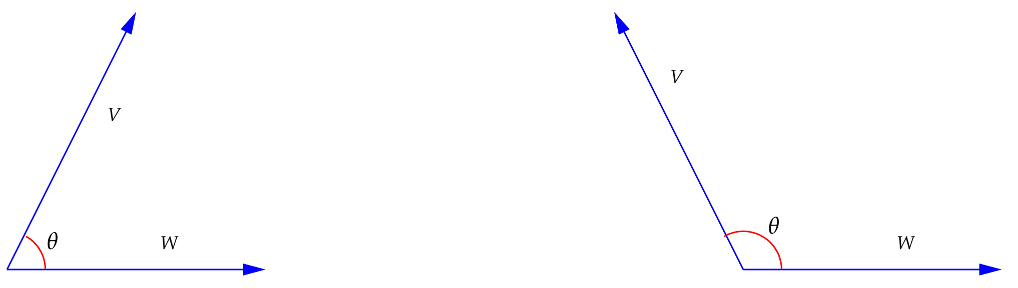


Figura 3.16. Ângulo entre dois vetores, agudo (à esquerda) e obtuso (à direita)

O ângulo entre dois vetores não nulos, V e W, é definido pelo ângulo θ determinado por V e W que satisfaz $0 \leq \theta \leq \pi$, quando eles estão representados com a mesma origem.

Quando o ângulo θ entre dois vetores V e W é reto ($\theta = 90^{\text{o}}$), ou um deles é o vetor nulo, dizemos que os vetores V e W são **ortogonais** ou **perpendiculares entre si**.

Vamos definir, agora, um produto entre dois vetores, cujo resultado é um escalar. Por isso ele é chamado **produto escalar**. Este produto tem aplicação, por exemplo, em Física: o trabalho realizado por uma força é o produto escalar do vetor força pelo vetor deslocamento, quando a força aplicada é constante.

Definição 3.1. O **produto escalar** ou **interno** de dois vetores V e W é definido por

$$V \cdot W = \begin{cases} 0, & \text{se } V \text{ ou } W \text{ é o vetor nulo,} \\ ||V||\,||W|| \cos\theta, & \text{caso contrário,} \end{cases}$$

em que θ é o ângulo entre eles.

Quando os vetores são dados em termos das suas componentes não sabemos diretamente o ângulo entre eles. Por isso, precisamos de uma forma de calcular o produto escalar que não necessite do ângulo entre os vetores.

Se V e W são dois vetores não nulos e θ é o ângulo entre eles, então pela lei dos cossenos,

$$||V - W||^2 = ||V||^2 + ||W||^2 - 2||V||\,||W|| \cos\theta.$$

Assim,

$$V \cdot W = ||V||\,||W|| \cos\theta = \frac{1}{2}\left(||V||^2 + ||W||^2 - ||V - W||^2\right). \tag{3.6}$$

Já temos então uma fórmula para calcular o produto escalar que não depende diretamente do ângulo entre eles. Substituindo-se as coordenadas dos vetores em (3.6) obtemos uma expressão mais simples para o cálculo do produto interno.

Por exemplo, se $V = (v_1, v_2, v_3)$ e $W = (w_1, w_2, w_3)$ são vetores no espaço, então substituindo-se $||V||^2 = v_1^2 + v_2^2 + v_3^2$, $||W||^2 = w_1^2 + w_2^2 + w_3^2$ e $||V - W||^2 = (v_1 - w_1)^2 + (v_2 - w_2)^2 + (v_3 - w_3)^2$ em (3.6) os termos v_i^2 e w_i^2 são cancelados e obtemos

$$V \cdot W = v_1 w_1 + v_2 w_2 + v_3 w_3.$$

Teorema 3.2. *O produto escalar ou interno, $V \cdot W$, entre dois vetores é dado por*

$$V \cdot W = v_1 w_1 + v_2 w_2,$$

se $V = (v_1, v_2)$ e $W = (w_1, w_2)$ são vetores no plano e por

$$V \cdot W = v_1 w_1 + v_2 w_2 + v_3 w_3,$$

se $V = (v_1, v_2, v_3)$ e $W = (w_1, w_2, w_3)$ são vetores no espaço.

Exemplo 3.8. Sejam $V = (0, 1, 0)$ e $W = (2, 2, 3)$. O produto escalar de V por W é dado por

$$V \cdot W = v_1 w_1 + v_2 w_2 + v_3 w_3 = 0 \cdot 2 + 1 \cdot 2 + 0 \cdot 3 = 2\,.$$

Podemos usar o Teorema 3.2 para determinar o ângulo entre dois vetores não nulos, V e W. O cosseno do ângulo entre V e W é, então, dado por

$$\boxed{\cos\theta = \frac{V \cdot W}{||V||\,||W||}\,.}$$

Se V e W são vetores não nulos e θ é o ângulo entre eles, então

(a) θ é agudo ($0 \le \theta < 90^o$) se, e somente se, $V \cdot W > 0$,

(b) θ é reto ($\theta = 90^o$) se, e somente se, $V \cdot W = 0$ e

(c) θ é obtuso ($90^o < \theta \leq 180^o$) se, e somente se, $V \cdot W < 0$.

Exemplo 3.9. Vamos determinar o ângulo entre uma diagonal de um cubo e uma de suas arestas. Sejam $V_1 = (1,0,0), V_2 = (0,1,0)$ e $V_3 = (0,0,1)$ (Figura 3.18). Uma diagonal do cubo é representada pelo vetor D dado por

$$D = V_1 + V_2 + V_3 = (1,1,1)\,.$$

Então o ângulo entre D e V_1 satisfaz

$$\cos\theta = \frac{D \cdot V_1}{||D||||V_1||} = \frac{1.1 + 0.1 + 0.1}{(\sqrt{1^2+1^2+1^2})(\sqrt{1^2+0^2+0^2})} = \frac{1}{\sqrt{3}}$$

ou seja,

$$\theta = \arccos(\frac{1}{\sqrt{3}}) \approx 54^o\,.$$

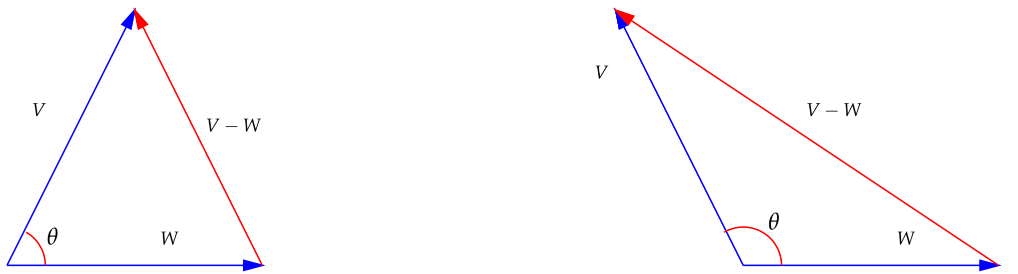


Figura 3.17. Triângulo formado por representantes de V, W e $V - W$. À esquerda o ângulo entre V e W é agudo e à direita é obtuso.

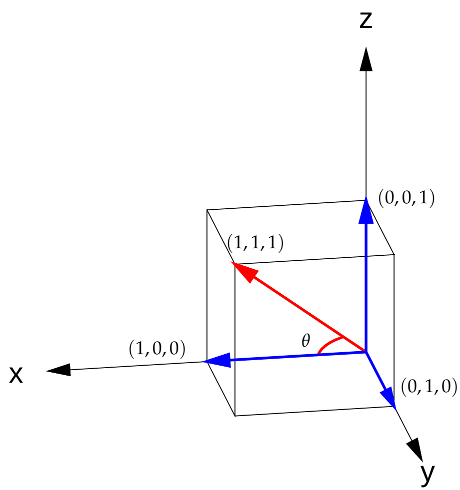


Figura 3.18. Ângulo entre a diagonal de um cubo e uma de suas arestas

Teorema 3.3. *Sejam U, V e W vetores e α um escalar. São válidas as seguintes propriedades:*

(a) *(comutatividade)* $U \cdot V = V \cdot U$;

(b) *(distributividade)* $U \cdot (V + W) = U \cdot V + U \cdot W$;

(c) *(associatividade)* $\alpha(U \cdot V) = (\alpha U) \cdot V = U \cdot (\alpha V)$;

(d) $V \cdot V = ||V||^2 \geq 0$, *para todo* V *e* $V \cdot V = 0$ *se, e somente se,* $V = \bar{0}$.

Demonstração. Sejam $U = (u_1, u_2, u_3)$, $V = (v_1, v_2, v_3)$ e $W = (w_1, w_2, w_3)$.

(a) $U \cdot V = u_1v_1 + u_2v_2 + u_3v_3 = v_1u_1 + v_2u_2 + v_3u_3 = V \cdot U$;

(b) $U \cdot (V + W) = (u_1, u_2, u_3) \cdot (v_1 + w_1, v_2 + w_2, v_3 + w_3) =$
$= u_1(v_1 + w_1) + u_2(v_2 + w_2) + u_3(v_3 + w_3) =$
$= (u_1v_1 + u_1w_1) + (u_2v_2 + u_2w_2) + (u_3v_3 + u_3w_3) =$
$= (u_1v_1 + u_2v_2 + u_3v_3) + (u_1w_1 + u_2w_2 + u_3w_3) = U \cdot V + U \cdot W$;

(c) $\alpha(U \cdot V) = \alpha(u_1v_1 + u_2v_2 + u_3v_3) = (\alpha u_1)v_1 + (\alpha u_2)v_2 + (\alpha u_3)v_3 = (\alpha U) \cdot V$;

(d) $V \cdot V = ||V||^2$ é uma soma de quadrados, por isso é sempre maior ou igual a zero e é zero se, e somente se, todas as parcelas são iguais a zero. ■

Projeção Ortogonal

Dados dois vetores V e W a **projeção ortogonal de** V **sobre** W denotada por

$$\text{proj}_W V$$

é o vetor que é paralelo a W tal que $V - \mathrm{proj}_W V$ seja ortogonal a W (Figura 3.19).

Proposição 3.4. *Seja W um vetor não nulo. Então, a projeção ortogonal de um vetor V em W é dada por*

$$\text{proj}_W V = \left(\frac{V \cdot W}{||W||^2}\right) W\,.$$

Demonstração. Sejam $V_1 = \text{proj}_W V$ e $V_2 = V - \text{proj}_W V$. Como V_1 é paralelo a W, então

$$V_1 = \alpha W. \tag{3.7}$$

Assim,

$$V_2 = V - \alpha W\,.$$

Multiplicando-se escalarmente V_2 por W e usando o Teorema 3.3 (d) obtemos

$$V_2 \cdot W = (V - \alpha W) \cdot W = V \cdot W - \alpha ||W||^2. \tag{3.8}$$

Mas, V_2 é ortogonal a W, então $V_2 \cdot W = 0$. Portanto, de (3.8) obtemos

$$\alpha = \frac{V \cdot W}{||W||^2}\,.$$

Substituindo este valor de α na equação (3.7) segue-se o resultado. ■

Exemplo 3.10. Sejam $V = (2, -1, 3)$ e $W = (4, -1, 2)$. Vamos encontrar dois vetores V_1 e V_2 tais que $V = V_1 + V_2$, V_1 é paralelo a W e V_2 é perpendicular a W (Figura 3.19). Temos que

$$V \cdot W = 2 \cdot 4 + (-1)(-1) + 3 \cdot 2 = 15$$
$$||W||^2 = 4^2 + (-1)^2 + 2^2 = 21\,.$$

$$V_1 = \mathrm{proj}_W V = \left(\frac{V \cdot W)}{||W||^2}\right) W = \left(\frac{15}{21}\right)(4,-1,2) = (\frac{20}{7}, -\frac{5}{7}, \frac{10}{7})$$

$$V_2 = V - V_1 = (2,-1,3) - (\frac{20}{7}, -\frac{5}{7}, \frac{10}{7}) = (-\frac{6}{7}, -\frac{2}{7}, \frac{11}{7}).$$

Exercícios Numéricos (respostas na página 504)

3.1.1. Determine o vetor X, tal que $3X - 2V = 15(X - U)$.

3.1.2. Determine o vetor X, tal que $\left\{\begin{array}{rcrcl} 6X & - & 2Y & = & U \\ 3X & + & Y & = & U + V \end{array}\right.$

3.1.3. Determine as coordenadas da extremidade do segmento orientado que representa o vetor $V = (3, 0, -3)$, sabendo-se que sua origem está no ponto $P = (2, 3, -5)$.

3.1.4. Quais são as coordenadas do ponto P', simétrico do ponto $P = (1, 0, 3)$ em relação ao ponto $M = (1, 2, -1)$? (Sugestão: o ponto P' é tal que o vetor $\overrightarrow{MP'} = -\overrightarrow{MP}$)

3.1.5. Verifique se os pontos dados a seguir são **colineares**, isto é, pertencem a uma mesma reta:

(a) $A = (5, 1, -3)$, $B = (0, 3, 4)$ e $C = (0, 3, -5)$;

(b) $A = (-1, 1, 3)$, $B = (4, 2, -3)$ e $C = (14, 4, -15)$;

3.1.6. Dados os pontos $A = (1, -2, -3)$, $B = (-5, 2, -1)$ e $C = (4, 0, -1)$. Determine o ponto D tal que A, B, C e D sejam vértices consecutivos de um paralelogramo.

3.1.7. Verifique se o vetor U é combinação linear (soma de múltiplos escalares) de V e W:

(a) $V = (9, -12, -6)$, $W = (-1, 7, 1)$ e $U = (-4, -6, 2)$;

(b) $V = (5, 4, -3)$, $W = (2, 1, 1)$ e $U = (-3, -4, 1)$;

3.1.8. Sejam $V = (1, 2, -3)$ e $W = (2, 1, -2)$. Determine vetores unitários paralelos aos vetores (a) $V + W$; (b) $V - W$; (c) $2V - 3W$.

3.1.9. Ache o vetor unitário da bissetriz do ângulo entre os vetores $V = (2, 2, 1)$ e $W = (6, 2, -3)$. (Sugestão: observe que a soma de dois vetores está na direção da bissetriz se, e somente se, os dois tiverem o mesmo comprimento. Portanto, tome múltiplos escalares de V e W de forma que eles tenham o mesmo comprimento e tome o vetor unitário na direção da soma deles.)

3.1.10. Determine o valor de x para o qual os vetores $V = (x, 3, 4)$ e $W = (3, 1, 2)$ são perpendiculares.

3.1.11. Demonstre que não existe x tal que os vetores $V = (x, 2, 4)$ e $W = (x, -2, 3)$ são perpendiculares.

3.1.12. Ache o ângulo entre os seguintes pares de vetores:

(a) $(2, 1, 0)$ e $(0, 1, -1)$; (b) $(1, 1, 1)$ e $(0, -2, -2)$; (c) $(3, 3, 0)$ e $(2, 1, -2)$.

3.1.13. Decomponha $W = (-1, -3, 2)$ como a soma de dois vetores W_1 e W_2, com W_1 paralelo ao vetor $(0, 1, 3)$ e W_2 ortogonal a este último. (Sugestão: revise o Exemplo 3.10 na página 176)

3.1.14. Sabe-se que o vetor X é ortogonal a $(1, 1, 0)$ e a $(-1, 0, 1)$, tem norma $\sqrt{3}$ e sendo θ o ângulo entre X e $(0, 1, 0)$, tem-se $\cos\theta > 0$. Ache X.

3.1.15. Mostre que $A = (3, 0, 2)$, $B = (4, 3, 0)$ e $C = (8, 1, -1)$ são vértices de um triângulo retângulo. Em qual dos vértices está o ângulo reto?

Exercícios usando o MATLAB®

Comandos do MATLAB®:

`>> V=[v1,v2,v3]` cria um vetor `V`, usando as componentes numéricas `v1, v2, v3`. Por exemplo `>> V=[1,2,3]` cria o vetor $V = (1, 2, 3)$;

`>> V+W` é a soma de `V` e `W`; `>> V-W` é a diferença `V` menos `W`; `>> num*V` é o produto do vetor `V` pelo escalar `num`;

`>> subs(expr,x,num)` substitui `x` por `num` na expressão `expr`;

`>> solve(expr)` determina a solução da equação `expr=0`;

Comandos numéricos do pacote GAAL:

`>> V=randi(1,3)` cria um vetor aleatório com componentes inteiras;

`>> no(V)` calcula a norma do vetor `V`.

`>> pe(V,W)` calcula o produto escalar do vetor `V` pelo vetor `W`.

Comandos gráficos do pacote GAAL:

`>> desvet(P,V)` desenha o vetor `V` com origem no ponto `P` e `>> desvet(V)` desenha o vetor `V` com origem no ponto $O = (0,0,0)$.

`>> po([P1;P2;...;Pn])` desenha os pontos `P1, P2, ..., Pn`.

`>> lineseg(P1,P2,'cor')` desenha o segmento de reta `P1P2`.

`>> eixos` desenha os eixos coordenados.

`>> box` desenha uma caixa em volta da figura.

`>> axiss` reescala os eixos com a mesma escala.

`>> rota` faz uma rotação em torno do eixo z.

`>> zoom3(fator)` amplifica a região pelo `fator`.

`>> tex(P,'texto')` coloca o `texto` no ponto `P`.

3.1.16. Digite no prompt

`demog21,`

(sem a vírgula!). Esta função demonstra as funções gráficas para vetores.

3.1.17. Coloque em duas variáveis V e W dois vetores do plano ou do espaço a seu critério

(a) Use a função `ilsvw(V,W)` para visualizar a soma dos dois vetores.

(b) Coloque em uma variável `a` um número e use a função `ilav(a,V)` para visualizar a multiplicação do vetor `V` pelo escalar `a`.

(c) Use a função `ilpv(V,W)` para visualizar o produto vetorial $V \times W$.

(d) Use a função `ilproj(W,V)` para visualizar a projeção de V em W.

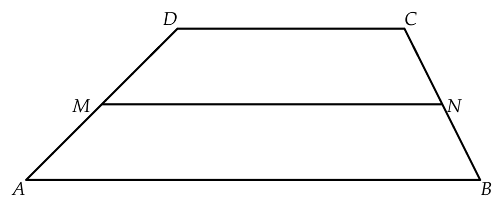
3.1.18. Use o MATLAB® para resolver os **Exercícios Numéricos** a partir do Exercício 1.3.

Exercícios Teóricos

3.1.19. Demonstre que o segmento que une os pontos médios dos lados não paralelos de um trapézio é paralelo às bases, e sua medida é a média aritmética das medidas das bases. (Sugestão: mostre que

$$\overrightarrow{MN} = \frac{1}{2}(\overrightarrow{AB} + \overrightarrow{DC})$$

e depois conclua que $\overrightarrow{MN}$ é um múltiplo escalar de $\overrightarrow{AB}$. Revise o Exemplo 3.3 na página 163)



3.1.20. Demonstrar que as diagonais de um paralelogramo se cortam ao meio. (Sugestão: Sejam M e N os pontos médios das duas diagonais do paralelogramo. Mostre que o vetor $\overrightarrow{MN} = \bar{0}$, então conclua que $M = N$.)

3.1.21. Sejam A, B e C pontos quaisquer com $A \neq B$. Prove que:

(a) Um ponto X pertence a reta determinada por A e B ($\overrightarrow{AX} = \lambda \overrightarrow{AB}$) se, e somente se,

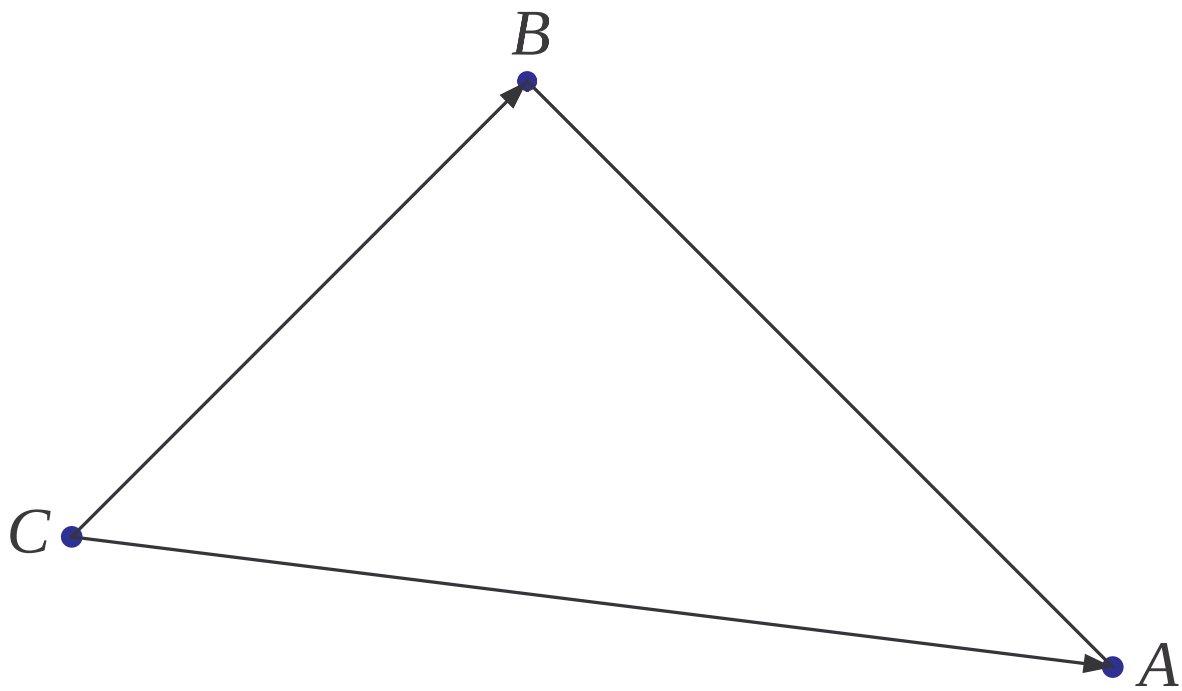
$$\overrightarrow{CX} = \alpha \overrightarrow{CA} + \beta \overrightarrow{CB}, \quad \text{com} \quad \alpha + \beta = 1.$$

(b) Um ponto X pertence ao interior do segmento AB ($\overrightarrow{AX} = \lambda \overrightarrow{AB}$, com $0 < \lambda < 1$) se, e somente se,

$$\overrightarrow{CX} = \alpha \overrightarrow{CA} + \beta \overrightarrow{CB}, \quad \text{com} \quad \alpha > 0,\ \beta > 0 \quad \text{e} \quad \alpha + \beta = 1.$$

(c) Um ponto X é um ponto interior ao triângulo ABC ($\overrightarrow{A'X} = \lambda\ \overrightarrow{A'B'}$, com $0 < \lambda < 1$, em que A' é um ponto interior ao segmento AC e B' é interior ao segmento CB) se, e somente se,

$$\overrightarrow{CX} = \alpha\ \overrightarrow{CA} + \beta\ \overrightarrow{CB}, \quad \text{com} \quad \alpha > 0,\ \beta > 0 \quad \text{e} \quad \alpha + \beta < 1.$$



3.1.22. Mostre que se $\alpha V = \bar{0}$, então $\alpha = 0$ ou $V = \bar{0}$.

3.1.23. Se $\alpha U = \alpha V$, então $U = V$? E se $\alpha \neq 0$?

3.1.24. Se $\alpha V = \beta V$, então $\alpha = \beta$? E se $V \neq \bar{0}$?

3.1.25. Mostre que $2V = V + V$.

3.1.26. Se $V \cdot W = V \cdot U$, então $W = U$?

3.1.27. Mostre que se V é ortogonal a W_1 e W_2, então V é ortogonal a $\alpha_1 W_1 + \alpha_2 W_2$.

3.1.28. Demonstre que as diagonais de um losango são perpendiculares. (Sugestão: mostre que $\overrightarrow{AC} \cdot \overrightarrow{BD} = 0$, usando o fato de que $\overrightarrow{AB} = \overrightarrow{DC}$ e $||\overrightarrow{AB}|| = ||\overrightarrow{BC}||$.)

3.1.29. Demonstre que, se V e W são vetores quaisquer, então:

(a) $V \cdot W = \dfrac{1}{4}\left(||V + W||^2 - ||V - W||^2\right)$;

(b) $||V||^2 + ||W||^2 = \dfrac{1}{2}\left(||V + W||^2 + ||V - W||^2\right)$.

(Sugestão: desenvolva os segundos membros das igualdades acima observando que $||V + W||^2 = (V + W) \cdot (V + W)$ e $||V - W||^2 = (V - W) \cdot (V - W)$)

3.1.30. Demonstre que se V e W são vetores quaisquer, então:

(a) $|V \cdot W| \leq ||V||\,||W||$;

(b) $||V + W|| \leq ||V|| + ||W||$;
(Sugestão: mostre que $||V + W||^2 = (V + W) \cdot (V + W) \leq (||V|| + ||W||)^2$, usando o item anterior)

(c) $\Big|\, ||V|| - ||W|| \,\Big| \leq ||V - W||$.
(Sugestão: defina $U = V - W$ e aplique o item anterior a U e W)

3.1.31. Sejam U_1, U_2 e U_3 três vetores unitários mutuamente ortogonais. Se $A = [\ U_1\ U_2\ U_3\ ]$ é uma matriz 3×3 cujas colunas são os vetores U_1, U_2 e U_3, então A é invertível e $A^{-1} = A^t$. (Sugestão: mostre que $A^t A = I_3$.)

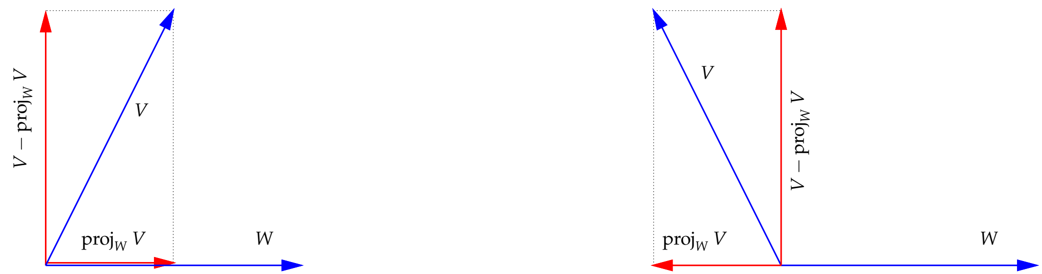


Figura 3.19. Projeção ortogonal do vetor V sobre o vetor W

3.2 Equações de Retas e Planos

3.2.1 Equação do Plano

Existe uma analogia entre uma reta no plano e um plano no espaço. No plano, a equação de uma reta é determinada se forem dados sua inclinação e um de seus pontos. No espaço, a inclinação de um plano é dada por um vetor perpendicular a ele e a equação de um plano é determinada se são dados um vetor perpendicular a ele e um de seus pontos.

Proposição 3.5. *A equação de um plano π que passa por um ponto $P_0 = (x_0, y_0, z_0)$ e é perpendicular ao vetor $N = (a, b, c)$ é*

$$ax + by + cz + d = 0\,, \tag{3.9}$$

*onde $d = -(ax_0 + by_0 + cz_0)$. A equação (3.9) é chamada **equação geral** do plano π e o vetor N é chamado **vetor normal** do plano.*

Demonstração. Um ponto $P = (x, y, z)$ pertence ao plano π se, e somente se, o vetor $\overrightarrow{P_0P}$ for perpendicular ao vetor N, ou seja,

$$N \cdot \overrightarrow{P_0P} = 0\,. \tag{3.10}$$

Como, $\overrightarrow{P_0P} = (x - x_0, y - y_0, z - z_0)$, a equação (3.10) pode ser reescrita como

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0,$$

ou seja,

$$ax + by + cz - (ax_0 + by_0 + cz_0) = 0\,.$$

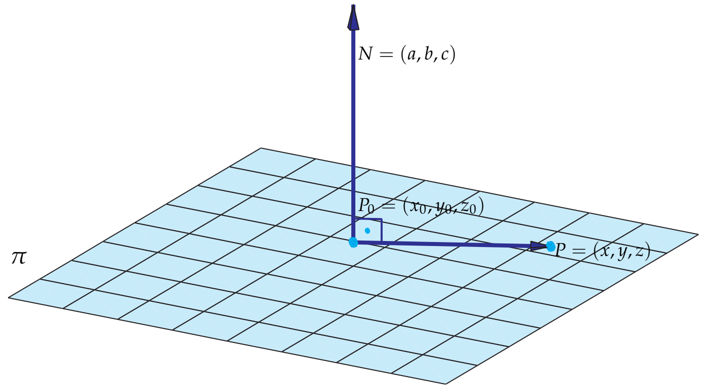


Figura 3.20. Plano perpendicular a $N = (a, b, c)$ e que passa por $P_0 = (x_0, y_0, z_0)$

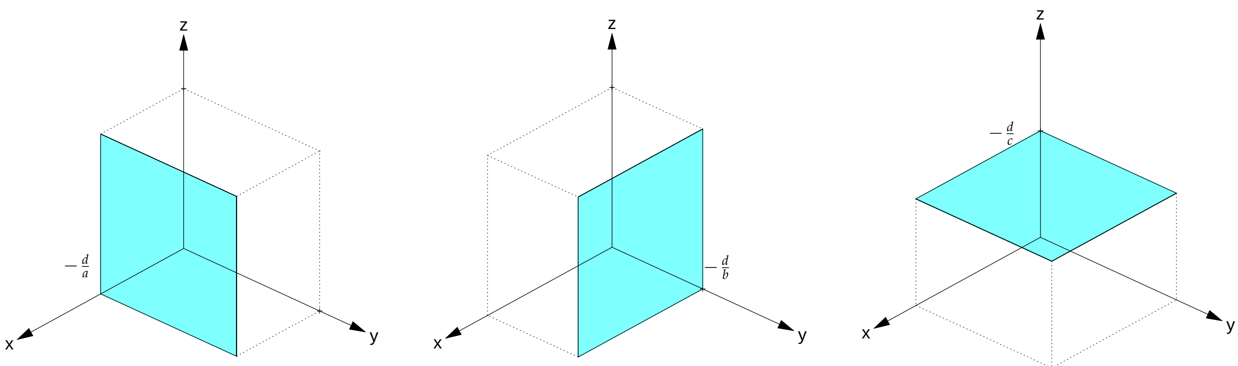


Figura 3.21. Planos $ax + d = 0$, $by + d = 0$ e $cz + d = 0$

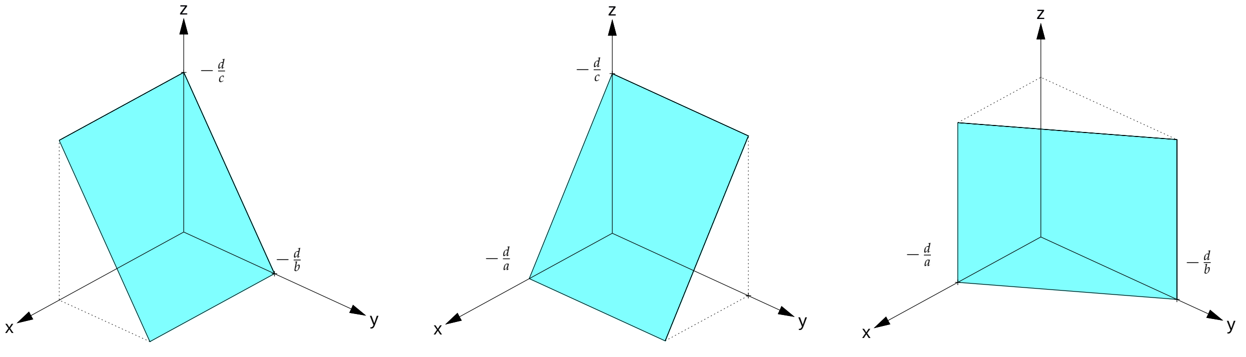


Figura 3.22. Planos $by + cz + d = 0$, $ax + cz + d = 0$ e $ax + by + d = 0$

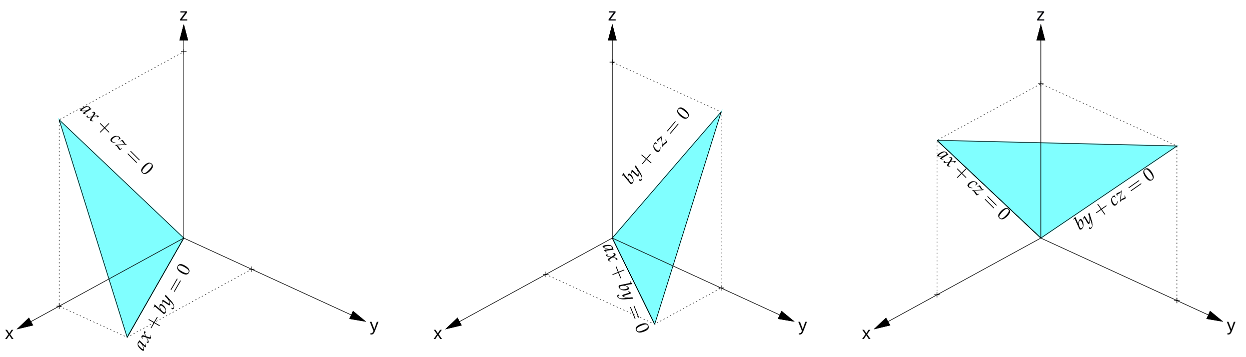


Figura 3.23. Planos $ax + by + cz = 0$

■

Exemplo 3.11. Vamos encontrar a equação do plano π que passa pelo ponto $P_0 = (3, -1, 7)$ e é perpendicular ao vetor $N = (4, 2, -5)$. Da proposição anterior, a equação do plano é da forma

$$ax + by + cz + d = 0\,,$$

onde os coeficientes de x, y e z são as componentes do vetor normal, ou seja, $a = 4$, $b = 2$ e $c = -5$. Assim, a equação de π é da forma

$$4x + 2y - 5z + d = 0\,.$$

Para determinar o coeficiente d, basta usarmos o fato de que $P_0 = (3, -1, 7)$ pertence a π. Mas, o ponto P_0 pertence a π se, e somente se, as suas coordenadas satisfazem a equação de π, ou seja,

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

De onde tiramos que $d = -12 + 2 + 35 = 25$. Finalmente, a equação do plano π é

$$4x + 2y - 5z + 25 = 0\,.$$

No plano, a equação de uma reta é determinada se forem dados dois pontos da reta. Analogamente, no espaço, a equação de um plano é determinada se são dados três pontos P_1, P_2 e P_3 não colineares (isto é, não pertencentes a uma mesma reta). Com os três pontos podemos "formar" os vetores $V = \overrightarrow{P_1P_2}$ e $W = \overrightarrow{P_1P_3}$ (Figura 3.25).

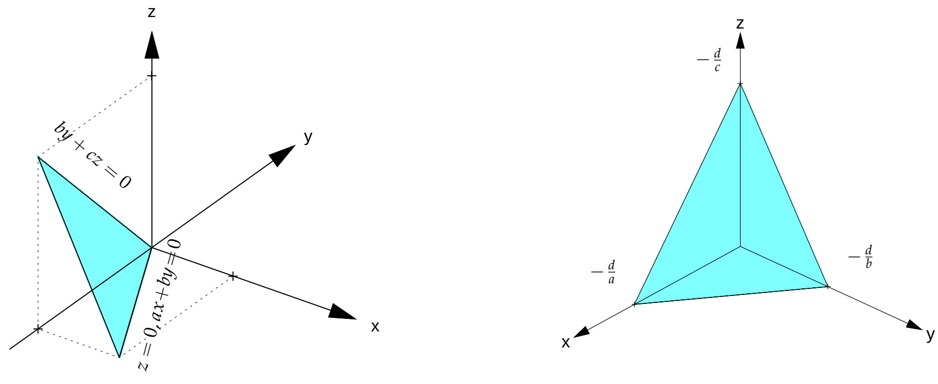


Figura 3.24. Planos $ax + by + cz = 0$ e $ax + by + cz + d = 0$

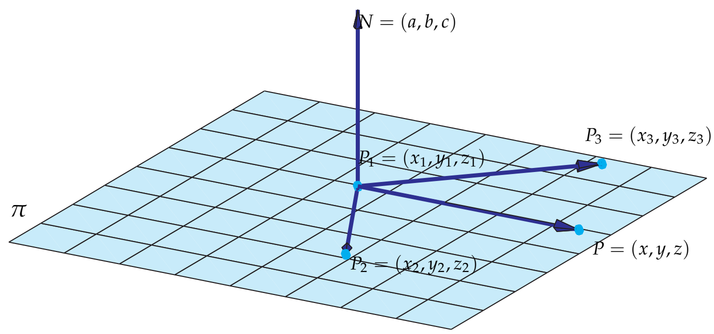


Figura 3.25. Plano que passa por três pontos

Exemplo 3.12. Vamos encontrar a equação do plano π que passa pelos pontos $P_1 = (1,2,-1)$, $P_2 = (2,3,1)$ e $P_3 = (3,-1,2)$. Com os três pontos podemos "formar" os vetores $V = \overrightarrow{P_1P_2} = (v_1, v_2, v_3)$ e $W = \overrightarrow{P_1P_3} = (w_1, w_2, w_3)$. O vetor normal do plano, $N = (a,b,c)$, é ortogonal a estes três vetores. De onde obtemos um sistema homogêneo com três equações ($N \cdot \overrightarrow{P_1P} = 0$, $N \cdot \overrightarrow{P_1P_2} = 0$ e $N \cdot \overrightarrow{P_1P_3} = 0$) e três incógnitas ($a$, b e c), que tem solução não trivial se, e somente se, o determinante da matriz do sistema é igual a zero (Teorema 2.15 na página 115), ou seja, se, e somente se,

$$\det \begin{bmatrix} x - x_1 & y - y_1 & z - z_1 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} = 0\,. \tag{3.11}$$

Mas,

$$\overrightarrow{P_1P} = (x-1, y-2, z-(-1)),$$

$$V = \overrightarrow{P_1P_2} = (1,1,2),$$

$$W = \overrightarrow{P_1P_3} = (2,-3,3).$$

Então, a equação do plano é

$$\det \begin{bmatrix} x-1 & y-2 & z+1 \\ 1 & 1 & 2 \\ 2 & -3 & 3 \end{bmatrix} = 9(x-1) + (y-2) - 5(z+1) = 9x + y - 5z - 16 = 0\,.$$

A equação do plano também é determinada se ao invés de serem dados três pontos, forem dados um ponto P_1 do plano e dois vetores paralelos ao plano, $V = (v_1, v_2, v_3)$ e $W = (w_1, w_2, w_3)$, desde que eles sejam não colineares. Ou ainda se forem dados dois pontos P_1 e P_2 do plano e um vetor paralelo ao plano $V = (v_1, v_2, v_3)$, já que neste caso podemos formar o vetor $W = \overrightarrow{P_1P_2} = (w_1, w_2, w_3)$ que é também paralelo ao plano.

Temos três vetores paralelos ao plano: $\overrightarrow{P_1P}= (x - x_1, y - y_1, z - z_1)$, V e W. O vetor normal do plano, $N = (a, b, c)$, é ortogonal a estes três vetores. De onde obtemos um sistema homogêneo com três equações ($N\cdot \overrightarrow{P_1P}= 0$, $N \cdot V = 0$ e $N \cdot W = 0$) e três incógnitas (a, b e c), que tem solução não trivial se, e somente se, o determinante da matriz do sistema é igual a zero (Teorema 2.15 na página 115), ou seja, se, e somente se,

$$\det \begin{bmatrix} x - x_1 & y - y_1 & z - z_1 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} = 0\,. \tag{3.12}$$

Assim, um ponto $P = (x, y, z)$ pertence a um plano π que passa pelo ponto $P_1 = (x_1, y_1, z_1)$ e é paralelo aos vetores $V = (v_1, v_2, v_3)$ e $W = (w_1, w_2, w_3)$ (não paralelos) se, e somente se, a equação (3.12) é verdadeira.

Observação. Não faz sentido dizer que um vetor pertence a um plano. Pois, por um lado, um plano é um conjunto de pontos e por outro, os vetores são “livres”, podem ser “colocados” em qualquer ponto. O correto é dizer que um vetor é paralelo a um plano.

3.2.2 Equações da Reta

Vamos supor que uma reta r é paralela a um vetor $V = (a, b, c)$ não nulo e que passa por um ponto $P_0 = (x_0, y_0, z_0)$. Um ponto $P = (x, y, z)$ pertence a reta r se, e somente se, o vetor $\overrightarrow{P_0P}$ é paralelo ao vetor V, isto é, se o vetor $\overrightarrow{P_0P}$ é um múltiplo escalar de V, ou seja,

$$\overrightarrow{P_0P}= t\, V\,. \tag{3.13}$$

Em termos de componentes, (3.13) pode ser escrito como

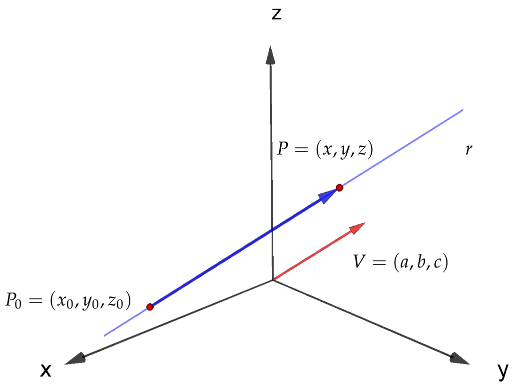
$$(x - x_0, y - y_0, z - z_0) = (ta, tb, tc)\,,$$

de onde segue-se que $x - x_0 = t\,a$, $y - y_0 = t\,b$ e $z - z_0 = t\,c$. Isto prova o resultado seguinte.

Proposição 3.6. *As equações*

$$\left\{\begin{array}{lcl} x & = & x_0 + t\,a \\ y & = & y_0 + t\,b \\ z & = & z_0 + t\,c \end{array}\right. \quad \text{para todo } t \in \mathbb{R} \tag{3.14}$$

são de uma reta r que passa por um ponto $P_0 = (x_0, y_0, z_0)$ e é paralela ao vetor $V = (a, b, c)$. As equações (3.14) são chamadas ***equações paramétricas da reta*** *r. O vetor $V = (a, b, c)$ é chamado* ***vetor diretor da reta*** *r.*



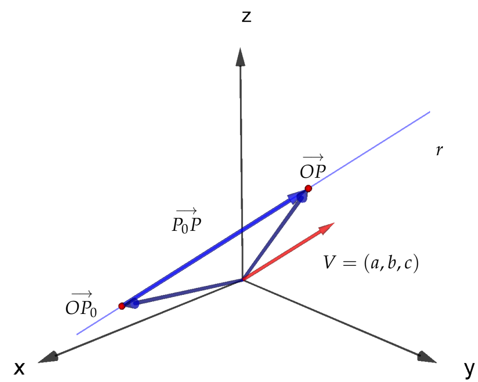


Figura 3.26. Reta paralela ao vetor $V = (a, b, c)$

O parâmetro t pode ser interpretado como o instante de tempo, se o ponto $P = (x, y, z)$ descreve o movimento de uma partícula em movimento retilíneo uniforme com vetor velocidade $V = (a, b, c)$. Observe que para $t = 1$, $P = (x, y, z) = (x_0 + a, y_0 + b, z_0 + c)$, para $t = 2$, $P = (x, y, z) = (x_0 + 2a, y_0 + 2b, z_0 + 2c)$ e assim por diante.

As equações (3.14), podem ser reescritas como $(x, y, z) = (x_0 + at, y_0 + bt, z_0 + ct)$.

Observação. Não faz sentido dizer que o vetor está contido na reta. Por um lado, a reta é um conjunto de pontos e por outro um vetor não tem posição fixa.

Exemplo 3.13. A reta que passa por $P_0 = (-3, 3/2, 4)$ e é paralela ao vetor $V = (-6, 1, 4)$ tem equações paramétricas

$$r : \begin{cases} x = -3 - 6\,t \\ y = \frac{3}{2} + t \\ z = 4 + 4t \end{cases} \quad \text{para } t \in \mathbb{R}$$

Podemos encontrar a interseção da reta r com os planos coordenados xy, yz e xz. A equação do plano xy é $z = 0$, do plano yz é $x = 0$ e do plano xz é $y = 0$. Substituindo $z = 0$ nas equações de r, obtemos $t = -1$, $x = 3$ e $y = 1/2$, ou seja, o ponto de interseção de r com o plano xy é

$$(x, y, z) = (3, \frac{1}{2}, 0).$$

De forma análoga obtemos o ponto de interseção de r com o plano yz é $(x, y, z) = (0, 1, 2)$, o ponto de interseção de r com o plano xz $(x, y, z) = (6, 0, -2)$.

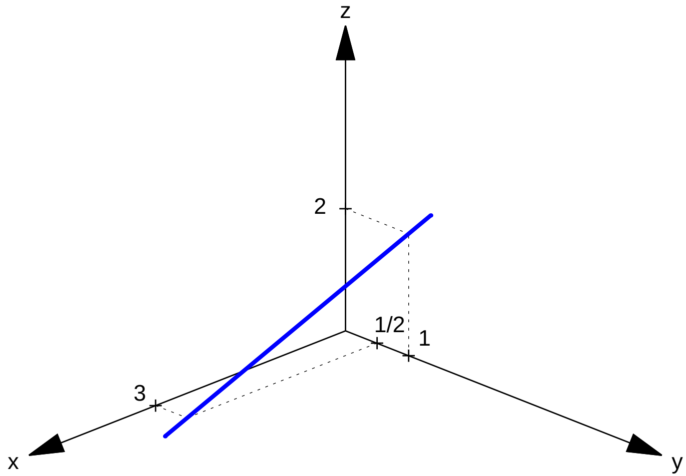


Figura 3.27. Reta que passa pelo ponto $P_0 = (-3, 3/2, 4)$ paralela ao vetor $V = (-6, 1, 4)$

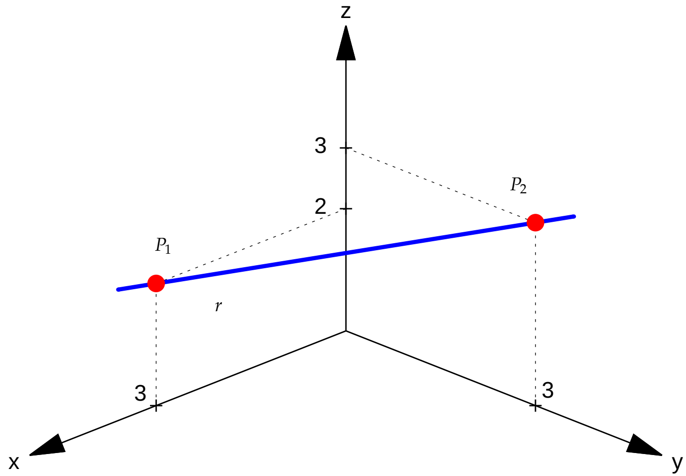


Figura 3.28. Reta que passa pelos pontos $P_1 = (3, 0, 2)$ e $P_2 = (0, 3, 3)$

Exemplo 3.14. Vamos encontrar as equações paramétricas da reta r que passa pelos pontos $P_1 = (3,0,2)$ e $P_2 = (0,3,3)$. O vetor

$$\overrightarrow{P_1P_2} = (0-3, 3-0, 3-2) = (-3, 3, 1)$$

é paralelo a r e o ponto $P_1 = (3,0,2)$ pertence a r. Portanto, as equações paramétricas de r são

$$\left\{\begin{array}{lcl} x & = & 3-3\,t \\ y & = & 3\,t \\ z & = & 2+t \end{array}\right. \quad \text{para } t \in \mathbb{R}.$$

Exemplo 3.15. Vamos encontrar as equações paramétricas da reta r, interseção dos planos

$$\begin{array}{lllcl} \pi_1 & : & 2x+y+4z-4 & = & 0 \\ \pi_2 & : & 2x-y+2z & = & 0. \end{array} \tag{3.15}$$

Podemos encontrar as equações paramétricas de r determinando a solução geral do sistema (3.15). Para isto devemos escalonar a matriz do sistema (3.15):

$$\left[\begin{array}{rrr|r} 2 & 1 & 4 & 4 \\ 2 & -1 & 2 & 0 \end{array}\right]$$

Precisamos "zerar" o outro elemento da 1ª coluna, que é a coluna do pivô, para isto, adicionamos à 2ª linha, menos a 1ª linha.

$\boxed{\text{-1ª linha} + \text{2ª linha} \longrightarrow \text{2ª linha}}$ $\quad \left[\begin{array}{rrr|r} 2 & 1 & 4 & 4 \\ 0 & -2 & -2 & -4 \end{array}\right]$

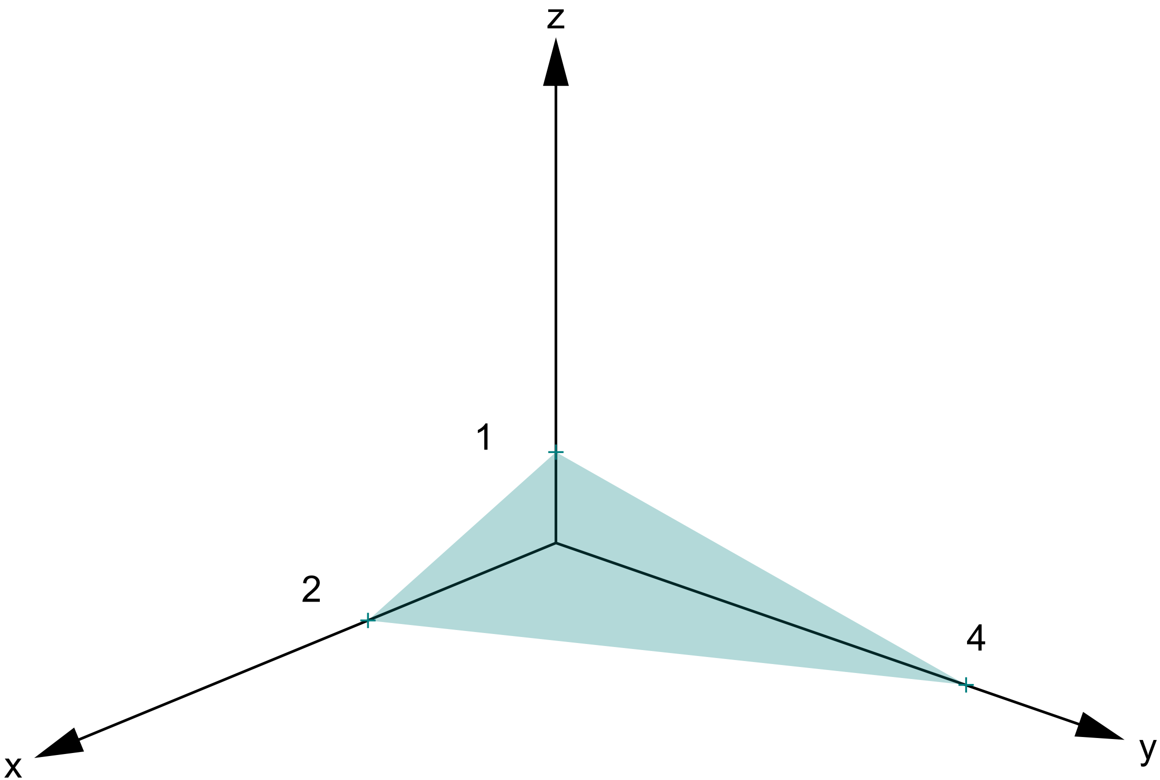


Figura 3.29. $\pi_1 : 2x + y + 4z - 4 = 0$

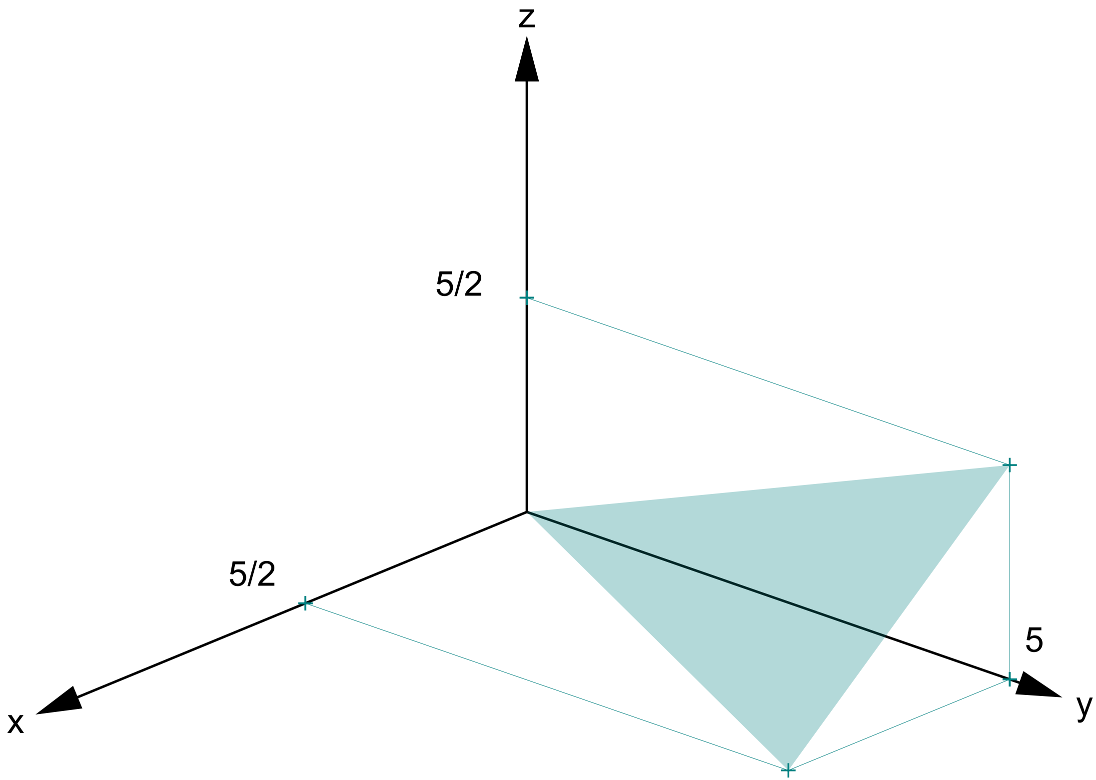


Figura 3.30. $\pi_2 : 2x - y + 2z = 0$

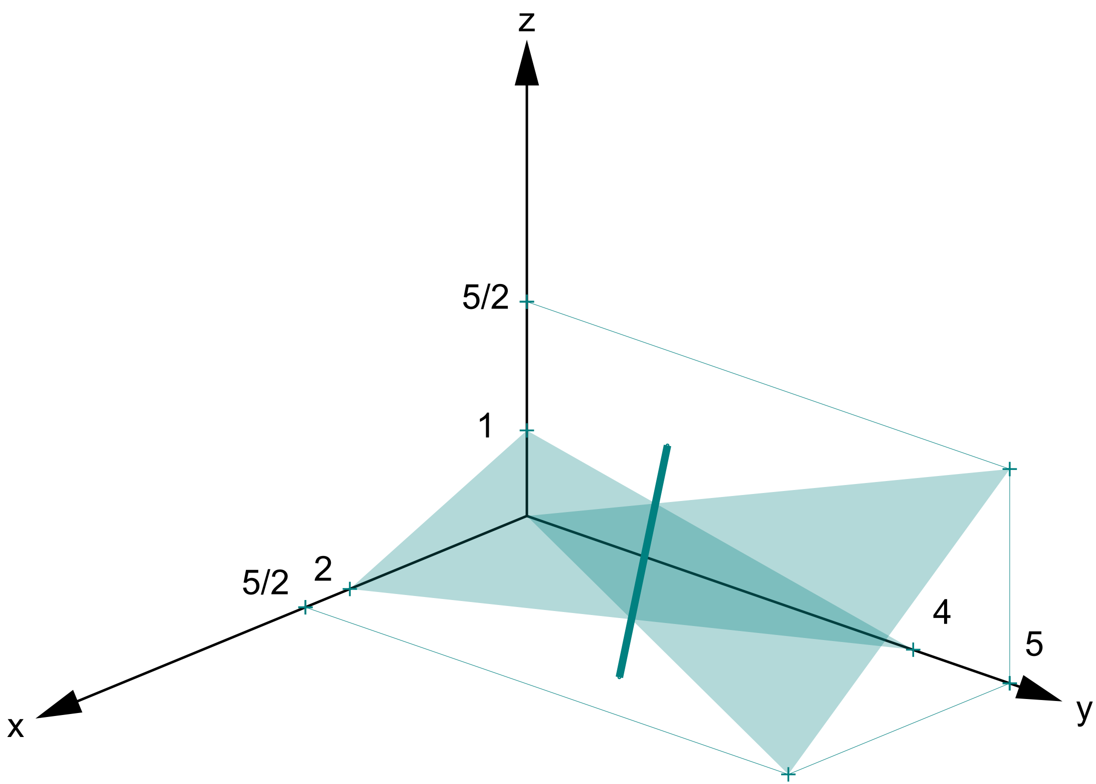


Figura 3.31. π_1, π_2 e $\pi_1 \cap \pi_2$

Agora, já podemos obter facilmente a solução geral do sistema dado, já que ele é equivalente ao sistema

$$\left\{\begin{array}{rcrcrcr} 2x & + & y & + & 4z & = & 4 \\ & - & 2y & - & 2z & = & -4 \end{array}\right.$$

A variável z é uma variável livre. Podemos dar a ela um valor arbitrário, digamos t, para $t \in \mathbb{R}$ qualquer. Assim, a solução geral do sistema dado é

$$\left\{\begin{array}{rcrcr} x & = & 1 & - & \frac{3}{2}t \\ y & = & 2 & - & t \\ z & = & & & t \end{array}\right. \quad \text{para todo } t \in \mathbb{R}. \tag{3.16}$$

Exercícios Numéricos (respostas na página 507)

3.2.1. Ache a equação do plano paralelo ao plano $2x - y + 5z - 3 = 0$ e que passa por $P = (1, -2, 1)$.

3.2.2. Encontre a equação do plano que passa pelo ponto $P = (2, 1, 0)$ e é perpendicular aos planos

$$x + 2y - 3z + 2 = 0 \quad \text{e} \quad 2x - y + 4z - 1 = 0.$$

3.2.3. Encontrar a equação do plano que passa pelos pontos $P = (1, 0, 0)$ e $Q = (1, 0, 1)$ e é perpendicular ao plano $y = z$.

3.2.4. Dadas as retas $r : (x, y, z) = (2 + 2t, 2t, t)$ e $s : (x, y, z) = (2 + t, t, t)$ obtenha uma equação geral para o plano determinado por r e s.

3.2.5. Sejam $P = (4, 1, -1)$ e $r : \ (x, y, z) = (2, 4, 1) + t\,(1, -1, 2)$.

(a) Mostre que $P \notin r$;

(b) Obtenha uma equação geral do plano determinado por r e P.

3.2.6. Dados os planos $\pi_1 : \ x - y + z + 1 = 0$ e $\pi_2 : \ x + y - z - 1 = 0$, determine o plano que contém $\pi_1 \cap \pi_2$ e é ortogonal ao vetor $(1, 1, 1)$.

3.2.7. Quais dos seguintes pares de planos se cortam segundo uma reta?

(a) $x + 2y - 3z - 4 = 0$ e $x - 4y + 2z + 1 = 0$;

(b) $2x - y + 4z + 3 = 0$ e $4x - 2y + 8z = 0$;

(c) $x - y = 0$ e $x + z = 0$.

3.2.8. Encontre as equações da reta que passa pelo ponto $Q = (1, 2, 1)$ e é perpendicular ao plano $x - y + 2z - 1 = 0$.

3.2.9. Ache a equação da reta que passa pelo ponto $P = (1, 0, 1)$ e é paralela aos planos $2x + 3y + z + 1 = 0$ e $x - y + z = 0$.

3.2.10. Seja r a reta determinada pela interseção dos planos $x + y - z = 0$ e $2x - y + 3z - 1 = 0$. Ache a equação do plano que passa por $A = (1, 0, -1)$ e contém a reta r.

Exercícios usando o MATLAB®

Comandos do MATLAB®:

`>> V=[v1,v2,v3]` cria um vetor `V`, usando as componentes numéricas `v1, v2, v3`. Por exemplo `>> V=[1,2,3]` cria o vetor $V = (1, 2, 3)$;

`>> V+W` é a soma de `V` e `W`; `>> V-W` é a diferença `V` menos `W`; `>> num*V` é o produto do vetor `V` pelo escalar `num`;

`>> subs(expr,x,num,)` substitui `x` por `num` na expressão `expr`;

`>> solve(expr)` determina a solução da equação `expr=0`;

Comandos numéricos do pacote GAAL:

`>> no(V)` calcula a norma do vetor `V`.

`>> pe(V,W)` calcula o produto escalar do vetor `V` pelo vetor `W`.

`>> pv(V,W)` calcula o produto vetorial do vetor `V` pelo vetor `W`.

Comandos gráficos do pacote GAAL:

`>> lin(P,V)` desenha a reta que passa por `P` com direção `V`.

`>> lin(P1,V1,P2,V2)` desenha retas que passam por `P1, P2`, direções `V1, V2`.

`>> plan(P,N)` desenha o plano que passa por `P` com normal `N`.

`>> plan(P1,N1,P2,N2)` desenha planos que passam por `P1, P2`, normais `N1, N2`.

`>> plan(P1,N1,P2,N2,P3,N3)` desenha planos que passam por `P1`, `P2` e `P3` com normais `N1`, `N2` e `N3`.

`>> poplan(P1,P2,N2)` desenha ponto `P1` e plano passando por `P2` com normal `N2`.

`>> poline(P1,P2,V2)` desenha ponto `P2` e reta passando por `P2` com direção `V2`.

`>> lineplan(P1,V1,P2,N2)` desenha reta passando por `P1` com direção `V1` e plano passando por `P2` com normal `N2`.

`>> axiss` reescala os eixos com a mesma escala.

`>> rota` faz uma rotação em torno do eixo z.

3.2.11. Digite no prompt `demog22`, (sem a vírgula!). Esta função demonstra as funções gráficas para visualização de retas e planos.

3.2.12. Use o MATLAB® para resolver os **Exercícios Numéricos**

Exercício Teórico

3.2.13. Seja $ax + by + cz + d = 0$ a equação de um plano π que não passa pela origem e corta os três eixos.

(a) Determine a interseção de π com os eixos;

(b) Se $P_1 = (p_1, 0, 0)$, $P_2 = (0, p_2, 0)$ e $P_3 = (0, 0, p_3)$ são as interseções de π com os eixos, a equação de π pode ser posta sob a forma

$$\frac{x}{p_1} + \frac{y}{p_2} + \frac{z}{p_3} = 1\,.$$

3.3 Os Espaços $\mathbb{R}^n$

Já vimos que os vetores no plano são identificados com os pares ordenados de números reais e que vetores no espaço são identificados com ternos ordenados de números reais. Muito do que estudamos sobre vetores no plano e no espaço pode ser estendido para n-uplas de números reais, em que n pode ser um número inteiro positivo.

Definição 3.2. Para cada inteiro positivo n, o **espaço (vetorial)** $\mathbb{R}^n$ é definido pelo conjunto de todas as n-uplas ordenadas $X = (x_1, \ldots, x_n)$ de números reais.

O conjunto $\mathbb{R}^1$ é simplesmente o conjunto dos números reais. O conjunto $\mathbb{R}^2$ é o conjunto dos pares de números reais e o $\mathbb{R}^3$ é o conjunto dos ternos de números reais.

No $\mathbb{R}^3$ o terno de números (x_1, x_2, x_3) pode ser interpretado geometricamente de duas maneiras: pode ser visto como um ponto, neste caso x_1, x_2 e x_3 são as coordenadas do ponto (Figura 3.32), ou como um vetor, neste caso x_1, x_2 e x_3 são as componentes do vetor (Figura 3.33). Também no $\mathbb{R}^n$ uma n-upla pode ser pensada como um vetor ou como um ponto. Por exemplo, a quíntupla $X = (1, -2, 3, 5, 4)$ pode ser pensada como um ponto no $\mathbb{R}^5$, quando consideramos X como um elemento do conjunto $\mathbb{R}^5$, ou como um vetor do $\mathbb{R}^5$, quando fazemos operações com X, como as que iremos definir adiante. Vamos chamar os elementos do $\mathbb{R}^n$ de pontos ou de vetores dependendo da situação.

Dois vetores $V = (v_1, \ldots, v_n)$ e $W = (w_1, \ldots, w_n)$ no $\mathbb{R}^n$ são considerados **iguais** se $v_1 = w_1, \ldots, v_n = w_n$. As operações de soma de vetores e multiplicação de vetor por escalar no $\mathbb{R}^n$ são definidas de maneira análoga ao que fizemos no plano e no espaço.

Definição 3.3. (a) A **soma** de dois vetores $V = (v_1, \ldots, v_n)$ e $W = (w_1, \ldots, w_n)$ do $\mathbb{R}^n$ é definida por

$$V + W = (v_1 + w_1, \ldots, v_n + w_n); \tag{3.17}$$

(b) A **multiplicação** de um vetor $V = (v_1, \ldots, v_n)$ do $\mathbb{R}^n$ por um escalar α é definida por

$$\alpha\, V = (\alpha\, v_1, \ldots, \alpha\, v_n). \tag{3.18}$$

O **vetor nulo** do $\mathbb{R}^n$ é denotado por $\bar{0}$ e é definido por $\bar{0} = (0, \ldots, 0)$. Se $V = (v_1, \ldots, v_n)$ é um vetor do $\mathbb{R}^n$, então o **simétrico de** V é denotado por $-V$ e é definido por $-V = (-v_1, \ldots, -v_n)$. A **diferença** de dois vetores no $\mathbb{R}^n$ é definida por $V - W = V + (-W)$. Se V e W são vetores do $\mathbb{R}^n$ tais que $W = \alpha V$, para algum escalar α, então dizemos que W é um **múltiplo escalar** de V.

Um vetor $V = (v_1, \ldots, v_n)$ do $\mathbb{R}^n$ pode também ser escrito na notação matricial como uma **matriz linha** ou como uma **matriz coluna**:

$$V = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \quad \text{ou} \quad V = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}.$$

Estas notações podem ser justificadas pelo fato de que as operações matriciais

$$V + W = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{bmatrix}, \quad \alpha V = \alpha \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \alpha v_1 \\ \vdots \\ \alpha v_n \end{bmatrix}$$

ou

$$V + W = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix} + \begin{bmatrix} w_1 & \ldots & w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 & \ldots & v_n + w_n \end{bmatrix},$$

$$\alpha V = \alpha \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix} = \begin{bmatrix} \alpha v_1 & \ldots & \alpha v_n \end{bmatrix}$$

produzem os mesmos resultados que as operações vetoriais

$$V + W = (v_1, \ldots, v_n) + (w_1, \ldots, w_n) = (v_1 + w_1, \ldots, v_n + w_n)$$

$$\alpha V = \alpha(v_1, \ldots, v_n) = (\alpha v_1, \ldots, \alpha v_n).$$

No teorema seguinte enunciamos as propriedades mais importantes da soma de vetores e multiplicação de vetores por escalar no $\mathbb{R}^n$.

Teorema 3.7. *Sejam $U = (u_1, \ldots, u_n)$, $V = (v_1, \ldots, v_n)$ e $W = (w_1, \ldots, w_n)$ vetores do $\mathbb{R}^n$ e α e β escalares. São válidas as seguintes propriedades:*

(a) $U + V = V + U$;

(b) $(U + V) + W = U + (V + W)$;

(c) $U + \bar{0} = U$;

(d) $U + (-U) = \bar{0}$;

(e) $\alpha(\beta U) = (\alpha\beta)U$;

(f) $\alpha(U + V) = \alpha U + \alpha V$;

(g) $(\alpha + \beta)U = \alpha U + \beta U$;

(h) $1U = U$.

Demonstração. Segue diretamente das propriedades da álgebra matricial (Teorema 1.1 na página 9). ■

O conceito de vetores pode ser generalizado ainda mais. Um conjunto não vazio onde estão definidas as operações de soma e multiplicação por escalar é chamado **espaço vetorial** se satisfaz as oito propriedades do Teorema 3.7 (Seção 4.3 na página 272).

3.3.1 Combinação Linear

Uma combinação linear de vetores $V_1, \ldots, V_k$, é simplesmente uma soma de múltiplos escalares de $V_1, \ldots, V_k$.

Definição 3.4. Um vetor $V \in \mathbb{R}^n$ é uma **combinação linear** dos vetores $V_1, \ldots, V_k \in \mathbb{R}^n$, se existem escalares $x_1, \ldots, x_k$ que satisfazem a equação

$$x_1 V_1 + x_2 V_2 + \ldots + x_k V_k = V \tag{3.19}$$

ou seja, se a equação vetorial (3.19) possui solução. Neste caso, dizemos também que V pode ser escrito como uma combinação linear de $V_1, \ldots, V_k$.

Se $k = 1$, então a equação (3.19) se reduz a $x_1 V_1 = V$, ou seja, V é uma combinação linear de V_1 se, e somente se, V é um múltiplo escalar de V_1.

Exemplo 3.16. Sejam $V_1 = (1,0,0)$ e $V_2 = (1,1,0)$, vetores de $\mathbb{R}^3$. O vetor $V = (2,3,2)$ **não** é uma combinação linear de V_1 e V_2, pois a equação

$$x_1V_1 + x_2V_2 = V, \tag{3.20}$$

que pode ser escrita como

$$x_1(1,0,0) + x_2(1,1,0) = (2,3,2),$$

ou ainda,

$$(x_1 + x_2, x_2, 0) = (2,3,2),$$

é equivalente ao sistema

$$\left\{\begin{array}{rcrcr} x_1 & + & x_2 & = & 2 \\ & & x_2 & = & 3 \\ & & 0 & = & 2 \end{array}\right.$$

que **não** possui solução.

Exemplo 3.17. O vetor $V = (2,3,0)$ é uma combinação linear de $V_1 = (1,0,0)$ e $V_2 = (1,1,0)$, pois a equação

$$x_1V_1 + x_2V_2 = V \tag{3.21}$$

ou

$$x_1(1,0,0) + x_2(1,1,0) = (2,3,0)$$

ou ainda,

$$(x_1 + x_2, x_2, 0) = (2,3,0),$$

é equivalente ao sistema

$$\left\{\begin{array}{rcrcr} x_1 & + & x_2 & = & 2 \\ & & x_2 & = & 3 \\ & & 0 & = & 0 \end{array}\right.$$

que possui solução.

Exemplo 3.18. O vetor nulo $\bar{0}$ é sempre combinação linear de quaisquer vetores $V_1, \ldots, V_k \in \mathbb{R}^n$, pois

$$\bar{0} = 0V_1 + \ldots + 0V_k.$$

Exemplo 3.19. Todo vetor $V = (a, b, c)$ do $\mathbb{R}^3$ é uma combinação linear de

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

Pois,

$$(a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) = a\vec{i} + b\vec{j} + c\vec{k}.$$

Para verificarmos se um vetor B é combinação linear de um conjunto de vetores $\{A_1, \ldots, A_n\}$, escrevemos a equação vetorial

$$x_1A_1 + x_2A_2 + \ldots + x_nA_n = B\,, \tag{3.22}$$

e verificamos se ela tem solução. Se $A_1, \ldots, A_n$ são vetores do $\mathbb{R}^m$, a equação (3.22), pode ser escrita como

$$x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

que é equivalente ao sistema linear

$$AX = B,$$

em que as colunas de A são os vetores A_i escritos como matrizes colunas, ou seja,

$$A = [A_1 \ldots A_n] \quad \text{e} \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Isto prova o seguinte resultado.

Proposição 3.8. *Sejam A uma matriz $m \times n$ e B uma matriz $m \times 1$. O vetor B é combinação linear das colunas de A se, e somente se, o sistema $AX = B$ tem solução.*

3.3.2 Independência Linear

Definição 3.5. Dizemos que um conjunto $\mathcal{S} = \{V_1, \ldots, V_k\}$ de vetores do $\mathbb{R}^n$ é **linearmente independente (L.I.)** se a equação vetorial

$$x_1V_1 + x_2V_2 + \ldots + x_kV_k = \bar{0} \tag{3.23}$$

só possui a solução trivial, ou seja, se a única forma de escrever o vetor nulo como combinação linear dos vetores $V_1, \ldots, V_k$ é aquela em que todos os escalares são iguais a zero. Caso contrário, isto é, se (3.23) possui solução **não** trivial, dizemos que o conjunto $\mathcal{S}$ é **linearmente dependente (L.D.).**

Exemplo 3.20. Um conjunto finito de vetores do $\mathbb{R}^n$ que contém o vetor nulo é L.D., pois se $\{V_1, \ldots, V_k\}$ é tal que $V_j = \bar{0}$, para algum j, então

$$0V_1 + \ldots + 0V_{j-1} + 1V_j + 0V_{j+1} + \ldots + 0V_k = \bar{0}.$$

Exemplo 3.21. Um conjunto formado por um único vetor do $\mathbb{R}^n$, $\{V_1\}$, **não nulo** é L.I., pois $x_1 V_1 = \bar{0}$ é equivalente a $x_1 = 0$ ou $V_1 = \bar{0}$. Mas, $V_1 \neq \bar{0}$; portanto $x_1 = 0$.

Exemplo 3.22. Se $\{V_1, \ldots, V_k\}$ é um conjunto de vetores L.D., então qualquer conjunto finito de vetores que contenha $V_1, \ldots, V_k$ é também L.D., pois a equação

$$x_1 V_1 + \ldots + x_k V_k + 0\, W_1 + \ldots + 0\, W_m = \bar{0}$$

admite solução não trivial.

Exemplo 3.23. Um conjunto formado por dois vetores do $\mathbb{R}^n$, $\{V_1, V_2\}$ é L.D. se, e somente se, a equação $x_1 V_1 + x_2 V_2 = \bar{0}$ possui solução não trivial. Mas se isto acontece, então um dos escalares x_1 ou x_2 pode ser diferente de zero. Se $x_1 \neq 0$, então $V_1 = (-x_2/x_1)V_2$ e se $x_2 \neq 0$, então $V_2 = (-x_1/x_2)V_1$. Ou seja, se $\{V_1, V_2\}$ é L.D., então um dos vetores é múltiplo escalar do outro.

Reciprocamente, se um vetor é múltiplo escalar do outro, digamos se $V_1 = \alpha V_2$, então $1\, V_1 - \alpha V_2 = \bar{0}$ e assim eles são L.D. Portanto, podemos dizer que dois vetores são L.D. se, e somente se, um é um múltiplo escalar do outro.

Por exemplo, o conjunto $\mathcal{S} = \{V_1, V_2\}$, em que $V_1 = (1, 0, 1)$ e $V_2 = (0, 1, 1)$, é L.I., pois um vetor não é múltiplo escalar do outro.

Exemplo 3.24. Um conjunto formado por três vetores de $\mathbb{R}^n$, $\{V_1, V_2, V_3\}$ é L.D. se, e somente se, a equação $x_1 V_1 + x_2 V_2 + x_3 V_3 = \bar{0}$ possui solução não trivial. Mas se isto acontece, então um dos escalares x_1 ou x_2 ou x_3 pode ser diferente de zero. Se $x_1 \neq 0$, então $V_1 = (-x_2/x_1)V_2 + (-x_3/x_1)V_3$, ou seja, o vetor V_1 é combinação linear de V_2 e V_3. De forma semelhante, se $x_2 \neq 0$, então V_2 é combinação linear de

V_1 e V_3 e se $x_3 \neq 0$, então V_3 é combinação linear de V_1 e V_2. Assim, se três vetores V_1, V_2 e V_3 do $\mathbb{R}^n$ são L.D., então um deles é uma combinação linear dos outros dois, ou seja, em deles é uma soma de múltiplos escalares dos outros dois. No $\mathbb{R}^3$ temos que se três vetores não nulos são L.D., então ou os três são paralelos (Figura 3.40), ou dois deles são paralelos (Figura 3.41) ou os três são coplanares, isto é, são paralelos a um mesmo plano (Figura 3.42).

Reciprocamente, se um vetor é uma combinação linear dos outros dois, digamos se $V_1 = \alpha V_2 + \beta V_3$, então $1\, V_1 - \alpha V_2 - \beta V_3 = \bar{0}$ e assim eles são L.D. Portanto, podemos dizer que três vetores são L.D. se, e somente se, um deles é uma combinação linear dos outros dois. No $\mathbb{R}^3$, se três vetores são L.I., então eles não são coplanares (Figura 3.43).

Exemplo 3.25. Vamos mostrar que os vetores

$$E_1 = (1, 0, \ldots, 0),\ E_2 = (0, 1, 0, \ldots, 0),\ \ldots,\ E_n = (0, \ldots, 0, 1)$$

são L.I. em particular os vetores $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$ e $\vec{k} = (0, 0, 1)$ são L.I. A equação

$$x_1 E_1 + \ldots + x_n E_n = \bar{0}$$

pode ser escrita como

$$x_1(1, 0, \ldots, 0) + \ldots + x_n(0, \ldots, 0, 1) = (0, \ldots, 0)\,.$$

Logo, $(x_1, \ldots, x_n) = (0, \ldots, 0)$, que é equivalente ao sistema

$$x_1 = 0,\ \ldots,\ x_n = 0\,.$$

Para descobrir se um conjunto de vetores $\{A_1, \ldots, A_n\}$ é L.I. precisamos saber se a equação vetorial

$$x_1 A_1 + x_2 A_2 + \ldots + x_n A_n = \bar{0} \tag{3.24}$$

tem somente a solução trivial. Se $A_1, \ldots, A_n$ são vetores do $\mathbb{R}^m$, a equação (3.24), pode ser escrita como

$$x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

que é equivalente ao sistema linear homogêneo $AX = \bar{0}$, em que as colunas de A são os vetores A_i escritos como matrizes colunas, ou seja,

$$A = [A_1 \ldots A_n] \quad \text{e} \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Isto prova o seguinte resultado.

Proposição 3.9. *Seja A uma matriz $m \times n$.*

(a) As colunas de A são linearmente independentes se, e somente se, o sistema $AX = \bar{0}$ tem somente a solução trivial.

(b) Se $m = n$, então as colunas de A são linearmente independentes se, e somente se,

$$\det(A) \neq 0.$$

Três ou mais vetores no $\mathbb{R}^2$, assim como quatro ou mais vetores no $\mathbb{R}^3$ e mais de n vetores no $\mathbb{R}^n$ são sempre L.D. Pois, nestes casos, o problema de verificar se eles são ou não L.I. leva a um sistema linear homogêneo com mais incógnitas do que equações, que pelo Teorema 1.6 na página 48 tem sempre solução não trivial.

Corolário 3.10. *Em $\mathbb{R}^n$ um conjunto com mais de n vetores é L.D.*

Exemplo 3.26. Considere os vetores $X_1 = (1,0,1)$, $X_2 = (0,1,1)$ e $X_3 = (1,1,1)$ de $\mathbb{R}^3$. Para sabermos se eles são L.I. ou L.D. escrevemos a equação

$$x_1X_1 + x_2X_2 + x_3X_3 = \bar{0}.$$

Esta equação vetorial é equivalente ao sistema linear $AX = \bar{0}$, em que

$$A = [\, X_1 \; X_2 \; X_3 \,] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Escalonando a matriz $[\, A \,|\, \bar{0} \,]$ podemos obter a sua forma escalonada reduzida

$$[\, R \,|\bar{0}\,] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

Concluímos, então que o sistema $A\,X = \bar{0}$ possui somente a solução trivial

$$x_1 = x_2 = x_3 = 0.$$

Portanto os vetores X_1, X_2 e X_3 são L.I.

Exemplo 3.27. Sejam $V_1 = (1,2,5)$, $V_2 = (7,-1,5)$ e $V_3 = (1,-1,-1)$ vetores do $\mathbb{R}^3$. Para sabermos se eles são L.I. ou L.D. escrevemos a equação

$$x_1V_1 + x_2V_2 + x_3V_3 = \bar{0}. \tag{3.25}$$

Esta equação vetorial é equivalente ao sistema linear $AX = \bar{0}$, em que

$$A = [\, V_1 \ V_2 \ V_3 \,] = \begin{bmatrix} 1 & 7 & 1 \\ 2 & -1 & -1 \\ 5 & 5 & -1 \end{bmatrix}.$$

A matriz $[\, A \,|\, \bar{0} \,]$ é equivalente por linhas à matriz escalonada reduzida

$$[\, R \,|\, \bar{0} \,] = \left[\begin{array}{rrr|r} 1 & 0 & -2/5 & 0 \\ 0 & 1 & 1/5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]. \tag{3.26}$$

Assim a variável x_3 pode ser uma variável livre que pode, portanto, assumir qualquer valor. Concluímos que o sistema $A\,X = \bar{0}$ e a equação vetorial (3.25) têm solução não trivial. Portanto, V_1, V_2 e V_3 são L.D.

A expressão "linearmente dependente" sugere que os vetores dependem uns dos outros em algum sentido. O teorema seguinte mostra que este realmente é o caso.

Teorema 3.11. *Um conjunto $\mathcal{S} = \{V_1, \ldots, V_k\}$ $(k > 1)$ de vetores de $\mathbb{R}^n$ é linearmente dependente (L.D.) se, e somente se, pelo menos um dos vetores, V_j, for combinação linear dos outros vetores de $\mathcal{S}$.*

Demonstração. Vamos dividir a demonstração em duas partes:

(a) Se V_j é uma combinação linear dos demais vetores do conjunto $\mathcal{S}$, isto é, se existem escalares $\alpha_1, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_k$ tais que

$$\alpha_1 V_1 + \ldots + \alpha_{j-1} V_{j-1} + \alpha_{j+1} V_{j+1} + \ldots + \alpha_k V_k = V_j,$$

então somando-se $-V_j$ a ambos os membros ficamos com

$$\alpha_1 V_1 + \ldots + \alpha_{j-1} V_{j-1} - V_j + \alpha_{j+1} V_{j+1} + \ldots + \alpha_k V_k = \bar{0}. \tag{3.27}$$

Isto implica que a equação $x_1 V_1 + \ldots + x_k V_k = \bar{0}$ admite solução não trivial, pois o coeficiente de V_j em (3.27) é -1. Portanto, $\mathcal{S}$ é L.D.

(b) Se $\mathcal{S}$ é L.D., então a equação

$$x_1 V_1 + x_2 V_2 + \ldots + x_k V_k = \bar{0} \tag{3.28}$$

admite solução não trivial, o que significa que pelo menos um x_j é diferente de zero. Então, multiplicando-se a equação (3.28) por $1/x_j$ e subtraindo-se $(\frac{x_1}{x_j})V_1 + \ldots + (\frac{x_k}{x_j})V_k$ obtemos

$$V_j = -\left(\frac{x_1}{x_j}\right) V_1 - \ldots - \left(\frac{x_{j-1}}{x_j}\right) V_{j-1} - \left(\frac{x_{j+1}}{x_j}\right) V_{j+1} - \ldots - \left(\frac{x_k}{x_j}\right) V_k\,.$$

Portanto, um vetor V_j é combinação linear dos outros vetores de $\mathcal{S}$. ■

Observação. Na demonstração da segunda parte, vemos que o vetor, cujo escalar na combinação linear, puder ser diferente de zero, pode ser escrito como combinação linear dos outros.

Exemplo 3.28. Sejam $V_1 = (1, 2, 5)$, $V_2 = (7, -1, 5)$ e $V_3 = (1, -1, -1)$ vetores do $\mathbb{R}^3$. Vamos escrever um dos vetores como combinação linear dos outros dois. Vimos no Exemplo 3.27 que estes vetores são L.D. De (3.26) segue-se que

$$x_1 V_1 + x_2 V_2 + x_3 V_3 = \bar{0}$$

se, e somente se, $x_1 = (2/5)\alpha$, $x_2 = -(1/5)\alpha$ e $x_3 = \alpha$, para todo $\alpha \in \mathbb{R}$. Substituindo-se os valores de x_1, x_2 e x_3 na equação acima, ficamos com

$$(2/5)\alpha V_1 - (1/5)\alpha V_2 + \alpha V_3 = \bar{0}$$

Tomando-se $\alpha = 1$, obtemos

$$(2/5)V_1 - (1/5)V_2 + V_3 = \bar{0}$$

multiplicando-se por -5 e somando-se $2V_1 + 5V_3$, temos que $V_2 = 2V_1 + 5V_3$. Observe que, neste exemplo, qualquer dos vetores pode ser escrito como combinação linear dos outros. O próximo exemplo mostra que isto nem sempre acontece.

Exemplo 3.29. Sejam $V_1 = (-2, -2, 2)$, $V_2 = (-3, 3/2, 0)$ e $V_3 = (-2, 1, 0)$. $\{V_1, V_2, V_3\}$ é L.D., mas V_1 não é combinação linear de V_2 e V_3 (Figura 3.41 na página 226).

Exercícios Numéricos (respostas na página 510)

3.3.1. Quais dos seguintes vetores são combinação linear de

$$X_1 = (4, 2, -3), X_2 = (2, 1, -2) \quad \text{e} \quad X_3 = (-2, -1, 0)?$$

(a) $(1, 1, 1)$;
(b) $(4, 2, -6)$;
(c) $(-2, -1, 1)$;
(d) $(-1, 2, 3)$.

3.3.2. Quais dos seguintes conjuntos de vetores são linearmente dependentes?

(a) $\{(1, 1, 2), (1, 0, 0), (4, 6, 12)\}$;
(b) $\{(1, -2, 3), (-2, 4, -6)\}$;
(c) $\{(1, 1, 1), (2, 3, 1), (3, 1, 2)\}$;
(d) $\{(4, 2, -1), (6, 5, -5), (2, -1, 3)\}$.

3.3.3. Para quais valores de λ o conjunto de vetores $\{(3, 1, 0), (\lambda^2 + 2, 2, 0)\}$ é L.D.?

3.3.4. Suponha que $\{V_1, V_2, V_3\}$ é um conjunto linearmente independente de vetores de $\mathbb{R}^n$. Responda se $\{W_1, W_2, W_3\}$ é linearmente dependente ou independente nos seguintes casos:

(a) $W_1 = V_1 + V_2$, $W_2 = V_1 + V_3$ e $W_3 = V_2 + V_3$;

(b) $W_1 = V_1$, $W_2 = V_1 + V_3$ e $W_3 = V_1 + V_2 + V_3$.

Exercício usando o MATLAB®

3.3.5. (a) Defina os vetores `V1=[1;2;3]`, `V2=[3;4;5]` e `V3=[5;6;7]`. Defina o vetor aleatório `V=randi(3,1)`. Verifique se `V` é combinação linear de `V1`, `V2` e `V3`.

(b) Defina a matriz aleatória `M=randi(3,5)`. Verifique se os vetores definidos pelas colunas de `M` são combinação linear de `V1`, `V2` e `V3`. Tente explicar o resultado.

(c) Verifique se `V1`, `V2` e `V3` são linearmente independentes. Se eles forem linearmente dependentes, escreva um deles como combinação linear dos outros e verifique o resultado.

Exercícios Teóricos

3.3.6. Suponha que $\{X_1, X_2, \ldots, X_k\}$, com $k \leq n$, é um conjunto de vetores do $\mathbb{R}^n$ linearmente independente. Mostre que se A é uma matriz $n \times n$ não singular, então $\{AX_1, AX_2, \ldots, AX_k\}$ também é um conjunto linearmente independente.

