
ÁLGEBRA LINEAR E APLICAÇÕES

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

Julho 2010

Álgebra Linear e Aplicações
Copyright © 2010 by Reginaldo de Jesus Santos (100808)

É proibida a reprodução desta publicação, ou parte dela, por qualquer meio, sem a prévia autorização, por escrito, do autor.

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

ISBN 85-7470-017-7

Ficha Catalográfica

S237a Santos, Reginaldo J.
Álgebra Linear e Aplicações / Reginaldo J. Santos - Belo Horizonte: Imprensa Universitária da UFMG, 2010.

1. Álgebra Linear I. Título

CDD: 512.5

Conteúdo

Prefácio	vi
1 Espaços Vetoriais	1
1.1 Definição e Exemplos	1
1.1.1 Os Espaços \mathbb{R}^n	1
1.1.2 Espaços Vetoriais Abstratos	5
1.2 Subespaços	14
1.2.1 Soma e Interseção de Subespaços	24
1.2.2 Conjunto de Geradores	31
1.3 Dependência Linear	44
1.3.1 Independência Linear de Funções	56
1.4 Base e Dimensão	67
1.4.1 Base	67
1.4.2 Dimensão	73
1.4.3 Aplicação: Frações Parciais	81
1.4.4 Aplicação: Interpolação por Splines	84

2	Espaços com Produto Interno	99
2.1	Produto Escalar e Norma	99
2.1.1	Produto Interno	99
2.1.2	Norma	106
2.1.3	Ortogonalidade	110
2.1.4	Projeção Ortogonal	116
2.1.5	Coeficientes de Fourier	124
2.2	Bases Ortonormais e Subespaços Ortogonais	132
2.2.1	Bases Ortonormais	132
2.2.2	Aplicação: Polinômios de Legendre	141
2.2.3	Complemento Ortogonal	146
2.2.4	Distância de um Ponto a um Subespaço	156
2.2.5	Aplicação: Séries de Fourier	164
3	Transformações Lineares	203
3.1	Definição, Exemplos e Propriedades	203
3.1.1	Definição e Exemplos	203
3.1.2	Propriedades	211
3.1.3	Aplicação: Matriz Jacobiana	216
3.2	A Imagem e o Núcleo	221
3.2.1	Espaço Linha e Espaço Coluna de uma Matriz	226
3.2.2	Injetividade e Sobrejetividade	229
3.3	Composição de Transformações Lineares	241
3.3.1	Matriz de uma Transformação Linear	243
3.3.2	Invertibilidade	252
3.3.3	Semelhança	256
3.3.4	Aplicação: Equações Diferenciais Lineares	260
3.4	A Adjunta	270
3.4.1	Aplicação: Problema de Quadrados Mínimos	277
4	Diagonalização	291

4.1	Diagonalização de Operadores	291
4.1.1	Motivação: Sistemas de Equações Diferenciais Lineares	291
4.1.2	Operadores e Matrizes Diagonalizáveis	294
4.1.3	Autovalores e Autovetores	297
4.1.4	Subespaços Invariantes e o Teorema de Cayley-Hamilton	311
4.1.5	Aplicação: Cálculo das Potências de uma Matriz	316
4.2	Operadores Auto-adjuntos e Normais	329
4.3	Aplicação na Identificação de Cônicas	345
4.4	Forma Canônica de Jordan	358
4.4.1	Autoespaço Generalizado	360
4.4.2	Ciclos de Autovetores Generalizados	372
4.4.3	Aplicação: Funções de Matrizes e Sistemas de Equações Diferenciais Lineares	384
	Respostas dos Exercícios	398
	Bibliografia	500
	Índice Alfabético	503

Prefácio

Este texto cobre o material para um curso de Álgebra Linear ministrado para estudantes da área de Ciências Exatas. O texto pode, mas **não** é necessário, ser acompanhado de um programa como o MATLAB[®] *, SciLab ou o Maxima.

O conteúdo é dividido em quatro capítulos. No Capítulo 1 o conceito de vetor do \mathbb{R}^n é introduzido e então são definidos os espaços vetoriais abstratos. Dependência Linear, conjunto de geradores e base são definidos o mais geral possível de forma a incluir espaços vetoriais que possuem uma base infinita, como o espaço dos polinômios, por exemplo. Splines são estudados neste capítulo como uma aplicação dos conceitos apresentados.

No Capítulo 2 são estudados os espaços com produto interno, bases ortonormais e complemento ortogonal. Contém também uma aplicação aos polinômios de Legendre e uma aplicação às séries de Fourier.

O Capítulo 3 aborda transformações lineares. Aqui é apresentada uma abordagem bastante geométrica deste tema. A matriz mudança de base aparece de maneira natural como a matriz da transformação identidade em relação a duas bases. A matriz jacobiana de uma transformação é apresentada como uma aplicação ao tema.

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

Aqui também é estudada a adjunta de uma transformação linear e é apresentada uma aplicação ao problema de quadrados mínimos.

O Capítulo 4 traz um estudo da diagonalização de operadores, incluindo a diagonalização de operadores normais e auto-adjuntos e uma aplicação na identificação de cônicas. A última seção traz a forma canônica de Jordan. Além de serem provados resultados sobre existência, é mostrado também como se obtém uma base de autovetores generalizados em relação a qual a matriz do operador está na forma canônica de Jordan. Como aplicações são apresentados: cálculo das potências de uma matriz, funções de matrizes e Sistemas de equações diferenciais.

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 funções pré-definidas, pacotes de funções para tarefas específicas. Um pacote chamado `gaa1` 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 `gaa1`. 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 ao professor Helder C. Rodrigues pelas frutíferas discussões e aos professores que colaboraram apresentando correções, críticas e sugestões, entre eles Joana Darc A. S. da Cruz, Francisco Satuf, Hamilton P. Bueno e Antônio J. Engler.

Histórico

Julho 2010 Várias correções. O texto foi completamente reformatado.

Março 2006 Várias correções. Foi acrescentado o Exemplo 3.40 na página 282. Foram acrescentados diagramas para a formação da base de Jordan no Capítulo 4.

Julho 2004 Na página 74 o Exemplo 1.65 foi substituído pela Proposição 1.11 com demonstração.

Julho 2003 Na página 31 foi definido espaço vetorial finitamente gerado. Foi acrescentado o Exemplo 1.65 na página 46. As Proposições 4.5 na página 306 e 4.20 na página 367 foram reescritas, assim como as suas demonstrações. Foram adicionados dois exercícios teóricos à seção 4.1.

Junho 2002 Criado a partir da Parte II do texto 'Geometria Analítica e Álgebra Linear' para ser usado numa disciplina de Álgebra Linear.

Sugestão de Cronograma

Capítulo 1	Seções 1.1 a 1.4	16 aulas
Capítulo 2	Seções 2.1 e 2.2	12 aulas
Capítulo 3	Seções 3.1 a 3.4	16 aulas
Capítulo 4	Seções 4.1 a 4.4	16 aulas
	Total	60 aulas

Capítulo 1

Espaços Vetoriais

1.1 Definição e Exemplos

Os vetores no plano são definidos por pares ordenados de números reais e que vetores no espaço são definidos por ternos ordenados de números reais. Muito do que é estudado sobre vetores em Geometria Analítica pode ser estendido para n -uplas de números reais, em que n pode ser um número inteiro positivo.

1.1.1 Os Espaços \mathbb{R}^n

Definição 1.1. O espaço \mathbb{R}^n (n é um inteiro positivo qualquer) é definido pelo conjunto de todas as n -uplas ordenadas $X = (x_1, \dots, x_n)$ 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 1.1), ou como um vetor, neste caso x_1 , x_2 e x_3 são as componentes do vetor (Figura 1.2). Também no \mathbb{R}^n uma n -upla pode ser pensada como um vetor ou como um ponto. Por exemplo, a quártupla $X = (1, -2, 3, 5, 4)$ 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 $V = (v_1, \dots, v_n)$ e $W = (w_1, \dots, w_n)$ no \mathbb{R}^n são considerados **iguais** se $v_1 = w_1, \dots, v_n = w_n$. 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 1.2. (a) A soma de dois vetores $V = (v_1, \dots, v_n)$ e $W = (w_1, \dots, w_n)$ do \mathbb{R}^n é definida por

$$V + W = (v_1 + w_1, \dots, v_n + w_n); \quad (1.1)$$

(b) A multiplicação de um vetor $V = (v_1, \dots, v_n)$ do \mathbb{R}^n por um escalar α é definida por

$$\alpha V = (\alpha v_1, \dots, \alpha v_n). \quad (1.2)$$

O **vetor nulo** do \mathbb{R}^n é denotado por $\vec{0}$ e é definido por $\vec{0} = (0, \dots, 0)$. Se $V = (v_1, \dots, v_n)$ é um vetor do \mathbb{R}^n , então o **simétrico de V** é denotado por $-V$ e é definido por $-V = (-v_1, \dots, -v_n)$. 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 $V = (v_1, \dots, v_n)$ do \mathbb{R}^n pode também ser escrito na notação matricial como uma **matriz linha** ou como uma **matriz coluna**:

$$V = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \quad \text{ou} \quad V = [v_1 \quad \dots \quad v_n].$$

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

$$V + W = [v_1 \quad \dots \quad v_n] + [w_1 \quad \dots \quad w_n] = [v_1 + w_1 \quad \dots \quad v_n + w_n],$$

$$\alpha V = \alpha [v_1 \quad \dots \quad v_n] = [\alpha v_1 \quad \dots \quad \alpha v_n]$$

produzem os mesmos resultados que as operações vetoriais

$$V + W = (v_1, \dots, v_n) + (w_1, \dots, w_n) = (v_1 + w_1, \dots, v_n + w_n),$$

$$\alpha V = \alpha(v_1, \dots, v_n) = (\alpha v_1, \dots, \alpha v_n).$$

No teorema seguinte enunciamos as propriedades mais importantes da soma de vetores e multiplicação de vetores por escalar no \mathbb{R}^n .

Teorema 1.1. *Sejam $U = (u_1, \dots, u_n)$, $V = (v_1, \dots, v_n)$ e $W = (w_1, \dots, w_n)$ 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 (ver por exemplo, [24]). ■

1.1.2 Espaços Vetoriais Abstratos

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 abstraíndo-se as propriedades mais importantes de vetores no \mathbb{R}^n . Assim, os vetores do \mathbb{R}^n satisfarão automaticamente estes axiomas.

Definição 1.3. Dizemos que um conjunto $\mathbb{V} \neq \emptyset$, 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 $\vec{0} \in \mathbb{V}$, tal que $V + \vec{0} = \vec{0} + V = V$, para todo $V \in \mathbb{V}$;

(4) Para cada $V \in \mathbb{V}$, existe um elemento $-V \in \mathbb{V}$ tal que $V + (-V) = (-V) + V = \vec{0}$;

(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 $V \in \mathbb{V}$, $1V = V$.

Os elementos de \mathbb{V} são chamados **vetores**. O vetor $\vec{0}$ 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 1.1. Para n um número inteiro positivo, o conjunto $\mathbb{V} = \mathbb{R}^n$ com as operações de soma e multiplicação por escalar definidas em (1.1) e (1.2) é um espaço vetorial sobre \mathbb{R} , pelo **Teorema 1.1 na página 4**. Em particular \mathbb{R} é um espaço vetorial sobre ele mesmo.

Exemplo 1.2. 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 1.3. Segue-se das propriedades da álgebra matricial, 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 1.4. Seja \mathcal{X} um conjunto não vazio qualquer. Seja $\mathcal{F}(\mathcal{X}; \mathbb{R})$ o conjunto das funções reais, $f : \mathcal{X} \rightarrow \mathbb{R}$. 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 $\{1, \dots, n\}$, então $\mathcal{F}(\mathcal{X}; \mathbb{R}) = \mathbb{R}^n$, pois podemos identificar cada vetor (x_1, \dots, x_n) de \mathbb{R}^n com a função $f : \{1, \dots, n\} \rightarrow \mathbb{R}$ definida por $f(1) = x_1, \dots, f(n) = x_n$.

Se \mathcal{X} é igual ao produto cartesiano $\{1, \dots, m\} \times \{1, \dots, n\}$, então $\mathcal{F}(\mathcal{X}; \mathbb{R}) = \mathcal{M}_{mn}$, pois podemos identificar cada matriz $(a_{ij})_{mn}$ com a função $f : \{1, \dots, m\} \times \{1, \dots, n\} \rightarrow \mathbb{R}$ definida por

$$f(1, 1) = a_{11}, \dots, f(1, n) = a_{1n}, \dots, f(m, 1) = a_{m1}, \dots, f(m, n) = a_{mn}.$$

Exemplo 1.5. O conjunto \mathbb{R}^∞ das seqüências de números reais, ou seja, o conjunto das listas infinitas $(x_1, x_2, \dots, x_n, \dots)$ tais que $x_n \in \mathbb{R}$, para $n = 1, 2, 3, \dots$, com as operações

$$\begin{aligned} (x_1, x_2, \dots, x_n, \dots) + (y_1, y_2, \dots, y_n, \dots) &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n, \dots) \\ \alpha(x_1, x_2, \dots, x_n, \dots) &= (\alpha x_1, \alpha x_2, \dots, \alpha x_n, \dots) \end{aligned}$$

é um espaço vetorial sobre \mathbb{R} . Pois $\mathbb{R}^\infty = \mathcal{F}(\{1, 2, 3, \dots\}; \mathbb{R})$, já que podemos identificar cada sequência (x_n) , com a função $f : \{1, \dots, n, \dots\} \rightarrow \mathbb{R}$ definida por $f(1) = x_1, \dots, f(n) = x_n, \dots$

Exemplo 1.6. 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

$$p(t) = a_0 + a_1t + a_2t^2 + \dots + a_nt^n + \dots = \sum_{j \in \mathbb{N}} a_j t^j,$$

em que existe um inteiro positivo n tal que $a_j = 0$, para todo inteiro $j > n$ e $a_0, a_1, \dots, a_n \in \mathbb{R}$. O polinômio identicamente nulo é aquele em que $a_j = 0$, para todo $j \in \mathbb{N}$.

Sejam $p(t) = a_0 + a_1t + \dots + a_mt^m + \dots = \sum_{j \in \mathbb{N}} a_j t^j$ e $q(t) = b_0 + b_1t + \dots + b_rt^r + \dots = \sum_{j \in \mathbb{N}} b_j t^j$ dois polinômios quaisquer. A soma é definida por

$$p(t) + q(t) = (a_0 + b_0) + (a_1 + b_1)t + \dots + (a_k + b_k)t^k + \dots = \sum_{j \in \mathbb{N}} (a_j + b_j)t^j.$$

A multiplicação por um escalar α é definida por

$$\alpha p(t) = (\alpha a_0) + (\alpha a_1)t + \dots + (\alpha a_n)t^n + \dots = \sum_{j \in \mathbb{N}} (\alpha a_j)t^j.$$

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.

$$(1) \quad p(t) + q(t) = \sum_{j \in \mathbb{N}} a_j t^j + \sum_{j \in \mathbb{N}} b_j t^j = \sum_{j \in \mathbb{N}} (a_j + b_j)t^j = \sum_{j \in \mathbb{N}} (b_j + a_j)t^j = q(t) + p(t).$$

$$(2) \quad p(t) + (q(t) + r(t)) = \sum_{j \in \mathbb{N}} a_j t^j + \left(\sum_{j \in \mathbb{N}} b_j t^j + \sum_{j \in \mathbb{N}} c_j t^j \right) = \sum_{j \in \mathbb{N}} [a_j + (b_j + c_j)] t^j = \\ \sum_{j \in \mathbb{N}} [(a_j + b_j) + c_j] t^j = \left(\sum_{j \in \mathbb{N}} a_j t^j + \sum_{j \in \mathbb{N}} b_j t^j \right) + \sum_{j \in \mathbb{N}} c_j t^j = (p(t) + q(t)) + r(t).$$

(3) Seja $\bar{0}(t)$ o polinômio nulo.

$$p(t) + \bar{0}(t) = \sum_{j \in \mathbb{N}} a_j t^j + \sum_{j \in \mathbb{N}} 0 t^j = \sum_{j \in \mathbb{N}} (a_j + 0) t^j = \sum_{j \in \mathbb{N}} a_j t^j = p(t).$$

(4) Defina o polinômio $(-p)(t) = \sum_{j \in \mathbb{N}} (-a_j) t^j$.

$$p(t) + (-p(t)) = \sum_{j \in \mathbb{N}} a_j t^j + \sum_{j \in \mathbb{N}} (-a_j) t^j = \sum_{j \in \mathbb{N}} (a_j + (-a_j)) t^j = \sum_{j \in \mathbb{N}} 0 t^j = \bar{0}(t).$$

$$(5) \quad \alpha(\beta p(t)) = \alpha \left(\sum_{j \in \mathbb{N}} \beta a_j t^j \right) = \sum_{j \in \mathbb{N}} (\alpha \beta a_j) t^j = (\alpha \beta) p(t).$$

$$(6) \quad \alpha(p(t) + q(t)) = \alpha \sum_{j \in \mathbb{N}} (a_j + b_j) t^j = \sum_{j \in \mathbb{N}} [\alpha(a_j + b_j)] t^j = \sum_{j \in \mathbb{N}} (\alpha a_j + \alpha b_j) t^j \\ = \sum_{j \in \mathbb{N}} (\alpha a_j) t^j + \sum_{j \in \mathbb{N}} (\alpha b_j) t^j = \alpha p(t) + \alpha q(t).$$

$$(7) \quad (\alpha + \beta)p(t) = \sum_{j \in \mathbb{N}} (\alpha + \beta) a_j t^j = \sum_{j \in \mathbb{N}} (\alpha a_j + \beta a_j) t^j = \sum_{j \in \mathbb{N}} (\alpha a_j) t^j + \sum_{j \in \mathbb{N}} (\beta a_j) t^j = \alpha p(t) + \beta p(t).$$

$$(8) \quad 1p(t) = \sum_{j \in \mathbb{N}} (1a_j) t^j = \sum_{j \in \mathbb{N}} a_j t^j = p(t).$$

Proposição 1.2. São válidas as seguintes propriedades em um espaço vetorial \mathbb{V} :

(a) $0V = \bar{0}$, para todo V em \mathbb{V} ;

- (b) $\alpha \vec{0} = \vec{0}$, para todo escalar α ;
 (c) Se $\alpha V = \vec{0}$, então $\alpha = 0$ ou $V = \vec{0}$;
 (d) $(-1)V = -V$, para todo V pertencente a \mathbb{V} .

Demonstração. (a) Usando-se o axioma (2) de espaço vetorial, temos que

$$0V + 0V = (0 + 0)V = 0V.$$

Somando-se o simétrico de $0V$ ao primeiro e ao último membro e usando os axiomas (2) e (4) temos que

$$-(0V) + 0V + 0V = -0V + 0V = \vec{0}.$$

Aplicando-se novamente o axioma (4) no primeiro membro, chegamos a $0V = \vec{0}$.

- (b) Este item se prova de forma inteiramente análoga ao anterior, mas a partir de $\alpha \vec{0} + \alpha \vec{0}$.
 (c) Se $\alpha \neq 0$, então pelos axiomas (8) e (5) e pelo item (b), temos que

$$V = 1V = \left(\frac{1}{\alpha}\alpha\right)V = \frac{1}{\alpha}(\alpha V) = \frac{1}{\alpha}\vec{0} = \vec{0}.$$

- (d) Usando-se os axiomas (8) e (7) e o item (a) temos que

$$(-1)V + V = (-1)V + 1V = (-1 + 1)V = 0V = \vec{0}$$

Somando-se $-V$ ao primeiro e ao último membro e usando os axiomas (2), (4), (3) temos que

$$(-1)V = \vec{0} + (-V) = -V.$$



Exercícios Numéricos (respostas na página 398)

- 1.1.1. Determine o vetor X , tal que $3X - 2V = 15(X - U)$, para vetores V e U fixos dados.
- 1.1.2. Determine o vetor X , tal que $\begin{cases} 6X - 2Y = U \\ 3X + Y = U + V \end{cases}$, para vetores V e U fixos dados.
- 1.1.3. Verifique que o polinômio $t^2 + 2t + 7$ é **combinação linear** (soma de múltiplos escalares) de $t^2 + 1$ e $t + 3$.
- 1.1.4. Verifique que a função constante igual a 3 é combinação linear de $g(t) = 5 \tan^2 t$ e $h(t) = \frac{2}{\cos^2 t}$.
- 1.1.5. 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) $(1, 1, 1)$; (c) $(-2, -1, 1)$;
 (b) $(4, 2, -6)$; (d) $(-1, 2, 3)$.
- 1.1.6. 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

- 1.1.7. 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.
- 1.1.8. Mostre que em um espaço vetorial o vetor nulo é único e para cada vetor V o simétrico $-V$ também é único.

- 1.1.9. Prove que em um espaço vetorial \mathbb{V} , $X + W = X + U$ implica que $W = U$.
- 1.1.10. Em um espaço vetorial, $\alpha X = \beta X$ implica que $\alpha = \beta$? E se $X \neq \vec{0}$?
- 1.1.11. Mostre que se V pertence a um espaço vetorial \mathbb{V} e n é um inteiro positivo, então $nV = V + \dots + V$ (n parcelas).

1.2 Subespaços

Definição 1.4. 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 1.3. *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 1.3 na página 5** 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 1.2**, $0V = \vec{0}$ e $-V = (-1)V$, ou seja, o vetor nulo $\vec{0}$ e o simétrico de V são múltiplos escalares de V , que por (0') pertence a \mathbb{W} . ■

Exemplo 1.7. Se \mathbb{V} é um espaço vetorial, então \mathbb{V} é um subespaço dele mesmo. E o subconjunto formado apenas pelo vetor nulo, $\mathbb{W} = \{\vec{0}\}$, é claramente um subespaço de \mathbb{V} . Assim, todo espaço vetorial $\mathbb{V} \neq \{\vec{0}\}$ possui pelo menos dois subespaços.

Exemplo 1.8. 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 1.9. Os subconjuntos

$$\mathcal{A} = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\} \text{ e } \mathcal{B} = \{(x, y) \in \mathbb{R}^2 \mid xy \geq 0\}$$

não são subespaços de \mathbb{R}^2 . Pois, para o primeiro, enquanto

$$V = (1, 1) \in \mathcal{A}, -V = (-1)V = (-1, -1) \notin \mathcal{A}.$$

Enquanto para o segundo,

$$V = (1, 0), W = (0, -1) \in \mathcal{B}, V + W = (1, -1) \notin \mathcal{B}.$$

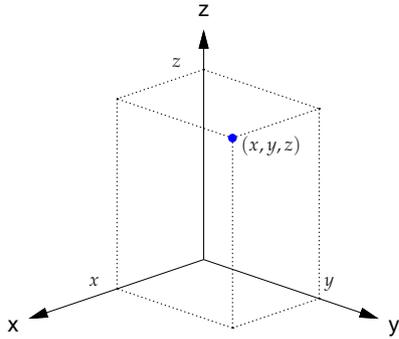


Figura 1.1: Ponto $(x, y, z) \in \mathbb{R}^3$

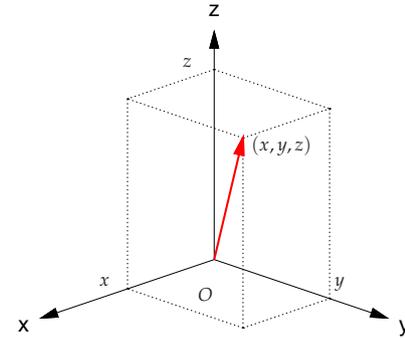


Figura 1.2: Vetor $(x, y, z) \in \mathbb{R}^3$

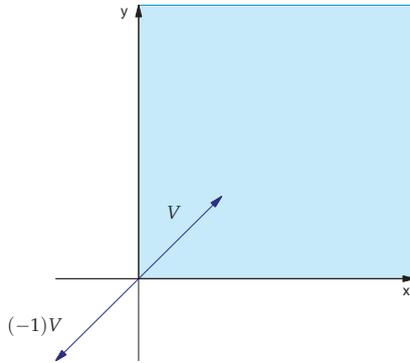


Figura 1.3: $\mathcal{A} = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}$

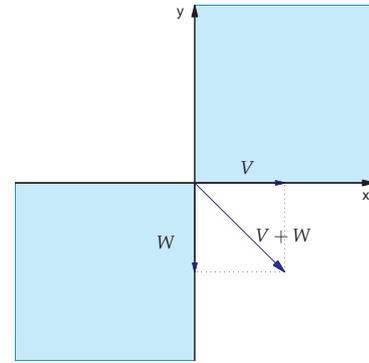


Figura 1.4: $\mathcal{B} = \{(x, y) \in \mathbb{R}^2 \mid xy \geq 0\}$

Exemplo 1.10. Seja $\mathbb{V} \neq \{0\}$ 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 1.11. Seja $N = (a_1, \dots, a_n)$ um vetor de \mathbb{R}^n fixo. O conjunto definido por

$$\mathbb{W} = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \dots + a_n x_n = 0\}$$

é um subespaço de \mathbb{R}^n .

- (0) Se $X = (x_1, \dots, x_n)$ e $Y = (y_1, \dots, y_n)$ pertencem a \mathbb{W} , então $a_1 x_1 + \dots + a_n x_n = 0$ e $a_1 y_1 + \dots + a_n y_n = 0$ e portanto

$$X + Y = (x_1 + y_1, \dots, x_n + y_n)$$

também pertence a \mathbb{W} , pois

$$a_1(x_1 + y_1) + \dots + a_n(x_n + y_n) = (a_1 x_1 + \dots + a_n x_n) + (a_1 y_1 + \dots + a_n y_n) = 0 + 0 = 0.$$

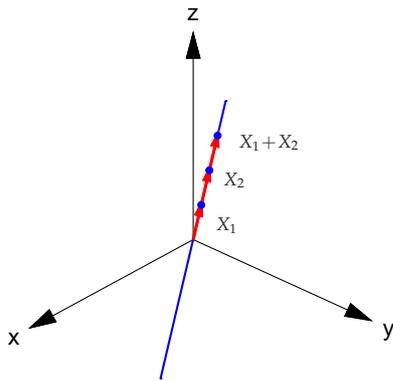


Figura 1.5: Soma de vetores da reta $X = \alpha V$

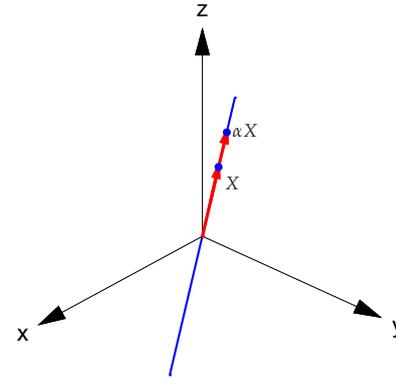


Figura 1.6: Multiplicação de vetor por escalar da reta $X = \alpha V$

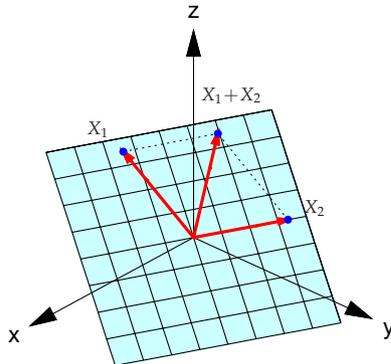


Figura 1.7: Soma de vetores do plano $a_1x + a_2y + a_3z = 0$

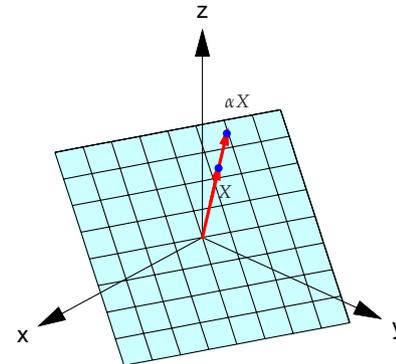


Figura 1.8: Multiplicação de vetor por escalar do plano $a_1x + a_2y + a_3z = 0$

(0') Se $X = (x_1, \dots, x_n)$ pertence a \mathbb{W} , então

$$\alpha X = (\alpha x_1, \dots, \alpha x_n)$$

também pertence a \mathbb{W} , pois

$$a_1(\alpha x_1) + \dots + a_n(\alpha x_n) = \alpha(a_1 x_1 + \dots + a_n x_n) = \alpha 0 = 0.$$

Por outro lado, suponha que o conjunto definido por

$$\mathbb{W} = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \dots + a_n x_n = c\}$$

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 $0X = \vec{0}$ também pertence a \mathbb{W} , ou seja, o subespaço tem que conter a origem. Substituindo-se $\vec{0} = (0, \dots, 0)$ na equação que define o conjunto, obtemos que $a_1 0 + \dots + a_n 0 = c$, ou seja, $c = 0$.

Se $N = (a_1, \dots, a_n) \neq \vec{0}$, 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 1.12. 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 1.13. 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 1.14. Seja $\mathbb{R}^{(\infty)}$ o conjunto das listas infinitas $(x_1, x_2, \dots, x_n, \dots)$ 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}^∞ , 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 1.15. O conjunto

$$\mathbb{W}_1 = \{f \in \mathcal{F}(\mathbb{R}; \mathbb{R}) \mid f(-x) = f(x) \text{ para todo } x \in \mathbb{R}\}$$

das funções, $f : \mathbb{R} \rightarrow \mathbb{R}$, pares e o conjunto

$$\mathbb{W}_2 = \{f \in \mathcal{F}(\mathbb{R}; \mathbb{R}) \mid f(-x) = -f(x) \text{ para todo } x \in \mathbb{R}\}$$

das funções, $f : \mathbb{R} \rightarrow \mathbb{R}$, 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 1.16. 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 1.17. 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.

Exemplo 1.18. Dados os números reais $x_1 < x_2 < \dots < x_n$. Seja \mathcal{S} o subconjunto de $\mathcal{C}^2[x_1, x_n]$ formado pelas funções que são polinômios de grau menor ou igual a 3 em cada subintervalo $[x_k, x_{k+1}]$, para $k = 1, \dots, n-1$. Este conjunto é chamado de **splines (cúbicos) em $[x_1, x_n]$ com pontos de quebra x_2, \dots, x_{n-1}** .*

Vamos mostrar que o conjunto \mathcal{S} é um subespaço de $\mathcal{C}^2[x_1, x_n]$. Sejam $f, g \in \mathcal{S}$ e α um escalar. Então

$$f(x) = \begin{cases} a_0^{(1)} + a_1^{(1)}x + a_2^{(1)}x^2 + a_3^{(1)}x^3, & \text{se } x_1 \leq x < x_2, \\ \vdots & \vdots \\ a_0^{(n-1)} + a_1^{(n-1)}x + a_2^{(n-1)}x^2 + a_3^{(n-1)}x^3, & \text{se } x_{n-1} \leq x \leq x_n, \end{cases}$$

$$g(x) = \begin{cases} b_0^{(1)} + b_1^{(1)}x + b_2^{(1)}x^2 + b_3^{(1)}x^3, & \text{se } x_1 \leq x < x_2, \\ \vdots & \vdots \\ b_0^{(n-1)} + b_1^{(n-1)}x + b_2^{(n-1)}x^2 + b_3^{(n-1)}x^3, & \text{se } x_{n-1} \leq x \leq x_n, \end{cases}$$

Assim as funções f e g são combinação linear de $4(n-1) = 4n - 4$ funções. Mas, os coeficientes não são independentes, pois f, g, f', g', f'' e g'' são contínuas nos pontos

*A motivação para estudar este conjunto \mathcal{S} vem do fato de que a equação da curva que descreve uma barra elástica que vai de x_1 a x_n sujeita a forças externas localizadas nos pontos x_2, \dots, x_{n-1} satisfaz $y^{(iv)}(x) = 0$ nos subintervalos (x_i, x_{i+1}) e $\lim_{x \rightarrow x_i^+} y'''(x) - \lim_{x \rightarrow x_i^-} y'''(x)$ proporcional à força aplicada no ponto x_i . O que leva a que $y(x)$ seja um polinômio de grau menor ou igual a 3 em cada subintervalo (x_i, x_{i+1}) com $y''(x)$ contínua no intervalo $[x_1, x_n]$.

de quebra x_2, \dots, x_{n-1} .

$$(f+g)(x) = \begin{cases} c_0^{(1)} + c_1^{(1)}x + c_2^{(1)}x^2 + c_3^{(1)}x^3, & \text{se } x_1 \leq x < x_2, \\ \vdots & \vdots \\ c_0^{(n-1)} + c_1^{(n-1)}x + c_2^{(n-1)}x^2 + c_3^{(n-1)}x^3, & \text{se } x_{n-1} \leq x \leq x_n, \end{cases}$$

em que $c_i^{(j)} = a_i^{(j)} + b_i^{(j)}$, para $i = 0, 1, 2, 3$ e $j = 1, \dots, n-1$.

$$(\alpha f)(x) = \begin{cases} \alpha a_0^{(1)} + \alpha a_1^{(1)}x + \alpha a_2^{(1)}x^2 + \alpha a_3^{(1)}x^3, & \text{se } x_1 \leq x < x_2, \\ \vdots & \vdots \\ \alpha a_0^{(n-1)} + \alpha a_1^{(n-1)}x + \alpha a_2^{(n-1)}x^2 + \alpha a_3^{(n-1)}x^3, & \text{se } x_{n-1} \leq x \leq x_n, \end{cases}$$

são splines, pois $f+g, \alpha f, (f+g)', (\alpha f)', (f+g)''$ e $(\alpha f)''$ também são contínuas nos pontos de quebra x_2, \dots, x_{n-1} .

Exemplo 1.19. Uma **equação diferencial** é uma equação envolvendo pelo menos uma das derivadas de uma função $y = y(t)$. Considere a equação diferencial

$$y'(t) = ay(t), \quad (1.3)$$

para $a \in \mathbb{R}, a \neq 0$.

Vamos determinar o conjunto solução da equação (1.3). Em primeiro lugar, suponhamos que a função $y : \mathbb{R} \rightarrow \mathbb{R}$ seja solução de (1.3). Vamos supor também que a função não se anula em nenhum ponto. Assim sendo, como y é contínua ela terá o mesmo sinal para todos os valores de $t \in \mathbb{R}$. Assim (1.3) é equivalente a

$$\frac{y'(t)}{y(t)} = a.$$

Integrando-se membro a membro obtemos

$$\ln |y(t)| = at + k, \quad \text{para uma constante } k \in \mathbb{R}.$$

Portanto,

$$y(t) = \pm e^k e^{at} = \alpha e^{at}.$$

Vamos mostrar que toda solução é desta forma, ou seja, que não existem outras soluções além das já encontradas. Seja $y(t)$ uma solução qualquer de (1.3). Defina

$$z(t) = e^{-at}y(t).$$

Derivando $z(t)$ obtemos,

$$z'(t) = (e^{-at})'y(t) + e^{-at}y'(t) = -ae^{-at}y(t) + ae^{-at}y(t) = 0.$$

O que implica que $z(t)$ é uma constante, ou seja, $z(t) = e^{-at}y(t) = \alpha$. O que implica que $y(t) = \alpha e^{at}$. Portanto o conjunto solução de (1.3) é o conjunto

$$\mathbb{W} = \{\alpha e^{at} \mid \alpha \in \mathbb{R}\},$$

que é um subespaço de $C^\infty(\mathbb{R})$, pois é o conjunto dos múltiplos escalares da função $f(t) = e^{at}$ (Exemplo 1.10 na página 17).

Exemplo 1.20. Considere a equação diferencial

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = f, \quad (1.4)$$

onde a_0, \dots, a_n e f são funções de t e $y^{(k)}$, denota a k -ésima derivada de y . Esta equação é chamada **linear**. Se as funções a_0, \dots, a_n forem constantes, dizemos que a

equação é **linear com coeficientes constantes**. Quando f é a função identicamente nula, a equação é chamada **linear homogênea**.

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = 0, \quad (1.5)$$

Vamos mostrar que o conjunto solução de uma equação diferencial linear homogênea com coeficientes constantes, \mathbb{W} , é um subespaço de $\mathcal{C}^\infty(\mathbb{R})$. Em primeiro lugar, \mathbb{W} é não vazio, pois a função identicamente nula é uma solução de (1.5)

(0) Sejam $y_1(t)$ e $y_2(t)$ duas soluções de (1.5). Vamos mostrar que $y(t) = y_1(t) + y_2(t)$ é também solução de (1.5).

$$\begin{aligned} & a_n (y_1 + y_2)^{(n)} + a_{n-1} (y_1 + y_2)^{(n-1)} + \dots + a_1 (y_1 + y_2)' + a_0 (y_1 + y_2) = \\ & (a_n y_1^{(n)} + a_{n-1} y_1^{(n-1)} + \dots + a_1 y_1' + a_0 y_1) + (a_n y_2^{(n)} + a_{n-1} y_2^{(n-1)} + \dots + a_1 y_2' + a_0 y_2) = \\ & 0 + 0 = 0 \end{aligned}$$

(0') Sejam $y(t)$ uma solução de (1.5) e α um escalar. Vamos mostrar que $z(t) = \alpha y(t)$ também é solução de (1.5).

$$\begin{aligned} & a_n (\alpha y)^{(n)} + a_{n-1} (\alpha y)^{(n-1)} + \dots + a_1 (\alpha y)' + a_0 (\alpha y) = \\ & \alpha (a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y) = \alpha 0 = 0 \end{aligned}$$

1.2.1 Soma e Interseção de Subespaços

Proposição 1.4. *Sejam \mathbb{W}_1 e \mathbb{W}_2 dois subespaços de um espaço vetorial \mathbb{V} . Então:*

- (a) $\mathbb{W}_1 \cap \mathbb{W}_2$ é um subespaço.
- (b) $\mathbb{W}_1 \cup \mathbb{W}_2$ é um subespaço se, e somente se, $\mathbb{W}_1 \subseteq \mathbb{W}_2$ ou $\mathbb{W}_2 \subseteq \mathbb{W}_1$.

Demonstração.

- (a) Sejam $V, W \in \mathbb{W}_1 \cap \mathbb{W}_2$ e α um escalar. Então, $V, W \in \mathbb{W}_1$ e $V, W \in \mathbb{W}_2$. O que implica que $V + W$ e αV pertencem a \mathbb{W}_1 e a \mathbb{W}_2 , ou seja, pertencem a $\mathbb{W}_1 \cap \mathbb{W}_2$.
- (b) Por contradição, suponha que exista $V \in \mathbb{W}_1$, que não pertença a \mathbb{W}_2 e $W \in \mathbb{W}_2$, que não pertença a \mathbb{W}_1 . Como a união $\mathbb{W}_1 \cup \mathbb{W}_2$, por hipótese, é um subespaço, então $U = V + W$ pertence a união $\mathbb{W}_1 \cup \mathbb{W}_2$, ou seja, U pertence a \mathbb{W}_1 ou a \mathbb{W}_2 . Se $U \in \mathbb{W}_1$, então $W = U - V$ pertenceria a \mathbb{W}_1 , contradizendo a hipótese feita inicialmente. Agora, se $U \in \mathbb{W}_2$, então $V = U - W$ pertenceria a \mathbb{W}_2 , contradizendo a hipótese feita inicialmente.



Exemplo 1.21. O conjunto solução de um sistema linear homogêneo com m equações e n incógnitas,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0 \end{cases}$$

onde a_{ij} são constantes reais, para $i = 1, \dots, m$ e $j = 1, \dots, n$, pode ser visto como a interseção de m subespaços de \mathbb{R}^n , que são hiperplanos que passam pela origem (Exemplo 1.11 na página 17).

Exemplo 1.22. Considere os sistemas lineares

$$(a) \begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (b) \begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ -2 & 4 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$(c) \begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Todos são sistemas homogêneos, portanto os conjuntos solução são subespaços de \mathbb{R}^3 .

- (a) A solução geral do primeiro sistema é $x = 2s - 3t$, $y = s$ e $z = t$ ou $x = 2y - 3z$, que é um plano que passa pela origem, com vetor normal $N = (1, -2, 3)$ (verifique!);
- (b) A solução geral do segundo sistema é $x = -5t$, $y = -t$ e $z = t$ que é a equação de uma reta que passa pela origem, com vetor diretor $V = (-5, -1, 1)$.
- (c) A solução do terceiro sistema é $x = 0$, $y = 0$ e $z = 0$, que é somente a origem $\{\vec{0}\}$.

O conjunto solução de um sistema homogêneo também é chamado de **espaço solução** do sistema homogêneo.

Definição 1.5. Sejam \mathbb{W}_1 e \mathbb{W}_2 dois subespaços de um espaço vetorial \mathbb{V} .

- (a) Definimos a **soma dos subespaços**, $\mathbb{W}_1 + \mathbb{W}_2$, como sendo o conjunto de todos os vetores de \mathbb{V} que são soma de um elemento de \mathbb{W}_1 com um elemento de \mathbb{W}_2 , ou seja,

$$\begin{aligned} \mathbb{W}_1 + \mathbb{W}_2 &= \{V_1 + V_2 \mid V_1 \in \mathbb{W}_1 \text{ e } V_2 \in \mathbb{W}_2\} \\ &= \{V \in \mathbb{V} \mid V = V_1 + V_2 \text{ com } V_1 \in \mathbb{W}_1 \text{ e } V_2 \in \mathbb{W}_2\} \end{aligned}$$

(b) Se o espaço \mathbb{V} é tal que

$$\begin{aligned}\mathbb{V} &= \mathbb{W}_1 + \mathbb{W}_2 \quad \text{e} \\ \mathbb{W}_1 \cap \mathbb{W}_2 &= \{\vec{0}\},\end{aligned}$$

dizemos que \mathbb{V} é **soma direta** de \mathbb{W}_1 e \mathbb{W}_2 e denotamos por $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2$.

Exemplo 1.23. Vimos no [Exemplo 1.12 na página 19](#) que 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 de \mathcal{M}_{nn} . Vamos mostrar que $\mathcal{M}_{nn} = \mathbb{W}_1 \oplus \mathbb{W}_2$. Seja A uma matriz qualquer $n \times n$. Sejam A_1 uma matriz simétrica e A_2 uma matriz anti-simétrica a serem determinadas tais que

$$A = A_1 + A_2. \quad (1.6)$$

Tomando a transposta da equação (1.6) e usando o fato de que $A_1^t = A_1$ e $A_2^t = -A_2$ obtemos

$$A^t = A_1 - A_2. \quad (1.7)$$

Tomando-se a soma e a diferença das equações (1.6) e (1.7) obtemos

$$2A_1 = A + A^t \quad \text{e} \quad 2A_2 = A - A^t,$$

ou seja, $A_1 = \frac{1}{2}(A + A^t)$ e $A_2 = \frac{1}{2}(A - A^t)$. Assim obtemos que a matriz A pode ser escrita como a soma de uma matriz simétrica e uma anti-simétrica da seguinte forma:

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t).$$

Assim, $\mathcal{M}_{nn} = \mathbb{W}_1 + \mathbb{W}_2$. Além disso, se uma matriz A é ao mesmo tempo simétrica e anti-simétrica, então $A = A^t = -A^t = -A$. O que implica que $A = \vec{0}$. Ou seja, $\mathbb{W}_1 \cap \mathbb{W}_2 = \{\vec{0}\}$.