3.3.7. Se os vetores não nulos U, V e W são L.D., então W é uma combinação linear de U e V?

z

(x, y, z)

x y

Figura 3.32. Ponto $(x, y, z) \in \mathbb{R}^3$

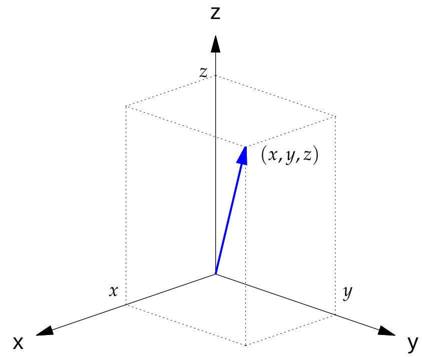


Figura 3.33. Vetor $(x, y, z) \in \mathbb{R}^3$

$V = (2, 3, 2)$

$V_1 = (1, 0, 0)$

$V_2 = (1, 1, 0)$

Figura 3.34. O vetor V **não** é combinação linear de V_1 e V_2

$V_1 = (1, 0, 0)$

$V_2 = (1, 1, 0)$

$V = (2, 3, 0)$

Figura 3.35. O vetor V é combinação linear de V_1 e V_2

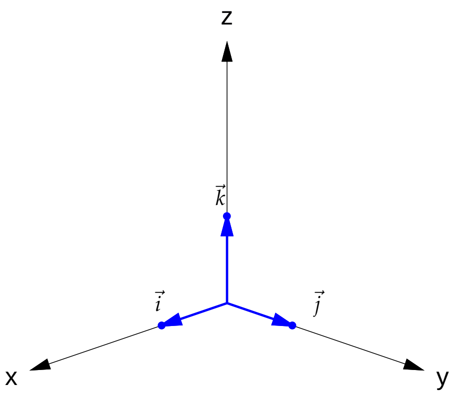


Figura 3.36. Vetores $\vec{i}$, $\vec{j}$ e $\vec{k}$

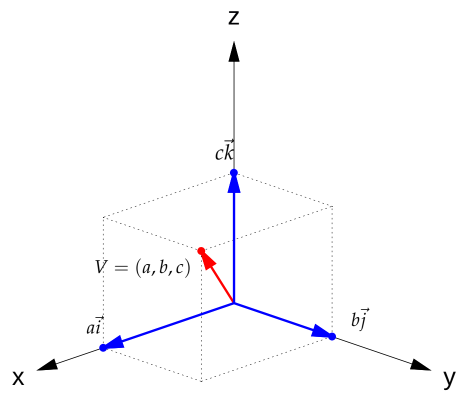


Figura 3.37. $V = (a, b, c) = a\vec{i} + b\vec{j} + c\vec{k}$

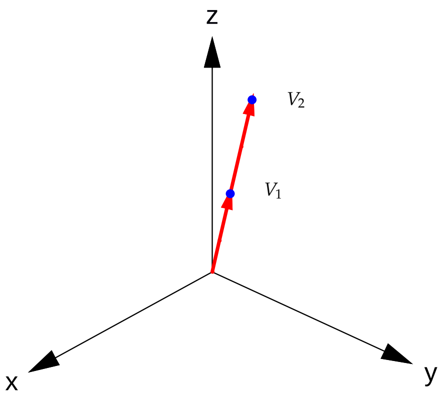


Figura 3.38. Dois vetores linearmente dependentes

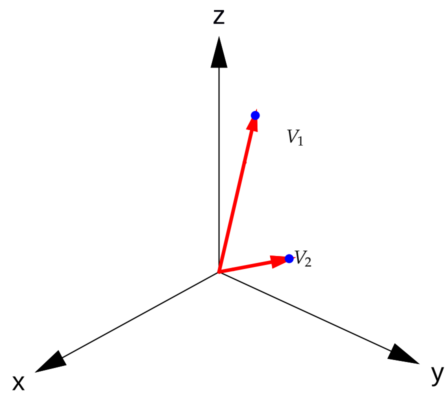


Figura 3.39. Dois vetores linearmente independentes

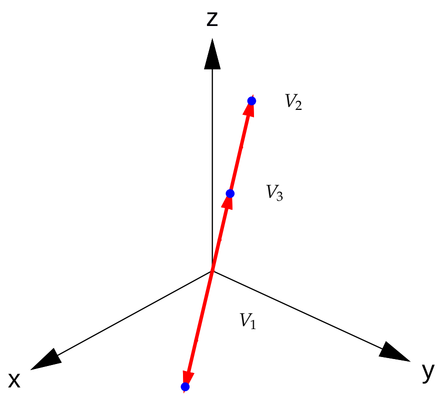


Figura 3.40. Três vetores linearmente dependentes (paralelos)

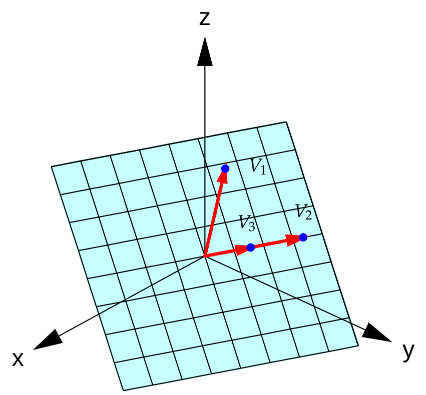


Figura 3.41. Três vetores linearmente dependentes (dois paralelos)

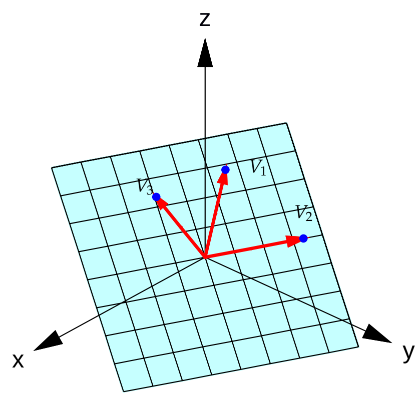


Figura 3.42. Três vetores linearmente dependentes (coplanares)

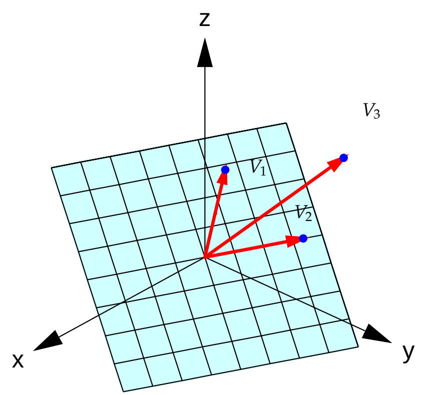


Figura 3.43. Três vetores linearmente independentes

4

Subespaços

4.1 Base e Dimensão

Sejam A uma matriz $m \times n$ e $\mathbb{W} \subseteq \mathbb{R}^n$ o conjunto solução do sistema linear homogêneo $AX = \bar{0}$. Já vimos na Proposição 1.7 na página 48 que o conjunto $\mathbb{W}$ satisfaz as seguintes propriedades:

(a) Se X e Y pertencem a $\mathbb{W}$, então $X + Y$ também pertence a $\mathbb{W}$.

(b) Se X pertence a $\mathbb{W}$, então αX também pertence a $\mathbb{W}$ para todo escalar α.

Revise como foi feita a demonstração dos itens (a) e (b) acima na Proposição 1.7 na página 48. Assim, se X e Y são soluções de um sistema homogêneo, então $X+Y$ e αX também o são. Portanto, combinações lineares de soluções de $AX = \bar{0}$ são também soluções de $AX = \bar{0}$.

O conjunto solução de um sistema homogêneo $AX = \bar{0}$ é chamado de **espaço solução do sistema homogêneo** $AX = \bar{0}$. Ele se comporta como se fosse um espaço, no sentido de que fazendo soma de vetores do conjunto ou multiplicando vetores do conjunto por escalar não saímos dele.

Um subconjunto não vazio de $\mathbb{R}^n$ que satisfaz as propriedades (a) e (b) acima é chamado de **subespaço** de $\mathbb{R}^n$. Com relação as operações de soma e multiplicação por escalar podemos "viver" nele sem termos que sair. Assim o espaço solução do sistema homogêneo $AX = \bar{0}$ é um subespaço de $\mathbb{R}^n$. Vale também a recíproca, todo subespaço é o espaço solução de um sistema homogêneo (Exercício 4.1.18 na página 244).

Exemplo 4.1. Os exemplos mais triviais de subespaços de $\mathbb{R}^n$ são o subespaço formado somente pelo vetor nulo, $\mathbb{W} = \{\bar{0}\}$ e $\mathbb{W} = \mathbb{R}^n$. Mas cuidado, o $\mathbb{R}^2$ não é subespaço de $\mathbb{R}^3$, pois o $\mathbb{R}^2$ (conjunto de pares de números reais) não é um subconjunto do $\mathbb{R}^3$ (conjunto de ternos de números reais). O plano

$$\mathbb{W} = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}$$

é um subespaço de $\mathbb{R}^3$ mas ele não é o $\mathbb{R}^2$.

Exemplo 4.2. Considere o sistema linear

$$\left\{\begin{array}{ccccccc} a_1x & + & b_1y & + & c_1z & = & 0 \\ a_2x & + & b_2y & + & c_2z & = & 0 \\ a_3x & + & b_3y & + & c_3z & = & 0 \end{array}\right.$$

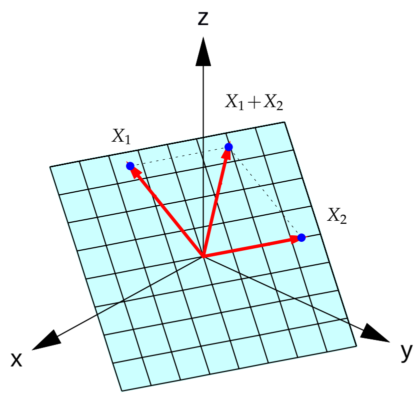


Figura 4.1. Soma de vetores do plano $ax + by + cz = 0$

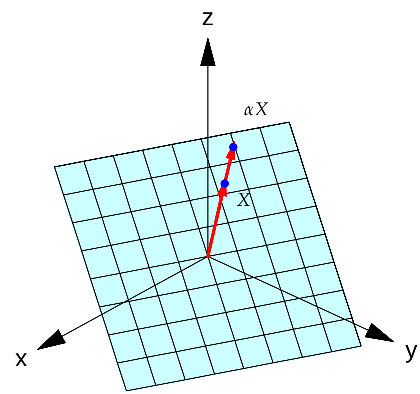


Figura 4.2. Multiplicação de vetor por escalar do plano $ax + by + cz = 0$

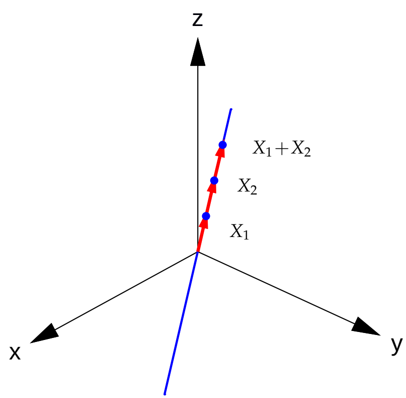


Figura 4.3. Soma de vetores da reta $(x, y, z) = (at, bt, ct)$

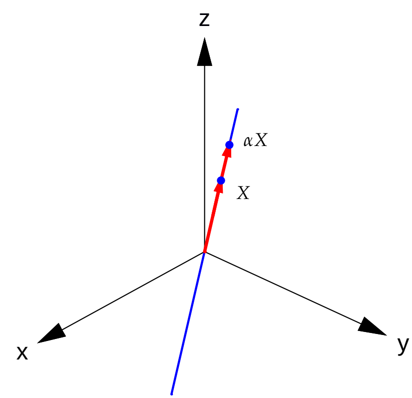


Figura 4.4. Multiplicação de vetor por escalar da reta $(x, y, z) = (at, bt, ct)$

Cada equação deste sistema é representada por um plano que passa pela origem. O conjunto solução é um subespaço de $\mathbb{R}^3$ e é a interseção dos planos definidos pelas equações, podendo ser:

(a) Somente um ponto que é a origem.

(b) Uma reta que passa pela origem.

(c) Um plano que passa pela origem.

Vamos escrever toda solução do sistema linear homogêneo $AX = \bar{0}$ como uma combinação linear de um número finito de vetores $V_1, \ldots, V_k$ que são também solução do sistema.

Exemplo 4.3. Considere o sistema linear homogêneo $AX = \bar{0}$, em que

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

Escalonando a matriz aumentada do sistema acima, obtemos a matriz escalonada reduzida

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

E assim a solução geral do sistema pode ser escrita como

$$x_1 = -\alpha - \gamma,\ x_2 = \gamma,\ x_3 = -\alpha + \beta,\ x_4 = \beta\ x_5 = \alpha$$

para todos os valores de $\alpha, \beta, \gamma \in \mathbb{R}$, ou seja, o conjunto solução do sistema $AX = \bar{0}$ é

$$\mathbb{W} = \{(x_1, x_2, x_3, x_4, x_5) = (-\alpha - \gamma, \gamma, -\alpha + \beta, \beta, \alpha) \mid \alpha, \beta, \gamma \in \mathbb{R}\}.$$

Agora, um elemento qualquer de $\mathbb{W}$ pode ser escrito como uma combinação linear de vetores de $\mathbb{W}$:

$$\begin{aligned}(-\alpha-\gamma,\gamma,-\alpha+\beta,\beta,\alpha) &= (-\alpha,0,-\alpha,0,\alpha)+(0,0,\beta,\beta,0)+(-\gamma,\gamma,0,0,0)\\ &= \alpha(-1,0,-1,0,1)+\beta(0,0,1,1,0)+\gamma(-1,1,0,0,0)\end{aligned}$$

Assim, todo vetor de $\mathbb{W}$ pode ser escrito como combinação linear dos vetores $V_1 = (-1,0,-1,0,1)$, $V_2 = (0,0,1,1,0)$ e $V_3 = (-1,1,0,0,0)$ pertencentes a $\mathbb{W}$ (V_1 é obtido fazendo-se $\alpha = 1$ e $\beta = \gamma = 0$, V_2 fazendo-se $\alpha = \gamma = 0$ e $\beta = 1$ e V_3 fazendo-se $\alpha = \beta = 0$ e $\gamma = 1$).

Neste caso dizemos que $V_1 = (-1,0,-1,0,1)$, $V_2 = (0,0,1,1,0)$ e $V_3 = (-1,1,0,0,0)$ **geram** o subespaço $\mathbb{W}$. Em geral temos a seguinte definição.

Definição 4.1. Seja $\mathbb{W}$ um subespaço de $\mathbb{R}^n$ (por exemplo, o espaço solução de um sistema linear homogêneo $AX = \bar{0}$). Dizemos que os vetores $V_1,\ldots,V_k$ pertencentes a $\mathbb{W}$, **geram** $\mathbb{W}$ ou que $\{V_1,\ldots,V_k\}$ é um **conjunto de geradores** de $\mathbb{W}$, se qualquer vetor de $\mathbb{W}$ é combinação linear de $V_1,\ldots,V_k$. Dizemos também que $\mathbb{W}$ é o **subespaço gerado por** $V_1,\ldots,V_k$.

Uma questão importante é encontrar o maior número possível de vetores linearmente independentes em um subespaço. O resultado a seguir responde a esta questão.

Teorema 4.1. *Seja $\mathbb{W}$ subespaço de $\mathbb{R}^n$ (por exemplo, o espaço solução de um sistema linear homogêneo $AX = \bar{0}$). Seja $\{V_1, \ldots, V_m\}$ um conjunto de vetores de $\mathbb{W}$*

(a) *linearmente independente (L.I.),*

(b) *que gera $\mathbb{W}$ (ou seja, todo vetor X de $\mathbb{W}$ é combinação linear de $V_1, \ldots, V_m$).*

Então, um conjunto com mais de m vetores em $\mathbb{W}$ é linearmente dependente (L.D.).

Demonstração. Seja $\{W_1, \ldots, W_p\}$ um subconjunto de $\mathbb{W}$, com $p > m$. Vamos mostrar que $\{W_1, \ldots, W_p\}$ é L.D. Vamos considerar a combinação linear nula de $W_1, \ldots, W_p$

$$x_1W_1 + x_2W_2 + \ldots + x_pW_p = \bar{0}. \tag{4.1}$$

Como qualquer elemento de $\mathbb{W}$ pode ser escrito como combinação linear de $V_1, \ldots, V_m$, em particular,

$$W_j = b_{1j}V_1 + b_{2j}V_2 + \ldots + b_{mj}V_m = \sum_{i=1}^{m} b_{ij}V_i\,, \quad \text{para } j = 1, \ldots, p\,. \tag{4.2}$$

Assim, substituindo (4.2) em (4.1) e agrupando os termos que contém V_i, para $i = 1, \ldots, m$, obtemos

$$(b_{11}x_1 + \ldots + b_{1p}x_p)V_1 + \ldots + (b_{m1}x_1 + \ldots + b_{mp}x_p)V_m = \bar{0}. \tag{4.3}$$

Como $\{V_1, \ldots, V_m\}$ é L.I., então os escalares na equação (4.3) são iguais a zero. Isto leva ao sistema linear

$$BX = \bar{0},$$

em que $B = (b_{ij})_{m \times p}$. Mas, este é um sistema homogêneo que tem mais incógnitas do que equações, portanto possui solução não trivial, (Teorema 1.6 na página 48), como queríamos provar. ■

O resultado anterior mostra que se podemos escrever todo elemento do subespaço $\mathbb{W}$ como uma combinação linear de vetores $V_1, \ldots, V_m$ L.I. pertencentes a $\mathbb{W}$, então m é o maior número possível de vetores L.I. em $\mathbb{W}$.

No Exemplo 4.3 os vetores
$V_1 = (-1, 0, -1, 0, 1)$, $V_2 = (0, 0, 1, 1, 0)$ e $V_3 = (-1, 1, 0, 0, 0)$ geram $\mathbb{W}$.
Além disso de

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

segue-se que V_1, V_2 e V_3 são L.I. (por que?)

Assim pelo Teorema 4.1 não podemos obter um número maior de vetores em $\mathbb{W}$ L.I. Neste caso dizemos que $\{V_1, V_2, V_3\}$ é uma **base** de $\mathbb{W}$. Em geral temos a seguinte definição.

Definição 4.2. Seja $\mathbb{W}$ um subespaço de $\mathbb{R}^n$ (por exemplo, o espaço solução de um sistema linear homogêneo $AX = \bar{0}$). Dizemos que um subconjunto $\{V_1, \ldots, V_k\}$ de $\mathbb{W}$ é uma **base** de $\mathbb{W}$, se

(a) $\{V_1, \ldots, V_k\}$ é um conjunto de geradores de $\mathbb{W}$ (ou seja, todo vetor de $\mathbb{W}$ é combinação linear de $V_1, \ldots, V_k$) e

(b) $\{V_1, \ldots, V_k\}$ é L.I.

Exemplo 4.4. Os vetores $E_1 = (1, 0, \ldots, 0), E_2 = (0, 1, 0, \ldots, 0), \ldots, E_n = (0, \ldots, 0, 1)$ formam uma base do $\mathbb{R}^n$. Pois, um vetor qualquer do $\mathbb{R}^n$ é da forma $V = (a_1, \ldots, a_n)$

e pode ser escrito como uma soma de vetores, sendo um vetor para cada parâmetro e cada vetor dependendo apenas de um parâmetro, obtendo

$$\begin{aligned} V = (a_1, \ldots, a_n) &= (a_1, 0, \ldots, 0) + (0, a_2, 0, \ldots, 0) + \ldots + (0, \ldots, 0, a_n) \\ &= a_1(1, 0, \ldots, 0) + a_2(0, 1, 0, \ldots, 0) + \ldots + a_n(0, \ldots, 0, 1). \end{aligned}$$

Assim, os vetores $E_1 = (1, 0, \ldots, 0), E_2 = (0, 1, 0, \ldots, 0), \ldots, E_n = (0, \ldots, 0, 1)$ geram o $\mathbb{R}^n$. Vimos no Exemplo 3.25 na página 216 que E_1, E_2, $\ldots E_n$ são L.I. Esses vetores formam a chamada **base canônica de** $\mathbb{R}^n$. No caso do $\mathbb{R}^3$, $E_1 = \vec{i}$, $E_2 = \vec{j}$ e $E_3 = \vec{k}$.

Exemplo 4.5. Seja $\mathbb{W} = \{(x, y, z) = t(a, b, c) \mid t \in \mathbb{R}\}$ uma reta que passa pela origem. Como o vetor diretor $V = (a, b, c)$ é não nulo e gera a reta, então $\{V\}$ é uma base de $\mathbb{W}$.

Exemplo 4.6. Seja $\mathbb{W} = \{(x, y, z) \in \mathbb{R}^3 \mid ax + by + cz = 0\}$ um plano que passa pela origem. Vamos supor que $a \neq 0$. Um ponto (x, y, z) satisfaz a equação

$$ax + by + cz = 0$$

se, e somente se,

$$z = \alpha, y = \beta, x = -\frac{1}{a}(c\alpha + b\beta), \quad \text{para todos } \alpha, \beta \in \mathbb{R}.$$

Assim, o plano $\mathbb{W}$ pode ser descrito como $\mathbb{W} = \{(-\frac{c}{a}\alpha - \frac{b}{a}\beta, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$. Assim, todo vetor de $\mathbb{W}$ pode ser escrito como uma soma de vetores, sendo um para cada parâmetro, obtendo

$$(-\frac{c}{a}\alpha - \frac{b}{a}\beta, \beta, \alpha) = (-\frac{c}{a}\alpha, 0, \alpha) + (-\frac{b}{a}\beta, \beta, 0) = \alpha(-\frac{c}{a}, 0, 1) + \beta(-\frac{b}{a}, 1, 0).$$

Assim, todo vetor de $\mathbb{W}$ pode ser escrito como uma combinação linear dos vetores $V_1 = (-\frac{c}{a}, 0, 1)$ e $V_2 = (-\frac{b}{a}, 1, 0)$ pertencentes a $\mathbb{W}$ (V_1 é obtido fazendo-se $\alpha = 1$ e $\beta = 0$ e V_2, fazendo-se $\alpha = 0$ e $\beta = 1$). Portanto, $V_1 = (-\frac{c}{a}, 0, 1)$ e $V_2 = (-\frac{b}{a}, 1, 0)$ geram o plano $\mathbb{W}$. Como V_1 e V_2 são L.I., pois um não é múltiplo escalar do outro, então $\{V_1, V_2\}$ é uma base do plano $\mathbb{W}$. Deixamos como exercício para o leitor encontrar uma base de $\mathbb{W}$ para o caso em que $b \neq 0$ e também para o caso em que $c \neq 0$.

Segue do Teorema 4.1 na página 232 que se $\mathbb{W} \neq \{\bar{0}\}$ é um subespaço, então qualquer base de $\mathbb{W}$ tem o mesmo número de elementos e este é o maior número de vetores L.I. que podemos ter em $\mathbb{W}$. O número de elementos de qualquer uma das bases de $\mathbb{W}$ é chamado de **dimensão de** $\mathbb{W}$. Se $\mathbb{W} = \{\bar{0}\}$ dizemos que $\mathbb{W}$ tem dimensão igual a 0.

Exemplo 4.7. A dimensão do $\mathbb{R}^n$ é n, pois como foi mostrado no Exemplo 4.4 na página 233, $E_1 = (1, 0, \ldots, 0)$, $E_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $E_n = (0, \ldots, 0, 1)$ formam uma base do $\mathbb{R}^n$.

Exemplo 4.8. Pelo Exemplo 4.5 na página 234 uma reta que passa pela origem tem dimensão 1 e pelo Exemplo 4.6 na página 234 um plano que passa pela origem tem dimensão 2.

Vamos mostrar a seguir que se a dimensão de um subespaço $\mathbb{W}$ é $m > 0$, então basta conseguirmos m vetores L.I. em $\mathbb{W}$, que teremos uma base.

Teorema 4.2. *Seja $\mathbb{W}$ um subespaço de dimensão $m > 0$. Se m vetores, $V_1, \ldots, V_m \in \mathbb{W}$, são L.I., então eles geram o subespaço $\mathbb{W}$ e portanto formam uma base de $\mathbb{W}$.*

Demonstração. Sejam $V_1, \ldots, V_m$ vetores L.I. e seja V um vetor qualquer do subespaço $\mathbb{W}$. Vamos mostrar que V é combinação linear de $V_1, \ldots, V_m$. Considere a equação vetorial

$$x_1V_1 + x_2V_2 + \ldots + x_mV_m + x_{m+1}V = \bar{0} \tag{4.4}$$

Pelo Teorema 4.1 na página 232, $V_1, \ldots, V_m, V$ são L.D., pois são $m+1$ vetores em um subespaço de dimensão m. Então a equação (4.4) admite solução não trivial, ou seja, pelo menos um $x_i \neq 0$. Mas, $x_{m+1} \neq 0$, pois caso contrário, $V_1, \ldots, V_m$ seriam L.D. Então, multiplicando-se a equação (4.4) por $\dfrac{1}{x_{m+1}}$ e subtraindo-se

$$\frac{x_1}{x_{m+1}}V_1 + \frac{x_2}{x_{m+1}}V_2 + \ldots + \frac{x_m}{x_{m+1}}V_m,$$

obtemos

$$V = -\left(\frac{x_1}{x_{m+1}}\right)V_1 - \ldots - \left(\frac{x_m}{x_{m+1}}\right)V_m\,.$$

■

Dos resultados anteriores, vemos que se a dimensão de um subespaço, $\mathbb{W}$, é $m > 0$, então basta conseguirmos m vetores L.I. em $\mathbb{W}$, que teremos uma base (Teorema 4.2) e não podemos conseguir mais do que m vetores L.I. (Teorema 4.1 na página 232).

Exemplo 4.9. Do Teorema 4.2 segue-se que n vetores L.I. do $\mathbb{R}^n$ formam uma base de $\mathbb{R}^n$. Por exemplo, 3 vetores L.I. do $\mathbb{R}^3$ formam uma base de $\mathbb{R}^3$.

Exemplo 4.10. Sejam $\mathbb{W}$ o plano $x+y+z=0$ e $\mathbb{V}$ o plano $4x-2y+z=0$.

Podemos encontrar as equações paramétricas da reta $\mathbb{V}\cap\mathbb{W}$, interseção dos planos determinando a solução geral do sistema (4.5)

$$\begin{array}{rrl} \mathbb{W}: & x+y+z & =0\,, \\ \mathbb{V}: & 4x-2y+z & =0\,. \end{array} \tag{4.5}$$

Para isto devemos escalonar a matriz do sistema (4.5):

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 4 & -2 & 1 & 0 \end{array}\right]$$

Precisamos "zerar" o outro elemento da 1ª coluna, que é a coluna do pivô, para isto, adicionamos à 2ª linha, -4 vezes a 1ª linha.

$$\boxed{-4{*}1^{\underline{a}}\text{ linha} + 2^{\underline{a}}\text{ linha} \longrightarrow 2^{\underline{a}}\text{ linha}} \qquad \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & -6 & -3 & 0 \end{array}\right]$$

Agora, já podemos obter facilmente a solução geral do sistema dado, já que ele é equivalente ao sistema

$$\left\{\begin{array}{rcrcrcl} x & + & y & + & z & = & 0 \\ & & -6y & - & 3z & = & 0 \end{array}\right.$$

A variável z é uma variável livre. Podemos dar a ela um valor arbitrário, digamos t, para $t\in\mathbb{R}$ qualquer. Assim, a solução geral do sistema (4.5) é

$$\left\{\begin{array}{rcr} x & = & -\frac{1}{2}\,t \\ y & = & -\frac{1}{2}\,t \\ z & = & t \end{array}\right. \qquad \text{para todo } t\in\mathbb{R}.$$

A reta que é a interseção, $\mathbb{V}\cap\mathbb{W}$, tem equação $(x,y,z)=t(-1/2,-1/2,1)$, para todo $t\in\mathbb{R}$ (revise o Exemplo 3.15 na página 199). Portanto, o vetor $V=(-1/2,-1/2,1)$

gera a interseção $\mathbb{V} \cap \mathbb{W}$. Como um vetor não nulo é L.I. o conjunto

$$\{V = (-1/2, -1/2, 1)\}$$

é uma base do subespaço que é a reta interseção de $\mathbb{V}$ com $\mathbb{W}$.

Observação. Como no exemplo anterior, em geral, o espaço solução de um sistema linear homogêneo pode ser visto como uma interseção de subespaços que são as soluções de sistemas formados por subconjuntos de equações do sistema inicial.

Exemplo 4.11. Considere o subespaço $\mathbb{W} = \{(a + c, b + c, a + b + 2c) \mid a, b, c \in \mathbb{R}\}$ de $\mathbb{R}^3$. Vamos encontrar um conjunto de geradores e uma base para $\mathbb{W}$.

Qualquer elemento V de $\mathbb{W}$ pode ser escrito como uma soma de vetores, sendo um vetor para cada parâmetro e cada vetor dependendo apenas de um parâmetro, obtendo

$$\begin{aligned} V = (a + c, b + c, a + b + 2c) &= (a, 0, a) + (0, b, b) + (c, c, 2c) \\ &= a(1, 0, 1) + b(0, 1, 1) + c(1, 1, 2). \end{aligned}$$

Logo, definindo $V_1 = (1, 0, 1)$, $V_2 = (0, 1, 1)$ e $V_3 = (1, 1, 2)$, temos que $\{V_1, V_2, V_3\}$ gera $\mathbb{W}$. Para sabermos se $\{V_1, V_2, V_3\}$ é base de $\mathbb{W}$, precisamos verificar se V_1, V_2 e V_3 são L.I. Para isto temos que saber se a equação vetorial

$$xV_1 + yV_2 + zV_3 = \bar{0} \tag{4.6}$$

ou equivalentemente,

$$A\,X = \bar{0}, \quad \text{em que} \quad A = [\ V_1\ V_2\ V_3\ ]$$

só possui a solução trivial. Escalonando a matriz A, obtemos

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

Logo 4.6 tem solução não trivial. Assim os vetores V_1, V_2 e V_3 são L.D. A solução de (4.6) é dada por $x = -\alpha, y = -\alpha$ e $z = \alpha$, para todo $\alpha \in \mathbb{R}$. Substituindo-se esta solução em (4.6) obtemos

$$-\alpha V_1 - \alpha V_2 + \alpha V_3 = \bar{0}$$

Tomando-se $\alpha = 1$ obtemos $V_3 = V_2 + V_1$. Assim o vetor V_3 pode ser descartado na geração de $\mathbb{W}$, pois ele é combinação linear dos outros dois. Logo, apenas V_1 e V_2 são suficientes para gerar $\mathbb{W}$. Como além disso, os vetores V_1 e V_2 são tais que um não é múltiplo escalar do outro, então eles são L.I. e portanto $\{V_1, V_2\}$ é uma base de $\mathbb{W}$. Observe que a mesma relação que vale entre as colunas de R vale entre as colunas de A (por que?).

Exemplo 4.12. Considere os vetores $V_1 = (-1, 1, 0, -3)$ e $V_2 = (-3, 3, 2, -1)$ linearmente independentes de $\mathbb{R}^4$. Vamos encontrar vetores V_3 e V_4 tais que $\{V_1, V_2, V_3, V_4\}$ formam uma base de $\mathbb{R}^4$. Escalonando a matriz cujas linhas são os vetores V_1 e V_2,

$$A = \begin{bmatrix} -1 & 1 & 0 & -3 \\ -3 & 3 & 2 & -1 \end{bmatrix}, \quad \text{obtemos} \quad R = \begin{bmatrix} 1 & -1 & 0 & 3 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

Vamos inserir linhas que são vetores da base canônica na matriz R até conseguir uma matriz 4×4 triangular superior com os elementos da diagonal diferentes de zero. Neste caso acrescentando as linhas $V_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}$ e $V_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$ em posições adequadas obtemos a matriz

$$\bar{R} = \begin{bmatrix} 1 & -1 & 0 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Vamos verificar que V_1, V_2, V_3 e V_4 são L.I.

$$x_1 V_1 + x_2 V_2 + x_3 V_3 + x_4 V_4 = \bar{0}$$

é equivalente ao sistema linear

$$CX = \bar{0}, \quad \text{em que } C = [\; V_1 \; V_2 \; V_3 \; V_4 \;].$$

Mas como $\det(\bar{R}) \neq 0$, então $\det(C) \neq 0$, pelo Teorema 2.13 na página 110, pois $\bar{R}$ pode ser obtida de C^t aplicando-se operações elementares. Logo $\{V_1, V_2, V_3, V_4\}$ é L.I. Como a dimensão do $\mathbb{R}^4$ é igual a 4 , então pelo Teorema 4.2 na página 236, $\{V_1, V_2, V_3, V_4\}$ é uma base de $\mathbb{R}^4$.

Exercícios Numéricos (respostas na página 515)

4.1.1. Encontre um conjunto de geradores para o espaço solução do sistema homogêneo $AX = \bar{0}$, em que

(a) $A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 3 & 1 \\ 2 & 1 & 3 & 1 \end{bmatrix}$;

(b) $A = \begin{bmatrix} 1 & 1 & 2 & -1 \\ 2 & 3 & 6 & -2 \\ -2 & 1 & 2 & 2 \end{bmatrix}$.

4.1.2. Encontre os valores de λ tais que o sistema homogêneo $(A - \lambda I_n)X = \bar{0}$ tem solução não trivial e para estes valores de λ, encontre uma base para o espaço solução, para as matrizes A dadas:

(a) $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{bmatrix}$;

(b) $A = \begin{bmatrix} 2 & 2 & 3 & 4 \\ 0 & 2 & 3 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$;

(c) $A = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$;

(d) $A = \begin{bmatrix} -1 & 2 & 2 & 0 \\ -1 & 2 & 1 & 0 \\ -1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(e) $A = \begin{bmatrix} 2 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$;

(f) $A = \begin{bmatrix} 2 & 3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$;

4.1.3. Determine uma base para a reta interseção dos planos $x - 7y + 5z = 0$ e $3x - y + z = 0$.

4.1.4. Sejam $V_1 = (4, 2, -3)$, $V_2 = (2, 1, -2)$ e $V_3 = (-2, -1, 0)$.

(a) Mostre que V_1, V_2 e V_3 são L.D.

(b) Mostre que V_1 e V_2 são L.I.

(c) Qual a dimensão do subespaço gerado por V_1, V_2 e V_3, ou seja, do conjunto das combinações lineares de V_1, V_2 e V_3.

(d) Descreva geometricamente o subespaço gerado por V_1, V_2 e V_3

4.1.5. Dados $V_1 = (2, 1, 3)$ e $V_2 = (2, 6, 4)$:

(a) Os vetores V_1 e V_2 geram o $\mathbb{R}^3$? Justifique.

(b) Seja V_3 um terceiro vetor do $\mathbb{R}^3$. Quais as condições sobre V_3, para que $\{V_1, V_2, V_3\}$ seja uma base de $\mathbb{R}^3$?

(c) Encontre um vetor V_3 que complete junto com V_1 e V_2 uma base do $\mathbb{R}^3$.

4.1.6. Seja $\mathbb{W}$ o plano $x + 2y + 4z = 0$. Obtenha uma base $\{V_1, V_2, V_3\}$ de $\mathbb{R}^3$ tal que V_1 e V_2 pertençam a $\mathbb{W}$.

4.1.7. Considere os seguintes subespaços de $\mathbb{R}^3$:

$$\mathbb{V} = [(-1,2,3),(1,3,4)] \quad \text{e} \quad \mathbb{W} = [(1,2,-1),(0,1,1)].$$

Encontre as equações paramétricas da reta $\mathbb{V} \cap \mathbb{W}$ e uma base para o subespaço $\mathbb{V} \cap \mathbb{W}$. A notação $[V_1, V_2]$ significa o subespaço gerado por V_1 e V_2, ou seja, o conjunto de todas as combinações lineares de V_1 e V_2.

4.1.8. Seja $\mathbb{V} = \{(3a + 4b - 4c, 2a - 4b - 6c, -2a - 4b + 2c) \mid a, b, c \in \mathbb{R}\}$ um subespaço de $\mathbb{R}^3$.

(a) Determine um conjunto de geradores para $\mathbb{V}$.

(b) Determine uma base para $\mathbb{V}$.

4.1.9. Dados $V_1 = (-3, 5, 2, 1)$ e $V_2 = (1, -2, -1, 2)$:

(a) Os vetores V_1 e V_2 geram o $\mathbb{R}^4$? Justifique.

(b) Sejam V_3 e V_4 vetores do $\mathbb{R}^4$. Quais as condições sobre V_3 e V_4 para que $\{V_1, V_2, V_3, V_4\}$ seja uma base de $\mathbb{R}^4$?

(c) Encontre vetores V_3 e V_4 que complete junto com V_1 e V_2 uma base do $\mathbb{R}^4$.

4.1.10. Dê exemplo de:

(a) Três vetores: V_1, V_2 e V_3, sendo $\{V_1\}$ L.I., $\{V_2, V_3\}$ L.I., V_2 e V_3 não são múltiplos de V_1 e $\{V_1, V_2, V_3\}$ L.D.

(b) Quatro vetores: V_1, V_2, V_3 e V_4, sendo $\{V_1, V_2\}$ L.I., $\{V_3, V_4\}$ L.I., V_3 e V_4 não são combinação linear de V_1 e V_2 e $\{V_1, V_2, V_3, V_4\}$ L.D.

Exercício usando o MATLAB®

4.1.11. Defina a matriz aleatória `A=triu(randi(4,4,3))`. Encontre os valores de λ tais que o sistema homogêneo $(A - \lambda I_4)X = \bar{0}$ tem solução não trivial e para estes valores de λ, encontre uma base para o espaço solução.

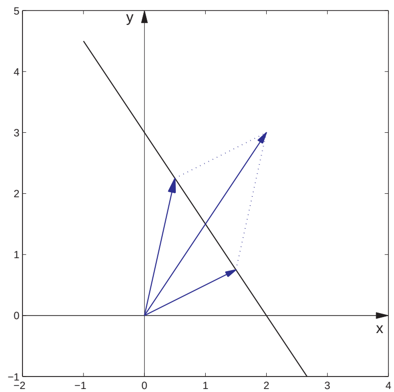
Exercícios Teóricos

4.1.12. Seja A uma matriz $m \times n$. Mostre que se o conjunto solução do sistema linear $AX = B$ é um subespaço, então $B = \bar{0}$, ou seja, o sistema linear é homogêneo. (Sugestão: se X é solução de $AX = B$, então $Y = 0\,X$ também o é.)

4.1.13. Determine uma base para o plano $ax + by + cz = 0$ no caso em que $b \neq 0$ e no caso em que $c \neq 0$.

4.1.14. Sejam V e W vetores do $\mathbb{R}^n$. Mostre que o conjunto dos vetores da forma $\alpha V + \beta W$ é um subespaço do $\mathbb{R}^n$.

4.1.15. Mostre que se uma reta em $\mathbb{R}^2$ ou em $\mathbb{R}^3$ não passa pela origem, então ela não é um subespaço. (Sugestão: se ela fosse um subespaço, então ...)



4.1.16. Sejam A uma matriz $m \times n$ e B uma matriz $m \times 1$. Mostre que o conjunto dos vetores B para os quais o sistema $A\,X = B$ tem solução é um subespaço de $\mathbb{R}^m$. Ou seja, mostre que o conjunto

$$\mathcal{I}(A) = \{B \in \mathbb{R}^m \,|\, B = A\,X, \text{ para algum } X \in \mathbb{R}^n\}$$

é um subespaço de $\mathbb{R}^m$.

4.1.17. Sejam $\mathbb{W}_1$ e $\mathbb{W}_2$ dois subespaços.

(a) Mostre que $\mathbb{W}_1 \cap \mathbb{W}_2$ é um subespaço.

(b) Mostre que $\mathbb{W}_1 \cup \mathbb{W}_2$ é um subespaço se, e somente se, $\mathbb{W}_1 \subseteq \mathbb{W}_2$ ou $\mathbb{W}_2 \subseteq \mathbb{W}_1$.

(c) Definimos a **soma dos subespaços** $\mathbb{W}_1$ e $\mathbb{W}_2$ por

$$\mathbb{W}_1 + \mathbb{W}_2 = \{V_1 + V_2 \mid V_1 \in \mathbb{W}_1 \text{ e } V_2 \in \mathbb{W}_2\}.$$

Mostre que $\mathbb{W}_1 + \mathbb{W}_2$ é um subespaço que contém $\mathbb{W}_1$ e $\mathbb{W}_2$.

4.1.18. Sejam $\mathbb{W}$ um subespaço de $\mathbb{R}^n$ e $\{W_1, \ldots, W_k\}$ uma base de $\mathbb{W}$. Defina a matriz $B = [\ W_1 \ldots W_k\ ]^t$, com $W_1, \ldots, W_k$ escritos como matrizes colunas. Sejam $\mathbb{W}^\perp$ o espaço solução do sistema homogêneo $BX = \bar{0}$ e $\{V_1, \ldots, V_p\}$ uma base de $\mathbb{W}^\perp$. Defina a matriz $A = [\ V_1 \ldots V_p\ ]^t$, com $V_1, \ldots, V_p$ escritos como matrizes colunas. Mostre que $\mathbb{W}$ é o espaço solução do sistema homogêneo $AX = \bar{0}$, ou seja,

$$\mathbb{W} = \{X \in \mathbb{R}^p \mid AX = \bar{0}\}.$$

4.1.19. Sejam A uma matriz $m \times n$ e B uma matriz $m \times 1$. Seja X_0 uma solução (particular) do sistema linear $A\,X = B$. Mostre que se $\{V_1, \ldots, V_k\}$ é uma base para o espaço solução do sistema homogêneo $AX = \bar{0}$, então toda solução de $A\,X = B$ pode ser escrita na forma

$$X = X_0 + \alpha_1 V_1 + \ldots + \alpha_k V_k,$$

em que $\alpha_1, \ldots, \alpha_k$ são escalares. (Sugestão: use o Exercício 1.2.21 na página 66)

4.1.20. Mostre que a dimensão do subespaço gerado pelas linhas de uma matriz escalonada reduzida é igual a dimensão do subespaço gerado pelas suas colunas.

4.1.21. Mostre que a dimensão do subespaço gerado pelas linhas de uma matriz é igual a dimensão do subespaço gerado pelas suas colunas. (Sugestão: Considere a forma escalonada reduzida da matriz A e use o exercício anterior.)

Apêndice III: Outros Resultados

Teorema 4.3. *Um subconjunto $\{V_1, V_2, \ldots, V_m\}$ de um subespaço $\mathbb{W}$ é uma base para $\mathbb{W}$ se, e somente se, todo vetor X de $\mathbb{W}$ é escrito de maneira única como combinação linear de $V_1, V_2, \ldots, V_m$.*

Demonstração. Em primeiro lugar, suponha que todo vetor X de $\mathbb{W}$ é escrito de maneira única como combinação linear de $V_1, \ldots, V_m$. Vamos mostrar que $\{V_1, V_2, \ldots, V_m\}$ é uma base de $\mathbb{W}$. Como todo vetor é escrito como combinação linear de $V_1, \ldots, V_m$, basta mostrarmos que $V_1, \ldots, V_m$ são L.I. Considere a equação

$$x_1V_1 + \ldots + x_mV_m = \bar{0}.$$

Como todo vetor é escrito de maneira única como combinação linear de $V_1, \ldots, V_m$, em particular temos que para $X = \bar{0}$,

$$x_1V_1 + \ldots + x_mV_m = \bar{0} = 0V_1 + \ldots + 0V_m,$$

o que implica que $x_1 = 0, \ldots, x_m = 0$, ou seja, $V_1, \ldots, V_m$ são linearmente independentes. Portanto, $\{V_1, V_2, \ldots, V_m\}$ é base de $\mathbb{W}$.

Suponha, agora, que $\{V_1, V_2, \ldots, V_m\}$ é base de $\mathbb{W}$. Seja X um vetor qualquer de $\mathbb{W}$. Se

$$x_1V_1 + \ldots + x_mV_m = X = y_1V_1 + \ldots + y_mV_m,$$

então

$$(x_1 - y_1)V_1 + \ldots + (x_m - y_m)V_m = \bar{0}.$$

Como $V_1, \ldots, V_m$ formam uma base de $\mathbb{W}$, então eles são L.I., o que implica que $x_1 = y_1, \ldots, x_m = y_m$. Portanto, todo vetor X de $\mathbb{W}$ é escrito de maneira única como combinação linear de $V_1, \ldots, V_m$. ■

Teorema 4.4. *Se $\mathcal{S} = \{V_1, \ldots, V_k\}$ é um conjunto de vetores que gera um subespaço $\mathbb{W}$, ou seja,*

$$\mathbb{W} = [\mathcal{S}] = [V_1, \ldots, V_k],$$

então existe um subconjunto de $\mathcal{S}$ que é base de $\mathbb{W}$.

Demonstração. Se $\mathcal{S}$ é L.I., então $\mathcal{S}$ é uma base de $\mathbb{W}$. Caso contrário, $\mathcal{S}$ é L.D. e pelo Teorema 3.11 na página 219, um dos vetores de $\mathcal{S}$ é combinação linear dos outros. Assim, o subconjunto de $\mathcal{S}$ obtido retirando-se este vetor continua gerando $\mathbb{W}$. Se esse subconjunto for L.I., temos uma base para $\mathbb{W}$, caso contrário, continuamos retirando vetores do subconjunto até obtermos um subconjunto L.I. e aí neste caso temos uma base para $\mathbb{W}$. ■

Vamos mostrar que se a dimensão de um subespaço $\mathbb{W}$ é m, então m vetores que geram o subespaço, $\mathbb{W}$, formam uma base (Corolário 4.5) e que não podemos ter menos que m vetores gerando o subespaço (Corolário 4.6).

São simples as demonstrações dos seguintes corolários, as quais deixamos como exercício.

Corolário 4.5. *Em um subespaço, $\mathbb{W}$, de dimensão $m > 0$, m vetores que geram o subespaço, são L.I. e portanto formam uma base.*

Corolário 4.6. *Em um subespaço, $\mathbb{W}$, de dimensão $m > 0$, um conjunto com menos de m vetores não gera o subespaço.*

Teorema 4.7. *Se $\mathcal{R} = \{V_1, \ldots, V_k\}$ é um conjunto de vetores L.I. em um subespaço $\mathbb{W}$ de $\mathbb{R}^n$, então o conjunto $\mathcal{R}$ pode ser completado até formar uma base de $\mathbb{W}$, ou seja, existe um conjunto $\mathcal{S} = \{V_1, \ldots, V_k, V_{k+1} \ldots, V_m\}$ ($\mathcal{R} \subseteq \mathcal{S}$), que é uma base de $\mathbb{W}$.*

Demonstração. Se $\{V_1, \ldots, V_k\}$ gera $\mathbb{W}$, então $\{V_1, \ldots, V_k\}$ é uma base de $\mathbb{W}$. Caso contrário, seja V_{k+1} um vetor que pertence a $\mathbb{W}$, mas não pertence ao subespaço gerado por $\{V_1, \ldots, V_k\}$. Então, o conjunto $\{V_1, \ldots, V_k, V_{k+1}\}$ é L.I., pois caso contrário $x_1 V_1 + \ldots + x_{k+1} V_{k+1} = \bar{0}$, implicaria que $x_{k+1} \neq 0$ (por que?) e assim, V_{k+1} seria combinação linear de $V_1, \ldots, V_k$, ou seja, V_{k+1} pertenceria ao subespaço $\mathbb{W}_k$. Se $\{V_1, \ldots, V_{k+1}\}$ gera $\mathbb{W}$, então $\{V_1, \ldots, V_{k+1}\}$ é uma base de $\mathbb{W}$. Caso contrário, o mesmo argumento é repetido para o subespaço gerado por $\{V_1, \ldots, V_k, V_{k+1}\}$.

Pelo Corolário 3.10 na página 218 este processo tem que parar, ou seja, existe um inteiro positivo $m \leq n$ tal que $\{V_1, \ldots, V_k, V_{k+1}, \ldots, V_m\}$ é L.I., mas $\{V_1, \ldots, V_k, V_{k+1}, \ldots, V_m, V\}$ é L.D. para qualquer vetor V de $\mathbb{W}$. O que implica que V é combinação linear de $\{V_1, \ldots, V_k, V_{k+1}, \ldots, V_m\}$ (por que?). Portanto, $\{V_1, \ldots, V_k, V_{k+1}, \ldots, V_m\}$ é uma base de $\mathbb{W}$. ■

Corolário 4.8. *Todo subespaço de $\mathbb{R}^n$ diferente do subespaço trivial $\{\bar{0}\}$ tem uma base e a sua dimensão é menor ou igual a n.*

Os próximos resultados são aplicações às matrizes.

Proposição 4.9. *Sejam A e B matrizes $m \times n$ equivalentes por linhas. Sejam $A_1, \ldots, A_n$ as colunas $1, \ldots, n$, respectivamente, da matriz A e $B_1, \ldots, B_n$ as colunas $1, \ldots, n$, respectivamente, da matriz B.*

(a) *$B_{j_1}, \ldots, B_{j_k}$ são L.I. se, e somente se, $A_{j_1}, \ldots, A_{j_k}$ também o são.*

(b) *Se existem escalares $\alpha_{j_1}, \ldots, \alpha_{j_k}$ tais que*

$$A_k = \alpha_{j_1} A_{j_1} + \cdots + \alpha_{j_k} A_{j_k},$$

então

$$B_k = \alpha_{j_1} B_{j_1} + \cdots + \alpha_{j_k} B_{j_k},$$

(c) *O subespaço gerado pelas linhas de A é igual ao subespaço gerado pelas linhas de B.*

Demonstração. Se B é equivalente por linhas a A, então B pode ser obtida de A aplicando-se uma sequência de operações elementares. Aplicar uma operação elementar a uma matriz corresponde a multiplicar a matriz à esquerda por uma matriz invertível (Teorema 1.8 na página 54). Seja M o produto das matrizes invertíveis correspondentes às operações elementares aplicadas na matriz A para se obter a matriz B. Então M é invertível e $B = MA$.

(a) Vamos supor que $B_{j_1}, \ldots, B_{j_k}$ são L.I. e vamos mostrar que $A_{j_1}, \ldots, A_{j_k}$ também o são. Se

$$x_{j_1} A_{j_1} + \cdots + x_{j_k} A_{j_k} = \bar{0},$$

então multiplicando-se à esquerda pela matriz M obtemos

$$x_{j_1} M A_{j_1} + \cdots + x_{j_k} M A_{j_k} = \bar{0}.$$

Como $MA_j = B_j$, para $j = 1, \ldots, n$ (Exercício 1.1.18 (a) na página 25), então

$$x_{j_1} B_{j_1} + \cdots + x_{j_k} B_{j_k} = \bar{0}.$$

Assim, se $B_{j_1}, \ldots, B_{j_k}$ são L.I., então $x_{j_1} = \ldots = x_{j_k} = 0$. O que implica que $A_{j_1}, \ldots, A_{j_k}$ também são L.I.

Trocando-se B por A o argumento acima mostra que se $A_{j_1}, \ldots, A_{j_k}$ são L.I., então $B_{j_1}, \ldots, B_{j_k}$ também o são.

(b) Sejam $\alpha_{j_1}, \ldots, \alpha_{j_k}$ escalares tais que

$$A_k = \alpha_{j_1} A_{j_1} + \cdots + \alpha_{j_k} A_{j_k},$$

então multiplicando-se à esquerda pela matriz M obtemos

$$MA_k = \alpha_{j_1} M A_{j_1} + \cdots + \alpha_{j_k} M A_{j_k}.$$

Como $MA_j = B_j$, para $j = 1, \ldots, n$ (Exercício 1.1.18 (a) na página 25), então

$$B_k = \alpha_{j_1} B_{j_1} + \cdots + \alpha_{j_k} B_{j_k}.$$

(c) A matriz B é obtida de A aplicando-se uma sequência de operações elementares às linhas de A. Assim, toda linha de B é uma combinação linear das linhas de A. Logo, o espaço gerado pelas linhas de B está contido no espaço gerado pelas linhas de A. Como toda operação elementar tem uma operação elementar inversa, o argumento anterior também mostra que o espaço gerado pelas linhas de A está contido no espaço gerado pelas linhas de B. Portanto, eles são iguais.

■

Somente agora vamos provar a unicidade da forma escalonada reduzida.

Teorema 4.10. *Se $R = (r_{ij})_{m \times n}$ e $S = (s_{ij})_{m \times n}$ são matrizes escalonadas reduzidas equivalentes por linhas a uma matriz $A = (a_{ij})_{m \times n}$, então $R = S$.*

Demonstração. Sejam S e R matrizes escalonadas reduzidas equivalentes a A. Sejam $R_1, \ldots, R_n$ as colunas de R e $S_1, \ldots, S_n$ as colunas de S. Seja r o número de linhas não nulas de R. Sejam $j_1, \ldots, j_r$ as colunas onde ocorrem os pivôs das linhas $1, \ldots, r$, respectivamente, da matriz R. Então R e S são equivalentes por linha, ou seja, existe uma sequência de operações elementares que podemos aplicar em R para chegar a S e uma outra sequência de operações elementares que podemos aplicar a S e chegar a R.

Assim, como as colunas $1, \ldots, j_1 - 1$ de R são nulas o mesmo vale para as colunas $1, \ldots, j_1 - 1$ de S. Logo o pivô da 1ª linha de S ocorre numa coluna maior ou igual a j_1. Trocando-se R por S e usando este argumento chegamos a conclusão que $R_{j_1} = S_{j_1}$ e assim $R_1 = S_1, \ldots, R_{j_1} = S_{j_1}$.

Vamos supor que $R_1 = S_1, \ldots, R_{j_k} = S_{j_k}$ e vamos mostrar que

$$R_{j_k+1} = S_{j_k+1}, \ldots, R_{j_{k+1}} = S_{j_{k+1}}, \quad \text{se } k < r \text{ ou}$$

$$R_{j_r+1} = S_{j_r+1}, \ldots, R_n = S_n, \quad \text{se } k = r.$$

Observe que para $j = j_k + 1, \ldots, j_{k+1} - 1$, se $k < r$, ou para $j = j_r + 1, \ldots, n$, se $k = r$, temos que

$$R_j = (r_{1j}, \ldots, r_{kj}, 0, \ldots, 0) = r_{1j} R_{j_1} + \ldots + r_{kj} R_{j_k},$$

o que implica pela Proposição 4.9 (b) na página 249 que

$$S_j = r_{1j}S_{j_1} + \ldots + r_{kj}S_{j_k}.$$

Mas por hipótese $R_{j_1} = S_{j_1}, \ldots, R_{j_k} = S_{j_k}$, então,

$$S_j = r_{1j}R_{j_1} + \ldots + r_{kj}R_{j_k} = R_j,$$

para $j = j_k + 1, \ldots, j_{k+1} - 1$, se $k < r$ ou para $j = j_r + 1, \ldots, n$, se $k = r$.

Logo, se $k < r$, o pivô da $(k+1)$-ésima linha de S ocorre numa coluna maior ou igual a j_{k+1}. Trocando-se R por S e usando o argumento anterior chegamos a conclusão que $R_{j_{k+1}} = S_{j_{k+1}}$ e assim $R_1 = S_1, \ldots, R_{j_r} = S_{j_r}$. E se $k = r$, então $R_1 = S_1, \ldots, R_n = S_n$.

Portanto $R = S$. ■

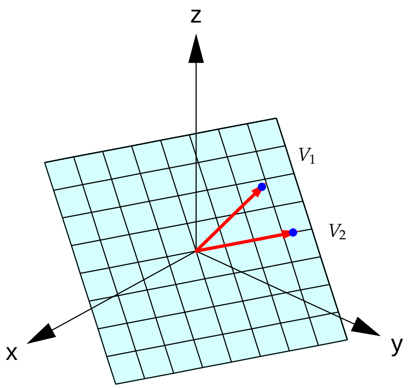


Figura 4.5. V_1 e V_2 que formam uma base para o plano

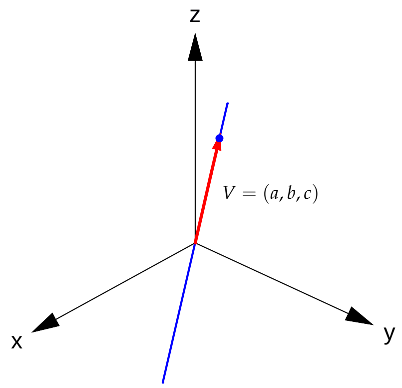


Figura 4.6. Vetor $V = (a, b, c)$ que é base para a reta $(x, y, z) = t(a, b, c)$

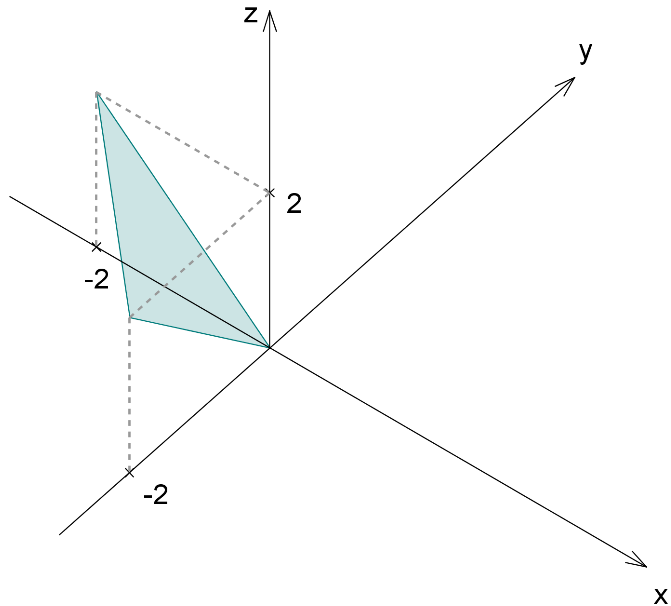


Figura 4.7. O subespaço $\mathbb{W}$ do Exemplo 4.10

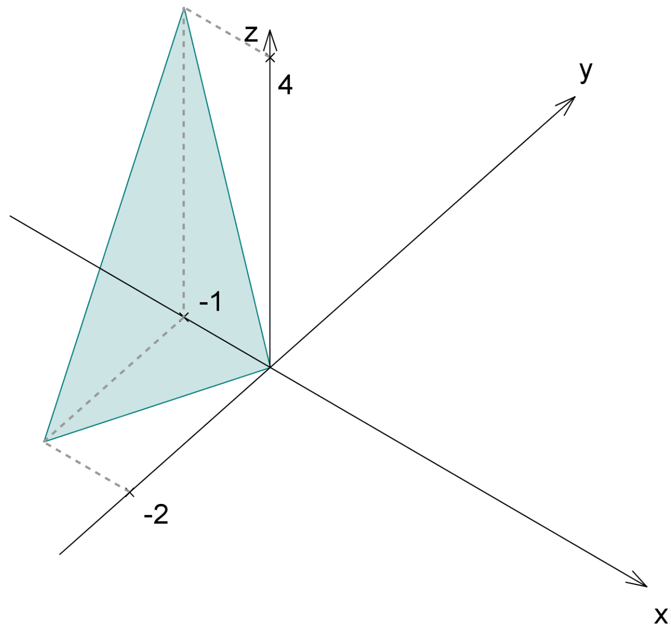


Figura 4.8. O subespaço $\mathbb{V}$ do Exemplo 4.10

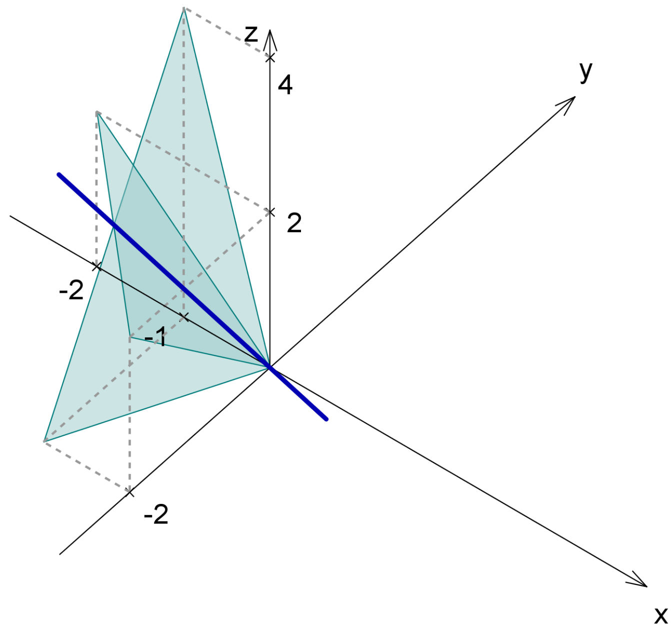


Figura 4.9. Os subespaços $\mathbb{W}$, $\mathbb{V}$ e $\mathbb{V} \cap \mathbb{W}$ do Exemplo 4.10

4.2 Espaço Linha e Espaço Coluna

Definição 4.3. Seja A uma matriz $m \times n$.

(a) O subespaço de $\mathbb{R}^n$ gerado pelas linhas de A é chamado **espaço linha de** A, ou seja, o conjunto de todas as combinações lineares das linhas de A.

(b) O subespaço de $\mathbb{R}^m$ gerado pelas colunas de A é chamado **espaço coluna de** A, ou seja, o conjunto de todas as combinações lineares das colunas de A.

Os espaços linha e coluna de uma matriz são diferentes, em geral, mesmo se a matriz é quadrada, como mostra o próximo exemplo.

Exemplo 4.13. Considere a matriz

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

O espaço linha de A é o subespaço gerado pelo vetor $(1,1)$, enquanto o espaço coluna de A é o subespaço gerado pelo vetor $(1,0)$.

Apesar dos espaços linha e coluna de uma matriz serem diferentes, em geral, eles possuem sempre a mesma dimensão.

Teorema 4.11. *Seja A uma matriz $m \times n$. O espaço linha e o espaço coluna de A possuem a mesma dimensão.*

Demonstração. Seja R uma matriz escalonada reduzida equivalente a matriz A.

(a) Pela Proposição 4.9 (c) na página 249 os espaços linha de A e de R são iguais.

(b) Pela Proposição 4.9 (a) na página 249 as colunas $j_1, \ldots, j_k$ da matriz R são L.I. se, somente se, as colunas $j_1, \ldots, j_k$ da matriz A também o são.

Pelo item (a) a dimensão do espaço linha de A é igual a dimensão do espaço linha de R e pelo item (b) a dimensão do espaço coluna de A é igual a dimensão do espaço coluna de R. Portanto, basta provarmos o teorema para a matriz escalonada reduzida R. Agora, a dimensão do espaço linha de R é igual ao número de linhas não nulas, pois estas são linearmente independentes (verifique!). A dimensão do espaço coluna de R é igual ao número de pivôs, pois as outras colunas são combinação linear das colunas dos pivôs e podem, portanto, ser descartadas para gerar o espaço coluna de R. Portanto, a dimensão dos dois espaços são iguais. ■

4.2.1 Posto e Nulidade

Definição 4.4. Seja A uma matriz $m \times n$.

(a) O **posto de** A é a dimensão do espaço linha ou do espaço coluna de A, ou seja, é o número máximo de linhas e colunas L.I. da matriz A.

(b) A **nulidade de** A é a dimensão do espaço solução de $A\,X = \bar{0}$.

Exemplo 4.14. Considere a matriz $A = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 4 & -3 & 0 \\ 1 & 2 & 1 & 5 \end{bmatrix}$.

A forma escalonada reduzida da matriz A é a matriz $R = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. As linhas não nulas de R, $V_1 = (1, 2, 0, 3)$ e $V_2 = (0, 0, 1, 2)$, formam uma base para o espaço linha de A. Portanto, o posto de A é igual a 2.

Quanto ao espaço coluna, sejam W_1, W_2, W_3 e W_4 as colunas de A. Sejam U_1, U_2, U_3 e U_4 as colunas de R. As colunas sem pivôs podem ser descartadas na geração do espaço coluna de R, pois elas são combinação linear das colunas dos pivôs. As colunas correspondentes de A podem, também, ser descartadas na geração do espaço coluna de A, pois os mesmos escalares que fazem a combinação linear nula de W_1, W_2, W_3 e W_4, fazem a combinação linear nula de U_1, U_2, U_3 e U_4. Assim, $W_1 = (1, 2, 1)$ e $W_3 = (-1, -3, 1)$ formam uma base para o espaço coluna de A.

Vamos apresentar a seguir uma outra forma de calcular o posto de uma matriz. Uma **submatriz** de uma matriz A é a própria matriz A ou qualquer matriz obtida de A retirando-se linha(s) e/ou coluna(s).

Teorema 4.12. *Seja A uma matriz $m \times n$.*

(a) *O posto de A é igual a $p = \min\{m, n\}$ se, e somente se, o determinante de uma submatriz $p \times p$ é diferente de zero.*

(b) *O posto de A é igual ao maior inteiro positivo r tal que alguma submatriz $r \times r$ de A possui determinante não nulo.*

Demonstração. (a) Podemos supor que $m \leq n$, já que o posto de A é igual ao posto de A^t. Neste caso, $p = m$ e o posto de A é m se, e somente se, existem m colunas linearmente independentes. Mas existem m colunas linearmente independentes se, e somente se, existe uma submatriz $m \times m$ cujas colunas são linearmente independentes, ou seja, com o seu determinante diferente de zero.

(b) Se as colunas de A são L.I., então posto$(A) = \min\{m, n\}$ e o resultado decorre do item anterior. Caso contrário, existe uma coluna de A que é combinação linear das outras e o posto de A é igual ao posto da submatriz obtida de A retirando-se esta coluna. Este processo pode ser continuado até se obter uma submatriz cujas colunas são linearmente independentes e cujo posto é igual ao de A. O posto desta submatriz é igual ao mínimo entre o seu número de linhas e o seu número de colunas e é também igual ao posto de A. Aplicando-se o item anterior a esta submatriz obtemos o resultado.

■

Exemplo 4.15. Considere a seguinte matriz $A = \begin{bmatrix} a & 3 & a \\ 3 & a & -a \end{bmatrix}$.

Se

$$\det \begin{bmatrix} a & 3 \\ 3 & a \end{bmatrix} = a^2 - 9 = (a-3)(a+3) = 0,$$

$$\det \begin{bmatrix} a & a \\ 3 & -a \end{bmatrix} = -a^2 - 3a = -a(a+3) = 0$$

e

$$\det \begin{bmatrix} 3 & a \\ a & -a \end{bmatrix} = a^2 + 3a = a(a+3) = 0,$$

então o posto de A é igual a 1. Assim, posto$(A) = 1$ se, e somente se, $a = -3$. Caso contrário, o posto de A é igual a 2.

4.2.2 Aplicação a Sistemas Lineares

Os conceitos de espaço linha e espaço coluna são úteis no estudo de sistemas lineares. O sistema $A\,X = B$ é equivalente a equação

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Assim, o sistema $A\,X = B$ tem solução se, e somente se, B é uma combinação linear das colunas de A, ou seja, se, e somente se, B pertence ao espaço coluna de A.

Proposição 4.13. *Sejam A uma matriz $m \times n$ e B uma matriz $m \times 1$. O sistema linear $A\,X = B$ tem solução se, e somente se, B pertence ao espaço coluna de A.*

O espaço solução do sistema $A\,X = \bar{0}$ é chamado de **núcleo de** A e é denotado por $\mathcal{N}(A)$.

Proposição 4.14. *Seja A uma matriz $m \times n$. O sistema linear $A\,X = B$, **para todo** $B \in \mathbb{R}^m$,*

(a) *tem solução se, e somente se, as colunas de A geram o $\mathbb{R}^m$ ($\text{posto}(A) = m$);*

(b) *tem no máximo uma solução se, e somente se, as colunas de A são linearmente independentes ($\mathcal{N}(A) = \{\bar{0}\}$).*

Demonstração. (a) Pela Proposição 4.13, o sistema tem solução para todo $B \in \mathbb{R}^m$ se, e somente se, o espaço coluna de A é igual ao $\mathbb{R}^m$, daí decorre o resultado.

(b) Se o sistema $AX = B$ tem no máximo uma solução para todo $B \in \mathbb{R}^m$, então o sistema homogêneo $AX = \bar{0}$ tem somente a solução trivial, daí segue-se que as colunas de A são linearmente independentes e $\mathcal{N}(A) = \{\bar{0}\}$. Se por outro lado, $\mathcal{N}(A) = \{\bar{0}\}$, ou seja, a única solução do sistema homogêneo $AX = \bar{0}$ é a solução trivial, e X_1 e X_2 são soluções de $AX = B$, então $X_1 - X_2$ é solução de $AX = \bar{0}$ (verifique!). Assim, $X_1 - X_2 = \bar{0}$, ou seja, $X_1 = X_2$.

■

Segue do item (a) da proposição anterior que um sistema linear $AX = B$ com mais incógnitas do que equações não pode ter solução para todo B. Segue também da proposição anterior, que o sistema $AX = B$ tem exatamente uma solução para todo $B \in \mathbb{R}^m$ se, e somente se, as colunas de A formam uma base do $\mathbb{R}^m$. E isto ocorre se, e somente se, $m = n$ e uma das duas condições ocorre: ou $\mathcal{N}(A) = \{\bar{0}\}$ ou $\text{posto}(A) = n = m$.

Teorema 4.15. *Sejam A uma matriz $m \times n$ e B uma matriz $m \times 1$. O sistema linear $A\,X = B$,*

(a) *tem solução se, e somente se,* $\text{posto}([\,A\,|\,B\,]) = \text{posto}(A)$;

(b) *tem solução única se, e somente se,* $\text{posto}([\,A\,|\,B\,]) = \text{posto}(A) = n$.

Demonstração. (a) Suponha, em primeiro lugar, que o sistema $AX = B$ tem solução. Então, B é combinação linear das colunas de A. Portanto, o espaço coluna de $[\,A\,|\,B\,]$ é igual ao espaço coluna de A, ou seja,

$$\text{posto}([\,A\,|\,B\,]) = \text{posto}(A).$$

Por outro lado, se $\text{posto}([\,A\,|\,B\,]) = \text{posto}(A)$, então B pertence ao espaço coluna de A, ou seja, B é combinação linear das colunas de A. Portanto, o sistema $AX = B$ tem solução.

(b) Do item anterior podemos assumir que $AX = B$ tem solução. Seja X_0 uma solução particular de $AX = B$. Então, $Y = X + X_0$ é solução de $AX = B$ se, e somente se, X é solução do sistema homogêneo $AX = \bar{0}$. Assim, $AX = B$ tem solução única se, e somente se, o sistema homogêneo $AX = \bar{0}$, tem somente a solução trivial. E isto acontece se, e somente se, as colunas de A são linearmente independentes, ou seja, as colunas de A formam uma base para o seu espaço coluna ou equivalentemente $\text{posto}(A) = n$.

■

Exemplo 4.16. Considere o sistema $\begin{cases} ax + 3y &= a \\ 3x + ay &= -a \end{cases}$.

Para este sistema, $A = \begin{bmatrix} a & 3 \\ 3 & a \end{bmatrix}$ e $B = \begin{bmatrix} a \\ -a \end{bmatrix}$.

O sistema tem solução única se, e somente se, posto$(A) = 2$ (neste caso, posto$(A) = 2$ implica que posto$([A|B]) = 2$). Agora, posto$(A) = 2$ se, e somente se, $\det(A) \neq 0$. Como $\det(A) = a^2 - 9$, então o sistema tem solução única se, e somente se, $a \neq \pm 3$.

O sistema tem infinitas soluções se, e somente se, posto$([A|B]) = 1$ (neste caso, posto$([A|B]) = 1$ implica que posto$(A) = 1$). Agora, posto$([A|B]) = 1$ se, e somente se, $\det(A) = 0, \det(A_1) = 0$ e $\det(A_2) = 0$, em que $A_1 = \begin{bmatrix} a & 3 \\ -a & a \end{bmatrix}$ e $A_2 = \begin{bmatrix} a & a \\ 3 & -a \end{bmatrix}$. Como

$$\det(A_1) = a^2 + 3a = a(a+3)$$

e

$$\det(A_2) = -a^2 - 3a = -a(a+3),$$

então o sistema tem infinitas soluções se, e somente se, $a = -3$.

O sistema não tem solução se, e somente se, posto$(A) = 1$ e posto$([A|B]) = 2$. Agora, posto$(A) = 1$ se, e somente se, $\det(A) = 0$. E posto$([A|B]) = 2$ se, e somente se, $\det(A) \neq 0$, ou $\det(A_1) \neq 0$ ou $\det(A_2) \neq 0$. Assim o sistema não tem solução se, e somente se, $a = 3$.

Proposição 4.16. *Sejam A uma matriz $m \times n$ e B uma matriz $m \times 1$. Seja X_0 uma solução (particular) do sistema linear $A\,X = B$. Se $V_1, \ldots, V_k$ formam uma base para o núcleo de A, então toda solução de $A\,X = B$ pode ser escrita na forma*

$$X = X_0 + \alpha_1 V_1 + \ldots + \alpha_k V_k, \tag{4.7}$$

em que $\alpha_1, \ldots, \alpha_k$ são escalares. Reciprocamente, para todos os escalares $\alpha_1, \ldots, \alpha_k$ a expressão (4.7) é solução do sistema $AX = B$.

Demonstração. Seja X uma solução qualquer do sistema $AX = B$. Então, $X - X_0$ é solução do sistema homogêneo, pois $A(X - X_0) = AX - AX_0 = B - B = \bar{0}$. Como $V_1, \ldots, V_k$ formam uma base para o núcleo de A, existem escalares $\alpha_1, \ldots, \alpha_k$ tais que

$$X - X_0 = \alpha_1 V_1 + \ldots + \alpha_k V_k.$$

Daí segue-se a equação (4.7). Por outro lado, se $\alpha_1, \ldots, \alpha_k$ são escalares, então

$$A(X_0 + \alpha_1 V_1 + \ldots + \alpha_k V_k) = AX_0 + \alpha_1 AV_1 + \ldots + \alpha_k AV_k = B + \bar{0} + \ldots + \bar{0} = B,$$

ou seja, a expressão (4.7) é solução de $AX = B$. ■

4.2.3 A Imagem de uma Matriz

Seja A uma matriz $m \times n$. Pela Proposição 4.13 na página 261, o espaço coluna de A é igual ao conjunto dos vetores $B \in \mathbb{R}^m$ tais que o sistema linear $AX = B$ tem solução, ou seja, é igual ao conjunto

$$\mathcal{I}(A) = \{B \in \mathbb{R}^m \mid A\,X = B \text{ para algum } X \in \mathbb{R}^n\},$$

chamado de **imagem de** A, por que é a imagem da função $f : \mathbb{R}^n \to \mathbb{R}^m$ que associa a cada vetor $X \in \mathbb{R}^n$ o vetor $A\,X \in \mathbb{R}^m$, $f(X) = AX$. De forma análoga, se vê que o espaço linha de A é igual a imagem de A^t. A função $f : \mathbb{R}^n \to \mathbb{R}^m$ definida por $f(X) = AX$, para uma matriz $m \times n$ é chamada **transformação linear associada à matriz** A.

Proposição 4.17. *Seja A uma matriz $m \times n$. O espaço coluna de A é igual a $\mathcal{I}(A)$ e o espaço linha é igual a $\mathcal{I}(A^t)$.*

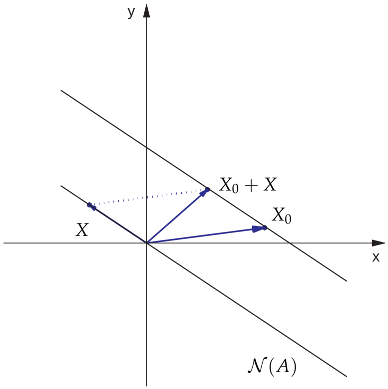


Figura 4.10. Solução de $AX = B$ e de $AX = \bar{0}$, se $\mathcal{N}(A) \neq \{\bar{0}\}$

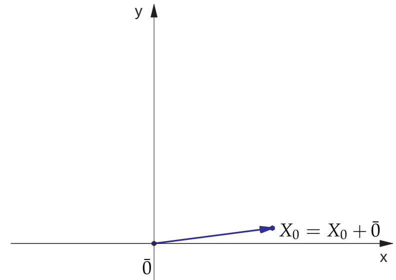


Figura 4.11. Solução de $AX = B$ e de $AX = \bar{0}$, se $\mathcal{N}(A) = \{\bar{0}\}$

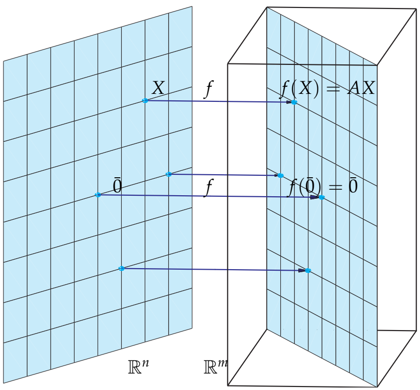


Figura 4.12. A função $f : \mathbb{R}^n \to \mathbb{R}^m$ dada por $f(X) = AX$

A dimensão do núcleo de A (nulidade), e a dimensão da imagem de A (posto) não são independentes um do outro, como mostra o próximo resultado.

Teorema 4.18 (da Dimensão do Núcleo e da Imagem). *Seja A uma matriz $m \times n$. A soma da dimensão do núcleo de A (nulidade) com a dimensão da imagem de A (posto) é igual ao número de colunas da matriz A, ou seja,*

$$\dim(\mathcal{N}(A)) + \dim(\mathcal{I}(A)) = n \quad ou \quad \text{nulidade}(A) + \text{posto}(A) = n.$$

Demonstração. Seja $\mathbb{V} = \mathbb{R}^n$. Sejam $V_1, \ldots, V_p$ vetores de $\mathbb{V}$, que formam uma base para o núcleo de A. Vamos estendê-la a uma base de $\mathbb{V}$. Sejam $V_{p+1}, \ldots, V_n$ vetores de $\mathbb{V}$ tais que $V_1, \ldots, V_n$ formam uma base de $\mathbb{V}$. Vamos mostrar que $AV_{p+1}, \ldots, AV_n$ formam uma base da imagem de A. Para isso, precisamos mostrar que eles geram a imagem de A e que eles são L.I.

Vamos mostrar, em primeiro lugar, que $AV_{p+1}, \ldots, AV_n$ geram a imagem de A. Seja $Y \in \mathcal{I}(A)$. Então existe $X \in \mathbb{V}$ tal que $A\,X = Y$. Como $V_1, \ldots, V_n$ é base de $\mathbb{V}$, existem escalares $\alpha_1, \ldots, \alpha_n$ tais que $X = \alpha_1 V_1 + \ldots + \alpha_n V_n$. Multiplicando à esquerda por A e usando que $AV_1 = \ldots = AV_p = \bar{0}$, obtemos que $\alpha_{p+1}AV_{p+1} + \ldots + \alpha_n AV_n = Y$, ou seja, $AV_{p+1}, \ldots, AV_n$ geram a imagem de A.

Vamos mostrar, agora, que $AV_{p+1}, \ldots, AV_n$ são linearmente independentes. Se $x_{p+1}AV_{p+1} + \ldots + x_n AV_n = \bar{0}$, então $A(x_{p+1}V_{p+1} + \ldots + x_n V_n) = \bar{0}$. Mas, isto implica que $x_{p+1}V_{p+1} + \ldots + x_n V_n$ pertence ao núcleo de A, ou seja, existem escalares $x_1, \ldots, x_p$ tais que $x_{p+1}V_{p+1} + \ldots + x_n V_n = x_1 V_1 + \ldots + x_p V_p$. Daí segue-se que $x_1V_1 + \ldots + x_pV_p - x_{p+1}V_{p+1} - \ldots - x_nV_n = \bar{0}$. Como $V_1, \ldots, V_n$ é base, então $x_1 = \ldots = x_p = x_{p+1} = \ldots = x_n = 0$, ou seja, $AV_{p+1}, \ldots, AV_n$ são L.I.

Portanto, a dimensão da imagem de A é igual a diferença entre n e a dimensão do núcleo de A. Daí segue-se o resultado. ■

Segue do Teorema da Dimensão do Núcleo e da Imagem que o número de variáveis livres na solução geral de um sistema linear $AX = B$ é igual a dimensão do núcleo de A.

Exercícios Numéricos (respostas na página 536)

4.2.1. Para cada uma das seguintes matrizes, encontre uma base para o espaço linha e para o espaço coluna.

(a) $\begin{bmatrix} 1 & 4 & 5 & 2 \\ 2 & 1 & 3 & 0 \\ -1 & 3 & 2 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -4 & -5 & 4 \\ -1 & 4 & 4 & -5 \\ 0 & 0 & 2 & 0 \end{bmatrix}$

4.2.2. Determine a dimensão do subespaço de $\mathbb{R}^3$ gerado pelos vetores:

(a) $(1,-2,2),(2,-2,4),(-3,3,6)$

(b) $(1,-3,4),(6,2,-1),(2,-2,3),(-4,-8,9)$

4.2.3. Seja $A = \begin{bmatrix} 1 & 2 & 2 & 3 & 1 & 4 \\ 2 & 4 & 5 & 5 & 4 & 9 \\ 3 & 6 & 7 & 8 & 5 & 9 \end{bmatrix}$.

(a) Determine a forma escalonada reduzida U da matriz A. Quais as colunas de U que correspondem às variáveis livres. Escreva cada uma destas colunas como uma combinação linear das colunas correspondentes aos pivôs.

(b) Quais as colunas de A que correspondem aos pivôs de U? Estas colunas formam uma base para o espaço coluna de A. Escreva cada uma das colunas restantes como combinação linear das colunas da base.

4.2.4. Determine o posto e a nulidade das seguintes matrizes:

(a) $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 3 \\ 3 & -1 & 4 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & -1 & 2 & 0 \\ 3 & 1 & 0 & 0 \\ -1 & 2 & 4 & 0 \end{bmatrix}$

4.2.5. Discuta como o posto de A varia com t.

(a) $A = \begin{bmatrix} 1 & 1 & t \\ 1 & t & 1 \\ t & 1 & 1 \end{bmatrix}$ (b) $A = \begin{bmatrix} t & 3 & -1 \\ 3 & 6 & -2 \\ -1 & -3 & -t \end{bmatrix}$

4.2.6. Encontre o maior valor possível para posto(A) e o menor valor possível para nulidade(A).
(a) A é 2×3 (b) A é 3×2 (c) A é 3×3 (d) A é $m \times n$.

4.2.7. Seja A uma matriz não nula. Encontre os valores de posto(A) e de posto$([A|B])$ para os quais o sistema linear $AX = B$ tem solução única, não tem solução e tem infinitas soluções.
(a) A é 2×3 (b) A é 3×2 (c) A é 3×3 (d) A é $m \times n$

Exercícios Teóricos

4.2.8. Seja A uma matriz $n \times n$.

(a) A matriz A é invertível se, e somente se, $\mathcal{N}(A) = \{\bar{0}\}$.

(b) Mostre que posto$(A) = n$ se, e somente se, as colunas de A são linearmente independentes.

(c) Mostre que o sistema homogêneo $AX = \bar{0}$ tem solução não trivial se, e somente se, o posto$(A) < n$.

(d) Mostre que o posto de A é n se, e somente se, $\det(A) \neq 0$.

4.2.9. Sejam $X = [\, x_1 \ldots x_m \,]$ e $Y = [\, y_1 \ldots y_n \,]$ matrizes $1 \times m$ e $1 \times n$, respectivamente. Seja $A = X^t Y$. Mostre que $\{X\}$ é uma base para o espaço coluna de A e que $\{Y\}$ é uma base para o espaço linha. Qual é a dimensão do $\mathcal{N}(A)$?

4.2.10. Mostre que se A é uma matriz, $m \times n$, de posto igual a 1, então existem matrizes $X = [\, x_1 \ldots x_m \,]$ e $Y = [\, y_1 \ldots y_n \,]$, $1 \times m$ e $1 \times n$, respectivamente, tais que $A = X^t Y$. (Sugestão: Tome X tal que $\{X\}$ é uma base para o espaço coluna de A.)

4.2.11. Sejam A e B matrizes $m \times p$ e $p \times n$, respectivamente. Mostre que AB pode ser escrita como uma soma de p matrizes de posto igual a 1.

4.2.12. Sejam A e B matrizes $m \times p$ e $p \times n$, respectivamente. Seja $C = AB$. Mostre que:

(a) O espaço coluna de C está contido no espaço coluna de A.

(b) O espaço linha de C está contido no espaço linha de B.

(c) posto$(C) \leq \min(\text{posto}(A), \text{posto}(B))$.

(d) Se as colunas de A e B são linearmente independentes, então as colunas de C também são linearmente independentes.

(e) Se as linhas de A e B são linearmente independentes, então as linhas de C também são linearmente independentes.

(f) Se as colunas de B são linearmente dependentes, então as colunas de C também são linearmente dependentes.

(g) Se as linhas de A são linearmente dependentes, então as linhas de C também são linearmente dependentes.

(h) O núcleo de B está contido no núcleo de C.

4.2.13. Seja A uma matriz $m \times n$. Se P e Q são matrizes invertíveis $m \times m$ e $n \times n$, respectivamente, então A, PA e AQ possuem o mesmo posto.

4.2.14. Sejam A e B matrizes $n \times n$. Mostre que $AB = \bar{0}$ se, e somente se, o espaço coluna de B está contido no núcleo de A.

4.2.15. Seja A uma matriz $n \times n$ e B uma matriz $n \times 1$. Para cada i, defina a matriz A_i como sendo a matriz que se obtém de A substituindo-se a i-ésima coluna por B.

(a) Mostre que se para algum i, $\det(A_i) \neq 0$, então $\text{posto}([A|B]) = n$.

(b) Suponha que o $\det(A) = 0$. Mostre que se para algum i, $\det(A_i) \neq 0$, então o sistema $AX = B$ não tem solução.

(c) Mostre que se $\det(A_i) = 0$ para $i = 1, \ldots, n$ e $\det(A) = 0$, então tanto pode ocorrer que o sistema $AX = B$ tenha infinitas soluções, como pode ocorrer que ele não tenha solução.

4.3 Espaços Vetoriais Abstratos (opcional)

Podemos generalizar o conceito de vetores ainda mais. Vamos estabelecer um conjunto de axiomas, os quais se forem satisfeitos por um conjunto de elementos, estes serão chamados de vetores. Os axiomas serão escolhidos abstraindo-se as propriedades mais importantes de vetores no $\mathbb{R}^n$. Assim, os vetores do $\mathbb{R}^n$ satisfarão automaticamente estes axiomas.

Definição 4.5. Dizemos que um conjunto $\mathbb{V} \neq \varnothing$, munido de duas operações, uma soma e uma multiplicação por escalar:

(0) Se $V, W \in \mathbb{V}$, então $V + W \in \mathbb{V}$;

(0') Se $V \in \mathbb{V}$ e $\alpha \in \mathbb{R}$ (ou $\mathbb{C}$), então $\alpha V \in \mathbb{V}$;

é um **espaço vetorial sobre $\mathbb{R}$ (ou $\mathbb{C}$)** se satisfaz os seguintes axiomas:

(1) Para todos os $V, W \in \mathbb{V}$, $V + W = W + V$;

(2) Para todos os $V, W, U \in \mathbb{V}$, $V + (W + U) = (V + W) + U$;

(3) Existe um elemento $\bar{0} \in \mathbb{V}$, tal que $V + \bar{0} = \bar{0} + V = V$, para todo $V \in \mathbb{V}$;

(4) Para cada $V \in \mathbb{V}$, existe um elemento $-V \in \mathbb{V}$ tal que $V + (-V) = (-V) + V = \bar{0}$;

(5) Para todo $V \in \mathbb{V}$ e todos os escalares α e β, $\alpha(\beta V) = (\alpha\beta)V$;

(6) Para todos os $V, W \in \mathbb{V}$ e todo escalar α, $\alpha(V + W) = \alpha V + \alpha W$;

(7) Para todo $V \in \mathbb{V}$ e todos os escalares α e β, $(\alpha + \beta)V = \alpha V + \beta V$;

(8) Para todo $V \in \mathbb{V}$, $1\,V = V$.

Os elementos de $\mathbb{V}$ são chamados **vetores**. O vetor $\bar{0}$ é chamado **vetor nulo** e para cada $V \in \mathbb{V}$ o vetor $-V$ é chamado o **simétrico** ou **inverso aditivo** de V. A **diferença** de dois vetores é definida por $V - W = V + (-W)$. Se V e W são vetores tais que $W = \alpha V$, para algum escalar α, então dizemos que W é um **múltiplo escalar** de V.

Exemplo 4.17. Para n um número inteiro positivo, o conjunto $\mathbb{V} = \mathbb{R}^n$ com as operações de soma e multiplicação por escalar definidas em (3.17) e (3.18) é um espaço vetorial sobre $\mathbb{R}$, pelo Teorema 3.7 na página 210. Em particular $\mathbb{R}$ é um espaço vetorial sobre ele mesmo.

Exemplo 4.18. O conjunto dos números complexos, $\mathbb{C}$, com as operações usuais é um espaço vetorial sobre ele mesmo, mas é também um espaço vetorial sobre $\mathbb{R}$.

Exemplo 4.19. Segue das propriedades da álgebra matricial, Teorema 1.1 na página 9, que o conjunto $\mathcal{M}_{mn}$ de todas as matrizes $m \times n$ com entradas que são números reais (números complexos) com as operações usuais de soma e multiplicação por escalar é um espaço vetorial sobre $\mathbb{R}$ (sobre $\mathbb{C}$).