Exemplo 1.24. Vimos no [Exemplo 1.15 na página 20](#) que o conjunto

$$\mathbb{W}_1 = \{f \in \mathcal{F}(\mathbb{R}; \mathbb{R}) \mid f(-x) = f(x) \text{ para todo } x \in \mathbb{R}\}$$

das funções, $f : \mathbb{R} \rightarrow \mathbb{R}$, pares e o conjunto

$$\mathbb{W}_2 = \{f \in \mathcal{F}(\mathbb{R}; \mathbb{R}) \mid f(-x) = -f(x) \text{ para todo } x \in \mathbb{R}\}$$

das funções, $f : \mathbb{R} \rightarrow \mathbb{R}$, ímpares são subespaços. Vamos mostrar que $\mathcal{F}(\mathbb{R}; \mathbb{R}) = \mathbb{W}_1 \oplus \mathbb{W}_2$. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função qualquer. Sejam f_1 uma função par e f_2 uma função ímpar a serem determinadas tais que

$$f(x) = (f_1 + f_2)(x) = f_1(x) + f_2(x). \quad (1.8)$$

Calculando-se $f(-x)$ usando a equação (1.8) e o fato de que $f_1(-x) = f_1(x)$ e $f_2(-x) = -f_2(x)$ obtemos

$$f(-x) = f_1(x) - f_2(x). \quad (1.9)$$

Tomando-se a soma e a diferença das equações (1.8) e (1.9) obtemos que

$$2f_1(x) = f(x) + f(-x) \quad \text{e} \quad 2f_2(x) = f(x) - f(-x),$$

ou seja, $f_1(x) = \frac{1}{2}(f(x) + f(-x))$ e $f_2(x) = \frac{1}{2}(f(x) - f(-x))$. Assim, toda função f pode ser escrita como a soma de uma função par e uma função ímpar da seguinte forma:

$$f(x) = \frac{1}{2}(f(x) + f(-x)) + \frac{1}{2}(f(x) - f(-x)).$$

Ou seja, $\mathcal{F}(\mathbb{R}; \mathbb{R}) = \mathbb{W}_1 + \mathbb{W}_2$. Agora, se $f \in \mathbb{W}_1 \cap \mathbb{W}_2$, então para todo $x \in \mathbb{R}$, $f(x) = f(-x) = -f(x)$, ou seja, $f(x) = 0$, para todo $x \in \mathbb{R}$. Portanto, f é a função identicamente nula e $\mathbb{W}_1 \cap \mathbb{W}_2 = \{\vec{0}\}$.

Proposição 1.5. *Sejam \mathbb{W}_1 e \mathbb{W}_2 subespaços de um espaço vetorial \mathbb{V} . Então:*

- (a) $\mathbb{W}_1 + \mathbb{W}_2$ é um subespaço que contém \mathbb{W}_1 e \mathbb{W}_2 .
- (b) $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2$ se, e somente se, todo elemento $V \in \mathbb{V}$ se escreve, de modo único, como soma $V = V_1 + V_2$, onde $V_1 \in \mathbb{W}_1$ e $V_2 \in \mathbb{W}_2$.

Demonstração.

(a) Claramente \mathbb{W}_1 e \mathbb{W}_2 estão contidos em $\mathbb{W}_1 + \mathbb{W}_2$. Vamos mostrar que $\mathbb{W}_1 + \mathbb{W}_2$ é um subespaço.

(0) Sejam $V = V_1 + V_2$ e $W = W_1 + W_2$, onde $V_1, W_1 \in \mathbb{W}_1$ e $V_2, W_2 \in \mathbb{W}_2$. Então,

$$V + W = (V_1 + V_2) + (W_1 + W_2) = \overbrace{(V_1 + W_1)}^{\in \mathbb{W}_1} + \overbrace{(V_2 + W_2)}^{\in \mathbb{W}_2}.$$

(0') Sejam $V = V_1 + V_2$, onde $V_1 \in \mathbb{W}_1$ e $V_2 \in \mathbb{W}_2$ e α um escalar.

$$\alpha V = \alpha(V_1 + V_2) = \overbrace{\alpha V_1}^{\in \mathbb{W}_1} + \overbrace{\alpha V_2}^{\in \mathbb{W}_2}.$$

Assim, pelo [Teorema 1.3 na página 14](#), $\mathbb{W}_1 + \mathbb{W}_2$ é um subespaço.

(b) Suponhamos, em primeiro lugar, que $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2$. Então, $\mathbb{W}_1 \cap \mathbb{W}_2 = \{\vec{0}\}$. Sejam $V_1, W_1 \in \mathbb{W}_1$ e $V_2, W_2 \in \mathbb{W}_2$ tais que

$$V_1 + V_2 = W_1 + W_2.$$

Somando-se $-W_1 - V_2$, obtemos

$$\underbrace{V_1 - W_1}_{\in \mathbb{W}_1} = \underbrace{W_2 - V_2}_{\in \mathbb{W}_2}.$$

O que implica que $V_1 - W_1 = W_2 - V_2 \in \mathbb{W}_1 \cap \mathbb{W}_2 = \{\bar{0}\}$. Logo, $V_1 = W_1$ e $V_2 = W_2$.

Por outro lado, suponhamos que todo elemento de $V \in \mathbb{V}$ se escreve, de modo único, como soma $V = V_1 + V_2$, onde $V_1 \in \mathbb{W}_1$ e $V_2 \in \mathbb{W}_2$. Seja $V \in \mathbb{W}_1 \cap \mathbb{W}_2$. Vamos mostrar que $V = \bar{0}$. Mas,

$$V = \underbrace{V}_{\in \mathbb{W}_1} + \underbrace{\bar{0}}_{\in \mathbb{W}_2} = \underbrace{\bar{0}}_{\in \mathbb{W}_1} + \underbrace{V}_{\in \mathbb{W}_2}.$$

Ou seja, se $V \neq \bar{0}$, teríamos duas formas de escrever V como uma soma de um elemento de \mathbb{W}_1 e um de \mathbb{W}_2 . Logo, $V = \bar{0}$ e $\mathbb{W}_1 \cap \mathbb{W}_2 = \{\bar{0}\}$. Portanto, como claramente $\mathbb{W} = \mathbb{W}_1 + \mathbb{W}_2$, temos que $\mathbb{W} = \mathbb{W}_1 \oplus \mathbb{W}_2$.

■

Na verdade $\mathbb{W}_1 + \mathbb{W}_2$ é o menor subespaço que contém \mathbb{W}_1 e \mathbb{W}_2 no sentido de que qualquer subespaço que contenha \mathbb{W}_1 e \mathbb{W}_2 tem que conter $\mathbb{W}_1 + \mathbb{W}_2$. Deixamos para o leitor como exercício a verificação deste fato.

Exemplo 1.25. Sejam \mathbb{W}_1 o conjunto dos polinômios da forma $p_1(t) = a + at + at^2$, para $a \in \mathbb{R}$ e \mathbb{W}_2 o conjunto dos polinômios da forma $p_2(t) = b + ct$, para $b, c \in \mathbb{R}$. Vamos mostrar que $\mathcal{P}_2 = \mathbb{W}_1 \oplus \mathbb{W}_2$. Seja $p(t) = \alpha + \beta t + \gamma t^2$ um polinômio qualquer de \mathcal{P}_2 . Vamos determinar $p_1 \in \mathbb{W}_1$ e $p_2 \in \mathbb{W}_2$ tais que

$$p = p_1 + p_2.$$

Agrupando os termos de mesmo grau obtemos

$$\alpha + \beta t + \gamma t^2 = (a + b) + (a + c)t + (a)t^2,$$

que é equivalente ao sistema

$$\begin{cases} a + b = \alpha \\ a + c = \beta \\ a = \gamma \end{cases}$$

que tem solução única $a = \gamma$, $b = \alpha - \gamma$ e $c = \beta - \gamma$. Como todo elemento de \mathcal{P}_2 se escreve de maneira única como a soma de um elemento de \mathbb{W}_1 e um de \mathbb{W}_2 , então pela **Proposição 1.5**, $\mathcal{P}_2 = \mathbb{W}_1 \oplus \mathbb{W}_2$.

1.2.2 Conjunto de Geradores

Definição 1.6. Seja \mathcal{X} um subconjunto não vazio de um espaço vetorial \mathbb{V} .

- (a) O conjunto $[\mathcal{X}]$ de todas as **combinações lineares** (somadas de múltiplos escalares) de vetores de \mathcal{X} , ou seja,

$$\begin{aligned} [\mathcal{X}] &= \{ \alpha_1 V_1 + \dots + \alpha_k V_k \mid \alpha_i \text{ são escalares e } V_i \in \mathcal{X}, i = 1, \dots, k \} \\ &= \{ V \in \mathbb{V} \mid V = \alpha_1 V_1 + \dots + \alpha_k V_k \text{ com } \alpha_i \text{ escalares e } V_i \in \mathcal{X}, i = 1, \dots, k \} \end{aligned}$$

é um subespaço de \mathbb{V} (verifique!) chamado de **espaço gerado por \mathcal{X}** .

- (b) Quando $[\mathcal{X}] = \mathbb{V}$, dizemos que \mathcal{X} é um **conjunto de geradores de \mathbb{V}** . Assim, \mathcal{X} é um conjunto de geradores de um espaço vetorial \mathbb{V} , se todo vetor V de \mathbb{V} pode ser escrito como combinação linear

$$V = \alpha_1 V_1 + \dots + \alpha_k V_k$$

de vetores V_1, \dots, V_k pertencentes a \mathcal{X} .

- (c) Se o conjunto de geradores \mathcal{X} de um espaço vetorial \mathbb{V} tem um número finito de elementos dizemos que o espaço vetorial é **finitamente gerado**.

Observação. A maior parte dos resultados que apresentaremos aqui refere-se a espaços vetoriais finitamente gerados.

Exemplo 1.26. Vamos verificar que os vetores $V_1 = (1, 1, 0)$, $V_2 = (0, 1, 1)$, $V_3 = (1, 0, 1)$ e $V_4 = (1, 2, 1)$ geram o \mathbb{R}^3 . A equação vetorial

$$x_1(1, 1, 0) + x_2(0, 1, 1) + x_3(1, 0, 1) + x_4(1, 2, 1) = (a, b, c) \quad (1.10)$$

ou

$$(x_1 + x_3 + x_4, x_1 + x_2 + 2x_4, x_2 + x_3 + x_4) = (a, b, c)$$

é equivalente ao sistema de equações lineares

$$\begin{cases} x_1 & & + x_3 & + x_4 & = a \\ x_1 & + x_2 & & + 2x_4 & = b \\ & & x_2 & + x_3 & + x_4 & = c \end{cases}, \quad (1.11)$$

Escalonando a matriz aumentada deste sistema

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & a \\ 1 & 1 & 0 & 2 & b \\ 0 & 1 & 1 & 1 & c \end{array} \right] \text{ obtemos a matriz } \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & \frac{a+b-c}{2} \\ 0 & 1 & 0 & 1 & \frac{b+c-a}{2} \\ 0 & 0 & 1 & 0 & \frac{a+c-b}{2} \end{array} \right].$$

Assim, o sistema (1.11) e a equação vetorial (1.10) possuem solução

$$(x_4 = \alpha, x_3 = \frac{a+c-b}{2}, x_2 = \frac{b+c-a}{2} - \alpha, x_1 = \frac{a+b-c}{2} - \alpha, \text{ para todo } \alpha \in \mathbb{R}).$$

Portanto, V_1, V_2, V_3 e V_4 geram o \mathbb{R}^3 .

Exemplo 1.27. Os vetores

$E_1 = (1, 0, \dots, 0), E_2 = (0, 1, 0, \dots, 0), \dots, E_n = (0, \dots, 0, 1)$ geram o \mathbb{R}^n . Vamos encontrar um conjunto de geradores para o \mathbb{R}^n . Um vetor qualquer do \mathbb{R}^n é da forma $V = (a_1, \dots, a_n)$ e pode ser escrito como uma soma de vetores, sendo um vetor para cada parâmetro e cada vetor depende apenas de um parâmetro, obtendo A equação

$$x_1 E_1 + \dots + x_n E_n = V$$

ou

$$x_1(1, 0, \dots, 0) + \dots + x_n(0, \dots, 0, 1) = (a_1, \dots, a_n)$$

ou ainda

$$(x_1, \dots, x_n) = (a_1, \dots, a_n)$$

tem solução $x_1 = a_1, \dots, x_n = a_n$. Em particular os vetores $\vec{i} = (1, 0, 0), \vec{j} = (0, 1, 0)$ e $\vec{k} = (0, 0, 1)$ geram o \mathbb{R}^3 .

Exemplo 1.28. Para $i = 1, \dots, m$ e $j = 1, \dots, n$ seja E_{ij} a matriz $m \times n$ cujo elemento na posição ij é igual a 1 e os demais elementos são iguais a zero. Vamos mostrar que as matrizes E_{ij} geram o espaço das matrizes $m \times n$. Seja $A = (a_{ij})$ uma matriz qualquer $m \times n$. A equação

$$x_{11}E_{11} + \dots + x_{1n}E_{1n} + \dots + x_{m1}E_{m1} + \dots + x_{mn}E_{mn} = A.$$

tem solução $x_{ij} = a_{ij}$. Assim toda matriz $m \times n$ é combinação linear das matrizes E_{ij} , que portanto geram \mathcal{M}_{mn} . Em particular as matrizes

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \text{ e } E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

geram o espaço das matrizes 2×2 .

Exemplo 1.29. O conjunto $\mathcal{X} = \{1, t, t^2, \dots, t^n, \dots\}$ é um conjunto de geradores para o espaço $\mathcal{P} = \mathbb{R}[t]$, pois todo polinômio

$$p(t) = a_0 + \dots + a_n t^n = a_0(1) + a_1(t) + \dots + a_n(t^n)$$

é combinação linear de elementos de \mathcal{X} . Além disso, $\mathcal{X}_n = \{1, t, t^2, \dots, t^n\}$ é um conjunto de geradores para o espaço \mathcal{P}_n , pois todo polinômio de grau no máximo n é combinação linear de elementos de \mathcal{X}_n .

Observação. O exemplo anterior mostra que um conjunto de geradores ser um conjunto infinito não significa que todo vetor do espaço tenha que ser escrito como uma “combinação linear infinita” dos geradores do espaço. Em caso de dúvida dê uma olhada novamente na definição de conjunto de geradores.

Exemplo 1.30. Vamos mostrar que as matrizes $M_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, $M_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, $M_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ e $M_4 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ geram o espaço das matrizes simétricas 2×2 . Seja $M = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ uma matriz simétrica qualquer. A equação matricial

$$x_1 M_1 + x_2 M_2 + x_3 M_3 + x_4 M_4 = M$$

é equivalente ao sistema linear

$$\begin{cases} x_1 & & + x_3 & + x_4 & = a \\ x_1 & + x_2 & & + 2x_4 & = b \\ x_1 & + x_2 & & + 2x_4 & = b \\ & x_2 & + x_3 & + x_4 & = c \end{cases} ,$$

que possui solução $x_1 = \frac{a+b-c}{2} - \alpha$, $x_2 = \frac{b+c-a}{2} - \alpha$, $x_3 = \frac{a+c-b}{2}$ e $x_4 = \alpha$, para $\alpha \in \mathbb{R}$. Portanto, M_1, M_2, M_3 e M_4 geram o espaço das matrizes simétricas 2×2 .

Exemplo 1.31. Considere o sistema linear homogêneo $AX = \vec{0}$, onde

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

Já vimos que o conjunto solução de um sistema homogêneo é um subespaço. Vamos encontrar um conjunto de vetores que gere este subespaço. Escalonando a matriz aumentada do sistema acima, obtemos a matriz escalonada reduzida

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

E assim a solução geral do sistema pode ser escrita como

$$x_1 = -\alpha - 2\beta, \quad x_2 = \alpha, \quad x_3 = \beta, \quad x_4 = \beta, \quad \text{para todos os valores de } \alpha, \beta \in \mathbb{R},$$

ou seja, o conjunto solução do sistema $AX = \vec{0}$ é

$$\mathbb{W} = \{(x_1, x_2, x_3, x_4) = (-\alpha - 2\beta, \alpha, \beta, \beta) \mid \alpha, \beta \in \mathbb{R}\}.$$

Agora, um elemento qualquer de \mathbb{W} pode ser escrito como uma soma de vetores, sendo um vetor para cada parâmetro e cada vetor depende apenas de um parâmetro, obtendo

$$(-\alpha - 2\beta, \alpha, \beta, \beta) = (-\alpha, \alpha, 0, 0) + (-2\beta, 0, \beta, \beta) = \alpha(-1, 1, 0, 0) + \beta(-2, 0, 1, 1).$$

Portanto, $X_1 = (-1, 1, 0, 0)$ e $X_2 = (-2, 0, 1, 1)$ geram \mathbb{W} .

Exemplo 1.32. Sejam \mathbb{W} o espaço gerado por $V_1 = (-1, 1, 0)$ e $V_2 = (-1, 0, 1)$ e \mathbb{V} o espaço gerado por $V_3 = (1, 0, -4)$ e $V_4 = (0, 1, -2)$. Devemos encontrar os vetores que são combinações lineares de V_1 e V_2 que são também combinações lineares de V_3 e V_4 , ou seja, devemos encontrar vetores V que satisfazem as duas equações:

$$V = xV_1 + yV_2 \tag{1.12}$$

$$V = zV_3 + wV_4 \tag{1.13}$$

Para isso, podemos resolver a equação

$$xV_1 + yV_2 = zV_3 + wV_4, \quad \text{ou}$$

$$xV_1 + yV_2 + z(-V_3) + w(-V_4) = \vec{0}.$$

Esta equação é equivalente ao sistema linear $AX = \vec{0}$, onde

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

A forma escalonada reduzida da matriz aumentada $[A | \vec{0}]$ é $\left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 2/3 & 0 \\ 0 & 0 & 1 & 1/3 & 0 \end{array} \right]$.

Assim, a solução do sistema linear é $w = t, z = -t/3, y = -2t/3$ e $x = t$, para todo $t \in \mathbb{R}$. Substituindo-se x e y em (1.12), obtemos que a interseção $\mathbb{V} \cap \mathbb{W}$ é formada por vetores da forma

$$V = tV_1 - \frac{2t}{3}V_2 = t(V_1 - \frac{2}{3}V_2) = t(-1/3, 1, -2/3)$$

ou substituindo-se z e w em (1.13),

$$V = -\frac{t}{3}V_3 + tV_4 = t(-\frac{1}{3}V_3 + V_4) = t(-1/3, 1, -2/3).$$

Assim, a reta que é a interseção, $\mathbb{V} \cap \mathbb{W}$, tem equação $(x, y, z) = t(-1, 3, -2)$, para todo $t \in \mathbb{R}$, que é também um subespaço.

Exemplo 1.33. Dados os números reais $x_1 < x_2 < \dots < x_n$, igualmente espaçados, isto é, $x_{k+1} - x_k = (x_n - x_1)/(n - 1)$, para $k = 1, \dots, n - 1$. Seja \mathcal{S} o espaço dos splines (cúbicos) em $[x_1, x_n]$ com pontos de quebra x_2, \dots, x_{n-1} (Exemplo 1.18 na página 21).

Seja f um elemento genérico de \mathcal{S} . Então

$$f(x) = \begin{cases} a_0^{(1)} + a_1^{(1)}x + a_2^{(1)}x^2 + a_3^{(1)}x^3, & \text{se } x_1 \leq x < x_2, \\ \vdots & \vdots \\ a_0^{(n-1)} + a_1^{(n-1)}x + a_2^{(n-1)}x^2 + a_3^{(n-1)}x^3, & \text{se } x_{n-1} \leq x \leq x_n, \end{cases}$$

Assim a função f é uma combinação linear de $4(n - 1) = 4n - 4$ funções. Mas, os coeficientes não são independentes, pois f, f' e f'' são contínuas nos pontos de quebra

x_2, \dots, x_{n-1} . Do fato de que f , f' e f'' são contínuas em x_2 obtemos as equações

$$\begin{cases} a_0^{(1)} + a_1^{(1)}x_2 + a_2^{(1)}x_2^2 + a_3^{(1)}x_2^3 - a_0^{(2)} - a_1^{(2)}x_2 - a_2^{(2)}x_2^2 - a_3^{(2)}x_2^3 = 0 \\ a_1^{(1)} + 2a_2^{(1)}x_2 + 3a_3^{(1)}x_2^2 - a_1^{(2)} - 2a_2^{(2)}x_2 - 3a_3^{(2)}x_2^2 = 0 \\ 2a_2^{(1)} + 3a_3^{(1)}x_2 - 2a_2^{(2)} - 6a_3^{(2)}x_2 = 0 \end{cases}$$

Do fato de que f , f' e f'' são contínuas em x_3 obtemos as equações

$$\begin{cases} a_0^{(2)} + a_1^{(2)}x_3 + a_2^{(2)}x_3^2 + a_3^{(2)}x_3^3 - a_0^{(3)} - a_1^{(3)}x_3 - a_2^{(3)}x_3^2 - a_3^{(3)}x_3^3 = 0 \\ a_1^{(2)} + 2a_2^{(2)}x_3 + 3a_3^{(2)}x_3^2 - a_1^{(3)} - 2a_2^{(3)}x_3 - 3a_3^{(3)}x_3^2 = 0 \\ 2a_2^{(2)} + 3a_3^{(2)}x_3 - 2a_2^{(3)} - 6a_3^{(3)}x_3 = 0 \end{cases}$$

Juntando os dois conjuntos de equações obtidos aos que podemos obter para os pontos de quebra restantes obtemos um sistema linear homogêneo triangular superior com $3(n-2) = 3n-6$ equações e $4n-4$ incógnitas. Como o sistema é triangular superior, então ao escalonarmos a matriz do sistema obteremos que a solução depende de $(4n-4) - (3n-6) = n+2$ parâmetros. Assim podemos escrever todo spline de \mathcal{S} como combinação linear de apenas $n+2$ splines. Ou seja, podemos ter um conjunto de geradores para \mathcal{S} com apenas $n+2$ splines.

Exercícios Numéricos (respostas na página 400)

1.2.1. Considere os seguintes conjuntos de vetores. Quais deles são subespaços de \mathbb{R}^3 ?

- (a) (x, y, z) ; tais que $z = x^3$
- (b) (x, y, z) , tais que $z = x + y$;
- (c) (x, y, z) , tais que $z \geq 0$;
- (d) (x, y, z) , tais que $z = 0$ e $xy \geq 0$;
- (e) (x, y, z) , tais que $x = z = 0$;
- (f) (x, y, z) , tais que $x = -z$;
- (g) (x, y, z) , tais que $y = 2x + 1$;
- (h) (x, y, z) , tais que $z^2 = x^2 + y^2$.

1.2.2. Considere os seguintes conjuntos de vetores. Quais deles são subespaços de \mathbb{R}^4 ?

- (a) (x, y, z, w) , tais que $x - y = 2$;
- (b) (x, y, z, w) , tais que $z = x = 2y$ e $w = x - 3y$;
- (c) (x, y, z, w) , tais que $x = y = 0$;
- (d) (x, y, z, w) , tais que $x = 0$ e $y = -w$;

1.2.3. Verifique se os seguintes conjuntos são espaços vetoriais.

- (a) O conjunto das funções f em $C^0[-1, 1]$ tais que $f(-1) = f(1)$.
- (b) O conjunto de todas as funções contínuas não decrescentes em $[0, 1]$.
- (c) O conjunto de todas as funções f em $C^0[-1, 1]$ tais que $f(-1) = 0$ ou $f(1) = 0$.
- (d) O conjunto de todas as funções f em $C^0[-1, 1]$ tais que $f(-1) = 0$ e $f(1) = 0$.
- (e) O conjunto de todos os polinômios de grau 3.

1.2.4. Seja A uma matriz $n \times n$ fixada. Determine se os conjuntos dados são ou não espaços vetoriais.

- (a) $\{B \in \mathcal{M}_{nn} \mid AB = BA\}$.
- (b) $\{B \in \mathcal{M}_{nn} \mid AB \neq BA\}$.
- (c) $\{B \in \mathcal{M}_{nn} \mid BA = \vec{0}\}$.

1.2.5. Seja \mathcal{X} um conjunto não vazio. Mostre que para qualquer $t_0 \in \mathcal{X}$ fixado, o conjunto

$$\{f \in \mathcal{F}(\mathcal{X}; \mathbb{R}) \mid f(t_0) = 0\}$$

é um subespaço de $\mathcal{F}(\mathcal{X}; \mathbb{R})$.

1.2.6. Sejam

$$\mathbb{W}_1 = \{(x, y, z) \in \mathbb{R}^3 \mid x = y = z\}$$

o subespaço de \mathbb{R}^3 formado pelos vetores que têm as três componentes iguais e

$$\mathbb{W}_2 = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}$$

o subespaço de \mathbb{R}^3 formado pelos vetores que têm a terceira componente igual a zero. Mostre que

$$\mathbb{R}^3 = \mathbb{W}_1 \oplus \mathbb{W}_2.$$

1.2.7. Quais dos seguintes conjuntos de vetores geram o \mathbb{R}^4 ?

- (a) $\{(1, 0, 0, 1), (0, 1, 0, 0), (1, 1, 1, 1), (1, 1, 1, 0)\}$;
- (b) $\{(1, 2, 1, 0), (1, 1, -1, 0), (0, 0, 0, 1)\}$;
- (c) $\{(6, 4, -2, 4), (2, 0, 0, 1), (3, 2, -1, 2), (5, 6, -3, 2), (0, 4, -2, -1)\}$;
- (d) $\{(1, 1, 0, 0), (1, 2, -1, 1), (0, 0, 1, 1), (2, 1, 2, 2)\}$;

1.2.8. Encontre um conjunto de vetores que gera o espaço solução do sistema homogêneo $AX = \vec{0}$, em que

$$(a) A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 3 & 1 \\ 2 & 1 & 3 & 1 \end{bmatrix}; \quad (b) A = \begin{bmatrix} 1 & 1 & 2 & -1 \\ 2 & 3 & 6 & -2 \\ -2 & 1 & 2 & 2 \end{bmatrix}.$$

1.2.9. Considere os seguintes subespaços de \mathbb{R}^3 : $\mathbb{V} = [(-1, 2, 3), (1, 3, 4)]$ e $\mathbb{W} = [(1, 2, -1), (0, 1, 1)]$. Encontre a equação paramétrica da reta $\mathbb{V} \cap \mathbb{W}$. 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 . (Sugestão: revise o [Exemplo 1.32 na página 36.](#))

1.2.10. Verifique que o espaço gerado pelo conjunto $\{\sin^2 t, \cos^2 t\}$ é igual ao espaço gerado por $\{1, \cos 2t\}$

1.2.11. Encontre conjuntos geradores para os seguintes subespaços:

(a) $\{(a, b, c) \in \mathbb{R}^3 \mid 3a - 5b + 2c = 0\}$

(c) $\{p \in \mathcal{P}_3 \mid p(2) = 0\}$

(b) $\{A \in \mathcal{M}_{22} \mid 3a_{11} = 2a_{12}\}$

(d) $\{p \in \mathcal{P}_3 \mid p(2) = p(-1)\}$

1.2.12. Mostre que $\{2, t + 1, t^2 + 1, \dots, t^n + 1, \dots\}$ é um conjunto de geradores para $\mathcal{P} = \mathbb{R}[t]$.

1.2.13. Encontre um conjunto de geradores para o subespaço dos polinômios pares, ou seja, para o subespaço dos polinômios $p(t)$ que satisfazem $p(-t) = p(t)$, para todo $t \in \mathbb{R}$.

Exercícios Teóricos

1.2.14. Mostre que o conjunto dos quocientes de polinômios chamado **frações racionais**,

$$\mathbb{R}(t) = \left\{ \frac{p(t)}{q(t)} \mid p(t), q(t) \in \mathbb{R}[t], q(t) \neq \bar{0} \right\},$$

é um espaço vetorial sobre \mathbb{R} .

1.2.15. Mostre que um subconjunto não vazio, \mathbb{W} , de uma espaço vetorial \mathbb{V} é um subespaço se, e somente se, $V + \alpha W$ pertence a \mathbb{W} , para quaisquer vetores V e W de \mathbb{W} e qualquer escalar α .

1.2.16. Mostre que $\mathbb{W}_1 + \mathbb{W}_2$ é o menor subespaço que contém \mathbb{W}_1 e \mathbb{W}_2 no sentido de que qualquer subespaço que contenha \mathbb{W}_1 e \mathbb{W}_2 tem que conter $\mathbb{W}_1 + \mathbb{W}_2$.

1.2.17. Seja \mathcal{X} um subconjunto não vazio de um espaço vetorial \mathbb{V} .

(a) Mostre que o conjunto, $[\mathcal{X}]$, de todas as combinações lineares

$$\alpha_1 V_1 + \dots + \alpha_k V_k$$

de vetores $V_1, \dots, V_k \in \mathcal{X}$ é um subespaço de \mathbb{V} .

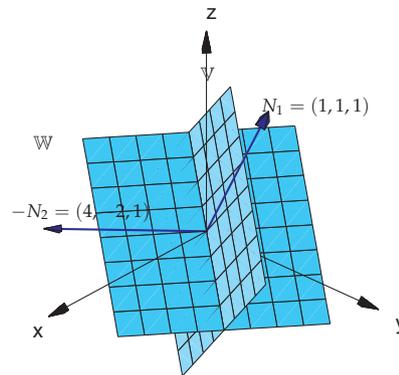
(b) Mostre que $[\mathcal{X}]$ é o menor subespaço de \mathbb{V} que contém \mathcal{X} , ou seja, se \mathbb{W} é um subespaço de \mathbb{V} e $\mathcal{X} \subseteq \mathbb{W}$, então $[\mathcal{X}] \subseteq \mathbb{W}$.

(c) Mostre que se \mathcal{X}_1 e \mathcal{X}_2 são subconjuntos de um espaço vetorial \mathbb{V} e $\mathcal{X}_1 \subseteq \mathcal{X}_2$, então $[\mathcal{X}_1] \subseteq [\mathcal{X}_2]$.

(d) Mostre que se \mathcal{X}_1 e \mathcal{X}_2 são subconjuntos de um espaço vetorial \mathbb{V} , então $[\mathcal{X}_1 \cup \mathcal{X}_2] = [\mathcal{X}_1] + [\mathcal{X}_2]$.

1.2.18. Sejam \mathcal{X} um subconjunto de um espaço vetorial \mathbb{V} e \mathcal{Y} um conjunto obtido de \mathcal{X} substituindo-se um de seus elementos V por $V + \alpha U$, para $U \in \mathcal{X}$ e $\alpha \in \mathbb{R}$. Mostre que $[\mathcal{X}] = [\mathcal{Y}]$.

1.2.19. Mostre que $\{V_1, \dots, V_k\}$ subconjunto de um espaço vetorial \mathbb{V} e $\{V_1, V_2 - V_1, \dots, V_k - V_1\}$ geram o mesmo subespaço de \mathbb{V} .

Figura 1.9: Os subespaços W, V e $V \cap W$

1.3 Dependência Linear

Definição 1.7. Um subconjunto \mathcal{X} não vazio de um espaço vetorial \mathbb{V} é chamado **conjunto linearmente dependente (L.D.)**, se existe um número finito de vetores $V_1, \dots, V_k \in \mathcal{X}$ e escalares $\alpha_1, \dots, \alpha_k$ *não todos nulos* tais que

$$\alpha_1 V_1 + \dots + \alpha_k V_k = \bar{0}.$$

Neste caso dizemos que os elementos de \mathcal{X} são **linearmente dependentes (L.D.)**. Se o subconjunto \mathcal{X} não é linearmente dependente, dizemos que ele é **linearmente independente (L.I.)**.

Proposição 1.6. *Um subconjunto \mathcal{X} não vazio de um espaço vetorial \mathbb{V} é linearmente independente se, e somente se, qualquer subconjunto finito de \mathcal{X} é linearmente independente.*

Demonstração. Pela definição dada acima, um subconjunto \mathcal{X} de um espaço vetorial \mathbb{V} é L.D. se, e somente se, existe um subconjunto finito de \mathcal{X} , V_1, \dots, V_k que é L.D. Portanto, \mathcal{X} é L.I. se, e somente se, ele não possui subconjunto finito L.D., ou seja, todo subconjunto finito de \mathcal{X} é L.I. ■

Exemplo 1.34. Um conjunto que contém o vetor nulo é L.D., pois o subconjunto $\{V_1 = \bar{0}\}$ é tal que $\alpha_1 V_1 = \bar{0}$, para todo escalar α_1 , em particular para $\alpha_1 \neq 0$.

Exemplo 1.35. Um conjunto formado por um único vetor, $\{V_1\}$, **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 1.36. Um conjunto formado por dois vetores, $\{V_1, V_2\}$ 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 $S = \{V_1, V_2\}$, 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 1.37. Um conjunto formado por três vetores, $\{V_1, V_2, V_3\}$ 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

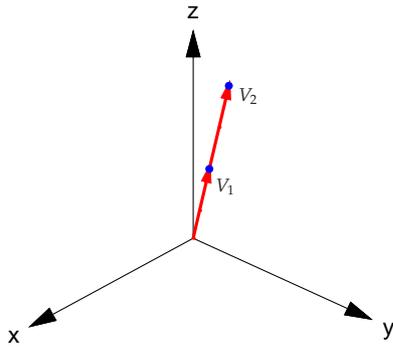


Figura 1.10: Dois vetores linearmente dependentes

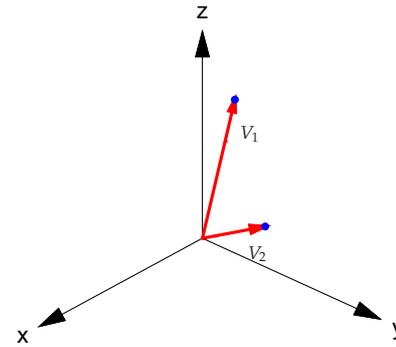


Figura 1.11: Dois vetores linearmente independentes

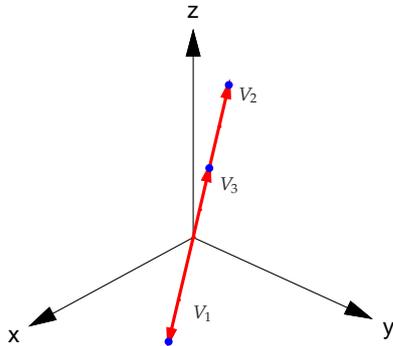


Figura 1.12: Três vetores linearmente dependentes (paralelos)

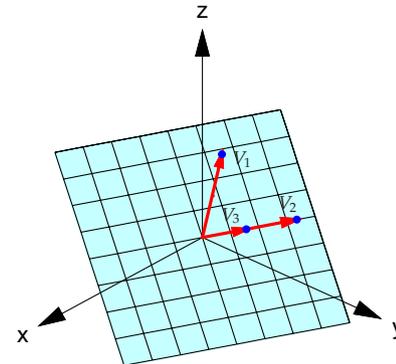


Figura 1.13: Três vetores linearmente dependentes (dois paralelos)

$x_3 \neq 0$, então V_3 é combinação linear de V_1 e V_2 . Assim, se três vetores V_1 , V_2 e V_3 de um espaço vetorial \mathbb{V} 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 1.12), ou dois deles são paralelos (Figura 1.13) ou os três são coplanares, isto é, são paralelos a um mesmo plano (Figura 1.14).

Reciprocamente, se um vetor é uma combinação linear dos outros dois, digamos se $V_1 = \alpha V_2 + \beta V_3$, então $1 V_1 - \alpha V_2 - \beta V_3 = \vec{0}$ 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 1.15).

Exemplo 1.38. 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 tem sempre solução não trivial.

Exemplo 1.39. 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

$$x_1(1, 2, 5) + x_2(7, -1, 5) + x_3(1, -1, -1) = \vec{0}. \quad (1.14)$$

Esta equação vetorial é equivalente ao sistema linear

$$\begin{cases} x_1 + 7x_2 + x_3 = 0 \\ 2x_1 - x_2 - x_3 = 0 \\ 5x_1 + 5x_2 - x_3 = 0 \end{cases} .$$

Escalonando a matriz aumentada deste sistema

$$\left[\begin{array}{ccc|c} 1 & 7 & 1 & 0 \\ 2 & -1 & -1 & 0 \\ 5 & 5 & -1 & 0 \end{array} \right] \text{ obbtemos a matriz } \left[\begin{array}{ccc|c} 1 & 0 & -2/5 & 0 \\ 0 & 1 & 1/5 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Assim a variável x_3 pode ser uma variável livre que pode, portanto, assumir qualquer valor. Concluímos que o sistema acima e a equação vetorial (1.14) têm solução não trivial. Portanto, V_1 , V_2 e V_3 são L.D.

Exemplo 1.40. Considere os polinômios

$$p_1(t) = 1 + t^2, p_2(t) = t + t^2 \text{ e } p_3(t) = 1 + t + t^2.$$

Para sabermos se eles são L.I. ou L.D. escrevemos a equação

$$x_1(1 + t^2) + x_2(t + t^2) + x_3(1 + t + t^2) = 0 \text{ para todo } t \in \mathbb{R} \quad (1.15)$$

Agrupando os termos de mesmo grau obtemos

$$(x_1 + x_3)(1) + (x_2 + x_3)t + (x_1 + x_2 + x_3)t^2 = 0 \text{ para todo } t \in \mathbb{R}.$$

Como um polinômio é identicamente nulo se, e somente se, todos os seus coeficientes são iguais a zero, a equação (1.15) é equivalente ao sistema linear

$$\begin{cases} x_1 & + & x_3 & = & 0 \\ & x_2 & + & x_3 & = & 0 \\ x_1 & + & x_2 & + & x_3 & = & 0 \end{cases},$$

Escalonando a matriz aumentada deste sistema

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array} \right] \text{ obtemos a matriz } \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right].$$

Concluímos, então que o sistema acima e a equação vetorial (1.15) possui somente a solução trivial $x_1 = x_2 = x_3 = 0$. Portanto os polinômios p_1 , p_2 e p_3 são L.I.

Exemplo 1.41. Vamos mostrar que as matrizes $M_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, $M_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ e

$M_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ são linearmente independentes no espaço das matrizes 2×2 . A

equação matricial

$$x_1M_1 + x_2M_2 + x_3M_3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

é equivalente ao sistema de equações lineares

$$\begin{cases} x_1 & & + & x_3 & = & 0 \\ x_1 & + & x_2 & & = & 0 \\ x_1 & + & x_2 & & = & 0 \\ & & x_2 & + & x_3 & = & 0 \end{cases},$$

que tem somente a solução trivial $x_1 = x_2 = x_3 = 0$. Portanto, M_1, M_2 e M_3 são linearmente independentes.

Exemplo 1.42. Vamos mostrar que os vetores

$$E_1 = (1, 0, \dots, 0), E_2 = (0, 1, 0, \dots, 0), \dots, E_n = (0, \dots, 0, 1)$$

são L.I. A equação

$$x_1E_1 + \dots + x_nE_n = \vec{0}$$

pode ser escrita como

$$x_1(1, 0, \dots, 0) + \dots + x_n(0, \dots, 0, 1) = (0, \dots, 0)$$

ou

$$(x_1, \dots, x_n) = (0, \dots, 0),$$

que é equivalente ao sistema

$$x_1 = 0, \dots, x_n = 0.$$

Em particular os vetores $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$ e $\vec{k} = (0, 0, 1)$ são L.I.

Exemplo 1.43. Para $i = 1, \dots, m$ e $j = 1, \dots, n$ seja E_{ij} a matriz $m \times n$ cujo elemento na posição ij é igual a 1 e os demais elementos são iguais a zero. As matrizes E_{ij} são linearmente independentes, pois a equação

$$x_{11}E_{11} + \dots + x_{1n}E_{1n} + \dots + x_{m1}E_{m1} + \dots + x_{mn}E_{mn} = \vec{0}.$$

tem somente a solução trivial $x_{11} = \dots = x_{1n} = \dots = x_{m1} = \dots = x_{mn} = 0$. Em particular as matrizes

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \text{ e } M_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

são L.I.

Exemplo 1.44. Vamos mostrar que o conjunto $\mathcal{X}_n = \{1, t, t^2, \dots, t^n\}$ é um conjunto linearmente independente no espaço dos polinômios \mathcal{P} . Vamos considerar a equação vetorial

$$x_0(1) + x_1(t) + \dots + x_n(t^n) = 0, \quad \forall t \in \mathbb{R}.$$

Do lado esquerdo temos um polinômio cujos coeficientes são os escalares x_0, \dots, x_n e do lado direito temos o polinômio nulo. Isto implica que os coeficientes do polinômio do lado esquerdo são todos iguais a zero. Assim, todos os escalares são iguais a zero e a equação vetorial acima tem somente a solução trivial. Portanto o conjunto \mathcal{X}_n é L.I.

Exemplo 1.45. O conjunto $\mathcal{X} = \{1, t, t^2, \dots, t^n, \dots\}$ é um conjunto linearmente independente no espaço dos polinômios \mathcal{P} . Já mostramos no [Exemplo 1.44](#) que o conjunto $\mathcal{X}_n = \{1, t, t^2, \dots, t^n\}$ é L.I. Pela [Proposição 1.6 na página 44](#) o conjunto \mathcal{X} é L.I. pois todo subconjunto finito de \mathcal{X} é um subconjunto de algum \mathcal{X}_n que é L.I.

Exemplo 1.46. Seja \mathcal{Y} um conjunto de polinômios $\{p_0, p_1, \dots, p_n, \dots\}$ tais que o grau de p_n é n , para $n = 0, 1, \dots$. Vamos mostrar que \mathcal{Y} é um conjunto linearmente

é linearmente independente. Considere a combinação linear nula

$$\sum_{i=1}^k \sum_{j=1}^{n_i} \alpha_{ij} p_{ij}(t) + \sum_{i=1}^l \sum_{j=1}^{m_i} \beta_{ij} P_{ij}(t) + \sum_{i=1}^l \sum_{j=1}^{m_i} \gamma_{ij} t P_{ij}(t) = 0.$$

Isolando o termo $\alpha_{1n_1} p_{1n_1}(t)$ obtemos

$$\alpha_{1n_1} p_{1n_1}(t) = - \sum_{j=1}^{n_1-1} \alpha_{1j} p_{1j}(t) - \sum_{i=2}^k \sum_{j=1}^{n_i} \alpha_{ij} p_{ij}(t) - \sum_{i=1}^l \sum_{j=1}^{m_i} (\beta_{ij} + \gamma_{ij} t) P_{ij}(t).$$

O polinômio $t - a_1$ é um fator do segundo membro da equação acima, mas não é de $p_{1n_1}(t)$ o que implica que $\alpha_{1n_1} = 0$ e assim o segundo membro é igual ao polinômio nulo. Se $n_1 > 1$, então dividindo-se o segundo membro por $t - a_1$ podemos isolar o termo $\alpha_{1(n_1-1)} p_{1(n_1-1)}(t) / (t - a_1)$ e pelo mesmo argumento anterior mostramos que $\alpha_{1(n_1-1)} = 0$. Repetindo-se isso obtemos que $\alpha_{ij} = 0$, para $j = 1, \dots, n_j$ e $i = 1, \dots, k$. A combinação linear inicial agora tornou-se

$$\sum_{i=1}^l \sum_{j=1}^{m_i} (\beta_{ij} + \gamma_{ij} t) P_{ij}(t) = 0.$$

Isolando o termo $(\beta_{1m_1} + \gamma_{1m_1} t) P_{1m_1}(t)$ obtemos

$$(\beta_{1m_1} + \gamma_{1m_1} t) P_{1m_1}(t) = - \sum_{j=1}^{m_1-1} (\beta_{1j} + \gamma_{1j} t) P_{1j}(t) - \sum_{i=2}^l \sum_{j=1}^{m_i} (\beta_{ij} + \gamma_{ij} t) P_{ij}(t).$$

O polinômio $t^2 + b_1 t + c_1$ é um fator do segundo membro da equação acima, mas não é de $P_{1m_1}(t)$ o que implica que $\beta_{1m_1} + \gamma_{1m_1} t$ é o polinômio nulo e assim o segundo membro é igual ao polinômio nulo. Se $m_1 > 1$, então dividindo-se o segundo membro por $q_1(t)$ podemos isolar o termo $(\beta_{1(m_1-1)} + \gamma_{1(m_1-1)} t) P_{1(m_1-1)}(t) / q_1(t)$ e pelo mesmo argumento anterior mostramos que $\beta_{1(m_1-1)} + \gamma_{1(m_1-1)} t$ é o polinômio nulo. Repetindo-se isso obtemos que $\beta_{ij} = \gamma_{ij} = 0$, para $j = 1, \dots, m_i$ e $i = 1, \dots, l$. Vejamos alguns exemplos:

(a) Se $g(t) = (t - a)(t - b)(t - c)$, com $a, b, c \in \mathbb{R}$ distintos, então os polinômios

$$p_1(t) = (t - b)(t - c), \quad p_2(t) = (t - a)(t - c), \quad p_3(t) = (t - a)(t - b)$$

são L.I.