Exemplo 4.20. Seja $\mathcal{X}$ um conjunto não vazio qualquer. Seja $\mathcal{F}(\mathcal{X};\mathbb{R})$ o conjunto das funções reais, $f : \mathcal{X} \to \mathbb{R}$. Para f e g funções de $\mathcal{F}(\mathcal{X};\mathbb{R})$ e α um escalar definimos a soma $f + g$ por

$$(f+g)(x) = f(x) + g(x), \quad \text{para todo } x \in \mathcal{X}$$

e a multiplicação de f pelo escalar α por

$$(\alpha\, f)(x) = \alpha\, f(x), \quad \text{para todo } x \in \mathcal{X}.$$

Vamos mostrar que o conjunto $\mathcal{F}(\mathcal{X};\mathbb{R})$ é um espaço vetorial sobre $\mathbb{R}$. Sejam $f, g, h \in \mathcal{F}(\mathcal{X};\mathbb{R})$ e α, β escalares.

(1) $(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x)$, para todo $x \in \mathcal{X}$;

(2) $[f+(g+h)](x) = f(x) + (g+h)(x) = f(x) + (g(x)+h(x)) = (f(x)+g(x)) + h(x) =$
$= (f+g)(x) + h(x) = [(f+g)+h](x)$, para todo $x \in \mathcal{X}$;

(3) Seja $\bar{0}$ a função identicamente nula. $(f+\bar{0})(x) = f(x) + \bar{0}(x) = f(x)$, para todo $x \in \mathcal{X}$;

(4) Dada a função f definimos a função $-f$ por $(-f)(x) = -f(x)$, para todo $x \in \mathcal{X}$.
$[f+(-f)](x) = f(x) + (-f(x) = 0 = \bar{0}(x)$, para todo $x \in \mathcal{X}$;

(5) $[\alpha(\beta f)](x) = \alpha(\beta f)(x) = \alpha(\beta f(x)) = (\alpha\beta) f(x) = [(\alpha\beta) f](x)$, para todo $x \in \mathcal{X}$;

(6) $[\alpha(f+g)](x) = \alpha(f+g)(x) = \alpha(f(x)+g(x)) = \alpha f(x) + \alpha g(x) = (\alpha f)(x) + (\alpha g)(x) = (\alpha f + \alpha g)(x)$, para todo $x \in \mathcal{X}$;

(7) $[(\alpha+\beta) f](x) = (\alpha+\beta) f(x) = \alpha f(x) + \beta f(x) = (\alpha f)(x) + (\beta f)(x) = [\alpha f + \beta f](x)$, para todo $x \in \mathcal{X}$;

(8) $(1f)(x) = 1f(x) = f(x)$, para todo $x \in \mathcal{X}$;

Variando o conjunto $\mathcal{X}$ obtemos vários exemplos de espaço vetorial.

Se $\mathcal{X}$ é igual a $\{1, \ldots, n\}$, então $\mathcal{F}(\mathcal{X};\mathbb{R}) = \mathbb{R}^n$, pois podemos identificar cada vetor $(x_1, \ldots, x_n)$ de $\mathbb{R}^n$ com a função $f : \{1, \ldots, n\} \to \mathbb{R}$ definida por $f(1) = x_1, \ldots, f(n) = x_n$.

Se $\mathcal{X}$ é igual ao produto cartesiano $\{1, \ldots, m\} \times \{1, \ldots, n\}$, então $\mathcal{F}(\mathcal{X};\mathbb{R}) = \mathcal{M}_{mn}$, pois podemos identificar cada matriz $(a_{ij})_{mn}$ com a função $f : \{1, \ldots, m\} \times \{1, \ldots, n\} \to \mathbb{R}$ definida por

$$f(1,1) = a_{11}, \ldots, f(1,n) = a_{1n}, \ldots, f(m,1) = a_{m1}, \ldots, f(m,n) = a_{mn}.$$

Exemplo 4.21. O conjunto $\mathbb{R}^\infty$ das sequencias de números reais, ou seja, o conjunto das listas infinitas $(x_1, x_2, \ldots, x_n, \ldots)$ tais que $x_n \in \mathbb{R}$, para $n = 1, 2, 3, \ldots$, com as operações

$$(x_1, x_2, \ldots, x_n, \ldots) + (y_1, y_2, \ldots, y_n, \ldots) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n, \ldots)$$
$$\alpha(x_1, x_2, \ldots, x_n, \ldots) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n, \ldots)$$

é um espaço vetorial sobre $\mathbb{R}$. Pois $\mathbb{R}^\infty = \mathcal{F}(\{1, 2, 3, \ldots\}; \mathbb{R})$, já que podemos identificar cada sequência (x_n), com a função $f : \{1, \ldots, n, \ldots\} \to \mathbb{R}$ definida por $f(1) = x_1, \ldots, f(n) = x_n, \ldots$.

Exemplo 4.22. Seja $\mathcal{P} = \mathbb{R}[\,t\,]$ o conjunto dos polinômios sobre $\mathbb{R}$ em uma variável t, ou seja, o conjunto das expressões da forma

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n + \ldots = \sum_{j \in \mathbb{N}} a_j t^j,$$

em que existe um inteiro positivo n tal que $a_j = 0$, para todo inteiro $j > n$ e $a_0, a_1, \ldots, a_n \in \mathbb{R}$. O polinômio identicamente nulo é aquele em que $a_j = 0$, para todo $j \in \mathbb{N}$.

Sejam $p(t) = a_0 + a_1 t + \cdots + a_m t^m + \ldots = \sum_{j \in \mathbb{N}} a_j t^j$ e $q(t) = b_0 + b_1 t + \cdots + b_r t^r + \ldots = \sum_{j \in \mathbb{N}} b_j t^j$ dois polinômios quaisquer. A soma é definida por

$$p(t) + q(t) = (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_k + b_k)t^k + \ldots = \sum_{j \in \mathbb{N}} (a_j + b_j) t^j.$$

A multiplicação por um escalar α é definida por

$$\alpha p(t) = (\alpha a_0) + (\alpha a_1)t + \cdots + (\alpha a_n)t^n + \ldots = \sum_{j \in \mathbb{N}} (\alpha a_j) t^j.$$

Vamos mostrar que o conjunto $\mathcal{P} = \mathbb{R}[\,t\,]$ é um espaço vetorial sobre $\mathbb{R}$.
Sejam $p(t) = \sum_{j \in \mathbb{N}} a_j t^j, q(t) = \sum_{j \in \mathbb{N}} b_j t^j, r(t) = \sum_{j \in \mathbb{N}} c_j t^j$ e α, β escalares.

(1) $p(t)+q(t)=\sum_{j\in\mathbb{N}} a_j t^j + \sum_{j\in\mathbb{N}} b_j t^j = \sum_{j\in\mathbb{N}} (a_j+b_j)t^j = \sum_{j\in\mathbb{N}} (b_j+a_j)t^j = q(t)+p(t).$

(2) $p(t)+(q(t)+r(t)) = \sum_{j\in\mathbb{N}} a_j t^j + \left(\sum_{j\in\mathbb{N}} b_j t^j + \sum_{j\in\mathbb{N}} c_j t^j\right) = \sum_{j\in\mathbb{N}} [a_j+(b_j+c_j)]t^j =$
$\sum_{j\in\mathbb{N}} [(a_j+b_j)+c_j]t^j = \left(\sum_{j\in\mathbb{N}} a_j t^j + \sum_{j\in\mathbb{N}} b_j t^j\right) + \sum_{j\in\mathbb{N}} c_j t^j = (p(t)+q(t))+r(t).$

(3) Seja $\bar{0}(t)$ o polinômio nulo.
$p(t)+\bar{0}(t) = \sum_{j\in\mathbb{N}} a_j t^j + \sum_{j\in\mathbb{N}} 0t^j = \sum_{j\in\mathbb{N}} (a_j+0)t^j = \sum_{j\in\mathbb{N}} a_j t^j = p(t).$

(4) Defina o polinômio $(-p)(t) = \sum_{j\in\mathbb{N}} (-a_j)t^j$.
$p(t)+(-p(t)) = \sum_{j\in\mathbb{N}} a_j t^j + \sum_{j\in\mathbb{N}} (-a_j)t^j = \sum_{j\in\mathbb{N}} (a_j+(-a_j))t^j = \sum_{j\in\mathbb{N}} 0t^j = \bar{0}(t).$

(5) $\alpha(\beta p(t)) = \alpha(\sum_{j\in\mathbb{N}} \beta a_j t^j) = \sum_{j\in\mathbb{N}} (\alpha\beta a_j)t^j = (\alpha\beta)p(t).$

(6) $\alpha(p(t)+q(t)) = \alpha\sum_{j\in\mathbb{N}} (a_j+b_j)t^j = \sum_{j\in\mathbb{N}} [\alpha(a_j+b_j)]t^j = \sum_{j\in\mathbb{N}} (\alpha a_j+\alpha b_j)t^j$
$= \sum_{j\in\mathbb{N}} (\alpha a_j)t^j + \sum_{j\in\mathbb{N}} (\alpha b_j)t^j = \alpha p(t)+\alpha q(t).$

(7) $(\alpha+\beta)p(t) = \sum_{j\in\mathbb{N}} (\alpha+\beta)a_j t^j = \sum_{j\in\mathbb{N}} (\alpha a_j+\beta a_j)t^j = \sum_{j\in\mathbb{N}} (\alpha a_j)t^j + \sum_{j\in\mathbb{N}} (\beta a_j)t^j = \alpha p(t)+\beta p(t).$

(8) $1p(t) = \sum_{j\in\mathbb{N}} (1a_j)t^j = \sum_{j\in\mathbb{N}} a_j t^j = p(t).$

Proposição 4.19. *São válidas as seguintes propriedades em um espaço vetorial* $\mathbb{V}$*:*

(a) $0\,V = \bar{0}$, *para todo* V *em* $\mathbb{V}$;

(b) $\alpha\,\bar{0} = \bar{0}$, *para todo escalar* α;

(c) *Se* $\alpha\,V = \bar{0}$, *então* $\alpha = 0$ *ou* $V = \bar{0}$;

(d) $(-1)V = -V$, *para todo* V *pertencente a* $\mathbb{V}$.

Demonstração. (a) Usando-se o axioma (2) de espaço vetorial, temos que

$$0\,V + 0\,V = (0+0)V = 0\,V.$$

Somando-se o simétrico de $0\,V$ ao primeiro e ao último membro e usando os axiomas (2) e (4) temos que

$$(-(0\,V) + 0\,V) + 0\,V = -0\,V + 0\,V = \bar{0}.$$

Aplicando-se novamente o axioma (4) no primeiro membro, chegamos a $0\,V = \bar{0}$.

(b) Este item se prova de forma inteiramente análoga ao anterior, mas a partir de $\alpha\,\bar{0} + \alpha\,\bar{0}$.

(c) Se $\alpha \neq 0$, então pelos axiomas (8) e (5) e pelo item (b), temos que

$$V = 1\,V = \left(\frac{1}{\alpha}\alpha\right)V = \frac{1}{\alpha}(\alpha V) = \frac{1}{\alpha}\bar{0} = \bar{0}.$$

(d) Usando-se os axiomas (8) e (7) e o item (a) temos que

$$(-1)V + V = (-1)V + 1\,V = (-1+1)V = 0\,V = \bar{0}$$

Somando-se $-V$ ao primeiro e ao último membro e usando os axiomas (2), (4), (3) temos que

$$(-1)V = \bar{0} + (-V) = -V.$$

■

Definição 4.6. Seja $\mathbb{V}$ um espaço vetorial. Dizemos que um subconjunto $\mathbb{W} \neq \emptyset$, de $\mathbb{V}$ é um **subespaço** de $\mathbb{V}$, se ele também é um espaço vetorial com relação às mesmas operações definidas em $\mathbb{V}$.

Para verificarmos se um subconjunto de um espaço vetorial é um subespaço não é necessária a verificação dos oito axiomas além dos dois que definem a soma e a multiplicação por escalar.

Teorema 4.20. *Seja $\mathbb{V}$ um espaço vetorial. Um subconjunto não vazio, $\mathbb{W} \subseteq \mathbb{V}$, é um subespaço de $\mathbb{V}$ se, e somente se, as operações de soma e multiplicação por escalar estão bem definidas, ou seja, se*

(0) Se $V, W \in \mathbb{W}$, então $V + W \in \mathbb{W}$;

(0') Se $V \in \mathbb{W}$ e α é um escalar, então $\alpha V \in \mathbb{W}$;

Demonstração. Se $\mathbb{W}$ é um subespaço, então obviamente as (0) e (0') são satisfeitas. Suponha, agora, que as condições (0) e (0') são verificadas para $\mathbb{W}$. Como $\mathbb{W}$ é um subconjunto de $\mathbb{V}$, então os Axiomas (1), (2), (5), (6), (7) e (8) da Definição 4.5 na página 272 são satisfeitos para os elementos de $\mathbb{W}$, pois são satisfeitos para todos os elementos de $\mathbb{V}$.

Vamos mostrar que os Axiomas (3) e (4) são também satisfeitos, se (0) e (0′) são verificados. Para qualquer elemento V de $\mathbb{W}$, pela Proposição 4.19, $0V = \bar{0}$ e $-V = (-1)V$, ou seja, o vetor nulo $\bar{0}$ e o simétrico de V são múltiplos escalares de V, que por (0′) pertence a $\mathbb{W}$. ■

Exemplo 4.23. Se $\mathbb{V}$ é um espaço vetorial, então $\mathbb{V}$ é um subespaço dele mesmo. E o subconjunto formado apenas pelo vetor nulo, $\mathbb{W} = \{\bar{0}\}$, é claramente um subespaço de $\mathbb{V}$. Assim, todo espaço vetorial $\mathbb{V} \neq \{\bar{0}\}$ possui pelo menos dois subespaços.

Exemplo 4.24. O conjunto $\mathbb{R}^2$ **não** é um subespaço de $\mathbb{R}^3$, pois $\mathbb{R}^2$ **não** é um subconjunto de $\mathbb{R}^3$.

Exemplo 4.25. Os subconjuntos

$$\mathcal{A} = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}$$

e

$$\mathcal{B} = \{(x, y) \in \mathbb{R}^2 \mid xy \geq 0\}$$

não são subespaços de $\mathbb{R}^2$. Pois, para o primeiro, enquanto $V = (1, 1) \in A$,

$$-V = (-1)V = (-1, -1) \notin A.$$

Enquanto para o segundo, $V = (1, 0), W = (0, -1) \in B$,

$$V + W = (1, -1) \notin B$$

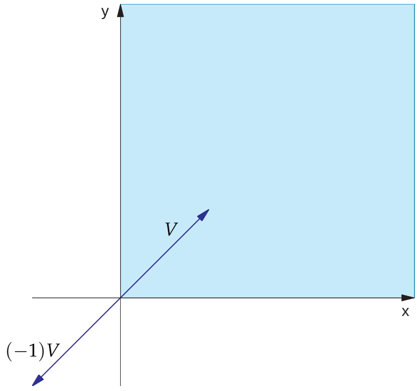


Figura 4.13. $\mathcal{A} = \{(x, y) \in \mathbb{R}^2 | x \geq 0, y \geq 0\}$

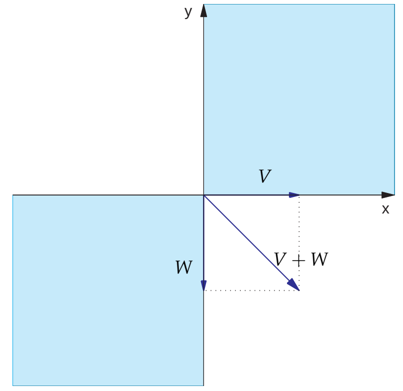


Figura 4.14. $\mathcal{B} = \{(x, y) \in \mathbb{R}^2 \mid xy \geq 0\}$

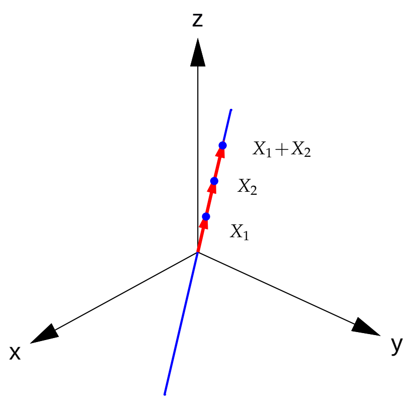


Figura 4.15. Soma de vetores da reta $X = \alpha V$

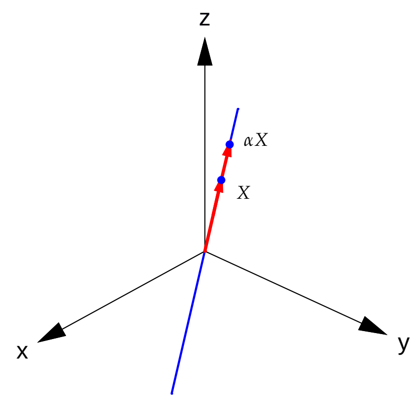


Figura 4.16. Multiplicação de vetor por escalar da reta $X = \alpha V$

Exemplo 4.26. Seja $\mathbb{V} \neq \{\bar{0}\}$ um espaço vetorial. Seja V um vetor não nulo de $\mathbb{V}$. Vamos mostrar que o conjunto dos múltiplos escalares de V,

$$\mathbb{W} = \{\alpha V \mid \alpha \text{ é um escalar}\},$$

é um subespaço de $\mathbb{V}$.

(0) Sejam V_1 e V_2 elementos de $\mathbb{W}$. Então existem escalares α_1 e α_2 tais que $V_1 = \alpha_1 V$ e $V_2 = \alpha_2 V$. Logo

$$V_1 + V_2 = \alpha_1 V + \alpha_2 V = (\alpha_1 + \alpha_2)V.$$

Assim, $V_1 + V_2$ é um múltiplo escalar de V e portanto pertence a $\mathbb{W}$.

(0′) Seja W um elemento de $\mathbb{W}$ e β um escalar. Então existe um escalar α tal que $W = \alpha V$. Logo

$$\beta W = \beta(\alpha V) = (\beta\alpha)V.$$

Assim, βW é um múltiplo escalar de V e portanto pertence a $\mathbb{W}$.

Exemplo 4.27. Seja $N = (a_1, \ldots, a_n)$ um vetor de $\mathbb{R}^n$ fixo. O conjunto definido por

$$\mathbb{W} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1x_1 + \ldots + a_nx_n = 0\}$$

é um subespaço de $\mathbb{R}^n$.

(0) Se $X = (x_1, \ldots, x_n)$ e $Y = (y_1, \ldots, y_n)$ pertencem a $\mathbb{W}$, então

$$a_1x_1 + \ldots + a_nx_n = 0$$

e

$$ay_1 + \ldots + a_ny_n = 0$$

e portanto
$$X + Y = (x_1 + y_1, \ldots, x_n + y_n)$$
também pertence a $\mathbb{W}$, pois
$$a_1(x_1 + y_1) + \ldots + a_n(x_n + y_n) = (a_1x_1 + \ldots + a_nx_n) + (a_1y_1 + \ldots + a_ny_n) = 0 + 0 = 0.$$

(0') Se $X = (x_1, \ldots, x_n)$ pertence a $\mathbb{W}$, então
$$\alpha X = (\alpha x_1, \ldots, \alpha x_n)$$
também pertence a $\mathbb{W}$, pois
$$a_1(\alpha x_1) + \ldots + a_n(\alpha x_n) = \alpha(a_1x_1 + \ldots + a_nx_n) = \alpha 0 = 0\,.$$

Por outro lado, suponha que o conjunto definido por
$$\mathbb{W} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1x_1 + \ldots + a_nx_n = c\}$$
seja um subespaço de $\mathbb{R}^n$, em que c é um número real fixado.

Se $\mathbb{W}$ é um subespaço e $X \in \mathbb{W}$, então $0X = \bar{0}$ também pertence a $\mathbb{W}$, ou seja, o subespaço tem que conter a origem. Substituindo-se $\bar{0} = (0, \ldots, 0)$ na equação que define o conjunto, obtemos que $a_1 0 + \ldots + a_n 0 = c$, ou seja, $c = 0$.

Se $N = (a_1, \ldots, a_n) \neq \bar{0}$, então $\mathbb{W}$ é chamado um **hiperplano de** $\mathbb{R}^n$. Para $n = 3$ os hiperplanos são planos e para $n = 2$ os hiperplanos são retas.

Exemplo 4.28. O conjunto das matrizes simétricas $n \times n$:
$$\mathbb{W}_1 = \{A \in \mathcal{M}_{nn} \mid A^t = A\}$$

e o conjunto das matrizes anti-simétricas $n \times n$:

$$\mathbb{W}_2 = \{A \in \mathcal{M}_{nn} \mid A^t = -A\}$$

são subespaços do espaço $\mathcal{M}_{nn}$ das matrizes $n \times n$, pois a soma de matrizes (anti-)simétricas é uma matriz (anti-)simétrica (verifique!). O mesmo ocorre com a multiplicação por escalar.

Exemplo 4.29. O conjunto $\mathcal{P}_n$ dos polinômios de **grau** (o maior índice j tal que $a_j \neq 0$) menor ou igual a n juntamente com o polinômio nulo é um subespaço do espaço dos polinômios $\mathcal{P}$. Pois, a soma de polinômios de grau menor ou igual a n é um polinômio de grau menor ou igual a n e a multiplicação de um polinômio por escalar é um polinômio de mesmo grau.

Exemplo 4.30. Seja $\mathbb{R}^{(\infty)}$ o conjunto das listas infinitas $(x_1, x_2, \ldots, x_n, \ldots)$ de números reais tais que $x_i \neq 0$ apenas para um número finito de índices i. $\mathbb{R}^{(\infty)}$ é um subespaço de $\mathbb{R}^\infty$, pois a soma de duas listas com um número finito de componentes não nulas é uma lista que também tem somente um número finito de componentes não nulas. O mesmo ocorre com a multiplicação por escalar.

Exemplo 4.31. Seja $\mathcal{X}$ um conjunto não vazio. O conjunto

$$\mathbb{W}_1 = \{f \in \mathcal{F}(\mathcal{X};\mathbb{R}) \mid f(-x) = f(x) \text{ para todo } x \in \mathcal{X}\}$$

das funções, $f : \mathcal{X} \to \mathbb{R}$, pares e o conjunto

$$\mathbb{W}_2 = \{f \in \mathcal{F}(\mathcal{X};\mathbb{R}) \mid f(-x) = -f(x) \text{ para todo } x \in \mathcal{X}\}$$

das funções, $f : \mathcal{X} \to \mathbb{R}$, ímpares são subespaços, pois a soma de funções (ím)pares e a multiplicação de uma função (ím)par por um escalar são também funções (ím)pares (verifique!).

Exemplo 4.32. O conjunto $\mathcal{C}^0(I)$ das funções reais contínuas, que são definidas no intervalo I, é um subespaço do espaço das funções reais $\mathcal{F}(I;\mathbb{R})$. Pois, a soma de funções contínuas é uma função contínua e o mesmo acontece com a multiplicação de uma função contínua por um escalar.

Exemplo 4.33. Seja $\mathcal{C}^n(I)$, para n inteiro positivo, o conjunto das funções reais que possuem a n-ésima derivada contínua no intervalo I. $\mathcal{C}^n(I)$ é um subespaço de $\mathcal{C}^m(I)$, para $0 \leq m \leq n$. E $\mathcal{C}^\infty(I)$, o conjunto das funções que possuem todas as derivadas, é um subespaço de $\mathcal{C}^n(I)$, para todo n inteiro positivo.

Exercícios Numéricos (respostas na página 539)

4.3.1. Determine o vetor X, tal que $3X - 2V = 15(X - U)$, para vetores V e U fixos dados.

4.3.2. Determine o vetor X, tal que $\left\{\begin{array}{ccccc} 6X & - & 2Y & = & U \\ 3X & + & Y & = & U+V \end{array}\right.$, para vetores V e U fixos dados.

4.3.3. Verifique que o polinômio $t^2 + 2t + 7$ é **combinação linear** (soma de múltiplos escalares) de $t^2 + 1$ e $t + 3$.

4.3.4. Verifique que a função constante igual a 3 é combinação linear de $g(t) = 5\tan^2 t$ e $h(t) = \dfrac{2}{\cos^2 t}$.

4.3.5. Quais dos seguintes vetores são combinação linear de $X_1 = (4, 2, -3)$, $X_2 = (2, 1, -2)$ e $X_3 = (-2, -1, 0)$?

(a) $(1, 1, 1)$;
(b) $(4, 2, -6)$;
(c) $(-2, -1, 1)$;
(d) $(-1, 2, 3)$.

4.3.6. Verifique se são espaços vetoriais os seguintes conjuntos:

(a) O $\mathbb{R}^2$ com a adição usual e a multiplicação por escalar definida por $\alpha(x, y) = (\alpha x, 0)$.

(b) O $\mathbb{R}^2$ com $(x_1, y_1) + (x_2, y_2) = (x_1 + 2x_2, y_1 + 2y_2)$ e a multiplicação por escalar usual.

(c) O $\mathbb{R}^2$ com $(x_1, y_1) + (x_2, y_2) = (y_1 + y_2, x_1 + x_2)$ e a multiplicação por escalar usual.

(d) O conjunto dos números reais positivos, com $x + y = xy$ e $\alpha x = x^\alpha$. Qual é o vetor nulo?

Exercícios Teóricos

4.3.7. Sejam $\mathcal{X}$ um conjunto não vazio e $\mathbb{V}$ um espaço vetorial. Mostre que, com as definições naturais de soma e multiplicação por escalar de funções, o conjunto das funções de $\mathcal{X}$ em $\mathbb{V}$, $\mathcal{F}(\mathcal{X}; \mathbb{V})$, é um espaço vetorial.

4.3.8. Mostre que em um espaço vetorial o vetor nulo é único e para cada vetor V o simétrico $-V$ também é único.

4.3.9. Prove que em um espaço vetorial $\mathbb{V}$, $X + W = X + U$ implica que $W = U$.

4.3.10. Em um espaço vetorial, $\alpha X = \beta X$ implica que $\alpha = \beta$? E se $X \neq \bar{0}$?

4.3.11. Mostre que se V pertence a um espaço vetorial $\mathbb{V}$ e n é um inteiro positivo, então $nV = V + \ldots + V$ (n parcelas).

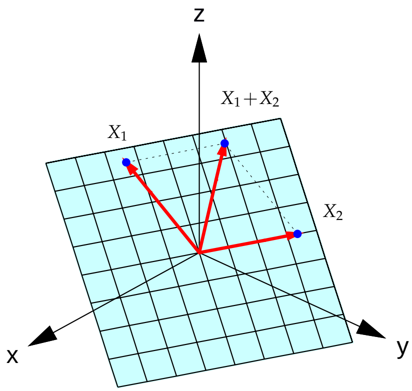


Figura 4.17. Soma de vetores do plano $a_1x + a_2y + a_3z = 0$

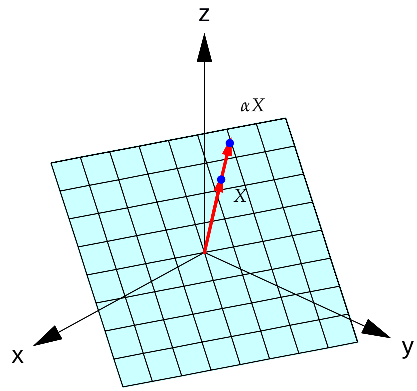


Figura 4.18. Multiplicação de vetor por escalar do plano $a_1x + a_2y + a_3z = 0$

Teste do Capítulo

Considere a matriz

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 4 & 2 \end{bmatrix}$$

1. Encontre os valores de λ tais que o sistema homogêneo $(A - \lambda I_3)X = \bar{0}$ tem solução não trivial.
2. Para os valores de λ encontrados no item anterior, encontre uma base para o espaço solução de
$$(A - \lambda I_3)X = \bar{0}.$$
3. Determine o espaço linha, o espaço coluna e o posto de A.

5

Ortogonalidade

5.1 Produto Escalar em $\mathbb{R}^n$

5.1.1 Produto Interno

Vimos que podemos estender a soma e a multiplicação de vetores por escalar para o $\mathbb{R}^n$. Podemos estender também os conceitos de produto escalar e ortogonalidade.

Definição 5.1. (a) Definimos o **produto escalar ou interno** de dois vetores

$$X = (x_1, \ldots, x_n) \quad \text{e} \quad Y = (y_1, \ldots, y_n) \in \mathbb{R}^n$$

por

$$X \cdot Y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i \,.$$

(b) Definimos a **norma** de um vetor $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ por

$$||X|| = \sqrt{X \cdot X} = \sqrt{x_1^2 + \ldots + x_n^2} = \sqrt{\sum_{i=1}^{n} x_i^2}\,.$$

Escrevendo os vetores como matrizes colunas, o produto interno de dois vetores

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

pode ser escrito em termos do produto de matrizes como

$$X \cdot Y = X^t Y.$$

Exemplo 5.1. Sejam $V = (1, -2, 4, 3, 5)$ e $W = (5, 3, -1, -2, 1)$ vetores do $\mathbb{R}^5$. O produto escalar entre V e W é dado por

$$V \cdot W = (1)(5) + (-2)(3) + (4)(-1) + (3)(-2) + (5)(1) = -6.$$

As normas de V e W são dadas por

$$||V|| = \sqrt{1^2 + (-2)^2 + 4^2 + 3^2 + 5^2} = \sqrt{55},$$

$$||W|| = \sqrt{5^2 + 3^2 + (-1)^2 + (-2)^2 + 1^2} = \sqrt{40}.$$

São válidas as seguintes propriedades para o produto escalar e a norma de vetores do $\mathbb{R}^n$.

Proposição 5.1. *Se X, Y e Z são vetores de $\mathbb{R}^n$ e α é um escalar, então*

(a) $X \cdot Y = Y \cdot X$ *(comutatividade);*

(b) $X \cdot (Y + Z) = X \cdot Y + X \cdot Z$ *(distributividade em relação à soma);*

(c) $(\alpha X) \cdot Y = \alpha(X \cdot Y) = X \cdot (\alpha Y)$;

(d) $X \cdot X = ||X||^2 \geq 0$ *e* $||X|| = 0$ *se, e somente se,* $X = \bar{0}$;

(e) $||\alpha X|| = |\alpha|\, ||X||$;

(f) $|X \cdot Y| \leq ||X||||Y||$ *(desigualdade de Cauchy-Schwarz);*

(g) $||X + Y|| \leq ||X|| + ||Y||$ *(desigualdade triangular).*

Demonstração. Sejam $X, Y, Z \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$. Usando o fato de que se os vetores são escritos como matrizes colunas, então o produto escalar pode ser escrito como o produto de matrizes, $X \cdot Y = X^t Y$, e as propriedades da álgebra matricial (Teorema 1.1 na página 9), temos que

(a) $X \cdot Y = x_1 y_1 + \cdots + x_n y_n = y_1 x_1 + \cdots + y_n x_n = Y \cdot X$.

(b) $X \cdot (Y + Z) = X^t(Y + Z) = X^t Y + X^t Z = X \cdot Y + X \cdot Z$.

(c) $\alpha(X \cdot Y) = \alpha(X^t Y) = (\alpha X^t)Y = (\alpha X)^t Y = (\alpha X) \cdot Y$. A outra igualdade é inteiramente análoga.

(d) $X \cdot X$ é uma soma de quadrados, por isso é sempre maior ou igual a zero e é zero se, e somente se, todas as parcelas são iguais a zero.

(e) $||\alpha X||^2 = (\alpha x_1)^2 + \cdots + (\alpha x_n)^2 = \alpha^2(x_1^2 + \cdots + x_n^2) = \alpha^2||X||^2$. Tomando a raiz quadrada, segue-se o resultado.

(f) A norma de $\lambda X + Y$ é maior ou igual a zero, para qualquer λ real. Assim,

$$0 \leq ||\lambda X + Y||^2 = (\lambda X + Y) \cdot (\lambda X + Y) = (||X||^2)\lambda^2 + (2X \cdot Y)\lambda + ||Y||^2,$$

para qualquer λ real. Logo, o discriminante deste trinômio tem que ser menor ou igual a zero. Ou seja, $\Delta = 4(X \cdot Y)^2 - 4||X||^2||Y||^2 \leq 0$. Logo,

$$|X \cdot Y| \leq ||X||\,||Y||.$$

(g) Pelo item anterior temos que

$$\begin{aligned}
||X + Y||^2 &= (X + Y) \cdot (X + Y) = ||X||^2 + 2X \cdot Y + ||Y||^2 \\
&\leq ||X||^2 + 2|X \cdot Y| + ||Y||^2 \\
&\leq ||X||^2 + 2||X||||Y|| + ||Y||^2 = (||X|| + ||Y||)^2.
\end{aligned}$$

Tomando a raiz quadrada, segue-se o resultado. ■

Dizemos que dois vetores X e Y são **ortogonais** se $X \cdot Y = 0$. As propriedades do produto escalar permitem introduzir o conceito de bases ortogonais no $\mathbb{R}^n$. Antes temos o seguinte resultado.

Proposição 5.2. *Se $V_1, \ldots, V_k$ são vetores **não nulos** de $\mathbb{R}^n$ **ortogonais**, isto é, $V_i \cdot V_j = 0$, para $i \neq j$, então*

(a) *O conjunto $\{V_1, \ldots, V_k\}$ é L.I.*

(b) *Se $V = \sum_{i=1}^{k} \alpha_i V_i$, então $\alpha_i = \dfrac{V \cdot V_i}{||V_i||^2}$.*

Demonstração. (a) Considere a equação vetorial

$$x_1 V_1 + \ldots + x_k V_k = \bar{0}\,. \tag{5.1}$$

Fazendo o produto escalar de ambos os membros de (5.1) com V_i, $i = 1, \ldots, k$ e aplicando as propriedades do produto escalar, obtemos

$$x_1(V_1 \cdot V_i) + \ldots + x_i(V_i \cdot V_i) + \ldots + x_k(V_k \cdot V_i) = 0\,. \tag{5.2}$$

Mas, $V_i \cdot V_j = 0$, se $i \neq j$. Assim, de (5.2) obtemos que

$$x_i ||V_i||^2 = 0\,.$$

Mas, como $V_i \neq \bar{0}$, então $||V_i|| \neq 0$ e $x_i = 0$, para $i = 1 \ldots, k$.

(b) Seja

$$V = \sum_{i=1}^{k} \alpha_i V_i. \tag{5.3}$$

Fazendo o produto escalar de V com V_j, para $j = 1, \ldots, k$, obtemos que

$$V \cdot V_j = \left(\sum_{i=1}^{k} \alpha_i V_i \right) \cdot V_j = \sum_{i=1}^{k} \left(\alpha_i \, V_i \cdot V_j \right) = \alpha_j \, ||V_j||^2.$$

Assim,

$$\alpha_j = \frac{V \cdot V_j}{||V_j||^2}, \quad \text{para } j = 1, \ldots, k.$$

■

Definimos a **projeção ortogonal** de um vetor V sobre um vetor não nulo W por

$$\boxed{\mathrm{proj}_W V = \left(\frac{V \cdot W}{||W||^2} \right) W.}$$

Observe que a projeção ortogonal de um vetor V sobre um vetor não nulo W é um múltiplo escalar do vetor W. Além disso temos o seguinte resultado.

Proposição 5.3. *Seja $W \in \mathbb{R}^n$ um vetor não nulo. Então, $V - \mathrm{proj}_W V$ é ortogonal a W, para qualquer vetor $V \in \mathbb{R}^n$.*

Demonstração. Precisamos calcular o produto escalar de W com $V - \mathrm{proj}_W V$:

$$(V - \mathrm{proj}_W V) \cdot W = V \cdot W - \left(\frac{V \cdot W}{||W||^2} \right) W \cdot W = 0.$$

Portanto, $V - \mathrm{proj}_W V$ é ortogonal a W. ■

O próximo resultado é uma generalização da Proposição 5.3.

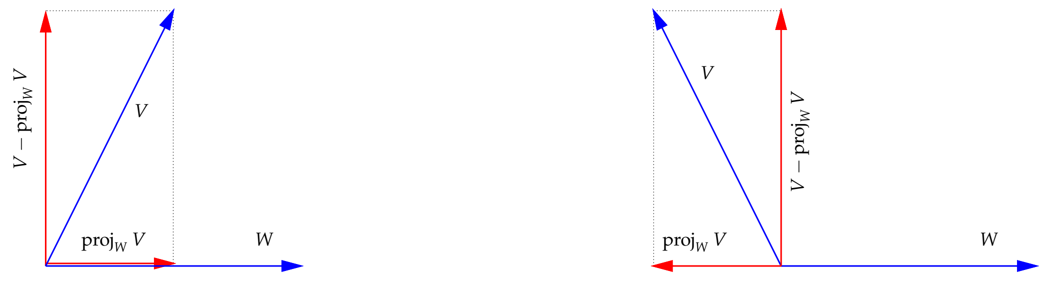


Figura 5.1. Projeção ortogonal do vetor V sobre o vetor W

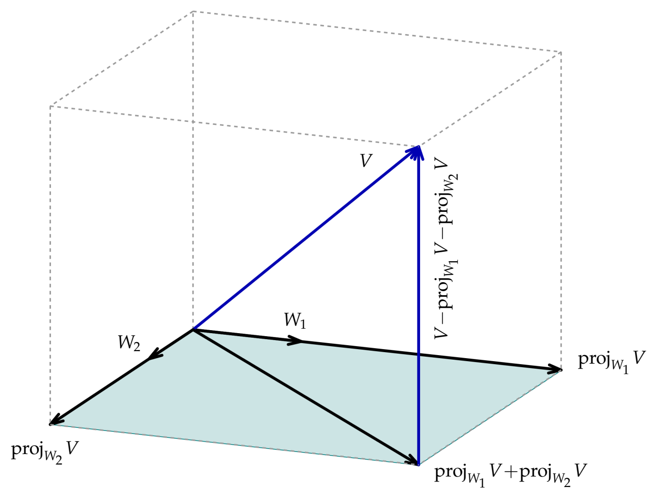


Figura 5.2. $V - \mathrm{proj}_{W_1} V - \mathrm{proj}_{W_2} V$ é ortogonal a W_1 e a W_2

Proposição 5.4. *Sejam $W_1, W_2, \ldots, W_k$ vetores não nulos do $\mathbb{R}^n$, ortogonais entre si, então para qualquer vetor V, $V - \mathrm{proj}_{W_1} V - \ldots - \mathrm{proj}_{W_k} V$ é ortogonal a W_i, para $i = 1, \ldots, k$.*

Demonstração. Vamos calcular o produto interno de $V - \mathrm{proj}_{W_1} V - \ldots - \mathrm{proj}_{W_k} V$ com W_j, para $j = 1, \ldots, k$.

$$\left(V - \sum_{i=1}^{k} \mathrm{proj}_{W_i} V \right) \cdot W_j = V \cdot W_j - \sum_{i=1}^{k} \left(\frac{V \cdot W_i}{||W_i||^2} \right) W_i \cdot W_j = V \cdot W_j - \left(\frac{V \cdot W_j}{||W_j||^2} \right) W_j \cdot W_j = 0,$$

pois $W_i \cdot W_j = 0$, se $i \neq j$ e $W_j \cdot W_j = ||W_j||^2$. ■

Exemplo 5.2. Considere o sistema linear homogêneo $AX = \bar{0}$, em que

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

Escalonando a matriz aumentada do sistema acima, obtemos a matriz escalonada reduzida

$$\left[\begin{array}{ccccc:c} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

E assim a solução geral do sistema pode ser escrita como

$$x_1 = -\alpha - \gamma,\ x_2 = \gamma,\ x_3 = -\alpha + \beta,\ x_4 = \beta\ x_5 = \alpha$$

para todos os valores de $\alpha, \beta, \gamma \in \mathbb{R}$, ou seja, o conjunto solução do sistema $AX = \bar{0}$ é

$$\mathbb{W} = \{(x_1, x_2, x_3, x_4, x_5) = (-\alpha - \gamma, \gamma, -\alpha + \beta, \beta, \alpha) \mid \alpha, \beta, \gamma \in \mathbb{R}\}.$$

Agora, um elemento qualquer de $\mathbb{W}$ pode ser escrito como uma combinação linear de vetores de $\mathbb{W}$:

$$\begin{aligned}(-\alpha - \gamma, \gamma, -\alpha + \beta, \beta, \alpha) &= (-\alpha, 0, -\alpha, 0, \alpha) + (0, 0, \beta, \beta, 0) + (-\gamma, \gamma, 0, 0, 0) \\ &= \alpha(-1, 0, -1, 0, 1) + \beta(0, 0, 1, 1, 0) + \gamma(-1, 1, 0, 0, 0)\end{aligned}$$

Assim, todo vetor de $\mathbb{W}$ pode ser escrito como combinação linear dos vetores $V_1 = (-1, 0, -1, 0, 1)$, $V_2 = (0, 0, 1, 1, 0)$ e $V_3 = (-1, 1, 0, 0, 0)$ pertencentes a $\mathbb{W}$ (V_1 é obtido fazendo-se $\alpha = 1$ e $\beta = \gamma = 0$, V_2 fazendo-se $\alpha = \gamma = 0$ e $\beta = 1$ e V_3 fazendo-se $\alpha = \beta = 0$ e $\gamma = 1$). Além disso segue da equação anterior que V_1, V_2 e V_3 são L.I. Logo $\{V_1, V_2, V_3\}$ é uma base de $\mathbb{W}$.

Vamos, agora, encontrar uma base ortonormal para $\mathbb{W}$. Para isso vamos aplicar a Proposição 5.3 na página 294.

$$\begin{aligned}W_1 &= V_1 = (-1, 0, -1, 0, 1); \\ W_2 &= V_2 - \operatorname{proj}_{W_1} V_2 = (0, 0, 1, 1, 0) + \frac{1}{3}(-1, 0, -1, 0, 1) = \frac{1}{3}(-1, 0, 2, 3, 1) \\ W_3 &= V_3 - \operatorname{proj}_{W_1} V_3 - \operatorname{proj}_{W_2} V_3 = (-1, 1, 0, 0, 0) - \frac{1}{3}(-1, 0, -1, 0, 1) - \frac{1}{15}(-1, 0, 2, 3, 1) \\ &= \frac{1}{5}(-3, 5, 1, -1, -2)\end{aligned}$$

Agora, vamos "dividir" cada vetor pela sua norma para obtermos vetores de norma igual a 1 (**unitários**).

$$
\begin{aligned}
U_1 &= \left(\frac{1}{||W_1||}\right) W_1 = (-\frac{1}{\sqrt{3}}, 0, -\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}) \\
U_2 &= \left(\frac{1}{||W_2||}\right) W_2 = (-\frac{1}{\sqrt{15}}, 0, \frac{2}{\sqrt{15}}, \frac{3}{\sqrt{15}}, \frac{1}{\sqrt{15}}) \\
U_3 &= \left(\frac{1}{||W_3||}\right) W_3 = (-\frac{3}{2\sqrt{10}}, \frac{5}{2\sqrt{10}}, \frac{1}{2\sqrt{10}}, -\frac{1}{2\sqrt{10}}, -\frac{1}{\sqrt{10}})
\end{aligned}
$$

5.1.2 Bases Ortogonais e Ortonormais

Definição 5.2. Seja $\{V_1, \ldots, V_k\}$ uma base de um subespaço de $\mathbb{R}^n$.

(a) Dizemos que $\{V_1, \ldots, V_k\}$ é uma **base ortogonal**, se $V_i \cdot V_j = 0$, para $i \neq j$, ou seja, se quaisquer dois vetores da base são ortogonais;

(b) Dizemos que $\{V_1, \ldots, V_k\}$ é uma **base ortonormal**, se além de ser uma base ortogonal, $||V_i|| = 1$, ou seja, o vetor V_i é **unitário**, para $i = 1, \ldots m$.

Exemplo 5.3. A **base canônica de** $\mathbb{R}^n$, que é formada pelos vetores

$$E_1 = (1, 0, \ldots, 0), \quad E_2 = (0, 1, 0, \ldots, 0), \quad \ldots \quad E_n = (0, \ldots, 0, 1)$$

é uma base ortonormal de $\mathbb{R}^n$.

Exemplo 5.4. No Exemplo 5.2, $\{W_1, W_2, W_3\}$ é uma base ortogonal de $\mathbb{W}$ e $\{U_1, U_2, U_3\}$ é uma base ortonormal de $\mathbb{W}$.

O resultado a seguir mostra que o procedimento usado no Exemplo 5.2 conhecido como **processo de ortogonalização de Gram-Schmidt** pode ser aplicado a qualquer subespaço de $\mathbb{R}^n$. Nas Figuras 5.3 e 5.4 vemos como isto é possível no caso em que o subespaço é o $\mathbb{R}^3$, já que o $\mathbb{R}^3$ é subespaço dele mesmo.

Teorema 5.5. *Seja $\{V_1, \ldots, V_k\}$ uma base de um subespaço $\mathbb{W}$ de $\mathbb{R}^n$. Então, existe uma base $\{U_1, \ldots, U_k\}$ de $\mathbb{W}$ que é ortonormal e tal que o subespaço gerado por $U_1, \ldots, U_j$ é igual ao subespaço gerado por $V_1, \ldots, V_j$ para $j = 1, \ldots, k$.*

Demonstração. (a) Sejam

$$\begin{array}{rcl}
W_1 & = & V_1\,, \\
W_2 & = & V_2 - \mathrm{proj}_{W_1} V_2\,, \\
W_3 & = & V_3 - \mathrm{proj}_{W_1} V_3 - \mathrm{proj}_{W_2} V_3\,, \\
& \ldots & \\
W_k & = & V_k - \mathrm{proj}_{W_1} V_k - \mathrm{proj}_{W_2} V_k \ldots - \mathrm{proj}_{W_{k-1}} V_k.
\end{array}$$

Pela Proposição 5.3, segue-se que W_2 é ortogonal a W_1 e $W_2 \neq \bar{0}$, pois V_1 e V_2 são L.I. Assim, W_1 e W_2 formam uma base ortogonal do subespaço gerado por V_1 e V_2. Agora, supondo que $W_1, \ldots, W_{k-1}$ seja uma base ortogonal do subespaço gerado por $V_1, \ldots, V_{k-1}$, segue-se da Proposição 5.4, que W_k é ortogonal a $W_1, \ldots, W_{k-1}$. $W_k \neq \bar{0}$, pois caso contrário, V_k pertenceria ao subespaço gerado por $W_1, \ldots, W_{k-1}$ que é igual ao subespaço gerado por $V_1, \ldots, V_{k-1}$ e assim $V_1, \ldots, V_k$ seriam L.D. Como $W_1, \ldots, W_k$ são ortogonais não nulos, pela Proposição 5.2 na página 293, eles são L.I. e portanto formam uma base do subespaço $\mathbb{W}$.

(b) Sejam, agora

$$U_1 = \left(\frac{1}{||W_1||}\right) W_1, \quad U_2 = \left(\frac{1}{||W_2||}\right) W_2, \quad \ldots, \quad U_k = \left(\frac{1}{||W_k||}\right) W_k\,.$$

Assim, $\{U_1, \ldots, U_k\}$ é uma base ortonormal para o subespaço $\mathbb{W}$.

■

Exercícios Numéricos (respostas na página 542)

5.1.1. Sejam $X = (1, 1, -2)$ e $Y = (a, -1, 2)$. Para quais valores de a, X e Y são ortogonais?

5.1.2. Sejam $X = (1/\sqrt{2}, 0, 1/\sqrt{2})$ e $Y = (a, 1/\sqrt{2}, -b)$. Para quais valores de a e b, o conjunto $\{X, Y\}$ forma uma base ortonormal do subespaço gerado por eles?

5.1.3. Use o processo de ortogonalização de Gram-Schmidt para encontrar uma base ortonormal para o subespaço de $\mathbb{R}^4$ que tem como base $\{(1, 1, -1, 0), (0, 2, 0, 1), (-1, 0, 0, 1)\}$.

5.1.4. Use o processo de ortogonalização de Gram-Schmidt para transformar a base do $\mathbb{R}^3$ $\{(1, 1, 1), (0, 1, 1), (1, 2, 3)\}$ em uma base ortonormal do $\mathbb{R}^3$.

5.1.5. Encontre uma base ortonormal para o subespaço de $\mathbb{R}^3$ que consiste de todos os vetores (a, b, c) tais que $a + b + c = 0$.

5.1.6. Encontre uma base ortonormal para o subespaço do $\mathbb{R}^4$ que consiste de todos os vetores (a, b, c, d) tais que $a - b - 2c + d = 0$.

5.1.7. Encontre uma base ortonormal para o espaço solução do sistema homogêneo

$$\left\{ \begin{array}{ccccccc} x & + & y & - & z & = & 0 \\ 2x & + & y & + & 2z & = & 0. \end{array} \right.$$

5.1.8. Considere as retas $(x, y, z) = t(1, 2, -3)$ e $(x, y, z) = (0, 1, 2) + s(2, 4, -6)$ em $\mathbb{R}^3$. Encontre a equação geral do plano que contém estas duas retas e ache uma base ortonormal para este plano. Complete esta base a uma base ortonormal de $\mathbb{R}^3$.

Exercícios Teóricos

5.1.9. Mostre que para todo vetor $V \in \mathbb{R}^n$ e todo escalar α, $||\alpha V|| = |\alpha| \, ||V||$.

5.1.10. Mostre que se V é ortogonal a W, então V é ortogonal a αW, para todo escalar α.

5.1.11. Mostre que se V é ortogonal a $W_1, \ldots, W_k$, então V é ortogonal a qualquer combinação linear de $W_1, \ldots, W_k$.

5.1.12. Sejam X, Y e Z vetores do $\mathbb{R}^n$. Prove que se $X \cdot Y = X \cdot Z$, então $Y - Z$ é ortogonal a X.

5.1.13. Mostre que se $W_1, \ldots, W_k$ são vetores não nulos ortogonais entre si e $X = \alpha_1 W_1 + \ldots + \alpha_k W_k$, então $X = \text{proj}_{W_1} X + \ldots + \text{proj}_{W_k} X$.

5.1.14. Sejam $V_1, \ldots, V_k$ vetores linearmente dependentes. Mostre que, aplicando-se o processo de ortogonalização de Gram-Schmidt aos vetores $V_1, \ldots, V_k$, se obtém um vetor W_i que é nulo, para algum $i = 1, \ldots, k$. (Sugestão: Seja V_i o primeiro vetor tal que $V_i \in [V_1, \ldots, V_{i-1}] = [W_1, \ldots, W_{i-1}]$ e use o exercício anterior.)

5.1.15. Seja $S = \{W_1, \ldots, W_k\}$ uma base ortogonal de um subespaço $\mathbb{W}$ de $\mathbb{R}^n$. Mostre que um todo vetor V de $\mathbb{W}$ pode ser escrito como

$$V = \frac{V \cdot W_1}{||W_1||^2} W_1 + \frac{V \cdot W_2}{||W_2||^2} W_2 + \ldots + \frac{V \cdot W_k}{||W_k||^2} W_k.$$

(Sugestão: escreva $V = x_1 W_1 + \ldots + x_k W_k$, faça o produto escalar de V com W_i e conclua que $x_i = \frac{V \cdot W_i}{||W_i||^2}$, para $i = 1, \ldots, k$.)

5.1.16. Mostre que o conjunto de todos os vetores do $\mathbb{R}^n$ ortogonais a um dado vetor $V = (a_1, \ldots, a_n)$,

$$\mathbb{W} = \{X = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid X \cdot V = 0\} \quad \text{é um subespaço do } \mathbb{R}^n.$$

5.1.17. Demonstre que, se V e W são vetores quaisquer do $\mathbb{R}^n$, então:

(a) $V \cdot W = \frac{1}{4}[||V + W||^2 - ||V - W||^2]$ (identidade polar);

(b) $||V||^2 + ||W||^2 = \frac{1}{2}(||V + W||^2 + ||V - W||^2)$ (lei do paralelogramo).

(Sugestão: desenvolva os segundos membros das igualdades acima observando que $||V + W||^2 = (V + W) \cdot (V + W)$ e $||V - W||^2 = (V - W) \cdot (V - W)$)

5.1.18. Seja $\{U_1, \ldots, U_n\}$ uma base ortonormal de $\mathbb{R}^n$. Se $A = [\, U_1 \ \ldots \ U_n \,]$ é uma matriz $n \times n$ cujas colunas são os vetores $U_1, \ldots, U_n$, então A é invertível e $A^{-1} = A^t$. (Sugestão: mostre que $A^t A = I_n$, usando o fato de que $U_i \cdot U_j = U_i^t U_j$.)

5.1.19. Mostre que o ângulo entre dois vetores não nulos $X = (x_1, \ldots, x_n)$ e $Y = (y_1, \ldots, y_n)$ de $\mathbb{R}^n$, que é definido como sendo o número real θ entre 0 e π tal que

$$\cos\theta = \frac{X \cdot Y}{||X||\,||Y||},$$

está bem definido, ou seja, que existe um tal número real θ e é único. (Sugestão: mostre, usando a desigualdade de Cauchy-Schwarz, que
$-1 \leq \dfrac{X \cdot Y}{||X||\,||Y||} \leq 1$.)

5.1.20. Seja $\mathbb{W}$ um subespaço de $\mathbb{R}^n$. Mostre que o conjunto de todos os vetores ortogonais a todos os vetores de $\mathbb{W}$ é um subespaço de $\mathbb{R}^n$. Este subespaço é chamado de **complemento ortogonal de** $\mathbb{W}$ e denotado por $\mathbb{W}^\perp$, ou seja,

$$\mathbb{W}^\perp = \{X \in \mathbb{R}^n \mid X \cdot Y = 0, \text{para todo } Y \in \mathbb{W}\}.$$

5.1.21. Mostre que todo subespaço $\mathbb{W}$ de $\mathbb{R}^n$ é o espaço solução de um sistema linear homogêneo. (Sugestão: seja $\{W_1, \ldots, W_k\}$ uma base de $\mathbb{W}^\perp$ tome $A = [\, W_1 \ldots W_k \,]^t$.)

5.1.22. Embora não exista o produto vetorial de dois vetores em $\mathbb{R}^n$, para $n > 3$, podemos definir o **produto vetorial de** $n-1$ **vetores**, $V_1 = (v_{11}, \ldots, v_{1n}), \ldots, V_{n-1} = (v_{(n-1)1}, \ldots, v_{(n-1)n})$ como

$$V_1 \times V_2 \times \cdots \times V_{n-1} = \left((-1)^{n+1}\det(v_{ij})_{j\neq 1}, (-1)^{n+2}\det(v_{ij})_{j\neq 2}, \ldots, (-1)^{2n}\det(v_{ij})_{j\neq n}\right).$$

Mostre que:

(a) $V_1 \times V_2 \times \cdots \times V_{n-1}$ é ortogonal a $V_1, \ldots, V_{n-1}$.

(b) $\alpha(V_1 \times V_2 \times \cdots \times V_{n-1}) = V_1 \times \cdots \alpha V_i \times \cdots \times V_{n-1}$

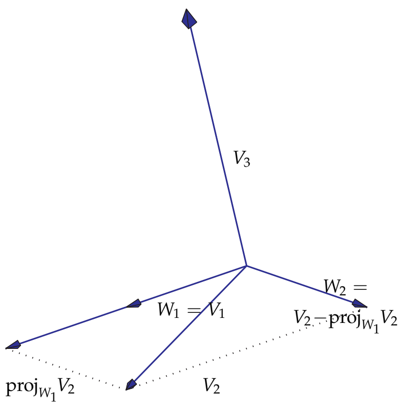


Figura 5.3. $W_1 = V_1$ e $W_2 = V_2 - \mathrm{proj}_{W_1} V_2$

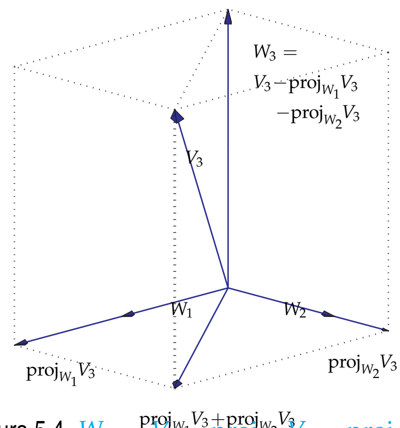


Figura 5.4. $W_3 = V_3 - \mathrm{proj}_{W_1} V_3 - \mathrm{proj}_{W_2} V_3$

5.2 Subespaços Ortogonais

Se $N = (a, b, c)$ é um vetor não nulo de $\mathbb{R}^3$, o conjunto dos vetores que são ortogonais a N, é um plano que passa pela origem e tem N como vetor normal. Neste caso dizemos que o plano é o subespaço ortogonal ao conjunto $\{N\}$.

Definição 5.3. Seja $\mathcal{S}$ um subconjunto não vazio de $\mathbb{R}^n$. O **complemento ortogonal de** $\mathcal{S}$, denotado por $\mathcal{S}^\perp$, é o conjunto de todos os vetores de $\mathbb{R}^n$ que são ortogonais a todo vetor de $\mathcal{S}$. Ou seja,

$$\mathcal{S}^\perp = \{X \in \mathbb{R}^n \mid X \cdot Y = 0 \text{ para todo } Y \in S\}.$$

Mesmo quando $\mathcal{S}$ não é um subespaço, $\mathcal{S}^\perp$ é um subespaço.

Proposição 5.6. *Seja $\mathcal{S}$ um subconjunto de $\mathbb{R}^n$. Então, o conjunto $\mathcal{S}^\perp$ é um subespaço.*

Demonstração. Vamos verificar as propriedades (a) e (b) na página 227 que definem um subespaço.

(a) Sejam X_1 e X_2 vetores de $\mathcal{S}^\perp$. Então,

$$(X_1 + X_2) \cdot Y = X_1 \cdot Y + X_2 \cdot Y = 0 + 0 = 0, \quad \text{para todo } Y \in S.$$

(b) Seja $X \in S^\perp$ e α um escalar. Então,

$$(\alpha X) \cdot Y = \alpha(X \cdot Y) = \alpha 0 = 0, \quad \text{para todo } Y \in S.$$

■

Exemplo 5.5. Se $\mathcal{S} = \{\bar{0}\} \subset \mathbb{R}^n$, então $\mathcal{S}^\perp = \mathbb{R}^n$. Se $\mathcal{S} = \mathbb{R}^n$, então $\mathcal{S}^\perp = \{\bar{0}\}$.

Exemplo 5.6. Seja $\mathcal{S} = \{N = (a_1, \ldots, a_n)\} \subset \mathbb{R}^n$. Então,

$$\mathcal{S}^\perp = \{X = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \ldots + a_n x_n = 0\}.$$

Exemplo 5.7. Seja $\mathcal{S} = \{V_1, \ldots, V_m\} \subset \mathbb{R}^n$, em que

$$V_1 = (a_{11}, \ldots, a_{1n}), \ldots, V_m = (a_{m1}, \ldots, a_{mn}).$$

Então,

$$\mathcal{S}^\perp = \{X \in \mathbb{R}^n \mid A\,X = \bar{0}\}, \quad \text{em que } A = (a_{ij})_{m \times n}.$$

Se $\mathcal{S} = \mathbb{W}$ é um subespaço, além de $\mathbb{W}^\perp$ ser um subespaço, são válidas as propriedades a seguir.

Proposição 5.7. *Sejam $\mathbb{W}$ um subespaço de $\mathbb{R}^n$ e $\mathbb{W}^\perp$ o seu complemento ortogonal. Então:*

(a) *Todo vetor de $V \in \mathbb{R}^n$ se decompõe de maneira única como uma soma de dois vetores V_1 e V_2, sendo V_1 pertencente a $\mathbb{W}$ e V_2 pertencente a $\mathbb{W}^\perp$, ou seja, para todo $V \in \mathbb{R}^n$ existe um único $V_1 \in \mathbb{W}$ e um único $V_2 \in \mathbb{W}^\perp$ tal que*

$$V = V_1 + V_2.$$

(b) *O subespaço ortogonal de $\mathbb{W}^\perp$ é $\mathbb{W}$, ou seja,*

$$(\mathbb{W}^\perp)^\perp = \mathbb{W}.$$

Demonstração. (a) Seja V um vetor qualquer de $\mathbb{V}$. Seja $W_1, \ldots, W_m$ uma base ortogonal de $\mathbb{W}$.

Defina $V_1 = \text{proj}_{W_1} V + \ldots + \text{proj}_{W_m} V$. Pela Proposição 5.4 na página 297, o vetor $V_2 = V - V_1$ é ortogonal a W_k, para $k = 1, \ldots, m$. Logo, V_2 é ortogonal a todo vetor de $\mathbb{W}$ e portanto $V_2 \in \mathbb{W}^\perp$. Assim, $V = V_1 + V_2$, com $V_1 \in \mathbb{W}$ e $V_2 \in \mathbb{W}^\perp$.

Sejam $X \in \mathbb{W}$ e $Y \in \mathbb{W}^\perp$ tais que $V = X + Y$. Então,

$$\begin{aligned} V_1 &= \text{proj}_{W_1} V + \ldots + \text{proj}_{W_m} V = \text{proj}_{W_1}(X+Y) + \ldots + \text{proj}_{W_m}(X+Y) \\ &= \text{proj}_{W_1} X + \ldots + \text{proj}_{W_m} X = X, \end{aligned}$$

pois como $Y \in \mathbb{W}^\perp$, $\text{proj}_{W_k}(X+Y) = \text{proj}_{W_k} X$, para $k = 1, \ldots, m$ e como $X \in \mathbb{W}$, pela Proposição 5.2 na página 293, $X = \text{proj}_{W_1} X + \ldots + \text{proj}_{W_m} X$.

E

$$Y = V - X = V - V_1 = V_2.$$

(b) Todo elemento de $\mathbb{W}$ claramente pertence a $(\mathbb{W}^\perp)^\perp$. Assim, $\mathbb{W} \subseteq (\mathbb{W}^\perp)^\perp$. Falta mostrar que $(\mathbb{W}^\perp)^\perp \subseteq \mathbb{W}$. Seja $X \in (\mathbb{W}^\perp)^\perp$. Então, $X = U + V$, em que $U \in \mathbb{W}$

e $V \in \mathbb{W}^\perp$. Assim,

$$0 = X \cdot V = (U + V) \cdot V = U \cdot V + V \cdot V = V \cdot V = ||V||^2.$$

Consequentemente, $V = \bar{0}$. Assim, $X \in \mathbb{W}$ e $(\mathbb{W}^\perp)^\perp \subseteq \mathbb{W}$. Portanto,

$$(\mathbb{W}^\perp)^\perp = \mathbb{W}.$$

.

■

Seja $\mathbb{W}$ um subespaço do $\mathbb{R}^n$. Dado um vetor $V \in \mathbb{R}^n$, em virtude da Proposição 5.7 existe uma única decomposição $V = V_1 + V_2$, com $V_1 \in \mathbb{W}$ e $V_2 \in \mathbb{W}^\perp$. O vetor V_1 é chamado **projeção ortogonal de V no subespaço** $\mathbb{W}$ e é denotado por $\text{proj}_{\mathbb{W}} V$.

Se $\{W_1, \ldots, W_m\}$ é uma base ortogonal de $\mathbb{W}$, então decorre da demonstração da Proposição 5.7 que

$$\boxed{\text{proj}_{\mathbb{W}} V = \text{proj}_{W_1} V + \ldots + \text{proj}_{W_m} V}.$$

Exemplo 5.8. Seja $\mathbb{W}$ o subespaço gerado pelos vetores

$$V_1 = (1, 0, 0, 0), V_2 = (1, 1, 0, 0), V_3 = (-4, -4, 1, 1).$$

Seja $X = (-2, -4, 0, 2)$. Vamos encontrar $Y \in \mathbb{W}$ e $Z \in \mathbb{W}^\perp$ tais que $X = Y + Z$. Basta tomarmos $Y = \text{proj}_{\mathbb{W}} X$ e $Z = X - Y$.

Para encontrar esta decomposição vamos aplicar o processo de ortogonalização de Gram-Schmidt aos vetores V_1, V_2 e V_3 obtendo W_1, W_2 e W_3 uma base ortogonal de $\mathbb{W}$.

$$W_1 = V_1, \quad W_2 = V_2 - \text{proj}_{W_1} V_2 = (0, 1, 0, 0),$$

$$W_3 = V_3 - \mathrm{proj}_{W_1} V_3 - \mathrm{proj}_{W_2} V_3 = (0,0,1,1).$$

Assim,

$$Y = \mathrm{proj}_{\mathbb{W}} X = \mathrm{proj}_{W_1} X + \mathrm{proj}_{W_2} X + \mathrm{proj}_{W_3} X = (-2,-4,1,1) \quad \text{e} \quad Z = X - Y = (0,0,-1,1)$$

5.2.1 Subespaços Fundamentais

Lembramos que a imagem de uma matriz A, $m \times n$, é o subespaço definido por

$$\mathcal{I}(A) = \{Y \in \mathbb{R}^m \mid A\,X = Y \text{para algum } X \in \mathbb{R}^n\},$$

que é igual ao espaço coluna de A (Proposição 4.17 na página 265). Lembramos também que o núcleo de A é definido por

$$\mathcal{N}(A) = \{X \in \mathbb{R}^n \mid A\,X = \bar{0}\,\}.$$

Já vimos que $\dim(\mathcal{N}(A)) + \dim(\mathcal{I}(A)) = n$ (Teorema 4.18 na página 267). Observe que enquanto $\mathcal{N}(A)$ é um subespaço do $\mathbb{R}^n$, $\mathcal{I}(A)$ é um subespaço do $\mathbb{R}^m$. Mas, $\mathcal{I}(A^t)$ é também um subespaço de $\mathbb{R}^n$ e $\mathcal{N}(A^t)$ é um subespaço do $\mathbb{R}^m$. Assim, $\mathcal{N}(A)$ e $\mathcal{I}(A^t)$ são subespaços do $\mathbb{R}^n$ e $\mathcal{N}(A^t)$ e $\mathcal{I}(A)$ são subespaços do $\mathbb{R}^m$ e são válidas as seguintes relações de ortogonalidade.

Teorema 5.8. *Se A é uma matriz $m \times n$, então*

(a) $\mathcal{N}(A) = \mathcal{I}(A^t)^\perp$.

(b) $\mathcal{N}(A^t) = \mathcal{I}(A)^\perp$.

Demonstração. (a) Um vetor V pertence ao núcleo de A se, e somente se, $AV = \bar{0}$. Isto acontece se, e somente se, AV é ortogonal a todo vetor $Y \in \mathbb{R}^m$, ou seja, se, e somente se, $Y^t AV = 0$ para todo vetor $Y \in \mathbb{R}^m$. Esta equação é equivalente a sua transposta, ou seja, $V^t A^t Y = 0$. Variando-se Y em $\mathbb{R}^m$, $A^t Y$ percorre toda a imagem de A^t. Assim, $V^t A^t Y = 0$, para todo $Y \in \mathbb{R}^m$ se, e somente se, $V \in \mathcal{I}(A^t)^\perp$. Portanto, $V \in \mathcal{N}(A)$ se, e somente se, $V \in \mathcal{I}(A^t)^\perp$.

(b) Basta aplicar o item anterior a A^t.

■

5.2.2 Problema de Quadrados Mínimos

Muitos problemas, quando modelados, levam a sistemas lineares $A\,X = B$, que são inconsistentes (isto é, não possuem solução), apesar dos problemas que os originaram requererem solução. A inconsistência vem com frequência devido a erros experimentais na matriz B. Uma forma de resolver esta inconsistência é resolver o **problema de quadrados mínimos** associado, ou seja,

$$\min ||A\,X - B||^2.$$

Apesar de não ser esta a única forma de resolver a inconsistência, pode-se mostrar que se os erros em B forem não viciados e os b_i tiverem a mesma variância (fixa), então a solução do problema de quadrados mínimos é a que tem a menor variância dentro de um certo conjunto de "soluções".

O teorema seguinte é a chave para a solução do problema de quadrados mínimos.

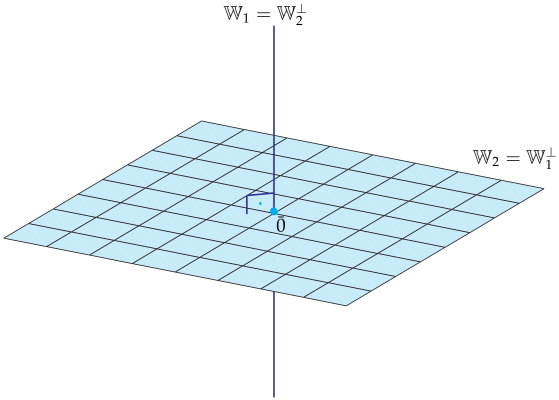


Figura 5.5. Complementos ortogonais

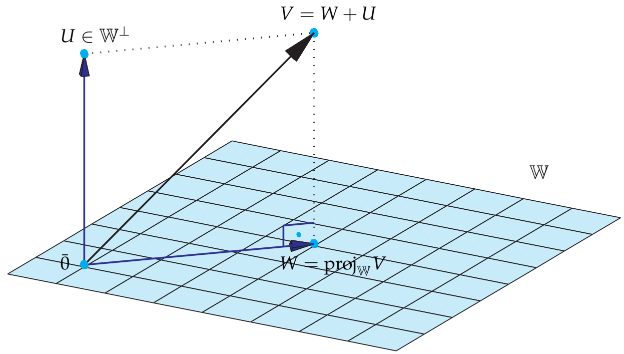


Figura 5.6. Decomposição de um ponto $X = Y + Z$, com $Y \in \mathbb{W}, Z \in \mathbb{W}^\perp$

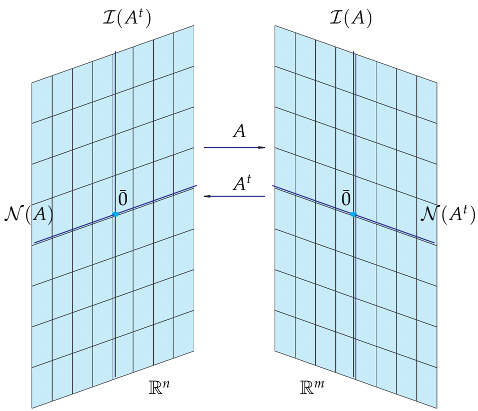


Figura 5.7. Subespaços $\mathcal{N}(A), \mathcal{I}(A^t)$, $\mathcal{I}(A)$ e $\mathcal{N}(A^t)$

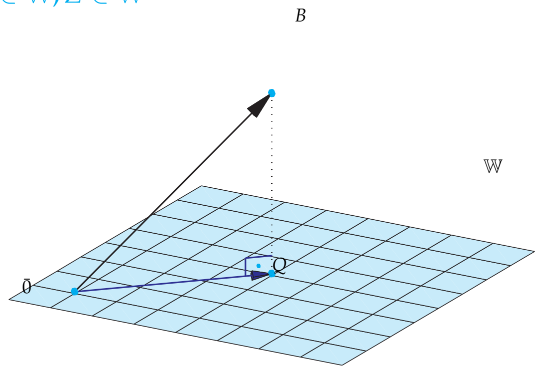


Figura 5.8. Ponto em um subespaço mais próximo do ponto B

Teorema 5.9. *Seja A uma matriz $m \times n$. O problema de quadrados mínimos:*

$$\min ||AX - B||^2$$

é equivalente a resolver o sistema linear consistente

$$A^t AX = A^t B,$$

chamado de ***sistema de equações normais****.*

Demonstração. O problema de quadrados mínimos

$$\min ||AX - B||^2$$

pode ser escrito como

$$\min_{Y \in \mathcal{I}(A)} ||Y - B||^2 \quad \text{e} \quad Y = A\,X. \tag{5.4}$$

Seja $\mathbb{W} = \mathcal{I}(A)$. Segue da Proposição 5.7 na página 307 que existe uma única decomposição de B como

$$B = Q + Z,$$

em que $Q \in \mathbb{W}$ e $Z \in \mathbb{W}^\perp$. Vamos mostrar que

$$\min_{Y \in \mathbb{W}} ||B - Y|| = ||B - Q||.$$

Seja Y um vetor qualquer de $\mathbb{W}$. Temos que

$$||Y - B||^2 = ||(Y - Q) + (Q - B)||^2 = ||Y - Q||^2 + 2(Y - Q) \cdot (Q - B) + ||B - Q||^2.$$

Mas,

$$(Y - Q) \cdot (Q - B) = Q \cdot (B - Q) - Y \cdot (B - Q) = 0,$$

pois $Y, Q \in \mathbb{W}$ e $B - Q = Z \in \mathbb{W}^\perp$. Logo,

$$||Y - B||^2 = ||(Y - Q) + (Q - B)||^2 = ||Y - Q||^2 + ||B - Q||^2. \tag{5.5}$$

Variando Y em $\mathbb{W}$, vemos de (5.5) que o mínimo de $||Y - B||$ ocorre somente para $Y = Q$, já que $||B - Q||^2$ permanece fixo em (5.5) quando variamos Y em $\mathbb{W}$. Portanto,

$$\min_{Y \in \mathbb{W}} ||B - Y|| = ||B - Q||,$$

em que o ponto Q é tal que $B - Q \in \mathbb{W}^\perp$. Assim, a solução do problema (5.4) é um ponto $X \in \mathbb{R}^n$ tal que $B - A\,X \in \mathcal{I}(A)^\perp$. Mas, Pelo Teorema 5.8 $\mathcal{I}(A)^\perp = \mathcal{N}(A^t)$. Assim, X é tal que

$$A^t(B - AX) = \bar{0}.$$

Ou seja, a solução do problema de quadrados mínimos é a solução do sistema linear

$$A^t A\,X = A^t B.$$

■

Exemplo 5.9. Vamos determinar a reta de equação $y = ax + b$ que melhor se ajusta aos pontos $P_1 = (-3, 6), P_2 = (0, 4), P_3 = (1, 0)$ e $P_4 = (2, 2)$ no sentido de quadrados mínimos, ou seja, tal que $\displaystyle\sum_{i=1}^{4}(y_i - ax_i - b)^2$ seja mínimo. Substituindo-se estes pontos na equação da reta obtemos o seguinte sistema

$$\left\{\begin{array}{rcrcr} -3a & + & b & = & 6 \\ & & b & = & 4 \\ a & + & b & = & 0 \\ 2a & + & b & = & 2 \end{array}\right.$$

Para este sistema temos que $A = \begin{bmatrix} -3 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 6 \\ 4 \\ 0 \\ 2 \end{bmatrix}$. Para encontrar a solução de quadrados mínimos deste sistema temos que resolver as equações normais $A^tAX = A^tB$. Neste caso,

$$A^tA = \begin{bmatrix} 14 & 0 \\ 0 & 4 \end{bmatrix} \quad \text{e} \quad A^tB = \begin{bmatrix} -14 \\ 12 \end{bmatrix}$$

Assim a solução de quadrados mínimos é $X = [-1\ 3]^t$, ou $a = -1, b = 3$. A reta $y = -x + 3$ é a reta procurada.

Exemplo 5.10. Vamos determinar a parábola de equação $y = ax^2 + bx + c$ que melhor se ajusta aos pontos $P_1 = (-2, 0), P_2 = (-1, 2), P_3 = (1, 2)$ e $P_4 = (2, 10)$ no sentido de quadrados mínimos, ou seja, tal que $\sum_{i=1}^{4}(y_i - ax_i^2 - bx_i - c)^2$ seja mínimo. Substituindo-se estes pontos na equação da parábola obtemos o seguinte sistema

$$\left\{ \begin{array}{rcrcrcr} 4a & - & 2b & + & c & = & 0 \\ a & - & b & + & c & = & 2 \\ a & + & b & + & c & = & 2 \\ 4a & + & 2b & + & c & = & 10 \end{array} \right.$$

Para este sistema temos que $A = \begin{bmatrix} 4 & -2 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 0 \\ 2 \\ 2 \\ 10 \end{bmatrix}$. Para encontrar a solução de quadrados mínimos deste sistema temos que resolver as equações normais $A^tAX = A^tB$. Aqui,

$$A^tA = \begin{bmatrix} 34 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 4 \end{bmatrix} \quad \text{e} \quad A^tB = \begin{bmatrix} 44 \\ 20 \\ 14 \end{bmatrix}$$

Escalonando a matriz aumentada $[A^tA|A^tB]$ obtemos que a solução de quadrados mínimos é $X = [1\ 2\ 1]^t$, ou $a = 1, b = 2$ e $c = 1$. E $y = x^2 + 2x + 1$ é a equação da parábola procurada.

Exemplo 5.11. Vamos determinar o círculo de equação $x^2 + y^2 = ax + by + c$ que melhor se ajusta aos pontos $P_1 = (-2,0), P_2 = (0,2), P_3 = (1,-3)$ e $P_4 = (3,1)$ no sentido de quadrados mínimos, ou seja, tal que $\sum_{i=1}^{4}(x_i^2 + y_i^2 - ax_i - by_i - c)^2$ seja mínimo. Substituindo-se estes pontos na equação do círculo obtemos o seguinte sistema

$$\left\{ \begin{array}{rcrcrcr} -2a & & & + & c & = & 4 \\ & + & 2b & + & c & = & 4 \\ a & - & 3b & + & c & = & 10 \\ 3a & + & b & + & c & = & 10 \end{array} \right.$$

Para este sistema temos que $A = \begin{bmatrix} -2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & -3 & 1 \\ 3 & 1 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 4 \\ 4 \\ 10 \\ 10 \end{bmatrix}$. Para encontrar a solução de quadrados mínimos deste sistema temos que resolver as equações normais $A^tAX = A^tB$. Aqui,

$$A^tA = \begin{bmatrix} 14 & 0 & 2 \\ 0 & 14 & 0 \\ 2 & 0 & 4 \end{bmatrix} \quad \text{e} \quad A^tB = \begin{bmatrix} 32 \\ -12 \\ 28 \end{bmatrix}$$

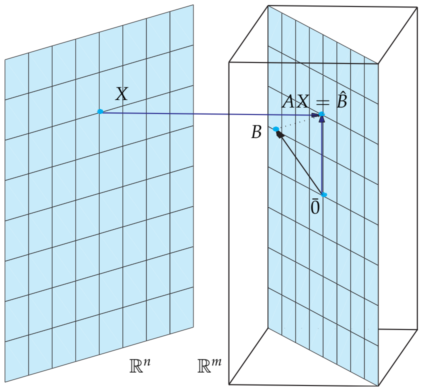


Figura 5.9. A solução de quadrados mínimos

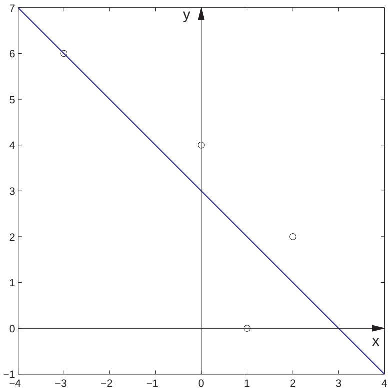


Figura 5.10. Reta que "melhor" se ajusta a quatro pontos

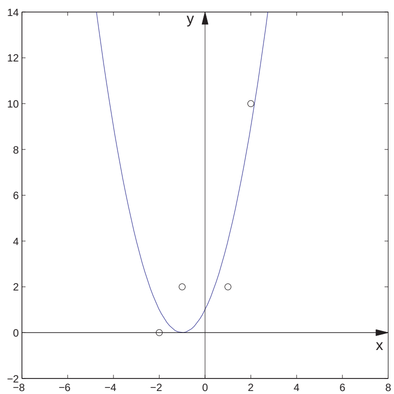


Figura 5.11. Parábola que "melhor" se ajusta a quatro pontos

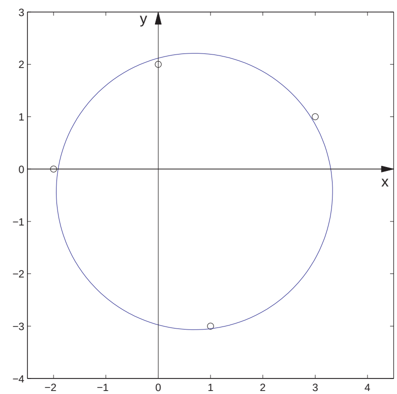


Figura 5.12. Círculo que "melhor" se ajusta a quatro pontos

Escalonando a matriz aumentada $[A^tA|A^tB]$ obtemos que a solução de quadrados mínimos é $X = [18/13 \;\; -6/7 \;\; 82/13]^t$, ou $a = 18/13, b = -6/7$ e $c = 82/13$. A equação do círculo procurado é $x^2 + y^2 - (18/13)x + (6/7)y = 82/13$. O centro do círculo $P_0 = (x_0, y_0)$ e o raio r são obtidos pelas equações $a = 2x_0, b = 2y_0$ e $r^2 = c + x_0^2 + y_0^2$. Assim, $x_0 = 9/13, y_0 = -3/7$ e $r = \sqrt{\frac{57724}{8281}} \approx 2,6$.

Exercícios Numéricos (respostas na página 546)

5.2.1. Para cada uma das seguintes matrizes determine uma base para cada um dos seguintes subespaços $\mathcal{I}(A^t), \mathcal{N}(A), \mathcal{I}(A)$ e $\mathcal{N}(A^t)$.

(a) $\begin{bmatrix} 4 & -2 \\ 1 & 3 \\ 2 & 1 \\ 3 & 4 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix}$

5.2.2. Seja $\mathbb{W}$ o subespaço de $\mathbb{R}^3$ gerado por $V = (1, -1, 1)$. Encontre uma base para $\mathbb{W}^\perp$ e dê uma interpretação geométrica para $\mathbb{W}$ e $\mathbb{W}^\perp$.

5.2.3. Seja $\mathbb{W}$ o subespaço do $\mathbb{R}^4$ gerado pelos vetores $V_1 = (1, 0, -2, 1)$ e $V_2 = (0, 1, 3, -2)$. Encontre uma base para $\mathbb{W}^\perp$.

5.2.4. Encontre a equação da parábola que melhor se ajusta aos pontos dados no sentido de quadrados mínimos, ou seja, tal que $\displaystyle\sum_{i=1}^{4}(y_i - ax_i^2 - bx_i - c)^2$ seja mínimo:

(a) $P_1 = (-2, 1), P_2 = (-1, 2), P_3 = (1, 0)$ e $P_4 = (2, 7)$.
(b) $P_1 = (-2, 1), P_2 = (-1, 3), P_3 = (1, 3)$ e $P_4 = (2, 11)$.

5.2.5. Encontre a equação do círculo que melhor se ajusta aos pontos dados no sentido de quadrados mínimos, ou seja, tal que $\displaystyle\sum_{i=1}^{4}(x_i^2 + y_i^2 - ax_i - by_i - c)^2$ seja mínimo:

(a) $P_1 = (-2, 0), P_2 = (0, 1), P_3 = (1, -2)$ e $P_4 = (2, 1)$.
(b) $P_1 = (-2, 1), P_2 = (-1, -2), P_3 = (0, 1)$ e $P_4 = (2, 0)$.

5.2.6. Encontre a solução de quadrados mínimos dos seguintes sistemas:

(a) $\left\{\begin{array}{rcrcr} x & + & 2y & = & 3 \\ 2x & + & 4y & = & 2 \\ -x & - & 2y & = & 1 \end{array}\right.$ (b) $\left\{\begin{array}{rcrcr} -x & + & y & = & 10 \\ 2x & + & y & = & 5 \\ x & - & 2y & = & 20 \end{array}\right.$ (c) $\left\{\begin{array}{rcrcrcr} x & + & y & + & z & = & 4 \\ -x & + & y & + & z & = & 2 \\ & - & y & + & z & = & 1 \\ x & & & + & z & = & 2 \end{array}\right.$

Exercícios usando o MATLAB®

5.2.7. (a) Use o comando `P=randi(5,2)`, para gerar 5 pontos com entradas inteiras e aleatórias entre -5 e 5. Os pontos estão armazenados nas linhas da matriz `P`.

(b) Use o MATLAB® para encontrar os coeficientes a, b, c e d da função polinomial $p(x) = ax^3 + bx^2 + cx + d$ que melhor se ajusta aos pontos dados pelas linhas da matriz `P`, no sentido de quadrados mínimos, ou seja, tal que $\sum(y_i - ax_i^3 - bx_i^2 - cx - d)^2$ seja mínimo. A matriz `A=matvand(P(:,1),3)` pode ser útil na solução deste problema, assim como a matriz `B=P(:,2)`.

(c) Desenhe os pontos e o gráfico do polinômio com os comandos `clf,po(P), syms x, plotf1(a*x^3+b*x^2+c*x+d,[-5,5])`, em que `a,b,c` e `d` são os coeficientes já encontrados. Desenhe os eixos coordenados com o comando `eixos`.

5.2.8. (a) Use o comando `P=randi(6,2)`, para gerar 6 pontos com entradas inteiras e aleatórias entre -5 e 5. Os pontos estão armazenados nas linhas da matriz `P`.

(b) Use o MATLAB® para encontrar os coeficientes a, b, c, d e e da cônica de equação $x^2 + axy + by^2 + cx + dy + e = 0$, cujo gráfico melhor se ajusta aos pontos dados pelas linhas da matriz `P`, no sentido de quadrados mínimos, ou seja, tal que $\sum(x_i^2 - ax_iy_i - by_i^2 - cx_i - dy_i - e)^2$ seja mínimo. As matrizes `M=matvand(P,2)`, `B=-M(:,1)` e `A=M(:,2:6)` podem ser úteis na solução deste problema.

(c) Desenhe os pontos e a cônica com os comandos `clf,po(P), syms x y, plotci(x^2+a*x*y+b*y^2+c*x+d*y+e,[-5,5],[-5,5])`, em que `a,b,c,d` e `e` são os coeficientes encontrados no item anterior. Desenhe os eixos coordenados com o comando `eixos`.

Exercícios Teóricos

5.2.9. Seja A_j uma coluna não nula de uma matriz A, $m \times n$. É possível que A_j pertença ao $\mathcal{N}(A^t)$?

5.2.10. Seja $\mathbb{W}$ um subespaço de $\mathbb{R}^n$ gerado pelos vetores $V_1, \ldots, V_k$. Mostre que $V \in \mathbb{W}^\perp$ se, e somente se, V é ortogonal a V_i, para $i = 1, \ldots, k$.

5.2.11. Sejam $\mathbb{W}_1$ e $\mathbb{W}_2$ subespaços de $\mathbb{R}^n$. Mostre que

$$(\mathbb{W}_1 + \mathbb{W}_2)^\perp = \mathbb{W}_1^\perp \cap \mathbb{W}_2^\perp$$

e que

$$(\mathbb{W}_1 \cap \mathbb{W}_2)^\perp = \mathbb{W}_1^\perp + \mathbb{W}_2^\perp.$$

5.2.12. Sejam $\mathcal{S}$ e $\mathcal{S}_0$ subconjuntos de $\mathbb{R}^n$. Mostre que $\mathcal{S}_0 \subseteq \mathcal{S}$ implica que $\mathcal{S}^\perp \subseteq \mathcal{S}_0^\perp$.

5.2.13. Se A é uma matriz $m \times n$ de posto igual a r, quais as dimensões de $\mathcal{N}(A)$ e $\mathcal{N}(A^t)$?

5.2.14. Seja A uma matriz $m \times n$.

(a) Mostre que se $AX = \bar{0}$, então $A^tAX = \bar{0}$;

(b) Mostre que se $A^tAX = \bar{0}$, então $AX = \bar{0}$; (Sugestão: Use o fato de que $A^tAX = \bar{0}$ se, e somente se, $AX \in \mathcal{N}(A^t)$.)

(c) Mostre que $\mathcal{N}(A^tA) = \mathcal{N}(A)$.

(d) Mostre que se A^tA é invertível, então as colunas de A são linearmente independentes.

(e) Mostre que se A é uma matriz cujas colunas são linearmente independentes, então A^tA é uma matriz invertível. Por que neste caso, $m \geq n$?

(f) Mostre que posto$(A) =$ posto(A^tA).

5.2.15. Sejam A uma matriz $m \times n$ com colunas linearmente independentes e B uma matriz $m \times 1$. Mostre que neste caso, a matriz A^tA é invertível e que vale a seguinte fórmula para a solução do problema de quadrados mínimos, min $||AX - B||^2$,

$$X = (A^tA)^{-1}A^tB.$$

5.3 Mudança de Coordenadas

Se as coordenadas de um ponto P no espaço são (x, y, z), então as componentes do vetor $\overrightarrow{OP}$ também são (x, y, z) e então podemos escrever

$$\begin{aligned} \overrightarrow{OP} &= (x, y, z) = (x, 0, 0) + (0, y, 0) + (0, 0, z) \\ &= x(1, 0, 0) + y(0, y, 0) + z(0, 0, 1) = x\vec{i} + y\vec{j} + z\vec{k}, \end{aligned}$$

em que $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$ e $\vec{k} = (0, 0, 1)$. Ou seja, as coordenadas de um ponto P são iguais aos escalares que aparecem ao escrevermos $\overrightarrow{OP}$ como uma combinação linear dos vetores canônicos. Assim, o ponto $O = (0, 0, 0)$ e os vetores $\vec{i}$, $\vec{j}$ e $\vec{k}$ determinam um sistema de coordenadas ortogonal, $\{O, \vec{i}, \vec{j}, \vec{k}\}$. Para resolver alguns problemas geométricos é necessário usarmos um segundo **sistema de coordenadas ortogonal** determinado por uma origem O' e por 3 vetores U_1, U_2 e U_3 ortonormais de $\mathbb{R}^3$.* Por exemplo, se $O' = (2, 3/2, 3/2)$, $U_1 = (\sqrt{3}/2, 1/2, 0)$, $U_2 = (-1/2, \sqrt{3}/2, 0)$ e $U_3 = (0, 0, 1) = \vec{k}$, então $\{O', U_1, U_2, U_3\}$ determina um novo sistema de coordenadas: aquele com origem no ponto O', cujos eixos x', y' e z' são retas que passam por O' orientadas com os sentidos e direções de U_1, U_2 e U_3, respectivamente (Figura 5.14).