(b) Se $g(t) = (t - a)^2(t - b)^3$, com $a, b \in \mathbb{R}$ distintos, então os polinômios

$$\begin{aligned} p_1(t) &= (t - a)(t - b)^3, & p_2(t) &= (t - b)^3, & p_3(t) &= (t - a)^2(t - b)^2, \\ p_4(t) &= (t - a)^2(t - b), & p_5(t) &= (t - a)^2 \end{aligned}$$

são L.I.

(c) Se $g(t) = (t - a)^2(t^2 + bt + c)$, com $a, b, c \in \mathbb{R}$ e $b^2 - 4c < 0$, então os polinômios

$$p_1(t) = (t - a)(t^2 + bt + c), \quad p_2(t) = t^2 + bt + c, \quad p_3(t) = (t - a)^2, \quad p_4(t) = t(t - a)^2$$

são L.I.

Teorema 1.7. *Um subconjunto \mathcal{X} não vazio de um espaço vetorial \mathbb{V} é linearmente dependente se, e somente se, $\mathcal{X} = \{\vec{0}\}$ ou existem vetores distintos V, V_1, \dots, V_k em \mathcal{X} tais que V é combinação linear de V_1, \dots, V_k .*

Demonstração. Seja $\mathcal{X} \neq \{\vec{0}\}$.

(a) Se \mathcal{X} é L.D., então existem vetores V, V_1, \dots, V_k em \mathcal{X} tais que a equação

$$x_0V + x_1V_1 + x_2V_2 + \dots + x_kV_k = \vec{0} \quad (1.16)$$

admite solução não trivial, o que significa que pelo menos um x_j é diferente de zero. Podemos supor sem perda de generalidade que $x_0 \neq 0$. Então, multiplicando-se a equação (1.16) por $1/x_0$ e subtraindo-se $(\frac{x_1}{x_0})V_1 + \dots + (\frac{x_k}{x_0})V_k$ obtemos

$$V = -\left(\frac{x_1}{x_0}\right)V_1 - \dots - \left(\frac{x_k}{x_0}\right)V_k.$$

Portanto, o vetor V é combinação linear de V_1, \dots, V_k .

- (b) Se existem vetores V, V_1, \dots, V_k em \mathcal{X} tais que V é uma combinação linear de V_1, \dots, V_k , isto é, se existem escalares $\alpha_1, \dots, \alpha_k$ tais que

$$\alpha_1 V_1 + \dots + \alpha_k V_k = V,$$

então somando-se $-V$ a ambos os membros ficamos com

$$-V + \alpha_1 V_1 + \dots + \alpha_k V_k = \bar{0}. \quad (1.17)$$

Isto implica que a equação $x_0 V + x_1 V_1 + \dots + x_k V_k = \bar{0}$ admite solução não trivial, pois o coeficiente de V em (1.17) é -1 . Portanto, \mathcal{X} é L.D. ■

Exemplo 1.48. Considere as funções $1, \cos 2t, \cos^2 t, \cos^3 t, \dots, \cos^n t, \dots$. A função $\cos 2t$ é combinação linear de $\cos^2 t$ e 1 , pois

$$\cos 2t = \cos^2 t - \sin^2 t = 2 \cos^2 t - 1,$$

o que implica, pelo Teorema 1.7, que elas são linearmente dependentes.

Exemplo 1.49. Sejam $V_1 = (1, 2, 5)$, $V_2 = (7, -1, 5)$ e $V_3 = (1, -1, -1)$ vetores do \mathbb{R}^3 . Vimos no Exemplo 1.39 na página 47 que estes vetores são L.D. Vamos escrever

um dos vetores como combinação linear dos outros dois. Já vimos que a equação vetorial

$$x_1 V_1 + x_2 V_2 + x_3 V_3 = \vec{0}$$

é equivalente ao sistema linear

$$\begin{cases} x_1 + 7x_2 + x_3 = 0 \\ 2x_1 - x_2 - x_3 = 0 \\ 5x_1 + 5x_2 - x_3 = 0 \end{cases} .$$

Escalonando a matriz aumentada deste sistema

$$\left[\begin{array}{ccc|c} 1 & 7 & 1 & 0 \\ 2 & -1 & -1 & 0 \\ 5 & 5 & -1 & 0 \end{array} \right] \text{ obbtemos a matriz } \left[\begin{array}{ccc|c} 1 & 0 & -2/5 & 0 \\ 0 & 1 & 1/5 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Assim a solução da equação vetorial acima é $x_1 = (2/5)\alpha$, $x_2 = -(1/5)\alpha$ 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

$$(2/5)\alpha V_1 - (1/5)\alpha V_2 + \alpha V_3 = \vec{0}$$

Tomando-se $\alpha = 1$, obtemos

$$(2/5)V_1 - (1/5)V_2 + V_3 = \vec{0}$$

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 1.50. 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 1.13 na página 46).

1.3.1 Independência Linear de Funções

Exemplo 1.51. Dados números reais $x_1 < x_2 < \dots < x_n$, igualmente espaçados, isto é, tais que $x_{k+1} - x_k = (x_n - x_1)/(n - 1)$, para $k = 1, \dots, n - 1$. Defina os splines (Exemplo 1.18 na página 21), $q_k(x) = 0$, se x está fora do intervalo $[x_{k-3}, x_{k+1}]$ e

$$q_k(x) = \begin{cases} p_1(t), & \text{em que } t = (x - x_{k-3})/h & \text{se } x_{k-3} \leq x < x_{k-2}, \\ p_2(t), & \text{em que } t = (x - x_{k-2})/h & \text{se } x_{k-2} \leq x < x_{k-1}, \\ p_2(1-t), & \text{em que } t = (x - x_{k-1})/h & \text{se } x_{k-1} \leq x < x_k, \\ p_1(1-t), & \text{em que } t = (x - x_k)/h & \text{se } x_k \leq x \leq x_{k+1}, \end{cases}$$

para $k = 1, \dots, n + 2$, em que

$$\begin{aligned} p_1(s) &= \frac{1}{4} s^3, \\ p_2(s) &= 1 - \frac{3}{4} (1+s)(1-s)^2 \end{aligned}$$

e $h = x_{k+1} - x_k = (x_n - x_1)/(n - 1)$. Vamos mostrar que o conjunto de splines $\{q_k \mid k = 1, \dots, n + 2\}$ é linearmente independente.

Vamos considerar a combinação linear nula dos splines q_k

$$\sum_{k=1}^{n+2} \alpha_k q_k(x) = 0. \quad (1.18)$$

Em cada intervalo $[x_k, x_{k+1}]$ somente as funções q_k, q_{k+1}, q_{k+2} e q_{k+3} podem ser diferentes de zero, e são dadas neste intervalo por

$$\begin{aligned} q_k(x) &= p_1(1-t), \\ q_{k+1}(x) &= p_2(1-t), \\ q_{k+2}(x) &= p_2(t), \\ q_{k+3}(x) &= p_1(t) \end{aligned}$$

em que $t = (x - x_k)/h$ com $h = x_{k+1} - x_k = (x_n - x_1)/(n - 1)$.

Derivando a equação (1.18) e substituindo nos pontos $x = x_1$ e $x = x_2$ obtemos as equações

$$\begin{cases} -\frac{3}{4}\alpha_1 & + \frac{3}{4}\alpha_3 & = 0 \\ & - \frac{3}{4}\alpha_2 & + \frac{3}{4}\alpha_4 & = 0 \end{cases}$$

Juntando com as equações correspondentes aos pontos x_3, \dots, x_n obtemos n equações que dão que

$$\alpha_1 = \alpha_3 = \dots = \alpha_{2k+1}, \quad \text{para } k = 1, \dots$$

$$\alpha_2 = \alpha_4 = \dots = \alpha_{2k}, \quad \text{para } k = 1, \dots$$

Substituindo $x = x_1$ e $x = x_2$ na equação (1.18) obtemos as equações

$$\begin{cases} \frac{1}{4}\alpha_1 + \alpha_2 + \frac{1}{4}\alpha_3 & = 0 \\ \frac{1}{4}\alpha_2 + \alpha_3 + \frac{1}{4}\alpha_4 & = 0 \end{cases}$$

Como $\alpha_3 = \alpha_1$ e $\alpha_4 = \alpha_2$, obtemos as equações

$$\begin{cases} \frac{1}{2}\alpha_1 + \alpha_2 & = 0 \\ \alpha_1 + \frac{1}{2}\alpha_2 & = 0 \end{cases}$$

o que dá que $\alpha_1 = \alpha_2 = \dots = \alpha_{n+2} = 0$. Portanto as funções q_k , para $k = 1, \dots, n + 2$, são L.I.

Provaremos a seguir que se as funções $f_1, \dots, f_n \in \mathcal{C}^{(n-1)}(\mathcal{I})$ forem tais que

$$W[f_1, \dots, f_n](t_0) = \det \begin{bmatrix} f_1(t_0) & f_2(t_0) & \cdots & f_n(t_0) \\ f_1'(t_0) & f_2'(t_0) & \cdots & f_n'(t_0) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(t_0) & f_2^{(n-1)}(t_0) & \cdots & f_n^{(n-1)}(t_0) \end{bmatrix} \neq 0$$

para algum ponto $t_0 \in \mathcal{I}$, então as funções f_1, \dots, f_n são linearmente independentes.

Proposição 1.8. *Sejam f_1, \dots, f_n funções de $\mathcal{C}^{(n-1)}(\mathcal{I})$, ou seja, funções que possuem a derivada de ordem $n - 1$ contínua no intervalo \mathcal{I} . Defina a função $W[f_1, \dots, f_n](t)$ no intervalo \mathcal{I} por*

$$W[f_1, \dots, f_n](t) = \det \begin{bmatrix} f_1(t) & f_2(t) & \cdots & f_n(t) \\ f_1'(t) & f_2'(t) & \cdots & f_n'(t) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(t) & f_2^{(n-1)}(t) & \cdots & f_n^{(n-1)}(t) \end{bmatrix},$$

chamada de **Wronskiano** de f_1, \dots, f_n . Se existe um ponto $t_0 \in \mathcal{I}$ tal que $W[f_1, \dots, f_n](t_0) \neq 0$, então f_1, \dots, f_n são linearmente independentes.

Demonstração. Sejam f_1, \dots, f_n funções de $\mathcal{C}^{(n-1)}(\mathcal{I})$. Vamos considerar a equação funcional

$$x_1 f_1(t) + \dots + x_n f_n(t) = \bar{0}(t) = 0, \text{ para todo } t \in \mathcal{I}. \quad (1.19)$$

Derivando a equação (1.19) $1, \dots, n - 1$ vezes obtemos o sistema

$$\begin{cases} x_1 f_1(t) + x_2 f_2(t) + \dots + x_n f_n(t) = 0 \\ x_1 f_1'(t) + x_2 f_2'(t) + \dots + x_n f_n'(t) = 0 \\ \vdots \\ x_1 f_1^{(n-1)}(t) + x_2 f_2^{(n-1)}(t) + \dots + x_n f_n^{(n-1)}(t) = 0 \end{cases}$$

que pode ser escrito como $A_t X = \bar{0}$, onde

$$A_t = \begin{bmatrix} f_1(t) & f_2(t) & \cdots & f_n(t) \\ f_1'(t) & f_2'(t) & \cdots & f_n'(t) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(t) & f_2^{(n-1)}(t) & \cdots & f_n^{(n-1)}(t) \end{bmatrix} \quad \text{e} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Vamos supor que para algum $t_0 \in \mathcal{I}$,

$$W[f_1, \dots, f_n](t_0) = \det(A_{t_0}) \neq 0.$$

Isto implica que a matriz do sistema é invertível, de onde segue-se que para t_0 o sistema $A_{t_0} X = \vec{0}$ só admite a solução trivial. Logo a equação (1.19) só admite a solução trivial e assim f_1, \dots, f_n são L.I. ■

Exemplo 1.52. As funções $f_1(t) = e^t$, $f_2(t) = e^{2t}$ e $f_3(t) = e^{3t}$ são L.I., pois pela [Proposição 1.8](#), temos que

$$W[f_1, f_2, f_3](t) = \det \begin{bmatrix} e^t & e^{2t} & e^{3t} \\ e^t & 2e^{2t} & 3e^{3t} \\ e^t & 4e^{2t} & 9e^{3t} \end{bmatrix} \quad \text{e} \quad W[f_1, f_2, f_3](0) = \det \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix} = 2 \neq 0.$$

Exemplo 1.53. As funções $f_1(t) = \cos t$, $f_2(t) = \sin t$, $f_3(t) = \cos 2t$ são L.I., pois pela [Proposição 1.8](#), temos que

$$W[f_1, f_2, f_3](t) = \det \begin{bmatrix} \cos t & \sin t & \cos 2t \\ -\sin t & \cos t & -2 \sin 2t \\ -\cos t & -\sin t & -4 \cos 2t \end{bmatrix} \quad \text{e}$$

$$W[f_1, f_2, f_3](0) = \det \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & -4 \end{bmatrix} = -3 \neq 0.$$

A recíproca da **Proposição 1.8** não é verdadeira, ou seja, mesmo que para todo t tenhamos $W(t) = 0$, não significa que as soluções dos sistemas $A_t X = \bar{0}$ na demonstração da **Proposição 1.8** sejam as mesmas para todo t , ou seja, não significa que as funções sejam linearmente dependentes. Vejamos o próximo exemplo.

Exemplo 1.54. Sejam $f_1(t) = t^2$ e $f_2(t) = t|t| = \begin{cases} t^2 & \text{se } t \geq 0 \\ -t^2 & \text{se } t < 0 \end{cases}$.

$$W[f_1, f_2](t) = \det \begin{bmatrix} t^2 & t|t| \\ 2t & 2|t| \end{bmatrix} = 0.$$

Apesar do Wronskiano ser zero para todo $t \in \mathbb{R}$ as funções f_1 e f_2 são L.I., pois uma função não é múltiplo escalar da outra. Para $t \geq 0$, $f_2(t) = f_1(t)$ e para $t < 0$, $f_2(t) = -f_1(t)$.

Exercícios Numéricos (respostas na página 406)

1.3.1. Quais dos seguintes conjuntos de vetores são linearmente dependentes?

(a) $\{(1, 1, 2, 1), (1, 0, 0, 2), (4, 6, 8, 6), (0, 3, 2, 1)\}$;

(b) $\{(1, -2, 3, -1), (-2, 4, -6, 2)\}$;

(c) $\{(1, 1, 1, 1), (2, 3, 1, 2), (3, 1, 2, 1), (2, 2, 1, 1)\}$;

(d) $\{(4, 2, -1, 3), (6, 5, -5, 1), (2, -1, 3, 5)\}$.

1.3.2. Para quais valores de a o conjunto de vetores $\{(3, 1, 0), (a^2 + 2, 2, 0)\}$ é L.D.?

1.3.3. Verifique se os polinômios seguintes são linearmente dependentes ou independentes.

(a) $t^2 - 2t + 3, 2t^2 + t + 8, t^2 + 8t + 7$

(b) $t^2 - 1, t + 1, t + 2$

1.3.4. Verifique se as funções seguintes são linearmente dependentes ou independentes.

(a) $t, \cos t, \sin t$ em $C^2[-\pi, \pi]$.

(b) $\cos t, 1, \sin^2(t/2)$ em $C^2[-\pi, \pi]$.

(c) $1, e^t + e^{-t}, e^t - e^{-t}$ em $C^2[-1, 1]$.

1.3.5. Verifique que as funções $e^t \cos 3t$ e $e^t \sin 3t$ são soluções linearmente independentes da equação diferencial $y'' - 2y' + 10y = 0$.

1.3.6. Suponha que $\mathcal{S} = \{X_1, X_2, X_3\}$ seja um conjunto linearmente independente de vetores de um espaço vetorial \mathbb{V} . Responda se $\mathcal{T} = \{Y_1, Y_2, Y_3\}$ é linearmente dependente ou independente nos seguintes casos:

(a) $Y_1 = X_1 + X_2, Y_2 = X_1 + X_3$ e $Y_3 = X_2 + X_3$;

(b) $Y_1 = X_1, Y_2 = X_1 + X_3$ e $Y_3 = X_1 + X_2 + X_3$.

- 1.3.7.** Para $k = 0, 1, \dots, n$, seja $p_k(t) = t^k + t^{k+1} + \dots + t^n$. Mostre que o conjunto $\{p_0(t), p_1(t), \dots, p_n(t)\}$ é linearmente independente em $\mathcal{P} = \mathbb{R}[t]$.
- 1.3.8.** Mostre que as funções $f_1(t) = te^{\lambda t}, f_2(t) = t^2e^{\lambda t}, \dots, f_k(t) = t^ke^{\lambda t}$, onde $\lambda \in \mathbb{R}$, são linearmente independentes.
- 1.3.9.** Mostre que as funções $f_1(t) = e^{\lambda_1 t}, f_2(t) = e^{\lambda_2 t}, \dots, f_k(t) = e^{\lambda_k t}$, onde $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, são linearmente independentes se, e somente se, $\lambda_i \neq \lambda_j$, para $i \neq j, i, j = 1, \dots, k$.
- 1.3.10.** Suponha que $\{X_1, X_2, \dots, X_n\}$ seja um conjunto de vetores do \mathbb{R}^n linearmente independente. Mostre que se A é uma matriz $n \times n$ não singular, então $\{AX_1, AX_2, \dots, AX_n\}$ também é um conjunto linearmente independente.
- 1.3.11.** Mostre que $\{2, t + 1, t^2 + 1, \dots, t^n + 1, \dots\}$ é um conjunto de polinômios linearmente independente.

Exercícios usando o MATLAB[®]

Comandos do MATLAB[®]:

>> clf limpa a janela gráfica.

>> hold on segura o gráfico atual de forma que o próximo gráfico será desenhado por cima do atual.

>> subs(expr, t, a) substitui na expressão expr a variável t por a.

>> axis([a,b,c,d]) define que será mostrada uma janela no plano cartesiano que vai de $x = a$ a $x = b$ e de $y = c$ a $y = d$.

Comandos do pacote GAAL:

>> plotf1(f, [a,b]) desenha o gráfico da função $f(x)$, em que $f(x)$ é dada pela expressão f, no intervalo [a,b].

>> `qk=spline1(k,x,nbp,a,b)` calcula o spline $q_k(x) = 0$, se x está fora do intervalo $[x_{k-3}, x_{k+1}]$ e

$$q_k(x) = \begin{cases} p_1(t), & \text{em que } t = (x - x_{k-3})/h, & \text{se } x_{k-3} \leq x < x_{k-2}, \\ p_2(t), & \text{em que } t = (x - x_{k-2})/h, & \text{se } x_{k-2} \leq x < x_{k-1}, \\ p_2(1-t), & \text{em que } t = (x - x_{k-1})/h, & \text{se } x_{k-1} \leq x < x_k, \\ p_1(1-t), & \text{em que } t = (x - x_k)/h, & \text{se } x_k \leq x \leq x_{k+1}, \end{cases}$$

para $k = 1, \dots, n+2$, em que

$$p_1(s) = \frac{1}{4}s^3,$$

$$p_2(s) = 1 - \frac{3}{4}(1+s)(1-s)^2$$

e $h = x_{k+1} - x_k = (x_n - x_1)/(n-1)$ no ponto x para um intervalo $[a, b]$ dividido em $\text{nbp}-1$ subintervalos e >> `qk=spline1(k,X,nbp,a,b)` calcula o spline q_k nos pontos dados pelas componentes do vetor coluna X .

>> `A=spline1(X,nbp,a,b)` cria a matriz $a_{ij} = q_j(X_i)$, em que X é um vetor coluna, para um intervalo $[a, b]$ dividido em $\text{nbp}-1$ subintervalos.

>> `plotspline1(C,nbp,a,b)` desenha o spline definido por $\sum_{k=1}^{\text{nbp}+2} c_k q_k(x)$.

- 1.3.12.** (a) Defina os vetores $V_1 = [1; 2; 3]$, $V_2 = [3; 4; 5]$ e $V_3 = [5; 6; 7]$. Defina o vetor $V = \text{randi}(3, 1)$. Verifique se V é combinação linear de V_1 , V_2 e V_3 .
- (b) Defina $M = \text{randi}(3, 5)$. Verifique se os vetores definidos pelas colunas de M são combinação linear de V_1 , V_2 e V_3 . Tente explicar o resultado.
- (c) Verifique se V_1 , V_2 e V_3 são linearmente independentes. Se eles forem linearmente dependentes, escreva um deles como combinação linear dos outros e verifique o resultado.

- 1.3.13. (a) Declare a variável t simbólica com o comando `>> syms t`. Defina as variáveis $p_1=t^3/4$ e $p_2=1-3*(1+t)*(1-t)^2/4$. Desenhe os gráficos das funções $p_1(t) = \frac{1}{4}t^3$ e $p_2(t) = 1 - \frac{3}{4}(1+t)(1-t)^2$ no intervalo $[0, 1]$ com os comandos

```
>> clf
>> plotf1(p1, [0, 1])
>> hold on
>> plotf1(p2, [0, 1])
```

- (b) Observe que $p_1(1) = p_2(0)$, $p_1'(1) = p_2'(0)$ e $p_1''(1) = p_2''(0)$. Verifique estas igualdades analiticamente. Usando p_1 e uma translação de p_2 defina um spline no intervalo $[0, 2]$. Use os comandos abaixo para desenhar o novo spline.

```
>> clf
>> plotf1(p1, [0, 1])
>> hold on
>> plotf1(subs(p2, t, t-1), [1, 2])
>> axis([0, 2, 0, 1])
```

- (c) Observe que colando o spline anterior com uma translação da reflexão deste spline em relação ao eixo y obtemos um novo spline no intervalo $[0, 4]$. Use os comandos abaixo para desenhar o novo spline.

```
>> plotf1(subs(subs(p2, t, 1-t), t, t-2), [2, 3])
>> plotf1(subs(subs(p1, t, 1-t), t, t-3), [3, 4])
>> axis([0, 4, 0, 1])
```

- (d) Considere os splines $q_k(x) = 0$, se x está fora do intervalo $[x_{k-3}, x_{k+1}]$ e

$$q_k(x) = \begin{cases} p_1(t), & \text{em que } t = (x - x_{k-3})/h, & \text{se } x_{k-3} \leq x < x_{k-2}, \\ p_2(t), & \text{em que } t = (x - x_{k-2})/h, & \text{se } x_{k-2} \leq x < x_{k-1}, \\ p_2(1-t), & \text{em que } t = (x - x_{k-1})/h, & \text{se } x_{k-1} \leq x < x_k, \\ p_1(1-t), & \text{em que } t = (x - x_k)/h, & \text{se } x_k \leq x \leq x_{k+1}, \end{cases}$$

para $k = 1, \dots, n + 2$, em que $p_1(s) = \frac{1}{4}s^3$, $p_2(s) = 1 - \frac{3}{4}(1+s)(1-s)^2$ e $h = x_{k+1} - x_k = (x_n - x_1)/(n - 1)$. Use os comandos abaixo para desenhar o spline q_1 , definidos no intervalo $[0, 1]$ com pontos de quebra $x_1 = 0, x_2 = 1/3, x_3 = 2/3$ e $x_4 = 1$.

```
>> I=eye(6);
>> plotspline1(I(:,1),4,0,1)
```

Substitua k no comando abaixo por 2, 3, 4, 5 e 6 para desenhar os splines q_2, q_3, q_4, q_5 e q_6 .

```
>> plotspline1(I(:,k),4,0,1)
```

(e) Mostre que os splines q_1, q_2, q_3, q_4, q_5 e q_6 do item anterior são linearmente independentes usando os seguintes comandos.

```
>> X=[0,1/5,2/5,3/5,4/5,1];
>> A=spline1(X,4,0,1);
>> det(A)
```

Exercícios Teóricos

- 1.3.14.** Sejam V_1, \dots, V_{k+1} vetores de um espaço vetorial \mathbb{V} , tais que $\{V_1, \dots, V_k\}$ é linearmente independente. Mostre que se V_{k+1} não pertence ao subespaço gerado por $\{V_1, \dots, V_k\}$, então $\{V_1, \dots, V_{k+1}\}$ é linearmente independente. (Sugestão: Considere a equação $x_1V_1 + \dots + x_{k+1}V_{k+1} = \vec{0}$. Separe em dois casos: $x_{k+1} = 0$ e $x_{k+1} \neq 0$.)
- 1.3.15.** Sejam V_1, \dots, V_n vetores de um espaço vetorial \mathbb{V} . Mostre que um vetor $V \in [V_1, \dots, V_n]$ pode ser escrito de maneira única como combinação linear de V_1, \dots, V_n se, e somente se, V_1, \dots, V_n são linearmente independentes.
- 1.3.16.** Sejam \mathcal{X} e \mathcal{Y} dois subconjuntos linearmente independentes de um espaço vetorial \mathbb{V} . Mostre que $\mathcal{X} \cup \mathcal{Y}$ é linearmente independente se, e somente se, $[\mathcal{X}] \cap [\mathcal{Y}] = \{\vec{0}\}$.
- 1.3.17.** Seja \mathbb{V} um espaço vetorial.
- (a) Mostre que se $\mathcal{X}_1 \subseteq \mathcal{X}_2 \subseteq \mathbb{V}$ e \mathcal{X}_1 é linearmente dependente, então \mathcal{X}_2 é linearmente dependente.

(b) Mostre que se $\mathcal{X}_1 \subseteq \mathcal{X}_2 \subseteq \mathbb{V}$ e \mathcal{X}_2 é linearmente independente, então \mathcal{X}_1 é linearmente independente.

1.3.18. Seja \mathcal{X} um conjunto de polinômios em que quaisquer dois polinômios pertencentes a \mathcal{X} têm graus diferentes. Mostre que \mathcal{X} é linearmente independente.

1.4 Base e Dimensão

Já vimos que, em um espaço vetorial \mathbb{V} , um conjunto de geradores pode ser linearmente independente ou linearmente dependente. Se o conjunto de geradores for linearmente dependente, então existe um vetor no conjunto que é combinação linear de outros elementos do conjunto. Então este elemento não é necessário na geração do espaço \mathbb{V} . Portanto, um conjunto de geradores linearmente dependente contém vetores que não são necessários para gerar \mathbb{V} .

1.4.1 Base

Definição 1.8. Um subconjunto \mathcal{B} de um espaço vetorial \mathbb{V} é uma **base de** \mathbb{V} , se

- (a) \mathcal{B} é um conjunto de geradores de \mathbb{V} e
 - (b) \mathcal{B} é um conjunto linearmente independente.
-

Exemplo 1.55. Os vetores

$$E_1 = (1, 0, \dots, 0), E_2 = (0, 1, 0, \dots, 0), \dots, E_n = (0, \dots, 0, 1)$$

formam uma base do \mathbb{R}^n . Pois, vimos no [Exemplo 1.27 na página 33](#) que E_1, \dots, E_n geram o \mathbb{R}^n e no [Exemplo 1.42 na página 49](#) que E_1, E_2, \dots, 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 1.56. Vamos determinar uma base para o espaço solução do sistema homogêneo

$$\begin{cases} 2x_1 + 2x_2 - x_3 + x_4 = 0 \\ -x_1 - x_2 + 2x_3 + x_4 = 0 \\ x_1 + x_2 - 2x_3 - x_4 = 0 \end{cases}$$

A matriz aumentada deste sistema é

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

Resolvendo o sistema pelo método de Gauss-Jordan, transformamos a matriz aumentada na sua forma reduzida escalonada, obtendo

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Portanto, o sistema dado é equivalente ao seguinte sistema:

$$\begin{cases} x_1 + x_2 + x_4 = 0 \\ x_3 + x_4 = 0 \end{cases}$$

cuja solução é dada por

$$(x_1, x_2, x_3, x_4) = (-\alpha - \beta, \alpha, -\beta, \beta),$$

para todos os números α e β reais. Assim, o espaço solução do sistema é

$$\mathbb{V} = \{(-\alpha - \beta, \alpha, -\beta, \beta) \mid \alpha, \beta \in \mathbb{R}\}.$$

Agora, vamos determinar uma base para este subespaço. Qualquer vetor V de \mathbb{V} pode ser escrito como uma soma de vetores de \mathbb{V} , sendo um vetor para cada

parâmetro e cada vetor depende apenas de um parâmetro, obtendo

$$\begin{aligned}(-\alpha - \beta, \alpha, -\beta, \beta) &= (-\alpha, \alpha, 0, 0) + (-\beta, 0, -\beta, \beta) \\ &= \alpha(-1, 1, 0, 0) + \beta(-1, 0, -1, 1)\end{aligned}$$

Assim, $V_1 = (-1, 1, 0, 0)$ e $V_2 = (-1, 0, -1, 1)$ geram \mathbb{V} . Além disso, eles são L.I., pois se

$$\alpha(-1, 1, 0, 0) + \beta(-1, 0, -1, 1) = (-\alpha - \beta, \alpha, -\beta, \beta) = (0, 0, 0, 0),$$

então $\alpha = 0$ e $\beta = 0$. Portanto, V_1 e V_2 formam uma base de \mathbb{V} .

Exemplo 1.57. Seja $\mathbb{V} = \{(a + c, b + c, a + b + 2c) \mid a, b, c \in \mathbb{R}\}$ um subespaço de \mathbb{R}^3 . Qualquer elemento V de \mathbb{V} pode ser escrito como uma soma de vetores de \mathbb{V} , sendo um vetor para cada parâmetro e cada vetor depende 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 $V_3 = (1, 1, 2)$, então $\{V_1, V_2, V_3\}$ gera \mathbb{V} . Para sabermos se $\{V_1, V_2, V_3\}$ é base de \mathbb{V} , precisamos verificar se V_1, V_2 e V_3 são L.I. Para isto temos que saber se a equação vetorial

$$xV_1 + yV_2 + zV_3 = \vec{0} \tag{1.20}$$

só possui a solução trivial, ou equivalentemente, se o sistema $AX = \vec{0}$ só possui a solução trivial, onde $A = [V_1 \ V_2 \ V_3]$. Escalonando a matriz $[A \mid \vec{0}]$, obtemos

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

A solução de (1.20) é dada por $x = -\alpha$, $y = \alpha$ e $z = \alpha$, para todo $\alpha \in \mathbb{R}$. Substituindo-se esta solução em (1.20) obtemos

$$-\alpha V_1 + \alpha V_2 + \alpha V_3 = \vec{0}$$

Tomando-se $\alpha = 1$ e somando-se $V_1 - V_2$ obtemos $V_3 = V_2 + V_1$. 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} .

Exemplo 1.58. Para $i = 1, \dots, m$ e $j = 1, \dots, n$ seja E_{ij} a matriz $m \times n$ cujo elemento na posição ij é igual a 1 e os demais elementos são iguais a zero. As matrizes E_{ij} formam uma base para o espaço das matrizes $m \times n$, pois mostramos no **Exemplo 1.28 na página 34** que elas geram \mathcal{M}_{mn} e no **Exemplo 1.43 na página 50** que elas são linearmente independentes.

Exemplo 1.59. O conjunto $\mathcal{X} = \{1, t, t^2, \dots, t^n, \dots\}$ é uma base para o espaço $\mathcal{P} = \mathbb{R}[t]$, pois já mostramos que todo polinômio é combinação linear de elementos de \mathcal{X} (**Exemplo 1.29 na página 34**) e que \mathcal{X} é um conjunto linearmente independente (**Exemplo 1.45 na página 50**). O conjunto $\mathcal{X}_n = \{1, t, t^2, \dots, t^n\}$ é uma base para o espaço \mathcal{P}_n , pois já mostramos que todo polinômio de grau no máximo n é combinação linear de elementos de \mathcal{X}_n e além disso é um conjunto linearmente independente, pois $\mathcal{X}_n \subset \mathcal{X}$.

Teorema 1.9. *Um subconjunto \mathcal{B} de um espaço vetorial \mathbb{V} é uma base de \mathbb{V} se, e somente se, cada vetor de \mathbb{V} se escreve de maneira única como combinação linear dos vetores de \mathcal{B} .*

Demonstração. Em primeiro lugar, suponha que todo vetor V de \mathbb{V} seja escrito de maneira única como combinação linear de elementos de \mathcal{B} . Vamos mostrar que \mathcal{B} é uma base de \mathbb{V} . Como todo vetor é escrito como combinação linear de elementos de \mathcal{B} , basta mostrarmos que \mathcal{B} é L.I. Sejam $V_1, \dots, V_m \in \mathcal{B}$. Considere a equação

$$x_1 V_1 + \dots + x_m V_m = \vec{0}.$$

Como todo vetor X de \mathbb{V} é escrito de maneira única como combinação linear de elementos de \mathcal{B} , em particular temos que para $X = \vec{0}$,

$$x_1 V_1 + \dots + x_m V_m = \vec{0} = 0V_1 + \dots + 0V_m,$$

o que implica que $x_1 = 0, \dots, x_m = 0$, ou seja, V_1, \dots, V_m são linearmente independentes. Como V_1, \dots, V_m são vetores quaisquer de \mathcal{B} , então \mathcal{B} é L.I. Portanto, \mathcal{B} é base de \mathbb{V} .

Suponha, agora, que \mathcal{B} seja base de \mathbb{V} . Seja V um vetor qualquer de \mathbb{V} . Se V é escrito de duas maneiras como combinação linear de elementos de \mathcal{B} , então podemos supor, sem perda de generalidade, que existem vetores $V_1, \dots, V_m \in \mathcal{B}$ tais que

$$x_1 V_1 + \dots + x_m V_m = V = y_1 V_1 + \dots + y_m V_m,$$

então

$$(x_1 - y_1) V_1 + \dots + (x_m - y_m) V_m = \vec{0}.$$

Como \mathcal{B} é uma base, então V_1, \dots, V_m são L.I. o que implica que

$$x_1 = y_1, \dots, x_m = y_m.$$

Portanto, todo vetor V de \mathbb{V} é escrito de maneira única como combinação linear de elementos de \mathcal{B} . ■

Teorema 1.10. (a) Se $\mathcal{B} = \{V_1, \dots, V_m\}$ é uma base de um espaço vetorial \mathbb{V} , então um subconjunto de \mathbb{V} com mais de m vetores é linearmente dependente.

(b) Se $\mathcal{B} = \{V_1, \dots, V_m\}$ e $\mathcal{B}' = \{W_1, \dots, W_n\}$ são duas bases de um espaço vetorial \mathbb{V} , então $m = n$.

Demonstração. (a) Seja $\{W_1, \dots, W_n\}$ um subconjunto de \mathbb{V} , com $n > m$. Vamos mostrar que $\{W_1, \dots, W_n\}$ é L.D. Considere a combinação linear nula de W_1, \dots, W_n

$$x_1 W_1 + x_2 W_2 + \dots + x_n W_n = \vec{0}. \quad (1.21)$$

Como $\{V_1, \dots, V_m\}$ é uma base, qualquer elemento do espaço pode ser escrito como combinação linear de V_1, \dots, V_m . Em particular,

$$W_j = a_{1j} V_1 + a_{2j} V_2 + \dots + a_{mj} V_m = \sum_{i=1}^m a_{ij} V_i, \quad \text{para } j = 1, \dots, n. \quad (1.22)$$

Assim, substituindo (1.22) em (1.21) e agrupando os termos que contém V_i , para $i = 1, \dots, m$, obtemos

$$(a_{11}x_1 + \dots + a_{1n}x_n)V_1 + \dots + (a_{m1}x_1 + \dots + a_{mn}x_n)V_m = \vec{0}. \quad (1.23)$$

Como $\{V_1, \dots, V_m\}$ é base, V_1, \dots, V_m são L.I. e portanto os escalares na equação (1.23) são iguais a zero. Isto leva ao sistema linear

$$AX = \vec{0},$$

onde $A = (a_{ij})_{m \times n}$. Mas, este é um sistema homogêneo que tem mais incógnitas do que equações, portanto possui solução não trivial, como queríamos provar.

- (b) Suponha por contradição que $n > m$. Pelo item anterior, segue-se que $\{W_1, \dots, W_n\}$ é L.D., o que é impossível. O caso $n < m$ pode ser tratado de forma análoga. ■

Exemplo 1.60. Segue do Teorema 1.10 e dos exemplos anteriores que mais de n vetores no \mathbb{R}^n , mais de $n + 1$ polinômios de \mathcal{P}_n e mais de mn matrizes $m \times n$ são linearmente dependentes.

1.4.2 Dimensão

Definição 1.9. Dizemos que um espaço vetorial \mathbb{V} tem **dimensão finita** se ele tem uma base consistindo de um número finito de vetores ou $\mathbb{V} = \{\vec{0}\}$. Se $\mathbb{V} \neq \{\vec{0}\}$, o número de elementos de uma de suas bases é chamado de **dimensão** de \mathbb{V} , denotado por $\dim(\mathbb{V})$ e $\dim(\{\vec{0}\}) = 0$. Quando um espaço não tem dimensão finita, dizemos que ele tem **dimensão infinita**.

Exemplo 1.61. A dimensão do \mathbb{R}^n é n , pois como foi mostrado no Exemplo 1.55 na página 67, $E_1 = (1, 0, \dots, 0)$, $E_2 = (0, 1, 0, \dots, 0)$, \dots , $E_n = (0, \dots, 0, 1)$ formam uma base do \mathbb{R}^n .

Exemplo 1.62. A dimensão do subespaço

$$\mathbb{W} = \{(a + c, b + c, a + b + 2c) \mid a, b, c \in \mathbb{R}\} \text{ de } \mathbb{R}^3$$

é 2 pois como foi mostrado no **Exemplo 1.57 na página 69**, os vetores $V_1 = (1, 0, 1)$ e $V_2 = (0, 1, 1)$ formam uma base de \mathbb{W} .

Exemplo 1.63. O espaço \mathcal{M}_{mn} das matrizes $m \times n$ tem dimensão mn , pois foi mostrado no **Exemplo 1.58 na página 70** que as matrizes E_{ij} , que tem o elemento i, j igual a 1 e todos os outros elementos iguais a zero formam uma base de \mathcal{M}_{mn} .

Exemplo 1.64. O espaço \mathcal{P}_n tem dimensão $n + 1$, pois como foi mostrado no **Exemplo 1.59 na página 70** o conjunto $\mathcal{X} = \{1, x, \dots, x^n\}$ é uma base de \mathcal{P}_n .

Proposição 1.11. *Seja \mathbb{V} um espaço vetorial de dimensão finita n . Se \mathbb{W} é um subespaço de \mathbb{V} , então \mathbb{W} é de dimensão finita e $\dim(\mathbb{W}) \leq \dim(\mathbb{V})$.*

Demonstração. Se $\mathbb{W} = \{\vec{0}\}$, então $\dim(\mathbb{W}) = 0 \leq n$.

Caso contrário \mathbb{W} contém um vetor não nulo V_1 . Então $\{V_1\}$ é um conjunto linearmente independente. Seja \mathbb{W}_1 o subespaço gerado por V_1 . Se $\mathbb{W}_1 = \mathbb{W}$, então $\{V_1\}$ é uma base de \mathbb{W} e $\dim(\mathbb{W}) = 1 \leq n$.

Caso contrário acrescentamos vetores V_2, \dots, V_k de forma que $\{V_1, \dots, V_k\}$ seja linearmente independente. Pelo **Teorema 1.10 (a) na página 72** $k \leq n$. Se $\{V_1, \dots, V_k\}$ gera \mathbb{W} , então $\{V_1, \dots, V_k\}$ é uma base de \mathbb{W} e $\dim(\mathbb{W}) = k \leq n$.

Caso contrário, seja V_{k+1} um vetor que pertence a \mathbb{W} , mas não pertence ao subespaço gerado por $\{V_1, \dots, V_k\}$. Então, o conjunto $\{V_1, \dots, V_k, V_{k+1}\}$ é L.I., pois caso contrário $x_1 V_1 + \dots + x_{k+1} V_{k+1} = \vec{0}$, implicaria que $x_{k+1} \neq 0$ (por que?) e assim, V_{k+1} seria combinação linear de V_1, \dots, V_k , ou seja, V_{k+1} pertenceria ao subespaço \mathbb{W}_k . Se $\{V_1, \dots, V_{k+1}\}$ gera \mathbb{W} , então $\{V_1, \dots, V_{k+1}\}$ é uma base de \mathbb{W} e $\dim(\mathbb{W}) = k + 1 \leq n$ pelo **Teorema 1.10 (a) na página 72**.