As coordenadas de um ponto P no sistema de coordenadas $\{O', U_1, U_2, U_3\}$ é definido como sendo os escalares que aparecem ao escrevermos $\overrightarrow{O'P}$ como combinação linear dos vetores U_1, U_2 e U_3, ou seja, se

$$\overrightarrow{O'P} = x'U_1 + y'U_2 + z'U_3,$$

*Em geral, um sistema de coordenadas (**não** necessariamente ortogonal) é definido por um ponto O' e três vetores V_1, V_2 e V_3 L.I. de $\mathbb{R}^3$ (não necessariamente ortonormais) (veja o Exercício 5.3.9 na página 333).

então as coordenadas de P no sistema $\{O', U_1, U_2, U_3\}$ são dadas por

$$[P]_{\{O',U_1,U_2,U_3\}} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}.$$

Vamos considerar inicialmente o caso em que $O = O'$. Assim, se $\overrightarrow{OP} = (x, y, z)$, então $x'U_1 + y'U_2 + z'U_3 = \overrightarrow{OP}$ é equivalente ao sistema linear

$$QX' = X, \quad \text{em que} \quad Q = [\, U_1\ U_2\ U_3 \,],\ X' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix},\ X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Como a matriz Q é invertível (por que?) a solução é dada por

$$X' = Q^{-1}X.$$

Mas, como U_1, U_2 e U_3 formam uma base ortonormal de $\mathbb{R}^3$, então

$$Q^tQ = \begin{bmatrix} U_1^t \\ U_2^t \\ U_3^t \end{bmatrix} [\, U_1\ U_2\ U_3 \,] = \begin{bmatrix} U_1^tU_1 & U_1^tU_2 & U_1^tU_3 \\ U_2^tU_1 & U_2^tU_2 & U_2^tU_3 \\ U_3^tU_1 & U_3^tU_2 & U_3^tU_3 \end{bmatrix} = \begin{bmatrix} U_1 \cdot U_1 & U_1 \cdot U_2 & U_1 \cdot U_3 \\ U_2 \cdot U_1 & U_2 \cdot U_2 & U_2 \cdot U_3 \\ U_3 \cdot U_1 & U_3 \cdot U_2 & U_3 \cdot U_3 \end{bmatrix} = I_3$$

Assim, a matriz $Q = [\, U_1\ U_2\ U_3 \,]$ é invertível e $Q^{-1} = Q^t$. Desta forma as coordenadas de um ponto P no espaço em relação ao sistema $\{O, U_1, U_2, U_3\}$, x', y' e z' estão unicamente determinados e

$$[P]_{\{O,U_1,U_2,U_3\}} = Q^t[P]_{\{O,\vec{i},\vec{j},\vec{k}\}} \quad \text{ou} \quad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = Q^t \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Também no plano temos o mesmo tipo de situação que é tratada de forma inteiramente análoga. As coordenadas de um ponto P no plano em relação a um sistema de coordenadas $\{O', U_1, U_2\}$, em que U_1 e U_2 são vetores que formam uma base ortonormal do $\mathbb{R}^2$, é definido como sendo os escalares que aparecem ao escrevermos $\overrightarrow{O'P}$ como combinação linear de U_1 e U_2, ou seja, se

$$\overrightarrow{O'P} = x'U_1 + y'U_2,$$

então as coordenadas de P no sistema $\{O', U_1, U_2\}$ são dadas por

$$[P]_{\{O',U_1,U_2\}} = \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

As coordenadas de um ponto P no plano em relação ao sistema $\{O, U_1, U_2, U_3\}$ estão bem definidas, ou seja, x' e y' estão unicamente determinados e são dados por

$$[P]_{\{O,U_1,U_2\}} = Q^t [P]_{\{O,E_1,E_2\}} \quad \text{ou} \quad \begin{bmatrix} x' \\ y' \end{bmatrix} = Q^t \begin{bmatrix} x \\ y \end{bmatrix},$$

em que $E_1 = (1, 0)$ e $E_2 = (0, 1)$. Observe que, tanto no caso do plano quanto no caso do espaço, a matriz Q satisfaz, $Q^{-1} = Q^t$. Uma matriz que satisfaz esta propriedade é chamada **matriz ortogonal**.

Exemplo 5.12. Considere o sistema de coordenadas no plano em que $O' = O$ e $U_1 = (\sqrt{3}/2, 1/2)$ e $U_2 = (-1/2, \sqrt{3}/2)$. Se $P = (2, 4)$, vamos determinar as coordenadas de P em relação ao novo sistema de coordenadas.

$$Q = [\ U_1\ U_2\ ] = \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix}.$$

Assim as coordenadas de P em relação ao novo sistema de coordenadas são dadas por

$$[P]_{\{O,U_1,U2\}} = Q^t \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} U_1^t \\ U_2^t \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} \sqrt{3}/2 & 1/2 \\ -1/2 & \sqrt{3}/2 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 2+\sqrt{3} \\ 2\sqrt{3}-1 \end{bmatrix}.$$

Exemplo 5.13. Considere o mesmo sistema de coordenadas do exemplo anterior, mas agora seja $P = (x, y)$ um ponto qualquer do plano. Vamos determinar as coordenadas de P em relação ao novo sistema de coordenadas.

As coordenadas de P em relação ao novo sistema de coordenadas são dadas por

$$[P]_{\{O,U_1,U2\}} = Q^t \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} U_1^t \\ U_2^t \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \sqrt{3}/2 & 1/2 \\ -1/2 & \sqrt{3}/2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (\sqrt{3}\,x+y)/2 \\ (-x+\sqrt{3}\,y)/2 \end{bmatrix}.$$

Exemplo 5.14. Vamos agora considerar um problema inverso àqueles apresentados nos exemplos anteriores. Suponha que sejam válidas as seguintes equações

$$\left\{ \begin{array}{lcl} x = \frac{1}{\sqrt{5}}x' & + & \frac{2}{\sqrt{5}}y' \\ y = \frac{2}{\sqrt{5}}x' & - & \frac{1}{\sqrt{5}}y' \end{array} \right. ,$$

ou equivalentemente

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

entre as coordenadas $\begin{bmatrix} x' \\ y' \end{bmatrix}$ de um ponto P em relação a um sistema de coordenadas $\{O, U_1, U_2\}$ e as coordenadas de P, $\begin{bmatrix} x \\ y \end{bmatrix}$, em relação ao sistema de coordenadas original

$$\{O, E_1 = (1,0), E_2 = (0,1)\}.$$

Queremos determinar quais são os vetores U_1 e U_2.

Os vetores U_1 e U_2 da nova base possuem coordenadas $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ e $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, respectivamente, em relação ao novo sistema de coordenadas, $\{O, U_1, U_2\}$. Pois, $U_1 = 1\,U_1 + 0\,U_2$ e $U_2 = 0\,U_1 + 1\,U_2$. Queremos saber quais as coordenadas destes vetores em relação ao sistema de coordenadas original, $\{O, E_1 = (1,0), E_2 = (0,1)\}$. Logo,

$$\begin{aligned} U_1 &= \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} \\ U_2 &= \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{bmatrix} \end{aligned}$$

Ou seja, U_1 e U_2 são as colunas da matriz $Q = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix}$.

5.3.1 Rotação

Suponha que o novo sistema de coordenadas $\{O, U_1, U_2\}$ seja obtido do sistema original $\{O, E_1 = (1,0), E_2 = (0,1)\}$ por uma rotação de um ângulo θ. Observando a

Figura 5.16, obtemos

$$\begin{array}{rcl} U_1 & = & (\cos\theta, \operatorname{sen}\theta) \\ U_2 & = & (-\operatorname{sen}\theta, \cos\theta) \end{array}$$

seja $P = (x, y)$ um ponto qualquer do plano. Vamos determinar as coordenadas de P em relação ao novo sistema de coordenadas.

A matriz

$$Q = [\ U_1\ U_2\ ] = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} = R_\theta$$

é chamada **matriz de rotação**.

As coordenadas de P em relação ao novo sistema de coordenadas são dadas por

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = R_\theta^t \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

O sistema de coordenadas que aparece nos dois primeiros exemplos desta seção podem ser obtidos por uma rotação de um ângulo $\theta = \pi/6$ em relação ao sistema original.

5.3.2 Translação

Vamos considerar, agora, o caso em que $O' \neq O$, ou seja, em que ocorre uma **translação** dos eixos coordenados.

Observando a Figura 5.17, obtemos

$$\overrightarrow{O'P} = \overrightarrow{OP} - \overrightarrow{OO'}\ . \tag{5.6}$$

Assim, se $\overrightarrow{OO'} = (h, k)$, então

$$\overrightarrow{O'P} = (x', y') = (x, y) - (h, k) = (x - h, y - k)$$

Logo, as coordenadas de P em relação ao novo sistema são dadas por

$$[P]_{\{O', E_1, E_2\}} = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x - h \\ y - k \end{bmatrix}. \tag{5.7}$$

O eixo x$'$ tem equação $y' = 0$, ou seja, $y = k$ e o eixo y$'$, $x' = 0$, ou seja, $x = h$.

5.3.3 Aplicação: Computação Gráfica - Projeção Ortográfica

Esta projeção é usada para fazer desenhos de objetos tridimensionais no papel ou na tela do computador. Com esta projeção os pontos no espaço são projetados ortogonalmente ao plano do desenho.

Para encontrar a projeção de um ponto P podemos encontrar as coordenadas de P em relação ao sistema $\mathcal{S}' = \{O', U_1, U_2, U_3\}$ e tomar as duas primeiras coordenadas.

Como a projeção em qualquer plano paralelo ao plano do desenho fornece as mesmas coordenadas podemos supor que $O' = O$, ou seja, que os dois sistemas têm a mesma origem.

A relação entre as coordenadas de um ponto nos dois sistemas

$$\mathcal{S}' = \{O, U_1, U_2, U_3\} \quad \text{e} \quad \mathcal{S} = \{O, \vec{i}, \vec{j}, \vec{k}\}$$

é dada por

$$X' = Q^t X, \quad \text{em que } Q = [\, U_1 \; U_2 \; U_3 \,]$$

Vamos encontrar os vetores U_1, U_2 e U_3 em função dos ângulos θ e ϕ. O vetor U_1 é paralelo ao plano xy e é perpendicular ao vetor $(\cos\theta, \operatorname{sen}\theta, 0)$, ou seja,

$$U_1 = (-\operatorname{sen}\theta, \cos\theta, 0).$$

Os vetores U_2 e U_3 estão no plano definido por $\vec{k}$ e $(\cos\theta, \operatorname{sen}\theta, 0)$.

$$\begin{aligned}
U_2 &= -\cos\phi(\cos\theta, \operatorname{sen}\theta, 0) + \operatorname{sen}\phi\vec{k} = (-\cos\phi\cos\theta, -\cos\phi\operatorname{sen}\theta, \operatorname{sen}\phi) \\
U_3 &= \cos\phi\vec{k} + \operatorname{sen}\phi(\cos\theta, \operatorname{sen}\theta, 0) = (\operatorname{sen}\phi\cos\theta, \operatorname{sen}\phi\operatorname{sen}\theta, \cos\phi)
\end{aligned}$$

Assim a relação entre as coordenadas de um ponto nos dois sistemas

$$\mathcal{S}' = \{O, U_1, U_2, U_3\} \quad \text{e} \quad \mathcal{S} = \{O, \vec{i}, \vec{j}, \vec{k}\}$$

é dada por

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} -\operatorname{sen}\theta & \cos\theta & 0 \\ -\cos\phi\cos\theta & -\cos\phi\operatorname{sen}\theta & \operatorname{sen}\phi \\ \operatorname{sen}\phi\cos\theta & \operatorname{sen}\phi\operatorname{sen}\theta & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

e a projeção é dada por

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -\operatorname{sen}\theta & \cos\theta & 0 \\ -\cos\phi\cos\theta & -\cos\phi\operatorname{sen}\theta & \operatorname{sen}\phi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Por exemplo para $\theta = 30^\circ$ e $\phi = 60^\circ$ temos que

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{4} & -\frac{1}{4} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \approx \begin{bmatrix} -0.50 & 0.87 & 0 \\ -0.43 & -0.25 & 0.87 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Usando esta projeção os vetores $\vec{i}$, $\vec{j}$ e $\vec{k}$ são desenhados como na figura abaixo.

Experimente desenhar o cubo que tem a origem $O = (0,0,0)$ como um dos vértices e como vértices adjacentes à origem $(1,0,0)$, $(0,1,0)$ e $(0,0,1)$. Observe que não é necessário calcular a projeção dos outros pontos (por que?)

Exercícios Numéricos (respostas na página 550)

5.3.1. Encontre as coordenadas do ponto P com relação ao sistema de coordenadas $\mathcal{S}$, nos seguintes casos:

(a) $\mathcal{S} = \{O, (1/\sqrt{2}, -1/\sqrt{2}), (1/\sqrt{2}, 1/\sqrt{2})\}$ e $P = (1, 3)$;

(b) $\mathcal{S} = \{O, (1/\sqrt{2}, -1/\sqrt{2}, 0), (0, 0, 1), (1/\sqrt{2}, 1/\sqrt{2}, 0)\}$ e $P = (2, -1, 2)$;

5.3.2. Encontre o ponto P, se as coordenadas de P em relação ao sistema de coordenadas $\mathcal{S}$, $[P]_{\mathcal{S}}$, são:

(a) $[P]_{\mathcal{S}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, em que $\mathcal{S} = \{O, (-1/\sqrt{2}, 1/\sqrt{2}), (1/\sqrt{2}, 1/\sqrt{2})\}$.

(b) $[P]_{\mathcal{S}} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$, em que $\mathcal{S} = \{O, (0, 1/\sqrt{2}, -1/\sqrt{2}), (1, 0, 0), (0, 1/\sqrt{2}, 1/\sqrt{2})\}$.

5.3.3. Sejam $[P]_{\mathcal{R}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ as coordenadas de um ponto P em relação ao sistema de coordenadas $\mathcal{R} = \{O, \vec{i}, \vec{j}, \vec{k}\}$ e $[P]_{\mathcal{S}} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$, em relação ao sistema de coordenadas $\mathcal{S} = \{O, U_1, U_2, U_3\}$. Suponha que temos a seguinte relação:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}.$$

Quais são os vetores U_1, U_2 e U_3?

5.3.4. Determine qual a rotação do plano em que as coordenadas do ponto $P = (\sqrt{3}, 1)$ são $\begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}$.

5.3.5. Considere o plano $\pi : \; 3x - \sqrt{3}y + 2z = 0$.

(a) Determine uma base ortonormal para o plano em que o primeiro vetor esteja no plano xy.

(b) Complete a base encontrada para se obter uma base ortonormal $\{U_1, U_2, U_3\}$ de $\mathbb{R}^3$.

(c) Determine as coordenadas dos vetores $\vec{i}$, $\vec{j}$ e $\vec{k}$ no sistema $\{O, U_1, U_2, U_3\}$.

5.3.6. Considere dois sistemas de coordenadas $\mathcal{R} = \{O, \vec{i}, \vec{j}, \vec{k}\}$ e $\mathcal{S} = \{O, \vec{i}, U_2, U_3\}$, em que o sistema $\mathcal{S}$ é obtido do sistema $\mathcal{R}$ por uma rotação do ângulo θ em torno do eixo x. Determine a relação entre as coordenadas, (x', y', z'), em relação ao sistema $\mathcal{S}$ e (x, y, z), em relação ao sistema $\mathcal{R}$

Exercícios Teóricos

5.3.7. Mostre que

(a) $R_{\theta_1} R_{\theta_2} = R_{\theta_1+\theta_2}$.

(b) $R_\theta^{-1} = R_{-\theta}$.

5.3.8. Seja B uma matriz quadrada 2×2.

(a) Verifique que $R_\theta B$ é a matriz obtida girando as colunas de B de θ.

(b) Verifique que BR_θ é a matriz obtida girando as linhas de B de $-\theta$.

(c) Quais as condições sobre B e θ para que $R_\theta B = BR_\theta$. Dê um exemplo.

5.3.9. Definimos coordenadas de pontos no espaço em relação a um sistema de coordenadas determinado por um ponto O' e três vetores V_1, V_2 e V_3 L.I. não necessariamente ortonormais do $\mathbb{R}^3$ da mesma forma como fizemos quando os vetores formam uma base ortonormal. As coordenadas de um ponto P no sistema de coordenadas $\{O', V_1, V_2, V_3\}$ é definido como sendo os escalares que aparecem ao escrevermos $\overrightarrow{O'P}$ como combinação linear dos vetores V_1, V_2 e V_3, ou seja, se

$$\overrightarrow{O'P} = x'V_1 + y'V_2 + z'V_3,$$

então as coordenadas de P no sistema $\{O', V_1, V_2, V_3\}$ são dadas por

$$[P]_{\{O',V_1,V_2,V_3\}} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}.$$

Assim, se $\overrightarrow{O'P} = (x, y, z)$, então $x'V_1 + y'V_2 + z'V_3 = \overrightarrow{O'P}$ pode ser escrito como

$$[\ V_1\ V_2\ V_3\ ] \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

(a) Mostre que a matriz $Q = [\, V_1\ V_2\ V_3 \,]$ é invertível.

(b) Mostre que as coordenadas de um ponto P no espaço em relação ao sistema $\{O', V_1, V_2, V_3\}$ estão bem definidas, ou seja, x', y' e z' estão unicamente determinados e são dados por

$$[P]_{\{O',V_1,V_2,V_3\}} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = Q^{-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = Q^{-1}[P]_{\{O',\vec{i},\vec{j},\vec{k}\}}.$$

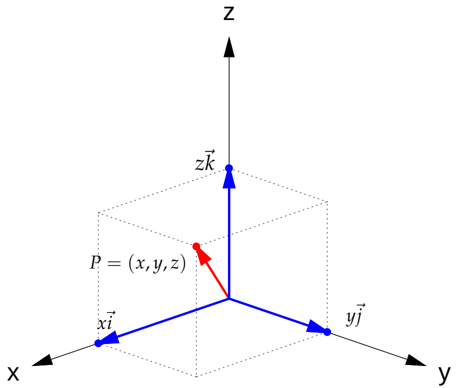


Figura 5.13. $\overrightarrow{OP} = x\vec{i} + y\vec{j} + z\vec{k}$

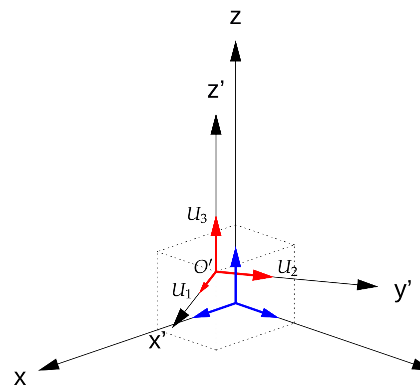


Figura 5.14. Dois sistemas de coordenadas ortogonais $\{O, \vec{i}, \vec{j}, \vec{k}\}$ e $\{O', U_1, U_2, U_3\}$

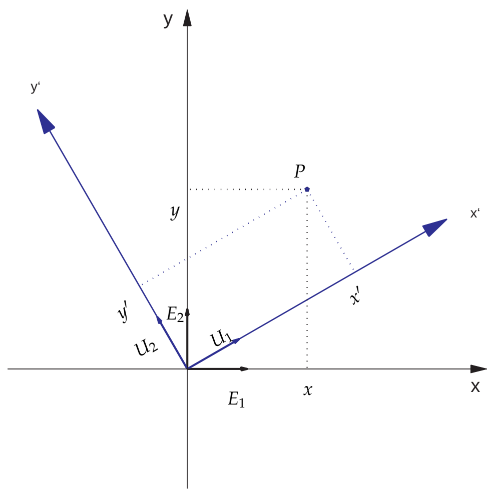


Figura 5.15. Coordenadas de um ponto P em dois sistemas

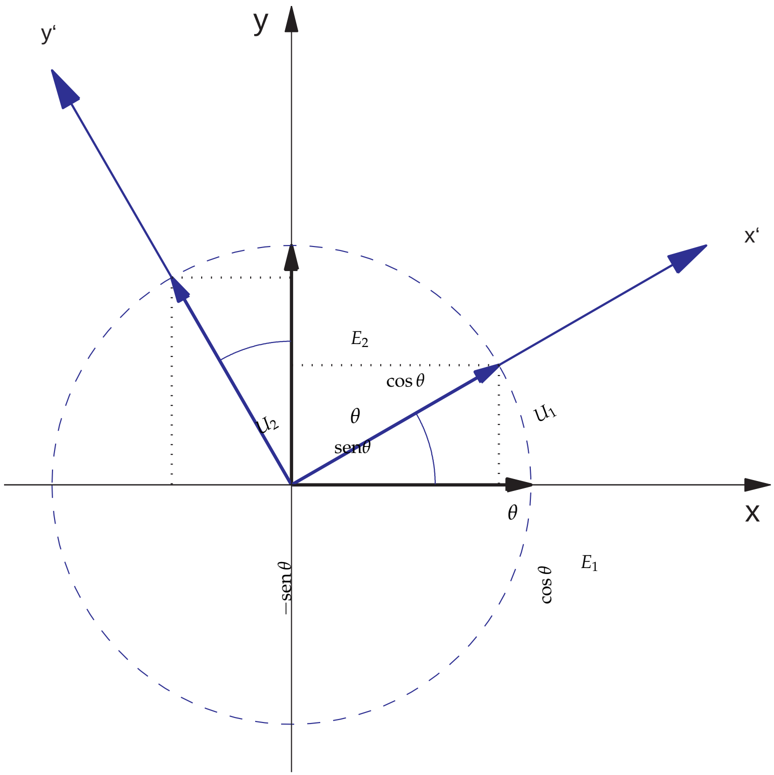


Figura 5.16. Rotação de um ângulo θ

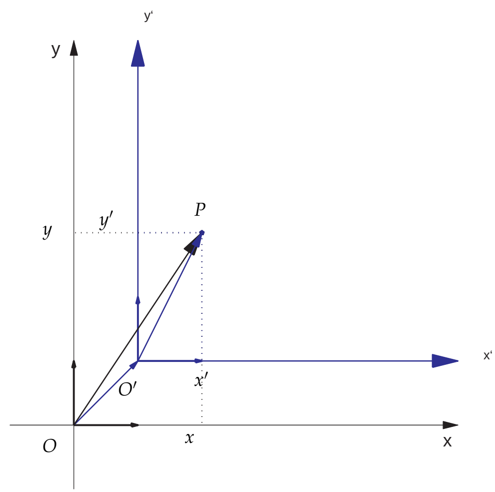


Figura 5.17. Coordenadas de um ponto P em dois sistemas (translação)

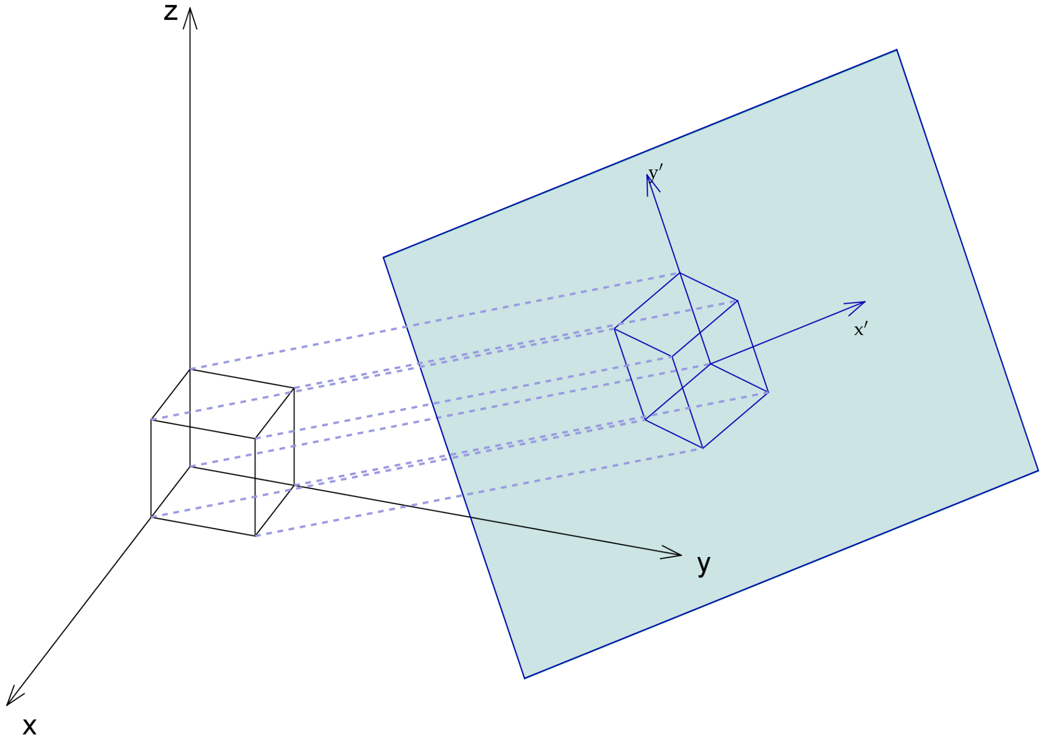


Figura 5.18. Projeção ortográfica de um cubo

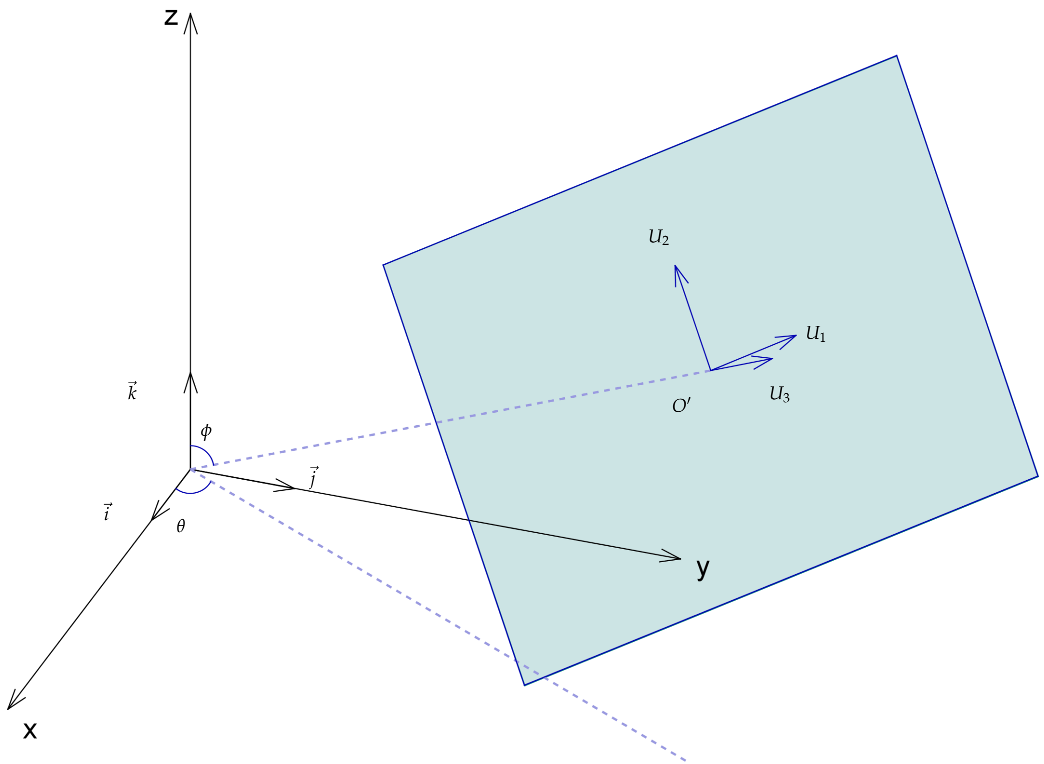


Figura 5.19. sistemas de coordenadas relacionados à projeção ortográfica

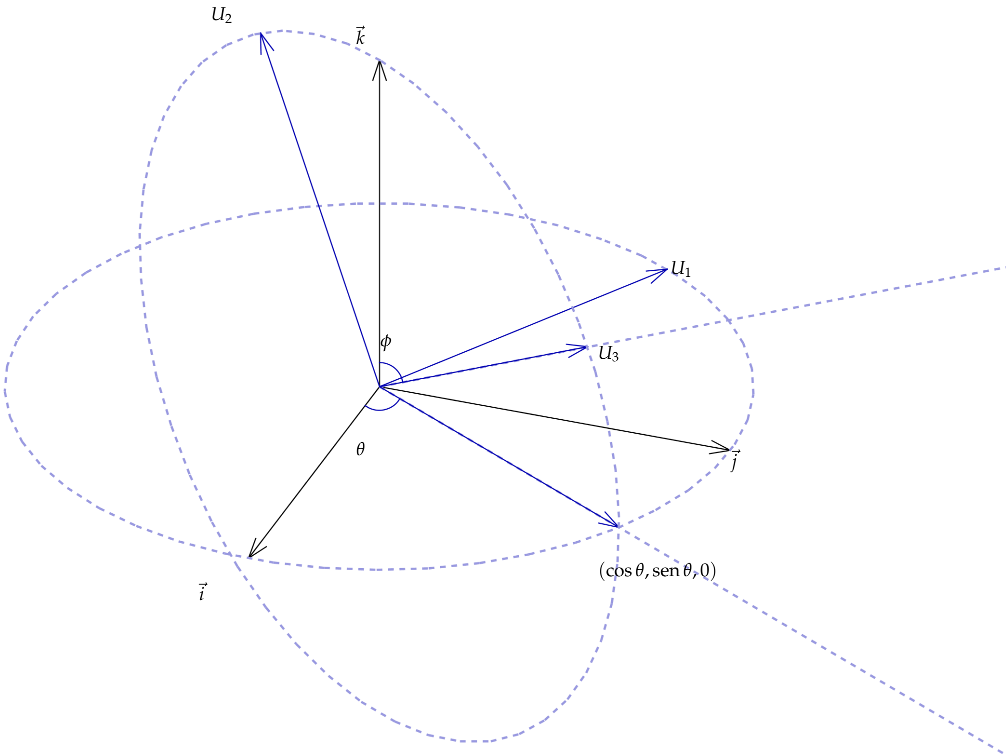


Figura 5.20. Bases relacionadas à projeção ortográfica

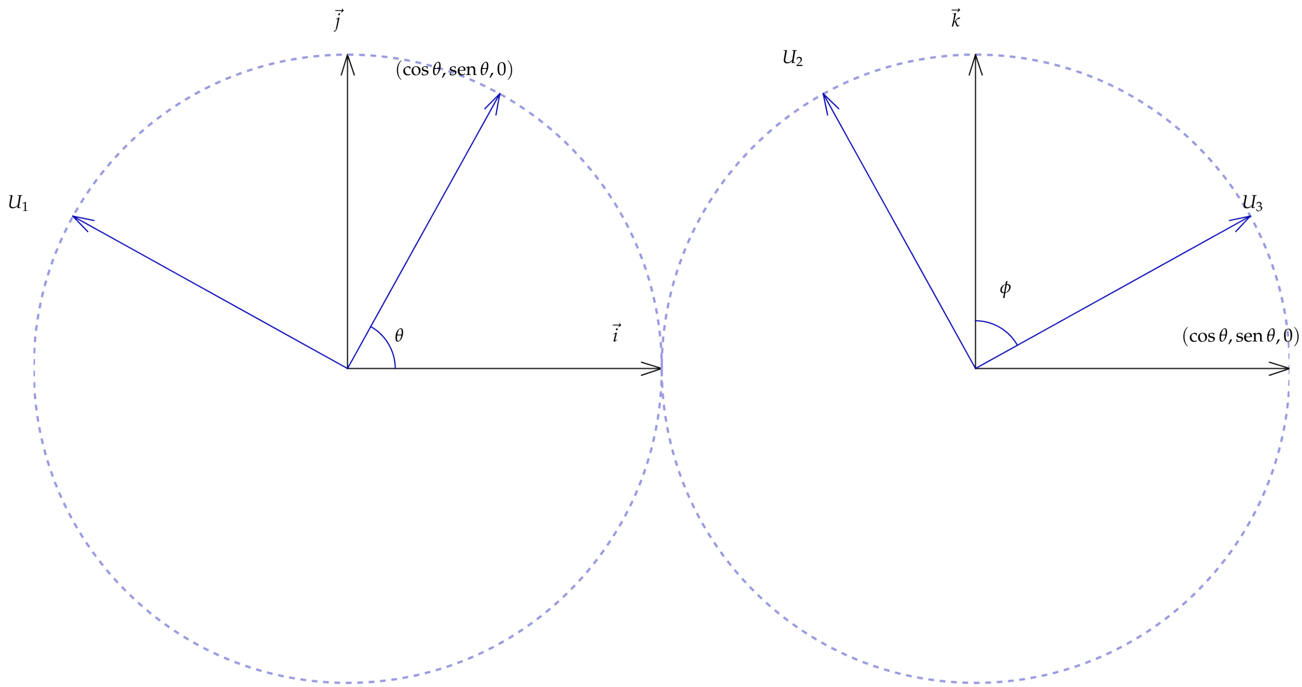


Figura 5.21. Relação entre os vetores das bases $\{U_1, U_2, U_3\}$ e $\{\vec{i}, \vec{j}, \vec{k}\}$

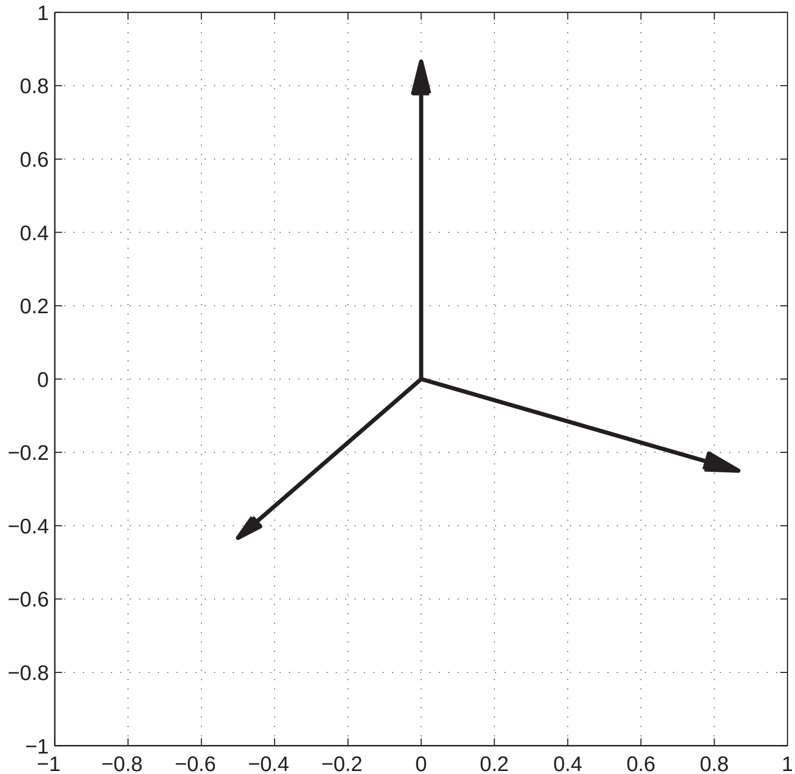


Figura 5.22. Vetores $\vec{i}$, $\vec{j}$ e $\vec{k}$ desenhados usando projeção ortográfica

6

Transformações Lineares (opcional)

6.1 Definição, Exemplos e Propriedades

6.1.1 Definição e Exemplos

Lembramos que uma **função** f **de um conjunto** A **em um conjunto** B, $f : A \to B$, é uma regra que associa a cada elemento do conjunto A, um único elemento do conjunto B. O conjunto A é chamado **domínio** e o conjunto B é chamado **contradomínio**. O subconjunto de B formado pelos elementos $b \in B$ tais que $f(a) = b$, para algum $a \in A$ é chamado **(conjunto) imagem** de f. Para todo elemento $a \in A$, $f(a)$ é chamado a **imagem de** a **por** f. Dizemos também que f **leva** a **em** $f(a)$.

Definição 6.1. Uma função $T : \mathbb{R}^n \to \mathbb{R}^m$ é uma **transformação linear** se

$$\boxed{T(\alpha X) = \alpha T(X)} \quad \text{e} \quad \boxed{T(X+Y) = T(X) + T(Y),} \tag{6.1}$$

para todos $X, Y \in \mathbb{R}^n$ e todos os escalares α.

Exemplo 6.1. A função O, que leva todo vetor de $\mathbb{R}^n$ no vetor nulo de $\mathbb{R}^m$, ou seja,

$$O(X) = \bar{0}, \quad \text{para todo } X \in \mathbb{R}^n$$

é claramente uma transformação linear e é chamada a **transformação linear nula**. Também a **transformação identidade**, I, de $\mathbb{R}^n$ em $\mathbb{R}^n$ que leva todo vetor de $\mathbb{R}^n$ nele mesmo, ou seja,

$$I(X) = X, \quad \text{para todo } X \in \mathbb{R}^n$$

é claramente uma transformação linear.

Exemplo 6.2. Sejam $P_x, P_y : \mathbb{R}^2 \to \mathbb{R}^2$ as funções que levam todo vetor nas suas projeções nos eixos x e y, respectivamente, ou seja,
$P_x(x, y) = (x, 0)$ e $P_y(x, y) = (0, y)$, para todo par $(x, y) \in \mathbb{R}^2$. Deixamos para o leitor a verificação de que P_x e P_y são transformações lineares.

Exemplo 6.3. Sejam $R_x, R_y : \mathbb{R}^2 \to \mathbb{R}^2$ as funções que levam todo vetor nas suas reflexões em relação aos eixos x e y, respectivamente, ou seja, $R_x(x, y) = (x, -y)$ e $R_y(x, y) = (-x, y)$, para todo par $(x, y) \in \mathbb{R}^2$. Deixamos para o leitor a verificação de que R_x e R_y são transformações lineares.

Exemplo 6.4. Considere a função, P_r, que faz a projeção ortogonal de todo vetor do plano numa reta que passa pela origem $r : (x, y) = t(a, b)$, ou seja, $P_r : \mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$P_r(x, y) = \text{proj}_{(a,b)}(x, y) = \frac{(a, b) \cdot (x, y)}{||(a, b)||^2}(a, b).$$

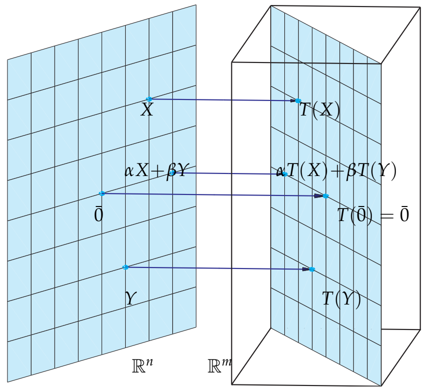


Figura 6.1. Transformação Linear $T : \mathbb{R}^n \to \mathbb{R}^m$

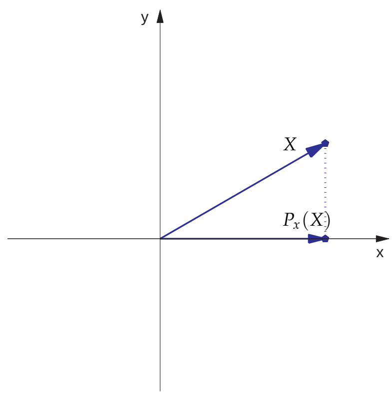


Figura 6.2. Projeção no eixo x

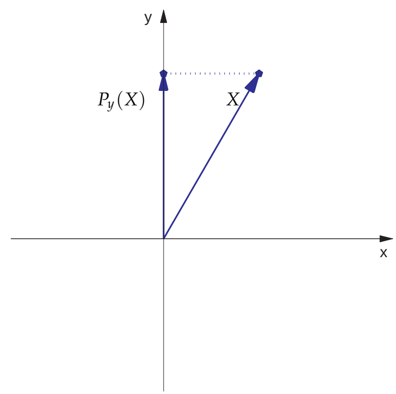


Figura 6.3. Projeção no eixo y

Ou seja,

$$P_r(x,y) = \left(\frac{a^2}{a^2+b^2}x + \frac{ab}{a^2+b^2}y, \frac{ab}{a^2+b^2}x + \frac{b^2}{a^2+b^2}y\right).$$

Esta transformação é um caso particular daquela que é tratada no Exemplo 6.6.

Exemplo 6.5. Considere a função, R_r, que faz a reflexão de todo vetor do plano em relação a uma reta que passa pela origem $r : (x,y) = t(a,b)$, ou seja, $R_r(x,y)$ é tal que $2P_r(x,y) = (x,y) + R_r(x,y)$. Assim,

$$\begin{aligned} R_r(x,y) &= 2P_r(x,y) - (x,y) \\ &= \left(\frac{a^2-b^2}{a^2+b^2}x + \frac{2ab}{a^2+b^2}y, \frac{2ab}{a^2+b^2}x + \frac{b^2-a^2}{a^2+b^2}y\right). \end{aligned}$$

Esta transformação é um caso particular daquela que é tratada no Exemplo 6.6.

Os quatro últimos exemplos são um caso particular do que é apresentado no próximo exemplo.

Exemplo 6.6. Considere a transformação $T : \mathbb{R}^n \to \mathbb{R}^m$, dada por

$$T(X) = T\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ a_{21}x_1 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{bmatrix}, \quad \text{para todo } X \in \mathbb{R}^n,$$

que pode ser escrita como

$$T(X) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\,X,$$

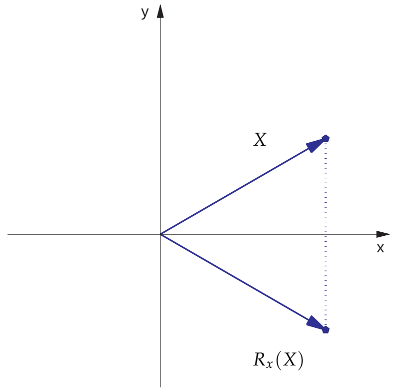


Figura 6.4. Reflexão em relação ao eixo x

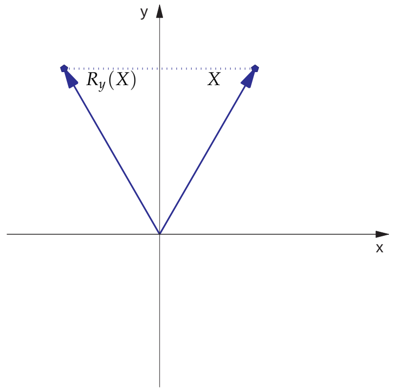


Figura 6.5. Reflexão em relação ao eixo y

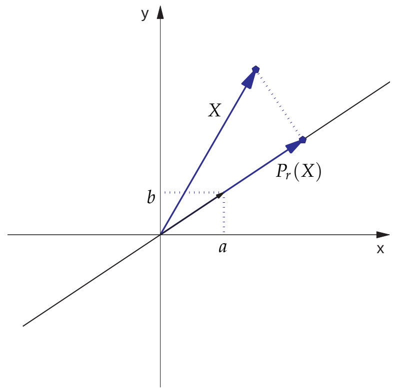


Figura 6.6. Projeção na reta r

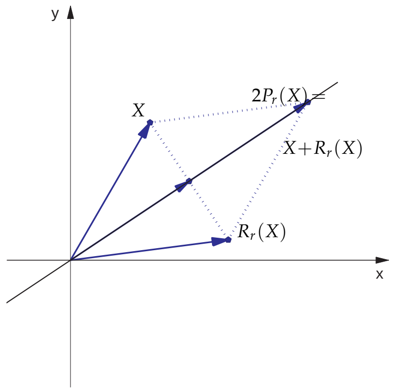


Figura 6.7. Reflexão em relação a reta r

em que $A = (a_{ij})_{m\times n}$. Segue das propriedades da álgebra matricial, Teorema 1.1 na página 9, que T é uma transformação linear. Pois,

$$T(X+Y) = A(X+Y) = AX + AY = T(X) + T(Y),$$

$$T(\alpha X) = A(\alpha X) = \alpha AX = \alpha T(X),$$

para todos $X, Y \in \mathbb{R}^n$ e escalares α.

Em termos de matrizes, as projeções nos eixos x e y podem ser escritas como

$$P_x \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad P_y \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

as reflexões em relação aos eixos x e y, como

$$R_x \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad R_y \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

a projeção ortogonal e a reflexão em relação a uma reta $r : (x, y) = t(a, b)$, como

$$P_r \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a^2}{a^2+b^2} & \frac{ab}{a^2+b^2} \\ \frac{ab}{a^2+b^2} & \frac{b^2}{a^2+b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad R_r \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a^2-b^2}{a^2+b^2} & \frac{2ab}{a^2+b^2} \\ \frac{2ab}{a^2+b^2} & \frac{b^2-a^2}{a^2+b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

6.1.2 Propriedades

Segue da Definição 6.1 que toda transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ leva o vetor nulo de $\mathbb{R}^n$ no vetor nulo de $\mathbb{R}^m$. Pois, se X é um vetor qualquer de $\mathbb{R}^n$, então

$$T(\bar{0}) = T(0\,X) = 0\,T(X) = \bar{0}.$$

Segue também da Definição 6.1 que uma função $T : \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear se, e somente se, $T(\alpha X + \beta Y) = \alpha T(X) + \beta T(Y)$, para todos os vetores $X, Y \in \mathbb{R}^n$ e todos os escalares α e β. Pois, se T é linear, então

$$T(\alpha X + \beta Y) = T(\alpha X) + T(\beta Y) = \alpha T(X) + \beta T(Y).$$

Por outro lado, se T é uma função tal que $T(\alpha X + \beta Y) = \alpha T(X) + \beta T(Y)$, para todos os vetores $X, Y \in \mathbb{R}^n$ e todos os escalares α e β, então fazendo $\alpha = 1, \beta = 1$ e depois $\beta = 0$ segue-se que T é uma transformação linear.

Se $T : \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear e $\mathcal{B} = \{V_1, \ldots, V_n\}$ é uma base de $\mathbb{R}^n$, então todo vetor $V \in \mathbb{R}^n$ pode ser escrito como combinação linear de vetores de $\mathcal{B}$, ou seja, existem escalares $\alpha_1, \ldots, \alpha_n$ tais que

$$V = \alpha_1 V_1 + \ldots + \alpha_n V_n = \sum_{i=1}^{n} \alpha_i V_i.$$

Então

$$T(V) = T\left(\sum_{i=1}^{n} \alpha_i V_i\right) = \sum_{i=1}^{n} T(\alpha_i V_i) = \sum_{i=1}^{n} \alpha_i T(V_i).$$

Por outro lado, se $U : \mathbb{R}^n \to \mathbb{R}^m$ é outra transformação linear tal que $U(V_i) = T(V_i)$ para $i = 1, \ldots, n$, então aplicando-se o raciocínio acima, temos que

$$U(V) = T(V), \quad \text{para todo } V \in \mathbb{R}^n.$$

Ou seja, $U = T$. Isto prova o seguinte teorema.

Teorema 6.1. *Uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ é totalmente caracterizada pelos seus valores em uma base de $\mathbb{R}^n$. Ou seja, se $\mathcal{B} = \{V_1, \ldots, V_n\}$ é uma base de $\mathbb{R}^n$ e uma função T está definida para valores em $\mathcal{B}$,*

$$T(V_i) = W_i, \quad \text{para } i = 1, \ldots, n.$$

Então, existe um única transformação linear definida em todo espaço $\mathbb{R}^n$, $T : \mathbb{R}^n \to \mathbb{R}^m$, tal que $T(V_i) = W_i$, para $i = 1, \ldots, n$.

Exemplo 6.7. Seja $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear definida na base canônica por (Figura 6.9)

$$\begin{aligned} R_\theta(E_1) &= \cos\theta\, E_1 + \operatorname{sen}\theta\, E_2 = (\cos\theta, \operatorname{sen}\theta) \\ R_\theta(E_2) &= -\operatorname{sen}\theta\, E_1 + \cos\theta\, E_2 = (-\operatorname{sen}\theta, \cos\theta). \end{aligned}$$

Assim, como $(x, y) = x(1, 0) + y(0, 1) = xE_1 + yE_2$, então

$$\begin{aligned} R_\theta \begin{bmatrix} x \\ y \end{bmatrix} &= xR_\theta(E_1) + yR_\theta(E_2) = x \begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix} + y \begin{bmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \end{aligned}$$

Se escrevemos $X = (x, y) = (r\cos\alpha, r\operatorname{sen}\alpha)$, então

$$\begin{aligned} R_\theta(X) &= R_\theta(r\cos\alpha, r\operatorname{sen}\alpha) = rR_\theta(\cos\alpha, \operatorname{sen}\alpha) \\ &= r(\cos\alpha\cos\theta - \operatorname{sen}\alpha\operatorname{sen}\theta, \cos\alpha\operatorname{sen}\theta + \operatorname{sen}\alpha\cos\theta) \\ &= r(\cos(\alpha+\theta), \operatorname{sen}(\alpha+\theta)) = (r\cos(\alpha+\theta), r\operatorname{sen}(\alpha+\theta)). \end{aligned}$$

Assim, segue-se da linearidade, que R_θ faz uma rotação de um ângulo θ em todo vetor $X = (x, y) \in \mathbb{R}^2$.

Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear tal que

$$T\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \ldots, \quad T\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Sejam $X = (x_1, \ldots, x_n)$ um vetor qualquer do $\mathbb{R}^n$ e

$$E_1 = (1, 0, \ldots, 0), E_2 = (0, 1, 0, \ldots, 0), \ldots, E_n = (0, \ldots, 0, 1).$$

Como $X = x_1 E_1 + \ldots + x_n E_n$, então

$$\begin{aligned} T(X) &= x_1 T(E_1) + \ldots + x_n T(E_n) = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\,X, \end{aligned}$$

em que as colunas de A são $T(E_1), \ldots, T(E_n)$, ou seja, $A = [\,T(E_1) \ldots T(E_n)\,]$, com $T(E_i)$, para $i = 1, \ldots, n$ escritos como matrizes colunas. Isto prova o seguinte resultado.

Proposição 6.2. *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Então, T é dada por*

$$T(X) = A\,X, \quad \textit{para todo } X \in \mathbb{R}^n,$$

em que $A = (a_{ij})_{m\times n} = [\, T(E_1) \ldots T(E_n)\,]$, *com* $T(E_i)$, *para* $i = 1, \ldots, n$, *escritos como matrizes colunas e*

$$E_1 = (1, 0, \ldots, 0), E_2 = (0, 1, 0, \ldots, 0), \ldots, E_n = (0, \ldots, 0, 1).$$

A matriz A *é chamada* ***matriz da transformação*** T ***(em relação às bases canônicas)***.

Exemplo 6.8. A matrizes de uma transformação linear pode ser obtida rapidamente aplicando-se a transformação nos vetores da base canônica. Por exemplo, as matrizes das projeções nos eixos x e y podem ser obtidas

$$\begin{bmatrix} P_x(E_1) & P_x(E_2) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} P_y(E_1) & P_y(E_2) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

as matrizes das reflexões em relação aos eixos x e y,

$$\begin{bmatrix} R_x(E_1) & R_x(E_2) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} R_y(E_1) & R_y(E_2) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$

as matrizes da projeção ortogonal e da reflexão em relação a uma reta $r : (x, y) = t(a, b)$, como

$$\begin{bmatrix} P_r(E_1) & P_r(E_2) \end{bmatrix} = \begin{bmatrix} \frac{a^2}{a^2+b^2} & \frac{ab}{a^2+b^2} \\ \frac{ab}{a^2+b^2} & \frac{b^2}{a^2+b^2} \end{bmatrix} \quad \begin{bmatrix} R_r(E_1) & R_r(E_2) \end{bmatrix} = \begin{bmatrix} \frac{a^2-b^2}{a^2+b^2} & \frac{2ab}{a^2+b^2} \\ \frac{2ab}{a^2+b^2} & \frac{b^2-a^2}{a^2+b^2} \end{bmatrix}$$

e a matriz da rotação de um ângulo θ,

$$\begin{bmatrix} R_\theta(E_1) & R_\theta(E_2) \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & cos\theta \end{bmatrix}.$$

Exercícios Numéricos (respostas na página 553)

6.1.1. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear para a qual sabemos que $T(1,1) = (2,-3)$ e $T(0,1) = (1,2)$.

(a) Determine $T(3,-2)$;

(b) Determine $T(a,b)$.

6.1.2. Determine a transformação $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que $T(1,1) = (3,2,1)$ e $T(0,-2) = (0,1,0)$. Encontre $T(1,0)$ e $T(0,1)$.

6.1.3. Determine expressões para as transformações lineares $P_{xy}, P_{yz}, P_{xz} : \mathbb{R}^3 \to \mathbb{R}^3$, que são projeções nos planos xy, yz e xz, respectivamente.

6.1.4. Considere o plano $\pi : x + 2y + 3z = 0$. Encontre expressões para as seguintes transformações lineares:

(a) A projeção ortogonal no plano π, $P_\pi : \mathbb{R}^3 \to \mathbb{R}^3$.

(b) A reflexão em relação ao plano π, $R_\pi : \mathbb{R}^3 \to \mathbb{R}^3$.

6.1.5. Determine expressões para as transformações lineares $R_{\pi/3,x}, R_{\pi/3,y}$ e $R_{\pi/3,z}$ que são rotações de $\pi/3$ em relação aos eixos x, y e z respectivamente.

6.1.6. Considere a reta $r : (x,y,z) = t(1,1,1)$.

(a) Seja $P_r : \mathbb{R}^3 \to \mathbb{R}^3$ a projeção ortogonal na reta r. Encontre uma expressão para $P_r(x,y,z)$.

(b) Seja $R_r : \mathbb{R}^3 \to \mathbb{R}^3$ a reflexão em relação à reta r. Encontre uma expressão para $R_r(x,y,z)$.

6.1.7. Existe uma única reflexão S do plano que transforma o ponto $(5,0)$ no ponto $(3,4)$. Determine a equação para o eixo da reflexão S. Verifique que ele passa pela origem. Calcule a matriz (em relação à base canônica de $\mathbb{R}^2$) da reflexão S.

Exercícios Teóricos

6.1.8. Considere o plano $\pi : ax + by + cz = 0$. Encontre expressões para as seguintes transformações lineares:

(a) A projeção ortogonal no plano π, $P_\pi : \mathbb{R}^3 \to \mathbb{R}^3$.

(b) A reflexão em relação ao plano π, $R_\pi : \mathbb{R}^3 \to \mathbb{R}^3$.

6.1.9. Determine expressões para as transformações lineares $R_{\theta,x}, R_{\theta,y}, R_{\theta,z} : \mathbb{R}^3 \to \mathbb{R}^3$, que são rotações de θ em relação aos eixos x, y e z, respectivamente.

6.1.10. Considere a reta $r : (x, y, z) = t(a, b, c)$. Encontre expressões para as seguintes transformações lineares:

(a) A projeção ortogonal na reta r, $P_r : \mathbb{R}^3 \to \mathbb{R}^3$.

(b) A reflexão em relação à reta r, $R_r : \mathbb{R}^3 \to \mathbb{R}^3$.

6.1.11. Seja $Y = (a, b, c) \in \mathbb{R}^3$. Determine a matriz do operador linear $T : \mathbb{R}^3 \to \mathbb{R}^3$, definido por

$$T(X) = T(x, y, z) = X \times Y = \left(\det \begin{bmatrix} y & z \\ b & c \end{bmatrix}, -\det \begin{bmatrix} x & z \\ a & c \end{bmatrix}, \det \begin{bmatrix} x & y \\ a & b \end{bmatrix}\right),$$

em relação à base canônica.

6.1.12. Seja c uma constante diferente de zero. Uma transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (cx, y)$, para todo $(x, y) \in \mathbb{R}^2$ é chamada **expansão ao longo do eixo** x se $c > 1$ e **contração ao longo do eixo** x se $0 < c < 1$. Uma transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x, cy)$, para todo $(x, y) \in \mathbb{R}^2$ é chamada **expansão ao longo do eixo** y se $c > 1$ e **contração ao longo do eixo** y se $0 < c < 1$.

Uma transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x + cy, y)$, para todo $(x, y) \in \mathbb{R}^2$ é chamada **cisalhamento ao longo do eixo** x. Uma transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x, y + cx)$, é chamada **cisalhamento ao longo do eixo** y.

(a) Mostre que a matriz elementar que corresponde a trocar duas linhas é a matriz de uma reflexão em relação à reta $y = x$.

(b) Mostre que a matriz elementar que corresponde a multiplicar uma linha por um escalar não nulo é a matriz de uma expansão, ou a matriz de uma contração, ou a matriz de uma reflexão em relação a um dos eixos coordenados, ou um produto de uma matriz de uma reflexão em relação a um dos eixos coordenados por uma matriz de uma expansão ou contração.

(c) Mostre que a matriz elementar que corresponde a somar a uma linha um múltiplo escalar de outra é a matriz de um cisalhamento ao longo de um dos eixos coordenados.

(Sugestão: veja no Exemplo 1.17 na página 53 as matrizes elementares 2×2 e compare com as matrizes das transformações definidas acima.)

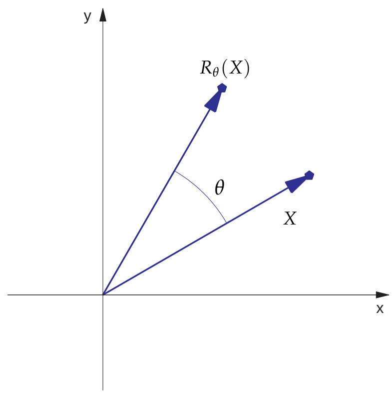


Figura 6.8. Transformação rotação de um ângulo θ

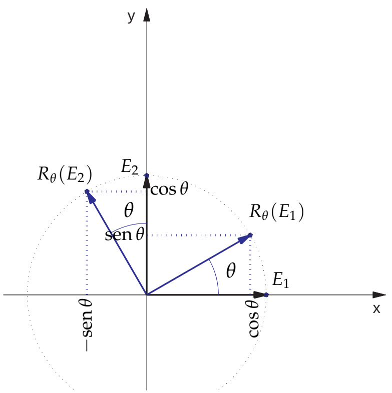


Figura 6.9. Transformação rotação sobre os vetores E_1 e E_2

6.2 A Imagem e o Núcleo

Definição 6.2. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear.

(a) O **núcleo** de T é definido pelo conjunto

$$\mathcal{N}(T) = \{X \in \mathbb{R}^n \mid T(X) = \bar{0}\}.$$

(b) A **imagem** de T é definida pelo conjunto

$$\mathcal{I}(T) = \{Y \in \mathbb{R}^m \mid Y = T(X), \text{para algum } X \in \mathbb{R}^n\}$$

(c) A dimensão do núcleo de T é chamada de **nulidade** de T e a dimensão da imagem de T é chamada **posto de** T.

Exemplo 6.9. Sejam $O : \mathbb{R}^n \to \mathbb{R}^m$ a transformação linear nula e $I : \mathbb{R}^n \to \mathbb{R}^n$ a transformação identidade. Então

$$\mathcal{N}(O) = \mathbb{R}^n, \quad \mathcal{I}(O) = \{\bar{0}\},$$

$$\mathcal{N}(I) = \{\bar{0}\}, \quad \mathcal{I}(I) = \mathbb{R}^n.$$

Teorema 6.3. *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. O núcleo, $\mathcal{N}(T)$, e a imagem, $\mathcal{I}(T)$ são subespaços de $\mathbb{R}^n$ e de $\mathbb{R}^m$, respectivamente.*

Demonstração. Para mostrar que um conjunto é um subespaço precisamos mostrar as propriedades (a) e (b) da página 227, ou seja, que com relação às operações de soma de vetores e multiplicação por escalar podemos "viver" nele sem termos que sair. Vamos mostrar em primeiro lugar que o núcleo de T é um subespaço.

(a) Se $X_1, X_2 \in \mathcal{N}(T)$, então $T(X_1) = T(X_2) = \bar{0}$. Logo,

$$T(X_1 + X_2) = T(X_1) + T(X_2) = \bar{0} + \bar{0} = \bar{0},$$

ou seja, $X_1 + X_2 \in \mathcal{N}(T)$;

(b) Se $X \in \mathcal{N}(T)$ e α é um escalar, então $T(\alpha X) = \alpha T(X) = \alpha \bar{0}$. Logo, $\alpha X \in \mathcal{N}(T)$;

Vamos mostrar, agora, que a imagem de T é um subespaço.

(a) Se $Y_1, Y_2 \in \mathcal{I}(T)$, então existem $X_1, X_2 \in \mathbb{R}^n$ tais que $T(X_1) = Y_1$ e $T(X_2) = Y_2$. Seja $X = X_1 + X_2$. Então, $T(X) = T(X_1 + X_2) = T(X_1) + T(X_2) = Y_1 + Y_2$. Logo, $Y_1 + Y_2 \in \mathcal{I}(T)$.

(b) Se $Y \in \mathcal{I}(T)$ e α é um escalar, então existe $X \in \mathbb{R}^n$ tal que $T(X) = Y$. Como T é linear, então $T(\alpha X) = \alpha T(X) = \alpha Y$. Logo, $\alpha Y \in \mathcal{I}(T)$.

■

Exemplo 6.10. A imagem de uma transformação linear de $\mathbb{R}^n$ em $\mathbb{R}$ não nula, $f : \mathbb{R}^n \to \mathbb{R}$, é o conjunto $\mathbb{R}$, pois $\{\bar{0}\}$ e $\mathbb{R}$ são os únicos subespaços do espaço vetorial $\mathbb{R}$.

Proposição 6.4. *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Se $\{V_1, \ldots, V_n\}$ é uma base de $\mathbb{R}^n$, então a imagem de T é gerada por $T(V_1), \ldots, T(V_n)$.*

Demonstração. Seja $W \in \mathcal{I}(T)$. Então, existe $V \in \mathbb{R}^n$ tal que $T(V) = W$. Como $\{V_1, \ldots, V_n\}$ é base de $\mathbb{R}^n$, existem escalares $\alpha_1, \ldots, \alpha_n$ tais que $V = \alpha_1 V_1 + \ldots + \alpha_n V_n$. Assim,

$$W = T(V) = T\left(\sum_{i=1}^{n} \alpha_i V_i\right) = \sum_{i=1}^{n} \alpha_i T(V_i).$$

Ou seja, $T(V_1), \ldots, T(V_n)$ geram $\mathcal{I}(T)$. ■

Exemplo 6.11. Vamos considerar as projeções nos eixos x e y (Figuras 6.2 e 6.3 na página 347)

$$P_x \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad P_y \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Geometricamente vemos que o núcleo de P_x é o eixo y, o núcleo de P_y é o eixo x, que são os pontos que são levados pelas transformações na origem. Vemos também que a imagem de P_x é o eixo x, pois todos os pontos são levados por P_x no eixo x. Analogamente, a imagem de P_y é o eixo y.

Exemplo 6.12. Vamos considerar as reflexões em relação aos eixos x e y (Figuras 6.4 e 6.5 na página 349)

$$R_x \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad R_y \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

Geometricamente vemos que o núcleo de R_x e o núcleo de R_y são iguais a origem, pois é o único ponto que é levado na origem pelas transformações. Vemos também que as imagens de R_x e de R_y são iguais a $\mathbb{R}^2$, pois todo ponto (x,y) é imagem do ponto (x,y) refletido pelas respectivas transformações.

Exemplo 6.13. Consideremos a projeção ortogonal e a reflexão em relação a uma reta

$$r : (x,y) = t(a,b)$$

(Figuras 6.6 e 6.7 na página 349)

$$P_r \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a^2}{a^2+b^2} & \frac{ab}{a^2+b^2} \\ \frac{ab}{a^2+b^2} & \frac{b^2}{a^2+b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad R_r \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a^2-b^2}{a^2+b^2} & \frac{2ab}{a^2+b^2} \\ \frac{2ab}{a^2+b^2} & \frac{b^2-a^2}{a^2+b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Geometricamente vemos que o núcleo de P_r é a reta $s : (x,y) = t(-b,a)$, perpendicular à reta r que passa pela origem, pois os pontos sobre a reta s são exatamente os pontos que são levados por P_r na origem. Vemos também que a imagem de P_r é a própria reta r, pois todos os pontos do $\mathbb{R}^2$ são levados na reta r. Geometricamente vemos que o núcleo de R_r é a origem, pois é o único ponto do $\mathbb{R}^2$ que é levado na origem. Vemos também que a imagem de R_r é o $\mathbb{R}^2$, pois todo ponto (x,y) é a imagem do ponto (x,y) refletido por r.

Exemplo 6.14. Geometricamente vemos que a rotação de um ângulo θ (Figura 6.8 na página 358)

$$R_\theta \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

tem núcleo igual a origem, pois é o único ponto que é levado na origem por R_θ. Vemos também que a imagem de R_θ é igual ao $\mathbb{R}^2$, pois todo ponto (x,y) é a imagem do ponto (x,y) girado de $-\theta$.

Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Pela Proposição 6.2 na página 353 a transformação T é dada por

$$T(X) = AX, \quad \text{para todo } X \in \mathbb{R}^n,$$

onde $A = (a_{ij})_{m\times n} = [\, T(E_1) \ldots T(E_n)\,]$, com $T(E_i)$, para $i = 1, \ldots, n$ escritos como matrizes colunas e

$$E_1 = (1, 0, \ldots, 0), E_2 = (0, 1, 0, \ldots, 0), \ldots, E_n = (0, \ldots, 0, 1).$$

Assim, o núcleo de T é o espaço solução do sistema homogêneo $AX = \bar{0}$ e a imagem de T é o subespaço gerado pelas colunas de A, ou seja, é o espaço coluna de A.

6.2.1 Injetividade e Sobrejetividade

Dizemos que uma função $f : A \to B$ é **sobrejetiva** se,
para todo $b \in B$ existe $a \in A$ tal que $f(a) = b$.
Ou seja, se a imagem de f é igual a B. No caso em que f é uma transformação linear, obtemos como consequência da Proposição 6.4 o seguinte resultado.

Corolário 6.5. *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Seja $\{V_1, \ldots, V_n\}$ base de $\mathbb{R}^n$. T é sobrejetiva se, e somente se, $T(V_1), \ldots, T(V_n)$ geram o $\mathbb{R}^m$.*

Exemplo 6.15. Toda transformação linear de $\mathbb{R}^n$ em $\mathbb{R}$ não nula, $f : \mathbb{R}^n \to \mathbb{R}$ é sobrejetiva, pois $\{\bar{0}\}$ e $\mathbb{R}$ são os únicos subespaços do espaço vetorial $\mathbb{R}$.

Exemplo 6.16. A reflexão em relação a uma reta que passa pela origem e a rotação são sobrejetivas enquanto a projeção ortogonal em uma reta que passa pela origem não é sobrejetiva (Exemplos 6.13 e 6.14 na página 362).

Dizemos que uma função $f : A \to B$ é **injetiva**, se $f(x) = f(y)$ implica que $x = y$.

Teorema 6.6. *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Então, T é injetiva se, e somente se,*

$$\mathcal{N}(T) = \{\bar{0}\}.$$

Demonstração. Suponha que T é injetiva. Seja $X \in \mathcal{N}(T)$. Então, como T é injetiva, $T(X) = T(\bar{0})$ implica que $X = \bar{0}$.

Agora, suponha que $\mathcal{N}(T) = \{\bar{0}\}$. Se $T(X) = T(Y)$, então $T(X - Y) = \bar{0}$, ou seja,

$$X - Y \in \mathcal{N}(T).$$

Como, $\mathcal{N}(T) = \{\bar{0}\}$, então $X - Y = \bar{0}$, ou seja, $X = Y$. ■

Exemplo 6.17. A reflexão em relação a uma reta que passa pela origem e a rotação são injetivas enquanto a projeção ortogonal em uma reta que passa pela origem **não** é injetiva (Exemplos 6.13 e 6.14 na página 362).

Teorema 6.7 (da Dimensão do Núcleo e da Imagem). *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Então a soma da dimensão do núcleo de T com a dimensão da imagem de T é igual a dimensão de $\mathbb{R}^n$, ou seja,*

$$\dim(\mathcal{N}(T)) + \dim(\mathcal{I}(T)) = \dim(\mathbb{R}^n) = n.$$

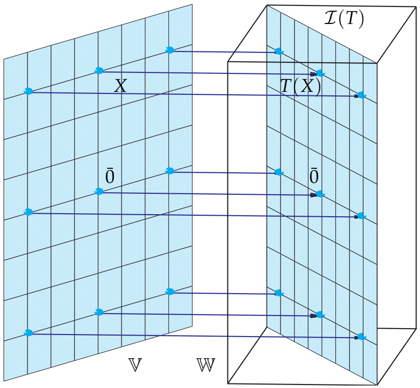


Figura 6.10. $\mathcal{N}(T) = \{\bar{0}\}$

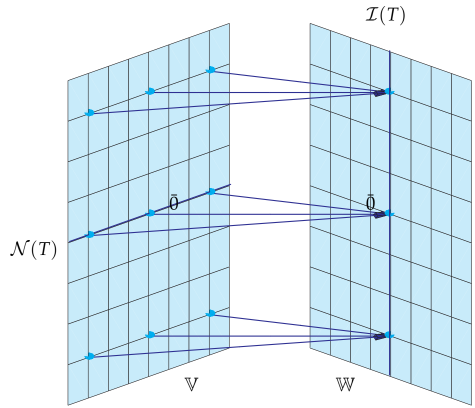


Figura 6.11. $\mathcal{N}(T) \neq \{\bar{0}\}$

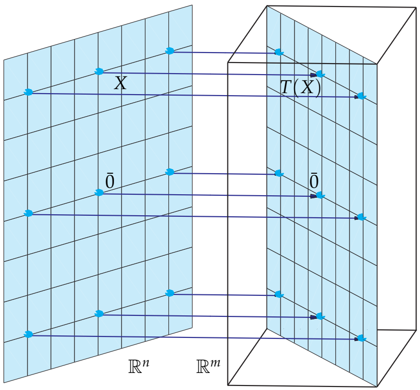


Figura 6.12. Transformação linear injetiva ($\mathcal{N}(T) = \{\bar{0}\}$)

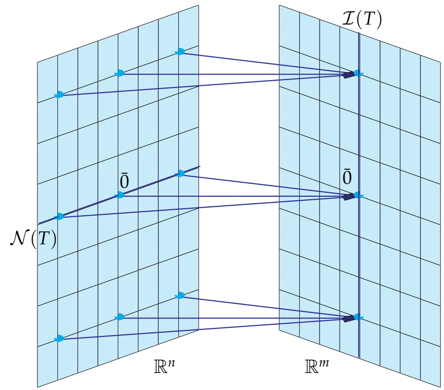


Figura 6.13. Transformação linear não injetiva ($\mathcal{N}(T) \neq \{\bar{0}\}$)

Demonstração. Vamos supor que $1 \leq \dim(\mathcal{N}(T)) < n$. Sejam $V_1, \ldots, V_p$ vetores de $\mathbb{R}^n$, que formam uma base para o núcleo de T. Vamos estendê-la a uma base de $\mathbb{R}^n$. Sejam $V_{p+1}, \ldots, V_n$ vetores de $\mathbb{R}^n$ tais que $V_1, \ldots, V_p, V_{p+1}, \ldots, V_n$ formam uma base de $\mathbb{R}^n$. Vamos mostrar que $T(V_{p+1}), \ldots, T(V_n)$ formam uma base da imagem de T. Para isso, precisamos mostrar que eles geram a imagem de T e que são L.I.

Pela Proposição 6.4 na página 361, $T(V_{p+1}), \ldots, T(V_n)$ geram a imagem de T, pois $T(V_1) = \ldots = T(V_p) = \bar{0}$. Vamos mostrar que $T(V_{p+1}), \ldots, T(V_n)$ são linearmente independentes. Considere a combinação linear nula:

$$x_{p+1}T(V_{p+1}) + \ldots + x_n T(V_n) = \bar{0}.$$

Pela linearidade de T segue-se que

$$T(x_{p+1}V_{p+1} + \ldots + x_n V_n) = \bar{0}.$$

Mas isto implica que $x_{p+1}V_{p+1} + \ldots + x_n V_n \in \mathcal{N}(T)$. Assim, existem escalares $y_1, \ldots, y_p$ tais que $x_{p+1}V_{p+1} + \ldots + x_n V_n = y_1 V_1 + \ldots + y_p V_p$. De onde segue-se que

$$y_1 V_1 + \ldots + y_p V_p - x_{p+1}V_{p+1} - \ldots - x_n V_n = \bar{0}.$$

Como $V_1, \ldots, V_n$ é base de $\mathbb{R}^n$, então $y_1 = \ldots = y_p = x_{p+1} = \ldots = x_n = 0$, ou seja, $T(V_{p+1}), \ldots, T(V_n)$ são L.I.

Portanto, a dimensão da imagem de T é igual a diferença entre a dimensão de $\mathbb{R}^n$ e a dimensão do núcleo de A, de onde segue-se o resultado. ■

Em geral, uma transformação linear pode ser injetiva sem ser sobrejetiva e sobrejetiva sem ser injetiva. Entretanto, se a dimensão do domínio for igual a dimensão do contradomínio temos o seguinte resultado, que é uma consequência imediata do Teorema 6.7 da Dimensão do Núcleo e da Imagem.

Corolário 6.8. *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Suponha que $n = m$.*

(a) Se T é sobrejetiva, então T é injetiva.

(b) Se T é injetiva, então T é sobrejetiva.

Se uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ é injetiva e sobrejetiva, então T é chamada **isomorfismo**.

Exercícios Numéricos (respostas na página 555)

6.2.1. Seja P a transformação de $\mathbb{R}^3$ em $\mathbb{R}^3$, definida por $P(x, y, z) = (x, y, 0)$. Se a imagem de uma reta r, por P, é um ponto, então quais são as equações paramétricas da reta r?

6.2.2. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ uma transformação linear dada por $T(x, y, z) = (z, x - y, -z)$.

(a) Encontre uma base para o núcleo de T.

(b) Encontre uma base para a imagem de T.

(c) Descreva geometricamente o núcleo e a imagem de T.

6.2.3. Seja $T : \mathbb{R}^n \to \mathbb{R}^5$ uma transformação linear.

(a) Se T é sobrejetiva e $\dim(\mathcal{N}(T)) = 2$, qual o valor de n?

(b) Se T é sobrejetiva e injetiva, qual o valor de n?

6.2.4. Dê exemplos de transformações lineares $T : \mathbb{R}^3 \to \mathbb{R}^3$ tais que

(a) $\mathcal{N}(T) = \{(x, y, z) \in \mathbb{R}^3 \,|\, z = -x\}$,

(b) $\mathcal{I}(T) = \{(x, y, z) \in \mathbb{R}^3 \,|\, x = y\}$.

6.2.5. Dê um exemplo de uma transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $\mathcal{N}(T) = \mathcal{I}(T)$.

Exercícios Teóricos

6.2.6. Determine o núcleo e a imagem do operador linear definido no Exercício 6.1.11 na página 356.

6.2.7. Considere o plano $\pi : ax + by + cz = 0$.

(a) Determine o núcleo e a imagem da projeção ortogonal no plano π, P_π. Responda se P_π é sobrejetiva e se é injetiva.

(b) Determine o núcleo e a imagem da reflexão em relação ao plano π, R_π. Responda se R_π é sobrejetiva e se é injetiva.

6.2.8. Considere a reta $r : (x, y, z) = t(a, b, c)$.

(a) Determine o núcleo e a imagem da projeção ortogonal na reta r, P_r. Responda se P_r é sobrejetiva e se é injetiva.

(b) Determine o núcleo e a imagem da reflexão em relação à reta r, R_r. Responda se R_r é sobrejetiva e se é injetiva.

6.2.9. Seja $f : \mathbb{R}^3 \to \mathbb{R}$ uma transformação linear.

(a) Mostre que existem escalares a, b, c tais que $f(x, y, z) = ax + by + cz$.

(b) Descreva geometricamente todas as possibilidades para o núcleo de f.

6.2.10. Sejam $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear e $\mathcal{B} = \{V_1, \ldots, V_n\}$ um conjunto de vetores de $\mathbb{R}^n$. Mostre que se $\mathcal{C} = \{T(V_1), \ldots, T(V_n)\}$ é L.I., então $\mathcal{B}$ também o é.

6.2.11. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Mostre que T é injetiva se, e somente se, $\dim(\mathcal{I}(T)) = n$.

6.2.12. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Mostre que T é injetiva se, e somente se, a imagem de todo conjunto de vetores linearmente independente é um conjunto de vetores linearmente independente.

6.2.13. Sejam $T : \mathbb{R}^n \to \mathbb{R}^n$ uma transformação linear. Mostre que T é injetiva se, e somente se, a imagem por T de uma base de $\mathbb{R}^n$ é uma base de $\mathbb{R}^n$.

6.3 Composição de Transformações Lineares

Sejam $T : \mathbb{R}^n \to \mathbb{R}^p$ e $S : \mathbb{R}^p \to \mathbb{R}^m$ transformações lineares. A **composição de S com T**, denotada por ST é a função de $\mathbb{R}^n$ em $\mathbb{R}^m$ definida por

$$(ST)(X) = S(T(X)), \quad \text{para todo } X \in \mathbb{R}^n.$$

Proposição 6.9. *Se $T : \mathbb{R}^n \to \mathbb{R}^p$ e $S : \mathbb{R}^p \to \mathbb{R}^m$ são transformações lineares, então a composição $ST : \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear.*

Demonstração. Sejam $X_1, X_2 \in \mathbb{R}^n$ e α, β escalares.

$$\begin{aligned} (ST)(\alpha X_1 + \beta X_2) &= S(T(\alpha X_1 + \beta X_2)) = S(\alpha T(X_1) + \beta T(X_2)) \\ &= \alpha S(T(X_1)) + \beta S(T(X_2)) = \alpha (ST)(X_1) + \beta (ST)(X_2) \end{aligned}$$

■

6.3.1 Matriz de uma Transformação Linear

Definição 6.3. Seja $\mathcal{B} = \{V_1, \ldots, V_n\}$ uma base de $\mathbb{R}^n$. Todo vetor $V \in \mathbb{R}^n$ se escreve de maneira única como combinação linear de $V_1, \ldots, V_n$ (Teorema 4.3 na página 246), ou seja, existem escalares $\alpha_1, \ldots, \alpha_n$ tais que

$$V = \alpha_1 V_1 + \ldots + \alpha_n V_n.$$

Definimos o **vetor de coordenadas de** V **em relação à base (ordenada)** $\mathcal{B} = \{V_1, \ldots, V_n\}$, por

$$[V]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

As coordenadas de um vetor V em relação à base $\mathcal{B}$ são os escalares que aparecem quando escrevemos V como combinação linear dos vetores da base $\mathcal{B}$. Por exemplo, para os vetores da base $\mathcal{B}$, $[V_1]_\mathcal{B} = E_1, [V_2]_\mathcal{B} = E_2, \ldots, [V_n]_\mathcal{B} = E_n$, em que $E_1, \ldots, E_n$ são os vetores da base canônica do $\mathbb{R}^n$. Pois, $V_i = 0V_1 + \ldots + 1V_i + \ldots + 0V_n$, para $i = 1, \ldots, n$.

Sejam $\mathcal{B} = \{V_1, \ldots, V_n\}$ e $\mathcal{C} = \{W_1, \ldots, W_m\}$ bases de $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Sejam

$$[V]_\mathcal{B} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad [T(V_1)]_\mathcal{C} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad [T(V_2)]_\mathcal{C} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \ldots, \quad [T(V_n)]_\mathcal{C} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Então,

$$\begin{aligned} T(V) &= T(x_1V_1 + \ldots + x_nV_n) = x_1T(V_1) + \ldots + x_nT(V_n) \\ &= x_1(a_{11}W_1 + \ldots a_{m1}W_m) + \ldots + x_n(a_{1n}W_1 + \ldots + a_{mn}W_m) \\ &= (x_1a_{11} + \ldots + x_na_{1n})W_1 + \ldots + (x_1a_{m1} + \ldots + x_na_{mn})W_m. \end{aligned}$$

Como escrevemos o vetor $T(V)$ como combinação linear dos vetores da base $\mathcal{C}$, então os escalares são as coordenadas de $T(V)$ em relação à base $\mathcal{C}$, ou seja,

$$[T(V)]_\mathcal{C} = \begin{bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ a_{21}x_1 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\,[V]_\mathcal{B},$$

em que $A = [\ [T(V_1)]_\mathcal{C}\ \ldots\ [T(V_n)]_\mathcal{C}\ ]$. Esta matriz é chamada **matriz da transformação linear T em relação às bases $\mathcal{B}$ e $\mathcal{C}$** e é denotada por $[T]_\mathcal{B}^\mathcal{C}$, ou seja,

$$[T]_\mathcal{B}^\mathcal{C} = [\ [T(V_1)]_\mathcal{C}\ \ldots\ [T(V_n)]_\mathcal{C}\ ].$$

Isto prova o seguinte resultado, que é uma generalização da Proposição 6.2 na página 353.

Teorema 6.10. *Sejam $\mathcal{B} = \{V_1, \ldots, V_n\}$ e $\mathcal{C} = \{W_1, \ldots, W_m\}$ bases de $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Então, a matriz $m \times n$*

$$[T]_\mathcal{B}^\mathcal{C} = [\ [T(V_1)]_\mathcal{C}\ \ldots\ [T(V_n)]_\mathcal{C}\ ],$$

é tal que

$$[T(V)]_\mathcal{C} = [T]_\mathcal{B}^\mathcal{C}\,[V]_\mathcal{B}, \quad \textit{para todo vetor } V \in \mathbb{R}^n.$$

Aqui $[V]_\mathcal{B}$ é o vetor de coordenadas de V em relação à base $\mathcal{B}$, $[T(V)]_\mathcal{C}$ é o vetor de coordenadas de $T(V)$ em relação à base $\mathcal{C}$ e a matriz $[T]_\mathcal{B}^\mathcal{C} = [\ [T(V_1)]_\mathcal{C}\ \ldots\ [T(V_n)]_\mathcal{C}\ ]$ é a matriz da transformação linear T em relação às bases $\mathcal{B}$ e $\mathcal{C}$.

Quando a transformação linear é a transformação identidade $I : \mathbb{R}^n \to \mathbb{R}^n$, definida por $I(X) = X$, para todo $X \in \mathbb{R}^n$, então aplicando o resultado anterior (Teorema 6.10) a esta transformação, obtemos uma relação entre os vetores de coordenadas de um vetor X em relação à duas bases.

$$[X]_{\mathcal{C}} = [I(X)]_{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}[X]_{\mathcal{B}} = P[X]_{\mathcal{B}},$$

em que $P = [\,[V_1]_{\mathcal{C}} \ldots [V_n]_{\mathcal{C}}\,]$ é chamada **matriz mudança de base de $\mathcal{B}$ para $\mathcal{C}$.**

Corolário 6.11. *Sejam $\mathcal{B} = \{V_1, \ldots, V_n\}$ e $\mathcal{C} = \{W_1, \ldots, W_n\}$ bases de $\mathbb{R}^n$. Então, Para todo $X \in \mathbb{R}^n$,*

$$[X]_{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}[X]_{\mathcal{B}},$$

em que $P = [I]_{\mathcal{B}}^{\mathcal{C}} = [\,[V_1]_{\mathcal{C}} \ldots [V_n]_{\mathcal{C}}\,]$ é a matriz da transformação linear identidade $I : \mathbb{R}^n \to \mathbb{R}^n$ em relação às bases $\mathcal{C}$ e $\mathcal{B}$.

Exemplo 6.18. Sejam

$$\mathcal{B} = \{E_1 = (1,0), E_2 = (0,1)\}$$

e

$$\mathcal{C} = \{V_1 = (1,1), V_2 = (1,-1)\}$$

bases do $\mathbb{R}^2$. Como $\mathcal{B}$ é a base canônica, temos que $[(x,y)]_{\mathcal{B}} = \begin{bmatrix} x \\ y \end{bmatrix}$. Vamos encontrar $[(x,y)]_{\mathcal{C}}$.

$$[(x,y)]_{\mathcal{C}} = [I_{\mathbb{R}^2}]_{\mathcal{B}}^{\mathcal{C}}[(x,y)]_{\mathcal{B}}$$

Para determinarmos $[I_{\mathbb{R}^2}]_{\mathcal{B}}^{\mathcal{C}} = [\ [E_1]_{\mathcal{C}}\ [E_2]_{\mathcal{C}}\ ]$ diretamente precisamos saber escrever E_1 e E_2 em termos da base $\mathcal{C}$. Para isto precisamos resolver as equações:

$$E_1 = x_1 V_1 + y_1 V_2 \quad \text{e} \quad E_2 = x_2 V_1 + y_2 V_2.$$

Temos que resolver dois sistemas lineares que têm a mesma matriz $A = [\ V_1\ V_2\ ]$. Como a matriz A é invertível e é fácil encontrar a inversa de uma matriz 2×2 (ver por exemplo Exemplo 2.17 na página 117), podemos obter as soluções dos sistemas como $A^{-1}E_1$ e $A^{-1}E_2$. Como

$$A^{-1} = \left(\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix},$$

então

$$[I]_{\mathcal{B}}^{\mathcal{C}} = [\ [E_1]_{\mathcal{C}}\ [E_2]_{\mathcal{C}}\ ] = [\ A^{-1}E_1\ \ A^{-1}E_2\ ] = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}.$$

Portanto

$$[(x,y)]_{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}[(x,y)]_{\mathcal{B}} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (x+y)/2 \\ (x-y)/2 \end{bmatrix}.$$

Teorema 6.12. *Sejam $T : \mathbb{R}^n \to \mathbb{R}^p$ e $S : \mathbb{R}^p \to \mathbb{R}^m$ transformações lineares. Sejam $\mathcal{B} = \{V_1, \ldots, V_n\}$, $\mathcal{C} = \{U_1, \ldots, U_p\}$ e $\mathcal{D} = \{W_1, \ldots, W_m\}$ bases de $\mathbb{R}^n, \mathbb{R}^p$ e $\mathbb{R}^m$ respectivamente. Então,*

$$[ST]_{\mathcal{B}}^{\mathcal{D}} = [S]_{\mathcal{C}}^{\mathcal{D}}[T]_{\mathcal{B}}^{\mathcal{C}}.$$

Ou seja, a matriz da composição de duas transformações lineares é o produto das matrizes das transformações lineares.

Demonstração. Sejam $A = [S]_{\mathcal{C}}^{\mathcal{D}}$, $B = [T]_{\mathcal{B}}^{\mathcal{C}}$ e $C = [ST]_{\mathcal{B}}^{\mathcal{D}}$. Vamos mostrar que $C = AB$.

$$\begin{aligned}(ST)(V_j) &= S(T(V_j)) = S\left(\sum_{k=1}^{p} b_{kj}U_k\right) = \sum_{k=1}^{p} b_{kj}S(U_k) \\ &= \sum_{k=1}^{p} b_{kj}\sum_{i=1}^{m} a_{ik}W_i = \sum_{k=1}^{p}\sum_{i=1}^{m} a_{ik}b_{kj}W_i = \sum_{i=1}^{m}\left(\sum_{k=1}^{p} a_{ik}b_{kj}\right)W_i\end{aligned}$$

Mas, por definição da matriz de uma transformação linear

$$(ST)(V_j) = \sum_{i=1}^{m} c_{ij}W_i.$$

Como os vetores $W_1, \ldots, W_m$ são L.I., então $c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}$, ou seja, $C = AB$, como queríamos provar. ■

6.3.2 Invertibilidade

Dizemos que uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ é **invertível** se, existe uma função $U : \mathbb{R}^m \to \mathbb{R}^n$ tal que $TU = I_{\mathbb{R}^m}$ e $UT = I_{\mathbb{R}^n}$. A função U é única (verifique!) e denotada por T^{-1}.