Pelo **Teorema 1.10 (a) na página 72** este processo tem que parar, ou seja, existe um inteiro positivo $m \leq n$ tal que $\{V_1, \dots, V_k, V_{k+1}, \dots, V_m\}$ é L.I., mas $\{V_1, \dots, V_k, V_{k+1}, \dots, V_m, V\}$ é L.D. para qualquer vetor V de \mathbb{W} . O que implica que V é combinação linear de $\{V_1, \dots, V_k, V_{k+1}, \dots, V_m\}$ (por que?). Portanto, $\{V_1, \dots, V_k, V_{k+1}, \dots, V_m\}$ é uma base de \mathbb{W} e $\dim(\mathbb{W}) = m \leq n$. ■

Teorema 1.12. *Seja \mathbb{V} um espaço vetorial de dimensão finita $n > 0$.*

- (a) *Todo subconjunto de \mathbb{V} linearmente independente que contém n elementos é uma base de \mathbb{V} .*
 - (b) *De todo conjunto de geradores de \mathbb{V} pode ser extraída uma base de \mathbb{V} .*
 - (c) *Todo conjunto de geradores de \mathbb{V} com n elementos é uma base de \mathbb{V} .*
 - (d) *Todo conjunto de geradores de \mathbb{V} contém pelo menos n elementos.*
 - (e) *Todo subconjunto de \mathbb{V} linearmente independente pode ser estendido a uma base de \mathbb{V} .*
-

Demonstração. (a) Sejam V_1, \dots, V_n vetores L.I. e seja V um vetor qualquer do espaço \mathbb{V} . Vamos mostrar que V é combinação linear de V_1, \dots, V_n . Considere a equação vetorial

$$x_1 V_1 + x_2 V_2 + \dots + x_n V_n + x_{n+1} V = \vec{0}. \quad (1.24)$$

Pela **Proposição 1.10 (b)**, V_1, \dots, V_n, V são L.D., pois são $n + 1$ vetores em um espaço de dimensão n . Então a equação acima admite solução não trivial, ou seja, pelo menos um $x_i \neq 0$. Mas, $x_{n+1} \neq 0$, pois caso contrário, V_1, \dots, V_n seriam L.D. Então, multiplicando-se a equação (1.24) por $1/x_{n+1}$ e subtraindo $(x_1/x_{n+1})V_1 + (x_2/x_{n+1})V_2 + \dots + (x_n/x_{n+1})V_n$, obtemos

$$V = - \left(\frac{x_1}{x_{n+1}} \right) V_1 - \dots - \left(\frac{x_n}{x_{n+1}} \right) V_n.$$

- (b) Seja \mathcal{B} um conjunto de geradores de \mathbb{V} . Se \mathcal{B} é L.I., então \mathcal{B} é uma base de \mathbb{V} . Caso contrário, \mathcal{B} é L.D. e pelo **Teorema 1.7 na página 53**, um dos vetores de \mathcal{B} é combinação linear dos outros. Assim, o subconjunto de \mathcal{B} obtido retirando-se este vetor continua gerando \mathbb{V} . Se esse subconjunto for L.I., temos uma base para \mathbb{V} , caso contrário, continuamos retirando vetores do subconjunto até obtermos um subconjunto L.I. e aí neste caso temos uma base para \mathbb{V} .
- (c) Se não fosse uma base então, pelo item anterior, poderíamos extrair deste conjunto uma base com menos de n vetores o que é impossível pela **Proposição 1.10 na página 72**.
- (d) Se existisse um conjunto de geradores com menos de n elementos, então poderíamos extrair deste conjunto uma base com menos de n vetores o que é impossível pela **Proposição 1.10 na página 72**.
- (e) Seja $\mathcal{B} = \{V_1, \dots, V_k\}$ um conjunto de vetores linearmente independente. Se $k = n$, pelo item (a), não há o que fazer. Se $k < n$, então seja $\mathbb{W}_k = [V_1, \dots, V_k]$, o subespaço gerado por \mathcal{B} . Seja V_{k+1} um vetor que pertence a \mathbb{V} , mas não

pertence a \mathbb{W}_k . Então, o conjunto $\{V_1, \dots, V_k, V_{k+1}\}$ é L.I., pois caso contrário $x_1 V_1 + \dots + x_{k+1} V_{k+1} = \vec{0}$, implicaria que $x_{k+1} \neq 0$ (por que?) e assim, V_{k+1} seria combinação linear de V_1, \dots, V_k , ou seja, V_{k+1} pertenceria ao subespaço \mathbb{W}_k . Se $n = k + 1$, então pelo item (a), $\{V_1, \dots, V_k, V_{k+1}\}$ é uma base de \mathbb{V} . Caso contrário, ou seja, se $n > k + 1$, então o mesmo argumento é repetido para o subespaço $\mathbb{W}_{k+1} = [V_1, \dots, V_k, V_{k+1}]$. Este processo pode ser continuado até que um conjunto $V_1, \dots, V_k, V_{k+1}, \dots, V_n$ de vetores L.I. seja obtido. Neste caso, pelo item (a), $\{V_1, \dots, V_k, V_{k+1}, \dots, V_n\}$ é uma base de \mathbb{V} .



Exemplo 1.65. Vamos determinar uma base para o espaço \mathbb{W} gerado pelas matrizes

$$M_1 = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 3 & 3 \\ -1 & -3 \end{bmatrix}, \quad M_3 = \begin{bmatrix} -1 & 0 \\ 1 & 4 \end{bmatrix} \quad \text{e} \quad M_4 = \begin{bmatrix} 5 & 6 \\ -1 & -2 \end{bmatrix}.$$

Como por definição do subespaço, M_1, M_2, M_3 e M_4 geram \mathbb{W} , precisamos saber se elas são L.I. Para isso, precisamos resolver a equação

$$xM_1 + yM_2 + zM_3 + wM_4 = \vec{0}. \quad (1.25)$$

que é equivalente ao sistema linear

$$\begin{cases} 2x + 3y - z + 5w = 0 \\ 3x + 3y + 6w = 0 \\ -y + z - w = 0 \\ x - 3y + 4z - w = 0 \end{cases}$$

cuja solução é dada por $x = -\alpha - \beta, y = -\alpha + \beta, z = \beta$ e $w = \alpha$, para todos $\alpha, \beta \in \mathbb{R}$. Portanto, as matrizes são L.D. Substituindo-se os valores encontrados de x, y, z e w na equação (1.25) temos que

$$(-\alpha - \beta)M_1 + (-\alpha + \beta)M_2 + \beta M_3 + \alpha M_4 = \vec{0}.$$

Tomando-se $\alpha = 1$ e $\beta = 0$ obtemos que $-M_1 - M_2 + M_4 = \bar{0}$ ou somando-se $M_1 + M_2$, $M_4 = M_1 + M_2$. Agora, tomando-se $\alpha = 0$ e $\beta = 1$ obtemos que $-M_1 + M_2 + M_3 = \bar{0}$ ou somando-se $M_1 - M_2$, $M_3 = M_1 - M_2$. Assim, as matrizes M_3 e M_4 podem ser descartadas na geração do subespaço \mathbb{W} . Como as matrizes M_1 e M_2 são tais que uma não é múltiplo escalar da outra, então elas são L.I. e formam, portanto, uma base para \mathbb{W} .

Exemplo 1.66. Vamos mostrar que os polinômios $p_1(t) = t^2 + 1$, $p_2(t) = t^2 - 1$ e $p_3(t) = t + 2$ formam uma base de \mathcal{P}_2 . Como sabemos que a dimensão de \mathcal{P}_2 é 3, basta mostrarmos que estes polinômios são L.I. Para isso, precisamos resolver a equação vetorial

$$x(t^2 + 1) + y(t^2 - 1) + z(t + 2) = 0, \quad \forall t \in \mathbb{R} \quad (1.26)$$

ou agrupando os termos de mesmo grau,

$$(x + y)t^2 + (z)t + (x - y + 2z) = 0, \quad \forall t \in \mathbb{R}.$$

Como o polinômio nulo tem todos os coeficientes iguais a zero, então (1.26) é equivalente ao sistema linear

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

que tem somente a solução trivial, $x = y = z = 0$. Portanto, p_1, p_2 e p_3 são L.I. e como a dimensão de \mathcal{P}_2 é 3, pelo Teorema 1.12 (a), eles formam uma base de \mathcal{P}_2 .

Exemplo 1.67. Considere o polinômio

$$g(t) = (t - a_1)^{n_1} \dots (t - a_k)^{n_k} (t^2 + b_1 t + c_1)^{m_1} \dots (t^2 + b_l t + c_l)^{m_l} \in \mathbb{R}[t],$$

com $a_i \in \mathbb{R}$ distintos para $i = 1, \dots, k$ e $(b_i, c_i) \in \mathbb{R}^2$ distintos tais que $b_i^2 - 4c_i < 0$, para $i = 1, \dots, l$. Seja $n = n_1 + \dots + n_k + 2m_1 + \dots + 2m_l$ o grau de $g(t)$. Considere os polinômios $p_{ij}(t) = g(t)/(t - a_i)^j$, para $j = 1, \dots, n_i$ e $i = 1, \dots, k$ e $P_{ij}(t) = g(t)/(t^2 + b_it + c_i)^j$, para $j = 1, \dots, m_i$ e $i = 1, \dots, l$. O conjunto de polinômios

$$\{p_{11}, \dots, p_{1n_1}, \dots, p_{k1}, \dots, p_{kn_k}, P_{11}, \dots, P_{1m_1}, \dots, P_{l1}, \dots, P_{lm_l}, tP_{l1}, \dots, tP_{lm_l}\}$$

é uma base para o espaço dos polinômios de grau menor ou igual a $n - 1$, \mathcal{P}_{n-1} , em que n é o grau de $g(t)$, pois são $n - 1$ polinômios de grau menor que n que foi mostrado no [Exemplo 1.47 na página 51](#) são L.I. Vejamos alguns exemplos:

(a) Se $g(t) = (t - a)(t - b)(t - c)$, com $a, b, c \in \mathbb{R}$ distintos, então os polinômios

$$p_1(t) = (t - b)(t - c), \quad p_2(t) = (t - a)(t - c), \quad p_3(t) = (t - a)(t - b)$$

formam uma base de \mathcal{P}_2 .

(b) Se $g(t) = (t - a)^2(t - b)^3$, com $a, b \in \mathbb{R}$ distintos, então os polinômios

$$p_1(t) = (t - a)(t - b)^3, \quad p_2(t) = (t - b)^3, \quad p_3(t) = (t - a)^2(t - b)^2, \\ p_4(t) = (t - a)^2(t - b), \quad p_5(t) = (t - a)^2$$

formam uma base de \mathcal{P}_4 .

(c) Se $g(t) = (t - a)^2(t^2 + bt + c)$, com $a, b, c \in \mathbb{R}$ e $b^2 - 4c < 0$, então os polinômios

$$p_1(t) = (t - a)(t^2 + bt + c), \quad p_2(t) = t^2 + bt + c, \quad p_3(t) = (t - a)^2, \quad p_4(t) = t(t - a)^2$$

formam uma base de \mathcal{P}_3 .

Vamos, agora, provar um resultado interessante sobre a dimensão da soma de subespaços.

Proposição 1.13. *Sejam \mathbb{W}_1 e \mathbb{W}_2 subespaços de dimensão finita de um espaço vetorial \mathbb{V} . Então*

$$\dim(\mathbb{W}_1 + \mathbb{W}_2) = \dim(\mathbb{W}_1) + \dim(\mathbb{W}_2) - \dim(\mathbb{W}_1 \cap \mathbb{W}_2).$$

Demonstração. Seja $\mathcal{B}_0 = \{U_1, \dots, U_k\}$ uma base de $\mathbb{W}_1 \cap \mathbb{W}_2$. Estenda-a a uma base $\mathcal{B}_1 = \{U_1, \dots, U_k, V_1, \dots, V_m\}$ de \mathbb{W}_1 e a uma base $\mathcal{B}_2 = \{U_1, \dots, U_k, W_1, \dots, W_p\}$ de \mathbb{W}_2 . Vamos mostrar que $\mathcal{B} = \{U_1, \dots, U_k, V_1, \dots, V_m, W_1, \dots, W_p\}$ é base de $\mathbb{W}_1 + \mathbb{W}_2$. Vamos mostrar em primeiro lugar que \mathcal{B} é um conjunto L.I. Considere a combinação linear nula

$$\alpha_1 U_1 + \dots + \alpha_k U_k + \beta_1 V_1 + \dots + \beta_m V_m + \gamma_1 W_1 + \dots + \gamma_p W_p = \vec{0}. \quad (1.27)$$

Somando-se $-\gamma_1 W_1 - \dots - \gamma_p W_p$, temos que

$$\overbrace{\alpha_1 U_1 + \dots + \alpha_k U_k + \beta_1 V_1 + \dots + \beta_m V_m}^{\in \mathbb{W}_1} = \overbrace{-\gamma_1 W_1 - \dots - \gamma_p W_p}^{\in \mathbb{W}_2}.$$

O que implica que $-\gamma_1 W_1 - \dots - \gamma_p W_p \in \mathbb{W}_1 \cap \mathbb{W}_2$ e portanto se escreve como combinação linear de U_1, \dots, U_k , ou seja, existem escalares x_1, \dots, x_k tais que

$$-\gamma_1 W_1 - \dots - \gamma_p W_p = x_1 U_1 + \dots + x_k U_k.$$

Somando-se $\gamma_1 W_1 + \dots + \gamma_p W_p$, temos que

$$\gamma_1 W_1 + \dots + \gamma_p W_p + x_1 U_1 + \dots + x_k U_k = \vec{0}$$

Daí segue-se que $x_1 = \dots = x_k = \gamma_1 = \dots = \gamma_p = 0$, pois $\mathcal{B}_2 = \{U_1, \dots, U_k, W_1, \dots, W_p\}$ é uma base. E assim, substituindo-se os valores de $\gamma_1, \dots, \gamma_p$ em (1.27) obtemos uma combinação linear nula de

$$\mathcal{B}_1 = \{U_1, \dots, U_k, V_1, \dots, V_m\},$$

o que implica que

$$\alpha_1 = \dots = \alpha_k = \beta_1 = \dots = \beta_m = 0.$$

Assim, \mathcal{B} é um conjunto L.I. Como além disso, todo vetor de $\mathbb{W}_1 + \mathbb{W}_2$ se escreve como combinação linear dos vetores de \mathcal{B} , então \mathcal{B} é uma base de $\mathbb{W}_1 + \mathbb{W}_2$, de onde segue-se o resultado. ■

Exemplo 1.68. Seja \mathbb{V} um espaço de dimensão n . Sejam \mathbb{W}_1 e \mathbb{W}_2 dois subespaços de \mathbb{V} com $\dim(\mathbb{W}_1) = n - 1$. Se $\mathbb{W}_2 \not\subset \mathbb{W}_1$, então $\dim(\mathbb{W}_1 \cap \mathbb{W}_2) < \dim(\mathbb{W}_2)$. Aplicando-se a **Proposição 1.13**, podemos afirmar que $\mathbb{V} = \mathbb{W}_1 + \mathbb{W}_2$. Além disso, se $\dim(\mathbb{W}_2) = 1$, então $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2$. Por exemplo, se $\mathbb{V} = \mathbb{R}^n$, \mathbb{W}_1 é igual a um hiperplano e \mathbb{W}_2 é igual ao subespaço gerado por um vetor $V \notin \mathbb{W}_1$, então $\mathbb{R}^n = \mathbb{W}_1 \oplus \mathbb{W}_2$.

Exemplo 1.69. Em virtude do exemplo anterior, se \mathbb{W}_2 for igual a um subespaço gerado por um polinômio de grau n , então $\mathcal{P}_n = \mathcal{P}_{n-1} \oplus \mathbb{W}_2$.

1.4.3 Aplicação: Frações Parciais

Vamos mostrar, usando Álgebra Linear, que o procedimento usado no cálculo para integrar frações racionais que se baseia na decomposição em frações parciais sempre funciona.

Considere a fração racional

$$F(t) = \frac{f(t)}{g(t)}$$

em que $f(t)$ e $g(t)$ são polinômios de $\mathbb{R}[t]$ com grau de $f(t)$ menor do que o grau de $g(t)$ e vamos supor que $g(t)$ possa ser decomposto da seguinte forma:

$$g(t) = (t - a_1)^{n_1} \dots (t - a_k)^{n_k} (t^2 + b_1t + c_1)^{m_1} \dots (t^2 + b_l t + c_l)^{m_l},$$

com $a_i \in \mathbb{R}$ distintos para $i = 1, \dots, k$ e $(b_i, c_i) \in \mathbb{R}^2$ distintos tais que $b_i^2 - 4c_i < 0$, para $i = 1, \dots, l$.

Vamos mostrar que existem escalares

$\alpha_{11}, \dots, \alpha_{1n_1}, \dots, \alpha_{k1}, \dots, \alpha_{kn_k}, \beta_{11}, \dots, \beta_{1m_1}, \dots, \beta_{l1}, \dots, \beta_{lm_l}, \gamma_{11}, \dots, \gamma_{1m_1}, \dots, \gamma_{lm_l}$
tais que

$$F(t) = \frac{f(t)}{g(t)} = \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{\alpha_{ij}}{(t - a_i)^j} + \sum_{i=1}^l \sum_{j=1}^{m_i} \frac{\beta_{ij} + \gamma_{ij}t}{(t^2 + b_i t + c_i)^j}. \quad (1.28)$$

Multiplicando-se a equação acima por $g(t)$ obtemos

$$f(t) = \sum_{i=1}^k \sum_{j=1}^{n_i} \alpha_{ij} p_{ij}(t) + \sum_{i=1}^l \sum_{j=1}^{m_i} \beta_{ij} P_{ij}(t) + \sum_{i=1}^l \sum_{j=1}^{m_i} \gamma_{ij} t P_{ij}(t), \quad (1.29)$$

em que

$$p_{ij}(t) = g(t)/(t - a_i)^j, \text{ para } j = 1, \dots, n_i \text{ e } i = 1, \dots, k$$

e

$$P_{ij}(t) = g(t)/(t^2 + b_i t + c_i)^j, \text{ para } j = 1, \dots, m_i \text{ e } i = 1, \dots, l.$$

Mostramos no [Exemplo 1.67 na página 78](#) que o conjunto

$$\{p_{11}, \dots, p_{1n_1}, \dots, p_{k1}, \dots, p_{kn_k}, P_{11}, \dots, P_{1m_1}, \dots, P_{l1}, \dots, P_{lm_l}, tP_{l1}, \dots, tP_{lm_l}\}$$

é uma base para o espaço dos polinômios de grau menor ou igual a $n - 1$, \mathcal{P}_{n-1} , em que n é o grau de $g(t)$. Como o grau de $f(t)$ é menor do que o grau de $g(t)$, então $f(t)$ pertence a \mathcal{P}_{n-1} e assim existem escalares $\alpha_{11}, \dots, \alpha_{1n_1}, \dots, \alpha_{k1}, \dots, \alpha_{kn_k}, \beta_{11}, \dots, \beta_{1m_1}, \dots, \beta_{l1}, \dots, \beta_{lm_l}, \gamma_{11}, \dots, \gamma_{1m_1}, \dots, \gamma_{lm_l}$ que satisfazem a equação (1.29) e portanto também a equação (1.28).

1.4.4 Aplicação: Interpolação por Splines

Dados números reais $x_1 < x_2 < \dots < x_n$, igualmente espaçados, isto é, tais que $x_{k+1} - x_k = (x_n - x_1)/(n - 1)$, para $k = 1, \dots, n - 1$. Defina os splines (**Exemplo 1.51 na página 56**) $q_k(x) = 0$, se x está fora do intervalo $[x_{k-3}, x_{k+1}]$ e

$$q_k(x) = \begin{cases} p_1(t), & \text{em que } t = (x - x_{k-3})/h & \text{se } x_{k-3} \leq x < x_{k-2}, \\ p_2(t), & \text{em que } t = (x - x_{k-2})/h & \text{se } x_{k-2} \leq x < x_{k-1}, \\ p_2(1-t), & \text{em que } t = (x - x_{k-1})/h & \text{se } x_{k-1} \leq x < x_k, \\ p_1(1-t), & \text{em que } t = (x - x_k)/h & \text{se } x_k \leq x \leq x_{k+1}, \end{cases}$$

para $k = 1, \dots, n + 2$, em que

$$\begin{aligned} p_1(s) &= \frac{1}{4}s^3, \\ p_2(s) &= 1 - \frac{3}{4}(1+s)(1-s)^2 \end{aligned}$$

e $h = x_{k+1} - x_k = (x_n - x_1)/(n - 1)$. Mostramos no **Exemplo 1.51 na página 56** que o conjunto de splines $\{q_k \mid k = 1, \dots, n + 2\}$ é linearmente independente. Isto mostra que a dimensão do espaço de splines (cúbicos) com n pontos de quebra igualmente espaçados é maior ou igual a $n + 2$ (**Teorema 1.12(e) na página 75**).

Por outro lado, no **Exemplo 1.33 na página 37** vimos que podemos escrever todo spline do espaço de splines (cúbicos) com $n - 2$ pontos de quebra, \mathcal{S} , como combinação linear de apenas $n + 2$ splines. Ou seja, podemos ter um conjunto de geradores para \mathcal{S} com apenas $n + 2$ splines. Isto mostra que a dimensão do espaço de splines (cúbicos) com $n - 2$ pontos de quebra é menor ou igual a $n + 2$ (**Teorema 1.12(b) na página 75**). Isto nos permite concluir que a dimensão do espaço de splines (cúbicos) com n pontos de quebra igualmente espaçados, \mathcal{S} , é $n + 2$.

Proposição 1.14. A dimensão do espaço de splines (cúbicos) com $n - 2$ pontos de quebra igualmente espaçados, \mathcal{S} , é $n + 2$.

Assim cada função $f \in \mathcal{S}$ tem uma única representação como uma combinação linear (Teorema 1.9 na página 70)

$$f(x) = \sum_{j=1}^{n+2} c_j q_j(x).$$

Usando a base $\{q_k \mid k = 1, \dots, n+2\}$ o problema de encontrar uma função de \mathcal{S} que se ajusta a um conjunto de pontos $(x_1, y_1), \dots, (x_{n+2}, y_{n+2})$ toma a forma

$$\sum_{j=1}^{n+2} c_j q_j(x_i) = f(x_i), \quad i = 1, \dots, n+2$$

ou

$$AX = B,$$

em que a matriz A é definida por $a_{ij} = q_j(x_i)$, o vetor B é dado por $b_i = y_i$ e X é o vetor dos coeficientes c_j , para $i, j = 1, \dots, n+2$.