Proposição 6.13. *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear invertível. Então, $T^{-1} : \mathbb{R}^m \to \mathbb{R}^n$ é também uma transformação linear.*

Demonstração. Sejam $Y_1, Y_2 \in \mathbb{R}^m$ e α, β escalares. Sejam $X_1 = T^{-1}(Y_1)$ e $X_2 = T^{-1}(Y_2)$. Então,

$$\begin{aligned} T^{-1}(\alpha Y_1 + \beta Y_2) &= T^{-1}(\alpha T(X_1) + \beta T(X_2) = T^{-1}(T(\alpha X_1 + \beta X_2)) \\ &= \alpha X_1 + \beta X_2 = \alpha T^{-1}(Y_1) + \beta T^{-1}(Y_2) \end{aligned}$$

o que prova que T^{-1} é uma transformação linear. ■

Lembramos que uma função $f : A \to B$ é invertível se, e somente se, é injetiva e sobrejetiva.

Teorema 6.14. *Sejam $\mathcal{B} = \{V_1, \ldots, V_n\}$ e $\mathcal{C} = \{W_1, \ldots, W_m\}$ bases dos espaços vetoriais $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente. Uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ é invertível se, e somente se, $[T]_{\mathcal{B}}^{\mathcal{C}}$ é invertível. Além disso, se T é invertível, então $[T^{-1}]_{\mathcal{C}}^{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{C}})^{-1}$.*

Demonstração. Suponha, em primeiro lugar, que T é invertível. Então T é injetiva e sobrejetiva, o que implica, pelo Teorema da Dimensão do Núcleo e da Imagem 6.7 na página 365, que $n = m$. Além disso, existe uma transformação linear, T^{-1}, tal que $TT^{-1} = I_{\mathbb{R}^m}$ e $T^{-1}T = I_{\mathbb{R}^n}$. Assim,

$$I_n = [I_{\mathbb{R}^n}]_{\mathcal{B}}^{\mathcal{B}} = [T^{-1}T]_{\mathcal{B}}^{\mathcal{B}} = [T^{-1}]_{\mathcal{C}}^{\mathcal{B}}[T]_{\mathcal{B}}^{\mathcal{C}}.$$

Portanto, a matriz $[T]_{\mathcal{B}}^{\mathcal{C}}$ é invertível e $([T]_{\mathcal{B}}^{\mathcal{C}})^{-1} = [T^{-1}]_{\mathcal{C}}^{\mathcal{B}}$.

Suponha, agora, que $A = [T]_{\mathcal{B}}^{\mathcal{C}}$ é uma matriz invertível. Então, A é uma matriz quadrada e $n = m$. Vamos mostrar que $\mathcal{N}(T) = \{\bar{0}\}$. Seja $V \in \mathcal{N}(T)$. Então, $[T(V)]_{\mathcal{C}} = A[V]_{\mathcal{B}} = \bar{0}$. Como A é invertível, então $[V]_{\mathcal{B}} = \bar{0}$. O que implica que

$V = \bar{0}$. Assim T é injetiva (Teorema 6.6 na página 364) e como $n = m$, então pelo Corolário 6.8 na página 368 segue-se que T é também sobrejetiva e portanto invertível. ■

Exemplo 6.19. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(x, y, z) = (x + y + 2z, y + 2z, z)$, para todo $(x, y, z) \in \mathbb{R}^3$. Vamos verificar se T é invertível. Seja $\mathcal{B} = \{E_1, E_2, E_3\}$. Vamos determinar a matriz de T em relação a $\mathcal{B}$. Para isto, vamos escrever o resultado da aplicação T em cada elemento de $\mathcal{B}$ como combinação linear dos elementos de $\mathcal{B}$.

$$\begin{array}{rclcl} T(E_1) &=& E_1 &=& 1E_1 + 0E_2 + 0E_3 \\ T(E_2) &=& E_1 + E_2 &=& 1E_1 + 1E_2 + 0E_3 \\ T(E_3) &=& 2E_1 + 2E_2 + E_3 &=& 2E_1 + 2E_2 + 1E_3 \end{array}$$

Assim, a matriz de T em relação a $\mathcal{B}$ é

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \left[\ [T(E_1)]_{\mathcal{B}} \quad [T(E_2)]_{\mathcal{B}} \quad [T(E_3)]_{\mathcal{B}}\ \right] = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Esta matriz é invertível (verifique!) e assim pelo Teorema 6.14 a transformação linear T é invertível e

$$[T^{-1}]_{\mathcal{B}}^{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}})^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Vamos determinar uma expressão para T^{-1}. Pelo Teorema 6.10 na página 374, temos que

$$[T^{-1}(x, y, z)]_{\mathcal{B}} = [T^{-1}]_{\mathcal{B}}^{\mathcal{B}}[(x, y, z)]_{\mathcal{B}} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y \\ y - 2z \\ z \end{bmatrix}.$$

Portanto, $T^{-1}(x, y, z) = (x - y, y - 2z, z)$.

Quando a transformação linear é a transformação identidade

$$I : \mathbb{R}^n \to \mathbb{R}^n, \text{definida por } I(X) = X, \text{ para todo } X \in \mathbb{R}^n,$$

então aplicando o resultado anterior (Teorema 6.14) a esta transformação, obtemos o seguinte resultado.

Corolário 6.15. *Sejam $\mathcal{B} = \{V_1, \ldots, V_n\}$ e $\mathcal{C} = \{W_1, \ldots, W_n\}$ bases de $\mathbb{R}^n$. A matriz mudança de base $P = [I]_{\mathcal{B}}^{\mathcal{C}}$ é invertível e*

$$P^{-1} = ([I]_{\mathcal{B}}^{\mathcal{C}})^{-1} = [I]_{\mathcal{C}}^{\mathcal{B}}$$

Exemplo 6.20. Vamos determinar a expressão da transformação linear que faz uma rotação de um ângulo θ no sentido anti-horário em torno de um eixo que passa pela origem e tem a direção e o sentido dados por um vetor **unitário** $U = (a, b, c)$. Seja $\mathcal{C} = \{U_1, U_2, U_3\}$, em que

$$\begin{aligned}
U_1 &= U = (a, b, c) \\
U_2 &= (-\frac{b}{\sqrt{a^2+b^2}}, \frac{a}{\sqrt{a^2+b^2}}, 0) \\
U_3 &= (-\frac{ac}{\sqrt{a^2+b^2}}, -\frac{bc}{\sqrt{a^2+b^2}}, \sqrt{a^2+b^2})
\end{aligned}$$

É fácil verificar que esta é uma base ortonormal, ou seja, uma base em que os seus vetores são unitários mutuamente ortogonais. Além disso, temos que

$$\begin{aligned}
R_\theta(U_1) &= U_1 = (a, b, c) \\
R_\theta(U_2) &= \cos\theta\, U_2 + \operatorname{sen}\theta\, U_3 = (\frac{-b\cos\theta - ac\operatorname{sen}\theta}{\sqrt{a^2+b^2}}, \frac{a\cos\theta - bc\operatorname{sen}\theta}{\sqrt{a^2+b^2}}, \sqrt{a^2+b^2}\operatorname{sen}\theta) \\
R_\theta(U_3) &= -\operatorname{sen}\theta\, U_2 + \cos\theta\, U_3 = (\frac{b\operatorname{sen}\theta - ac\cos\theta}{\sqrt{a^2+b^2}}, \frac{-a\operatorname{sen}\theta - bc\cos\theta}{\sqrt{a^2+b^2}}, \sqrt{a^2+b^2}\cos\theta).
\end{aligned}$$

Se $\mathcal{B} = \{E_1, E_2, E_3\}$ é a base canônica de $\mathbb{R}^3$, então

$$R_{\theta,U}(X) = [R_{\theta,U}(X)]_{\mathcal{B}} = [R_{\theta,U}]_{\mathcal{B}}^{\mathcal{B}}[X]_{\mathcal{B}}$$

Podemos escrever $R_{\theta,U} = R_{\theta,U} I$ e assim

$$[R_{\theta,U}]_{\mathcal{B}}^{\mathcal{B}} = [R_{\theta,U} I]_{\mathcal{B}}^{\mathcal{B}} = [R_{\theta,U}]_{\mathcal{C}}^{\mathcal{B}}[I]_{\mathcal{B}}^{\mathcal{C}}$$

Agora,

$$[R_{\theta,U}]_{\mathcal{C}}^{\mathcal{B}} = [\ [R_{\theta,U}(U_1)]_{\mathcal{B}}\ [R_{\theta,U}(U_2)]_{\mathcal{B}}\ [R_{\theta,U}(U_3)]_{\mathcal{B}}\ ] = [\ R_{\theta,U}(U_1)\ R_{\theta,U}(U_2)\ R_{\theta,U}(U_3)\ ],$$

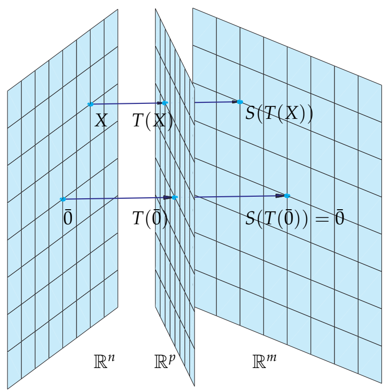


Figura 6.14. Composição das Transformações Lineares $T : \mathbb{R}^n \to \mathbb{R}^p$ e $S : \mathbb{R}^p \to \mathbb{R}^m$

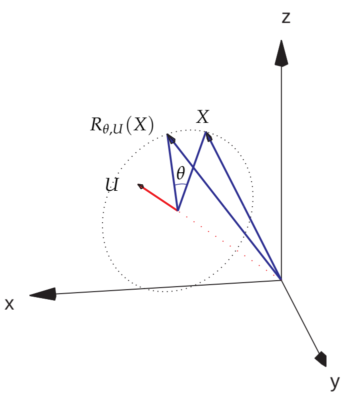


Figura 6.15. Rotação de um ângulo θ em torno de um eixo determinado pelo vetor U

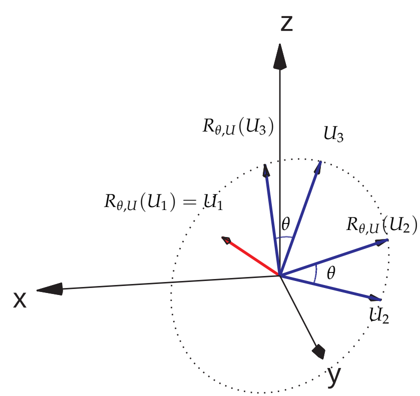


Figura 6.16. $R_{\theta,U}(U_1), R_{\theta,U}(U_2)$ e $R_{\theta,U}(U_3)$

e

$$[I]_{\mathcal{B}}^{\mathcal{C}} = ([I]_{\mathcal{C}}^{\mathcal{B}})^{-1} = [\ [U_1]_{\mathcal{B}}\ [U_2]_{\mathcal{B}}\ [U_3]_{\mathcal{B}}\ ]^{-1} = [\ U_1\ U_2\ U_3\ ]^t,$$

pois $\mathcal{B}$ é a base canônica e os vetores U_1, U_2 e U_3 são unitários mutuamente ortogonais. Assim,

$$R_{\theta,U}(X) = [R_{\theta,U}]_{\mathcal{C}}^{\mathcal{B}}[X]_{\mathcal{C}} = [\ R_{\theta,U}(U_1)\ R_{\theta,U}(U_2)\ R_{\theta,U}(U_3)\ ][\ U_1\ U_2\ U_3\ ]^t X.$$

Mas,

$$[\ R_{\theta,U}(U_1)\ R_{\theta,U}(U_2)\ R_{\theta,U}(U_3)\ ][\ U_1\ U_2\ U_3\ ]^t \quad =$$

$$\begin{bmatrix} a & \dfrac{-b\cos\theta - ac\operatorname{sen}\theta}{\sqrt{a^2+b^2}} & \dfrac{b\operatorname{sen}\theta - ac\cos\theta}{\sqrt{a^2+b^2}} \\ b & \dfrac{a\cos\theta - bc\operatorname{sen}\theta}{\sqrt{a^2+b^2}} & \dfrac{-a\operatorname{sen}\theta - bc\cos\theta}{\sqrt{a^2+b^2}} \\ c & \sqrt{a^2+b^2}\operatorname{sen}\theta & \sqrt{a^2+b^2}\cos\theta \end{bmatrix} \begin{bmatrix} a & b & c \\ -\dfrac{b}{\sqrt{a^2+b^2}} & \dfrac{a}{\sqrt{a^2+b^2}} & 0 \\ -\dfrac{ac}{\sqrt{a^2+b^2}} & -\dfrac{bc}{\sqrt{a^2+b^2}} & \sqrt{a^2+b^2} \end{bmatrix} =$$

$$\begin{bmatrix} a^2(1-\cos\theta)+\cos\theta & ab(1-\cos\theta)-c\operatorname{sen}\theta & ac(1-\cos\theta)+b\operatorname{sen}\theta \\ ab(1-\cos\theta)+c\operatorname{sen}\theta & b^2(1-\cos\theta)+\cos\theta & bc(1-\cos\theta)-a\operatorname{sen}\theta \\ ac(1-\cos\theta)-b\operatorname{sen}\theta & bc(1-\cos\theta)+a\operatorname{sen}\theta & c^2(1-\cos\theta)+\cos\theta \end{bmatrix}$$

que é a matriz de $R_{\theta,U}$ em relação à base canônica. Finalmente,

$$R_{\theta,U}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a^2(1-\cos\theta)+\cos\theta & ab(1-\cos\theta)-c\operatorname{sen}\theta & ac(1-\cos\theta)+b\operatorname{sen}\theta \\ ab(1-\cos\theta)+c\operatorname{sen}\theta & b^2(1-\cos\theta)+\cos\theta & bc(1-\cos\theta)-a\operatorname{sen}\theta \\ ac(1-\cos\theta)-b\operatorname{sen}\theta & bc(1-\cos\theta)+a\operatorname{sen}\theta & c^2(1-\cos\theta)+\cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

6.3.3 Semelhança

Corolário 6.16. *Sejam $\mathcal{B} = \{V_1, \ldots, V_n\}$ e $\mathcal{C} = \{W_1, \ldots, W_n\}$ bases do $\mathbb{R}^n$. Se $T : \mathbb{R}^n \to \mathbb{R}^n$ é uma transformação linear, então*

$$[T]_{\mathcal{C}}^{\mathcal{C}} = P^{-1}\,[T]_{\mathcal{B}}^{\mathcal{B}}\,P.$$

Demonstração. Pelo Teorema 6.12 na página 376 temos que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}[T]_{\mathcal{B}}^{\mathcal{B}}[I]_{\mathcal{C}}^{\mathcal{B}}.$$

Mas pelo Corolário 6.15 na página 380 a matriz $P = [I]_{\mathcal{C}}^{\mathcal{B}}$ é invertível e $P^{-1} = [I]_{\mathcal{B}}^{\mathcal{C}}$.
De onde segue-se o resultado. ■

Uma transformação linear do $\mathbb{R}^n$ nele mesmo é chamada um **operador linear**. Sejam A e B matrizes $n \times n$. Dizemos que B é **semelhante a** A se existe uma matriz invertível P tal que $B = P^{-1}AP$. Observe que com esta terminologia o Corolário 6.16 pode ser estabelecido da seguinte forma:

Se $T : \mathbb{R}^n \to \mathbb{R}^n$ é uma transformação linear, $\mathcal{B} = \{V_1, \ldots, V_n\}$ e $\mathcal{C} = \{W_1, \ldots, W_n\}$ são bases de $\mathbb{R}^n$, então $[T]_{\mathcal{C}}^{\mathcal{C}}$ é semelhante a $[T]_{\mathcal{B}}^{\mathcal{B}}$.

O traço de uma matriz quadrada A, denotado por $\text{tr}(A)$, é definido como sendo a soma dos elementos da sua diagonal principal. Como $\text{tr}(AB) = \text{tr}(BA)$ (Exercício 1.1.26 na página 27), então

$$\text{tr}(P^{-1}AP) = \text{tr}(A).$$

Assim, em virtude do Corolário 6.16, se $\mathbb{R}^n$ é um espaço vetorial de dimensão finita, podemos definir o **traço de um operador linear** $T : \mathbb{R}^n \to \mathbb{R}^n$ como sendo

$$\text{tr}(T) = \text{tr}([T]_{\mathcal{B}}^{\mathcal{B}}),$$

onde $\mathcal{B}$ é uma base de $\mathbb{R}^n$.

De forma análoga, como $\det(AB) = \det(A)\det(B) = \det(BA)$ (Teorema 2.13 na página 110), então

$$\det(P^{-1}AP) = \det(A).$$

Assim, em virtude do Corolário 6.16, se $\mathbb{R}^n$ é um espaço vetorial de dimensão finita, podemos definir o **determinante de um operador linear** $T : \mathbb{R}^n \to \mathbb{R}^n$ como sendo

$$\det(T) = \det([T]_{\mathcal{B}}^{\mathcal{B}}),$$

onde $\mathcal{B}$ é uma base de $\mathbb{R}^n$ qualquer.

Exemplo 6.21. Vamos obter uma expressão para a reflexão na reta $r : y = 2x$, $R_r : \mathbb{R}^2 \to \mathbb{R}^2$, usando o Corolário 6.16. Vamos escolher uma base do $\mathbb{R}^2$, tal que a avaliação de R_r nos elementos desta base seja fácil de se obter. Por exemplo, $\mathcal{C} = \{V_1 = (1,2), V_2 = (-2,1)\}$.

$$\begin{array}{rcrcrcl} R_r(V_1) &=& R_r(1,2) &=& (1,2) &=& 1\,V_1 + 0\,V_2 \\ R_r(V_2) &=& R_r(-2,1) &=& (2,-1) &=& 0\,V_1 - 1\,V_2. \end{array}$$

Assim,

$$B = [R_r]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

A matriz mudança de base, da base $\mathcal{C}$ para a base canônica $\mathcal{B} = \{(1,0),(0,1)\}$ é dada por

$$P = [I_{\mathbb{R}^2}]_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}.$$

Pelo Corolário 6.16, a matriz $A = [R_r]_{\mathcal{B}}^{\mathcal{B}}$ é obtida através da equação matricial

$$A = [R_r]_{\mathcal{B}}^{\mathcal{B}} = [I_{\mathbb{R}^2}]_{\mathcal{C}}^{\mathcal{B}}[R_r]_{\mathcal{C}}^{\mathcal{C}}[I_{\mathbb{R}^2}]_{\mathcal{B}}^{\mathcal{C}} = PBP^{-1}.$$

Vamos enunciar uma versão mais geral do Corolário 6.16, cuja demonstração é inteiramente análoga e deixamos como exercício para o leitor.

Corolário 6.17. *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Sejam $\mathcal{B} = \{V_1, \ldots, V_n\}$ e $\mathcal{B}' = \{V_1', \ldots, V_n'\}$ bases de $\mathbb{R}^n$ e $\mathcal{C} = \{W_1, \ldots, W_m\}$ e $\mathcal{C}' = \{W_1', \ldots, W_m'\}$ bases de $\mathbb{R}^m$. Então,*

$$[T]_{\mathcal{B}'}^{\mathcal{C}'} = P^{-1}[T]_{\mathcal{B}}^{\mathcal{C}}Q,$$

onde P é a matriz mudança de base de $\mathcal{C}'$ para $\mathcal{C}$ e Q, de $\mathcal{B}'$ para $\mathcal{B}$.

Exercícios Numéricos (respostas na página 556)

6.3.1. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear dada por $T(x, y) = (2x + y, x - 3y)$, para todo $(x, y) \in \mathbb{R}^2$. Seja $\mathcal{C} = \{(1, 1), (1, 2)\}$. Determine $[T]_{\mathcal{C}}^{\mathcal{C}}$.

6.3.2. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(X) = AX$, para todo $X \in \mathbb{R}^3$, onde

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

Sejam $V_1 = (1, 0, 0)$, $V_2 = (1, 2, 0)$ e $V_3 = (0, -2, 1)$.

(a) Encontre a matriz mudança de base de $\mathcal{C} = \{V_1, V_2, V_3\}$ para a base canônica $\mathcal{B} = \{E_1, E_2, E_3\}$;

(b) Use a matriz obtida no item anterior para determinar a matriz B que representa T com relação à base $\{V_1, V_2, V_3\}$.

6.3.3. Considere a reta $r : (x, y, z) = t(1, 1, 1)$. Sejam $\mathcal{B} = \{E_1, E_2, E_3\}$ a base canônica do $\mathbb{R}^3$ e $\mathcal{C} = \{U_1, U_2, U_3\}$ a base ortonormal de $\mathbb{R}^3$ definida por $U_1 = 1/\sqrt{3}(1, 1, 1)$, $U_2 = 1/\sqrt{2}(-1, 1, 0)$ e $U_3 = 1/\sqrt{6}(-1, -1, 2)$.

(a) Seja $P_r : \mathbb{R}^3 \to \mathbb{R}^3$ a projeção ortogonal na reta r. Encontre $[P_r]_{\mathcal{C}}^{\mathcal{C}}$ e $[P_r]_{\mathcal{B}}^{\mathcal{B}}$.

(b) Seja $R_r : \mathbb{R}^3 \to \mathbb{R}^3$ a reflexão em relação à reta r. Encontre $[R_r]_{\mathcal{C}}^{\mathcal{C}}$ e $[R_r]_{\mathcal{B}}^{\mathcal{B}}$.

6.3.4. Considere a reta $r : (x, y, z) = t(1, 1, 0)$. Sejam $\mathcal{B} = \{E_1, E_2, E_3\}$ a base canônica do $\mathbb{R}^3$ e $\mathcal{C} = \{U_1, U_2, U_3\}$ a base ortonormal de $\mathbb{R}^3$ definida por $U_1 = 1/\sqrt{2}(1, 1, 0)$, $U_2 = (0, 0, 1)$ e $U_3 = 1/\sqrt{2}(1, -1, 0)$. Seja $R_{\pi/2,r} : \mathbb{R}^3 \to \mathbb{R}^3$ a transformação linear, que é uma rotação de um ângulo de $\pi/2$ em torno da reta r. Determine $[R_{\pi/2,r}]_{\mathcal{C}}^{\mathcal{C}}$ e $[R_{\pi/2,r}]_{\mathcal{B}}^{\mathcal{B}}$.

6.3.5. Para cada uma das transformações lineares T verifique se T é invertível e calcule a inversa, T^{-1}, se ela existe.

(a) $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(x, y, z) = (x + 2y + z, y + 2z, z)$.

(b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(a, b, c) = (a, -2a + b, -2a - 4b + c)$.

(c) $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(a, b, c) = (a + b + c, a + 2b + c, a + 2c)$.

(d) $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(a, b, c) = (a + b + c, -a + c, a + c)$.

Exercícios Teóricos

6.3.6. Mostre que se $T : \mathbb{R}^2 \to \mathbb{R}^2$ é uma transformação linear invertível, então T é uma composição de expansões, compressões, cisalhamentos e reflexões. (Sugestão: use o fato de que toda matriz invertível é o produto de matrizes elementares e o Exercício 6.1.12 na página 356.)

6.3.7. Seja A uma matriz triangular superior $n \times n$ com todos os elementos da diagonal iguais a zero. Mostre que $A^n = \bar{0}$. (Sugestão: considere o operador $T : \mathbb{R}^n \to \mathbb{R}^n$ cuja matriz na base canônica é igual a A e determine a matriz de T^n.)

6.3.8. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Sejam $\mathcal{B} = \{V_1, \ldots, V_n\}$ e $\mathcal{B}' = \{V'_1, \ldots, V'_n\}$ bases de $\mathbb{R}^n$ e $\mathcal{C} = \{W_1, \ldots, W_m\}$ e $\mathcal{C}' = \{W'_1, \ldots, W'_m\}$ bases de $\mathbb{R}^m$. Mostre que

$$[T]_{\mathcal{B}'}^{\mathcal{C}'} = P^{-1}[T]_{\mathcal{B}}^{\mathcal{C}}Q,$$

onde P é a matriz mudança de base de $\mathcal{C}'$ para $\mathcal{C}$ e Q, de $\mathcal{B}'$ para $\mathcal{B}$. (Sugestão: siga a demonstração do Corolário 6.16 na página 384.)

6.3.9. Seja $\mathcal{B} = \{V_1, \ldots, V_n\}$ uma base de um espaço vetorial $\mathbb{R}^n$. Seja P uma matriz $n \times n$ invertível. Mostre que $\mathcal{C} = \{W_1, \ldots, W_n\}$ é uma base de $\mathbb{R}^n$, onde $W_j = \sum_{i=1}^{n} p_{ij}V_i$, para $j = 1, \ldots, n$. Assim, P é a matriz mudança de base de $\mathcal{B}$ para $\mathcal{C}$.

6.3.10. Um operador $T : \mathbb{R}^n \to \mathbb{R}^n$ é chamado **nilpotente** se $T^n = O$, a transformação linear nula, para algum $n \in \mathbb{N}$. Seja T um operador nilpotente. Mostre que existe um vetor $V \neq \bar{0}$ tal que $T(V) = \bar{0}$.

7

Diagonalização

7.1 Diagonalização de Matrizes

7.1.1 Motivação

Certos processos são descritos em cada estágio por uma matriz A quadrada e em k estágios pela potência k da matriz A, A^k, em que k é um número inteiro positivo. Suponha que desejamos saber a matriz que corresponde a k estágios, para k um inteiro positivo qualquer. Se a matriz A é diagonal,

$$A = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}, \quad \text{então} \quad A^k = \begin{bmatrix} \lambda_1^k & 0 & \ldots & 0 \\ 0 & \lambda_2^k & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n^k \end{bmatrix}.$$

Se a matriz A não é diagonal, mas existe uma matriz P tal que

$$A = PDP^{-1}, \quad \text{em que} \quad D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix},$$

então

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PD^2P^{-1}.$$

Agora, supondo que $A^{k-1} = PD^{k-1}P^{-1}$, temos que

$$\begin{aligned} A^k &= A^{k-1}A = (PDP^{-1})^{k-1}(PDP^{-1}) \\ &= (PD^{k-1}P^{-1})(PDP^{-1}) = PD^{k-1}(P^{-1}P)DP^{-1} \\ &= PD^kP^{-1} = P \begin{bmatrix} \lambda_1^k & 0 & \ldots & 0 \\ 0 & \lambda_2^k & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n^k \end{bmatrix} P^{-1}. \end{aligned}$$

Assim, podemos facilmente encontrar a k-ésima potência de A.

Exemplo 7.1. Seja

$$A = \begin{bmatrix} 1 & -1 \\ -4 & 1 \end{bmatrix}.$$

mostraremos no Exemplo 7.6 na página 406 que

$$P = \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$$

são tais que

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

Assim,

$$\begin{aligned} A^k &= PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 3^k & 0 \\ 0 & (-1)^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 3^k & (-1)^k \\ -2\,3^k & 2(-1)^k \end{bmatrix} \frac{1}{4} \begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} 2(3^k + (-1)^k) & (-1)^k - 3^k \\ 4((-1)^k - 3^k) & 2(3^k + (-1)^k) \end{bmatrix} \end{aligned}$$

Vamos descobrir, a seguir, como podemos determinar matrizes P e D, quando elas existem, tais que $A = PDP^{-1}$, ou multiplicando à esquerda por P^{-1} e à direita por P, $D = P^{-1}AP$, com D sendo uma matriz diagonal. Chamamos **diagonalização** ao processo de encontrar as matrizes P e D.

7.1.2 Autovalores e Autovetores

Definição 7.1. Dizemos que uma matriz A, $n \times n$, é **diagonalizável**, se existem matrizes P e D tais que

$$A = PDP^{-1},$$

ou equivalentemente, $D = P^{-1}AP$, em que D é uma matriz diagonal.

Exemplo 7.2. Toda matriz diagonal

$$A = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}$$

é diagonalizável, pois

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

Vamos supor inicialmente que a matriz A seja diagonalizável. Então existe uma matriz P tal que

$$P^{-1}AP = D\,, \tag{7.1}$$

em que D é uma matriz diagonal. Vamos procurar tirar conclusões sobre as matrizes P e D.

Multiplicando à esquerda por P ambos os membros da equação anterior, obtemos

$$AP = PD\,. \tag{7.2}$$

Sejam

$$D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix} \quad \text{e} \quad P = \begin{bmatrix} V_1 & V_2 & \ldots & V_n \end{bmatrix},$$

em que V_j é a coluna j de P. Por um lado

$$AP = A\left[\begin{array}{cccc} V_1 & V_2 & \ldots & V_n \end{array}\right] = \left[\begin{array}{cccc} AV_1 & AV_2 & \ldots & AV_n \end{array}\right]$$

(Exercício 1.1.18 na página 25) e por outro lado

$$PD = \left[\begin{array}{cccc} V_1 & V_2 & \ldots & V_n \end{array}\right] \left[\begin{array}{cccc} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{array}\right] = \left[\begin{array}{cccc} \lambda_1 V_1 & \lambda_2 V_2 & \ldots & \lambda_n V_n \end{array}\right]$$

(Exercício 1.1.17 na página 25) Assim, (7.2) pode ser reescrita como,

$$\left[\begin{array}{cccc} AV_1 & AV_2 & \ldots & AV_n \end{array}\right] = \left[\begin{array}{cccc} \lambda_1 V_1 & \lambda_2 V_2 & \ldots & \lambda_n V_n \end{array}\right].$$

Logo,

$$AV_j = \lambda_j V_j,$$

para $j = 1, \ldots n$. Ou seja, as colunas de P, V_j, e os elementos da diagonal de D, λ_j, satisfazem a equação

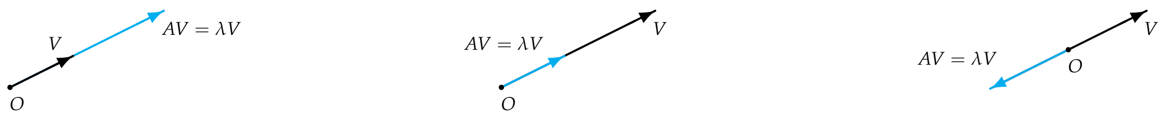
$$AX = \lambda X,$$

em que λ e X são incógnitas. Isto motiva a seguinte definição.

Definição 7.2. Seja A uma matriz $n \times n$. Um número real λ é chamado **autovalor** (real) de A, se existe um vetor *não nulo* $V = \left[\begin{array}{c} v_1 \\ \vdots \\ v_n \end{array}\right]$ de $\mathbb{R}^n$, tal que

$$AV = \lambda V\,. \tag{7.3}$$

Um vetor *não nulo* que satisfaça (7.3), é chamado de **autovetor** de A.



$\lambda > 1$ $\qquad$ $0 < \lambda < 1$ $\qquad$ $\lambda < 0$

Observe que, usando o fato de que a matriz identidade

$$I_n = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & 1 \end{bmatrix}$$

é tal que $I_n V = V$, a equação (7.3) pode ser escrita como

$$AV = \lambda I_n V$$

ou

$$(A - \lambda I_n)V = \bar{0}\,. \tag{7.4}$$

Como os autovetores são vetores não nulos, os autovalores são os valores de λ, para os quais o sistema $(A - \lambda I_n)X = \bar{0}$ tem solução não trivial. Mas, este sistema homogêneo tem solução não trivial se, e somente se, $\det(A - \lambda I_n) = 0$ (Teorema 2.15 na página 115). Assim temos um método para encontrar os autovalores e os autovetores de uma matriz A.

Proposição 7.1. *Seja A uma matriz $n \times n$.*

(a) *Os autovalores (reais) de A são as raízes* ***reais*** *do polinômio*

$$p(t) = \det(A - t\, I_n) \tag{7.5}$$

(b) *Para cada autovalor λ, os autovetores associados a λ são os vetores não nulos da solução do sistema*

$$(A - \lambda I_n)X = \bar{0}\,. \tag{7.6}$$

Definição 7.3. Seja A uma matriz $n \times n$. O polinômio

$$p(t) = \det(A - t\, I_n) \tag{7.7}$$

é chamado **polinômio característico de** A.

Assim, para determinarmos os autovalores de uma matriz A precisamos determinar as raízes reais do seu polinômio característico, que tem a forma

$$p(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0.$$

Um resultado sobre polinômios que muitas vezes é útil, é o que diz que se $a_0, a_1, \ldots, a_{n-1}$ são inteiros, então as suas raízes racionais (se existirem) são números inteiros e divisores do coeficiente do termo de grau zero a_0. Por exemplo, se $p(t) = -t^3 + 6t^2 - 11t + 6$, então as possíveis raízes racionais são $\pm 1, \pm 2, \pm 3$ e ± 6. Substituindo estes valores em $p(t)$, vemos que $p(1) = 0$, ou seja, 1 é uma raiz de $p(t)$. Finalmente, dividindo $p(t)$ por $t-1$, obtemos que $p(t) = (t-1)(-t^2+5t-6)$. Como as raízes de $-t^2+5t-6$ são 2 e 3, então as raízes de $p(t)$, são $1, 2$ e 3.

Exemplo 7.3. Vamos determinar os autovalores e autovetores da matriz

$$A = \begin{bmatrix} 1 & -1 \\ -4 & 1 \end{bmatrix}$$

Para esta matriz o polinômio característico é

$$p(t) = \det(A - tI_2) = \det \begin{bmatrix} 1-t & -1 \\ -4 & 1-t \end{bmatrix} = (1-t)^2 - 4 = t^2 - 2t - 3\,.$$

Como os autovalores de A são as raízes de $p(t)$, temos que os autovalores de A são $\lambda_1 = 3$ e $\lambda_2 = -1$.

Agora, vamos determinar os autovetores associados aos autovalores $\lambda_1 = 3$ e $\lambda_2 = -1$. Para isto vamos resolver os sistemas $(A - \lambda_1 I_2)X = \bar{0}$ e $(A - \lambda_2 I_2)X = \bar{0}$. Como

$$A - \lambda_1 I_2 = \begin{bmatrix} -2 & -1 \\ -4 & -2 \end{bmatrix},$$

então

$$(A - \lambda_1 I_2)X = \bar{0}$$

é

$$\begin{bmatrix} -2 & -1 \\ -4 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ou

$$\left\{ \begin{array}{rcrcr} -2x & - & y & = & 0 \\ -4x & - & 2y & = & 0 \end{array} \right.$$

cuja solução geral é

$$\mathbb{W}_1 = \{(\alpha, -2\alpha) \mid \alpha \in \mathbb{R}\}.$$

que é o conjunto de todos os autovetores associados a $\lambda_1 = 3$ acrescentado o vetor nulo. Agora,

$$(A - \lambda_2 I_2)X = \bar{0}$$

é

$$\begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

cuja solução geral é

$$\mathbb{W}_2 = \{(\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\},$$

que é o conjunto de todos os autovetores associados a $\lambda_2 = -1$ acrescentado o vetor nulo.

Exemplo 7.4. Vamos determinar os autovalores e autovetores da matriz

$$A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

Para esta matriz o polinômio característico é

$$\begin{aligned}
p(t) &= \det(A - t\,I_3) = \det\begin{bmatrix} 4-t & 2 & 2 \\ 2 & 4-t & 2 \\ 2 & 2 & 4-t \end{bmatrix} \\
&= (4-t)\det\begin{bmatrix} 4-t & 2 \\ 2 & 4-t \end{bmatrix} - 2\det\begin{bmatrix} 2 & 2 \\ 2 & 4-t \end{bmatrix} + 2\det\begin{bmatrix} 2 & 4-t \\ 2 & 2 \end{bmatrix} \\
&= (4-t)\det\begin{bmatrix} 4-t & 2 \\ 2 & 4-t \end{bmatrix} - 4\det\begin{bmatrix} 2 & 2 \\ 2 & 4-t \end{bmatrix} \\
&= (4-t)[(4-t)^2 - 4] - 8(2-t) = -t^3 + 12t^2 - 36t + 32
\end{aligned}$$

Substituindo-se $t = \pm 1$ obtemos

$$p(1) = -1 + 12 - 36 + 32 > 0, \quad p(-1) = 1 + 12 + 36 + 32 > 0.$$

Substituindo-se $t = 2$ obtemos $p(2) = 0$. Dividindo-se $p(t)$ por $t - 2$ obtemos que

$$p(t) = (t-2)(-t^2 + 10t - 16) = (t-2)^2(8-t).$$

Portanto os autovalores de A são $\lambda_1 = 2$ e $\lambda_2 = 8$. Agora, vamos determinar os autovetores associados aos autovalores λ_1 e λ_2. Para isto vamos resolver os sistemas $(A - \lambda_1 I_3)X = \bar{0}$ e $(A - \lambda_2 I_3)X = \bar{0}$. Como

$$(A - \lambda_1 I_3)X = \bar{0} \quad \text{é} \quad \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

A forma escalonada reduzida da matriz aumentada do sistema é

$$\left[\begin{array}{ccc:c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Assim, a solução geral do sistema $(A - \lambda_1 I_3)X = \bar{0}$ é

$$\mathbb{W}_1 = \{(-\alpha - \beta, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\},$$

que é o conjunto de todos os autovetores associados a $\lambda_1 = 2$ acrescentado o vetor nulo.

Com relação ao autovalor $\lambda_2 = 8$, o sistema $(A - \lambda_2 I_3)X = \bar{0}$ é

$$\begin{bmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

A forma escalonada reduzida da matriz aumentada do sistema é

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Assim, a solução geral do sistema $(A - \lambda_2 I_3)X = \bar{0}$ é

$$\mathbb{W}_2 = \{(\alpha, \alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

Nos exemplos anteriores, para cada autovalor encontramos todos os autovetores associados a ele. Podemos observar que para cada autovalor λ, o conjunto dos autovetores associados a ele acrescentado o vetor nulo é um subespaço, já que é o conjunto solução de um sistema linear homogêneo $(A - \lambda I_n)X = \bar{0}$. Este subespaço recebe o nome de **autoespaço associado ao autovalor** λ.

7.1.3 Diagonalização

Vamos enunciar e demonstrar o resultado principal deste capítulo. Já vimos que se uma matriz A é diagonalizável, então as colunas da matriz P, que faz a diagonalização, são autovetores associados a autovalores, que por sua vez são elementos da matriz diagonal D. Como a matriz P é invertível, estes n autovetores são L.I. Vamos mostrar, a seguir, que se a matriz A tem n autovetores L.I., então ela é diagonalizável.

Teorema 7.2. *Seja A uma matriz $n \times n$ que tem n autovetores L.I. $V_1, \ldots, V_n$ associados a $\lambda_1, \ldots, \lambda_n$, respectivamente. Então as matrizes*

$$P = \begin{bmatrix} V_1 & V_2 & \ldots & V_n \end{bmatrix} \quad e \quad D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}.$$

são tais que

$$D = P^{-1}AP,$$

ou seja A é diagonalizável. Reciprocamente, se A é diagonalizável, então ela possui n autovetores linearmente independentes.

Demonstração. Suponha que $V_1, \ldots, V_n$ são n autovetores linearmente independentes associados a $\lambda_1, \ldots, \lambda_n$, respectivamente. Vamos definir as matrizes

$$P = \begin{bmatrix} V_1 & V_2 & \ldots & V_n \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}.$$

Como $AV_j = \lambda_j V_j$, para $j = 1, \ldots, n$, então

$$\begin{aligned} AP &= A\left[\begin{array}{cccc} V_1 & V_2 & \ldots & V_n \end{array}\right] = \left[\begin{array}{cccc} AV_1 & AV_2 & \ldots & AV_n \end{array}\right] \\ &= \left[\begin{array}{cccc} \lambda_1 V_1 & \lambda_2 V_2 & \ldots & \lambda_n V_n \end{array}\right] = \left[\begin{array}{cccc} V_1 & V_2 & \ldots & V_n \end{array}\right] \left[\begin{array}{cccc} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{array}\right] = PD. \end{aligned}$$

Como $V_1, \ldots, V_n$ são L.I., a matriz P é invertível. Assim, multiplicando a equação anterior por P^{-1} à esquerda obtemos

$$D = P^{-1}AP.$$

Ou seja, a matriz A é diagonalizável.

Vamos, agora, provar que se A é diagonalizável, então ela possui n autovetores L.I. Se a matriz A é diagonalizável, então existe uma matriz P tal que

$$P^{-1}AP = D\,, \tag{7.8}$$

em que D é uma matriz diagonal. Multiplicando à esquerda por P ambos os membros da equação anterior, obtemos

$$AP = PD\,. \tag{7.9}$$

Sejam

$$D = \left[\begin{array}{cccc} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{array}\right] \quad \text{e} \quad P = \left[\begin{array}{cccc} V_1 & V_2 & \ldots & V_n \end{array}\right],$$

em que V_j é a coluna j de P. Usando as definições de P e D temos que

$$AP = A\left[\begin{array}{cccc} V_1 & V_2 & \ldots & V_n \end{array}\right] = \left[\begin{array}{cccc} AV_1 & AV_2 & \ldots & AV_n \end{array}\right]$$

$$PD = \begin{bmatrix} V_1 & V_2 & \ldots & V_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 V_1 & \lambda_2 V_2 & \ldots & \lambda_n V_n \end{bmatrix}$$

Assim, de (7.9) segue-se que

$$AV_j = \lambda_j V_j,$$

para $j = 1, \ldots n$. Como a matriz P é invertível, pela Proposição 3.9 na página 217, os autovetores $V_1, \ldots, V_n$ são L.I.

■

Assim, se uma matriz A é diagonalizável e $D = P^{-1}AP$, então os autovalores de A formam a diagonal de D e n autovetores linearmente independentes associados aos autovalores formam as colunas de P.

O resultado que vem a seguir, garante que se conseguirmos para cada autovalor, autovetores L.I., então ao juntarmos todos os autovetores obtidos, eles continuarão sendo L.I.

Proposição 7.3. *Seja A uma matriz $n \times n$. Se $V_1^{(1)}, \ldots, V_{n_1}^{(1)}$ são autovetores L.I. associados a λ_1, $V_1^{(2)}, \ldots, V_{n_2}^{(2)}$ são autovetores L.I. associados a λ_2, ..., $V_1^{(k)}, \ldots, V_{n_k}^{(k)}$ são autovetores L.I. associados a λ_k, com $\lambda_1, \ldots, \lambda_k$ distintos, então $\{V_1^{(1)}, \ldots, V_{n_1}^{(1)}, \ldots, V_1^{(k)}, \ldots, V_{n_k}^{(k)}\}$ é um conjunto L.I.*

Demonstração. Vamos demonstrar apenas para o caso em que temos dois autovalores diferentes. O caso geral é inteiramente análogo. Sejam $V_1^{(1)}, \ldots, V_{n_1}^{(1)}$ autovetores L.I. associados a λ_1 e $V_1^{(2)}, \ldots, V_{n_2}^{(2)}$ autovetores L.I. associados a λ_2. Precisamos mostrar que a única solução da equação

$$x_1^{(1)} V_1^{(1)} + \ldots + x_{k_1}^{(1)} V_{n_1}^{(1)} + x_1^{(2)} V_1^{(2)} + \ldots + x_{k_2}^{(2)} V_{n_2}^{(2)} = \bar{0} \tag{7.10}$$

é a solução trivial. Multiplicando a equação (7.10) por A e usando o fato de que os $V_i^{(j)}$ são autovetores, obtemos

$$x_1^{(1)} \lambda_1 V_1^{(1)} + \ldots + x_{n_1}^{(1)} \lambda_1 V_{n_1}^{(1)} + x_1^{(2)} \lambda_2 V_1^{(2)} + \ldots + x_{n_2}^{(2)} \lambda_2 V_{n_2}^{(2)} = \bar{0} \tag{7.11}$$

Multiplicando a equação (7.10) por λ_1, obtemos

$$x_1^{(1)} \lambda_1 V_1^{(1)} + \ldots + x_{n_1}^{(1)} \lambda_1 V_{n_1}^{(1)} + x_1^{(2)} \lambda_1 V_1^{(2)} + \ldots + x_{n_2}^{(2)} \lambda_1 V_{n_2}^{(2)} = \bar{0}\,. \tag{7.12}$$

Subtraindo a equação (7.11) da equação (7.12), obtemos

$$x_1^{(2)} (\lambda_2 - \lambda_1) V_1^{(2)} + \ldots + x_{n_2}^{(2)} (\lambda_2 - \lambda_1) V_{n_2}^{(2)} = \bar{0}\,.$$

Como $V_1^{(2)}, \ldots, V_{n_2}^{(2)}$ são L.I., temos que $x_1^{(2)} = \ldots = x_{n_2}^{(2)} = 0$. Agora, multiplicando a equação (7.10) por λ_2 e subtraindo da equação (7.12) obtemos

$$x_1^{(1)} (\lambda_2 - \lambda_1) V_1^{(1)} + \ldots + x_{n_1}^{(1)} (\lambda_2 - \lambda_1) V_{n_1}^{(1)} = \bar{0}\,.$$

Como $V_1^{(1)}, \ldots, V_{n_1}^{(1)}$ são L.I., temos que $x_1^{(1)} = \ldots = x_{n_1}^{(1)} = 0$. O que prova que todos os autovetores juntos são L.I. ■

Exemplo 7.5. Considere a matriz

$$A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

Já vimos no Exemplo 7.4 na página 398 que seu polinômio característico é

$$p(t) = (t-2)(-t^2 + 10t - 16) = (t-2)^2(8-t),$$

os seus autovalores são $\lambda_1 = 2$ e $\lambda_2 = 8$ e os autoespaços correspondentes são

$$\mathbb{W}_1 = \{(-\alpha - \beta, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\},$$

$$\mathbb{W}_2 = \{(\alpha, \alpha, \alpha) \mid \alpha \in \mathbb{R}\},$$

respectivamente. Vamos encontrar, para cada autoespaço, o maior número possível de autovetores L.I., ou seja, vamos encontrar uma base para cada autoespaço. E o teorema anterior garante que se juntarmos todos estes autovetores eles vão continuar sendo L.I.

Para $\mathbb{W}_1$, temos que

$$(-\alpha - \beta, \beta, \alpha) = \alpha(-1, 0, 1) + \beta(-1, 1, 0).$$

Assim, os vetores $V_1 = (-1, 0, 1)$ e $V_2 = (-1, 1, 0)$ geram $\mathbb{W}_1$. Como além disso, eles são L.I. (um não é múltiplo escalar do outro), então eles formam uma base para $\mathbb{W}_1$. Assim, não podemos ter um número maior de autovetores L.I. associados a $\lambda_1 = 2$ (Teorema 4.1 na página 232).

Para $\mathbb{W}_2$, temos que o conjunto $\{V_3 = (1, 1, 1)\}$ é uma base para $\mathbb{W}_2$, pois como

$$(\alpha, \alpha, \alpha) = \alpha(1, 1, 1),$$

V_3 gera $\mathbb{W}_2$ e um vetor não nulo é L.I. Assim, não podemos ter um número maior de autovetores L.I. associados a $\lambda_2 = 8$ (Teorema 4.1 na página 232).

Como V_1 e V_2 são autovetores L.I. associados a λ_1 e V_3 é um autovetor L.I. associado a λ_2, então pela Proposição 7.3 na página 403 os autovetores juntos V_1, V_2 e V_3 são L.I. Assim, a matriz A é diagonalizável e as matrizes

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix} \quad \text{e} \quad P = [\, V_1 \; V_2 \; V_3] = \begin{bmatrix} -1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

são tais que

$$D = P^{-1}AP.$$

Exemplo 7.6. Considere a matriz

$$A = \begin{bmatrix} 1 & -1 \\ -4 & 1 \end{bmatrix}$$

Já vimos no Exemplo 7.3 na página 397 que o seu polinômio característico é

$$p(t) = \det(A - t\, I_2) = t^2 - 2t - 3,$$

que os seus autovalores são $\lambda_1 = 3$ e $\lambda_2 = -1$ e que os autoespaços correspondentes são

$$\mathbb{W}_1 = \{(\alpha, -2\alpha) \mid \alpha \in \mathbb{R}\} \quad \text{e} \quad \mathbb{W}_2 = \{(\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\},$$

respectivamente.

Para $\lambda_1 = 3$, temos que $\{V_1 = (1, -2)\}$ é uma base de $\mathbb{W}_1$. Assim, não podemos ter mais autovetores L.I. associados a λ_1. De forma análoga para $\lambda_2 = -1$, $\{V_2 = (1, 2)\}$

é um conjunto com o maior número possível de autovetores L.I. associados a λ_2. Assim, a matriz A é diagonalizável e as matrizes

$$P = [\, V_1 \; V_2 \,] = \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$$

são tais que

$$D = P^{-1}AP.$$

Exemplo 7.7. Considere a matriz

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

O seu polinômio característico é $p(t) = \det(A - t\, I_2) = t^2$, assim A possui um único autovalor: $\lambda_1 = 0$. Agora, vamos determinar os autovetores associados ao autovalor $\lambda_1 = 0$. Para isto vamos resolver o sistema $(A - \lambda_1 I_2)X = \bar{0}$. Como

$$A - \lambda_1 I_2 = A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

então

$$(A - \lambda_1 I_2)X = \bar{0}$$

é

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ou

$$\left\{ \begin{array}{rcl} y & = & 0 \\ 0 & = & 0 \end{array} \right.$$

cuja solução geral é

$$\mathbb{W}_1 = \{(\alpha, 0) \mid \alpha \in \mathbb{R}\}.$$

que é o autoespaço correspondente a $\lambda_1 = 0$. Assim, para $\lambda_1 = 0$, temos que $\{V_1 = (1, 0)\}$ é uma base de $\mathbb{V}_1$. Portanto, não podemos ter mais autovetores L.I. associados a λ_1 e como só temos um autovalor não podemos ter mais autovetores L.I. Portanto, pelo Teorema 7.2 na página 401, a matriz A **não** é diagonalizável, ou seja, não existem matrizes P e D tais que $D = P^{-1}AP$.

Exemplo 7.8. Vamos retomar a cadeia de Markov do Exemplo 1.9 na página 14. Vamos supor que uma população é dividida em três estados (por exemplo: ricos, classe média e pobres) e que em cada unidade de tempo a probabilidade de mudança de um estado para outro seja constante no tempo, só dependa dos estados.

Seja t_{ij} a probabilidade de mudança do estado j para o estado i em uma unidade de tempo (geração). A matriz de transição é dada por

$$T = \begin{array}{c} \begin{array}{ccc} (1) & (2) & (3) \end{array} \\ \left[\begin{array}{ccc} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{array}\right] \end{array} \begin{array}{c} \\ (1) \\ (2) \\ (3) \end{array}$$

Vamos considerar a matriz de transição

$$T = \begin{array}{c} \begin{array}{ccc} (1) & (2) & (3) \end{array} \\ \left[\begin{array}{ccc} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{array}\right] \end{array} \begin{array}{c} \\ (1) \\ (2) \\ (3) \end{array}$$

Vamos calcular potências k de T, para k um inteiro positivo qualquer. Para isto vamos diagonalizar a matriz T. Para isso precisamos determinar seus os autovalores e autovetores. Para esta matriz o polinômio característico é

$$\begin{aligned}
p(t) &= \det(T - t\, I_3) = \det \begin{bmatrix} \frac{1}{2}-t & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2}-t & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2}-t \end{bmatrix} \\
&= (\frac{1}{2}-t)\det \begin{bmatrix} \frac{1}{2}-t & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2}-t \end{bmatrix} - \frac{1}{4}\det \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2}-t \end{bmatrix} \\
&= (\frac{1}{2}-t)\left[(\frac{1}{2}-t)^2 - \frac{1}{8}\right] - \frac{1}{8}(\frac{1}{2}-t) \\
&= -t^3 + \frac{3}{2}t^2 - \frac{1}{2}t = t(-t^2 + \frac{3}{2}t - \frac{1}{2}) = -t(t-1)(t-\frac{1}{2})
\end{aligned}$$

Portanto os autovalores de T são $\lambda_1 = 0, \lambda_2 = 1/2$ e $\lambda_3 = 1$. Agora, vamos determinar os autovetores associados aos autovalores λ_1, λ_2 e λ_3. Para isto vamos resolver os sistemas $(T - \lambda_1 I_3)X = \bar{0}, (T - \lambda_2 I_3)X = \bar{0}$ e $(T - \lambda_3 I_3)X = \bar{0}$. Como

$$(T - \lambda_1 I_3)X = TX = \bar{0} \quad \text{é} \quad \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

A forma escalonada reduzida da matriz aumentada do sistema é

$$\left[\begin{array}{ccc:c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Assim, a solução geral do sistema $(T - \lambda_1 I_3)X = \bar{0}$ é

$$\mathbb{W}_1 = \{(\alpha, -2\alpha, \alpha) \mid \alpha \in \mathbb{R}\},$$

que é o conjunto de todos os autovetores associados a $\lambda_1 = 0$ acrescentado o vetor nulo. O conjunto $\{V_1 = (1, -2, 1)\}$ é uma base para $\mathbb{W}_1$, pois como $(\alpha, -2\alpha, \alpha) = \alpha(1, -2, 1)$, V_1 gera $\mathbb{W}_1$ e um vetor não nulo é L.I.

Com relação ao autovalor $\lambda_2 = 1/2$, o sistema $(T - \lambda_2 I_3)X = \bar{0}$ é

$$\begin{bmatrix} 0 & \frac{1}{4} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

A forma escalonada reduzida da matriz aumentada do sistema é

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Assim, a solução geral do sistema $(T - \lambda_2 I_3)X = \bar{0}$ é

$$\mathbb{W}_2 = \{(\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}.$$

O conjunto $\{V_2 = (1, 0, 1)\}$ é uma base para $\mathbb{W}_2$, pois como $(\alpha, 0, \alpha) = \alpha(1, 0, 1)$, V_3 gera $\mathbb{W}_2$ e um vetor não nulo é L.I.

Com relação ao autovalor $\lambda_3 = 1$, o sistema $(T - \lambda_3 I_3)X = \bar{0}$ é

$$\begin{bmatrix} -\frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

A forma escalonada reduzida da matriz aumentada do sistema é

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Assim, a solução geral do sistema $(T - \lambda_3 I_3)X = \bar{0}$ é

$$\mathbb{W}_3 = \{(\alpha, 2\alpha, \alpha) \mid \alpha \in \mathbb{R}\},$$

que é o conjunto de todos os autovetores associados a $\lambda_3 = 1$ acrescentado o vetor nulo. O conjunto $\{V_1 = (1,2,1)\}$ é uma base para $\mathbb{W}_1$, pois como $(\alpha, 2\alpha, \alpha) = \alpha(1,2,1)$, V_1 gera $\mathbb{W}_1$ e um vetor não nulo é L.I.

Como V_1, V_2 e V_3 são autovetores associados a λ_1, λ_2 e λ_3, respectivamente, então pela Proposição 7.3 na página 403 os autovetores juntos V_1, V_2 e V_3 são L.I. Assim, a matriz A é diagonalizável e as matrizes

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad Q = [\, V_1\ V_2\ V_3] = \begin{bmatrix} 1 & -1 & 1 \\ -2 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix}$$

são tais que

$$D = Q^{-1}TQ \quad \text{ou} \quad T = QDQ^{-1}.$$

Assim,

$$\begin{aligned} T^k &= QD^kQ^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ -2 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & (\frac{1}{2})^k & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4} + (\frac{1}{2})^{k+1} & \frac{1}{4} & \frac{1}{4} - (\frac{1}{2})^{k+1} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} - (\frac{1}{2})^{k+1} & \frac{1}{4} & \frac{1}{4} + (\frac{1}{2})^{k+1} \end{bmatrix} \end{aligned}$$

Esta é a matriz que dá a transição entre k unidades de tempo (gerações).

7.1.4 Diagonalização de Operadores (opcional)

Definição 7.4. Dizemos que um operador linear $T : \mathbb{R}^n \to \mathbb{R}^n$ é **diagonalizável**, se existe uma base $\mathcal{C}$ de $\mathbb{R}^n$ tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é uma matriz diagonal.

O próximo resultado mostra que um operador linear é diagonalizável se, e somente se, a matriz dele em relação a uma base é diagonalizável.

Proposição 7.4. *Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ um operador linear. Seja $\mathcal{B} = \{E_1, \ldots, E_n\}$ a base canônica de $\mathbb{R}^n$. T é um operador diagonalizável se, e somente se, $[T]_{\mathcal{B}}^{\mathcal{B}}$ é uma matriz diagonalizável.*

Demonstração. Se um operador linear $T : \mathbb{R}^n \to \mathbb{R}^n$ é diagonalizável, então existe uma base $\mathcal{C}$ de $\mathbb{R}^n$ tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é diagonal. Então, pelo Corolário 6.16 na página 384 existe uma matriz P tal que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}^{\mathcal{B}}P,$$

ou seja, a matriz $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ é diagonalizável.

Reciprocamente, se a matriz $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ é diagonalizável, em que $\mathcal{B} = \{E_1, \ldots, E_n\}$, então existe uma matriz P tal que

$$D = P^{-1}AP$$

é uma matriz diagonal. Seja $\mathcal{C} = \{W_1, \ldots, W_n\}$, em que $W_j = \sum_{i=1}^{n} p_{ij} E_i$, para $j = 1, \ldots, n$. $\mathcal{C}$ é uma base de $\mathbb{R}^n$ (Exercício 6.3.9 na página 389) e

$$[T]_{\mathcal{C}}^{\mathcal{C}} = P^{-1} [T]_{\mathcal{B}}^{\mathcal{B}} P = D.$$

Portanto, T é um operador diagonalizável. ■

7.1.5 Forma Canônica de Jordan (opcional)

Já vimos (Exemplo 7.7 na página 407) que nem todo operador $T : \mathbb{R}^n \to \mathbb{R}^n$ é diagonalizável, ou seja, nem sempre existe uma base $\mathcal{C}$ de $\mathbb{R}^n$ tal que a matriz $[T]_{\mathcal{C}}^{\mathcal{C}}$ é diagonal. Entretanto, para várias aplicações, é suficiente que exista uma base $\mathcal{C}$ tal que a matriz $[T]_{\mathcal{C}}^{\mathcal{C}}$ tenha uma forma bastante próxima da forma diagonal chamada **forma canônica de Jordan**.

Definição 7.5. Uma matriz J, $n \times n$, está na **forma canônica de Jordan**, se ela é da forma

$$J = \begin{bmatrix} J_{\lambda_1} & \bar{0} & \cdots & \bar{0} \\ \bar{0} & J_{\lambda_2} & \cdots & \bar{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{0} & \bar{0} & \cdots & J_{\lambda_m} \end{bmatrix}, \quad \text{em que } J_{\lambda_j} = \begin{bmatrix} \lambda_j & 0 & \cdots & 0 & 0 \\ 1 & \lambda_j & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_j & 0 \\ 0 & 0 & \cdots & 1 & \lambda_j \end{bmatrix},$$

para $j = 1, \ldots, m$. J_{λ_j} é chamado **bloco de Jordan**.

Exemplo 7.9. As matrizes

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 5 & 0 & 0 & 0 \\ 1 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & -3 \end{bmatrix}, \quad \begin{bmatrix} -4 & 0 & 0 & 0 \\ 1 & -4 & 0 & 0 \\ 0 & 1 & -4 & 0 \\ 0 & 0 & 1 & -4 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 7 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 7 \end{bmatrix}$$

estão na forma canônica de Jordan. A primeira é formada de dois blocos de Jordan, o primeiro sendo 3×3 e o segundo 1×1. A segunda matriz é formada de dois blocos de Jordan 2×2. A terceira, por somente um bloco e a última por 4 blocos 1×1. A matriz

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

não está na forma canônica de Jordan. Pois como os elementos da diagonal não são iguais, ela teria que ser formada por pelo menos dois blocos de Jordan e $[-1]$ deveria ser um bloco de Jordan 1×1. Entretanto, a entrada imediatamente acima de -1 não é igual a 0.

Autoespaço Generalizado

Vamos supor que exista uma base $\mathcal{C} = \{V_1^{(1)}, \ldots, V_{m_1}^{(1)}, \ldots, V_1^{(k)}, \ldots, V_{m_k}^{(k)}\}$ de $\mathbb{R}^n$ $(m_1 + \ldots + m_k = n)$ tal que a matriz do operador linear $T : \mathbb{R}^n \to \mathbb{R}^n$, esteja na forma canônica de Jordan

$$[T]_{\mathcal{C}}^{\mathcal{C}} = J = \begin{bmatrix} J_{\lambda_1} & \bar{0} & \ldots & \bar{0} \\ \bar{0} & J_{\lambda_2} & \ldots & \bar{0} \\ \vdots & & \ddots & \vdots \\ \bar{0} & \ldots & \bar{0} & J_{\lambda_k} \end{bmatrix} \quad \text{em que} \quad J_{\lambda_j} = \begin{bmatrix} \lambda_j & 0 & 0 & \cdots & 0 & 0 \\ 1 & \lambda_j & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_j & 0 \\ 0 & 0 & 0 & \cdots & 1 & \lambda_j \end{bmatrix}_{m_j \times m_j}.$$

Assim,

$$\begin{aligned} T(V_i^{(j)}) &= \lambda_j V_i^{(j)} + V_{i+1}^{(j)} \quad \text{ou} \quad (T - \lambda_j I)(V_i^{(j)}) = V_{i+1}^{(j)} \quad \text{e} \\ T(V_{m_j}^{(j)}) &= \lambda_j V_{m_j}^{(j)} \end{aligned} \tag{7.13}$$

para $i = 1, \ldots, m_j - 1$ e $j = 1, \ldots, k$. Ou seja,

$$\begin{aligned} (T - \lambda_j I)(V_{m_j}^{(j)}) &= \bar{0}, \\ (T - \lambda_j I)^2(V_{m_j-1}^{(j)}) &= (T - \lambda_j I)(V_{m_j}^{(j)}) = \bar{0}, \\ \vdots \quad & \quad \vdots \\ (T - \lambda_j I)^{m_j}(V_1^{(j)}) &= (T - \lambda_j I)^{m_j-1}(V_2^{(j)}) = \bar{0}. \end{aligned} \tag{7.14}$$

Portanto, os vetores da base $\mathcal{C}$, $V_i^{(j)}$, e os elementos da diagonal de J, λ_j, satisfazem as equações

$$(T - \lambda I)^k(V) = \bar{0},$$

para $k = 1, \ldots, m_j$. Isto motiva a seguinte definição.

Definição 7.6. Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ um operador linear. Um vetor $V \in \mathbb{R}^n$ não nulo é chamado **autovetor generalizado** de T associado ao escalar λ se

$$(T - \lambda I)^k(V) = \bar{0}, \quad \text{para algum inteiro positivo } k.$$

Observe que se V é um autovetor generalizado associado a λ e p é o menor inteiro positivo tal que $(T - \lambda I)^p(V) = \bar{0}$, então $(T - \lambda I)^{p-1}(V)$ é um autovetor de T associado a λ. Portanto, λ é um autovalor de T. Além disso, $(T - \lambda I)^q(V) = \bar{0}$, para todo $q \geq p$.

Definição 7.7. Seja λ um autovalor de um operador $T : \mathbb{R}^n \to \mathbb{R}^n$. O **autoespaço generalizado** associado a λ, é o subconjunto definido por

$$\mathbb{W}_\lambda = \{V \in \mathbb{R}^n \mid (T - \lambda I)^k(V) = \bar{0}, \text{ para algum inteiro } k > 0\}.$$

Observe que $\mathbb{W}_\lambda$ consiste dos autovetores generalizados acrescentado o vetor nulo e contém o autoespaço associado a λ. Além disso $\mathbb{W}_\lambda$ é um subespaço (verifique!) . Assim, temos a seguinte sequência de subespaços encaixados:

$$\mathcal{N}(T - \lambda I_n) \subseteq \mathcal{N}(T - \lambda I_n)^2 \subseteq \cdots \subseteq \mathbb{W}_\lambda \subseteq \mathbb{R}^n.$$

O seguinte resultado é muito útil na determinação das possíveis formas canônicas de Jordan de um operador. A demonstração pode ser encontrada por exemplo em [26].

Teorema 7.5. *Seja $T : \mathbb{V} \to \mathbb{V}$ um operador tal que seu polinômio característico é da forma*

$$p(t) = (-1)^n (t - \lambda_1)^{n_1} \ldots (t - \lambda_k)^{n_k},$$

com $\lambda_1, \ldots, \lambda_k$ distintos. Então, $\dim(\mathbb{W}_i) = n_i$, para $i = 1, \ldots, k$.

Exemplo 7.10. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ o operador definido por $T(X) = AX$, para todo $X \in \mathbb{R}^3$, em que

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

O polinômio característico de T é dado por $p(\lambda) = \det(A - \lambda I_3) = -(\lambda - 3)^3$. Assim, o único autovalor de T é $\lambda = 3$. A forma escalonada reduzida de

$$A - 3I_3 = \begin{bmatrix} 2 & 1 & 0 \\ -4 & -2 & 0 \\ -2 & -1 & 0 \end{bmatrix} \quad \text{é} \quad \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, $\mathcal{N}(T - 3I) = \{(-\beta, 2\beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$.
Agora, $(T - 3I)^2 = O$, a transformação linear nula. Portanto,

$$\mathbb{W}_3 = \mathcal{N}(A - 3I_3)^2 = \mathbb{R}^3.$$

Assim existe uma matriz P invertível tal que

$$P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

que está na forma canônica de Jordan, ou seja, a menos de ordenação dos blocos esta é a única forma canônica de Jordan do operador T.

Exemplo 7.11. Seja $T : \mathbb{R}^4 \to \mathbb{R}^4$ o operador definido por $T(X) = AX$, para todo $X \in \mathbb{R}^4$, em que

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

O polinômio característico de T é dado por $p(\lambda) = \det(A - \lambda I_4) = (\lambda - 3)(\lambda - 2)^3$. Assim os autovalores de T são $\lambda_1 = 3$ e $\lambda_2 = 2$. A forma escalonada reduzida de

$$A - 3I_4 = \begin{bmatrix} -1 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \quad \text{é} \quad \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Assim, pelo Teorema 7.5, $\mathbb{W}_3 = \mathcal{N}(T - 3I) = \{(\alpha, 0, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. A forma escalonada reduzida de

$$A - 2I_4 = \begin{bmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \quad \text{é} \quad \begin{bmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Assim, $\mathcal{N}(T - 2I) = \{(\beta, \alpha, \alpha, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$. A forma escalonada reduzida de

$$(A - 2I_4)^2 = \begin{bmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{bmatrix} \quad \text{é} \quad \begin{bmatrix} 0 & 1 & -1/2 & -1/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Assim, pelo Teorema 7.5, $\mathbb{W}_2 = \mathcal{N}(T-2I)^2 = \{(\gamma, \alpha+\beta, 2\beta, 2\alpha) \mid \alpha, \beta, \gamma \in \mathbb{R}\}$.
Assim existe uma matriz P tal que

$$P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

que está na forma canônica de Jordan. E a menos da ordem dos blocos esta é a única forma canônica de Jordan possível do operador T.

Exercícios Numéricos (respostas na página 563)

7.1.1. Ache o polinômio característico, os autovalores e os autovetores de cada matriz:

(a) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 0 & 0 \\ -1 & 3 & 0 \\ 3 & 2 & -2 \end{bmatrix}$

(e) $\begin{bmatrix} 2 & -2 & 3 \\ 0 & 3 & -2 \\ 0 & -1 & 2 \end{bmatrix}$

(f) $\begin{bmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{bmatrix}$

7.1.2. Ache bases para os auto-espaços associados a cada autovalor

(a) $\begin{bmatrix} 2 & 0 & 0 \\ 3 & -1 & 0 \\ 0 & 4 & 3 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

(d) $\begin{bmatrix} 2 & 2 & 3 & 4 \\ 0 & 2 & 3 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

7.1.3. Verifique quais das matrizes são diagonalizáveis:

(a) $\begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 1 & -2 \\ 4 & 0 & 4 \\ 1 & -1 & 4 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$

7.1.4. Ache para cada matriz A, se possível, uma matriz não-singular P tal que $P^{-1}AP$ seja diagonal:

(a) $\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix}$

(b) $\begin{bmatrix} 4 & 2 & 3 \\ 2 & 1 & 2 \\ -1 & -2 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 2 & 1 & 2 \end{bmatrix}$

(d) $\begin{bmatrix} 3 & -2 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

7.1.5. Sabendo-se que $V_1 = (-4, -4, -1)$, $V_2 = (5, 4, 1)$ e $V_3 = (5, 3, 1)$ são autovetores da matriz

$$A = \begin{bmatrix} -\frac{1}{3} & -\frac{5}{6} & \frac{20}{3} \\ -\frac{2}{3} & -\frac{1}{6} & \frac{16}{3} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{11}{6} \end{bmatrix}$$

(a) Sem obter o polinômio característico determine os autovalores correspondentes a estes autovetores.

(b) A matriz é diagonalizável? Justifique?

7.1.6. Dê exemplo de:

(a) Uma matriz que não tem autovalor (real) (Sugestão: use o Exercício 29 na página 426).

(b) Uma matriz que tem um autovalor e não é diagonalizável (em $\mathbb{R}^n$).

(c) Uma matriz que tem dois autovalores e não é diagonalizável (em $\mathbb{R}^n$).

7.1.7. (Relativo a subseção 7.1.5) Quais das seguintes matrizes estão na forma canônica de Jordan.

(a) $\begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$

7.1.8. (Relativo a subseção 7.1.5) Determine as possíveis formas canônicas de Jordan.

(a) $\begin{bmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} -1 & 1 & -1 & -3 & -1 & 7 \\ 0 & -1 & 1 & 2 & 3 & 2 \\ 0 & 0 & -1 & 0 & -2 & 1 \\ 0 & 0 & 0 & -1 & 1 & -2 \\ 0 & 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 & -4 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 1 & 0 & 0 & -1 & 0 & 4 & 0 \\ 0 & 1 & 1 & -1 & -1 & -3 & 3 & -4 \\ 0 & 0 & 1 & 0 & 1 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & -4 & -5 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$

Exercícios usando o MATLAB®

`>> A=[A1,...,An]` cria uma matriz `A` formada pelas matrizes, definidas anteriormente, `A1, ..., An` colocadas uma ao lado da outra;

`>> solve(expr)` determina a solução da equação `expr=0`. Por exemplo, `>> solve(x^2-4)` determina as soluções da equação $x^2 - 4 = 0$;

`>> subs(expr,x,num)` substitui na expressão `expr` a variável `x` por `num`.

`>> [P,D]=eig(A)` determina matrizes `P` e `D` (diagonal) tais que `AP=PD`.

`inv(A)` calcula a inversa da matriz `A`.

`A=sym(A)` converte a matriz `A` numa matriz em que os elementos são armazenados no formato simbólico. A função `numeric` faz o processo inverso.

Comandos do pacote GAAL:

`>> A=randi(n)` ou `>> A=randi(m,n)` cria uma matriz `n` por `n` ou `m` por `n`, respectivamente, com elementos inteiros aleatórios.

`>> escalona(A)` calcula passo a passo a forma reduzida escalonada da matriz `A`.

`>> null(A)` determina uma base para o espaço solução do sistema homogêneo $AX = \bar{0}$.

7.1.9. Defina as matrizes `B=sym(randi(2))` e `A=[B-B',zeros(2,1);zeros(1,2),randi]`. A matriz `A` é diagonalizável? Por que?

7.1.10. Defina as matrizes `L=[eye(2),zeros(2,1);randi(1,2),0]` e `A=sym(L*L')`. Determine o polinômio característico de `A`, os autovalores e um conjunto de autovetores linearmente independentes com o maior número possível de vetores. Encontre matrizes `P` e `D` (diagonal) tais que `inv(P)*A*P=D`, se possível. Verifique o resultado. Use o comando `[P,D]=eig(A)` e compare com as matrizes que você encontrou.

7.1.11. Defina `a=randi,b=randi` e `A=sym([2*a,a-b,a-b;0,a+b,b-a;0,b-a,a+b])`. Determine o polinômio característico de `A`, os autovalores e um conjunto de autovetores linearmente independentes com o maior número possível de vetores. Encontre matrizes `P` e `D` (diagonal) tais que `inv(P)*A*P=D`, se possível. Verifique o resultado. Use o comando `[P,D]=eig(A)` e compare com as matrizes que você encontrou.

7.1.12. Defina `a=randi,b=randi` e `A=sym([a,0,b;2*b,a-b,2*b;b,0,a])`. Determine o polinômio característico de `A`, os autovalores e um conjunto de autovetores linearmente independentes com o maior número possível de vetores. Encontre matrizes `P` e `D` (diagonal) tais que `inv(P)*A*P=D`, se possível. Verifique o resultado. Use o comando `[P,D]=eig(A)` e compare com as matrizes que você encontrou.

Exercícios Teóricos

7.1.13. Dizemos que uma matriz B, $n \times n$, é **semelhante** a uma matriz A, $n \times n$, se existir uma matriz P não singular tal que $B = P^{-1}AP$. Demonstre:

(a) A é semelhante a A;

(b) Se A é semelhante a B, então B é semelhante a A;

(c) Se A é semelhante a B e B é semelhante a C, então A é semelhante a C.

7.1.14. Seja λ um autovalor (fixo) de A. Demonstre que o conjunto formado por todos os autovetores de A associados a λ, juntamente com o vetor nulo, é um subespaço de $\mathbb{R}^n$. Este subespaço é chamado de **autoespaço associado a** λ. Em outras palavras, combinação linear de autovetores associados a um mesmo autovalor é um autovetor associado a esse mesmo autovalor.

7.1.15. Demonstre que se A e B são semelhantes, então possuem os mesmos polinômios característicos e portanto os mesmos autovalores.

7.1.16. Demonstre que se A é uma matriz triangular superior, então os autovalores de A são os elementos da diagonal principal de A.

7.1.17. Demonstre que A e A^t possuem os mesmos autovalores. O que podemos dizer sobre os autovetores de A e A^t?

7.1.18. Seja λ um autovalor de A com autovetor associado X. Demonstre que λ^k é um autovalor de $A^k = A \ldots A$ associado a X, em que k é um inteiro positivo.

7.1.19. Uma matriz A é chamada **nilpotente** se $A^k = \bar{0}$, para algum inteiro positivo k. Demonstre que se A é nilpotente, então o único autovalor de A é 0. (Sugestão: use o exercício anterior)

7.1.20. Seja A uma matriz $n \times n$.

(a) Mostre que o determinante de A é o produto de todas as raízes do polinômio característico de A; (Sugestão: $p(t) = \det(A - t\,I_n) = (-1)^n(t - \lambda_1)\ldots(t - \lambda_n)$.)

(b) Mostre que A é singular se, e somente se, 0 for um autovalor de A.

7.1.21. Seja λ um autovalor da matriz não-singular A com autovetor associado X. Mostre que $1/\lambda$ é um autovalor de A^{-1} com autovetor associado X.

7.1.22. Seja $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Ache as condições necessárias e suficientes para que A seja diagonalizável.

7.1.23. Se V e W são autovetores associados a um autovalor λ, então $W - \text{proj}_V W$ é também um autovetor associado a λ? E se V e W forem autovetores associados a autovalores diferentes?

7.1.24. Sejam A e B matrizes $n \times n$. Mostre que AB e BA possuem os mesmos autovalores. (Sugestão: Separe em dois casos: $\lambda = 0$ e $\lambda \neq 0$. No segundo caso, mostre que se V é autovetor de AB, então BV é autovetor de BA.)

7.1.25. Seja A uma matriz $n \times n$ diagonalizável. Mostre que o traço de A é igual a soma das raízes do seu polinômio característico, incluindo as multiplicidades. (Sugestão: use o fato de que $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.)

7.1.26. Suponha que duas matrizes $n \times n$ A e B são tais que $B = \alpha A$, para um escalar $\alpha \neq 0$. Mostre que se λ é autovalor de uma matriz A, então $\alpha\lambda$ é autovalor de B.

7.1.27. Seja A uma matriz $n \times n$ com n autovalores diferentes. Mostre que A é diagonalizável.

7.1.28. (a) Mostre que se V é autovetor de A, então V é autovetor de A^k. Com qual autovalor?

(b) E se V é autovetor de A^k, então V é autovetor de A? (Sugestão: veja o que acontece com uma matriz nilpotente)

7.1.29. Dado um polinômio

$$p(t) = (-1)^n(t^n + a_{n-1}t^{n-1} + \cdots + a_0)$$

Verifique que a matriz

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}_{n\times n},$$

é tal que o seu polinômio característico é $p(t)$. Esta matriz é chamada **matriz companheira** do polinômio $p(t)$. (Sugestão: verifique para $n = 2$ e depois supondo que seja verdade para matrizes $(n-1) \times (n-1)$ mostre que é verdade para matrizes $n \times n$ expandindo em cofatores em relação a primeira coluna)

7.2 Diagonalização de Matrizes Simétricas

7.2.1 Motivação

O problema da identificação de uma **cônica** (curva no plano descrita por uma equação de 2º grau em x e y) através da sua equação é facilmente resolvido se a equação **não** possui um termo em que aparece o produto xy. Mas, ao contrário, se aparece este termo misto, temos que fazer uma mudança de coordenadas de forma que nas novas coordenadas ele não apareça. Vejamos o exemplo seguinte.

Exemplo 7.12. Considere o problema de identificar uma cônica representada pela equação

$$3x^2 + 2xy + 3y^2 = 4\,. \tag{7.15}$$

Usando matrizes, esta equação pode ser escrita como

$$[3x + y \quad x + 3y] \begin{bmatrix} x \\ y \end{bmatrix} = 4$$

ou

$$[x \quad y] \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 4$$

ou ainda,

$$X^t AX = 4\,, \tag{7.16}$$

em que

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \text{ e } X = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Como veremos adiante (Exemplo 7.14 na página 433), podemos escrever

$$A = PDP^t$$

em que

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ e } D = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}.$$

Assim, a equação (7.16) pode ser escrita como

$$(X^t P)D(P^t X) = (P^t X)^t D(P^t X) = 4\,.$$

Se fazemos a mudança de variáveis (ou de coordenadas) $X = PX'$, então como $P^t P = I_2$, a equação (7.16) se transforma em

$$X'^t D X' = 4$$

ou

$$[x' \quad y'] \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = 4$$

que pode ser reescrita como,

$$2x'^2 + 4y'^2 = 4\,,$$

ou dividindo por 4, como

$$\frac{x'^2}{2} + \frac{y'^2}{1} = 1$$

que é a equação da elipse mostrada na Figura 7.2. Veremos na próxima seção como traçar esta elipse.

A matriz P, tem a propriedade de que a sua inversa é simplesmente a sua transposta, $P^{-1} = P^t$. Uma matriz que satisfaz esta propriedade é chamada de **matriz ortogonal**. O que possibilitou a identificação da cônica, no exemplo anterior, foi o fato de que a matriz A é diagonalizável através de uma matriz ortogonal P. Ou seja, existe uma matriz P tal que $A = PDP^{-1}$ e $P^{-1} = P^t$.

Já vimos que nem toda matriz é diagonalizável (Exemplo 7.7 na página 407). Vamos ver que se uma matriz A é simétrica, então ela é diagonalizável, isto é, existe uma matriz diagonal D e uma matriz invertível P tal que $D = P^{-1}AP$. Além disso, para matrizes simétricas, existe uma matriz P tal que $D = P^tAP$. Isto porque existe uma matriz ortogonal P que faz a diagonalização, ou seja, que tem a propriedade $P^{-1} = P^t$. Em algumas aplicações a diagonalização com uma tal matriz é necessária, como por exemplo na identificação de cônicas.

Vamos em primeiro lugar, caracterizar as matrizes ortogonais.

7.2.2 Matrizes Ortogonais

Uma matriz P tal que $P^{-1} = P^t$ é chamada de matriz **ortogonal**.

Proposição 7.6. *Uma matriz P é ortogonal se, e somente se, as suas colunas formam um conjunto ortonormal de vetores.*

Demonstração. Vamos escrever $P = [U_1 \; \ldots \; U_n]$. Ou seja, $U_1, \ldots, U_n$ são as colunas de P. A inversa de P é P^t se, e somente se, $P^tP = I_n$. Mas,

$$P^tP = \begin{bmatrix} U_1^t \\ \vdots \\ U_n^t \end{bmatrix} [U_1 \; \ldots \; U_n] = \begin{bmatrix} U_1^tU_1 & U_1^tU_2 & \ldots & U_1^tU_n \\ U_2^tU_1 & U_2^tU_2 & \ldots & U_2^tU_n \\ \vdots & & \ddots & \vdots \\ U_n^tU_1 & U_n^tU_2 & \ldots & U_n^tU_n \end{bmatrix}$$

Logo, $P^tP = I_n$ se, e somente se, $U_i^tU_j = U_i \cdot U_j = 0$ para $i \neq j$ e $U_i^tU_i = U_i \cdot U_i = 1$ para $i = 1, \ldots n$. Ou seja, $P^tP = I_n$ se, e somente se, $U_1, \ldots, U_n$ são ortonormais. ■

Vamos supor que uma matriz A é diagonalizável através de uma matriz ortogonal, ou seja, que existe uma matriz P tal que $D = P^t AP$ é uma matriz diagonal. Como a matriz P é uma matriz cujas colunas são autovetores de A, deduzimos da proposição anterior que uma matriz A é diagonalizável através de uma matriz ortogonal se, e somente se, ela possui um conjunto ortonormal de autovetores. Como veremos, as matrizes simétricas possuem esta característica.

Proposição 7.7. *Para uma matriz A simétrica, os autovetores associados a autovalores diferentes são ortogonais.*

Demonstração. Sejam V_1 e V_2 autovetores de A associados aos autovalores λ_1 e λ_2, respectivamente, com $\lambda_1 \neq \lambda_2$. Então, $AV_1 = \lambda_1 V_1$ e $AV_2 = \lambda_2 V_2$.

Agora, se escrevemos os vetores como matrizes colunas, o produto escalar é simplesmente o produto matricial da transposta da primeira matriz pela segunda. Assim,

$$AV_1 \cdot V_2 = (AV_1)^t V_2 = V_1^t A^t V_2 = V_1 \cdot A^t V_2 \,. \tag{7.17}$$

Como A é simétrica $A^t = A$ e como V_1 e V_2 são autovetores de A, temos de (7.17) que

$$\lambda_1 V_1 \cdot V_2 = \lambda_2 V_1 \cdot V_2$$

ou

$$(\lambda_1 - \lambda_2) V_1 \cdot V_2 = 0 \,.$$

Como $\lambda_1 \neq \lambda_2$, concluímos que $V_1 \cdot V_2 = 0$, ou seja, V_1, V_2 são ortogonais. ■

Como autovetores associados a autovalores diferentes já são ortogonais, para diagonalizarmos uma matriz simétrica A através de uma matriz P ortogonal, precisamos encontrar, para cada autovalor, autovetores ortonormais associados a eles. Para isso, podemos aplicar o processo de ortogonalização de Gram-Schmidt a cada conjunto de autovetores L.I. associados a cada um dos autovalores.

Exemplo 7.13. Considere a matriz

$$A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

Esta é a matriz do Exemplo 7.5 na página 405. Para esta matriz o polinômio característico é

$$p(t) = \det(A - t\, I_3) = (t-2)^2(8-t)$$

Portanto os autovalores de A (raízes reais do polinômio característico) são $\lambda_1 = 2$ e $\lambda_2 = 8$.

Os autovetores associados aos autovalores $\lambda_1 = 2$ e $\lambda_2 = 8$ são as soluções de

$$(A - \lambda_1 I_3)X = \bar{0} \quad \text{e} \quad (A - \lambda_2 I_3)X = \bar{0},$$

respectivamente.

A forma escalonada reduzida de

$$A - 2I_3 = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} \quad \text{é} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Portanto o autoespaço associado a $\lambda_1 = 2$ é

$$\mathbb{W}_1 = \{(-\alpha - \beta, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\},$$

Agora, $(-\alpha - \beta, \beta, \alpha) = \alpha(-1, 0, 1) + \beta(-1, 1, 0)$. Assim, os vetores $V_1 = (-1, 0, 1)$ e $V_2 = (-1, 1, 0)$ geram $\mathbb{W}_1$. Como além disso, eles são L.I. (um não é múltiplo escalar do outro), então eles formam uma base para $\mathbb{W}_1$.

Para encontrar dois autovetores ortonormais associados a $\lambda_1 = 2$ vamos aplicar o processo de ortogonalização de Gram-Schmidt aos vetores V_1 e V_2.

$$W_1 = V_1 = (-1, 0, 1); \qquad W_2 = V_2 - \text{proj}_{W_1} V_2 = (-\frac{1}{2}, 1, -\frac{1}{2})$$

$$U_1 = \left(\frac{1}{||W_1||}\right) W_1 = (-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$$

$$U_2 = \left(\frac{1}{||W_2||}\right) W_2 = (-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}})$$

Com relação ao autovalor $\lambda_2 = 8$, temos que a forma escalonada reduzida da matriz

$$A - 8I_3 = \begin{bmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{bmatrix} \quad \text{é} \quad \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Assim, o autoespaço associado a $\lambda_2 = 8$ é

$$\mathbb{W}_2 = \{(\alpha, \alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

O conjunto $\{V_3 = (1, 1, 1)\}$ é uma base para $\mathbb{W}_2$, pois como $(\alpha, \alpha, \alpha) = \alpha(1, 1, 1)$, V_3 gera $\mathbb{W}_2$ e um vetor não nulo é L.I. Assim, o vetor

$$U_3 = \left(\frac{1}{||V_3||}\right) V_3 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$$

forma uma base ortonormal para $\mathbb{W}_2$.

Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais. Portanto, U_1, U_2 e U_3 são ortonormais e assim a matriz

$$P = [U_1 U_2 U_3] = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

satisfaz $D = P^t A P$, em que

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

Exemplo 7.14. Considere a matriz

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

O seu polinômio característico é

$$p(t) = \det(A - t\,I_2) = t^2 - 6t + 8 = (t-2)(t-4).$$

Portanto os autovalores de A são $\lambda_1 = 2$ e $\lambda_2 = 4$. Os autovetores associados aos autovalores $\lambda_1 = 2$ e $\lambda_2 = 4$ são as soluções de $(A - \lambda_1 I_2)X = \bar{0}$ e $(A - \lambda_2 I_2)X = \bar{0}$ respectivamente.

A solução geral do sistema $(A - 2I_2)X = \bar{0}$ é o autoespaço

$$\mathbb{W}_1 = \{(\alpha, -\alpha) \mid \alpha \in \mathbb{R}\}.$$

Como $(\alpha, -\alpha) = \alpha(1, -1)$, então $V_1 = (1, -1)$ gera $\mathbb{W}_1$ e como um vetor não nulo é L.I., $\{V_1\}$ é uma base de $\mathbb{W}_1$. Assim,

$$U_1 = \left(\frac{1}{||W_1||}\right) = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

é uma base ortonormal de $\mathbb{W}_1$.

A solução geral do sistema $(A - 4I_2)X = \bar{0}$ é o autoespaço

$$\mathbb{W}_2 = \{(\alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

Como $(\alpha, \alpha) = \alpha(1, 1)$, então $V_2 = (1, 1)$ gera $\mathbb{W}_2$ e como um vetor não nulo é L.I., $\{V_2\}$ é uma base de $\mathbb{W}_2$. Assim,

$$U_2 = \left(\frac{1}{||W_2||}\right) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

é uma base ortonormal de $\mathbb{W}_2$.

Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais. Portanto

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ e } D = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}.$$

são tais que $D = P^t AP$.

O próximo resultado, que não será demonstrado no momento (Apêndice IV na página 439), garante que o procedimento seguido nos dois exemplos anteriores sempre funciona, ou seja, que toda matriz simétrica é diagonalizável através de uma matriz ortogonal.

Teorema 7.8. *Se A é uma matriz simétrica, então existe uma matriz P ortogonal e uma matriz diagonal D tal que*

$$D = P^t A P\,.$$

Assim, se A é simétrica, então ela é diagonalizável.

Exercícios Numéricos (respostas na página 584)

7.2.1. Diagonalize cada matriz dada A por meio de uma matriz ortogonal, ou seja, ache uma matriz ortogonal P tal que P^tAP seja diagonal:

(a) $\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(f) $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

(g) $\begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{bmatrix}$

(h) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

7.2.2. Seja A uma matriz simétrica. Sabendo-se que

$$V_1 = (0, 2, -2, 1) \quad \text{e} \quad V_2 = (2, 1, -2, 3)$$

são autovetores de A associados a $\lambda_1 = 2$ e

$$V_3 = (-2, 0, 1, 2) \quad \text{e} \quad V_4 = (-3, -2, -1, 2)$$

são autovetores associados a $\lambda_2 = 4$ determine, se possível, uma matriz P e uma matriz diagonal D tais que $A = PDP^t$.

Exercícios Teóricos

7.2.3. Mostre que se A é uma matriz ortogonal, então $\det(A) = \pm 1$.

7.2.4. Mostre que se A e B são matrizes ortogonais, então AB é ortogonal.

7.2.5. (a) Verifique se a matriz $\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ é ortogonal;

(b) Mostre que $X = (x, y)$ é ortogonal a $V = (a, b) \neq \bar{0}$ com $||X|| = ||V||$ se, e somente se, $X = (-b, a)$ ou $X = (b, -a)$.

(c) Mostre que se A é uma matriz ortogonal 2×2, então existe um número real θ tal que

$$A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{ou} \quad A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}.$$

A primeira matriz tem determinante igual a 1 e é chamada **matriz de rotação**.

(Sugestão: Comece com uma matriz $(a_{ij})_{2\times 2}$ e use o fato de que as colunas são ortonormais. Uma das equações será $a_{11}^2 + a_{21}^2 = 1$. Faça $a_{11} = \cos\theta$ e $a_{21} = \operatorname{sen}\theta$. Use o item anterior.)

7.2.6. Mostre que se uma matriz A é diagonalizável por uma matriz ortogonal (isto é, existem P e D, com $P^{-1} = P^t$ e D diagonal, tais que $D = P^t AP$), então A é uma matriz simétrica.

7.2.7. Dizemos que uma matriz simétrica A, $n \times n$, é **(definida) positiva** se $X^t AX > 0$, para todo $X \in \mathbb{R}^n$, $X \neq \bar{0}$, X escrito como matriz coluna. Mostre que são equivalentes as seguintes afirmações:

(a) A matriz A é definida positiva.

(b) A é simétrica e todos os autovalores de A são positivos.

(c) Existe uma matriz definida positiva B tal que $A = B^2$. A matriz B é chamada a **raiz quadrada** de A.

(Sugestão: Mostre que (a)⇒(b)⇒(c)⇒(a). Na parte (b)⇒(c) faça primeiro o caso em que A é uma matriz diagonal)

7.2.8. Seja A uma matriz invertível $n \times n$. Mostre que existe uma matriz simétrica definida positiva P e uma matriz ortogonal U, tal que $A = PU$. Esta decomposição é única chamada de **decomposição polar de** A. (Sugestão: Sejam $P = (AA^t)^{\frac{1}{2}}$ e $U = P^{-1}A$. Mostre que $UU^t = I_n$.)

7.2.9. Seja A uma matriz $n \times n$. Para $k = 1, \ldots, n$, seja A_k a submatriz obtida de A eliminando-se as últimas $n - k$ linhas e colunas. A_k é chamada **submatriz principal de** A **de ordem** k. Mostre que se A é uma matriz simétrica definida positiva $n \times n$, então

(a) A é não singular;

(b) $\det(A) > 0$;

(c) as submatrizes principais $A_1, \ldots, A_n$ são todas definidas positivas. (Sugestão: considere vetores X_k tais que os últimos $n - k$ elementos são nulos.)

Apêndice IV: Autovalores Complexos

Vamos provar que toda matriz simétrica é diagonalizável através de uma matriz ortogonal. Para isto, precisamos trabalhar com matrizes cujas entradas são números complexos. Vamos chamar o conjunto das matrizes $m \times n$ cujas entradas são números complexos de $\mathcal{M}_{mn}(\mathbb{C})$.

Para uma matriz $A = (a_{ij}) \in \mathcal{M}_{mn}(\mathbb{C})$, definimos o **conjugado da matriz** A, denotado por $\overline{A}$ como sendo a matriz $B = (b_{ij}) \in \mathcal{M}_{mn}(\mathbb{C})$ dada por $b_{ij} = \bar{a}_{ij}$, em que, se $a_{ij} = \alpha_{ij} + i\beta_{ij}$, então $\bar{a}_{ij} = \alpha_{ij} - i\beta_{ij}$.

Para as matrizes de $\mathcal{M}_{mn}(\mathbb{C})$ além das propriedades que já foram demonstradas no Teorema 1.1 na página 9 são válidas as seguintes propriedades, cuja demonstração deixamos a cargo do leitor:

(p) Se $A \in \mathcal{M}_{mp}(\mathbb{C})$ e $B \in \mathcal{M}_{pn}(\mathbb{C})$, então

$$\overline{AB} = \overline{A}\,\overline{B}.$$

(q) Se $A \in \mathcal{M}_{mn}(\mathbb{C})$ e $\alpha \in \mathbb{C}$, então

$$\overline{\alpha A} = \bar{\alpha}\overline{B}.$$

Proposição 7.9. *Seja A uma matriz $n \times n$ com entradas reais. Se $Z \in \mathcal{M}_{n1}(\mathbb{C})$, é um autovetor de A associado a um autovalor complexo $\lambda = \alpha + i\beta$ com $\beta \neq 0$, ou seja, se $AZ = \lambda Z$, então $\overline{Z}$ também é um autovetor de A associado a $\overline{\lambda} = \alpha - i\beta$.*

Demonstração.

$$A\overline{Z} = \overline{A}\,\overline{Z} = \overline{(AZ)} = \overline{\lambda Z} = \bar{\lambda}\,\overline{Z}.$$

■

Teorema 7.10. *Toda matriz simétrica, cujas entradas são números reais, possui autovalor real.*

Demonstração. Seja A uma matriz simétrica, cujas entradas são números reais. Vamos mostrar que as raízes do seu polinômio característico são reais. Seja λ uma raiz do polinômio característico de A. Então o sistema linear $(A - \lambda I_n)Z = \bar{0}$ tem solução não trivial $Z \in \mathcal{M}_{n1}(\mathbb{C})$. O que implica que

$$AZ = \lambda Z.$$

Como A é uma matriz cujas entradas são números reais, pela Proposição 7.9 temos que $A\overline{Z} = \bar{\lambda}\,\overline{Z}$. Por um lado,

$$\overline{Z}^t AZ = \overline{Z}^t \lambda Z = \lambda \overline{Z}^t Z = \lambda \sum_{i=1}^{n} |z_i|^2.$$

Por outro lado

$$\overline{Z}^t AZ = \overline{Z}^t A^t Z = (A\overline{Z})^t Z = \bar{\lambda}\,\overline{Z}^t Z = \bar{\lambda} \sum_{i=1}^{n} |z_i|^2.$$

Logo, $\overline{\lambda} = \lambda$, ou seja, λ é um número real. ■

Demonstração do Teorema 7.8 na página 435. O resultado é obvio se $n = 1$. Vamos supor que o resultado seja verdadeiro para matrizes $(n-1) \times (n-1)$ e vamos provar que ele é verdadeiro para matrizes $n \times n$. Pelo Teorema 7.10 a matriz A tem um autovalor λ_1. Isto significa que existem autovetores associados a λ_1. Seja V_1 um autovetor de norma igual a 1 associado a λ_1. Sejam $V_2, \ldots, V_n$ vetores tais que $\{V_1, \ldots, V_n\}$ é uma base ortonormal de $\mathbb{R}^n$ (isto pode ser conseguido aplicando-se o processo de ortogonalização de Gram-Schmidt a uma base de $\mathbb{R}^n$ que contenha V_1.) Seja $P_1 = [\, V_1 \ \ldots \ V_n \,]$. Como $AV_1 = \lambda_1 V_1$ e $AV_2, \ldots, AV_n$ são combinações lineares de $V_1, \ldots, V_n$, temos que

$$AP_1 = [\, AV_1 \ \ldots \ AV_n \,] = [\, V_1 \ \ldots \ V_n \,] M = P_1 M, \tag{7.18}$$

em que $M = \left[\begin{array}{c|ccc} \lambda_1 & * & \ldots & * \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right]$. Multiplicando-se à esquerda (7.18) por P_1^t obtemos $M = P_1^t A P_1$. Mas, $M^t = (P_1^t A P_1)^t = P_1^t A^t P_1 = P_1^t A P_1 = M$, ou seja, a matriz M é simétrica. Portanto,

$$M = \left[\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right]$$

com B uma matriz simétrica $(n-1) \times (n-1)$. Como estamos supondo o resultado verdadeiro para matrizes $(n-1) \times (n-1)$, então existe uma matriz ortogonal $\tilde{P}_2$, $(n-1) \times (n-1)$, tal que $D_2 = \tilde{P}_2^t B \tilde{P}_2$ é diagonal. Seja $P_2 = \left[\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P}_2 & \\ 0 & & & \end{array}\right]$.

Seja $P = P_1 P_2$. P é ortogonal (verifique!) e pela equação (7.18)

$$AP = (AP_1)P_2 = P_1 M P_2 = P_1 \left[\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & B\tilde{P}_2 & \\ 0 & & & \end{array}\right]$$

Mas, $B\tilde{P}_2 = \tilde{P}_2 D_2$ e assim,

$$AP = P_1 P_2 \left[\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & D_2 & \\ 0 & & & \end{array}\right] = PD,$$

em que $D = \left[\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & D_2 & \\ 0 & & & \end{array}\right]$. Multiplicando-se à direita por P^t obtemos o resultado. ■

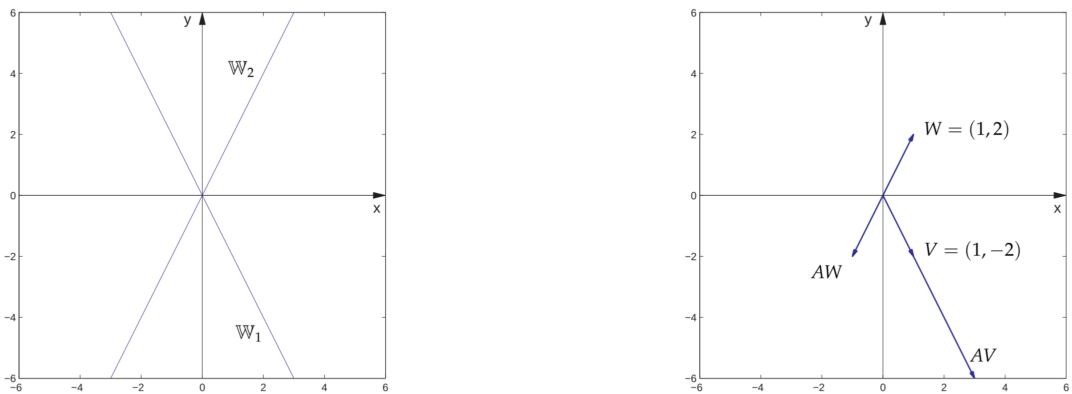


Figura 7.1. Autovetores associados a $\lambda_1 = 3$ e a $\lambda_2 = -1$ da matriz do Exemplo 7.3

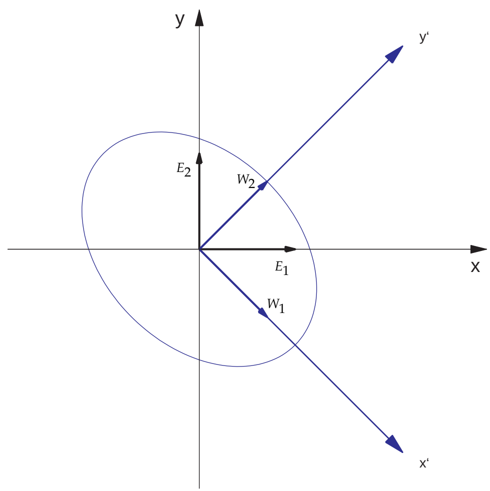


Figura 7.2. Elipse do Exemplo 7.12

7.3 Aplicação na Identificação de Cônicas

Uma **equação quadrática** nas variáveis x e y tem a forma

$$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

em que a, b, c, d, e e f são números reais, com a, b e c não simultaneamente nulos. Esta equação representa uma **(seção) cônica**, por poder ser obtida da interseção de um cone circular com um plano. As cônicas mais importantes são elipses, hipérboles e parábolas, que são chamadas de **cônicas não degeneradas**. As outras que incluem um único ponto, um par de retas, são chamadas **cônicas degeneradas**.

Dizemos que a equação de uma cônica não degenerada está na forma padrão se ela tem uma das formas dadas na Figura 7.5 na página 454.

Nesta seção veremos como a diagonalização de matrizes simétricas pode ser usada na identificação das cônicas cujas equações não estão na forma padrão.

Vamos estudar alguns exemplos.

Exemplo 7.15. Considere a cônica C cuja equação é

$$5x^2 - 4xy + 8y^2 - 36 = 0.$$

Esta equação pode ser escrita como

$$X^t AX - 36 = 0\,, \tag{7.19}$$

em que

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}.$$

O polinômio característico de A é

$$p(\lambda) = \det(A - \lambda I_2) = \det \begin{bmatrix} 5-\lambda & -2 \\ -2 & 8-\lambda \end{bmatrix} = \lambda^2 - 13\lambda + 36\,.$$

Logo, os autovalores de A são $\lambda_1 = 4$ e $\lambda_2 = 9$. Os autovetores associados a $\lambda_1 = 4$ são as soluções não nulas do sistema

$$(A - 4I_2)X = \bar{0}$$

ou

$$\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

cuja solução é

$$\mathbb{V}_1 = \{(2\alpha, \alpha) \mid \alpha \in \mathbb{R}\}\,.$$

Assim, $V_1 = (2, 1)$ é uma base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ e é L.I. E $W_1 = \frac{V_1}{||V_1||} = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})$ é uma base ortonormal para $\mathbb{V}_1$.

Os autovetores associados a $\lambda_2 = 9$ são as soluções não nulas do sistema

$$(A - 9I_2)X = \bar{0}$$

ou

$$\begin{bmatrix} -4 & -2 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

cuja solução é

$$\mathbb{V}_2 = \{(-\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}\,.$$

Assim, $V_2 = (-1, 2)$ é uma base para $\mathbb{V}_2$, pois gera $\mathbb{V}_2$ e é L.I. E

$$W_2 = \frac{V_2}{||V_2||} = (\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$$

é uma base ortonormal para $\mathbb{V}_2$. Portanto,

$$D = P^t A P$$

em que,

$$D = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}, \text{ e } P = [W_1\ W_2] = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}.$$

Vamos fazer a mudança de variáveis $X = PX'$, em que $X' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ na equação (7.19).

Substituindo $X = PX'$ na equação (7.19), obtemos

$$X'^t (P^t A P) X' - 36 = 0,$$

ou

$$X'^t D X' - 36 = 0,$$

ou

$$4x'^2 + 9y'^2 - 36 = 0,$$

ou ainda

$$\frac{x'^2}{9} + \frac{y'^2}{4} = 1 \tag{7.20}$$

que é a equação de uma elipse cujo esboço é mostrado na Figura 7.3. Para fazer o esboço do gráfico, em primeiro lugar temos que traçar os eixos x′ e y′. O eixo x′ passa pela origem, é paralelo e possui o mesmo sentido do vetor W_1, que tem coordenadas $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ em relação ao sistema de coordenadas x′y′. Assim, $W_1 = P \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, que é a primeira coluna de P. O eixo y′ passa pela origem, é paralelo e possui o mesmo sentido de W_2 que tem coordenadas $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ em relação ao sistema de coordenadas

$x'y'$. Assim, $W_2 = P \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, que é a segunda coluna de P. Depois, a partir da equação (7.20), verificamos na Figura 7.5 na página 454 a forma da curva em relação aos eixos x' e y'.

Exemplo 7.16. Considere a cônica cuja equação é dada por

$$5x^2 - 4xy + 8y^2 + \frac{20}{\sqrt{5}}x - \frac{80}{\sqrt{5}}y + 4 = 0\,.$$

Esta equação pode ser escrita como

$$X^t AX + KX + 4 = 0\,, \tag{7.21}$$

em que

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix} \ \text{e} \ K = \begin{bmatrix} \frac{20}{\sqrt{5}} & -\frac{80}{\sqrt{5}} \end{bmatrix}.$$

A matriz A é a mesma do exemplo anterior. Assim, temos que

$$D = P^t AP$$

em que,

$$D = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}, \ \text{e} \ P = [W_1\ W_2] = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}.$$

Vamos fazer a mudança de variáveis $X = PX'$, em que $X' = \begin{bmatrix} x' \\ y' \end{bmatrix}$.

Substituindo $X = PX'$ na equação (7.21), obtemos

$$X'^t(P^t AP)X' + KPX' + 4 = 0$$

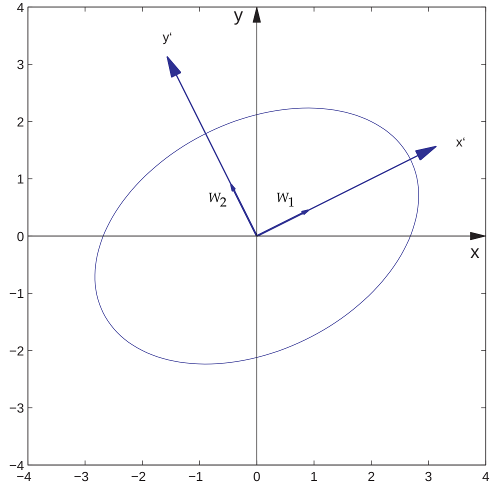


Figura 7.3. Elipse do Exemplo 7.15

ou

$$X'^t D X' + K P X' + 4 = 0,$$

ou

$$4x'^2 + 9y'^2 - 8x' - 36y' + 4 = 0\,.$$

ou ainda,

$$4(x'^2 - 2x') + 9(y'^2 - 4y') + 4 = 0$$

Completando os quadrados, obtemos

$$4[(x'^2 - 2x' + 1) - 1] + 9[(y'^2 - 4y' + 4) - 4] + 4 = 0$$

ou

$$4(x' - 1)^2 + 9(y' - 2)^2 - 36 = 0.$$

Fazendo mais uma mudança de variáveis

$$x'' = x' - 1 \quad \text{e} \tag{7.22}$$

$$y'' = y' - 2 \tag{7.23}$$

obtemos

$$4x''^2 + 9y''^2 - 36 = 0$$

ou

$$\frac{x''^2}{9} + \frac{y''^2}{4} = 1 \tag{7.24}$$

que é a equação de uma elipse cujo esboço é mostrado na Figura 7.4. Para fazer o esboço do gráfico, em primeiro lugar temos que traçar os eixos x″ e y″, que por sua vez são translações dos eixos x′ e y′. O eixo x′ tem a direção e o sentido do vetor $W_1 = P \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (a primeira coluna de P). O eixo y′ tem a direção e o sentido do vetor $W_2 = P \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ (a segunda coluna de P). O eixo x″ tem equação $y'' = 0$. Usando a equação (7.22) obtemos $y' = 2$. O eixo y$''$ tem equação $x'' = 0$. Usando a equação (7.23) obtemos $x' = 1$. Depois, a partir da equação (7.24), verificamos na Figura 7.5 na página 454 a forma da curva em relação aos eixos x$''$ e y$''$.

Os exemplos anteriores são casos particulares do próximo teorema, cuja demonstração é feita da mesma forma que fizemos com os exemplos e por isso deixamos para o leitor a tarefa de escrevê-la.

Teorema 7.11. *Considere a equação*

$$ax^2 + bxy + cy^2 + dx + ey + f = 0, \tag{7.25}$$

com $a, b, c, d, e, f \in \mathbb{R}$, sendo a, b e c não simultaneamente nulos. Então existe um sistema de coordenadas ortogonal x$'$y$'$, em que a equação (7.25) tem a forma

$$\lambda_1 x'^2 + \lambda_2 y'^2 + d'x' + e'y' + f = 0\,,$$

em que λ_1, λ_2 são os autovalores de

$$A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}.$$

Mais ainda,

$$X = PX'\,,$$

em que $X' = \begin{bmatrix} x' \\ y' \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \end{bmatrix}$ e P é uma matriz ortogonal ($P^{-1} = P^t$).

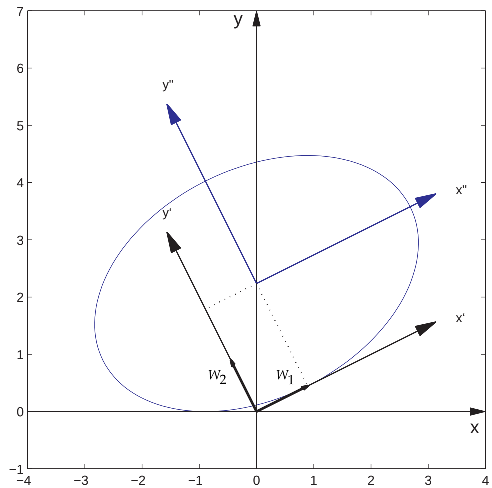
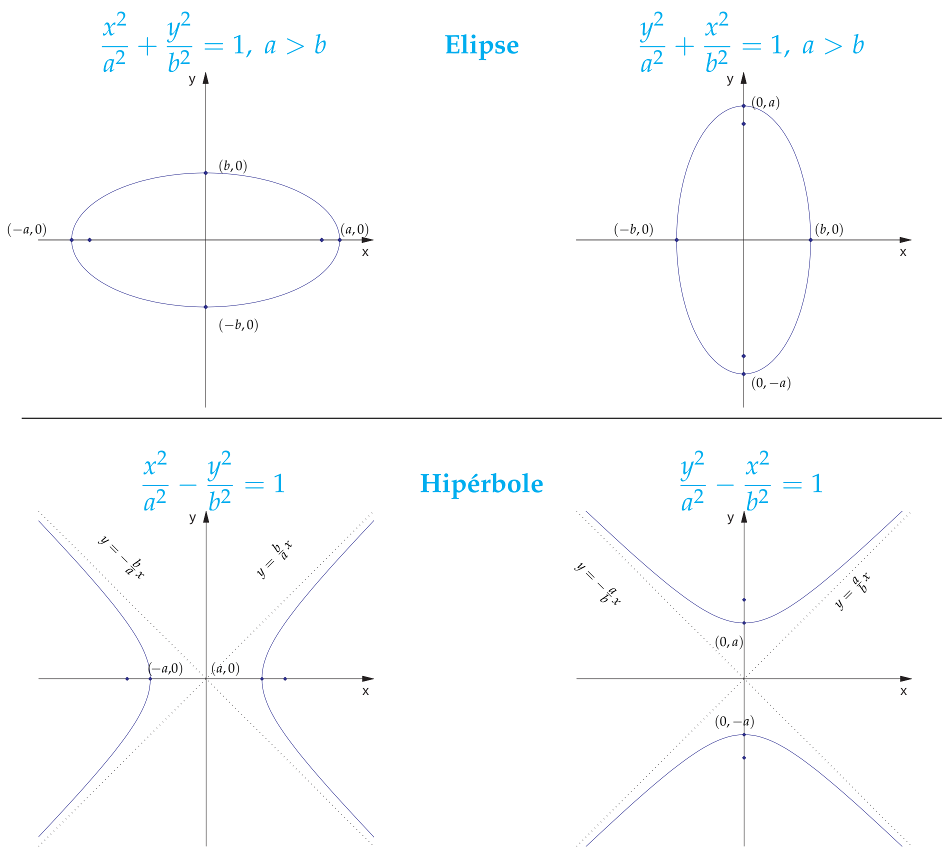


Figura 7.4. Elipse do Exemplo 7.16



$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \ a > b$
Elipse
$\frac{y^2}{a^2} + \frac{x^2}{b^2} = 1, \ a > b$
y
x
$(b, 0)$
$(-a, 0)$
$(a, 0)$
$(-b, 0)$
$(0, a)$
$(-b, 0)$
$(b, 0)$
$(0, -a)$
$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$
Hipérbole
$\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1$
$y = -\frac{b}{a}x$
$y = \frac{b}{a}x$
$(-a, 0)$
$(a, 0)$
$y = -\frac{a}{b}x$
$y = \frac{a}{b}x$
$(0, a)$
$(0, -a)$

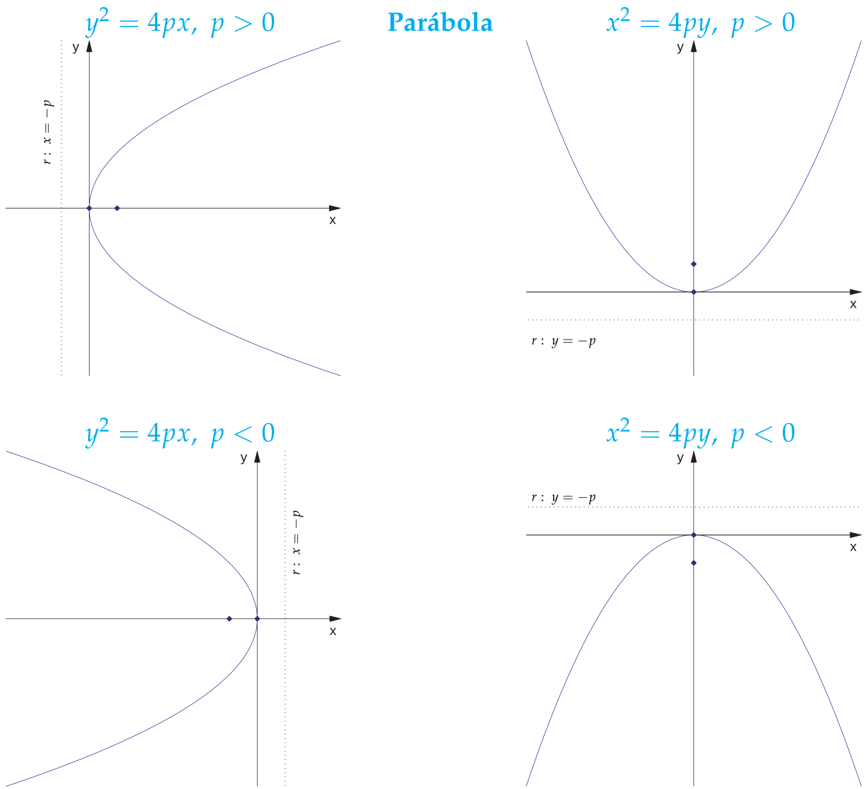


Figura 7.5. Cônicas não degeneradas com equações na forma padrão

Exercícios Numéricos (respostas na página 594)

Identificar a cônica, achar a equação no último sistema de coordenadas utilizado e fazer um esboço do gráfico.

7.3.1. $9x^2 - 4xy + 6y^2 = 30$;

7.3.2. $3x^2 - 8xy - 12y^2 + 81 = 0$;

7.3.3. $2x^2 - 4xy - y^2 = -24$;

7.3.4. $21x^2 + 6xy + 13y^2 - 132 = 0$;

7.3.5. $4x^2 - 20xy + 25y^2 - 15x - 6y = 0$;

7.3.6. $9x^2 + y^2 + 6xy - 10\sqrt{10}x + 10\sqrt{10}y + 90 = 0$;

7.3.7. $5x^2 + 5y^2 - 6xy - 30\sqrt{2}x + 18\sqrt{2}y + 82 = 0$;

7.3.8. $5x^2 + 12xy - 12\sqrt{13}x = 36$;

7.3.9. $6x^2 + 9y^2 - 4xy - 4\sqrt{5}x - 18\sqrt{5}y = 5$;

7.3.10. $x^2 - y^2 + 2\sqrt{3}xy + 6x = 0$;

7.3.11. $8x^2 + 8y^2 - 16xy + 33\sqrt{2}x - 31\sqrt{2}y + 70 = 0$;

7.3.12. $x^2 - 6xy - 7y^2 + 10x + 2y + 9 = 0$;

Exercícios usando o MATLAB®

Comandos do pacote GAAL:

`>> [P,D]=diagonal(A)` diagonaliza a matriz `A`, de forma que `AP=PD`, em que `D` é uma matriz diagonal e `P` é uma matriz ortogonal.

`>> subst(expr,[x;y],[a;b])` substitui na expressão `expr` as variáveis `x,y` por `a,b`, respectivamente.

`>> elipse(a,b)` desenha a elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

`>> elipse(a,b,[U1 U2])` desenha a elipse $\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1$, em que x' e y' são as coordenadas em relação à base ortonormal `U1` e `U2`.

`>> elipse(a,b,[U1 U2],X0)` desenha a elipse $\frac{x''^2}{a^2} + \frac{y''^2}{b^2} = 1$, em que x'' e y'' são as coordenadas em relação ao sistema de coordenadas determinado pela base ortonormal `U1` e `U2` e pelo ponto `X0`.

`>> hiperbx(a,b)` desenha a hipérbole $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

`>> hiperbx(a,b,[U1 U2])` desenha a hipérbole $\frac{x'^2}{a^2} - \frac{y'^2}{b^2} = 1$, em que x' e y' são as coordenadas em relação à base ortonormal `U1` e `U2`.

`>> hiperbx(a,b,[U1 U2],X0)` desenha a hipérbole $\frac{x''^2}{a^2} - \frac{y''^2}{b^2} = 1$, em que x'' e y'' são as coordenadas em relação ao sistema de coordenadas determinado pela base ortonormal `U1` e `U2` e pelo ponto `X0`.

`>> hiperby(a,b)` desenha a hipérbole $\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1$.

`>> hiperby(a,b,[U1 U2])` desenha a hipérbole $\frac{y'^2}{a^2} - \frac{x'^2}{b^2} = 1$, em que x' e y' são as coordenadas em relação à base ortonormal `U1` e `U2`.

`>> hiperby(a,b,[U1 U2],X0)` desenha a hipérbole $\frac{y''^2}{a^2} - \frac{x''^2}{b^2} = 1$, em que x'' e y'' são as coordenadas em relação ao sistema de coordenadas determinado pela base ortonormal `U1` e `U2` e pelo ponto `X0`.

`>> parabx(p)` desenha a parábola $y^2 = 4px$.

`>> parabx(p,[U1 U2])` desenha a parábola $y'^2 = 4px'$, em que x' e y' são as coordenadas em relação à base ortonormal `U1` e `U2`.

`>> parabx(p,[U1 U2],X0)` desenha a parábola $y''^2 = 4px''$, em que x'' e y'' são as coordenadas em relação ao sistema de coordenadas determinado pela base ortonormal `U1` e `U2` e por `X0`.

`>> paraby(p)` desenha a parábola $x^2 = 4py$.

`>> paraby(p,[U1 U2])` desenha a parábola $x'^2 = 4py'$, em que x' e y' são as coordenadas em relação à base ortonormal `U1` e `U2`.

`>> paraby(p,[U1 U2],X0)` desenha a parábola $x''^2 = 4py''$, em que x'' e y'' são as coordenadas em relação ao sistema de coordenadas determinado pela base ortonormal `U1` e `U2` e por `X0`.

7.3.13. Use o MATLAB® para resolver os **Exercícios Numéricos**

Exercícios Teóricos

7.3.14. Demonstre o Teorema 7.11 na página 451.

7.3.15. Seja C o conjunto dos pontos do plano que satisfazem a equação

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

Consideremos a matriz $A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$. Sejam λ e μ os autovalores de A.

(a) Mostre que $\lambda\mu = ac - b^2/4$.

(b) Mostre que se $\lambda\mu > 0$, então C é uma elipse, um ponto ou o conjunto vazio.

(c) Mostre que se $\lambda\mu < 0$, então C é uma hipérbole, ou um par de retas concorrentes.

(d) Mostre que se $\lambda\mu = 0$, então temos duas possibilidades:

i. Se $\lambda \neq 0$ ou $\mu \neq 0$, então C é uma parábola, um par de retas paralelas, uma reta ou o conjunto vazio.

ii. Se $\lambda = \mu = 0$, então C é uma reta. Observe que neste caso C não é uma cônica (por que?).

Teste do Capítulo

---

1. (a) Encontre matrizes P e D tais que

$$D = P^t AP,$$

em que

$$A = \begin{bmatrix} 8 & -8 \\ -8 & 8 \end{bmatrix}.$$

(b) Identificar a cônica, achar a equação no último sistema de coordenadas utilizado e fazer um esboço do gráfico.

$$8x^2 + 8y^2 - 16xy + 33\sqrt{2}x - 31\sqrt{2}y + 70 = 0$$

---

2. Verifique quais das matrizes seguintes são diagonalizáveis:

(a) $\begin{bmatrix} a & b \\ 3b & c \end{bmatrix}$ (b) $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

---

3. (a) Seja $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Calcule D^{10}.

(b) Sabendo-se que $A = P^{-1}DP$, calcule A^{10}.

---

4. Diga se é verdadeiro ou falso cada item abaixo, justificando.

(a) Se A é uma matriz 2×2 com somente 1 autovalor, então A não é diagonalizável;

(b) Se V e W são autovetores associados a um autovalor λ, então $W - \text{proj}_V W$ é também um autovetor associado a λ.

(c) Se A não é singular, então 0 não é autovalor de A;

(d) As matrizes A e A^2 possuem os mesmos autovetores;

Respostas dos Exercícios

1.1. Matrizes (página 17)

1.1.1.
```
>> A=[2,0;6,7]; B=[0,4;2,-8]; C=[-6,9,-7;7,-3,-2];
>> D=[-6,4,0;1,1,4;-6,0,6]; E=[6,9,-9;-1,0,-4;-6,0,-1];
>> A*B-B*A
   -24   -20
    58    24
>> 2*C-D
??? Error using ==> - Matrix dimensions must agree.
>> 2*D-3*E
   -30   -19    27
     5     2    20
     6     0    15
>> D*(D-E)
    80    34   -22
   -10    -4    45
```

```
    72    30   -12
```

No item (c) foram usadas as propriedades (l) e (n) do Teorema 1.1 na página 9 e no item (d) foi usada a propriedade (i).

1.1.2. $A(B+C) = AB+AC$, $B^tA^t = (AB)^t$, $C^tA^t = (AC)^t$, $(ABA)C = (AB)(AC)$.

1.1.3. (a)
```
>> A=[-3,2,1;1,2,-1];B=[2,-1;2,0;0,3];
>> C=[-2,1,-1;0,1,1;-1,0,1];
>> syms d1 d2 d3
>> D=diag([d1,d2,d3]);
>> E1=[1;0;0];E2=[0;1;0];E3=[0;0;1];
>> B*A
    -7     2     3
    -6     4     2
     3     6    -3
>> A*B
    -2     6
     6    -4
```

(b)
```
>> [A*E1-A(:,1),A*E2-A(:,2),A*E3-A(:,3)]
     0     0     0
     0     0     0
>> E1.'*B-B(1,:)
     0     0
>> E2.'*B-B(2,:)
     0     0
>> E3.'*B-B(3,:)
     0     0
```

(c)
```
>> C1=C(:,1);C2=C(:,2);C3=C(:,3);
>> C*D-[d1*C1,d2*C2,d3*C3]
[ 0, 0, 0]
```

```
[ 0, 0, 0]
[ 0, 0, 0]
```

(d)
```
>> C1=C(1,:);C2=C(2,:);C3=C(3,:);
>> D*C-[d1*C1;d2*C2;d3*C3]
[ 0, 0, 0]
[ 0, 0, 0]
[ 0, 0, 0]
```

(e)
```
>> B1=B(:,1);B2=B(:,2);
>> A*B-A*[B1,B2]
     0     0
     0     0
```

(f)
```
>> A1=A(1,:);A2=A(2,:);
>> A*B-[A1;A2]*B
     0     0
     0     0
```

1.1.4.
```
>> syms x y z
>> A=[1,-3,0;0,4,-2]; X=[x;y;z];
>> A*X
[   x-3*y]
[ 4*y-2*z]
>> x*A(:,1)+y*A(:,2)+z*A(:,3)
[   x-3*y]
[ 4*y-2*z]
```

1.1.5.
```
>> syms x
>> A=[x,4,-2]; B=[2,-3,5];
>> solve(A*B.')
11
```

1.1.6.
```
>> syms y
>> A=[1,1/y;y,1];
>> A^2-2*A
[ 0, 0]
[ 0, 0]
```

1.1.7.
```
>> syms x y z w
>> X=[x,y;z,w]; M=[0,1;-1,0];
>> X*M-M*X
[ -y-z,  x-w]
[  x-w,  z+y]
>> syms a b c d
>> A=[x,y;-y,x]; B=[a,b;-b,a];
>> A*B-B*A
[ 0, 0]
[ 0, 0]
```

1.1.8. (a) Sejam $A = \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}$ e $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

```
>> syms x y z w
>> syms a b c d
>> A=[x,0;0,y];B=[a,b;c,d];
>> A*B
[ x*a, x*b]
[ y*c, y*d]
>> B*A
[ x*a, b*y]
[ c*x, y*d]
```

Como $yb = xb$, para todo b, em particular para $b = 1$, obtemos que $y = x$. Assim, a matriz A que além de ser diagonal tem os elementos da diagonal iguais.

(b) Sejam $A = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ e $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

```
>> A=[x,y;z,w];B=[a,b;c,d];
>> A*B
[ x*a+y*c, x*b+y*d]
[ z*a+w*c, z*b+w*d]
>> B*A
[ x*a+z*b, a*y+b*w]
[ c*x+d*z, y*c+w*d]
```

Comparando os elementos de posição 1,1 obtemos que $cy = bz$, para todos os valores de b e c. Em particular para $b = 0$ e $c = 1$, obtemos que $y = 0$ e para $b = 1$ e $c = 0$, obtemos que $z = 0$. Ou seja, a matriz A tem que ser diagonal. Assim, pelo item anterior temos que a matriz A tem que ser diagonal com os elementos da diagonal iguais.

1.1.9.
```
>> A=[0,1,0;0,0,1;0,0,0];
>> A^2,A^3
ans=0     0     1
    0     0     0
    0     0     0
ans =0     0     0
     0     0     0
     0     0     0
```

1.1.10. (a)
```
>> A=[1,1/2;0,1/3]
A =
    1.0000    0.5000
         0    0.3333
>> A^2,A^3,A^4,A^5
ans =
    1.0000    0.6667
         0    0.1111
```

```
ans =
    1.0000    0.7222
         0    0.0370
ans =
    1.0000    0.7407
         0    0.0123
ans =
    1.0000    0.7469
         0    0.0041
>> A^6,A^7,A^8,A^9
ans =
    1.0000    0.7490
         0    0.0014
ans =
    1.0000    0.7497
         0    0.0005
ans =
    1.0000    0.7499
         0    0.0002
ans =
    1.0000    0.7500
         0    0.0001
```

A sequência parece estar convergindo para a matriz $\begin{bmatrix} 1 & 0.75 \\ 0 & 0 \end{bmatrix}$.

(b)
```
>> A=[1/2,1/3;0,-1/5]
A =
    0.5000    0.3333
         0   -0.2000
>> A^2,A^3,A^4,A^5
ans =
```

```
    0.2500    0.1000
         0    0.0400
ans =
    0.1250    0.0633
         0   -0.0080
ans =
    0.0625    0.0290
         0    0.0016
ans =
    0.0312    0.0150
         0   -0.0003
>> A^6,A^7,A^8,A^9
ans =
    0.0156    0.0074
         0    0.0001
ans =
    0.0078    0.0037
         0    0.0000
ans =
    0.0039    0.0019
         0    0.0000
ans =
    0.0020    0.0009
         0    0.0000
```

A sequência parece estar convergindo para a matriz nula $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

1.1.11. (a)

```
>> A=[0,0,1;1,0,0;0,1,0];
>> A=sym(A)
[ 0, 0, 1]
[ 1, 0, 0]
```

```
[ 0, 1, 0]
>> A^2
[ 0, 1, 0]
[ 0, 0, 1]
[ 1, 0, 0]
>> A^3
[ 1, 0, 0]
[ 0, 1, 0]
[ 0, 0, 1]
```

Para $k = 3$, $A^k = I_3$.

(b)
```
>> A=[0,1,0,0;-1,0,0,0;0,0,0,1;...
0,0,1,0];
>> A=sym(A)
[  0,  1,  0,  0]
[ -1,  0,  0,  0]
[  0,  0,  0,  1]
[  0,  0,  1,  0]
>> A^2
[ -1,  0,  0,  0]
[  0, -1,  0,  0]
[  0,  0,  1,  0]
[  0,  0,  0,  1]
>> A^3
[  0, -1,  0,  0]
[  1,  0,  0,  0]
[  0,  0,  0,  1]
[  0,  0,  1,  0]
>> A^4
[ 1, 0, 0, 0]
```

```
[ 0, 1, 0, 0]
[ 0, 0, 1, 0]
[ 0, 0, 0, 1]
```

Para $k = 4$, $A^k = I_4$.

(c)
```
>> A=[0,1,0,0;0,0,1,0;0,0,0,1;0,0,0,0];
>> A=sym(A)
[ 0, 1, 0, 0]
[ 0, 0, 1, 0]
[ 0, 0, 0, 1]
[ 0, 0, 0, 0]
>> A^2
[ 0, 0, 1, 0]
[ 0, 0, 0, 1]
[ 0, 0, 0, 0]
[ 0, 0, 0, 0]
>> A^3
[ 0, 0, 0, 1]
[ 0, 0, 0, 0]
[ 0, 0, 0, 0]
[ 0, 0, 0, 0]
>> A^4
[ 0, 0, 0, 0]
[ 0, 0, 0, 0]
[ 0, 0, 0, 0]
[ 0, 0, 0, 0]
```

Para $k = 4$, $A^k = \bar{0}$.

1.1.12. Concluímos que é muito raro encontrar matrizes cujo produto comute.

1.1.13. Concluímos que matrizes diagonais em geral comutam. Pode-se mostrar que elas sempre comutam

(Exercício 28 na página 27).

1.1.14. Se a matriz `A` for diagonal, então o produto comuta, se os elementos da diagonal de `A` são iguais. (ver Exercício 17 na página 24). A probabilidade de um tal par de matrizes comute é aproximadamente igual a probabilidade de que a primeira matriz tenha os elementos da sua diagonal iguais, ou seja, $11/11^3 = 1/11^2 \approx 1\%$.

1.2. Sistemas Lineares (página 57)

1.2.1. As matrizes que estão na forma reduzida escalonada são A e C.

1.2.2. (a) $X = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 8+7\alpha \\ 2-3\alpha \\ -5-\alpha \\ \alpha \end{bmatrix}$, $\forall \alpha \in \mathbb{R}$.

(b) $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2-3\alpha+6\beta \\ \beta \\ 7-4\alpha \\ 8-5\alpha \\ \alpha \end{bmatrix}$, $\forall \alpha, \beta \in \mathbb{R}$.

(c) $X = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 2-\alpha \\ \alpha \end{bmatrix}$, $\forall \alpha \in \mathbb{R}$.

(d) $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -3+8\alpha-7\beta \\ \beta \\ 5-6\alpha \\ 9-3\alpha \\ \alpha \end{bmatrix}$, $\forall \alpha, \beta \in \mathbb{R}$.

1.2.3. (a)
```
>> A=[1,1,2,8;-1,-2,3,1;3,-7,4,10];
>> escalona(A)
eliminação 1:
1*linha 1 + linha 2 ==> linha 2
-3*linha 1 + linha 3 ==> linha 3
[   1,   1,   2,   8]
[   0,  -1,   5,   9]
```

```
[   0, -10,  -2, -14]
eliminação 2:
-1*linha 2 ==> linha 2
[   1,   1,   2,   8]
[   0,   1,  -5,  -9]
[   0, -10,  -2, -14]
-1*linha 2 + linha 1 ==> linha 1
10*linha 2 + linha 3 ==> linha 3
[    1,    0,    7,   17]
[    0,    1,   -5,   -9]
[    0,    0,  -52, -104]
eliminação 3:
-1/52*linha 3 ==> linha 3
[  1,  0,  7, 17]
[  0,  1, -5, -9]
[  0,  0,  1,  2]
-7*linha 3 + linha 1 ==> linha 1
5*linha 3 + linha 2 ==> linha 2
[ 1, 0, 0, 3]
[ 0, 1, 0, 1]
[ 0, 0, 1, 2]
```

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}.$$

(b)
```
>> A=[2,2,2,0;-2,5,2,1;8,1,4,-1];
>> escalona(A)
eliminação 1:
1/2*linha 1 ==> linha 1
[  1,  1,  1,  0]
[ -2,  5,  2,  1]
```

```
[  8,  1,  4, -1]
2*linha 1 + linha 2 ==> linha 2
-8*linha 1 + linha 3 ==> linha 3
[  1,  1,  1,  0]
[  0,  7,  4,  1]
[  0, -7, -4, -1]
eliminação 2:
1/7*linha 2 ==> linha 2
[   1,   1,   1,   0]
[   0,   1, 4/7, 1/7]
[   0,  -7,  -4,  -1]
-1*linha 2 + linha 1 ==> linha 1
7*linha 2 + linha 3 ==> linha 3
[    1,    0,  3/7, -1/7]
[    0,    1,  4/7,  1/7]
[    0,    0,    0,    0]
```

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{7} - \frac{3}{7}\alpha \\ \frac{1}{7} - \frac{4}{7}\alpha \\ \alpha \end{bmatrix}, \forall \alpha \in \mathbb{R}.$$

(c)

```
>> A=[0,-2,3,1;3,6,-3,-2;6,6,3,5]
>> escalona(A)
eliminação 1:
linha 2 <==> linha 1
[  3,  6, -3, -2]
[  0, -2,  3,  1]
[  6,  6,  3,  5]
1/3*linha 1 ==> linha 1
[    1,    2,   -1, -2/3]
[    0,   -2,    3,    1]
```

```
[    6,    6,    3,    5]
-6*linha 1 + linha 3 ==> linha 3
[    1,    2,   -1, -2/3]
[    0,   -2,    3,    1]
[    0,   -6,    9,    9]
eliminação 2:
-1/2*linha 2 ==> linha 2
[    1,    2,   -1, -2/3]
[    0,    1, -3/2, -1/2]
[    0,   -6,    9,    9]
-2*linha 2 + linha 1 ==> linha 1
6*linha 2 + linha 3 ==> linha 3
[    1,    0,    2,  1/3]
[    0,    1, -3/2, -1/2]
[    0,    0,    0,    6]
```

O sistema **não** tem solução!

1.2.4.
```
>> A=[1,-2,1;2,-5,1;3,-7,2];
>> B1=[1;-2;-1];B2=[2;-1;2];
>> escalona([A,B1,B2])
eliminação 1:
-2*linha 1 + linha 2 ==> linha 2
-3*linha 1 + linha 3 ==> linha 3
[  1, -2,  1,  1,  2]
[  0, -1, -1, -4, -5]
[  0, -1, -1, -4, -4]
eliminação 2:
-1*linha 2 ==> linha 2
[  1, -2,  1,  1,  2]
[  0,  1,  1,  4,  5]
```

```
[  0, -1, -1, -4, -4]
2*linha 2 + linha 1 ==> linha 1
1*linha 2 + linha 3 ==> linha 3
[  1,  0,  3,  9, 12]
[  0,  1,  1,  4,  5]
[  0,  0,  0,  0,  1]
```

(a) $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9-3\alpha \\ 4-\alpha \\ \alpha \end{bmatrix}, \forall \alpha \in \mathbb{R}.$

(b) O sistema **não** tem solução!

1.2.5. (a)
```
>> A=[1,0,5;1,1,1;0,1,-4];
>> B=A+4*eye(3);
>> escalona([B,zeros(3,1)])
eliminação 1:
linha 2 <==> linha 1
[ 1, 5, 1, 0]
[ 5, 0, 5, 0]
[ 0, 1, 0, 0]
(-5)*linha 1 + linha 2 ==> linha 2
[   1,   5,   1,   0]
[   0, -25,   0,   0]
[   0,   1,   0,   0]
eliminação 2:
linha 3 <==> linha 2
[   1,   5,   1,   0]
[   0,   1,   0,   0]
[   0, -25,   0,   0]
(-5)*linha 2 + linha 1 ==> linha 1
(25)*linha 2 + linha 3 ==> linha 3
```

```
[ 1, 0, 1, 0]
[ 0, 1, 0, 0]
[ 0, 0, 0, 0]
```

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\alpha \\ 0 \\ \alpha \end{bmatrix}, \forall \alpha \in \mathbb{R}.$$

(b)
```
>> B=A-2*eye(3);
>> escalona([B,zeros(3,1)])
eliminação 1:
(-1)*linha 1 ==> linha 1
[  1,  0, -5,  0]
[  1, -1,  1,  0]
[  0,  1, -6,  0]
(-1)*linha 1 + linha 2 ==> linha 2
[  1,  0, -5,  0]
[  0, -1,  6,  0]
[  0,  1, -6,  0]
eliminação 2:
(-1)*linha 2 ==> linha 2
[  1,  0, -5,  0]
[  0,  1, -6,  0]
[  0,  1, -6,  0]
(-1)*linha 2 + linha 3 ==> linha 3
[  1,  0, -5,  0]
[  0,  1, -6,  0]
[  0,  0,  0,  0]
```

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5\alpha \\ 6\alpha \\ \alpha \end{bmatrix}, \forall \alpha \in \mathbb{R}.$$

1.2.6. (a)
```
>> syms a
>> A=[1,2,-3,4;3,-1,5,2;4,1,a^2-14,a+2];
>> escalona(A)
eliminação 1:
-3*linha 1 + linha 2 ==> linha 2
-4*linha 1 + linha 3 ==> linha 3
[     1,     2,    -3,     4]
[     0,    -7,    14,   -10]
[     0,    -7, a^2-2,  a-14]
eliminação 2:
-1/7*linha 2 ==> linha 2
[     1,     2,    -3,     4]
[     0,     1,    -2,  10/7]
[     0,    -7, a^2-2,  a-14]
-2*linha 2 + linha 1 ==> linha 1
7*linha 2 + linha 3 ==> linha 3
```

$$\begin{bmatrix} 1 & 0 & 1 & 8/7 \\ 0 & 1 & -2 & 10/7 \\ 0 & 0 & a^2-16 & a-4 \end{bmatrix}$$

i. Se $a^2 - 16 = 0$ e $a - 4 = 0$, então o sistema tem infinitas soluções. Neste caso, $a = 4$;

ii. Se $a^2 - 16 = 0$ e $a - 4 \neq 0$, então o sistema não tem solução. Neste caso, $a = -4$;

iii. Se $a^2 - 16 \neq 0$, então o sistema tem solução única. Neste caso, $a \neq \pm 4$;

(b)
```
>> A=[1,1,1,2;2,3,2,5;2,3,a^2-1,a+1];
>> escalona(A)
eliminação 1:
-2*linha 1 + linha 2 ==> linha 2
-2*linha 1 + linha 3 ==> linha 3
[     1,     1,     1,     2]
```

```
[      0,      1,      0,      1]
[      0,      1, a^2-3,    a-3]
eliminação 2:
-1*linha 2 + linha 1 ==> linha 1
-1*linha 2 + linha 3 ==> linha 3
```

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & a^2-3 & a-4 \end{bmatrix}$$

i. Se $a^2 - 3 = 0$ e $a - 4 = 0$, então o sistema tem infinitas soluções. Este caso não pode ocorrer;

ii. Se $a^2 - 3 = 0$ e $a - 4 \neq 0$, então o sistema não tem solução. Neste caso, $a = \pm\sqrt{3}$;

iii. Se $a^2 - 3 \neq 0$, então o sistema tem solução única. Neste caso, $a \neq \pm\sqrt{3}$;

1.2.7.

$$\begin{array}{r c} & \begin{array}{ccc} \text{X} & \text{Y} & \text{Z} \end{array} \\ \begin{array}{r} \text{gramas de A/kg} \\ \text{gramas de B/kg} \\ \text{preço/kg} \end{array} & \begin{bmatrix} 2 & 1 & 3 \\ 1 & 3 & 5 \\ 3 & 2 & 4 \end{bmatrix} \end{array}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \begin{array}{l} \text{kg de X} \\ \text{kg de Y} \\ \text{kg de Z} \end{array} \qquad \begin{bmatrix} 1900 \\ 2400 \\ 2900 \end{bmatrix} \quad \begin{array}{l} \text{gramas de A} \\ \text{gramas de B} \\ \text{arrecadação} \end{array}$$

$$\begin{bmatrix} 2 & 1 & 3 \\ 1 & 3 & 5 \\ 3 & 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1900 \\ 2400 \\ 2900 \end{bmatrix}$$

```
>> A=[2,1,3,1900;1,3,5,2400;3,2,4,2900];
>> escalona(A)
eliminação 1:
linha 2 <==> linha 1
[     1,     3,     5, 2400]
```

```
[    2,    1,    3, 1900]
[    3,    2,    4, 2900]
(-2)*linha 1 + linha 2 ==> linha 2
(-3)*linha 1 + linha 3 ==> linha 3
[     1,     3,     5,  2400]
[     0,    -5,    -7, -2900]
[     0,    -7,   -11, -4300]
eliminação 2:
(-1/5)*linha 2 ==> linha 2
[     1,     3,     5,  2400]
[     0,     1,   7/5,   580]
[     0,    -7,   -11, -4300]
(-3)*linha 2 + linha 1 ==> linha 1
(7)*linha 2 + linha 3 ==> linha 3
[    1,    0,  4/5,  660]
[    0,    1,  7/5,  580]
[    0,    0, -6/5, -240]
eliminação 3:
(-5/6)*linha 3 ==> linha 3
[   1,   0, 4/5, 660]
[   0,   1, 7/5, 580]
[   0,   0,   1, 200]
(-4/5)*linha 3 + linha 1 ==> linha 1
(-7/5)*linha 3 + linha 2 ==> linha 2
[   1,   0,   0, 500]
[   0,   1,   0, 300]
[   0,   0,   1, 200]
```

Foram vendidos 500 kg do produto X, 300 kg do produto Y e 200 kg do produto Z.

1.2.8. Substituindo os pontos na função obtemos:

$$\left\{\begin{array}{rcrcrcrcr} & & & & & & d & = & 10 \\ a & + & b & + & c & + & d & = & 7 \\ 27a & + & 9b & + & 3c & + & d & = & -11 \\ 64a & + & 16b & + & 4c & + & d & = & -14 \end{array}\right. .$$

Substituindo $d = 10$ nas outras equações e escalonando a matriz aumentada do sistema correspondente:

```
>> escalona([1,1,1,-3;27,9,3,-21;64,16,4,-24])
eliminação 1:
-27*linha 1 + linha 2 ==> linha 2
-64*linha 1 + linha 3 ==> linha 3
[   1,   1,   1,  -3]
[   0, -18, -24,  60]
[   0, -48, -60, 168]
eliminação 2:
-1/18*linha 2 ==> linha 2
[     1,      1,      1,    -3]
[     0,      1,    4/3, -10/3]
[     0,    -48,    -60,   168]
-1*linha 2 + linha 1 ==> linha 1
48*linha 2 + linha 3 ==> linha 3
[     1,      0,   -1/3,   1/3]
[     0,      1,    4/3, -10/3]
[     0,      0,      4,     8]
eliminação 3:
1/4*linha 3 ==> linha 3
[     1,      0,   -1/3,   1/3]
[     0,      1,    4/3, -10/3]
[     0,      0,      1,     2]
1/3*linha 3 + linha 1 ==> linha 1
-4/3*linha 3 + linha 2 ==> linha 2
[  1,  0,  0,  1]
```

```
[  0,  1,  0, -6]
[  0,  0,  1,  2]
```

Assim, os coeficientes são $a = 1, b = -6, c = 2$ e $d = 10$ e o polinômio $p(x) = x^3 - 6x^2 + 2x + 10$.

1.2.9. Substituindo os pontos na equação do círculo obtemos:

$$\left\{ \begin{array}{rcrcrcrcl} -2a & + & 7b & + & c & = & -[(-2)^2 + 7^2] & = & -53 \\ -4a & + & 5b & + & c & = & -[(-4)^2 + 5^2] & = & -41 \\ 4a & - & 3b & + & c & = & -[4^2 + 3^2] & = & -25 \end{array} \right. .$$

```
>> A=[-2,7,1,-53;-4,5,1,-41;4,-3,1,-25];
>> escalona(A)
eliminação 1:
-1/2*linha 1 ==> linha 1
[    1, -7/2, -1/2, 53/2]
[   -4,    5,    1,  -41]
[    4,   -3,    1,  -25]
4*linha 1 + linha 2 ==> linha 2
-4*linha 1 + linha 3 ==> linha 3
[    1, -7/2, -1/2, 53/2]
[    0,   -9,   -1,   65]
[    0,   11,    3, -131]
eliminação 2:
-1/9*linha 2 ==> linha 2
[     1,  -7/2,  -1/2,  53/2]
[     0,     1,   1/9, -65/9]
[     0,    11,     3,  -131]
7/2*linha 2 + linha 1 ==> linha 1
-11*linha 2 + linha 3 ==> linha 3
[      1,      0,   -1/9,   11/9]
[      0,      1,    1/9,  -65/9]
```

```
[      0,      0,   16/9, -464/9]
eliminação 3:
9/16*linha 3 ==> linha 3
[     1,     0,  -1/9,  11/9]
[     0,     1,   1/9, -65/9]
[     0,     0,     1,   -29]
1/9*linha 3 + linha 1 ==> linha 1
-1/9*linha 3 + linha 2 ==> linha 2
[   1,   0,   0,  -2]
[   0,   1,   0,  -4]
[   0,   0,   1, -29]
```

Os coeficientes são $a = -2, b = -4$ e $c = -29$ e a equação do círculo é $x^2 + y^2 - 2x - 4y - 29 = 0$.

1.2.10. (a)

```
>> syms b1 b2 b3
>> A=[1,-2,5,b1;4,-5,8,b2;-3,3,-3,b3];
>> escalona(A)
eliminação 1:
-4*linha 1 + linha 2 ==> linha 2
3*linha 1 + linha 3 ==> linha 3
[ 1, -2,   5,      b1]
[ 0,  3, -12, b2-4*b1]
[ 0, -3,  12, b3+3*b1]
eliminação 2:
1/3*linha 2 ==> linha 2
[ 1, -2,  5,            b1]
[ 0,  1, -4, 1/3*b2-4/3*b1]
[ 0, -3, 12,       b3+3*b1]
2*linha 2 + linha 1 ==> linha 1
3*linha 2 + linha 3 ==> linha 3
[ 1, 0, -3, -5/3*b1+2/3*b2]
```

```
[ 0, 1, -4,  1/3*b2-4/3*b1]
[ 0, 0,  0,       b3-b1+b2]
```

O sistema é consistente se, e somente se, $b_3 - b_1 + b_2 = 0$.

(b)
```
>> syms b1 b2 b3
>> A=[1,-2,-1,b1;-4,5,2,b2;-4,7,4,b3];
>> escalona(A)
eliminação 1:
4*linha 1 + linha 2 ==> linha 2
4*linha 1 + linha 3 ==> linha 3
[ 1, -2, -1,      b1]
[ 0, -3, -2, b2+4*b1]
[ 0, -1,  0, b3+4*b1]
eliminação 2:
linha 3 <==> linha 2
[ 1, -2, -1,      b1]
[ 0, -1,  0, b3+4*b1]
[ 0, -3, -2, b2+4*b1]
-1*linha 2 ==> linha 2
[ 1, -2, -1,       b1]
[ 0,  1,  0, -b3-4*b1]
[ 0, -3, -2,  b2+4*b1]
2*linha 2 + linha 1 ==> linha 1
3*linha 2 + linha 3 ==> linha 3
[ 1, 0, -1,   -7*b1-2*b3]
[ 0, 1,  0,     -b3-4*b1]
[ 0, 0, -2, b2-8*b1-3*b3]
```

O sistema é consistente para todos os valores reais de b_1, b_2 e b_3.

1.2.11. `>> A=[0,1,7,8;1,3,3,8;-2,-5,1,-8];`

```
>> escalona(A)
eliminação 1:
linha 2 <==> linha 1
[  1,  3,  3,  8]
[  0,  1,  7,  8]
[ -2, -5,  1, -8]
2*linha 1 + linha 3 ==> linha 3
[ 1, 3, 3, 8]
[ 0, 1, 7, 8]
[ 0, 1, 7, 8]
eliminação 2:
-3*linha 2 + linha 1 ==> linha 1
-1*linha 2 + linha 3 ==> linha 3
[   1,   0, -18, -16]
[   0,   1,   7,   8]
[   0,   0,   0,   0]
>> I=eye(3);E=oe(-1,2,3,I),...
F=oe(-3,2,1,I),G=oe(2,1,3,I),H=oe(I,1,2)
E =[  1,  0,  0]F =[  1, -3,  0]
   [  0,  1,  0]   [  0,  1,  0]
   [  0, -1,  1]   [  0,  0,  1]
G =[ 1, 0, 0]H =[ 0, 1, 0]
   [ 0, 1, 0]   [ 1, 0, 0]
   [ 2, 0, 1]   [ 0, 0, 1]
>> E*F*G*H*A
[   1,   0, -18, -16]
[   0,   1,   7,   8]
[   0,   0,   0,   0]
```

1.2.12. (a)
```
>> A=[1,2,0,-3,1,0,2;1,2,1,-3,1,2,3;...
1,2,0,-3,2,1,4;3,6,1,-9,4,3,9]
```

```
>> escalona(A)
[  1,  2,  0, -3,  0, -1,  0]
[  0,  0,  1,  0,  0,  2,  1]
[  0,  0,  0,  0,  1,  1,  2]
[  0,  0,  0,  0,  0,  0,  0]
```

$$\left\{\begin{array}{rrrrrrr} x_1 + 2x_2 & & -\ 3x_4 & & - & x_6 & = 0 \\ & x_3 & & & + & 2x_6 & = 1 \\ & & & x_5 & + & x_6 & = 2 \end{array}\right.$$

$X = [\alpha + 3\beta - 2\gamma \quad \gamma \quad 1 - 2\alpha \quad \beta \quad 2 - \alpha \quad \alpha]^t$,
$\forall \alpha, \beta, \gamma \in \mathbb{R}$

(b)
```
>> A=[1,3,-2,0,2,0,0;2,6,-5,-2,4,-3,-1;...
0,0,5,10,0,15,5;2,6,0,8,4,18,6]
>> escalona(A)
[    1,    3,    0,    4,    2,    0,    0]
[    0,    0,    1,    2,    0,    0,    0]
[    0,    0,    0,    0,    0,    1,  1/3]
[    0,    0,    0,    0,    0,    0,    0]
```

$$\left\{\begin{array}{rrrrrr} x_1 + 3x_2 & & +\ 4x_4 & + & 2x_5 & = 0 \\ & x_3 & +\ 2x_4 & & & = 0 \\ & & & & x_6 & = \frac{1}{3} \end{array}\right.$$

$X = [-2\alpha - 4\beta - 3\gamma \quad \gamma \quad -2\beta \quad \beta \quad \alpha \quad 1/3]^t$,
$\forall \alpha, \beta, \gamma \in \mathbb{R}$

1.2.13.
```
>> syms a, B=[4,3,1,6]';
>> A=[1,1,1,1;1,3,-2,a;
2,2*a-2,-a-2,3*a-1;3,a+2,-3,2*a+1]
>> escalona([A,B])
[ 1,  0,  0,  0, (4*a-11)/(a-5)]
[ 0,  1,  0,  0,       -4/(a-5)]
```

```
[ 0,  0,  1,  0,       -4/(a-5)]
[ 0,  0,  0,  1,       -1/(a-5)]
>> solve(-3/2*a+5/4+1/4*a^2,a)
ans = [ 1][ 5]
```

Se $a \neq 1$ e $a \neq 5$, então $X = [\frac{4a-11}{a-5} \; \frac{-4}{a-5} \; \frac{-4}{a-5} \; \frac{-1}{a-5}]^t$.

```
>> C=subs(A,a,1)
>> escalona([C,B])
[ 1, 0, 0, 1, 2]
[ 0, 1, 0, 0, 1]
[ 0, 0, 1, 0, 1]
[ 0, 0, 0, 0, 0]
```

Se $a = 1$, então $X = [2-\alpha, 1, 1, \alpha]^t \; \forall \alpha \in \mathbb{R}$.

```
>> D=subs(A,a,5)
>> escalona([D,B])
[    1,    0,  5/2,   -1,    0]
[    0,    1, -3/2,    2,    0]
[    0,    0,    0,    0,    1]
[    0,    0,    0,    0,    0]
```

Se $a = 5$, então o sistema não tem solução.

1.2.14. (a)
```
>> A=[1,2,3,1,8;1,3,0,1,7;1,0,2,1,3];
>> escalona(A)
[ 1, 0, 0, 1, 1]
[ 0, 1, 0, 0, 2]
[ 0, 0, 1, 0, 1]
```

$$\{(1-\alpha, 2, 1, \alpha) \mid \alpha \in \mathbb{R}\}$$

(b)
```
>> A=[1,1,3,-3,0;0,2,1,-3,3;1,0,2,-1,-1];
>> escalona(A)
[  1,  0,  0,  1,  1]
[  0,  1,  0, -1,  2]
[  0,  0,  1, -1, -1]
```

$$\{(1-\alpha, 2+\alpha, -1+\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$$

(c)
```
>> A=[1,2,3,0;1,1,1,0;1,1,2,0;1,3,3,0];
>> escalona(A)
[ 1, 0, 0, 0]
[ 0, 1, 0, 0]
[ 0, 0, 1, 0]
[ 0, 0, 0, 0]
```

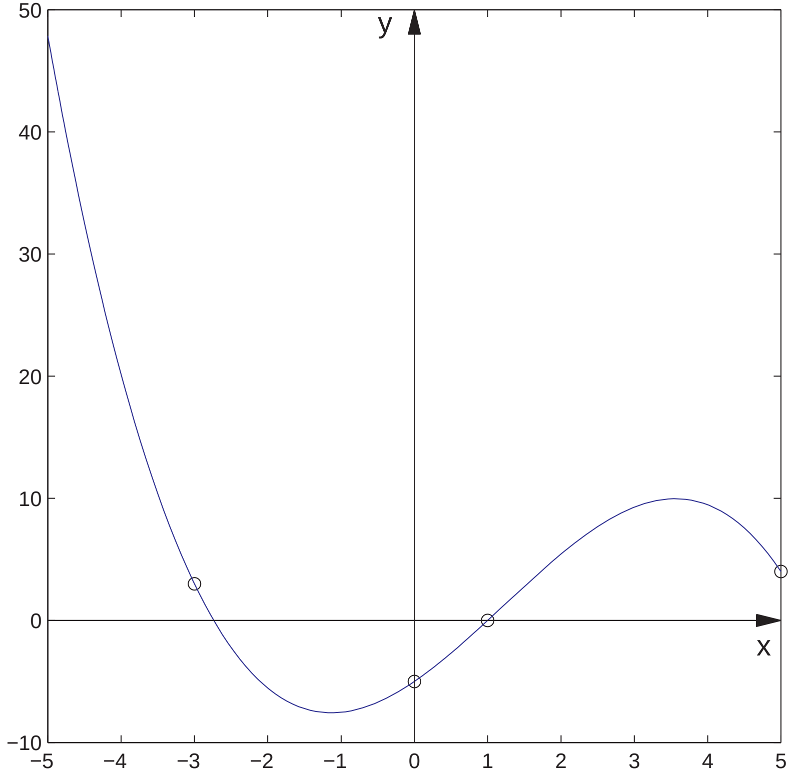
$$\{(0,0,0)\}$$

1.2.15.
```
>> P=randi(4,2)
P =  5     4
    -3     3
     1     0
     0    -5
>> A=matvand(P(:,1),3),B=P(:,2)
A =125    25     5     1
   -27     9    -3     1
     1     1     1     1
     0     0     0     1
B =  4
     3
     0
```

```
    -5
>> R=escalona([A,B])
R  = [ 1,  0,  0,  0, -163/480]
     [ 0,  1,  0,  0,    99/80]
     [ 0,  0,  1,  0, 1969/480]
     [ 0,  0,  0,  1,       -5]

>> p=poly2sym(R(:,5),x)
p = -163/480*x^3+99/80*x^2+1969/480*x-5
>> clf,po(P),syms x,plotf1(p,[-5,5])
>> eixos
```

Pode não ser possível encontrar o polinômio, se mais de um ponto tiver a mesma abscissa x_i.

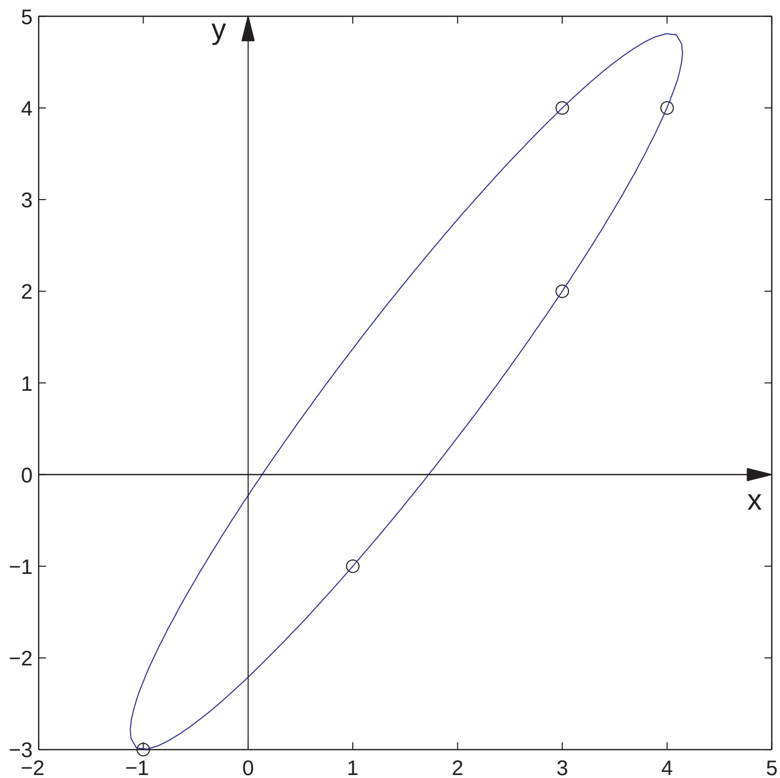


Observação. A sua resposta pode ser diferente da que está aqui.

1.2.16.
```
>> P=randi(5,2)
P =  3     2
    -1    -3
     1    -1
```

```
     3     4
     4     4
>> A=matvand(P,2)
A =  9     6     4     3     2     1
     1     3     9    -1    -3     1
     1    -1     1     1    -1     1
     9    12    16     3     4     1
    16    16    16     4     4     1
>> R=escalona([A,zeros(5,1)])
R  = [1,     0,     0,     0,     0, -35/8,     0]
     [0,     1,     0,     0,     0,  45/8,     0]
     [0,     0,     1,     0,     0,    -2,     0]
     [0,     0,     0,     1,     0,  65/8,     0]
     [0,     0,     0,     0,     1, -39/8,     0]

>> p=poly2sym2([-R(:,6);1],x,y)
p =35/8*x^2-45/8*x*y-65/8*x+1+2*y^2+39/8*y
>> clf,po(P),syms x y,
>> plotci(p,[-5,5],[-5,5])
>> eixos
```



Observação. A sua resposta pode ser diferente da que está aqui.

1.2.17. (a) A inversa da operação elementar de trocar duas linhas é ela mesma.

(b) A inversa da operação elementar de multiplicar uma linha por um escalar, $\alpha \neq 0$, é a operação de multiplicar a mesma linha pelo escalar $1/\alpha$.

(c) A inversa de somar à linha k, α vezes a linha l, é somar à linha k, $-\alpha$ vezes a linha l.

1.2.18. (a) Basta multiplicar qualquer linha da matriz pelo escalar 1.

(b) Pelo exercício anterior cada operação elementar, e, tem uma operação elementar inversa, e^{-1}, do mesmo tipo que desfaz o que a operação e fez. Se aplicando as operações elementares $e_1, \ldots, e_k$ na matriz A chegamos na matriz B, então aplicando-se as operações elementares $e_k^{-1}, \ldots, e_1^{-1}$ na matriz B chegamos na matriz A.

(c) Se aplicando as operações elementares $e_1, \ldots, e_k$ na matriz A chegamos na matriz B e aplicando as operações elementares $e_{k+1}, \ldots, e_l$ na matriz B chegamos na matriz C, então aplicando-se as operações elementares $e_1, \ldots, e_l$ na matriz A chegamos na matriz C.

2.1. Matriz Inversa (página 95)

2.1.1. A matriz é singular, pois o sistema homogêneo tem solução não trivial (Teorema 2.8 na página 84).

2.1.2. (a)
```
>> A=[1,2,3;1,1,2;0,1,2];
>> B=[A,eye(3)];
>> escalona(B)
[1, 0, 0, 0, 1,-1]
[0, 1, 0, 2,-2,-1]
[0, 0, 1,-1, 1, 1]
```

(b)
```
[1, 0, 0, 3, 2,-4]
[0, 1, 0,-1, 0, 1]
[0, 0, 1, 0,-1, 1]
```

(c)
```
[1,  0,  0,   0, 7/3,-1/3,-1/3,-2/3]
[0,  1,  0,   0, 4/9,-1/9,-4/9, 1/9]
[0,  0,  1,   0,-1/9,-2/9, 1/9, 2/9]
[0,  0,  0,   1,-5/3, 2/3, 2/3, 1/3]
```

(d)
```
[1, 0,  0,  1, -1,   0]
[0, 1,  0,3/2,1/2,-3/2]
[0, 0,  1, -1,  0,   1]
```

(e)
```
[ 1  0  1   1  0  -2 ]
[ 0  1  1   0  0   1 ]
[ 0  0  0  -1  1   1 ]

Continua ? (s/n) n
```

(f)
```
[1, 0,  0,1/4, 5/4,-3/4, 1/2, 0]
[0, 1,  0,1/2,-1/2, 1/2,   0, 0]
[0, 0,  1,1/4, 1/4, 1/4,-1/2, 0]
[0, 0,  0,  0,  -2,  -1,  -2, 1]
```

```
    Continua ? (s/n) n
```

2.1.3.
```
>> syms a
>> A=[1,1,0;1,0,0;1,2,a];
>> escalona(A)
```

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

```
Continua ?  (s/n) n
```
Para valores de a diferentes de zero a matriz A tem inversa.

2.1.4.
```
>> invA=[3,2;1,3]; invB=[2,5;3,-2];
>> invAB=invB*invA
invAB =     11            19
             7             0
```

2.1.5.
```
>> invA=[2,3;4,1]; B=[5;3];
>> X=invA*B
X =     19
        23
```

2.1.6.

$$
\begin{aligned}
A^k &= PD^kP^{-1} \\
&= \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 3^k & 0 \\ 0 & (-1)^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 3^k & (-1)^k \\ -2\,3^k & 2(-1)^k \end{bmatrix} \frac{1}{4} \begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix} \\
&= \frac{1}{4} \begin{bmatrix} 2(3^k + (-1)^k) & (-1)^k - 3^k \\ 4((-1)^k - 3^k) & 2(3^k + (-1)^k) \end{bmatrix}
\end{aligned}
$$

2.1.7.

```
>> A=[1,2,3;2,1,2;0,1,2];
>> escalona([A,eye(3)])
eliminação 1:
(-2)*linha 1 + linha 2 ==> linha 2
[  1,  2,  3,  1,  0,  0]
[  0, -3, -4, -2,  1,  0]
[  0,  1,  2,  0,  0,  1]
eliminação 2:
linha 3 <==> linha 2
[  1,  2,  3,  1,  0,  0]
[  0,  1,  2,  0,  0,  1]
[  0, -3, -4, -2,  1,  0]
(-2)*linha 2 + linha 1 ==> linha 1
(3)*linha 2 + linha 3 ==> linha 3
[  1,  0, -1,  1,  0, -2]
[  0,  1,  2,  0,  0,  1]
[  0,  0,  2, -2,  1,  3]
eliminação 3:
(1/2)*linha 3 ==> linha 3
[   1,   0,  -1,   1,   0,  -2]
```

```
[   0,   1,   2,   0,   0,   1]
[   0,   0,   1,  -1, 1/2, 3/2]
(1)*linha 3 + linha 1 ==> linha 1
(-2)*linha 3 + linha 2 ==> linha 2
[    1,    0,    0,    0,  1/2, -1/2]
[    0,    1,    0,    2,   -1,   -2]
[    0,    0,    1,   -1,  1/2,  3/2]
>> I=eye(3);E1=oe(2,1,2,I);E2=oe(I,2,3);...
E3=oe(2,2,1,I);E4=oe(-3,2,3,I);...
E5=oe(2,3,I);E6=oe(-1,3,1,I);E7=oe(2,3,2,I);
>> E1*E2*E3*E4*E5*E6*E7
     1     2     3
     2     1     2
     0     1     2
```

2.1.8.

```
>> menc=lerarq('menc1.txt'); key=lerarq('key.txt');
>> y=char2num(menc); M=char2num(key);
>> N=escalona([M,eye(3)])
>> N=N(:,6:10)
>> x=N*y;
>> num2char(x)
ans =
Desejo boa sorte a todos que estudam Álgebra Linear !
>> menc=lerarq('menc2.txt');
>> y=char2num(menc);
>> x=N*y;
>> num2char(x)
ans = Buda tinha este nome por que vivia setado!
```

Deve ser uma matriz com entradas entre 0 e 118 invertível de forma que a sua inversa seja uma matriz com entradas inteiras.

2.2. Determinantes (página 130)

2.2.1. $\det(A^2) = 9$; $\det(A^3) = -27$; $\det(A^{-1}) = -1/3$; $\det(A^t) = -3$.

2.2.2. $\det(A^t B^{-1}) = \det(A)/\det(B) = -2/3$.

2.2.3. (a) $\det\begin{bmatrix} a_{11} & a_{12} & a_{13}+a_{12} \\ a_{21} & a_{22} & a_{23}+a_{22} \\ a_{31} & a_{32} & a_{33}+a_{32} \end{bmatrix} =$

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

$\det\begin{bmatrix} a_{11} & a_{12} & a_{12} \\ a_{21} & a_{22} & a_{22} \\ a_{31} & a_{32} & a_{32} \end{bmatrix} = \det(A) + 0 = 3$

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

$\det\begin{bmatrix} a_{11} & a_{11} & a_{13} \\ a_{21} & a_{21} & a_{23} \\ a_{31} & a_{31} & a_{33} \end{bmatrix} +$

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

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

$\det\begin{bmatrix} a_{12} & -a_{12} & a_{13} \\ a_{22} & -a_{22} & a_{23} \\ a_{32} & -a_{32} & a_{33} \end{bmatrix} = -2\det(A) = -6$

2.2.4. (a) $\det \begin{bmatrix} e^{rt} & te^{rt} \\ re^{rt} & (1+rt)e^{rt} \end{bmatrix} =$
$e^{2rt} \det \begin{bmatrix} 1 & t \\ r & (1+rt) \end{bmatrix} = e^{2rt}$

(b) $\det \begin{bmatrix} \cos \beta t & \operatorname{sen} \beta t \\ \alpha \cos \beta t - \beta \operatorname{sen} \beta t & \alpha \operatorname{sen} \beta t + \beta \cos \beta t \end{bmatrix} = \alpha \det \begin{bmatrix} \cos \beta t & \operatorname{sen} \beta t \\ \cos \beta t & \operatorname{sen} \beta t \end{bmatrix} +$
$\beta \det \begin{bmatrix} \cos \beta t & \operatorname{sen} \beta t \\ -\operatorname{sen} \beta t & \cos \beta t \end{bmatrix} = \beta$

2.2.5. (a)
```
>> A=[1,-2,3,1;5,-9,6,3;-1,2,-6,-2;2,8,6,1];
>> detopelp(A)
[  1, -2,  3,  1]
[  5, -9,  6,  3]
[ -1,  2, -6, -2]
[  2,  8,  6,  1]
eliminação 1:
-5*linha 1 + linha 2 ==> linha 2
1*linha 1 + linha 3 ==> linha 3
-2*linha 1 + linha 4 ==> linha 4
[  1, -2,  3,  1]
[  0,  1, -9, -2]
[  0,  0, -3, -1]
[  0, 12,  0, -1]
eliminação 2:
-12*linha 2 + linha 4 ==> linha 4
[   1,   -2,    3,    1]
[   0,    1,   -9,   -2]
[   0,    0,   -3,   -1]
[   0,    0,  108,   23]
eliminação 3:
```

```
-1/3*linha 3 ==> linha 3
[   1,  -2,   3,   1]
[   0,   1,  -9,  -2]
[   0,   0,   1, 1/3]
[   0,   0, 108,  23]
det(A) = -3*det(A)
-108*linha 3 + linha 4 ==> linha 4
[   1,  -2,   3,   1]
[   0,   1,  -9,  -2]
[   0,   0,   1, 1/3]
[   0,   0,   0, -13]
ans = 39
```

(b)
```
>> A=[2,1,3,1;1,0,1,1;0,2,1,0;0,1,2,3];
>> detopelp(A)
[ 2, 1, 3, 1]
[ 1, 0, 1, 1]
[ 0, 2, 1, 0]
[ 0, 1, 2, 3]
eliminação 1:
linha 2 <==> linha 1
[ 1, 0, 1, 1]
[ 2, 1, 3, 1]
[ 0, 2, 1, 0]
[ 0, 1, 2, 3]
det(A) = (-1)*det(A)
-2*linha 1 + linha 2 ==> linha 2
[  1,  0,  1,  1]
[  0,  1,  1, -1]
[  0,  2,  1,  0]
[  0,  1,  2,  3]
```

```
eliminação 2:
-2*linha 2 + linha 3 ==> linha 3
-1*linha 2 + linha 4 ==> linha 4
[  1,  0,  1,  1]
[  0,  1,  1, -1]
[  0,  0, -1,  2]
[  0,  0,  1,  4]
eliminação 3:
-1*linha 3 ==> linha 3
[  1,  0,  1,  1]
[  0,  1,  1, -1]
[  0,  0,  1, -2]
[  0,  0,  1,  4]
det(A) = (-1)*(-1)*det(A)
-1*linha 3 + linha 4 ==> linha 4
[  1,  0,  1,  1]
[  0,  1,  1, -1]
[  0,  0,  1, -2]
[  0,  0,  0,  6]
ans = 6
```

2.2.6. (a)
```
>> A=[0,1,2;0,0,3;0,0,0];
>> p=det(A-x*eye(3))
p =-x^3
>> solve(p)
[0][0][0]
```

(b) `p =(1-x)*(3-x)*(-2-x)  [ 1][ 3][-2]`

(c) `p =(2-x)*(4-5*x+x^2)  [2][4][1]`

(d) `p =-8-2*x+5*x^2-x^3  [ 2][ 4][-1]`

2.2.7. (a)
```
>> A=[2,0,0;3,-1,0;0,4,3];
>> B=A-x*eye(3);
>> p=det(B)
p =(2-x)*(-1-x)*(3-x)
>> solve(p)
[ 2][-1][ 3]
```

(b)
```
p =(2-x)^2*(1-x)  [2][2][1]
```

(c)
```
p =(1-x)*(2-x)*(-1-x)*(3-x)  [ 1][ 2][-1][ 3]
```

(d)
```
p =(2-x)^2*(1-x)^2  [2][2][1][1]
```

2.2.8. (a)
```
>> Bm1=subs(B,x,-1);
>> escalona(Bm1)
[1, 0, 0]
[0, 1, 1]
[0, 0, 0]
```

$$\mathbb{W}_{-1} = \{ \begin{bmatrix} 0 \\ -\alpha \\ \alpha \end{bmatrix} | \alpha \in \mathbb{R}\}.$$

```
>> B2=subs(B,x,2);
>> escalona(B2)
[1, 0, 1/4]
[0, 1, 1/4]
[0, 0,   0]
```

$$\mathbb{W}_{2} = \{ \begin{bmatrix} -\alpha \\ -\alpha \\ 4\alpha \end{bmatrix} | \alpha \in \mathbb{R}\}.$$

```
>> B3=subs(B,x,3);
```

```
>> escalona(B3)
[1, 0, 0]
[0, 1, 0]
[0, 0, 0]
```

$$\mathbb{W}_3 = \{ \left[\begin{array}{c} 0 \\ 0 \\ \alpha \end{array}\right] | \alpha \in \mathbb{R}\}.$$

(b)
```
[1, 3, 0]
[0, 0, 1]
[0, 0, 0]
```

$$\mathbb{W}_1 = \{ \left[\begin{array}{c} -3\alpha \\ \alpha \\ 0 \end{array}\right] \mid \alpha \in \mathbb{R}\}.$$

```
[0, 1, 0]
[0, 0, 0]
[0, 0, 0]
```

$$\mathbb{W}_2 = \{ \left[\begin{array}{c} \alpha \\ 0 \\ \beta \end{array}\right] \mid \alpha, \beta \in \mathbb{R}\}.$$

(c)
```
[1, 1, 0, 0]
[0, 0, 1, 0]
[0, 0, 0, 1]
[0, 0, 0, 0]
```

$$\mathbb{W}_{-1} = \{ \left[\begin{array}{cccc} -\alpha & \alpha & 0 & 0 \end{array}\right]^t \mid \alpha \in \mathbb{R}\}.$$

```
[0, 1, 0, 0]
[0, 0, 1, 0]
[0, 0, 0, 1]
[0, 0, 0, 0]
```

$$\mathbb{W}_1 = \{ \begin{bmatrix} \alpha & 0 & 0 & 0 \end{bmatrix}^t \mid \alpha \in \mathbb{R} \}.$$

```
[1, 0, 0, 29/3]
[0, 1, 0,  7/3]
[0, 0, 1,    3]
[0, 0, 0,    0]
```

$$\mathbb{W}_2 = \{ \begin{bmatrix} -29\alpha & -7\alpha & -9\alpha & 3\alpha \end{bmatrix}^t \mid \alpha \in \mathbb{R} \}.$$

```
[1, 0, -9/4, 0]
[0, 1, -3/4, 0]
[0, 0,    0, 1]
[0, 0,    0, 0]
```

$\mathbb{W}_3 = \{ \begin{bmatrix} 9\alpha & 3\alpha & 4\alpha & 0 \end{bmatrix}^t \mid \alpha \in \mathbb{R} \}.$

(d)
```
[1, 0, -3, 0]
[0, 1,  3, 0]
[0, 0,  0, 1]
[0, 0,  0, 0]
```

$$\mathbb{W}_1 = \{ \begin{bmatrix} 3\alpha & -3\alpha & \alpha & 0 \end{bmatrix}^t \mid \alpha \in \mathbb{R} \}.$$

```
[0, 1, 0, 0]
[0, 0, 1, 0]
[0, 0, 0, 1]
[0, 0, 0, 0]
```

$$\mathbb{W}_2 = \{ \begin{bmatrix} \alpha & 0 & 0 & 0 \end{bmatrix}^t \mid \alpha \in \mathbb{R} \}.$$

2.2.9. Concluímos que é muito raro encontrar matrizes invertíveis.

3.1. Vetores no Plano e no Espaço (página 178)

3.1.1. A equação $3X - 2V = 15(X - U)$ é equivalente a $3X - 2V = 15X - 15U$. Somando-se $-15X + 2V$ obtemos $-15X + 3X = 2V - 15U$ ou $-12X = 2V - 15U$ multiplicando-se por $-\frac{1}{12}$ obtemos $X = \frac{5}{4}U - \frac{1}{6}V$.

3.1.2. Multiplicando-se a segunda equação por 2 e somando-se a primeira, obtemos $12X = 3U + 2V$ ou $X = \frac{1}{4}U + \frac{1}{6}V$. Substituindo-se X na primeira equação obtemos, $\frac{3}{2}U + V - 2Y = U$ ou $2Y = \frac{1}{2}U + V$ ou $Y = \frac{1}{4}U + \frac{1}{2}V$.

3.1.3.
```
>> OP=[  2,  3, -5]; V=[  3,  0, -3];
>> OQ=OP+V
OQ =      5      3     -8
```
As coordenadas da extremidade do segmento orientado são $(5, 3, -8)$.

3.1.4.
```
>> OP=[1,0,3]; OM=[1,2,-1];
>> MP=OP-OM;  OPlinha=OM-MP
OPlinha =      1      4     -5
```
As coordenadas de P' são $(1, 4, -5)$.

3.1.5. (a)
```
>> OA=[5,1,-3];OB=[0,3,4];OC=[0,3,-5];
>> AB=OB-OA, AC=OC-OA,
AB =     -5      2      7
AC =     -5      2     -2
```
Os pontos não são colineares, pois $\overrightarrow{AC} \neq \lambda \overrightarrow{AB}$.

(b)
```
>> OA=[-1,1,3];OB=[4,2,-3];OC=[14,4,-15];
>> AB=OB-OA, AC=OC-OA,
AB =      5      1     -6
AC =     15      3    -18
```
Os pontos são colineares, pois $\overrightarrow{AC} = 3\overrightarrow{AB}$.

3.1.6.
```
>> OA=[1,-2,-3];OB=[-5,2,-1];OC=[4,0,-1];
```

```
>> DC=OB-OA, OD=OC-DC
DC =     -6      4      2
OD =     10     -4     -3
```

O ponto é $D = (10, -4, -3)$.

3.1.7. (a) A equação $xV + yW = U$ é equivalente ao sistema $\left\{\begin{array}{rcrcr} 9x & - & y & = & -4 \\ -12x & + & 7y & = & -6 \\ -6x & + & y & = & 2 \end{array}\right.$, cuja matriz aumentada é a matriz que tem colunas V, W e U.

```
>> V=[9,-12,-6];W=[-1,7,1];U=[-4,-6,2];
>> escalona([V;W;U]')
[    1,     0, -2/3]
[    0,     1,   -2]
[    0,     0,    0]
```

Assim, $U = -2/3V - 2W$.

(b)
```
>> V=[5,4,-3];W=[2,1,1];U=[-3,-4,1];
>> escalona([V;W;U]')
[     1,      0,  -5/3]
[     0,      1,   8/3]
[     0,      0, -20/3]
```

Assim, U não é combinação linear de V e W.

3.1.8.
```
>> V=[1,2,-3]; W=[2,1,-2];
>> Va=(V+W)/no(V+W), Vb=(V-W)/no(V-W),...
>> Vc=(2*V-3*W)/no(2*V-3*W)
```

$$Va = \left[\begin{array}{ccc} \frac{3}{43}\sqrt{43} & \frac{3}{43}\sqrt{43} & -\frac{5}{43}\sqrt{43} \end{array}\right]$$
$$Vb = \left[\begin{array}{ccc} -\frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} & -\frac{1}{3}\sqrt{3} \end{array}\right]$$
$$Vc = \left[\begin{array}{ccc} -\frac{4}{17}\sqrt{17} & \frac{1}{17}\sqrt{17} & 0 \end{array}\right]$$

3.1.9.
```
>> V=[2,2,1]; W=[6,2,-3];
```

```
>> X=V/no(V)+W/no(W), U=X/no(X)
X=[32/21, 20/21, -2/21]
```

$\left[\begin{array}{ccc} \frac{16}{357}\sqrt{17}\sqrt{21} & \frac{10}{357}\sqrt{17}\sqrt{21} & -\frac{1}{357}\sqrt{17}\sqrt{21} \end{array}\right]$

3.1.10.

```
>> syms x
>> V=[x,3,4];W=[3,1,2];
>> solve(pe(V,W))
-11/3
```

Para $x = -11/3$, V e W são perpendiculares.

3.1.11.

```
>> V=[x,2,4];W=[x,-2,3];
>> pe(V,W)
x^2+8
```

A equação $x^2 + 8$ não tem solução real.

3.1.12.

```
>> Va=[2,1,0];Wa=[0,1,-1];Vb=[1,1,1];
>> Wb=[0,-2,-2];Vc=[3,3,0];Wc=[2,1,-2];
>> cosVaWa=pe(Va,Wa)/(no(Va)*no(Wa)),...
>> cosVbWb=pe(Vb,Wb)/(no(Vb)*no(Wb)),...
>> cosVcWc=pe(Vc,Wc)/(no(Vc)*no(Wc))
```

`cosVaWa=`$\frac{1}{10}\sqrt{5}\sqrt{2}$, `cosVbWb=`$-\frac{1}{3}\sqrt{3}\sqrt{2}$, `cosVcWc=`$\frac{1}{2}\sqrt{2}$. O ângulo entre Va e Wa é $\arccos(\sqrt{10}/10)$ entre Vb e Wb é $\arccos(-\sqrt{6}/3)$ e entre Vc e Wc é $\arccos(\sqrt{2}/2) = \pi/4$.

3.1.13.

```
>> W=[-1,-3,2]; V=[0,1,3];
>> W1=(pe(W,V)/pe(V,V))*V, W2=W-W1
W1 =       0          3/10        9/10
W2 =      -1        -33/10       11/10
```

3.1.14.

```
>> X=[x,y,z]; V=[1,1,0]; W=[-1,0,1]; U=[0,1,0];
>> expr1=pe(X,V), expr2=pe(X,W),...
>> expr3=pe(X,X)-3, expr4=pe(X,U)
expr1=x+y,expr2=z-x,expr3=x^2+y^2+z^2-3,expr4=y
```

```
>> solve(expr1,expr2,expr3)
S =  x: [2x1 sym]  y: [2x1 sym]  z: [2x1 sym]
>> S.x, S.y, S.z
ans =[ -1][  1] ans =[  1][ -1] ans =[ -1][  1]
```

Como y tem que ser maior que zero, $X = (-1, 1, -1)$.

3.1.15.

```
>> A=[3,0,2];B=[4,3,0];C=[8,1,-1];
>> pe(B-A,C-A), pe(A-B,C-B), pe(A-C,B-C)
14,0,21
```

Portanto o ângulo reto está no vértice B.

3.2. Equações de Retas e Planos (página 204)

3.2.1.

```
>> syms x y z
>> N=[2,-1,5]; P=[1,-2,1]; X=[x,y,z];
>> PX=X-P; expr=pe(PX,N)
expr =2*x-9-y+5*z
```

A equação do plano é $2x - y + 5z - 9 = 0$.

3.2.2.

```
>> X=[x,y,z]; P=[2,1,0]; PX=X-P
PX =[x-2, y-1, z]
>> M=[PX;1,2,-3;2,-1,4], expr=det(M)
M =[x-2, y-1, z]
   [  1,   2,-3]
   [  2,  -1, 4]  expr = 5*x-10*y-5*z
```

A equação do plano é $5x - 10y - 5z = 0$.

3.2.3.

```
>> P=[1,0,0]; Q=[1,0,1]; N1=[0,1,-1];
>> X=[x,y,z]; PQ=Q-P, PX=X-P
PQ =[0, 0, 1],PX =[x-1, y, z]
>> M=[PX;PQ;N1], expr=det(M)
M =[x-1, y, z]
```

```
   [  0, 0, 1]
   [  0, 1,-1] expr = -x+1
```

A equação do plano é $-x + 1 = 0$.

3.2.4.
```
>> V1=[2,2,1]; V2=[1,1,1]; P1=[2,0,0];
>> X=[x,y,z]; P1X=X-P1
P1X =[x-2, y, z]
>> M=[P1X;V1;V2], expr=det(M)
M =[x-2, y, z]
   [  2, 2, 1]
   [  1, 1, 1] expr = x-2-y
```

A equação do plano é $x - y - 2 = 0$.

3.2.5. (a)
```
>> solve('4=2+t'), solve('1=4-t'),...
>> solve('-1=1+2*t')
ans = 2 ans = 3 ans = -1
```

Logo não existe um valor de t tal que $P = (2, 4, 1) + t(1, -1, 2)$.

(b)
```
>> P=[4,1,-1]; Q=[2,4,1]; V=[1,-1,2];
>> X=[x,y,z];
>> PX=X-P, PQ=Q-P
PX = [x-4, y-1, z+1] PQ = [-2, 3, 2]
>> M=[PX;PQ;V], expr=dete(M)
M =[x-4,y-1,z+1]
   [ -2,  3,  2]
   [  1, -1,  2] expr = 8*x-39+6*y-z
```

A equação do plano é $8x + 6y - z - 39 = 0$.

3.2.6. Fazendo $z = 0$ nas equações dos planos π_1 e π_2 e resolvendo o sistema resultante, obtemos
```
>> expr1=x-y+1;expr2=x+y-1;
>> S=solve(expr1,expr2)
```

```
>> S.x, S.y
ans = 0 ans = 1
```

Portanto, o ponto $P = (0, 1, 0)$ pertence a π_1 e a π_2.

```
>> P=[0,1,0]; N=[1,1,1]; X=[x,y,z];
>> PX=X-P, expr=pe(PX,N)
PX =[x, y-1, z] expr = x+y-1+z
```

A equação do plano é $x + y + z - 1 = 0$.

3.2.7. (a)
```
>> N1=[1,2,-3]; N2=[1,-4,2]; V=pv(N1,N2)
V =    -8    -5    -6
```

Os planos se interceptam segundo uma reta cujo vetor diretor é $V = (-8, -5, -6)$.

(b)
```
>>  N1=[2,-1,4]; N2=[4,-2,8]; V=pv(N1,N2)
V =    0     0     0
```

Os planos são paralelos.

(c)
```
>>  N1=[1,-1,0]; N2=[1,0,1]; V=pv(N1,N2)
V =    -1    -1     1
```

Os planos se interceptam segundo uma reta cujo vetor diretor é $V = (-1, -1, 1)$.

3.2.8. $(x, y, z) = (1, 2, 1) + t(1, -1, 2)$.

3.2.9.
```
>> pv([2,3,1],[1,-1,1])
     4            -1            -5
```

$(x, y, z) = (1, 0, 1) + t(4, -1, -5)$.

3.2.10.
```
>> escalona([1,1,-1,0;2,-1,3,1])
 1      0      2/3      1/3
 0      1     -5/3     -1/3
```

A reta interseção dos planos é $(x, y, z) = (1/3, -1/3, 0) + t(-2/3, 5/3, 1)$.

```
>> A=[1,0,-1]; P=[1/3,-1/3,0];
>> V=[-2/3,5/3,1]; X=[x,y,z];
```

```
>> AX=X-A, AP=P-A
AX = [x-1, y, z+1] AP = [-2/3, -1/3, 1]
>> M=[AX;AP;V], expr=dete(M)
M =[ x-1,    y, z+1]
   [-2/3, -1/3,   1]
   [-2/3,  5/3,   1]  expr = -2*x+2/3-4/3*z
```

A equação do plano é $6x + 4z - 2 = 0$.

3.3. Os Espaços $\mathbb{R}^n$ (página 222)

3.3.1.
```
>> x1=[4,2,-3];x2=[2,1,-2];x3=[-2,-1,0];
>> x=[1,1,1];
>> A=[x1;x2;x3;x].'
     4     2    -2     1
     2     1    -1     1
    -3    -2     0     1
>> R=escalona(A)
     1     0    -2     0
     0     1     3     0
     0     0     0     1
>> x=[4,2,-6];
>> A=[x1;x2;x3;x].'
     4     2    -2     4
     2     1    -1     2
    -3    -2     0    -6
>> R=escalona(A)
     1     0    -2    -2
     0     1     3     6
     0     0     0     0
>> x=[-2,-1,1];
>> A=[x1;x2;x3;x].'
```

```
     4     2    -2    -2
     2     1    -1    -1
    -3    -2     0     1
>> R=escalona(A)
     1     0    -2    -1
     0     1     3     1
     0     0     0     0
>> x=[-1,2,3];
>> A=[x1;x2;x3;x].'
     4     2    -2    -1
     2     1    -1     2
    -3    -2     0     3
>> R=escalona(A)
     1     0    -2     0
     0     1     3     0
     0     0     0     1
```

Assim, os vetores das letras (b) e (c) são combinação linear de X_1, X_2 e X_3.

3.3.2. (a)

```
>> v1=[1,1,2];v2=[1,0,0];
>> v3=[4,6,12]
>> A=[v1;v2;v3;zeros(1,3)].'
     1     1     4     0
     1     0     6     0
     2     0    12     0
>> R=escalona(A)
     1     0     6     0
     0     1    -2     0
     0     0     0     0
```

Logo, a equação $x(1,1,2)+y(1,0,0)+z(4,6,12)=\bar{0}$ admite solução não trivial. Isto implica que os vetores do item (a) são L.D.

(b)
```
>> v1=[1,-2,3];v2=[-2,4,-6];
>> A=[v1;v2;zeros(1,3)].'
     1    -2     0
    -2     4     0
     3    -6     0
>> R=escalona(A)
     1    -2     0
     0     0     0
     0     0     0
```

Logo, a equação $x(1,-2,3)+y(-2,4,-6) = \bar{0}$ admite solução não trivial. Isto implica que os vetores da item (b) são L.D. Observe que o segundo vetor é -2 vezes o primeiro.

(c)
```
>> v1=[1,1,1];v2=[2,3,1];
>> v3=[3,1,2];
>> A=[v1;v2;v3;zeros(1,3)].'
     1     2     3     0
     1     3     1     0
     1     1     2     0
>> R=escalona(A)
     1     0     0     0
     0     1     0     0
     0     0     1     0
```

Logo, a equação $x(1,1,1)+y(2,3,1)+z(3,1,2) = \bar{0}$ só admite a solução trivial. Isto implica que os vetores do item (c) são L.I.

(d)
```
>> v1=[4,2,-1];v2=[6,5,-5];v3=[2,-1,3];
>> A=[v1;v2;v3;zeros(1,3)].'
     4     6     0
     2     5     0
    -1    -5     0
```

```
>> R=escalona(A)
     1     0     2
     0     1    -1
     0     0     0
```

Logo, o sistema $x(4,2,-1) + y(2,3,1) + z(2,-1,3) = \bar{0}$ admite solução não trivial. Isto implica que os vetores da item (d) são L.D.

3.3.3.
```
>> syms a
>> A=[3,1,0;a^2+2,2,0;0,0,0]
A =
[3, a^2+2, 0]
[1,     2, 0]
[0,     0, 0]
>> escalona(A)
eliminação 1:
linha 2 <==> linha 1
  [ 1     2    0 ]
  [              ]
  [     2        ]
  [ 3  a   + 2  0 ]
  [              ]
  [ 0     0    0 ]
Continua ? (s/n) s
-(3)*linha 1 + linha 2 ==> linha 2
  [ 1     2    0 ]
  [              ]
  [     2        ]
  [ 0  a   - 4  0 ]
  [              ]
  [ 0     0    0 ]
```

```
Continua ? (s/n) n
>> solve(a^2-4)
ans = [ 2][-2]
```

Para $\lambda = \pm 2$ o conjunto de vetores é L.D.

3.3.4. (a) $x_1 W_1 + x_2 W_2 + x_3 W_3 = x_1(V_1 + V_2) + x_2(V_1 + V_3) + x_3(V_2 + V_3) = (x_1 + x_2)V_1 + (x_1 + x_3)V_2 + (x_2 + x_3)V_3 = \bar{0}$. Como V_1, V_2 e V_3 são por hipótese L.I., os escalares que os estão multiplicando têm que ser iguais a zero. O que leva ao sistema $\left\{\begin{array}{ccccccc} x_1 & + & x_2 & & & = & 0 \\ x_1 & & & + & x_3 & = & 0 \\ & & x_2 & + & x_3 & = & 0 \end{array}\right.$

```
>> A=[1,1,0;1,0,1;0,1,1]
>> escalona(A)
[ 1, 1, 0]
[ 1, 0, 1]
[ 0, 1, 1]

[ 1, 0, 0]
[ 0, 1, 0]
[ 0, 0, 1]
```

Assim, o sistema e a equação vetorial inicial têm somente a solução trivial $x_1 = x_2 = x_3 = 0$. Portanto os vetores W_1, W_2 e W_3 são L.I.

(b) $x_1 W_1 + x_2 W_2 + x_3 W_3 = x_1 V_1 + x_2(V_1 + V_3) + x_3(V_1 + V_2 + V_3) = (x_1 + x_2 + x_3)V_1 + x_3 V_2 + (x_2 + x_3)V_3 = \bar{0}$ Como V_1, V_2 e V_3 são por hipótese L.I., os escalares que os estão multiplicando têm que ser iguais a zero. O que leva ao sistema $\left\{\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 0 \\ & & & & x_3 & = & 0 \\ & & x_2 & + & x_3 & = & 0 \end{array}\right.$ Assim, o sistema e a equação vetorial inicial têm somente a solução trivial $x_1 = x_2 = x_3 = 0$. Portanto os vetores W_1, W_2 e W_3 são L.I.

3.3.5. (a)

```
>> V1=[1;2;3]; V2=[3;4;5]; V3=[5;6;7];
```

```
>> V=randi(3,1)
V =   0
      4
      3
>> escalona([V1,V2,V3,V])
ans = 1     0     -1     0
      0     1      2     0
      0     0      0     1
```

Assim, V **não** é combinação linear de V1, V2 e V3.

(b)
```
>> M=randi(3,5)
M =  -2    -4     1     -5     5
      3    -3    -3      3     0
     -5    -3    -3     -1    -1
>> escalona([V1,V2,V3,M])
1  0 -1  0  37/13 -101/26 173/26  -96/13
0  1  2  0 -29/13   37/26 -85/26   51/13
0  0  0  1   1/13   -4/13  12/13   -4/13
```

Assim, nenhuma das colunas de M é combinação linear de V1, V2 e V3. Como as colunas de M foram geradas aleatoriamente, o mais provável é que elas não pertençam ao plano gerado por V1, V2 e V3.

(c) V3=-V1+2V2, que é a mesma relação que é válida entre as colunas de forma escalonada reduzida da matriz [V1,V2,V3,M].

4.1. Base e Dimensão (página 241)

4.1.1. (a)
```
>> A=[1,0,1,0,0;1,2,3,1,0;2,1,3,1,0]
     1     0     1     0     0
     1     2     3     1     0
     2     1     3     1     0
>> R=escalona(A)
```

```
     1     0     1     0     0
     0     1     1     0     0
     0     0     0     1     0
```

Encontramos a forma reduzida escalonada da matriz $[A \mid \bar{0}]$, que corresponde ao sistema

$$\left\{ \begin{array}{ccccccc} x_1 & & + & x_3 & & = & 0 \\ & x_2 & + & x_3 & & = & 0 \\ & & & & x_4 & = & 0 \end{array} \right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(-\alpha, -\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}\,.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(-\alpha, -\alpha, \alpha, 0) = \alpha(-1, -1, 1, 0)\,.$$

Logo, $\{V = (-1, -1, 1, 0)\}$ gera $\mathbb{W}$.

(b)
```
>> A=[1,1,2,-1,0;2,3,6,-2,0;-2,1,2,2,0]
     1     1     2    -1     0
     2     3     6    -2     0
    -2     1     2     2     0
>> R=escalona(A)
     1     0     0    -1     0
     0     1     2     0     0
     0     0     0     0     0
```

Encontramos a forma reduzida escalonada da matriz $[A \mid \bar{0}]$, que corresponde ao sistema

$$\left\{ \begin{array}{ccccccc} x_1 & & + & & - & x_4 & = & 0 \\ & x_2 & + & 2x_3 & & & = & 0 \end{array} \right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(\alpha, -2\beta, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\}\,.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$\begin{aligned}(\alpha, -2\beta, \beta, \alpha) &= \\ &= (\alpha, 0, 0, \alpha) + (0, -2\beta, \beta, 0) \\ &= \alpha(1, 0, 0, 1) + \beta(0, -2, 1, 0)\,.\end{aligned}$$

Logo, $\mathcal{B} = \{V_1 = (1, 0, 0, 1), V_2 = (0, -2, 1, 0)\}$ gera $\mathbb{W}$.

4.1.2. (a)
```
>> syms x
>> A=[0,0,1;1,0,-3;0,1,3];
>> B=A-x*eye(3)
[-x,  0,   1]
[ 1, -x,  -3]
[ 0,  1, 3-x]
>> solve(det(B))
ans = [1][1][1]
>> B1=subs(B,x,1)
    -1     0     1
     1    -1    -3
     0     1     2
>> escalona([B1,zeros(3,1)])
     1     0    -1     0
     0     1     2     0
     0     0     0     0
```

$$\left\{\begin{array}{rcrcl} x_1 & & - & x_3 & = 0 \\ & x_2 & + & 2x_3 & = 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(\alpha, -2\alpha, \alpha) \mid \alpha \in \mathbb{R}\}\,.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(\alpha, -2\alpha, \alpha) = \alpha(1, -2, 1)\,.$$

Logo, $\mathcal{B} = \{V = (1, -2, 1)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

(b)
```
>> A=[2,2,3,4;0,2,3,2;0,0,1,1;0,0,0,1]
>> B=A-x*eye(4)
[2-x,   2,   3,   4]
[  0, 2-x,   3,   2]
[  0,   0, 1-x,   1]
[  0,   0,   0, 1-x]
>> solve(det(B))
ans = [2][2][1][1]
>> B1=subs(B,x,1)
     1     2     3     4
     0     1     3     2
     0     0     0     1
     0     0     0     0
>> escalona([B1,zeros(4,1)])
     1     0    -3     0     0
     0     1     3     0     0
     0     0     0     1     0
     0     0     0     0     0
```

$$\left\{\begin{array}{rcrcrcrcl} x_1 & & & - & 3x_3 & & & = & 0 \\ & & x_2 & + & 3x_3 & & & = & 0 \\ & & & & & & x_4 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(3\alpha, -3\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(3\alpha, -3\alpha, \alpha, 0) = \alpha(3, -3, 1, 0).$$

Logo, $\mathcal{B} = \{V = (3, -3, 1, 0)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

```
>> B2=subs(B,x,2)
      0      2      3      4
      0      0      3      2
      0      0     -1      1
      0      0      0     -1
>> escalona([B2,zeros(4,1)])
      0      1      0      0      0
      0      0      1      0      0
      0      0      0      1      0
      0      0      0      0      0
```

$$\left\{\begin{array}{ccccc} x_2 & & & = & 0 \\ & x_3 & & = & 0 \\ & & x_4 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(\alpha, 0, 0, 0) \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(\alpha, 0, 0, 0) = \alpha(1, 0, 0, 0).$$

Logo, $\mathcal{B} = \{V = (1, 0, 0, 0)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

(c)
```
>> A=[1,1,-2;-1,2,1;0,1,-1]
>> B=A-x*eye(3)
[1-x,   1,   -2]
[ -1, 2-x,    1]
[  0,   1, -1-x]
>> solve(det(B))
ans = [ 1][ 2][-1]
>> Bm1=subs(B,x,-1)
     2     1    -2
    -1     3     1
     0     1     0
>> escalona([Bm1,zeros(3,1)])
     1     0    -1    0
     0     1     0    0
     0     0     0    0
```

$$\left\{ \begin{array}{ccccccc} x_1 & & - & 3x_3 & & = & 0 \\ & x_2 & & & & = & 0 \end{array} \right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}\,.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(\alpha, 0, \alpha) = \alpha(1, 0, 1)\,.$$

Logo, $\mathcal{B} = \{V = (1, 0, 1)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

```
>> B1=subs(B,x,1)
     0     1    -2
```

```
    -1      1      1
     0      1     -2
>> escalona([B1,zeros(3,1)])
     1      0     -3    0
     0      1     -2    0
     0      0      0    0
```

$$\left\{\begin{array}{rcrcrcl} x_1 & & & - & 3x_3 & = & 0 \\ & & x_2 & - & 2x_3 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(3\alpha, 2\alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(3\alpha, 2\alpha, \alpha) = \alpha(3, 2, 1).$$

Logo, $\mathcal{B} = \{V = (3, 2, 1)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

```
>> B2=subs(B,x,2)
    -1      1     -2
    -1      0      1
     0      1     -3
>> escalona([B2,zeros(3,1)])
     1      0     -1    0
     0      1     -3    0
     0      0      0    0
```

$$\left\{\begin{array}{rcrcrcl} x_1 & & & - & x_3 & = & 0 \\ & & x_2 & - & 3x_3 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(\alpha, 3\alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(\alpha, 3\alpha, \alpha) = \alpha(1, 3, 1).$$

Logo, $\mathcal{B} = \{V = (1,3,1)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

(d)
```
>> A=[-1,2,2,0;-1,2,1,0;-1,1,2,0;0,0,0,1];
>> B=A-x*eye(4)
B =
[ -1-x,    2,    2,    0]
[   -1,  2-x,    1,    0]
[   -1,    1,  2-x,    0]
[    0,    0,    0,  1-x]
>> solve(det(B))
ans = [ 1][ 1][ 1][ 1]
>> B1=subs(B,x,1);
>> escalona(B1)
[ -2,  2,  2,  0]
[ -1,  1,  1,  0]
[ -1,  1,  1,  0]
[  0,  0,  0,  0]
eliminação 1:
linha 2 <==> linha 1
[ -1,  1,  1,  0]
[ -2,  2,  2,  0]
[ -1,  1,  1,  0]
[  0,  0,  0,  0]
(-1)*linha 1 ==> linha 1
[  1, -1, -1,  0]
[ -2,  2,  2,  0]
[ -1,  1,  1,  0]
```

```
[  0,  0,  0,  0]
(2)*linha 1 + linha 2 ==> linha 2
(1)*linha 1 + linha 3 ==> linha 3
[  1, -1, -1,  0]
[  0,  0,  0,  0]
[  0,  0,  0,  0]
[  0,  0,  0,  0]
```

$$\left\{ \begin{array}{ccccccc} x_1 & - & x_2 & - & x_3 & = & 0 \end{array} \right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(\beta+\gamma, \gamma, \beta, \alpha) \,|\, \alpha, \beta, \gamma \in \mathbb{R}\}.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(\beta+\gamma, \gamma, \beta, \alpha) = \alpha(0,0,0,1) + \beta(1,0,1,0) + \gamma(1,1,0,0).$$

Logo, $\mathcal{B} = \{V_1 = (0,0,0,1), V_2 = (1,0,1,0), V_3 = ((1,1,0,0)\}$ gera $\mathbb{W}$. Como

$$\begin{aligned} (0,0,0,0) &= (\beta+\gamma, \gamma, \beta, \alpha) \\ &= \alpha(0,0,0,1) + \beta(1,0,1,0) + \gamma(1,1,0,0) \end{aligned}$$

implica que $\alpha = \beta = \gamma = 0$, então $\mathcal{B}$ é base para $\mathbb{W}$.

(e)
```
>> A=[2,3,0;0,1,0;0,0,2]
>> B=A-x*eye(3)
B = [ 2-x,   3,   0]
    [   0, 1-x,   0]
    [   0,   0, 2-x]
>> solve(det(B))
[ 2] [ 2] [ 1]
```

```
>> B1=subs(B,x,1)
B1 = [ 1, 3 ,0]
     [ 0, 0, 0]
     [ 0, 0, 1]
```

$$\left\{\begin{array}{rcrcrcr} x_1 & + & 3x_2 & & & = & 0 \\ & & & & x_3 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(-3\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}\,.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(-3\alpha, \alpha, 0) = \alpha(-3, 1, 0)\,.$$

Logo, $\mathcal{B} = \{V = (-3, 1, 0)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

```
>> B2=subs(B,x,2)
B2 =[  0,  3,  0]
    [  0, -1,  0]
    [  0,  0,  0]
```

$$\left\{\begin{array}{rcr} 3x_2 & = & 0 \\ -x_2 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(\alpha, 0, \beta) \mid \alpha, \beta \in \mathbb{R}\}\,.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(\alpha, 0, \beta) = \alpha(1, 0, 0) + \beta(0, 0, 1)\,.$$

Logo, $\mathcal{B} = \{V_1 = (1, 0, 0), V_2 = (0, 0, 1)\}$ gera $\mathbb{W}$. Como um vetor não é múltiplo escalar do outro, o conjunto $\mathcal{B}$ é L.I. Assim, $\mathcal{B}$ é base para $\mathbb{W}$.

(f)
```
>> A=[2,3,0;0,2,0;0,0,2]
>> B=A-x*eye(3)
B =[ 2-x,   3,   0]
   [   0, 2-x,   0]
   [   0,   0, 2-x]
>> solve(det(B))
[ 2][ 2][ 2]
>> B2=subs(B,x,2)
B2 =[ 0, 3, 0]
    [ 0, 0, 0]
    [ 0, 0, 0]
```

$$\left\{ \begin{array}{rcl} 3x_2 & = & 0 \end{array} \right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(\alpha, 0, \beta) \mid \alpha, \beta \in \mathbb{R}\}.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(\alpha, 0, \beta) = \alpha(1, 0, 0) + \beta(0, 0, 1).$$

Logo, $\mathcal{B} = \{V_1 = (1, 0, 0), V_2 = (0, 0, 1)\}$ gera $\mathbb{W}$. Como um vetor não é múltiplo escalar do outro, o conjunto $\mathcal{B}$ é L.I. Assim, $\mathcal{B}$ é base para $\mathbb{W}$.

4.1.3.
```
>> N1=[1,-7,5]; N2=[3,-1,1];
>> escalona([N1,0;N2,0])
[  1, -7,  5,  0]
[  3, -1,  1,  0]
eliminação 1:
(-3)*linha 1 + linha 2 ==> linha 2
[   1,  -7,   5,   0]
```

```
[   0,  20, -14,   0]
Continua ? (s/n) s
eliminação 2:
(1/20)*linha 2 ==> linha 2
[     1,    -7,     5,     0]
[     0,     1, -7/10,     0]
Continua ? (s/n) s
(7)*linha 2 + linha 1 ==> linha 1
[     1,     0,  1/10,     0]
[     0,     1, -7/10,     0]
```

Fazendo $z = 10t$ obtemos que $y = 7t$ e $x = -t$. Assim a reta interseção dos dois subespaços é $(x, y, z) = t(-1, 7, 10)$, para todo $t \in \mathbb{R}$. Assim, $\{V = (-1, 7, 10)\}$ é uma base para a reta.

4.1.4. (a)

```
>> v1=[4,2,-3];v2=[2,1,-2];v3=[-2,-1,0];
>> escalona([v1;v2;v3;zeros(1,3)]')
[  4,  2, -2,  0]
[  2,  1, -1,  0]
[ -3, -2,  0,  0]
eliminação 1:
(1/4)*linha 1 ==> linha 1
[    1,  1/2, -1/2,    0]
[    2,    1,   -1,    0]
[   -3,   -2,    0,    0]
(-2)*linha 1 + linha 2 ==> linha 2
(3)*linha 1 + linha 3 ==> linha 3
[    1,  1/2, -1/2,    0]
[    0,    0,    0,    0]
[    0, -1/2, -3/2,    0]
eliminação 2:
linha 3 <==> linha 2
```

```
[    1,  1/2, -1/2,    0]
[    0, -1/2, -3/2,    0]
[    0,    0,    0,    0]
(-2)*linha 2 ==> linha 2
[    1,  1/2, -1/2,    0]
[    0,    1,    3,    0]
[    0,    0,    0,    0]
(-1/2)*linha 2 + linha 1 ==> linha 1
[  1,  0, -2,  0]
[  0,  1,  3,  0]
[  0,  0,  0,  0]
```

Os vetores V_1, V_2 e V_3 são L.D., pois a equação $xV_1 + yV_2 + zV_3 = \bar{0}$ admite solução não trivial.

(b) Os vetores V_1 e V_2 são L.I. pois um vetor não é múltiplo escalar do outro.

(c) A dimensão do subespaço gerado por V_1, V_2 e V_3, é 2, pois, pelos itens anteriores, V_1 e V_2 formam uma base para ele.

(d) Este subespaço é um plano que passa pela origem paralelo aos vetores V_1 e V_2.

4.1.5. (a) Não, pois basta tomarmos um vetor que não está no subespaço gerado por V_1 e V_2 (plano que passa pela origem), que ele não será combinação linear de V_1 e V_2.

(b) Para que V_1, V_2 e V_3 formem uma base de $\mathbb{R}^3$ basta que V_1, V_2 e V_3 sejam L.I. Para isso $V_3 = (a, b, c)$ deve ser um vetor que não seja combinação linear de V_1 e V_2.

(c)
```
>> V1=[2,1,3];V2=[2,6,4];
>> escalona([V1;V2])
[ 2, 1, 3]
[ 2, 6, 4]
eliminação 1:
(1/2)*linha 1 ==> linha 1
[   1, 1/2, 3/2]
```

```
[   2,   6,   4]
(-2)*linha 1 + linha 2 ==> linha 2
[   1, 1/2, 3/2]
[   0,   5,   1]
eliminação 2:
(1/5)*linha 2 ==> linha 2
[   1, 1/2, 3/2]
[   0,   1, 1/5]
(-1/2)*linha 2 + linha 1 ==> linha 1
[   1,   0, 7/5]
[   0,   1, 1/5]
```

Pela Proposição 4.9 (c) o subespaço gerado pelos vetores V_1 e V_2 é igual ao subespaço gerado por $W_1 = (1, 0, 7/5)$ e $W_2 = (0, 1, 1/5)$ (as linhas da forma escalonada reduzida, R). Inserindo, por exemplo, a linha $V_3 = (0, 0, 1)$ na matriz R ela continua escalonada reduzida. Então V_3 não pertence ao subespaço gerado por V_1, V_2.

$$x_1 V_1 + x_2 V_2 + x_3 V_3 = \bar{0}$$

então

$$x_3 V_3 = -x_1 V_1 - x_2 V_2$$

o que implica que $x_3 = 0$ e $x_1 = 0$ e $x_2 = 0$. Ou seja, V_1, V_2 e V_3 são L.I. Como a dimensão do $\mathbb{R}^3$ é igual a 3, então pelo Teorema 4.2 na página 236 $V_3 = (0, 0, 1)$, é tal que V_1, V_2 e V_3 formam uma base de $\mathbb{R}^3$.

4.1.6. Fazendo $z = \alpha$ e $y = \beta$ na equação do plano obtemos que

$$x = -2\beta - 4\alpha.$$

Assim, os pontos do plano $x + 2y + 4z = 0$ são da forma

$$(x, y, z) = (-2\beta - 4\alpha, \beta, \alpha), \ \forall \alpha, \beta \in \mathbb{R},$$

ou seja, são da forma

$$(x, y, z) = \alpha(-4, 0, 1) + \beta(-2, 1, 0) = \alpha V_1 + \beta V_2 \; \forall \alpha, \beta \in \mathbb{R},$$

em que $V_1 = (-4, 0, 1)$ e $V_2 = (-2, 1, 0)$.

Assim, V_1 e V_2 formam uma base do plano $\mathbb{W}$, pois são L.I. (um não é múltiplo escalar do outro) e geram $\mathbb{W}$ (todo vetor de $\mathbb{W}$ é combinação linear deles).

Para estender V_1 e V_2 a uma base de $\mathbb{R}^3$, precisamos acrescentar um vetor que não seja combinação linear de V_1 e V_2. Uma maneira de se conseguir isso é tomar um vetor que não pertença ao plano, ou seja, um vetor (a, b, c) tal que $a + 2b + 4z \neq 0$. Por exemplo $V_3 = (1, 0, 0)$.

4.1.7. Devemos encontrar vetores que são combinações lineares de $V_1 = (-1, 2, 3)$ e $V_2 = (1, 3, 4)$ e ao mesmo tempo sejam combinação lineares de $V_3 = (1, 2, 1)$ e $V_4 = (0, 1, 1)$, ou seja, que satisfaçam as equações

$$V = x_1 V_+ x_2 V_2 = x_3 V_3 + x_4 V_4$$

ou seja,

$$x_1 V_1 + x_2 V_2 - x_3 V_3 - x_4 V_4 = \bar{0}$$

```
>> V1=[-1,2,3];V2=[1,3,4];V3=[1,2,-1];V4=[0,1,1];
>> escalona([V1',V2',-V3',-V4',zeros(3,1)])
[ -1,  1, -1,  0,  0]
[  2,  3, -2, -1,  0]
[  3,  4,  1, -1,  0]
eliminação 1:
(-1)*linha 1 ==> linha 1
[  1, -1,  1,  0,  0]
[  2,  3, -2, -1,  0]
[  3,  4,  1, -1,  0]
(-2)*linha 1 + linha 2 ==> linha 2
(-3)*linha 1 + linha 3 ==> linha 3
[  1, -1,  1,  0,  0]
```

```
[  0,  5, -4, -1,  0]
[  0,  7, -2, -1,  0]
eliminação 2:
(1/5)*linha 2 ==> linha 2
[    1,    -1,     1,     0,     0]
[    0,     1, -4/5, -1/5,     0]
[    0,     7,    -2,    -1,     0]
(1)*linha 2 + linha 1 ==> linha 1
(-7)*linha 2 + linha 3 ==> linha 3
[    1,     0,  1/5, -1/5,     0]
[    0,     1, -4/5, -1/5,     0]
[    0,     0, 18/5,  2/5,     0]
eliminação 3:
(5/18)*linha 3 ==> linha 3
[    1,     0,  1/5, -1/5,     0]
[    0,     1, -4/5, -1/5,     0]
[    0,     0,     1,  1/9,     0]
(-1/5)*linha 3 + linha 1 ==> linha 1
(4/5)*linha 3 + linha 2 ==> linha 2
[    1,     0,     0, -2/9,     0]
[    0,     1,     0, -1/9,     0]
[    0,     0,     1,  1/9,     0]
```

Fazendo $x_4 = 9t$, então $x_3 = -t$, $x_2 = t$, $x_1 = 2t$. Ou seja, $V = x_1V_1 + x_2V_2 = 2tV_1 + tV_2 = 2t(-1,2,3) + t(1,3,4) = (-t,7t,10t) = t(-1,7,10)$ A reta interseção dos dois subespaços é $(x,y,z) = t(-1,7,10)$, para qualquer $t \in \mathbb{R}$. Uma base para a reta é $\{V = (-1,7,10)\}$.

4.1.8. (a)

$$\begin{aligned} V &= (3a + 4b - 4c, 2a - 4b - 6c, -2a - 4b + 2c) \\ &= (3a, 2a, -2a) + (4b, -4b, -4b) + (-4c, -6c, 2c) \\ &= a(3,2,-2) + b(4,-4,-4) + c(-4,-6,2). \end{aligned}$$

Logo, definindo $V_1 = (3, 2, -2)$, $V_2 = (4, -4, -4)$ e $V_3 = (-4, -6, 2)$, então $\{V_1, V_2, V_3\}$ gera $\mathbb{V}$.

(b)
```
>> V1=[3,2,-2];V2=[4,-4,-4];V3=[-4,-6,2];
>> escalona([V1;V2;V3]')
[  3,  4, -4]
[  2, -4, -6]
[ -2, -4,  2]
eliminação 1:
(1/3)*linha 1 ==> linha 1
[    1,  4/3, -4/3]
[    2,   -4,   -6]
[   -2,   -4,    2]
(-2)*linha 1 + linha 2 ==> linha 2
(2)*linha 1 + linha 3 ==> linha 3
[     1,   4/3,  -4/3]
[     0, -20/3, -10/3]
[     0,  -4/3,  -2/3]
eliminação 2:
(-3/20)*linha 2 ==> linha 2
[    1,  4/3, -4/3]
[    0,    1,  1/2]
[    0, -4/3, -2/3]
(-4/3)*linha 2 + linha 1 ==> linha 1
(4/3)*linha 2 + linha 3 ==> linha 3
[   1,   0,  -2]
[   0,   1, 1/2]
[   0,   0,   0]
```

Como a terceira coluna da matriz escalonada reduzida equivalente a $A = [V_1\ V_2\ V_3]$ é igual a soma de -2 vezes a primeira com $1/2$ vezes a segunda, pela Proposição 4.9 (b) a mesma relação vale para as colunas correspondentes da matriz A, ou seja, $V_3 = -2V_2 + \frac{1}{2}V_1$. Assim o vetor V_3 pode ser descartado na geração de $\mathbb{V}$, pois ele é combinação linear dos outros dois. Logo, apenas V_1 e V_2 são suficientes para gerar $\mathbb{V}$. Como além disso, os vetores V_1 e V_2 são tais que um não é múltiplo escalar do outro, então eles são L.I. e portanto $\{V_1, V_2\}$ é uma base de $\mathbb{V}$.

4.1.9. (a) Não pois são necessários 4 vetores L.I. para se obter uma base de $\mathbb{R}^4$ (Teorema 4.2 na página 236).

(b) V_3 e V_4 devem ser L.I. e não pertencerem ao subespaço gerado por V_1 e V_2.

(c) Escalonando a matriz cujas linhas são os vetores V_1 e V_2,

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

obtemos

$$R = \begin{bmatrix} 1 & 0 & 1 & -12 \\ 0 & 1 & 1 & -7 \end{bmatrix}$$

Pela Proposição 4.9 (c) o subespaço gerado pelos vetores V_1 e V_2 (as linhas de A) é igual ao subespaço gerado por $W_1 = (1, 0, 1, -12)$ e $W_2 = (0, 1, 1, -7)$ (as linhas de R). Vamos inserir linhas na matriz R até conseguir uma matriz 4×4 de forma que ela continue escalonada reduzida. Por exemplo, podemos obter a matriz

$$\bar{R} = \begin{bmatrix} 1 & 0 & 1 & -12 \\ 0 & 1 & 1 & -7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Sejam $V_3 = (0, 0, 1, 0)$ e $V_4 = (0, 0, 0, 1)$. Vamos verificar que

$$V_1 = (-3, 5, 2, 1), \quad V_2 = (1, -2, -1, 2),$$

$$V_3 = (0, 0, 1, 0) \quad \text{e} \quad V_4 = (0, 0, 0, 1)$$

são L.I.

$$x_1V_1 + x_2V_2 + x_3V_3 + x_4V_4 = \bar{0}$$

implica que

$$x_1V_1 + x_2V_2 = -x_3V_3 - x_4V_4$$

Como V_3 e V_4 não pertencem ao subespaço gerado por V_1 e V_2 que é o mesmo subespaço gerado por W_1 e W_2, então

$$x_1V_1 + x_2V_2 = \bar{0} \quad \text{e} \quad -x_3V_3 - x_4V_4 = \bar{0}$$

como $\{V_1, V_2\}$ é L.I. assim como $\{V_3, V_4\}$ então $x_1 = x_2 = x_3 = x_4 = 0$. Logo $\{V_1, V_2, V_3, V_4\}$ é L.I. Assim pelo Teorema 4.2 na página 236 V_1, V_2, V_3, V_4 formam uma base de $\mathbb{R}^4$.

4.1.10.
```
>> A=triu(randi(4,4,3))
A = -1    -2     1     1
     0     2    -2    -2
     0     0    -1     2
     0     0     0     0
>> B=A-x*eye(4)
B =
[ -1-x,   -2,    1,    1]
[    0,  2-x,   -2,   -2]
[    0,    0, -1-x,    2]
[    0,    0,    0,   -x]
>> solve(det(B))
[ -1][ -1][  2][  0]
>> Bm1=subs(B,x,-1)
Bm1 =
[  0, -2,  1,  1]
[  0,  3, -2, -2]
[  0,  0,  0,  2]
[  0,  0,  0,  1]
>> escalona(Bm1)
[ 0, 1, 0, 0]
[ 0, 0, 1, 0]
```

```
[ 0, 0, 0, 1]
[ 0, 0, 0, 0]
```

$$\left\{\begin{array}{ccccccc} & x_2 & & & = & 0 \\ & & x_3 & & = & 0 \\ & & & x_4 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(\alpha, 0, 0, 0) \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(\alpha, 0, 0, 0) \quad = \quad \alpha(1, 0, 0, 0).$$

Logo, $\mathcal{B} = \{V = (1, 0, 0, 0)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

```
>> B2=subs(B,x,2)
B2 =
[ -3, -2,  1,  1]
[  0,  0, -2, -2]
[  0,  0, -3,  2]
[  0,  0,  0, -2]
>> escalona(B2)
[   1, 2/3,   0,   0]
[   0,   0,   1,   0]
[   0,   0,   0,   1]
[   0,   0,   0,   0]
```

$$\left\{\begin{array}{ccccccc} x_1 & + & 2/3x_2 & & & = & 0 \\ & & & x_3 & & = & 0 \\ & & & & x_4 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(-2\alpha, 3\alpha, 0, 0) \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(-2\alpha, 3\alpha, 0, 0) \quad = \quad \alpha(-2, 3, 0, 0).$$

Logo, $\mathcal{B} = \{V = (-2, 3, 0, 0)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

```
>> B0=subs(B,x,0)
B0 =
[ -1, -2,  1,  1]
[  0,  2, -2, -2]
[  0,  0, -1,  2]
[  0,  0,  0,  0]
>> escalona(B0)
[  1,  0,  0,  3]
[  0,  1,  0, -3]
[  0,  0,  1, -2]
[  0,  0,  0,  0]
```

$$\left\{\begin{array}{ccccccc} x_1 & & & & 3x_4 & = & 0 \\ & x_2 & & & -3x_4 & = & 0 \\ & & x_3 & - & 2x_4 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(-3\alpha, 3\alpha, 2\alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(-3\alpha, 3\alpha, 2\alpha, \alpha) \quad = \quad \alpha(-3, 3, 2, 1).$$

Logo, $\mathcal{B} = \{V = (-3, 3, 2, 1)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então $\mathcal{B}$ é base para $\mathbb{W}$.

4.2. Espaço Linha e Espaço Coluna (página 269)

4.2.1. (a)
```
>> A=[1,4,5,2;2,1,3,0;-1,3,2,2];
>> escalona(A)
[    1,    0,    1, -2/7]
[    0,    1,    1,  4/7]
[    0,    0,    0,    0]
```

$V_1 = (1, 0, 1, -2/7), V_2 = (0, 1, 1, 4/7)$ formam uma base para o espaço linha e $W_1 = (1, 2, -1), W_2 = (4, 1, 3)$ formam uma base para o espaço coluna de A.

(b)
```
>> A=[1,-4,-5,4;-1,4,4,-5;0,0,2,0];
>> escalona(A)
[  1, -4,  0,  0]
[  0,  0,  1,  0]
[  0,  0,  0,  1]
```

$V_1 = (1, -4, 0, 0), V_2 = (0, 0, 1, 0), V_3 = (0, 0, 0, 1)$ formam uma base para o espaço linha de A e $W_1 = (1, -1, 0), W_2 = (-5, 4, 2), W_3 = (4, -5, 0)$ formam uma base para o espaço coluna de A.

4.2.2. (a)
```
>> A=[1,-2,2;2,-2,4;-3,3,6];
>> escalona(A)
[  1,  0,  2]
[  0,  1,  0]
[  0,  0, 12]
```

A dimensão é 3, ou seja, o subespaço é o $\mathbb{R}^3$.

(b)
```
>> A=[1,-3,4;6,2,-1;2,-2,3;-4,-8,9].'
>> escalona(A)
```

```
[   1,    0, 4/5,   2]
[   0,    1, 1/5,  -1]
[   0,    0,   0,   0]
```

A dimensão é 2.

4.2.3. (a)
```
>> A=[1,2,2,3,1,4;2,4,5,5,4,9;3,6,7,8,5,9]
>> U=escalona(A)
U=[  1,  2,  0,  5, -3,  0]
  [  0,  0,  1, -1,  2,  0]
  [  0,  0,  0,  0,  0,  1]
```

As colunas U_2, U_4 e U_5 correspondem às variáveis livres. $U_2 = 2U_1, U_4 = 5U_1 - U_3, U_5 = -3U_1 + 2U_3$.

(b) As colunas A_1, A_3 e A_6 formam uma base para o espaço coluna de A. $A_2 = 2A_1, A_4 = 5A_1 - A_3, A_5 = -3A_1 + 2A_3$.

4.2.4. (a) O posto é 2, pois as duas linhas são L.I.

(b) O posto é 1, pois as duas linhas são L.D.

(c) O posto é 2, pois o determinante da matriz é igual a zero e existem pelo menos duas linhas L.I.

(d) O posto é 3, pois o determinante da submatriz formada pelas primeiras três colunas da matriz é diferente de zero.

4.2.5. (a)
```
>> syms t; A=[1,1,t;1,t,1;t,1,1]
>> escalona(A)
[     1,     1,     t]
[     0,   t-1,   1-t]
[     0,   1-t, 1-t^2]
Continua? s
[        1,        0,      t+1]
[        0,        1,       -1]
```

```
[        0,        0, -t^2-t+2]
```

Se $t = 1$, então o posto de A é igual a 1. Se $t \neq 1$, então o posto de A é igual a 2 ou 3. Se além disso, $-t^2 - t + 2 = -(t+2)(t-1) = 0$, então o posto de A é igual a 2, ou seja, o posto de A é igual a 2, se $t = -2$ e é igual a 3, se $t \neq 1, -2$.

(b)
```
>> A=[t,3,-1;3,6,-2;-1,-3,-t]
>> escalona(A);
[          1,          0,     -2*t-2]
[          0,          1,      2/3+t]
[          0,          0, -3+2*t^2-t]
```

O posto de A é igual a 2, se $2t^2 - t - 3 = 0$, ou seja, se $t = -1, 3/2$. Caso contrário, o posto de A é igual a 3.

4.2.6. (a) $\text{posto}(A) \leq 2$ e $\text{nulidade}(A) \geq 1$.

(b) $\text{posto}(A) \leq 2$ e $\text{nulidade}(A) \geq 0$.

(c) $\text{posto}(A) \leq 3$ e $\text{nulidade}(A) \geq 0$.

(d) $\text{posto}(A) \leq \min\{m, n\}$ e $\text{nulidade}(A) \geq n - \min\{m, n\}$.

4.2.7. (a) Em nenhuma hipótese o sistema $AX = B$ tem solução única. Se $\text{posto}(A) = \text{posto}([A|B]) = 1$ ou $\text{posto}(A) = \text{posto}([A|B]) = 2$, então o sistema $AX = B$ tem infinitas soluções. Se $\text{posto}(A) = 1$ e $\text{posto}([A|B]) = 2$, então o sistema $AX = B$ não tem solução.

(b) Se $\text{posto}(A) = \text{posto}([A|B]) = 2$, então o sistema $AX = B$ tem solução única. Se $\text{posto}(A) = \text{posto}([A|B]) = 1$, então o sistema $AX = B$ tem infinitas soluções. Se $\text{posto}(A) = 1$ e $\text{posto}([A|B]) = 2$ (observe que neste caso não pode ocorrer $\text{posto}([A|B]) = 3$), então o sistema $AX = B$ não tem solução.

(c) Se $\text{posto}(A) = \text{posto}([A|B]) = 3$, então o sistema $AX = B$ tem solução única. Se $\text{posto}(A) = \text{posto}([A|B]) = 1$ ou $\text{posto}(A) = \text{posto}([A|B]) = 2$, então o sistema $AX = B$ tem infinitas soluções.

Se posto$(A) = 1$ e posto$([A|B]) = 2, 3$ ou se posto$(A) = 2$ e posto$([A|B]) = 3$, então o sistema $AX = B$ não tem solução.

(d) Se posto$(A) =$ posto$([A|B]) = n$, então o sistema $AX = B$ tem solução única. Se posto$(A) =$ posto$([A|B]) = 1, \ldots, n-1$, então o sistema $AX = B$ tem infinitas soluções. Se posto$(A) = i$ e posto$([A|B]) > i$, para $i = 1, \ldots, n-1$, então o sistema $AX = B$ não tem solução.

4.3. Espaços Vetoriais (página 285)

4.3.1. A equação $3X - 2V = 15(X - U)$ é equivalente a $3X - 2V = 15X - 15U$. Somando-se $-15X + 2V$ obtemos $-15X + 3X = 2V - 15U$ ou $-12X = 2V - 15U$ multiplicando-se por $-\frac{1}{12}$ obtemos $X = \frac{5}{4}U - \frac{1}{6}V$.

4.3.2. Multiplicando-se a segunda equação por 2 e somando-se a primeira, obtemos $12X = 3U + 2V$ ou $X = \frac{1}{4}U + \frac{1}{6}V$. Substituindo-se X na primeira equação obtemos, $\frac{3}{2}U + V - 2Y = U$ ou $2Y = \frac{1}{2}U + V$ ou $Y = \frac{1}{4}U + \frac{1}{2}V$.

4.3.3. Vamos verificar que existem escalares a e b tais que
$t^2 + 2t + 7 = a(t^2 + 1) + b(t + 3) = at^2 + bt + (a + 3b)$
que é equivalente ao sistema

$$\left\{ \begin{array}{ccccc} a & & & = & 1 \\ & & b & = & 2 \\ a & + & 3b & = & 7 \end{array} \right.$$

que tem solução $a = 1$ e $b = 2$.

4.3.4. Vamos verificar que existem escalares a e b tais que $3 = a(5\tan^2 t) + 2b\sec^2 t = 5a(\sec^2 t - 1) + 2b\sec^2 t = -5a + (5a + 2b)\sec^2 t$. Basta tomarmos escalares a e b tais que $3 = -5a$ e $5a + 2b = 0$, ou seja, basta que $a = -3/5$ e $b = 3/2$.

4.3.5.
```
>> x1=[4,2,-3];x2=[2,1,-2];x3=[-2,-1,0];
>> x=[1,1,1];
>> A=[x1;x2;x3;x].'
     4     2    -2     1
```

```
     2     1    -1     1
    -3    -2     0     1
>> R=escalona(A)
     1     0    -2     0
     0     1     3     0
     0     0     0     1
>> x=[4,2,-6];
>> A=[x1;x2;x3;x].'
     4     2    -2     4
     2     1    -1     2
    -3    -2     0    -6
>> R=escalona(A)
     1     0    -2    -2
     0     1     3     6
     0     0     0     0
>> x=[-2,-1,1];
>> A=[x1;x2;x3;x].'
     4     2    -2    -2
     2     1    -1    -1
    -3    -2     0     1
>> R=escalona(A)
     1     0    -2    -1
     0     1     3     1
     0     0     0     0
>> x=[-1,2,3];
>> A=[x1;x2;x3;x].'
     4     2    -2    -1
     2     1    -1     2
    -3    -2     0     3
>> R=escalona(A)
```

```
1      0     -2      0
0      1      3      0
0      0      0      1
```

Assim, os vetores das letras (b) e (c) são combinação linear de X_1, X_2 e X_3.

4.3.6. (a) Não é espaço vetorial pois $1(1,1) = (1,0) \neq (1,1)$, ou seja, o axioma (8) não é satisfeito.

(b) Não é espaço vetorial pois $(1,1) + (2,2) = (5,5) \neq (2,2) + (1,1) = (4,4)$, ou seja, o axioma (1) não é satisfeito.

(c) Não é espaço vetorial pois $(0,1) + ((0,2) + 0,3)) = (0,1) + (5,0) = (1,5) \neq ((0,1) + (0,2)) + (0,3) = (3,0) + (0,3) = (3,3)$

(d) É espaço vetorial, pois:

i. $x + y = xy = yx = y + x$;

ii. $x + (y + z) = x(yz) = (xy)z = (x + y) + z$;

iii. Se $\bar{0} = 1$, então $x = 1x = \bar{0} + x = x + \bar{0}$;

iv. Se $-x = x^{-1}$, então $(-x) + x = x + (-x) = xx^{-1} = 1 = \bar{0}$;

v. $\alpha(\beta x) = \alpha(x^{\alpha}) = (x^{\alpha})^{\beta} = x^{\alpha\beta} = (\alpha\beta)x$;

vi. $\alpha(x + y) = (xy)^{\alpha} = x^{\alpha}y^{\alpha} = \alpha x + \alpha y$;

vii. $(\alpha + \beta)x = x^{(\alpha+\beta)} = x^{\alpha}x^{\beta} = \alpha x + \beta x$;

viii. $1x = x^1 = x$.

4.3.7. (1) $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$, para todo $x \in \mathcal{X}$;

(2) $[f + (g + h)](x) = f(x) + (g + h)(x) = f(x) + (g(x) + h(x)) = (f(x) + g(x)) + h(x) = (f + g)(x) + h(x) = [(f + g) + h](x)$, para todo $x \in \mathcal{X}$;

(3) Seja $\bar{0}$ a função identicamente nula. $(f + \bar{0})(x) = f(x) + \bar{0}(x) = f(x)$, para todo $x \in \mathcal{X}$;

(4) Dada a função f definimos a função $-f$ por $(-f)(x) = -f(x)$, para todo $x \in \mathcal{X}$. $[f + (-f)](x) = f(x) + (-f(x) = 0 = \bar{0}(x)$, para todo $x \in \mathcal{X}$;

(5) $[\alpha(\beta f)](x) = \alpha(\beta f)(x) = \alpha(\beta f(x)) = (\alpha\beta) f(x) = [(\alpha\beta) f](x)$, para todo $x \in \mathcal{X}$;

(6) $[\alpha(f + g)](x) = \alpha(f + g)(x) = \alpha(f(x) + g(x)) = \alpha f(x) + \alpha g(x) = (\alpha f)(x) + (\alpha g)(x) = (\alpha f + \alpha g)(x)$, para todo $x \in \mathcal{X}$;

(7) $[(\alpha + \beta) f](x) = (\alpha + \beta) f(x) = \alpha f(x) + \beta f(x) = (\alpha f)(x) + (\beta f)(x) = [\alpha f + \beta f](x)$, para todo $x \in \mathcal{X}$;

(8) $(1f)(x) = 1f(x) = f(x)$, para todo $x \in \mathcal{X}$;

5.1. Produto Escalar em $\mathbb{R}^n$ (página 302)

5.1.1.
```
>> syms a
>> x=[1,1,-2];y=[a,-1,2];
>> solve(pe(x,y))
ans = 5
```

5.1.2.
```
>> syms a b
>> x=[1/2^(1/2),0,1/2^(1/2)];y=[a,1/2^(1/2),-b];
>> sol=solve(pe(x,y),no(y)-1)
sol =
    a: [2x1 sym]
    b: [2x1 sym]
>> sol.a, sol.b
ans = [  1/2] [ -1/2] ans = [  1/2] [ -1/2]
```

5.1.3.
```
>> v1=[1,1,-1,0];v2=[0,2,0,1];v3=[-1,0,0,1];
>> w1=v1; w2=v2-proj(w1,v2)
w2 = [-2/3, 4/3, 2/3, 1]
>> w3=v3-proj(w1,v3)-proj(w2,v3)
w3 = [-4/11, -3/11, -7/11, 6/11]
```

```
>> u1=w1/no(w1),u2=w2/no(w2),u3=w3/no(w3)
```

$$u_1 = \left[\begin{array}{cccc} \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} & -\frac{1}{3}\sqrt{3} & 0 \end{array}\right]$$
$$u2 = \left[\begin{array}{cccc} -\frac{2}{33}\sqrt{11}\sqrt{3} & \frac{4}{33}\sqrt{11}\sqrt{3} & \frac{2}{33}\sqrt{11}\sqrt{3} & \frac{1}{11}\sqrt{11}\sqrt{3} \end{array}\right]$$
$$u3 = \left[\begin{array}{cccc} -\frac{2}{55}\sqrt{110} & -\frac{3}{110}\sqrt{110} & -\frac{7}{110}\sqrt{110} & \frac{3}{55}\sqrt{110} \end{array}\right]$$

5.1.4.
```
>> v1=[1,1,1];v2=[0,1,1];v3=[1,2,3];
>> w1=v1; w2=v2-proj(w1,v2)
w2 =  [-2/3, 1/3, 1/3]
>> w3=v3-proj(w1,v3)-proj(w2,v3)
w3 = [0, -1/2, 1/2]
>> u1=w1/no(w1),u2=w2/no(w2),u3=w3/no(w3)
```

$$u_1 = \left[\begin{array}{ccc} \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} \end{array}\right]$$
$$u_2 = \left[\begin{array}{ccc} -\frac{1}{3}\sqrt{2}\sqrt{3} & \frac{1}{6}\sqrt{2}\sqrt{3} & \frac{1}{6}\sqrt{2}\sqrt{3} \end{array}\right]$$
$$u_3 = \left[\begin{array}{ccc} 0 & -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{array}\right]$$

5.1.5. Este subespaço consiste dos vetores da forma:

$$\begin{aligned} (-\alpha-\beta,\beta,\alpha) &= (-\alpha,0,\alpha)+(-\beta,\beta,0) \\ &= \alpha(-1,0,1)+\beta(-1,1,0) \end{aligned}$$

```
>> v1=[-1,0,1];v2=[-1,1,0];
>> w1=v1; w2=v2-proj(w1,v2);
>> u1=w1/no(w1), u2=w2/no(w2)
```

$$u_1 = \left[\begin{array}{ccc} -\frac{1}{2}\sqrt{2} & 0 & \frac{1}{2}\sqrt{2} \end{array}\right]$$
$$u_2 = \left[\begin{array}{ccc} -\frac{1}{6}\sqrt{3}\sqrt{2} & \frac{1}{3}\sqrt{3}\sqrt{2} & -\frac{1}{6}\sqrt{3}\sqrt{2} \end{array}\right]$$

5.1.6. Este subespaço consiste dos vetores da forma:

$$\begin{aligned}(-\alpha + 2\beta + \gamma, \gamma, \beta, \alpha) &= \\ (-\alpha, 0, 0, \alpha) + (2\beta, 0, \beta, 0) + (\gamma, \gamma, 0, 0) &= \\ \alpha(-1, 0, 0, 1) + \beta(2, 0, 1, 0) + \gamma(1, 1, 0, 0)\end{aligned}$$

```
>> v1=[-1,0,0,1];v2=[2,0,1,0];v3=[1,1,0,0];
>> w1=v1; w2=v2-proj(w1,v2);
>> w3=v3-proj(w1,v3)-proj(w2,v3);
>> u1=w1/no(w1), u2=w2/no(w2), u3=w3/no(w3)
```

$$u_1 = \begin{bmatrix} -\frac{1}{2}\sqrt{2} & 0 & 0 & \frac{1}{2}\sqrt{2} \end{bmatrix}$$
$$u_2 = \begin{bmatrix} \frac{1}{3}\sqrt{3} & 0 & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} \end{bmatrix}$$
$$u_3 = \begin{bmatrix} \frac{1}{42}\sqrt{42} & \frac{1}{7}\sqrt{42} & -\frac{1}{21}\sqrt{42} & \frac{1}{42}\sqrt{42} \end{bmatrix}$$

5.1.7.
```
>> A=[1,1,-1,0;2,1,2,0];
>> escalona(A)
  1     0      3    0
  0     1     -4    0
```

$$\left\{\begin{array}{ccccc} x_1 & & + & 3x_3 & = & 0 \\ & x_2 & - & 4x_3 & = & 0 \end{array}\right.$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(-3\alpha, 4\alpha, \alpha) \,|\, \alpha \in \mathbb{R}\}\,.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(-3\alpha, 4\alpha, \alpha) = \alpha(-3, 4, 1)\,.$$

Logo, $S = \{V = (-3, 4, 1)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então S é base para $\mathbb{W}$.

```
>> v=[-3,4,1];
>> u=v/no(v)
```

$$u = \left[\begin{array}{ccc} -\frac{3}{26}\sqrt{26} & \frac{2}{13}\sqrt{26} & \frac{1}{26}\sqrt{26} \end{array}\right]$$

5.1.8.
```
>> V1=[1,2,-3]; P1=[0,0,0];
>> V2=[2,4,-6]; P2=[0,1,2];
>> pv(V1,V2)
ans =   0     0     0
>> syms x y z; X=[x,y,z];
>> M=[X-P1;V1;P2-P1], expr=det(M)
M =[  x,  y,  z]
   [  1,  2, -3]
   [  0,  1,  2] expr = 7*x-2*y+z
```

Como o produto vetorial de V_1 e V_2 (os dois vetores diretores das retas) é igual ao vetor nulo, então as retas são paralelas. Neste caso, os vetores V_1 e $\overrightarrow{P_1P_2}$ são não colineares e paralelos ao plano procurado. Assim, $7x - 2y + z = 0$ é a equação do plano, que passa pela origem, logo é um subespaço. Este subespaço consiste dos vetores da forma:

$$\begin{aligned} (\alpha, \beta, -7\alpha + 2\beta) &= (\alpha, 0, -7\alpha) + (0, \beta, 2\beta) \\ &= \alpha(1, 0, -7) + \beta(0, 1, 2) \end{aligned}$$

```
>> V1=[1,0,-7];V2=[0,1,2];
>> W1=V1; W2=V2-proj(W1,V2)
W2 =[ 7/25,    1, 1/25]
>> U1=W1/no(W1), U2=W2/no(W2)
```

$$U_1 = \left[\begin{array}{ccc} 1/10\sqrt{2} & 0 & -\frac{7}{10}\sqrt{2} \end{array}\right]$$

$$U_2 = \left[\begin{array}{ccc} \frac{7}{45}\sqrt{3} & 5/9\sqrt{3} & 1/45\sqrt{3} \end{array}\right]$$

Para completarmos a uma base ortonormal de $\mathbb{R}^3$, basta acrescentarmos $U_3 = U_1 \times U_2$.

```
>> U3=pv(U1,U2)
```

$$U_3 = \left[\begin{array}{ccc} \frac{7}{18}\sqrt{2}\sqrt{3} & -1/9\sqrt{2}\sqrt{3} & 1/18\sqrt{2}\sqrt{3} \end{array} \right]$$

5.2. Subespaços Ortogonais (página 319)

5.2.1.
```
>> v1=[1,1,-1,0];v2=[0,2,0,1];v3=[-1,0,0,1];
>> w1=v1; w2=v2-proj(w1,v2)
w2 = [-2/3, 4/3, 2/3, 1]
>> w3=v3-proj(w1,v3)-proj(w2,v3)
w3 = [-4/11, -3/11, -7/11, 6/11]
>> u1=w1/no(w1),u2=w2/no(w2),u3=w3/no(w3)
```

$$u_1 = \left[\begin{array}{cccc} \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} & -\frac{1}{3}\sqrt{3} & 0 \end{array} \right]$$

$$u2 = \left[\begin{array}{cccc} -\frac{2}{33}\sqrt{11}\sqrt{3} & \frac{4}{33}\sqrt{11}\sqrt{3} & \frac{2}{33}\sqrt{11}\sqrt{3} & \frac{1}{11}\sqrt{11}\sqrt{3} \end{array} \right]$$

$$u3 = \left[\begin{array}{cccc} -\frac{2}{55}\sqrt{110} & -\frac{3}{110}\sqrt{110} & -\frac{7}{110}\sqrt{110} & \frac{3}{55}\sqrt{110} \end{array} \right]$$

5.2.2.
```
>> v1=[1,1,1];v2=[0,1,1];v3=[1,2,3];
>> w1=v1; w2=v2-proj(w1,v2)
w2 =  [-2/3, 1/3, 1/3]
>> w3=v3-proj(w1,v3)-proj(w2,v3)
w3 = [0, -1/2, 1/2]
>> u1=w1/no(w1),u2=w2/no(w2),u3=w3/no(w3)
```

$$u_1 = \left[\begin{array}{ccc} \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} \end{array} \right]$$

$$u_2 = \left[\begin{array}{ccc} -\frac{1}{3}\sqrt{2}\sqrt{3} & \frac{1}{6}\sqrt{2}\sqrt{3} & \frac{1}{6}\sqrt{2}\sqrt{3} \end{array} \right]$$

$$u_3 = \left[\begin{array}{ccc} 0 & -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{array} \right]$$

5.2.3. Este subespaço consiste dos vetores da forma:

$$\begin{aligned}(-\alpha-\beta,\beta,\alpha) &= (-\alpha,0,\alpha)+(-\beta,\beta,0)\\ &= \alpha(-1,0,1)+\beta(-1,1,0)\end{aligned}$$

```
>> v1=[-1,0,1];v2=[-1,1,0];
>> w1=v1; w2=v2-proj(w1,v2);
>> u1=w1/no(w1), u2=w2/no(w2)
```

$$u_1 = \left[\begin{array}{ccc} -\frac{1}{2}\sqrt{2} & 0 & \frac{1}{2}\sqrt{2}\end{array}\right]$$

$$u_2 = \left[\begin{array}{ccc} -\frac{1}{6}\sqrt{3}\sqrt{2} & \frac{1}{3}\sqrt{3}\sqrt{2} & -\frac{1}{6}\sqrt{3}\sqrt{2}\end{array}\right]$$

5.2.4. Este subespaço consiste dos vetores da forma:

$$\begin{aligned}(-\alpha+2\beta+\gamma,\gamma,\beta,\alpha) &=\\ (-\alpha,0,0,\alpha)+(2\beta,0,\beta,0)+(\gamma,\gamma,0,0) &=\\ \alpha(-1,0,0,1)+\beta(2,0,1,0)+\gamma(1,1,0,0)\end{aligned}$$

```
>> v1=[-1,0,0,1];v2=[2,0,1,0];v3=[1,1,0,0];
>> w1=v1; w2=v2-proj(w1,v2);
>> w3=v3-proj(w1,v3)-proj(w2,v3);
>> u1=w1/no(w1), u2=w2/no(w2), u3=w3/no(w3)
```

$$u_1 = \left[\begin{array}{cccc} -\frac{1}{2}\sqrt{2} & 0 & 0 & \frac{1}{2}\sqrt{2}\end{array}\right]$$

$$u_2 = \left[\begin{array}{cccc} \frac{1}{3}\sqrt{3} & 0 & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3}\end{array}\right]$$

$$u_3 = \left[\begin{array}{cccc} \frac{1}{42}\sqrt{42} & \frac{1}{7}\sqrt{42} & -\frac{1}{21}\sqrt{42} & \frac{1}{42}\sqrt{42}\end{array}\right]$$

5.2.5.
```
>> A=[1,1,-1,0;2,1,2,0];
>> escalona(A)
 1    0     3   0
```

```
  0     1     -4    0
```

$$\begin{cases} x_1 & & + & 3x_3 & = & 0 \\ & x_2 & - & 4x_3 & = & 0 \end{cases}$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(-3\alpha, 4\alpha, \alpha) \,|\, \alpha \in \mathbb{R}\}\,.$$

Agora, para qualquer elemento de $\mathbb{W}$ temos:

$$(-3\alpha, 4\alpha, \alpha) = \alpha(-3, 4, 1)\,.$$

Logo, $S = \{V = (-3, 4, 1)\}$ gera $\mathbb{W}$. Como um conjunto formado por um único vetor não nulo é sempre L.I., então S é base para $\mathbb{W}$.

```
>> v=[-3,4,1];
>> u=v/no(v)
```

$$u = \left[\; -\frac{3}{26}\sqrt{26} \quad \frac{2}{13}\sqrt{26} \quad \frac{1}{26}\sqrt{26} \;\right]$$

5.2.6.
```
>> V1=[1,2,-3]; P1=[0,0,0];
>> V2=[2,4,-6]; P2=[0,1,2];
>> pv(V1,V2)
ans =   0     0     0
>> syms x y z; X=[x,y,z];
>> M=[X-P1;V1;P2-P1], expr=det(M)
M =[  x,   y,   z]
   [  1,   2,  -3]
   [  0,   1,   2] expr = 7*x-2*y+z
```

Como o produto vetorial de V_1 e V_2 (os dois vetores diretores das retas) é igual ao vetor nulo, então as retas são paralelas. Neste caso, os vetores V_1 e $\overrightarrow{P_1P_2}$ são não colineares e paralelos ao plano procurado. Assim, $7x - 2y + z = 0$ é a equação do plano, que passa pela origem, logo é um subespaço. Este subespaço consiste dos vetores da forma:

$$\begin{aligned}(\alpha,\beta,-7\alpha+2\beta) &= (\alpha,0,-7\alpha)+(0,\beta,2\beta)\\ &= \alpha(1,0,-7)+\beta(0,1,2)\end{aligned}$$

```
>> V1=[1,0,-7];V2=[0,1,2];
>> W1=V1; W2=V2-proj(W1,V2)
W2 =[ 7/25,    1, 1/25]
>> U1=W1/no(W1), U2=W2/no(W2)
```

$$U_1 = \left[\begin{array}{ccc} 1/10\sqrt{2} & 0 & -\frac{7}{10}\sqrt{2}\end{array}\right]$$

$$U_2 = \left[\begin{array}{ccc} \frac{7}{45}\sqrt{3} & 5/9\sqrt{3} & 1/45\sqrt{3}\end{array}\right]$$

Para completarmos a uma base ortonormal de $\mathbb{R}^3$, basta acrescentarmos $U_3 = U_1 \times U_2$.

```
>> U3=pv(U1,U2)
```

$$U_3 = \left[\begin{array}{ccc} \frac{7}{18}\sqrt{2}\sqrt{3} & -1/9\sqrt{2}\sqrt{3} & 1/18\sqrt{2}\sqrt{3}\end{array}\right]$$

5.2.7.
```
>> syms x y z d
>> expr1=2*x+2*y+2*z+d;
>> P1=[0,0,-d/2]; N=[2,2,2]; P=[1,1,1];
>> expr2=abs(pe(P-P1,N))/no(N)
```

$$\mathtt{expr2} = 1/6\,|6+d|\,\sqrt{3}$$

```
>> solve(expr2-sqrt(3),d)
ans = [   0][ -12]
```

Os planos $2x+2y+2z = 0$ e $2x+2y+2z-12 = 0$ satisfazem as condições do exercício. Apenas o primeiro plano é um subespaço. Este subespaço consiste dos vetores da forma:

$$\begin{aligned}(\alpha,\beta,-\alpha-\beta) &= (\alpha,0,-\alpha)+(0,\beta,-\beta)\\ &= \alpha(1,0,-1)+\beta(0,1,-1)\end{aligned}$$

```
>> V1=[1,0,-1];V2=[0,1,-1];
>> W1=V1; W2=V2-proj(W1,V2)
W2 = [ -1/2,    1, -1/2]
>> U1=W1/no(W1), U2=W2/no(W2)
```

$$U_1 = \left[\begin{array}{ccc} 1/2\sqrt{2} & 0 & -1/2\sqrt{2} \end{array}\right]$$

$$U_2 = \left[\begin{array}{ccc} -1/6\sqrt{3}\sqrt{2} & 1/3\sqrt{3}\sqrt{2} & -1/6\sqrt{3}\sqrt{2} \end{array}\right].$$

5.3. Mudança de Coordenadas (página 331)

5.3.1. (a)

```
>> v1=sym([1/sqrt(2),-1/sqrt(2)]);
>> v2=sym([1/sqrt(2),1/sqrt(2)]);
>> p=[1,3];
>> A=[v1;v2;p].'
>> escalona(A)
[1,         0,  -2^(1/2)]
[0,         1, 2*2^(1/2)]
```

Assim, as coordenadas de P em relação ao sistema $\mathcal{S}$ são:

$$\left[\begin{array}{c} -\sqrt{2} \\ 2\sqrt{2} \end{array}\right]$$

(b)

```
>> v1=sym([1/sqrt(2),-1/sqrt(2),0]);
>> v2=sym([0,0,1]);
>> v3=sym([1/sqrt(2),1/sqrt(2),0]);
>> p=[2,-1,2]; A=[v1;v2;v3;p].';
>> escalona(A)
[ 1,           0,           0, 3/2*2^(1/2)]
[ 0,           1,           0,           2]
[ 0,           0,           1, 1/2*2^(1/2)]
```

Assim, as coordenadas de P em relação ao sistema $\mathcal{S}$ são:

$$\begin{bmatrix} 3\sqrt{2}/2 \\ 2 \\ \sqrt{2}/2 \end{bmatrix}$$

5.3.2. (a)
```
>> v1=sym([-1/sqrt(2),1/sqrt(2)]);
>> v2=sym([1/sqrt(2),1/sqrt(2)]);
>> v=2*v1+v2
```
$\begin{bmatrix} -\sqrt{2}/2 & 3\sqrt{2}/2 \end{bmatrix}$

(b)
```
>> v1=sym([0,1/sqrt(2),-1/sqrt(2)]);
>> v2=sym([1,0,0]);
>> v3=sym([0,1/sqrt(2),1/sqrt(2)]);
>> v=-v1+v2+2*v3
v =      3            1            3
```
$\begin{bmatrix} 1 & \sqrt{2}/2 & 3\sqrt{2}/2 \end{bmatrix}$

5.3.3. As coordenadas de U_1, U_2 e U_3 em relação ao sistema $\mathcal{S} = \{O, U_1, U_2, U_3\}$ são dadas por $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ e $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, respectivamente. Assim, $U_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $U_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/2 \\ \sqrt{3}/2 \end{bmatrix}$ e $U_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\sqrt{3}/2 \\ 1/2 \end{bmatrix}$

5.3.4.
```
>> p=sym([sqrt(3),1]).'; pr=sym([sqrt(3),-1]).';
>> A=[cos(th),-sin(th);sin(th),cos(th)];
>> expr=A*pr-p
expr = [ cos(th)*3^(1/2)+sin(th)-3^(1/2)]
       [       sin(th)*3^(1/2)-cos(th)-1]
```

```
>> solve(expr(1,1),expr(2,1),th)
ans = 1/3*pi
```

A rotação é de $\pi/3$.

6.1. Definição, Exemplos e Propriedades (página 355)

6.1.1. (a)
```
>> V1=[1;1];V2=[0;1]; A=[V1,V2];
>> escalona([A,[3;-2]])
[  1,  0,  3]
[  0,  1, -5]
```

$(3,-2) = 3V_1 - 5V_2$. Assim, $T(3,-2) = 3T(V_1) - 5T(V_2) = 3(2,-3) - 5(1,2) = (1,-19)$.

(b)
```
>> syms a b
>> escalona([A,[a;b]])
[   1,   0,   a]
[   0,   1, b-a]
```

$(a,b) = aV_1 + (b-a)V_2$. Assim, $T(a,b) = aT(V_1) + (b-a)T(V_2) = a(2,-3) + (b-a)(1,2) = (a+b,-5a+2b)$.

6.1.2.
```
>> V1=[1;1];V2=[0;-2]; A=[V1,V2];
>> escalona([A,[1;0],[0;1]])
[    1,    0,    1,    0]
[    0,    1,  1/2, -1/2]
```

$(1,0) = V_1 + 1/2V_2$ e $(0,1) = -1/2V_2$. Assim, $T(1,0) = T(V_1) + 1/2T(V_2) = (3,2,1) + 1/2(0,1,0) = (3,5/2,1)$ e $T(0,1) = -1/2T(V_2) = -1/2(0,1,0) = (0,-1/2,0)$.

6.1.3. $P_{xy}(x,y,z) = (x,y,0)$, $P_{yz}(x,y,z) = (0,y,z)$ e $P_{xz}(x,y,z) = (x,0,z)$.

6.1.4. (a) $P_\pi(x,y,z) = (x,y,z) - P_{(1,2,3)}(x,y,z) = \frac{1}{14}(13x-2y-3z,-2x+10y-6z,-3x-6y+5z)$.

(b) $R_\pi(x,y,z) = 2P_\pi(x,y,z) - (x,y,z) = (x,y,z) - 2P_{(1,2,3)}(x,y,z) = 1/7(6x-2y-3z,3y-2x-6z,-2z-3x-6y)$.

6.1.5. $R_{\pi/3,x}(E_1) = E_1, R_{\pi/3,x}(E_2) = 1/2E_2 + \sqrt{3}/2E_3, R_{\pi/3,x}(E_3) = -\sqrt{3}/2E_2 + 1/2E_3$. Portanto,

$$R_{\pi/3,x}\begin{bmatrix} x\\ y\\ z\end{bmatrix} = \begin{bmatrix} 1 & 0 & 0\\ 0 & 1/2 & -\sqrt{3}/2\\ 0 & \sqrt{3}/2 & 1/2\end{bmatrix}\begin{bmatrix} x\\ y\\ z\end{bmatrix}.$$

$R_{\pi/3,y}(E_2) = E_2, R_{\pi/3,y}(E_1) = 1/2E_1 - \sqrt{3}/2E_3, R_{\pi/3,y}(E_3) = \sqrt{3}/2E_1 + 1/2E_3$. Portanto,

$$R_{\pi/3,y}\begin{bmatrix} x\\ y\\ z\end{bmatrix} = \begin{bmatrix} 1/2 & 0 & \sqrt{3}/2\\ 0 & 1 & 0\\ -\sqrt{3}/2 & 0 & 1/2\end{bmatrix}\begin{bmatrix} x\\ y\\ z\end{bmatrix}.$$

$R_{\pi/3,z}(E_3) = E_3, R_{\pi/3,z}(E_1) = 1/2E_1 + \sqrt{3}/2E_2, R_{\pi/3,z}(E_2) = -\sqrt{3}/2E_1 + 1/2E_2$. Portanto,

$$R_{\pi/3,z}\begin{bmatrix} x\\ y\\ z\end{bmatrix} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 & 0\\ \sqrt{3}/2 & 1/2 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} x\\ y\\ z\end{bmatrix}.$$

6.1.6. (a) $P_r(x,y,z) = \text{proj}_{(1,1,1)}(x,y,z) = \frac{\langle (x,y,z),(1,1,1)\rangle}{||(1,1,1)||^2} = \frac{x+y+z}{3}(1,1,1) = (\frac{x+y+z}{3}, \frac{x+y+z}{3}, \frac{x+y+z}{3})$.

(b) $R_r(x,y,z) = 2\text{proj}_{(1,1,0)}(x,y,z) - (x,y,z) = 2(\frac{x+y+z}{3}, \frac{x+y+z}{3}, \frac{x+y+z}{3}) - (x,y,z) = (\frac{-x+2y+2z}{3}, \frac{2x-y+2z}{3}, \frac{2x+2y-z}{3})$.

6.1.7. Sejam $P = (5,0)$ e $Q = (3,4)$. O eixo da reflexão é uma reta r, perpendicular ao vetor $\overrightarrow{PQ} = (-2,4)$ que passa pelo ponto médio, $M = \frac{Q+P}{2} = (4,2)$. Assim, o eixo é uma reta de equação $-2x + 4y + c = 0$. Substituindo-se o ponto M na equação da reta é $-2x + 4y = 0$. $R_r(5,0) = (3,4)$ e $R_r(4,2) = (4,2)$.

```
>> V1=[5;0];V2=[4;2]; A=[V1,V2];
>> escalona([A,[1;0],[0;1]])
[    1,    0,  1/5, -2/5]
[    0,    1,    0,  1/2]
```

Assim, $R_r(1,0) = 1/5R_r(5,0) = 1/5(3,4) = (3/5,4/5)$ e $R_r(0,1) = -2/5R_r(5,0) + 1/2R_r(4,2) = -2/5(3,4) + 1/2(4,2) = (4/5,-3/5)$. A matriz de R_r em relação à base canônica é $[R_r]_{\mathcal{B}}^{\mathcal{B}} =$

$$\begin{bmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{bmatrix}.$$

6.2. A Imagem e o Núcleo (página 369)

6.2.1. Esta transformação é a projeção no plano xy. Se a imagem da reta r é um ponto, então a reta é perpendicular ao plano xy. Assim, as equações paramétricas da reta r são da forma $x = x_0, y = y_0, z = z_0 + ct, \forall t \in \mathbb{R}$, em que $(x_0, y_0, z_0) = P$.

6.2.2. (a) Seja $\mathcal{B} = \{E_1, E_2, E_3\}$ a base canônica. $T(X) = [T(E_1)T(E_2)T(E_3)]X = \begin{bmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} X.$

```
A=[0,0,1;1,-1,0;0,0,-1];
>> escalona(A)
[  1, -1,  0]
[  0,  0,  1]
[  0,  0,  0]
```

$\mathcal{N}(T) = \{\alpha(1,1,0) \mid \alpha \in \mathbb{R}\}$. $\{(1,1,0)\}$ é base para o núcleo de T, pois é L.I. e gera o núcleo de T.

(b) A imagem de T é gerada por $T(E_1), T(E_2)$ e $T(E_3)$, ou seja, pelas colunas da matriz $[T]_{\mathcal{B}}^{\mathcal{B}}$. As colunas associadas aos pivôs (primeira e terceira) são L.I. Assim, $(0,1,0)$ e $(1,0,-1)$ formam uma base para a imagem de T.

(c) O núcleo de T é uma reta que passa pela origem e tem vetor diretor $(-1,1,0)$ e a imagem de T é o plano que passa pela origem que é paralelo aos vetores $(0,1,0)$ e $(1,0,-1)$.

6.2.3. (a) $n = \dim(\mathcal{N}(T)) + \dim(\mathcal{I}(T)) = 2 + 5 = 7$.

(b) $n = \dim(\mathcal{N}(T)) + \dim(\mathcal{I}(T)) = 0 + 5 = 5$.

6.2.4. (a) $\mathcal{N}(T) = \{(x, y, -x) \mid x, y \in \mathbb{R}\}$. $\{(1,0,-1), (0,1,0)\}$ é uma base para o núcleo de T. Seja T uma transformação linear tal que $T(1,0,-1) = (0,0,0), T(0,1,0) = (0,0,0)$. Seja V um vetor que não pertence ao espaço gerado por $(1,0,-1)$ e $(0,1,0)$. Defina $T(V) = W$, onde $W \neq (0,0,0)$. Por exemplo, tomando $V = (1,0,0)$ e $T(1,0,0) = (1,0,0)$. A transformação T está totalmente caracterizada por seus valores em uma base do domínio.

(b) $\mathcal{I}(T) = \{(x,x,z) \mid x,z \in \mathbb{R}\}$. $\{(1,1,0),(0,0,1)\}$ é uma base para a imagem de T. Seja T uma transformação linear tal que $T(E_1) = (1,1,0), T(E_2) = (0,0,1)$. Seja W um vetor que pertence ao espaço gerado por $(1,1,0)$ e $(0,0,1)$. Defina $T(E_3) = W$. Por exemplo: $W = \bar{0}$, então $T(E_3) = \bar{0}$. A transformação T está totalmente caracterizada por seus valores em uma base do domínio.

6.2.5. Seja $\{V_1, V_2\}$ uma base de $\mathbb{R}^2$. Defina $T(V_1) = \bar{0}$ e $T(V_2) = V_1$. Ex.: $T(1,0) = (0,0)$ e $T(0,1) = (1,0)$.

6.3. Composição de Transformações Lineares (página 387)

6.3.1. Seja $\mathcal{B} = \{E_1, E_2\}$ a base canônica de $\mathbb{R}^2$. Como $P = [I]^{\mathcal{B}}_{\mathcal{C}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, $A = [T]^{\mathcal{B}}_{\mathcal{B}} = [T(E_1)T(E_2)] = \begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}$ e $P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$. Então, $[T]^{\mathcal{C}}_{\mathcal{C}} = [I]^{\mathcal{C}}_{\mathcal{B}}[T]^{\mathcal{B}}_{\mathcal{B}}[I]^{\mathcal{B}}_{\mathcal{C}} = P^{-1}AP = \begin{bmatrix} 8 & 13 \\ -5 & -9 \end{bmatrix}$.

6.3.2. (a) $P = [I]^{\mathcal{B}}_{\mathcal{C}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 1 \end{bmatrix}$.

(b) $B = [T]^{\mathcal{C}}_{\mathcal{C}} = [I]^{\mathcal{C}}_{\mathcal{B}}[T]^{\mathcal{B}}_{\mathcal{B}}[I]^{\mathcal{B}}_{\mathcal{C}} = P^{-1}AP$. Multiplicando-se à esquerda por P, $PB = AP$. Seja $B = [X_1 X_2 X_3]$. Para encontrar B basta resolvermos os sistemas lineares $PX_1 = AP, PX_2 = AP, PX_3 = AP$ que podem ser resolvidos simultaneamente escalonando a matriz aumentada $[P|AP]$.

```
>> A=[3,-1,-2;0,0,-2;0,0,-1];
>> P=[1,1,0;0,2,-2;0,0,1];
>> escalona([P,A*P])
[  1,  1,  0,  3,  1,  0]
[  0,  2, -2,  0,  0, -2]
[  0,  0,  1,  0,  0, -1]
eliminação 2:
(1/2)*linha 2 ==> linha 2
[  1,  1,  0,  3,  1,  0]
```

```
[  0,  1, -1,  0,  0, -1]
[  0,  0,  1,  0,  0, -1]
(-1)*linha 2 + linha 1 ==> linha 1
[  1,  0,  1,  3,  1,  1]
[  0,  1, -1,  0,  0, -1]
[  0,  0,  1,  0,  0, -1]
eliminação 3:
(-1)*linha 3 + linha 1 ==> linha 1
(1)*linha 3 + linha 2 ==> linha 2
[  1,  0,  0,  3,  1,  2]
[  0,  1,  0,  0,  0, -2]
[  0,  0,  1,  0,  0, -1]
```

Assim, $B = [X_1 X_2 X_3] = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & -1 \end{bmatrix}$.

6.3.3. (a) $P_r(U_1) = U_1 = 1U_1 + 0U_2 + 0U_3$,
$P_r(U_2) = \bar{0} = 0U_1 + 0U_2 + 0U_3$ e
$P_r(U_3) = \bar{0} = 0U_1 + 0U_2 + 0U_3$.

Assim, $[P_r]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Seja $P = [I]_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} \sqrt{3}/3 & \sqrt{6}/3 & 0 \\ \sqrt{3}/3 & -\sqrt{6}/6 & \sqrt{2}/2 \\ \sqrt{3}/3 & -\sqrt{6}/3 & -\sqrt{2}/2 \end{bmatrix}$.

$P_r(E_1) = \text{proj}_{(1,1,1)} E_1 = \frac{\langle E_1,(1,1,1)\rangle}{||(1,1,1)||^2}(1,1,1) = (1/3, 1/3, 1/3) = 1/3E_1 + 1/3E_2 + 1/3E_3$
$P_r(E_2) = \text{proj}_{(1,1,1)} E_2 = \frac{\langle E_2,(1,1,1)\rangle}{||(1,1,1)||^2}(1,1,1) = (1/3, 1/3, 1/3) = 1/3E_1 + 1/3E_2 + 1/3E_3$

$P_r(E_3) = \text{proj}_{(1,1,1)} E_3 = \frac{\langle E_3,(1,1,1)\rangle}{||(1,1,1)||^2}(1,1,1) = (1/3, 1/3, 1/3) = 1/3E_1 + 1/3E_2 + 1/3E_3$
Assim,

$$[P_r]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

(b) $R_r(U_1) = U_1 = 1U_1 + 0U_2 + 0U_3$,
$R_r(U_2) = -U_2 = 0U_1 - 1U_2 + 0U_3$ e
$R_r(U_3) = -U_3 = 0U_1 + 0U_2 - 1U_3$.

Assim, $[R_r]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

$R_r(E_1) = 2\text{proj}_{(1,1,1)} E_1 - E_1 = 2\frac{\langle E_1,(1,1,1)\rangle}{||(1,1,1)||^2}(1,1,1) - (1,0,0) = (-1/3, 2/3, 2/3) = -1/3E_1 + 2/3E_2 + 2/3E_3$
$R_r(E_2) = 2\text{proj}_{(1,1,1)} E_2 - E_2 = 2\frac{\langle E_2,(1,1,1)\rangle}{||(1,1,1)||^2}(1,1,1) - (0,1,0) = (2/3, -1/3, 2/3) = 2/3E_1 - 1/3E_2 + 2/3E_3$
$R_r(E_3) = 2\text{proj}_{(1,1,1)} E_3 - E_3 = 2\frac{\langle E_3,(1,1,1)\rangle}{||(1,1,1)||^2}(1,1,1) - (0,0,1) = (2/3, 2/3, -1/3) = 2/3E_1 + 2/3E_2 - 1/3E_3$

Assim, $[R_r]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$.

6.3.4. $R_{\pi/2,r}(U_1) = U_1 = 1U_1 + 0U_2 + 0U_3$,

$R_{\pi/2,r}(U_2) = U_3 = 0U_1 + 0U_2 + 1U_3$,

$R_{\pi/2,r}(U_3) = -U_2 = 0U_1 - 1U_2 + 0U_3$.

Assim, $[R_{\pi/2,r}]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$.

$E_1 = \langle E_1, U_1 \rangle U_1 + \langle E_1, U_2 \rangle U_2 + \langle E_1, U_3 \rangle U_3 = 1/\sqrt{2}U_1 + 1/\sqrt{2}U_3$

$R_{\pi/2,r}(E_1) = 1/\sqrt{2}R_{\pi/2,r}(U_1) + 1/\sqrt{2}R_{\pi/2,r}(U_3) = 1/\sqrt{2}U_1 - 1/\sqrt{2}U_2 = 1/\sqrt{2}(1/\sqrt{2}, 1/\sqrt{2}, 0) - 1/\sqrt{2}(0,0,1) = (1/2, 1/2, -1/\sqrt{2})$

$E_2 = \langle E_2, U_1 \rangle U_1 + \langle E_2, U_2 \rangle U_2 + \langle E_2, U_3 \rangle U_3 = 1/\sqrt{2}U_1 - 1/\sqrt{2}U_3$

$R_{\pi/2,r}(E_2) = 1/\sqrt{2}R_{\pi/2,r}(U_1) - 1/\sqrt{2}R_{\pi/2,r}(U_3) = 1/\sqrt{2}U_1 + 1/\sqrt{2}U_3 = 1/\sqrt{2}(1/\sqrt{2}, 1/\sqrt{2}, 0) + 1/\sqrt{2}(0,0,1) = (1/2, 1/2, 1/\sqrt{2})$

$E_3 = \langle E_3, U_1 \rangle U_1 + \langle E_3, U_2 \rangle U_2 + \langle E_3, U_3 \rangle U_3 = U_2$

$R_{\pi/2,r}(E_3) = R_{\pi/2,r}(U_2) = U_3 = (1/\sqrt{2}, -1/\sqrt{2}, 0)$

Assim, $[R_{\pi/2,r}]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 1/2 & 1/2 & 1/\sqrt{2} \\ 1/2 & 1/2 & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}$.

6.3.5. (a) A matriz de T em relação à base canônica é $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Assim,

```
≫ A=[1,2,1;0,1,2;0,0,1];
≫ escalona([A,eye(3)])
[ 1, 2, 1, 1, 0, 0]
[ 0, 1, 2, 0, 1, 0]
[ 0, 0, 1, 0, 0, 1]
eliminação 2:
(-2)*linha 2 + linha 1 ==> linha 1
[  1,  0, -3,  1, -2,  0]
```

```
[  0,  1,  2,  0,  1,  0]
[  0,  0,  1,  0,  0,  1]
eliminação 3:
(3)*linha 3 + linha 1 ==> linha 1
(-2)*linha 3 + linha 2 ==> linha 2
[  1,  0,  0,  1, -2,  3]
[  0,  1,  0,  0,  1, -2]
[  0,  0,  1,  0,  0,  1]
```

Portanto, a inversa de T é dada por $T^{-1}(x,y,z) = (x - 2y + 3z, y - 2z, z)$.

(b) A matriz de T em relação à base $\{E_1, E_2, E_3\}$ é $A = \begin{bmatrix} 1 & -2 & -2 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}$. Assim,

```
≫ A=[1,-2,-2;0,1,-4;0,0,1];
≫ escalona([A,eye(3)])
[  1, -2, -2,  1,  0,  0]
[  0,  1, -4,  0,  1,  0]
[  0,  0,  1,  0,  0,  1]
eliminação 2:
(2)*linha 2 + linha 1 ==> linha 1
[   1,   0, -10,   1,   2,   0]
[   0,   1,  -4,   0,   1,   0]
[   0,   0,   1,   0,   0,   1]
eliminação 3:
(10)*linha 3 + linha 1 ==> linha 1
(4)*linha 3 + linha 2 ==> linha 2
[  1,  0,  0,  1,  2, 10]
[  0,  1,  0,  0,  1,  4]
[  0,  0,  1,  0,  0,  1]
```

Assim, $T^{-1}(a,b,c) = (a + 2b + 10c, b + 4c, c)$.

(c) A matriz de T em relação à base $\{E_1, E_2, E_3\}$ é $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix}$. Assim,

```
≫ A=[1,1,1;1,2,1;1,0,2];
≫ escalona([A,eye(3)])
[ 1, 1, 1, 1, 0, 0]
[ 1, 2, 1, 0, 1, 0]
[ 1, 0, 2, 0, 0, 1]
eliminação 1:
(-1)*linha 1 + linha 2 ==> linha 2
(-1)*linha 1 + linha 3 ==> linha 3
[  1,  1,  1,  1,  0,  0]
[  0,  1,  0, -1,  1,  0]
[  0, -1,  1, -1,  0,  1]
eliminação 2:
(-1)*linha 2 + linha 1 ==> linha 1
(1)*linha 2 + linha 3 ==> linha 3
[  1,  0,  1,  2, -1,  0]
[  0,  1,  0, -1,  1,  0]
[  0,  0,  1, -2,  1,  1]
eliminação 3:
(-1)*linha 3 + linha 1 ==> linha 1
[  1,  0,  0,  4, -2, -1]
[  0,  1,  0, -1,  1,  0]
[  0,  0,  1, -2,  1,  1]
```

Portanto, a inversa de T é dada por $T^{-1}(a, b, c) = (4a - 2b - c, -a + b, -2a + b + c)$.

(d) A matriz de T em relação à base $\{E_1, E_2, E_3\}$ é $\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. Assim,

```
≫ A=[1,-1,1;1,0,0;1,1,1];
≫ escalona([A,eye(3)])
[  1, -1,  1,  1,  0,  0]
[  1,  0,  0,  0,  1,  0]
[  1,  1,  1,  0,  0,  1]
eliminação 1:
(-1)*linha 1 + linha 2 ==> linha 2
(-1)*linha 1 + linha 3 ==> linha 3
[  1, -1,  1,  1,  0,  0]
[  0,  1, -1, -1,  1,  0]
[  0,  2,  0, -1,  0,  1]
eliminação 2:
(1)*linha 2 + linha 1 ==> linha 1
(-2)*linha 2 + linha 3 ==> linha 3
[  1,  0,  0,  0,  1,  0]
[  0,  1, -1, -1,  1,  0]
[  0,  0,  2,  1, -2,  1]
eliminação 3:
(1/2)*linha 3 ==> linha 3
[   1,   0,   0,   0,   1,   0]
[   0,   1,  -1,  -1,   1,   0]
[   0,   0,   1, 1/2,  -1, 1/2]
(1)*linha 3 + linha 2 ==> linha 2
[    1,    0,    0,    0,    1,    0]
[    0,    1,    0, -1/2,    0,  1/2]
[    0,    0,    1,  1/2,   -1,  1/2]
```

Portanto, a inversa de T é dada por $T^{-1}(a, b, c) = (b, c/2 - a/2, a/2 - b + c/2)$.

7.1. Diagonalização de Matrizes (página 421)

7.1.1.

(a)
```
>> A=[1,1;1,1];
>> B=A-x*eye(2)
[1-x,   1]
[  1, 1-x]
>> p=det(B)
p =-2*x+x^2
>> solve(p)
[0][2]
>> B0=subs(B,x,0)
[1, 1]
[1, 1]
>> escalona(B0)
     1     1
     0     0
>> B2=subs(B,x,2)
[-1,  1]
[ 1, -1]
>> escalona(B2)
     1    -1
     0     0
```

$$\mathbb{V}_0 = \{(-\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$$
$$\mathbb{V}_2 = \{(\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$$

(b)
```
>> A=[1,-1;2,4];
>> B=A-x*eye(2)
[1-x,  -1]
[  2, 4-x]
>> p=det(B)
p =6-5*x+x^2
>> solve(p)
[3][2]
>> B2=subs(B,x,2)
[-1, -1]
[ 2,  2]
>> escalona(B2)
     1     1
     0     0
>> B3=subs(B,x,3)
[-2, -1]
[ 2,  1]
>> escalona(B3)
      1           1/2
      0            0
```

$$\mathbb{V}_2 = \{(-\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$$
$$\mathbb{V}_3 = \{(-\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$$

(c)

```
>> A=[0,1,2;0,0,3;0,0,0];
>> B=A-x*eye(3)
[-x,  1,  2]
[ 0, -x,  3]
[ 0,  0, -x]
>> p=det(B)
p=-x^3
>> solve(p)
[0][0][0]
```

```
>> B0=subs(B,x,0)
[0, 1, 2]
[0, 0, 3]
[0, 0, 0]
>> escalona(B0)
[0, 1, 0]
[0, 0, 1]
[0, 0, 0]
```

$$\mathbb{V}_0 = \{(\alpha, 0, 0) \mid \alpha \in \mathbb{R}\}$$

(d)

```
>> A=[1,0,0;-1,3,0;3,2,-2];
>> B=A-x*eye(3)
[1-x,   0,    0]
[ -1, 3-x,    0]
[  3,   2, -2-x]
>> p=det(B)
p =(1-x)*(3-x)*(-2-x)
>> solve(p)
[ 1][ 3][-2]
>> B1=subst(B,x,1)
[ 0, 0,  0]
[-1, 2,  0]
[ 3, 2, -3]
>> escalona(B1)
[1, 0, -3/4]
[0, 1, -3/8]
[0, 0,    0]
```

```
>> Bm2=subs(B,x,-2)
[ 3, 0, 0]
[-1, 5, 0]
[ 3, 2, 0]
>> escalona(Bm2)
[1, 0, 0]
[0, 1, 0]
[0, 0, 0]
>> B3=subs(B,x,3)
[-2, 0,  0]
[-1, 0,  0]
[ 3, 2, -5]
>> escalona(B3)
[1, 0,    0]
[0, 1, -5/2]
[0, 0,    0]
```

$$\mathbb{V}_{-2} = \{(0, 0, \alpha) \mid \alpha \in \mathbb{R}\}$$

$$\mathbb{V}_1 = \{(6\alpha, 3\alpha, 8\alpha) \mid \alpha \in \mathbb{R}\}$$

$$\mathbb{V}_3 = \{(0, 5\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$$

(e)

```
>> A=[2,-2,3;0,3,-2;0,-1,2];
>> B=A-x*eye(3)
[2-x,  -2,   3]
[  0, 3-x,  -2]
[  0,  -1, 2-x]
>> p=det(B)
p =(2-x)*(4-5*x+x^2)
>> solve(p)
[2][4][1]
>> B1=subs(B,x,1)
[1, -2,  3]
[0,  2, -2]
[0, -1,  1]
>> escalona(B1)
[1, 0,  1]
[0, 1, -1]
[0, 0,  0]
```

```
>> B2=subs(B,x,2)
[0, -2,  3]
[0,  1, -2]
[0, -1,  0]
>> escalona(B2)
[0, 1, 0]
[0, 0, 1]
[0, 0, 0]
>> B4=subs(B,x,4)
[-2, -2,  3]
[ 0, -1, -2]
[ 0, -1, -2]
>> escalona(B4)
[1, 0, -7/2]
[0, 1,    2]
[0, 0,    0]
```

$$\mathbb{V}_1 = \{(-\alpha, \alpha, \alpha) \mid \alpha \in \mathbb{R}\}$$

$$\mathbb{V}_2 = \{(\alpha, 0, 0) \mid \alpha \in \mathbb{R}\}$$

$$\mathbb{V}_4 = \{(7\alpha, -4\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$$

(f)

```
>> A=[2,2,3;1,2,1;2,-2,1];
>> B=A-x*eye(3)
[2-x,   2,   3]
[  1, 2-x,   1]
[  2,  -2, 1-x]
>> p=det(B)
p =-8-2*x+5*x^2-x^3
>> solve(p)
[ 2][ 4][-1]
>> Bm1=subs(B,x,-1)
[3,  2, 3]
[1,  3, 1]
[2, -2, 2]
>> escalona(Bm1)
[1, 0, 1]
[0, 1, 0]
[0, 0, 0]
```

```
>> B2=subs(B,x,2)
[0,  2,  3]
[1,  0,  1]
[2, -2, -1]
>> escalona(B2)
[1, 0,   1]
[0, 1, 3/2]
[0, 0,   0]
>> B4=subs(B,x,4)
[-2,  2,  3]
[ 1, -2,  1]
[ 2, -2, -3]
>> escalona(B4)
[1, 0,   -4]
[0, 1, -5/2]
[0, 0,    0]
```

$$\mathbb{V}_{-1} = \{(-\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}, \mathbb{V}_2 = \{(-2\alpha, -3\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\} \text{ e } \mathbb{V}_4 = \{(8\alpha, 5\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$$

7.1.2. (a)

```
>> A=[2,0,0;3,-1,0;0,4,3];
>> B=A-x*eye(3)
[2-x,    0,    0]
[  3, -1-x,    0]
[  0,    4, 3-x]
>> p=det(B)
p =(2-x)*(-1-x)*(3-x)
>> solve(p)
[ 2][-1][ 3]
>> Bm1=subs(B,x,-1)
[3, 0, 0]
[3, 0, 0]
[0, 4, 4]
>> escalona(Bm1)
[1, 0, 0]
[0, 1, 1]
[0, 0, 0]
```

```
>> B2=subs(B,x,2)
[0,  0, 0]
[3, -3, 0]
[0,  4, 1]
>> escalona(B2)
[1, 0, 1/4]
[0, 1, 1/4]
[0, 0,   0]
>> B3=subst(B,x,3)
[-1,  0, 0]
[ 3, -4, 0]
[ 0,  4, 0]
>> escalona(B3)
[1, 0, 0]
[0, 1, 0]
[0, 0, 0]
```

$\mathbb{V}_{-1} = \{(0, -\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{(0, -1, 1)\}$ é base para $\mathbb{V}_{-1}$, pois gera $\mathbb{V}_{-1}$ $((0, -\alpha, \alpha) = \alpha(0, -1, 1))$ e um vetor não nulo é L.I.

$\mathbb{V}_2 = \{(-\alpha, -\alpha, 4\alpha) \mid \alpha \in \mathbb{R}\}$. $\{(-1, -1, 4)\}$ é base para $\mathbb{V}_2$, pois gera $\mathbb{V}_2$ $((-\alpha, -\alpha, 4\alpha) = \alpha(-1, -1, 4))$ e um vetor não nulo é L.I.

$\mathbb{V}_3 = \{(0, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{(0, 0, 1)\}$ é base para $\mathbb{V}_3$, pois gera $\mathbb{V}_3$ $((0, 0, \alpha) = \alpha(0, 0, 1))$ e um vetor não nulo é L.I.

(b)

```
>> A=[2,3,0;0,1,0;0,0,2];
>> B=A-x*eye(3)
[2-x,   3,   0]
[  0, 1-x,   0]
[  0,   0, 2-x]
>> p=det(B)
p =(2-x)^2*(1-x)
>> solve(p)
[2][2][1]
>> B1=subs(B,x,1)
[1, 3, 0]
[0, 0, 0]
[0, 0, 1]
```

```
>> escalona(B1)
[1, 3, 0]
[0, 0, 1]
[0, 0, 0]
>> B2=subs(B,x,2)
[0,  3, 0]
[0, -1, 0]
[0,  0, 0]
>> escalona(B2)
[0, 1, 0]
[0, 0, 0]
[0, 0, 0]
```

$\mathbb{V}_1 = \{(-3\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}$. $\{(-3, 1, 0)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ ($(-3\alpha, \alpha, 0) = \alpha(-3, 1, 0)$) e um vetor não nulo é L.I.

$\mathbb{V}_2 = \{(\alpha, 0, \beta) \mid \alpha, \beta \in \mathbb{R}\}$. $\{V_1 = (1, 0, 0), V_2 = (0, 0, 1)\}$ é base para $\mathbb{V}_2$, pois gera $\mathbb{V}_2$ ($(\alpha, 0, \beta) = \alpha(1, 0, 0) + \beta(0, 0, 1)$) e é L.I. ($xV_1 + yV_2 = \bar{0}$ se, e somente se, $(x, 0, y) = (0, 0, 0)$ ou $x = 0$ e $y = 0$).

(c)

```
>> A=[1,2,3,4;0,-1,3,2;0,0,3,3;0,0,0,2];
>> B=A-x*eye(4)
[1-x,    2,   3,   4]
[  0, -1-x,   3,   2]
[  0,    0, 3-x,   3]
[  0,    0,   0, 2-x]
>> p=det(B)
p =(1-x)*(2-x)*(-1-x)*(3-x)
>> solve(p)
[ 1][ 2][-1][ 3]
```

```
>> Bm1=subs(B,x,-1)
[2, 2, 3, 4]
[0, 0, 3, 2]
[0, 0, 4, 3]
[0, 0, 0, 3]
>> escalona(Bm1)
[1, 1, 0, 0]
[0, 0, 1, 0]
[0, 0, 0, 1]
[0, 0, 0, 0]
```

```
>> B1=subs(B,x,1)
[0,  2, 3, 4]
[0, -2, 3, 2]
[0,  0, 2, 3]
[0,  0, 0, 1]
>> escalona(B1)
[0, 1, 0, 0]
[0, 0, 1, 0]
[0, 0, 0, 1]
[0, 0, 0, 0]
```

```
>> B2=subs(B,x,2)
[-1,  2, 3, 4]
[ 0, -3, 3, 2]
[ 0,  0, 1, 3]
[ 0,  0, 0, 0]
>> escalona(B2)
[1, 0, 0, 29/3]
[0, 1, 0,  7/3]
[0, 0, 1,    3]
[0, 0, 0,    0]
```

```
>> B3=subst(B,x,3)
[-2,  2, 3,  4]
[ 0, -4, 3,  2]
[ 0,  0, 0,  3]
[ 0,  0, 0, -1]
>> escalona(B3)
[1, 0, -9/4, 0]
[0, 1, -3/4, 0]
[0, 0,    0, 1]
[0, 0,    0, 0]
```

$\mathbb{V}_{-1} = \{(-\alpha,\alpha,0,0) \mid \alpha \in \mathbb{R}\}$. $\{(-1,1,0,0)\}$ é base para $\mathbb{V}_{-1}$, pois gera $\mathbb{V}_{-1}$ ($(-\alpha,\alpha,0,0) = \alpha(-1,1,0,0)$) e um vetor não nulo é L.I.

$\mathbb{V}_1 = \{(\alpha,0,0,0) \mid \alpha \in \mathbb{R}\}$. $\{(1,0,0,0)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ ($(\alpha,0,0,0) = \alpha(1,0,0,0)$) e um vetor não nulo é L.I.

$\mathbb{V}_2 = \{(-29\alpha,-7\alpha,-9\alpha,3\alpha) \mid \alpha \in \mathbb{R}\}$. $\{(-29,-7,-9,3)\}$ é base para $\mathbb{V}_2$, pois gera $\mathbb{V}_2$ ($(-29\alpha,-7\alpha,-9\alpha,3\alpha) = \alpha(-29,-7,-9,3)$) e um vetor não nulo é L.I.

$\mathbb{V}_3 = \{(9\alpha,3\alpha,4\alpha,0) \mid \alpha \in \mathbb{R}\}$. $\{(9,3,4,0)\}$ é base para $\mathbb{V}_3$, pois gera $\mathbb{V}_3$ ($(9\alpha,3\alpha,4\alpha,0) = \alpha(9,3,4,0)$) e um vetor não nulo é L.I.

(d)

```
>> A=[2,2,3,4;0,2,3,2;0,0,1,1;0,0,0,1];
>> B=A-x*eye(4)
[2-x,   2,   3,   4]
[  0, 2-x,   3,   2]
[  0,   0, 1-x,   1]
[  0,   0,   0, 1-x]
>> p=det(B)
p =(2-x)^2*(1-x)^2
>> solve(p)
[2][2][1][1]
```

```
>> B1=subs(B,x,1)
[1, 2, 3, 4]
[0, 1, 3, 2]
[0, 0, 0, 1]
[0, 0, 0, 0]
>> escalona(B1)
[1, 0, -3, 0]
[0, 1,  3, 0]
[0, 0,  0, 1]
[0, 0,  0, 0]
```

```
>> B2=subs(B,x,2)
[0, 2,  3,  4]
[0, 0,  3,  2]
[0, 0, -1,  1]
[0, 0,  0, -1]
>> escalona(B2)
[0, 1, 0, 0]
[0, 0, 1, 0]
[0, 0, 0, 1]
[0, 0, 0, 0]
```

$\mathbb{V}_1 = \{(3\alpha, -3\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}$. $\{(3, -3, 1, 0)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ ($(3\alpha, -3\alpha, \alpha, 0) = \alpha(3, -3, 1, 0)$) e um vetor não nulo é L.I.

$\mathbb{V}_2 = \{(\alpha, 0, 0, 0) \mid \alpha \in \mathbb{R}\}$. $\{(1, 0, 0, 0)\}$ é base para $\mathbb{V}_2$, pois gera $\mathbb{V}_2$ ($(\alpha, 0, 0, 0) = \alpha(1, 0, 0, 0)$) e um vetor não nulo é L.I.

7.1.3.

(a)
```
>> A=[1,4;1,-2];
>> B=A-x*eye(2)
[1-x,    4]
[  1, -2-x]
>> p=det(B)
p =-6+x+x^2
>> solve(p)
[ 2][-3]
```

A matriz A possui dois autovalores diferentes, logo possui dois autovetores L.I. (Proposição 7.3 na página 403). A matriz A é diagonalizável pois, é 2×2 e possui dois autovetores L.I. (Teorema 7.2 na página 401).

(b)
```
>> A=[1,0;-2,1];
>> B=A-x*eye(2)
[1-x,   0]
[ -2, 1-x]
>> p=det(B)
p =(1-x)^2
>> solve(p)
[1][1]
>> B1=subs(B,x,1)
[ 0, 0]
[-2, 0]
>> escalona(numeric(B1))
[1, 0]
[0, 0]
```

$$\mathbb{V}_1 = \{(\alpha, 0) \mid \alpha \in \mathbb{R}\}$$

A matriz A não é diagonalizável pois, não possui dois autovetores L.I. (Teorema 7.2 na página 401).

(c)
```
>> A=[1,1,-2;4,0,4;1,-1,4]
A =   1            1            -2
      4            0             4
      1           -1             4
>> B=A-x*eye(3); p=det(B)
p =5*x^2-6*x-x^3
>> solve(p)
ans =[0][2][3]
```

A matriz A possui três autovalores diferentes, logo possui três autovetores L.I. (Proposição 7.3 na página 403). A matriz A é diagonalizável pois, é 3×3 e possui três autovetores L.I. (Teorema 7.2 na página 401).

(d)
```
>> A=[1,2,3;0,-1,2;0,0,2];
>> B=A-x*eye(3)
[1-x,    2,   3]
[  0, -1-x,   2]
[  0,    0, 2-x]
```

```
>> p=det(B)
p =(1-x)*(-1-x)*(2-x)
>> solve(p)
[ 1][-1][ 2]
```

A matriz A possui três autovalores diferentes, logo possui três autovetores L.I. (Proposição 7.3 na página 403). A matriz A é diagonalizável pois, é 3×3 e possui três autovetores L.I. (Teorema 7.2 na página 401).

7.1.4.

(a)
```
>> A=[1,1,2;0,1,0;0,1,3];
>> B=A-x*eye(3)
[1-x,   1,   2]
[  0, 1-x,   0]
[  0,   1, 3-x]
>> p=det(B)
p =(1-x)^2*(3-x)
>> solve(p)
[1][1][3]
```

```
>> B1=subs(B,x,1)
[0, 1, 2]
[0, 0, 0]
[1, 1, 2]
>> escalona(B1)
[ 0, 1, 2]
[ 0, 0, 0]
[ 0, 0, 0]
```

```
>> B3=subs(B,x,3)
[ -2,  1,  2]
[  0, -2,  0]
[  0,  1,  0]
>> escalona(B3)
[  1,  0, -1]
[  0,  1,  0]
[  0,  0,  0]
```

$\mathbb{V}_1 = \{(\beta, -2\alpha, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$. $\{(1,0,0),(0,-2,1)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ ($(\beta,-2\alpha,\alpha) = \alpha(0,-2,1) + \beta(1,0,0)$) e são L.I. (um vetor não é múltiplo escalar do outro)

$\mathbb{V}_3 = \{((\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{(1,0,1)\}$ é base para $\mathbb{V}_3$, pois gera $\mathbb{V}_3$ ($(\alpha,0,\alpha) = \alpha(1,0,1)$) e um vetor não nulo é L.I.

$$P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

(b)
```
>> A=[4,2,3;2,1,2;-1,-2,0];
>> B=A-x*eye(3)
[4-x,   2,  3]
[  2, 1-x,  2]
[ -1,  -2, -x]
>> p=det(B)
p =-7*x+5*x^2+3-x^3
>> solve(p)
[3][1][1]
```

```
>> B1=subs(B,x,1)
[ 3,  2,  3]
[ 2,  0,  2]
[-1, -2, -1]
>> escalona(B1)
[1, 0, 1]
[0, 1, 0]
[0, 0, 0]
```

```
>> B3=subs(B,x,3)
[ 1,  2,  3]
[ 2, -2,  2]
[-1, -2, -3]
>> escalona(B3)
[1, 0, 5/3]
[0, 1, 2/3]
[0, 0,   0]
```

$\mathbb{V}_1 = \{(-\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{(-1, 0, 1)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ ($(-\alpha, 0, \alpha) = \alpha(-1, 0, 1)$) e um vetor não nulo é L.I.

$\mathbb{V}_2 = \{(-5\alpha, -2\alpha, 3\alpha) \mid \alpha \in \mathbb{R}\}$. $\{(-5, -2, 3)\}$ é base para $\mathbb{V}_2$, pois gera $\mathbb{V}_2$ ($(-5\alpha, -2\alpha, 3\alpha) = \alpha(-5, -2, 3)$) e um vetor não nulo é L.I.

A matriz não é diagonalizável pois só possui dois autovalores e cada um deles só possui um autovetor L.I. associado (Teorema 7.2 na página 401).

(c)
```
>> A=[1,2,3;0,1,0;2,1,2];
>> B=A-x*eye(3)
[1-x,   2,   3]
[  0, 1-x,   0]
[  2,   1, 2-x]
>> p=det(B)
p =-4+x+4*x^2-x^3
>> solve(p)
[ 1][ 4][-1]
>> Bm1=subs(B,x,-1)
[2, 2, 3]
[0, 2, 0]
[2, 1, 3]
>> escalona(Bm1)
[1, 0, 3/2]
[0, 1,   0]
[0, 0,   0]
>> B1=subst(B,x,1)
[0, 2, 3]
[0, 0, 0]
[2, 1, 1]
>> escalona(B1)
[1, 0, -1/4]
[0, 1,  3/2]
[0, 0,    0]
>> B4=subs(B,x,4)
[-3,  2,  3]
[ 0, -3,  0]
[ 2,  1, -2]
>> escalona(B4)
[1, 0, -1]
[0, 1,  0]
[0, 0,  0]
```

$\mathbb{V}_{-1} = \{(-3\alpha, 0, 2\alpha) \mid \alpha \in \mathbb{R}\}$. $\{(-3, 0, 2)\}$ é base para $\mathbb{V}_{-1}$, pois gera $\mathbb{V}_{-1}$ ($(-3\alpha, 0, 2\alpha) = \alpha(-3, 0, 2)$) e um vetor não nulo é L.I.

$\mathbb{V}_1 = \{(\alpha, -6\alpha, 4\alpha) \mid \alpha \in \mathbb{R}\}$. $\{(1, -6, 4)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ ($(\alpha, -6\alpha, 4\alpha) = \alpha(1, -6, 4)$) e um vetor não nulo é L.I.

$\mathbb{V}_4 = \{(\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{(1, 0, 1)\}$ é base para $\mathbb{V}_4$, pois gera $\mathbb{V}_4$ ($(\alpha, 0, \alpha) = \alpha(1, 0, 1)$) e um vetor não nulo é L.I.

$$P = \begin{bmatrix} -3 & 1 & 1 \\ 0 & -6 & 0 \\ 2 & 4 & 1 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

(d)
```
>> A=[3,-2,1;0,2,0;0,0,0];
>> B=A-x*eye(3)
[3-x,  -2,   1]
[  0, 2-x,   0]
[  0,   0,  -x]
>> p=det(B)
p =-(3-x)*(2-x)*x
>> solve(p)
[3][2][0]
>> B2=subs(B,x,2)
[1, -2,  1]
[0,  0,  0]
[0,  0, -2]
>> escalona(B2)
[1, -2, 0]
[0,  0, 1]
[0,  0, 0]
```

```
>> B0=subs(B,x,0)
[3, -2, 1]
[0,  2, 0]
[0,  0, 0]
>> escalona(B0)
[1, 0, 1/3]
[0, 1,   0]
[0, 0,   0]
>> B3=subs(B,x,3)
[0, -2,  1]
[0, -1,  0]
[0,  0, -3]
>> escalona(B3)
[0, 1, 0]
[0, 0, 1]
[0, 0, 0]
```

$\mathbb{V}_0 = \{(-\alpha, 0, 3\alpha) \mid \alpha \in \mathbb{R}\}$. $\{(-1, 0, 3)\}$ é base para $\mathbb{V}_0$, pois gera $\mathbb{V}_0$ $((-\alpha, 0, 3\alpha) = \alpha(-1, 0, 3))$ e um vetor não nulo é L.I.

$\mathbb{V}_2 = \{(2\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}$. $\{(2, 1, 0)\}$ é base para $\mathbb{V}_2$, pois gera $\mathbb{V}_2$ $((2\alpha, \alpha, 0) = \alpha(2, 1, 0))$ e um vetor não nulo é L.I.

$\mathbb{V}_3 = \{(\alpha, 0, 0) \mid \alpha \in \mathbb{R}\}$. $\{(1, 0, 0)\}$ é base para $\mathbb{V}_3$, pois gera $\mathbb{V}_3$ $((\alpha, 0, 0) = \alpha(1, 0, 0))$ e um vetor não nulo é L.I.

$$P = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 0 \\ 3 & 0 & 0 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

7.1.5. (a)
```
>> V1=[-4,-4,-1]';
>> V2=[5,4,1]';
```

```
>> V3=[5,3,1]';
>> A=sym([-1/3,-5/6,20/3;-2/3,-1/6,16/3;-1/6,-1/6,11/6]);
>> [A*V1,A*V2,A*V3]
[   -2,  5/3,  5/2] [   -2,  4/3,  3/2] [ -1/2,  1/3,  1/2]
```

Se V é autovetor de A então $AV = \lambda V$, ou seja, AV é um múltiplo escalar de V. Assim, concluímos que V_1 é autovetor associado a $\lambda_1 = 1/2$, V_2 é autovetor associado a $\lambda_2 = 1/3$ e V_3 é autovetor associado a $\lambda_3 = 1/2$.

(b) V_1 e V_3 são autovetores associados a $1/2$ e V_2 associado a $1/3$. Como $\{V_1, V_3\}$ é L.I. (um não é múltiplo escalar do outro) e ao juntarmos autovetores L.I. associados a diferentes autovalores eles continuam L.I., então a matriz A tem 3 autovetores L.I. Como ela é 3×3, é portanto diagonalizável.

7.1.6.

7.1.7. As matrizes dos itens (a) e (b) estão na forma canônica de Jordan.

7.1.8. (a) O polinômio característico de A é claramente $p(\lambda) = -(\lambda - 2)^2(\lambda + 1)^3$. Assim, os autovalores da matriz são $\lambda_1 = 2$ e $\lambda_2 = -1$.

```
>> A=[2,5,0,0,0;0,2,0,0,0;0,0,-1,0,-1;
0,0,0,-1,0;0,0,0,0,-1]; A=sym(A);
>> N21=null(A-2*eye(5))
N21 = -1
       0
       0
       0
       0
>> N22=null((A-2*eye(5))^2)
N22 =1     0
     0     1
     0     0
     0     0
```

```
     0     0
>> Nm11=null(A+eye(5))
Nm11 =0     0
      0     0
     -1     0
      0     1
      0     0
>> Nm12=null((A+eye(5))^2)
Nm12 =0     0     0
      0     0     0
      1     0     0
      0     1     0
      0     0     1
```

A menos da ordem dos blocos a única forma canônica de Jordan possível é

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

(b) O polinômio característico de A é claramente $p(\lambda) = (\lambda - 2)^5(\lambda + 1)$. Assim, os autovalores da matriz são $\lambda_1 = 2$ e $\lambda_2 = -1$.

```
>> A=[2,0,0,0,0,0;1,2,0,0,0,0;
-1,0,2,0,0,0;0,1,0,2,0,0;
1,1,1,1,2,0;0,0,0,0,1,-1]; A=sym(A);
>> N21=null(A-2*eye(6))
N21 =[  0,  0]
     [  0,  0]
     [  0, -1]
```

```
     [  0,  1]
     [  3,  0]
     [  1,  0]
>> N22=null((A-2*eye(6))^2)
N22 =[  0,  0,  0]
     [  0,  0,  0]
     [ -9, -1,  3]
     [  0,  1,  0]
     [  0,  0,  1]
     [  1,  0,  0]
>> N23=null((A-2*eye(6))^3)
[      0,      0,     0,     0]
[   -3/2, -27/2,  -3/2,   9/2]
[      0,      0,     1,     0]
[      1,      0,     0,     0]
[      0,      0,     0,     1]
[      0,      1,     0,     0]
>> N24=null((A-2*eye(6))^4)
N24 =[     1,     0,     0,     0,     0]
     [  -5/3,   9/2, -27/2,  -3/2,  -3/2]
     [     0,     0,     0,     0,     1]
     [     0,     0,     0,     1,     0]
     [     0,     1,     0,     0,     0]
     [     0,     0,     1,     0,     0]
```

A menos da ordem dos blocos a única forma canônica de Jordan possível é

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

(c) O polinômio característico de A é claramente $p(\lambda) = (\lambda+1)^5(\lambda+4)$. Assim, os autovalores da matriz são $\lambda_1 = -1$ e $\lambda_2 = -4$.

```
>> A=[-1,1,-1,-3,-1,7;0,-1,1,2,3,2;
0,0,-1,0,-2,1;0,0,0,-1,1,-2;
0,0,0,0,-1,3;0,0,0,0,0,-4]; A=sym(A);
>> Nm11=null(A+eye(6))
[  0,  1]
[  1,  0]
[ -2,  0]
[  1,  0]
[  0,  0]
[  0,  0]
>> Nm12=null((A+eye(6))^2)
[  0,  1,  0,  0]
[  0,  0,  1,  0]
[ -2,  0,  0, -2]
[  1,  0,  0,  0]
[  0,  0,  0,  1]
[  0,  0,  0,  0]
>> Nm13=null((A+eye(6))^3)
[ 0, 0, 0, 0, 1]
```

```
[ 0, 0, 1, 0, 0]
[ 0, 1, 0, 0, 0]
[ 0, 0, 0, 1, 0]
[ 1, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
```

A menos da ordem dos blocos as possíveis formas canônicas de Jordan são

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

(d) O polinômio característico de A é claramente $p(\lambda) = (\lambda - 1)^7(\lambda - 3)$. Assim, os autovalores da matriz são $\lambda_1 = 1$ e $\lambda_2 = 3$.

```
>> A=[1,1,0,0,-1,0,4,0;
0,1,1,-1,-1,-3,3,-4;
0,0,1,0,1,1,-2,1;
0,0,0,1,1,1,-4,-5;
0,0,0,0,1,0,-1,-5;
0,0,0,0,0,1,1,-1;
0,0,0,0,0,0,1,-2;
0,0,0,0,0,0,0,3];A=sym(A);
>> N11=null(A-eye(8))
[  1,  0,  0]
[  0,  0, -1]
[  0,  1,  2]
[  0,  1,  0]
[  0,  0, -1]
```

```
[  0,  0,  1]
[  0,  0,  0]
[  0,  0,  0]
>> N12=null((A-eye(8))^2)
[  0,  1,  0,  0,  0,  0]
[  0,  0,  0,  1,  0,  0]
[  1,  0,  1,  0,  3, -4]
[  0,  0,  1,  0,  0,  0]
[  1,  0,  0,  0,  0,  0]
[  0,  0,  0,  0,  1,  0]
[  0,  0,  0,  0,  0,  1]
[  0,  0,  0,  0,  0,  0]
>> N13=null((A-eye(8))^3)
[ 0, 0, 0, 1, 0, 0, 0]
[ 0, 1, 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 1, 0]
[ 0, 0, 0, 0, 1, 0, 0]
[ 1, 0, 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 0, 1]
[ 0, 0, 1, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 0, 0]
```

A menos da ordem dos blocos as possíveis formas canônicas de Jordan são

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

7.1.9.
```
>> B=randi(2), A=[B-B',zeros(2,1);zeros(1,2),randi]
B =  5    -1
     3     0
A =  0    -4     0
     4     0     0
     0     0    -3
>> syms x, p=det(A-x*eye(3)), solve(p)
p = -3*x^2-x^3-48-16*x
ans = [   -3][  4*i][ -4*i]
>> escalona(A+3*eye(3))
ans =[ 1, 0, 0]
     [ 0, 1, 0]
     [ 0, 0, 0]
```

A matriz A não é diagonalizável pois ela só tem um autovalor e auto espaço associado a este autovalor tem dimensão 2. Assim, não é possível encontrar 3 autovetores L.I.

7.1.10.
```
>> L=[eye(2),zeros(2,1);randi(1,2),0]; A=L*L'
A =  1     0     2
     0     1    -2
     2    -2     8
>> syms x, p=det(A-x*eye(3)), solve(p)
p = -9*x+10*x^2-x^3
```

```
ans = [ 0][ 1][ 9]
>> escalona(A)
ans =[  1,  0,  2]
     [  0,  1, -2]
     [  0,  0,  0]
```

O autoespaço associado ao autovalor $\lambda = 0$ é

$$\mathbb{V}_0 = \{(-2\alpha, 2\alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

Assim, $\{V_1 = (-2, 2, 1)\}$ é um conjunto com o maior número possível de autovetores L.I. associado a $\lambda = 0$.

```
>> escalona(A-eye(3))
ans =[  1, -1,  0]
     [  0,  0,  1]
     [  0,  0,  0]
```

O autoespaço associado ao autovalor $\lambda = 1$ é

$$\mathbb{V}_1 = \{(\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}.$$

Assim, $\{V_2 = (1, 1, 0)\}$ é um conjunto com o maior número possível de autovetores L.I. associado a $\lambda = 1$.

```
>> escalona(A-9*eye(3))
ans =[    1,    0, -1/4]
     [    0,    1,  1/4]
     [    0,    0,    0]
```

O autoespaço associado ao autovalor $\lambda = 9$ é

$$\mathbb{V}_9 = \{(\alpha, -\alpha, 4\alpha) \mid \alpha \in \mathbb{R}\}.$$

Assim, $\{V_3 = (1, -1, 4)\}$ é um conjunto com o maior número possível de autovetores L.I. associado a $\lambda = 9$.

```
>> V1=[-2,2,1];V2=[1,1,0];V3=[1,-1,4];
>> P=[V1',V2',V3'], D=diag([0,1,9])
P = -2     1     1
     2     1    -1
     1     0     4
D =  0     0     0
     0     1     0
     0     0     9
>> inv(P)*A*P
ans =    0    0    0
         0    1    0
         0    0    9
>> [P,D]=eig(sym(A))
P =[ -1, -2,  1]
   [  1,  2,  1]
   [ -4,  1,  0]
D =[ 9, 0, 0]
   [ 0, 0, 0]
   [ 0, 0, 1]
```

Os elementos da diagonal da matriz D têm que ser os autovalores de A. As matrizes D podem diferir na ordem com que os autovalores aparecem. As colunas de P são autovetores associados aos autovalores que aparecem nas colunas correspondentes de D. Assim, fazendo uma reordenação das colunas das matrizes P e D de forma que as matrizes D sejam iguais, as colunas de uma matriz P são múltiplos escalares das colunas correspondentes da outra matriz P.

7.2. Diagonalização de Matrizes Simétricas (página 436)

7.2.1.

(a)
```
>> A=[2,2;2,2];
>> B=A-x*eye(2)
```

```
[2-x,   2]
[  2, 2-x]
>> p=det(B)
p =-4*x+x^2
>> solve(p)
[0][4]
```

```
>> B0=subs(B,x,0)
[2, 2]
[2, 2]
>> escalona(B0)
[1, 1]
[0, 0]
```

```
>> B4=subs(B,x,4)
[-2,  2]
[ 2, -2]
>> escalona(B4)
[1, -1]
[0,  0]
```

$\mathbb{V}_0 = \{(-\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_1 = (-1, 1)\}$ é base para $\mathbb{V}_0$, pois gera $\mathbb{V}_0$ ($(-\alpha, \alpha) = \alpha(-1, 1)$) e um vetor não nulo é L.I. Seja $W_1 = (1/||V_1||)V_1 = (-1/\sqrt{2}, 1/\sqrt{2})$. $\{W_1 = (-1/\sqrt{2}, 1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_0$.

$\mathbb{V}_4 = \{(\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_2 = (1, 1)\}$ é base para $\mathbb{V}_4$, pois gera $\mathbb{V}_4$ ($(\alpha, \alpha) = \alpha(1, 1)$) e um vetor não nulo é L.I. Seja $W_2 = (1/||V_2||)V_2 = (1/\sqrt{2}, 1/\sqrt{2})$. $\{W_2 = (1/\sqrt{2}, 1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_4$.

$$P = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$$

(b)
```
>> A=[2,1;1,2];
>> B=A-x*eye(2)
[2-x,   1]
[  1, 2-x]
>> p=det(B)
p =3-4*x+x^2
>> solve(p)
```

```
[3][1]

>> B1=subs(B,x,1)
[1, 1]
[1, 1]
>> escalona(numeric(B1))
[1, 1]
[0, 0]
```

```
>> B3=subs(B,x,3)
[-1,  1]
[ 1, -1]
>> escalona(B3)
[1, -1]
[0,  0]
```

$\mathbb{V}_1 = \{(-\alpha,\alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_1 = (-1,1)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ ($(-\alpha,\alpha) = \alpha(-1,1)$) e um vetor não nulo é L.I. Seja $W_1 = (1/||V_1||)V_1 = (-1/\sqrt{2}, 1/\sqrt{2})$. $\{W_1 = (-1/\sqrt{2}, 1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_1$.

$\mathbb{V}_3 = \{(\alpha,\alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_2 = (1,1)\}$ é base para $\mathbb{V}_3$, pois gera $\mathbb{V}_3$ ($(\alpha,\alpha) = \alpha(1,1)$) e um vetor não nulo é L.I. Seja $W_2 = (1/||V_2||)V_2 = (1/\sqrt{2}, 1/\sqrt{2})$. $\{W_2 = (1/\sqrt{2}, 1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_3$.

$$P = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

(c)
```
>> A=[0,0,1;0,0,0;1,0,0];
>> B=A-x*eye(3)
[-x,  0,  1]
[ 0, -x,  0]
[ 1,  0, -x]
>> p=det(B)
p =-x^3+x
>> solve(p)
[ 0][-1][ 1]
```

```
>> B0=subs(B,x,0)
[0, 0, 1]
[0, 0, 0]
[1, 0, 0]
>> escalona(B0)
[1, 0, 0]
[0, 0, 1]
[0, 0, 0]
```

```
>> Bm1=subs(B,x,-1)
[1, 0, 1]
[0, 1, 0]
[1, 0, 1]
>> escalona(Bm1)
[1, 0, 1]
[0, 1, 0]
[0, 0, 0]
```

```
>> B1=subs(B,x,1)
[-1,  0,  1]
[ 0, -1,  0]
[ 1,  0, -1]
>> escalona(B1)
[1, 0, -1]
[0, 1,  0]
[0, 0,  0]
```

$\mathbb{V}_0 = \{(0, \alpha, 0) \mid \alpha \in \mathbb{R}\}$. $\{V_1 = (0,1,0)\}$ é base para $\mathbb{V}_0$, pois gera $\mathbb{V}_0$ $((0,\alpha,0) = \alpha(0,1,0))$ e um vetor não nulo é L.I. $\{V_1 = (0,1,0)\}$ é base ortonormal de $\mathbb{V}_0$, pois $||V_1|| = 1$.

$\mathbb{V}_{-1} = \{(-\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_2 = (-1,0,1)\}$ é base para $\mathbb{V}_{-1}$, pois gera $\mathbb{V}_{-1}$ $((-\alpha,0,\alpha) = \alpha(-1,0,1))$ e um vetor não nulo é L.I. Seja $W_2 = (1/||V_2||)V_2 = (-1/\sqrt{2}, 0, 1/\sqrt{2})$. $\{W_2 = (-1/\sqrt{2}, 0, 1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_{-1}$.

$\mathbb{V}_1 = \{(\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_3 = (1,0,1)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ $((\alpha,0,\alpha) = \alpha(1,0,1))$ e um vetor não nulo é L.I. Seja $W_3 = (1/||V_3||)V_3 = (1/\sqrt{2}, 0, 1/\sqrt{2})$. $\{W_3 = (1/\sqrt{2}, 0, 1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_1$.

Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais (Proposição 7.7 na página 430). Portanto, $\{W_1, W_2, W_3\}$ é uma base ortonormal de autovetores de A.

$$P = \begin{bmatrix} 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(d)
```
>> A=[0,0,0;0,2,2;0,2,2];
>> B=A-x*eye(3)
[-x,   0,   0]
[ 0, 2-x,   2]
[ 0,   2, 2-x]
>> p=det(B)
p =-x*(-4*x+x^2)
>> solve(p)
[0][0][4]
>> B0=subs(B,x,0)
[0, 0, 0]
[0, 2, 2]
[0, 2, 2]
```

```
>> escalona(B0)
[0, 1, 1]
[0, 0, 0]
[0, 0, 0]
>> B4=subs(B,x,4)
[-4,  0,  0]
[ 0, -2,  2]
[ 0,  2, -2]
>> escalona(B4)
[1, 0,  0]
[0, 1, -1]
[0, 0,  0]
```

$\mathbb{V}_0 = \{(\alpha, -\beta, \beta) \mid \alpha, \beta \in \mathbb{R}\}$. $\{V_1 = (1,0,0), V_2 = (0,-1,1)\}$ é base para $\mathbb{V}_0$, pois gera $\mathbb{V}_0$ ($(\alpha,-\beta,\beta) = \alpha(1,0,0) + \beta(0,-1,1)$) e é L.I. ($xV_1 + yV_2 = \bar{0}$ se, e somente se, $(x,-y,y) = (0,0,0)$ ou $x = 0$ e $y = 0$). Sejam $W_1 = V_1$, $W_2 = V_2 - \text{proj}_{W_1} V_2 = V_2 - \bar{0} = V_2$. Sejam $U_1 = (1/||W_1||)W_1 = W_1 = V_1 = (1,0,0)$ e $U_2 = (1/||W_2||)W_2 = (0, -1/\sqrt{2}, 1/\sqrt{2})$. $\{U_1 = (1,0,0), U_2 = ((0,-1/\sqrt{2}, 1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_0$.

$\mathbb{V}_4 = \{(0,\alpha,\alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_3 = (0,1,1)\}$ é base para $\mathbb{V}_4$, pois gera $\mathbb{V}_4$ ($(0,\alpha,\alpha) = \alpha(0,1,1)$) e um vetor não nulo é L.I. Seja $U_3 = (1/||V_3||)V_3 = (0, 1/\sqrt{2}, 1/\sqrt{2})$. $\{U_3 = (0,1/\sqrt{2},1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_4$. Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais (Proposição 7.7 na página 430). Portanto, $\{U_1, U_2, U_3\}$ é uma base ortonormal de autovetores de A.

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

(e)
```
>> A=[1,1,0;1,1,0;0,0,1];
>> B=A-x*eye(3)
[1-x,   1,   0]
[  1, 1-x,   0]
[  0,   0, 1-x]
>> p=det(B)
p =-2*x+3*x^2-x^3
>> solve(p)
[0][1][2]
>> B1=subs(B,x,1)
[0, 1, 0]
[1, 0, 0]
[0, 0, 0]
>> escalona(B1)
[1, 0, 0]
[0, 1, 0]
[0, 0, 0]
```

```
>> B0=subs(B,x,0)
[1, 1, 0]
[1, 1, 0]
[0, 0, 1]
>> escalona(B0)
[1, 1, 0]
[0, 0, 1]
[0, 0, 0]
>> B2=subs(B,x,2)
[-1,  1,  0]
[ 1, -1,  0]
[ 0,  0, -1]
>> escalona(B2)
[1, -1, 0]
[0,  0, 1]
[0,  0, 0]
```

$\mathbb{V}_0 = \{(-\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}$. $\{V_1 = (-1, 1, 0)\}$ é base para $\mathbb{V}_0$, pois gera $\mathbb{V}_0$ ($(-\alpha, \alpha, 0) = \alpha(-1, 1, 0)$) e um vetor não nulo é L.I. Seja $U_1 = (1/||V_1||)V_1 = (-1/\sqrt{2}, 1/\sqrt{2}, 0)$. $\{U_1 = (-1/\sqrt{2}, 1/\sqrt{2}, 0)\}$ é base ortonormal de $\mathbb{V}_0$.

$\mathbb{V}_1 = \{(0, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_2 = (0, 0, 1)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ ($(0, 0, \alpha) = \alpha(0, 0, 1)$) e um vetor não nulo é L.I. Seja $W_2 = (1/||V_2||)V_2 = (0, 0, 1)$. $\{W_2 = (0, 0, 1)\}$ é base ortonormal de $\mathbb{V}_1$.

$\mathbb{V}_2 = \{(\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}$. $\{V_3 = (1, 1, 0)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ ($(\alpha, \alpha, 0) = \alpha(1, 1, 0)$) e um vetor não nulo é L.I. Seja $W_3 = (1/||V_3||)V_3 = (1/\sqrt{2}, 1/\sqrt{2}, 0)$. $\{W_3 = (1/\sqrt{2}, 1/\sqrt{2}, 0)\}$ é base ortonormal de $\mathbb{V}_1$.

Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais (Proposição 7.7 na página 430). Portanto, $\{W_1, W_2, W_3\}$ é uma base ortonormal de autovetores de

A.

$$P = \begin{bmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

(f)
```
>> A=[2,1,1;1,2,1;1,1,2];
>> B=A-x*eye(3)
[2-x,   1,   1]
[  1, 2-x,   1]
[  1,   1, 2-x]
>> p=det(B)
p =4-9*x+6*x^2-x^3
>> solve(p)
[4][1][1]
>> B1=subs(B,x,1)
[1, 1, 1]
[1, 1, 1]
[1, 1, 1]
```

```
>> escalona(B1)
[1, 1, 1]
[0, 0, 0]
[0, 0, 0]
>> B4=subst(B,x,4)
[-2,  1,  1]
[ 1, -2,  1]
[ 1,  1, -2]
>> escalona(B4)
[1, 0, -1]
[0, 1, -1]
[0, 0,  0]
```

$\mathbb{V}_1 = \{(-\alpha-\beta,\alpha,\beta) \mid \alpha,\beta \in \mathbb{R}\}$. $\{V_1 = (-1,1,0), V_2 = (-1,0,1)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_0$ ($(-\alpha-\beta,\alpha,\beta) = \alpha(-1,1,0)+\beta(-1,0,1)$) e é L.I.(um vetor não é múltiplo escalar do outro). Sejam $W_1 = V_1$, $W_2 = V_2 - \text{proj}_{W_1} V_2 = V_2 - (-1/2,1/2,0) = (-1/2,-1/2,1)$. Sejam $U_1 = (1/||W_1||)W_1 = (-1/\sqrt{2},1/\sqrt{2},0)$ e $U_2 = (1/||W_2||)W_2 = (-\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\frac{\sqrt{6}}{3})$. $\{U_1,U_2\}$ é base ortonormal de $\mathbb{V}_1$.

$\mathbb{V}_4 = \{(\alpha,\alpha,\alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_3 = (1,1,1)\}$ é base para $\mathbb{V}_4$, pois gera $\mathbb{V}_4$ ($(\alpha,\alpha,\alpha) = \alpha(1,1,1)$) e um vetor não nulo é L.I. Seja $U_3 = (1/||V_3||)V_3 = (1/\sqrt{3},1/\sqrt{3},1/\sqrt{3})$. $\{U_3 = (1/\sqrt{3},1/\sqrt{3},1/\sqrt{3})\}$ é base ortonormal de $\mathbb{V}_4$. Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais (Proposição 7.7 na página 430). Portanto, $\{U_1,U_2,U_3\}$ é uma base ortonormal de autovetores de A.

$$P = \begin{bmatrix} -\sqrt{2}/2 & -\sqrt{6}/6 & \sqrt{3}/3 \\ \sqrt{2}/2 & -\sqrt{6}/6 & \sqrt{3}/3 \\ 0 & \sqrt{6}/3 & \sqrt{3}/3 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

(g)
```
>> A=[1,2,0,0;2,1,0,0;0,0,1,2;0,0,2,1];
>> B=A-x*eye(4)
[1-x,   2,   0,   0]
[  2, 1-x,   0,   0]
[  0,   0, 1-x,   2]
[  0,   0,   2, 1-x]
>> p=det(B)
p =9+12*x-2*x^2-4*x^3+x^4
>> solve(p)
[-1][-1][ 3][ 3]
```

```
>> Bm1=subs(B,x,-1)
[2, 2, 0, 0]
[2, 2, 0, 0]
[0, 0, 2, 2]
[0, 0, 2, 2]
>> escalona(Bm1)
[1, 1, 0, 0]
[0, 0, 1, 1]
[0, 0, 0, 0]
[0, 0, 0, 0]
```

```
>> B3=subs(B,x,3)
[-2,  2,  0,  0]
[ 2, -2,  0,  0]
[ 0,  0, -2,  2]
[ 0,  0,  2, -2]
>> escalona(B3)
[1, -1, 0,  0]
[0,  0, 1, -1]
[0,  0, 0,  0]
[0,  0, 0,  0]
```

$\mathbb{V}_{-1} = \{(-\alpha, \alpha, -\beta, \beta) \mid \alpha, \beta \in \mathbb{R}\}$. $\{V_1 = (-1, 1, 0, 0), V_2 = (0, 0, -1, 1)\}$ é base para $\mathbb{V}_{-1}$, pois gera $\mathbb{V}_{-1}$ ($(-\alpha, \alpha, -\beta, \beta) = \alpha(-1, 1, 0, 0) + \beta(0, 0, -1, 1)$) e é L.I.(um vetor não é múltiplo escalar do outro). Sejam $W_1 = V_1$, $W_2 = V_2 - \mathrm{proj}_{W_1} V_2 = V_2 - \bar{0} = V_2$. Sejam $U_1 = (1/||W_1||)W_1 = (-1/\sqrt{2}, 1/\sqrt{2}, 0, 0)$ e $U_2 = (1/||W_2||)W_2 = (0, 0, -1/\sqrt{2}, 1/\sqrt{2})$. $\{U_1, U_2\}$ é base ortonormal de

$\mathbb{V}_{-1}$.

$\mathbb{V}_3 = \{(\alpha, \alpha, \beta, \beta) \mid \alpha, \beta \in \mathbb{R}\}$. $\{V_3 = (1,1,0,0), V_4 = (0,0,1,1)\}$ é base para $\mathbb{V}_3$, pois gera $\mathbb{V}_{-1}$ $((\alpha, \alpha, \beta, \beta) = \alpha(1,1,0,0) + \beta(0,0,1,1))$ e é L.I.(um vetor não é múltiplo escalar do outro). Sejam $W_3 = V_3$, $W_4 = V_4 - \text{proj}_{W_3} V_4 = V_4 - \bar{0} = V_4$. Sejam $U_3 = (1/||W_3||)W_3 = (1/\sqrt{2}, 1/\sqrt{2}, 0, 0)$ e $U_4 = (1/||W_4||)W_4 = (0, 0, 1/\sqrt{2}, 1/\sqrt{2})$. $\{U_1, U_2\}$ é base ortonormal de $\mathbb{V}_3$. Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais (Proposição 7.7 na página 430). Portanto, $\{U_1, U_2, U_3, U_4\}$ é uma base ortonormal de autovetores de A.

$$P = \begin{bmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 0 & -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

(h)
```
>> A=[0,0,0,0;0,0,0,0;0,0,0,1;0,0,1,0];
>> B=A-x*eye(4)
[-x,  0,  0,  0]
[ 0, -x,  0,  0]
[ 0,  0, -x,  1]
[ 0,  0,  1, -x]
>> p=det(B)
p =x^2*(x^2-1)
>> solve(p)
[ 0][ 0][ 1][-1]
```

```
>> B0=subs(B,x,0)
[0, 0, 0, 0]
[0, 0, 0, 0]
[0, 0, 0, 1]
[0, 0, 1, 0]
>> escalona(B0)
[0, 0, 1, 0]
[0, 0, 0, 1]
[0, 0, 0, 0]
[0, 0, 0, 0]
```

```
>> Bm1=subs(B,x,-1)
[1, 0, 0, 0]
[0, 1, 0, 0]
[0, 0, 1, 1]
[0, 0, 1, 1]
>> escalona(Bm1)
[1, 0, 0, 0]
[0, 1, 0, 0]
[0, 0, 1, 1]
[0, 0, 0, 0]
```

```
>> B1=subs(B,x,1)
B1 =
[-1,  0,  0,  0]
[ 0, -1,  0,  0]
[ 0,  0, -1,  1]
[ 0,  0,  1, -1]
>> escalona(B1)
[1, 0, 0,  0]
[0, 1, 0,  0]
[0, 0, 1, -1]
[0, 0, 0,  0]
```

$\mathbb{V}_0 = \{(\alpha, \beta, 0, 0) \mid \alpha, \beta \in \mathbb{R}\}$. $\{V_1 = (1, 0, 0, 0), V_2 = (0, 1, 0, 0)\}$ é base para $\mathbb{V}_0$, pois gera $\mathbb{V}_{-1}$ $((\alpha, \beta, 0, 0) = \alpha(1, 0, 0, 0) + \beta(0, 1, 0, 0))$ e é L.I.(um vetor não é múltiplo escalar do outro). Claramente $V_1 \cdot V_2 = 0$ e possuem norma igual a 1. Sejam $U_1 = V_1$ e $U_2 = V_2$. $\{U_1, U_2\}$ é base ortonormal de $\mathbb{V}_0$.

$\mathbb{V}_1 = \{(0, 0, -\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_3 = (0, 0, -1, 1)\}$ é base para $\mathbb{V}_1$, pois gera $\mathbb{V}_1$ $((0, 0, -\alpha, \alpha) = \alpha(0, 0, -1, 1))$ e um vetor não nulo é L.I. Seja $U_3 = (1/||V_3||)V_3 = (0, 0, -1/\sqrt{2}, 1/\sqrt{2})$. $\{U_3 = (0, 0, -1/\sqrt{2}, 1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_1$.

$\mathbb{V}_{-1} = \{(0, 0, \alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_4 = (0, 0, 1, 1)\}$ é base para $\mathbb{V}_{-1}$, pois gera $\mathbb{V}_{-1}$ $((0, 0, \alpha, \alpha) = \alpha(0, 0, 1, 1))$ e um vetor não nulo é L.I. Seja $U_4 = (1/||V_4||)V_4 = (0, 0, 1/\sqrt{2}, 1/\sqrt{2})$. $\{U_4 = (0, 0, 1/\sqrt{2}, 1/\sqrt{2})\}$ é base ortonormal de $\mathbb{V}_{-1}$. Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais (Proposição 7.7 na página 430). Portanto, $\{U_1, U_2, U_3, U_4\}$ é uma base ortonormal de autovetores de A.

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

7.2.2. Como a matriz é simétrica ela é diagonalizável. Além disso, também por ser simétrica, autovetores associados a autovalores diferentes são ortogonais. Assim basta ortogonalizar dentro de cada autoespaço.

```
>> W1=V1;
>> W2=V2-proj(W1,V2)
W2 = [  2, -1,  0,  2]
>> W3=V3;
>> W4=V4-proj(W3,V4)
W4 = [ -1, -2, -2,  0]
>> U1=W1/no(W1)
U1 = [    0,  2/3, -2/3,  1/3]
>> U2=W2/no(W2)
U2 = [  2/3, -1/3,    0,  2/3]
>> U3=W3/no(W3)
U3 = [ -2/3,    0,  1/3,  2/3]
```

$$P = [\, U_1\ U_2\ U_3\ U_4 \,] = \begin{bmatrix} 0 & 2/3 & -2/3 & 1/3 \\ 2/3 & -1/3 & 0 & 2/3 \\ -2/3 & 0 & 1/3 & 2/3 \\ 1/3 & 2/3 & 2/3 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

7.3. Aplicação ao Estudo de Cônicas (página 455)

7.3.1.
```
>> A=[9,-2;-2,6];
>> syms x y; X=[x;y];
>> expr=simplify(X.'*A*X-30)
```

$9\,x^2 - 4\,xy + 6\,y^2 - 30$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} \sqrt{5}/5 & -2\sqrt{5}/5 \\ 2\sqrt{5}/5 & \sqrt{5}/5 \end{bmatrix}$$

```
D=[5, 0]
```

```
  [0,10]
>> syms x1 y1; X1=[x1;y1];
>> expr=subst(expr,X,P*X1)
```

$5\,{x_1}^2+10\,{y_1}^2-30$

```
>> expr=expr/30
```

${x_1}^2/6+{y_1}^2/3-1$

```
>> elipse(sqrt(6),sqrt(3),P)
```

7.3.2.
```
>> A=[3,-4;-4,-12];
>> expr=simplify(X.'*A*X+81)
```

$$3\,x^2 - 8\,xy - 12\,y^2 + 81$$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} \sqrt{17}/17 & -4\sqrt{17}/17 \\ 4\sqrt{17}/17 & \sqrt{17}/17 \end{bmatrix}$$

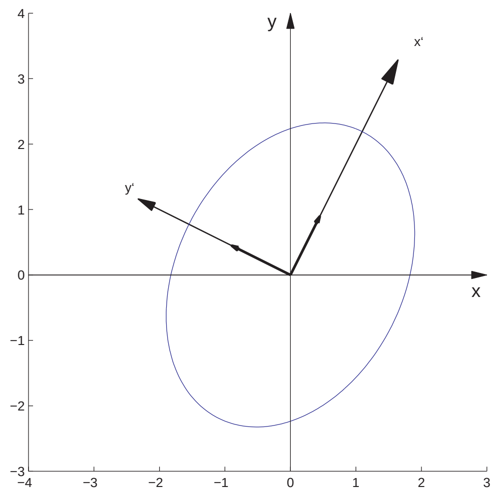
```
D=[-13,0]
  [  0,4]
>> expr=subst(expr,X,P*X1)
```

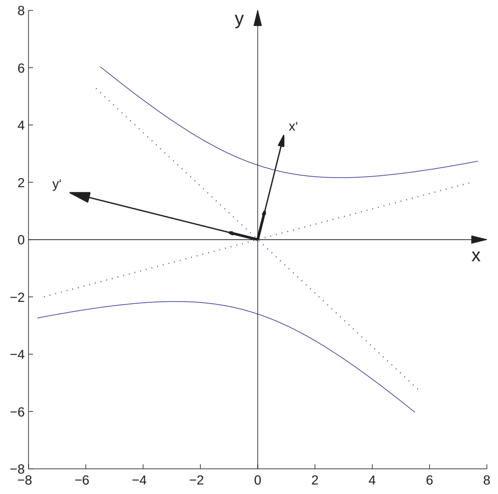
$$-13\,{x_1}^2 + 4\,{y_1}^2 + 81$$

```
>> expr=expr/81
```

$$-\frac{13}{81}\,{x_1}^2 + \frac{4}{81}\,{y_1}^2 + 1$$

```
>> hiperbx(9/sqrt(13),9/2,P)
```


y
x'
y'
x
4
3
2
1
0
−1
−2
−3
−4
−3
−2
−1
0
1
2
3


y
x'
y'
x
8
6
4
2
0
−2
−4
−6
−8
−8
−6
−4
−2
0
2
4
6
8

7.3.3.
```
>> A=[2,-2;-2,-1];
>> expr=simplify(X.'*A*X+K*X+24)
```

$2\,x^2 - 4\,xy - y^2 + 24$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} \sqrt{5}/5 & -2\sqrt{5}/5 \\ 2\sqrt{5}/5 & 1\sqrt{5}/5 \end{bmatrix}$$

```
D =[-2, 0]
   [ 0, 3]
>> expr=subst(expr,X,P*X1)
```

$-2\,{x_1}^2 + 3\,{y_1}^2 + 24$

```
>> expr=expr/24
```

$-{x_1}^2/12 + {y_1}^2/8 + 1$

```
>> hiperbx(sqrt(12),sqrt(8),P)
```

7.3.4.
```
>> A=[21,3;3,13];
>> expr=simplify(X.'*A*X-132)
```

$21\,x^2 + 6\,xy + 13\,y^2 - 132$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} \sqrt{10}/10 & 3\sqrt{10}/10 \\ -3\sqrt{10}/10 & \sqrt{10}/10 \end{bmatrix}$$

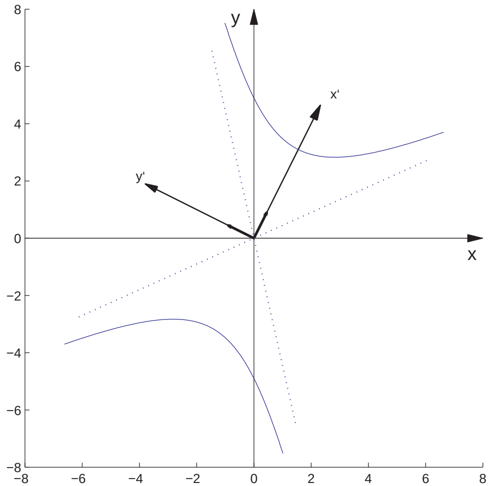
```
D=[12, 0]
  [ 0,22]
>> expr=subst(expr,X,P*X1)
```

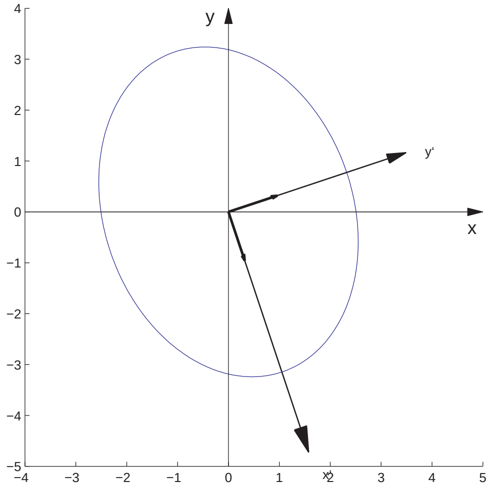
$12\,{x_1}^2 + 22\,{y_1}^2 - 132$

```
>> expr=expr/132
```

${x_1}^2/11 + {y_1}^2/6 - 1$

```
>> elipse(sqrt(11),sqrt(6),P)
```


y
x'
y'
x
8
6
4
2
0
−2
−4
−6
−8
−8 −6 −4 −2 0 2 4 6 8


y
y'
x
x'
4
3
2
1
0
−1
−2
−3
−4
−5
−4 −3 −2 −1 0 1 2 3 4 5

7.3.5.
```
>> A=[4,-10;-10,25];
>> K=[-15,-6];
>> expr=simplify(X.'*A*X+K*X)
```

$$4\,x^2 - 20\,xy + 25\,y^2 - 15\,x - 6\,y$$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} \frac{5}{29}\sqrt{29} & -\frac{2}{29}\sqrt{29} \\ \frac{2}{29}\sqrt{29} & \frac{5}{29}\sqrt{29} \end{bmatrix}$$

```
D =[0,  0]
   [0, 29]
>> expr=subst(expr,X,P*X1)
```

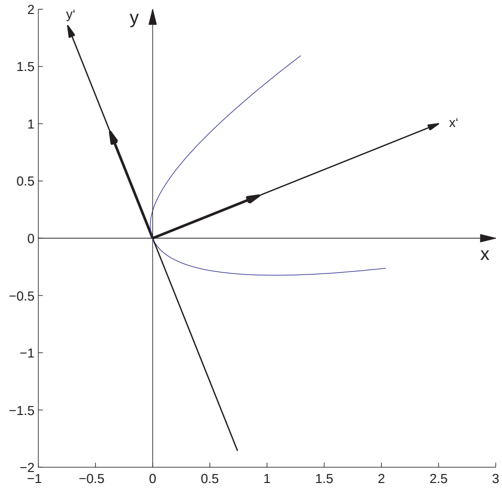
$$29\,{y_1}^2 - 3\,\sqrt{29}x_1$$

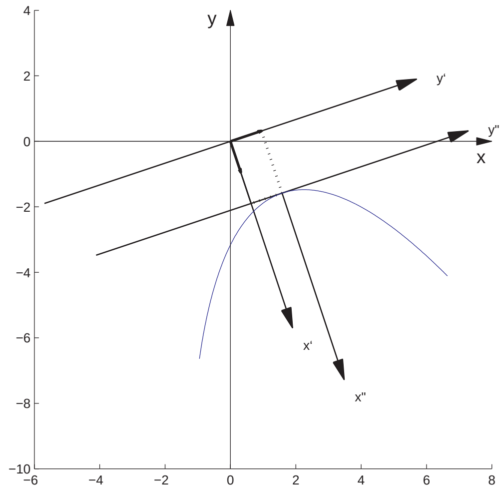
```
>> expr=expr/29
```

$${y_1}^2 - \frac{3}{29}\,\sqrt{29}x_1$$

```
>> parabx(3/(4*sqrt(29)),P)
```

7.3.6.
```
>> A=[9,3;3,1]; K=[-10*10^(1/2),10*10^(1/2)];
>> expr=simplify(X.'*A*X+K*X+90)
```
$9\,x^2+6\,xy+y^2-10\,\sqrt{10}x+10\,\sqrt{10}y+90$
```
>> [P,D]=diagonal(A)
```
$$P=\begin{bmatrix} \sqrt{10}/10 & 3\sqrt{10}/10 \\ -3\sqrt{10}/10 & \sqrt{10}/10 \end{bmatrix}$$
```
D =[0,  0]
   [0, 10]
>> expr=subst(expr,X,P*X1)
```
$10\,{y_1}^2-20\,y_1-40\,x_1+90$
```
EDU>> expr=subst(expr,y1,y2+1)
```
$10\,{y_2}^2+80-40\,x_1$
```
>> expr=subst(expr,x1,x2+2)
```
$10\,{y_2}^2-40\,x_2$
```
>> expr=expr/10
```
${y_2}^2-4\,x_2$
```
>> paraby(1,P,[2;1])
```


y'
y
x'
x
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
−1
−0.5
0
0.5
1
1.5
2
2.5
3


y
y'
y''
x
x'
x''
4
2
0
−2
−4
−6
−8
−10
−6
−4
−2
0
2
4
6
8

7.3.7.
```
>> A=[5,-3;-3,5];
>> K=[-30*(2)^(1/2),18*(2)^(1/2)];
>> expr=simplify(X.'*A*X+K*X+82)
```

$$5\,x^2 - 6\,xy + 5\,y^2 - 30\,\sqrt{2}x + 18\,\sqrt{2}y + 82$$

```
>> [P,D]=diagonal(A)
```

$$P = \left[\begin{array}{cc} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{array}\right]$$

```
D =[2, 0]
   [0, 8]
>> expr=subst(expr,X,P*X1)
```

$$2\,{x_1}^2 + 8\,{y_1}^2 - 12\,x_1 + 48\,y_1 + 82$$

```
>> X0=[3;-3];
>> expr=subst(expr,X1,X2+X0)
```

$$2\,{x_2}^2 - 8 + 8\,{y_2}^2$$

```
>> expr=expr/8
```

$${x_2}^2/4 - 1 + {y_2}^2$$

```
>> elipse(2,1,P,X0)
```

7.3.8.
```
>> A=[5,6;6,0];
>> K=[-12*(13)^(1/2),0];
>> expr=simplify(X.'*A*X+K*X-36)
```

$$5\,x^2 + 12\,xy - 12\,\sqrt{13}x - 36$$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} 2/\sqrt{13} & 3/\sqrt{13} \\ -3/\sqrt{13} & 2/\sqrt{13} \end{bmatrix}$$

```
D =[-4, 0]
   [ 0, 9]
>> expr=subst(expr,X,P*X1)
```

$$-4\,{x_1}^2 + 9\,{y_1}^2 - 24\,x_1 - 36\,y_1 - 36$$

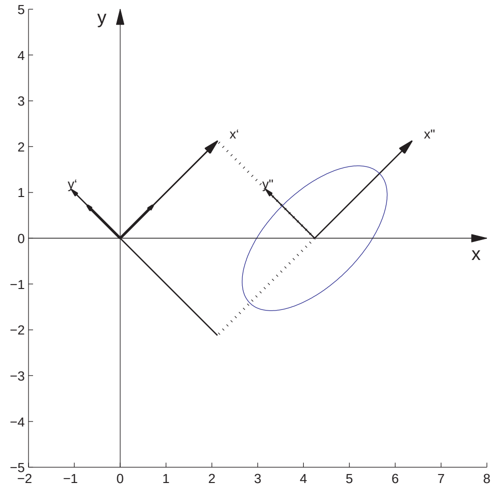
```
>> X0=[-3;2];
>> expr=subst(expr,X1,X2+X0)
```

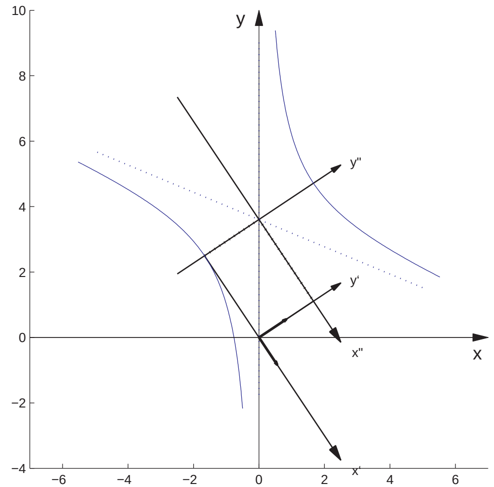
$$-4\,{x_2}^2 - 36 + 9\,{y_2}^2$$

```
>> expr=expr/36
```

$$-{x_2}^2/9 - 1 + {y_2}^2/4$$

```
>> hiperby(2,3,P,X0)
```


y
x
x'
y'
x"
y"
5
4
3
2
1
0
−1
−2
−3
−4
−5
−2
−1
0
1
2
3
4
5
6
7
8


y
x
y"
y'
x"
x'
10
8
6
4
2
0
−2
−4
−6
−4
−2
0
2
4
6

7.3.9.
```
>> A=[6,-2;-2,9];
>> K=[-4*5^(1/2),-18*5^(1/2)];
>> expr=simplify(X.'*A*X+K*X-5)
```

$$6\,x^2 - 4\,xy + 9\,y^2 - 4\,\sqrt{5}x - 18\,\sqrt{5}y - 5$$

```
>> [P,D]=diagonal(A)
```

$$P = \left[\begin{array}{cc} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{array}\right]$$

```
D =[5,  0]
   [0, 10]
>> expr=subst(expr,X,P*X1)
```

$$5\,{x_1}^2 + 10\,{y_1}^2 - 26\,x_1 - 32\,y_1 - 5$$

```
>> X0=[26/10;32/20];
>> expr=subst(expr,X1,X2+X0)
```

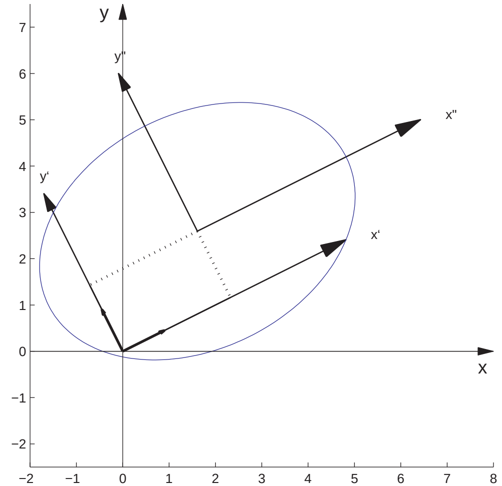
$$5\,{x_2}^2 - \frac{322}{5} + 10\,{y_2}^2$$

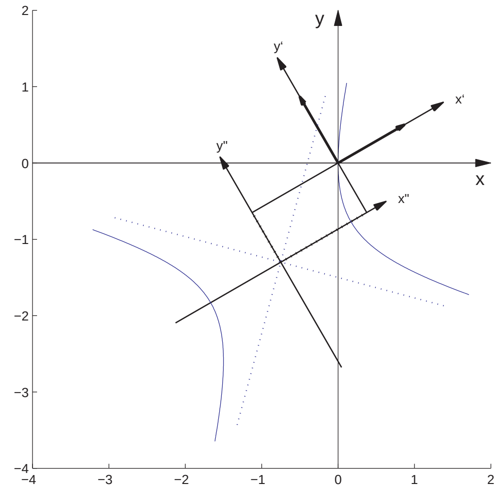
```
>> expr=expr*5/322
```

$$\frac{25}{322}\,{x_2}^2 - 1 + \frac{25}{161}\,{y_2}^2$$

```
>> elipse(sqrt(322)/5,sqrt(161)/5,P,X0)
```

7.3.10.
```
>> A=[1,3^(1/2);3^(1/2),-1];
>> K=[6,0];
>> expr=simplify(X.'*A*X+K*X)
```
$x^2 + 2\,xy\sqrt{3} - y^2 + 6\,x$
```
>> [P,D]=diagonal(A)
```
$$P = \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix}$$
```
D =[ 2, 0]
   [ 0,-2]
>> expr=subst(expr,X,P*X1)
```
$2\,{x_1}^2 - 2\,{y_1}^2 + 3\,\sqrt{3}x_1 - 3\,y_1$
```
>> X0=[-3*3^(1/2)/4;-3/4];
>> expr=subst(expr,X1,X2+X0)
```
$2\,{x_2}^2 - 9/4 - 2\,{y_2}^2$
```
>> expr=expr*4/9
```
$\frac{8}{9}\,{x_2}^2 - 1 - \frac{8}{9}\,{y_2}^2$
```
>> hiperbx(3/sqrt(8),3/sqrt(8),P,X0)
```


y
y"
x"
y'
x'
x
7
6
5
4
3
2
1
0
−1
−2
−2
−1
0
1
2
3
4
5
6
7
8


y
y'
x'
y"
x"
x
2
1
0
−1
−2
−3
−4
−4
−3
−2
−1
0
1
2

7.3.11.
```
>> A=[8,-8;-8,8];
>> K=[33*2^(1/2),-31*2^(1/2)];
>> expr=simplify(X.'*A*X+K*X+70)
```
$$8\,x^2 - 16\,xy + 8\,y^2 + 33\,\sqrt{2}x - 31\,\sqrt{2}y + 70$$
```
>> [P,D]=diagonal(A)
```
$$P = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$
```
D =[0,  0]
   [0, 16]
>> expr=subst(expr,X,P*X1)
```
$$16\,{y_1}^2 + 2\,x_1 - 64\,y_1 + 70$$
```
>> expr=subst(expr,y1,y2+2)
```
$$16\,{y_2}^2 + 6 + 2\,x_1$$
```
>> expr=subst(expr,x1,x2-3)
```
$$16\,{y_2}^2 + 2\,x_2$$
```
>> expr=expr/16
```
$${y_2}^2 + x_2/8$$
```
>> parabx(-1/32,P,[-3;2])
```

7.3.12.
```
>> A=[1,-3;-3,-7];
>> K=[10,2];
>> expr=simplify(X.'*A*X+K*X+9)
```

$x^2 - 6\,xy - 7\,y^2 + 10\,x + 2\,y + 9$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} 1/\sqrt{10} & -3/\sqrt{10} \\ 3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix}$$

```
D =[-8,  0]
   [ 0,  2]
>> expr=subst(expr,X,P*X1)
```

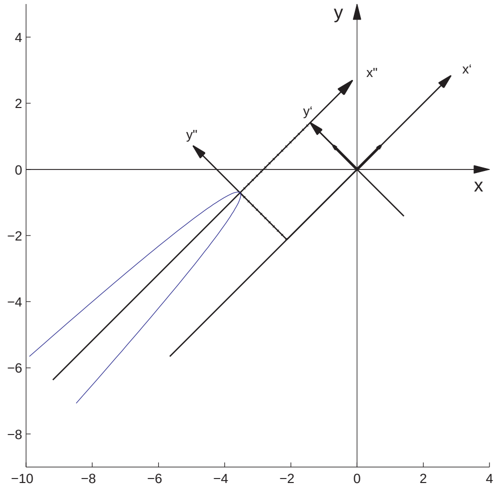
$-8\,{x_1}^2 + 2\,{y_1}^2 + \frac{8}{5}\,\sqrt{10}x_1 - \frac{14}{5}\,\sqrt{10}y_1 + 9$

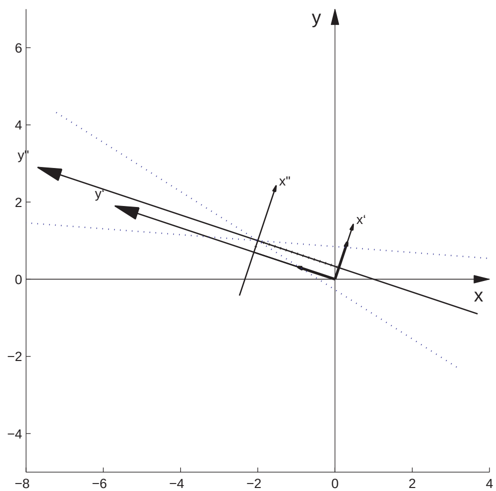
```
>> X0=[1/10^(1/2);7/10^(1/2)];
>> expr=subst(expr,X1,X2+X0)
```

$-8\,{x_2}^2 + 2\,{y_2}^2$

```
>> hiperby(4,1,P,X0,'d')
```

Esta é uma cônica degenerada. A equação representa as duas retas $y''^2 = 4x''^2$, ou $y'' = \pm 2x''$.


y
x''
x'
y'
y''
x
4
2
0
−2
−4
−6
−8
−10
−8
−6
−4
−2
0
2
4


y
y''
y'
x''
x'
x
6
4
2
0
−2
−4
−8
−6
−4
−2
0
2
4

Bibliografia

[1] Howard Anton e Chris Rorres. *Álgebra Linear com Aplicações*. Bookman, São Paulo, 8a. edição, 2000.

[2] José L. Boldrini, Sueli I. R. Costa, Vera L. Figueiredo, e Henry G. Wetzler. *Álgebra Linear*. Ed. Harbra Ltda., São Paulo, 3a. edição, 1986.

[3] Frederico F. C., filho. *Introdução ao MATLAB*. Departamento de Ciência da Computação - UFMG, Belo Horizonte, Fevereiro de 2000.

[4] Carlos A. Callioli, Hygino H. Domingues, e Roberto C. F. Costa. *Álgebra Linear e Aplicações*. Atual Editora, São Paulo, 6a. edição, 1995.

[5] Adilson Gonçalves e Rita M. L. de Souza. *Introdução à Álgebra Linear*. Edgard Blücher, Rio de Janeiro, 1977.

[6] Alésio de Caroli, Carlos A. Callioli, e Miguel O. Feitosa. *Matrizes, Vetores, Geometria Analítica*. Nobel, São Paulo, 1976.

[7] João Pitombeira de Carvalho. *Álgebra Linear - Introdução*. Livros Técnicos e Científicos Editora S.A., Rio de Janeiro, 2a. edição, 1977.

[8] Nathan M. dos Santos. *Vetores e Matrizes*. Livros Técnicos e Científicos Ed. S.A., Rio de Janeiro, 3a. edição, 1988.

[9] C. H. Edwards, Jr e D. E. Penney. *Introdução à Álgebra Linear*. Livros Técnicos e Científicos Editora S.A., Rio de Janeiro, 2000.

[10] John B. Fraleigh e Raymond A. Beauregard. *Linear Algebra*. Addison Wesley, Reading, Massachusetts, 3a. edição, 1995.

[11] Stephen H. Friedberg, Arnold J. Insel, e Lawrence E. Spence. *Linear Algebra*. Prentice Hall, Upper Saddle River, New Jersey, 3a. edição, 1997.

[12] G. H. Golub e C. F. Van Loan. *Matrix Computations*. Johns Hopkins U.P., Baltimore, 3a. edição, 1996.

[13] Stanley I. Grossman. *Elementary Linear Algebra*. Saunders College Publishing, New York, 5a. edição, 1994.

[14] David R. Hill e David E. Zitarelli. *Linear Algebra Labs with MATLAB*. Macmillan Publishing Company, New York, 1994.

[15] Kenneth Hoffman e Ray Kunze. *Álgebra Linear*. Livros Técnicos e Científicos Ed. S.A., Rio de Janeiro, 3a. edição, 1979.

[16] Bernard Kolman. *Introdução à Álgebra Linear com Aplicações*. Prentice Hall do Brasil, Rio de Janeiro, 6a. edição, 1998.

[17] Serge Lang. *Introduction to Linear Algebra*. Springer, New York, 2a. edição, 1986.

[18] David C. Lay. *Álgebra Linear e suas Aplicações*. Livros Técnicos e Científicos Editora S.A., Rio de Janeiro, 2a. edição, 1999.

[19] Steven Leon, Eugene Herman, e Richard Faulkenberry. *ATLAST Computer Exercises for Linear Algebra*. Prentice Hall, Upper Saddle River, New Jersey, 1996.

[20] Steven J. Leon. *Álgebra Linear com Aplicações*. Livros Técnicos e Científicos Editora S.A., Rio de Janeiro, 5a. edição, 1998.

[21] Emília Giraldes, Vitor H. Fernandes, e Maria P. M Smith. *Curso de Álgebra Linear e Geometria Analítica*. Mc Graw Hill, Lisboa, 1995.

[22] Elon L. Lima. *Álgebra Linear*. IMPA, Rio de Janeiro, 2a. edição, 1996.

[23] Seymour Lipschutz. *Álgebra Linear*. McGraw-Hill, São Paulo, 3a. edição, 1994.

[24] Mathworks Inc. *Student Edition of MATLAB Version 5 for Windows*. Prentice Hall, Upper Saddle River, New Jersey, 1997.

[25] Ben Noble e James W. Daniel. *Applied Linear Algebra*. Prentice Hall, Upper Saddle River, New Jersey, 3a. edição, 1988.

[26] Reginaldo J. Santos. *Álgebra Linear e Aplicações*. Imprensa Universitária da UFMG, Belo Horizonte, 2006.

[27] Shayle R. Searle. *Matrix Algebra Useful for Statistics*. John Wiley and Sons, New York, 1982.

[28] Carl P. Simon e Lawrence Blume. *Mathematics for Economists*. W. W. Norton and Co. Inc., New York, 1994.

[29] Gilbert Strang. *Linear Algebra and its Applications*. Harcourt Brace Jovanovich, Inc., Orlando, 3a. edição, 1988.

[30] Gilbert Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, Wellesley, 1993.

Índice Alfabético