Exemplo 1.70. Considere o seguinte conjunto de dados

x	-1	0	1	2	3
y	1	4	1	3	0

Dividindo-se o intervalo $[-1, 3]$ em dois subintervalos e usando a base $\{q_1, q_2, q_3, q_4, q_5\}$, o problema de encontrar um spline

$$f(x) = c_1 q_1(x) + c_2 q_2(x) + c_3 q_3(x) + c_4 q_4(x) + c_5 q_5(x)$$

que melhor se ajusta ao conjunto de pontos $(-1, 1)$, $(0, 4)$, $(1, 1)$, $(2, 3)$, $(3, 0)$ toma a forma

$$c_1q_1(x_i) + c_2q_2(x_i) + c_3q_3(x_i) + c_4q_4(x_i) + c_5q_5(x_i) = f(x_i), \quad i = 1, 2, 3, 4, 5$$

ou

$$AX = B,$$

em que a matriz A é definida por $a_{ij} = q_j(x_i)$, B por $b_j = y_j$ e X por $x_j = c_j$, para $i = 1, \dots, 5, j = 1, \dots, 5$. Neste caso

$$A = \begin{bmatrix} \frac{1}{4} & 1 & \frac{1}{4} & 0 & 0 \\ \frac{1}{32} & \frac{23}{32} & \frac{23}{32} & \frac{1}{32} & 0 \\ 0 & \frac{1}{4} & 1 & \frac{1}{4} & 0 \\ 0 & \frac{1}{32} & \frac{23}{32} & \frac{23}{32} & \frac{1}{32} \\ 0 & 0 & \frac{1}{4} & 1 & \frac{1}{4} \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 4 \\ 1 \\ 3 \\ 0 \end{bmatrix}, \quad X = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix},$$

Os coeficientes c_j obtidos resolvendo o sistema linear são

$$c_1 = -103/9, \quad c_2 = 95/9, \quad c_3 = -35/9, \quad c_4 = 9, \quad c_5 = -289/9$$

Exercícios Numéricos (respostas na página 413)

1.4.1. Quais dos seguintes conjuntos de vetores formam uma base para o \mathbb{R}^4 ?

- (a) $\{(1, 0, 0, 1), (0, 1, 0, 0), (1, 1, 1, 1), (0, 1, 1, 1)\}$;
- (b) $\{(1, -1, 0, 2), (3, -1, 2, 1), (1, 0, 0, 1)\}$;
- (c) $\{(0, 0, 1, 1), (-1, 1, 1, 2), (1, 1, 0, 0), (2, 1, 2, 1)\}$;

1.4.2. Encontre uma base para os seguintes subespaços do \mathbb{R}^3 :

- (a) Todos os vetores da forma (a, b, c) , onde $b = a$;
- (b) Todos os vetores da forma (a, b, c) , onde $a = 0$;
- (c) Todos os vetores da forma $(a - b, b + c, 2a - b + c)$.

1.4.3. Encontre as dimensões dos seguintes subespaços do \mathbb{R}^4 :

- (a) Todos os vetores da forma (a, b, c, d) , onde $d = a + b$;
- (b) Todos os vetores da forma (a, b, c, d) , onde $c = a - b$ e $d = a + b$;
- (c) Todos os vetores da forma $(a + c, a - b, b + c, -a + b)$.

1.4.4. Determine os valores de a para os quais $\{(a^2, 0, 1), (0, a, 2), (1, 0, 1)\}$ é uma base do \mathbb{R}^3 .

1.4.5. Encontre os valores de λ tais que o sistema homogêneo $(A - \lambda I_n)X = \vec{0}$ tem solução não trivial e para estes valores de λ , encontre uma base para o espaço solução, para as matrizes A dadas:

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

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

$$(c) A = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix};$$

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

- 1.4.6.** Sejam $V_1 = (2, 1, 3)$, $V_2 = (3, -1, 4)$ e $V_3 = (2, 6, 4)$.
- (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 , $[V_1, V_2, V_3]$.
 - (d) Dê uma interpretação geométrica para o subespaço $[V_1, V_2, V_3]$.
- 1.4.7.** Dados $V_1 = (1, 1, 1)$ e $V_2 = (3, -1, 4)$:
- (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 .
- 1.4.8.** Seja \mathbb{W} o subespaço de \mathbb{R}^3 formado pelos vetores $V = (x, y, z)$ tais que $x + 2y + 4z = 0$. Obtenha uma base $\{V_1, V_2, V_3\}$ de \mathbb{R}^3 tal que V_1 e V_2 pertençam a \mathbb{W} .
- 1.4.9.** Mostre que os polinômios $1, t - 1, t^2 - 3t + 1$ formam uma base de \mathcal{P}_2 . Exprima o polinômio $2t^2 - 5t + 6$ como combinação linear dos elementos desta base.
- 1.4.10.** Em $\mathcal{C}^0[-\pi, \pi]$, encontre a dimensão do subespaço gerado por $1, \cos 2t, \cos^2 t$.
- 1.4.11.** Seja \mathbb{W}_1 o subespaço de \mathcal{P}_3 que consiste de todos os polinômios $p(t)$ tais que $p(0) = 0$, e seja \mathbb{W}_2 o subespaço de \mathcal{P}_3 dos polinômios $q(t)$ tais que $q(1) = 0$. Encontre bases para
- (a) \mathbb{W}_1 ;
 - (b) \mathbb{W}_2 ;
 - (c) $\mathbb{W}_1 \cap \mathbb{W}_2$.
- 1.4.12.** Determine uma base e a dimensão de cada um dos subespaços de \mathcal{M}_{nn} abaixo:
- (a) Matrizes simétricas.

- (b) Matrizes anti-simétricas.
- (c) Matrizes de traço igual a zero.
- (d) Matrizes que têm a primeira linha igual a última coluna.
- (e) Matrizes em que a soma dos elementos da primeira linha é igual a soma dos elementos da última coluna.

1.4.13. Para $a \in \mathbb{R}$ fixado, determine a dimensão do subespaço de \mathcal{P}_n definido por

$$\mathbb{W} = \{p \in \mathcal{P}_n \mid p(a) = 0\}.$$

1.4.14. Se V_1, V_2 e V_3 são vetores de um espaço vetorial \mathbb{V} e \mathbb{W} é o subespaço gerado por V_1, V_2 e V_3 , então a dimensão de \mathbb{W} é igual a 3?

1.4.15. Suponha que $\{X_1, X_2, \dots, X_n\}$ seja uma base do \mathbb{R}^n . Mostre que se A é uma matriz $n \times n$ não singular, então $\{AX_1, AX_2, \dots, AX_n\}$ também é uma base de \mathbb{R}^n . E se A for singular?

1.4.16. Mostre que se \mathbb{V} e \mathbb{W} são subespaços de dimensão 2 de um espaço vetorial de dimensão 3, então

$$\mathbb{V} \cap \mathbb{W} \neq \{\vec{0}\}.$$

O mesmo seria verdade se estes fossem subespaços de um espaço vetorial de dimensão 4? (Sugestão: $x_1V_1 + x_2V_2 = y_1W_1 + y_2W_2 \in \mathbb{V} \cap \mathbb{W}$ se, e somente se, $x_1V_1 + x_2V_2 - y_1W_1 - y_2W_2 = \vec{0}$.)

1.4.17. Mostre que $\{2, t + 1, t^2 + 1, \dots, t^n + 1, \dots\}$ é uma base para $\mathcal{P} = \mathbb{R}[t]$.

1.4.18. Mostre que $\{1, t^2, t^4, \dots, t^{2n}, \dots\}$ é uma base para o subespaço dos polinômios que satisfazem $p(-t) = p(t)$, para todo $t \in \mathbb{R}$.

Exercícios usando o MATLAB[®]

Comandos do pacote GAAL:

>> A=randi(m,n) cria a matriz $m \times n$ com entradas inteiras e aleatórias entre -5 e 5 . >> A=randi(m,n,p) cria a matriz $m \times n$ com entradas inteiras e aleatórias entre $-p$ e p .

>> A=spline1(X,nbp,a,b) cria a matriz $a_{ij} = q_j(x_i)$, se $X=[x_1, \dots, x_n]$, para um intervalo $[a, b]$ dividido em nbp-1 subintervalos.

>> plotspline1(C,nbp,a,b) desenha o spline definido por $\sum_{k=1}^{nbp+2} c_k q_k(x)$.

- 1.4.19.** Defina $A=\text{randi}(3,2)*\text{randi}(2,5,2)$. Verifique se as colunas de A são linearmente independentes. Se elas forem linearmente dependentes, escreva três das colunas de A como combinação linear das outras duas.
- 1.4.20.** Defina a matriz $A=\text{randi}(4,3)*\text{randi}(3,5,2)$. Considere o subespaço gerado pelas colunas de A . Extraia das colunas de A uma base para este subespaço.
- 1.4.21.** Defina a matriz $A=\text{randi}(4,2)$. Verifique que as colunas de A são L.I. Considere o conjunto formado pelas colunas de A . Complete este conjunto até obter uma base do \mathbb{R}^4 .
- 1.4.22.** (a) Defina a matriz $A=\text{randi}(4,3)*\text{randi}(3,5,2)$. Considere o subespaço gerado pelas colunas de A . Obtenha uma base para este subespaço, cujo primeiro vetor é a soma das duas primeiras colunas de A .
- (b) Defina a matriz $B=A*\text{randi}(5,2)$. Sejam v_1, v_2 as colunas de B , complete a uma base do subespaço gerado pelas colunas de A .
- 1.4.23.** (a) Use o comando $P=\text{randi}(5,2)$, 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 um spline com 3 pontos de quebra cujo gráfico passa pelos pontos cujas coordenadas são dadas pelas linhas da matriz P (veja o [Exemplo 1.70 na página 85](#)). O comando $A=\text{spline1}(P(:,1),3,-5,5)$ cria a matriz cuja coluna j é $q_j(P(:,1))$. O comando $C=A \setminus P(:,2)$ resolve o sistema linear $AX=P(:,2)$. Se não conseguiu, repita o item anterior. Por que pode não ser possível?
- (c) Desenhe os pontos e o spline com os comandos `clf, po(P), plotspline1(C,3,-5,5)`.
- (d) Desenhe os eixos coordenados com o comando `eixos`.

Exercícios Teóricos

- 1.4.24.** Seja \mathcal{P}_n o subespaço dos polinômios de grau menor ou igual a n .
- Mostre que existe um subespaço \mathbb{W}_2 de \mathcal{P} , tal que $\mathcal{P} = \mathcal{P}_n \oplus \mathbb{W}_2$.
 - Seja \mathbb{W}_3 um subespaço de \mathcal{P}_n . Seja $\{p_0, \dots, p_k\}$ uma base de \mathbb{W}_3 e $\{p_0, \dots, p_k, p_{k+1}, \dots, p_n\}$ uma base de \mathcal{P}_n . Mostre que $\{p_0, \dots, p_n, x^{n+1}, x^{n+2}, \dots\}$ é uma base de \mathcal{P} .
 - Mostre que existe um subespaço \mathbb{W}_4 de \mathcal{P} , tal que $\mathcal{P} = \mathbb{W}_3 \oplus \mathbb{W}_4$.
- 1.4.25.** Seja \mathcal{X} um conjunto infinito. Para cada $a \in \mathcal{X}$, seja $f_a : \mathcal{X} \rightarrow \mathbb{R}$ a função tal que $f_a(a) = 1$ e $f_a(x) = 0$, se $x \neq a$. Mostre que o conjunto $\mathcal{G} \subset \mathcal{F}(\mathcal{X}; \mathbb{R})$ formado por estas funções é linearmente independente. Mostre ainda que \mathcal{G} não gera $\mathcal{F}(\mathcal{X}; \mathbb{R})$.
- 1.4.26.** Sejam \mathbb{W}_1 e \mathbb{W}_2 subespaços de dimensão finita. Seja $\mathbb{V} = \mathbb{W}_1 + \mathbb{W}_2$. Mostre que \mathbb{V} é soma direta de \mathbb{W}_1 e \mathbb{W}_2 se, e somente se, $\dim(\mathbb{V}) = \dim(\mathbb{W}_1) + \dim(\mathbb{W}_2)$.
- 1.4.27.** Sejam \mathbb{W}_1 e \mathbb{W}_2 subespaços de um espaço vetorial \mathbb{V} .
- Seja $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2$. Se \mathcal{B}_1 é uma base de \mathbb{W}_1 e \mathcal{B}_2 é uma base de \mathbb{W}_2 , mostre que $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$ e $\mathcal{B}_1 \cup \mathcal{B}_2$ é uma base de \mathbb{V} .
 - Reciprocamente, se \mathcal{B}_1 e \mathcal{B}_2 são bases disjuntas de \mathbb{W}_1 e \mathbb{W}_2 , respectivamente e $\mathcal{B}_1 \cup \mathcal{B}_2$ é uma base de \mathbb{V} , mostre que $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2$.
- 1.4.28.** Seja \mathbb{V} um espaço vetorial de dimensão finita. Mostre que se \mathbb{W}_1 é um subespaço de \mathbb{V} , então existe um subespaço \mathbb{W}_2 , tal que $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2$.

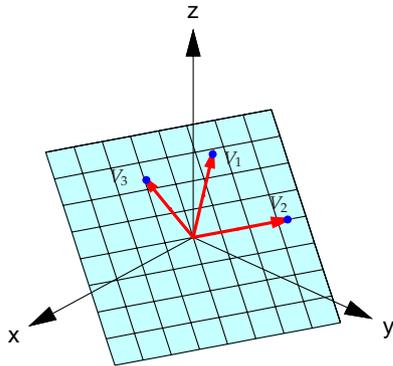


Figura 1.14: Três vetores linearmente dependentes (coplanares)

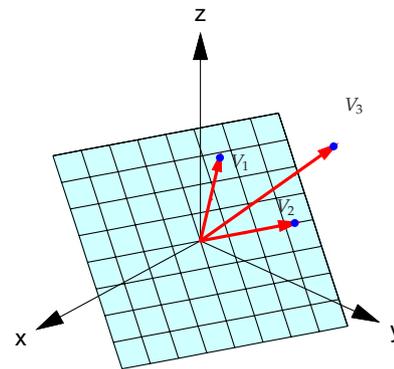


Figura 1.15: Três vetores linearmente independentes

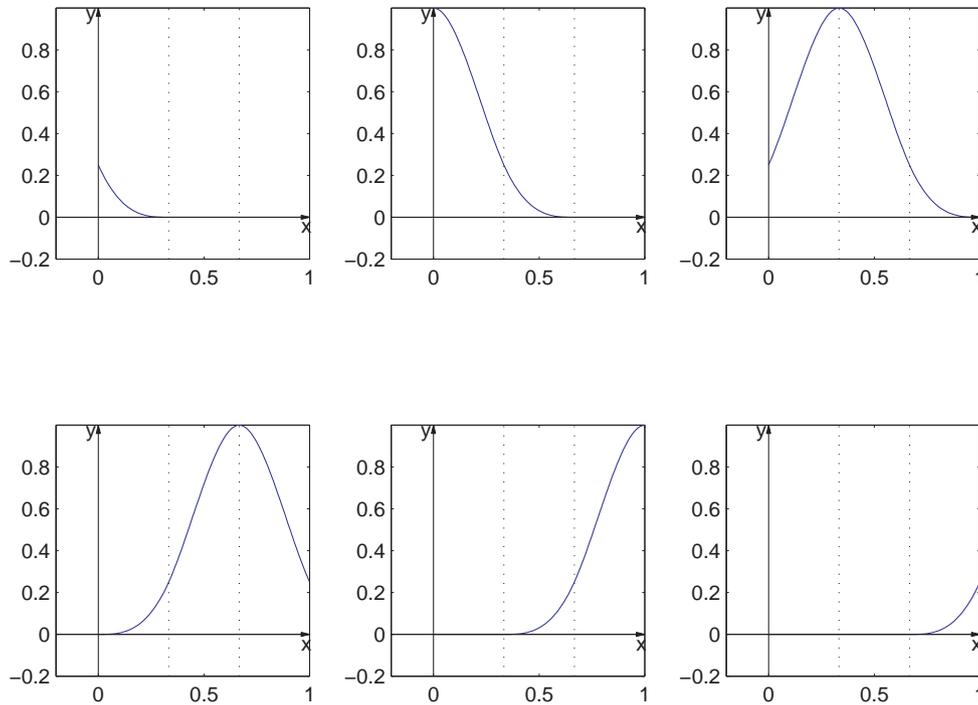


Figura 1.16: Funções q_k , para $k = 1, \dots, 6$ no intervalo $[0, 1]$ dividido em 3 subintervalos

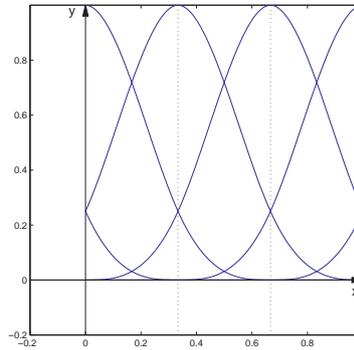


Figura 1.17: Funções q_k , para $k = 1, \dots, 6$ no intervalo $[0, 1]$ dividido em 3 subintervalos

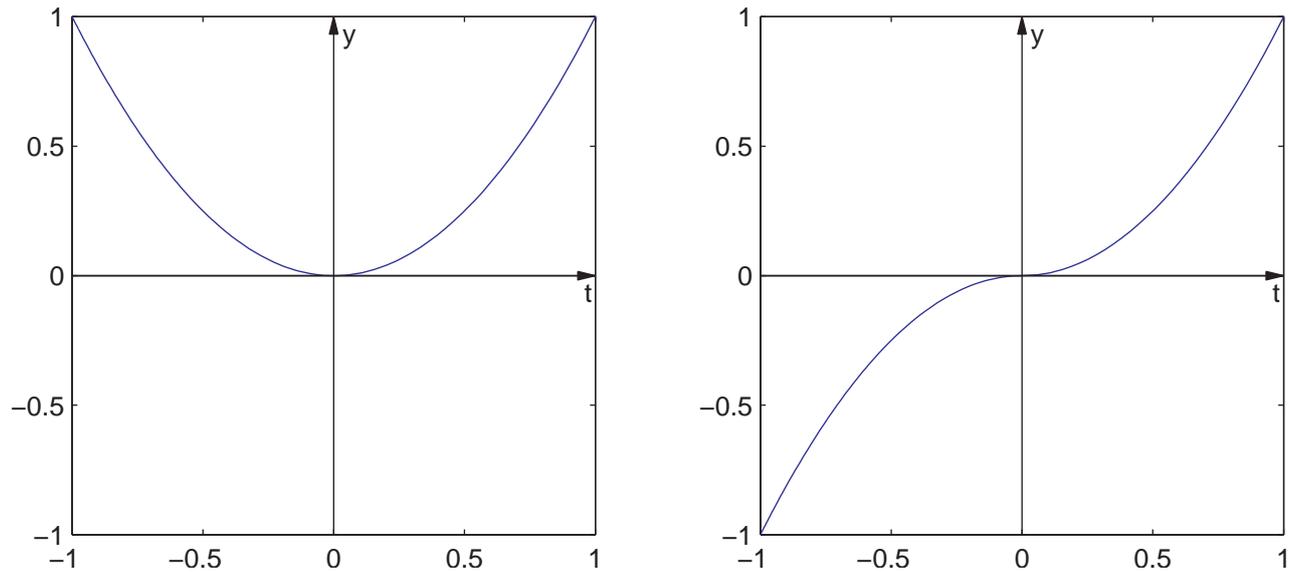


Figura 1.18: $f_1(t) = t^2$ e $f_2(t) = t|t|$ são L.I. mas o wronskiano é igual a zero para todo t

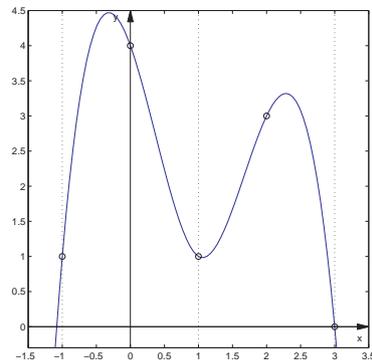


Figura 1.19: Ajuste dos dados do Exemplo 1.70 por splines dividindo-se o intervalo $[-1, 3]$ em dois subintervalos

Teste do Capítulo

- Determine uma base e a dimensão de cada um dos subespaços abaixo:
 - Das matrizes $n \times n$ que têm a primeira linha igual a última linha.
 - Das funções que possuem a segunda derivada identicamente nula.

- Mostre que $\{t, t^3, t^5, \dots, t^{2n+1}, \dots\}$ é uma base para o subespaço dos polinômios que satisfazem $p(-t) = -p(t)$, para todo $t \in \mathbb{R}$.

- Mostre que $y_1(t) = e^{\lambda t}$, $y_2(t) = te^{\lambda t}$ são soluções linearmente independentes da equação diferencial $y'' - 2\lambda y' + \lambda^2 y = 0$.

- Seja $\mathcal{S} = \{p_0, p_1, \dots, p_n\}$ um conjunto de polinômios tais que o grau de p_k é igual a k , para $k = 0, \dots, n$. Mostre que \mathcal{S} é uma base para o espaço dos polinômios de grau menor ou igual a n .

Capítulo 2

Espaços com Produto Interno

2.1 Produto Escalar e Norma

2.1.1 Produto Interno

Produto Escalar em \mathbb{R}^n

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 2.1. Definimos o **produto escalar** ou **interno** de dois vetores $X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in \mathbb{R}^n$ por

$$X \cdot Y = x_1y_1 + x_2y_2 + \dots + x_ny_n = \sum_{i=1}^n x_iy_i.$$

Exemplo 2.1. 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

$$V \cdot W = (1)(5) + (-2)(3) + (4)(-1) + (3)(-2) + (5)(1) = -6.$$

São válidas as seguintes propriedades para o produto escalar de vetores do \mathbb{R}^n .

Proposição 2.1. Se X, Y e Z são vetores de \mathbb{R}^n e α é um escalar, então

(a) $(X + Y) \cdot Z = X \cdot Z + Y \cdot Z;$

(b) $(\alpha X) \cdot Y = \alpha(X \cdot Y) = X \cdot (\alpha Y);$

(c) $X \cdot Y = Y \cdot X;$

(d) Se $X \in \mathbb{R}^n, X \neq \vec{0}$, então $X \cdot X > 0$.

Demonstração. Escrevendo os vetores como matrizes colunas, o produto interno de

dois vetores $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ e $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ pode ser escrito em termos do produto de

matrizes como $X \cdot Y = X^t Y$. Sejam $X, Y, Z \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$. Usando as propriedades da álgebra matricial, temos que

(a) $X \cdot (Y + Z) = X^t(Y + Z) = X^t Y + X^t Z = X \cdot Y + X \cdot Z;$

(b) $\alpha(X \cdot Y) = \alpha(X^t Y) = (\alpha X^t) Y = (\alpha X)^t Y = (\alpha X) \cdot Y;$ a outra igualdade é inteiramente análoga;

(c) $X \cdot Y = X^t Y = (X^t Y)^t = Y^t X = Y \cdot X;$ pois $X^t Y$ é uma matriz 1×1 que é igual a sua transposta.

(d) Se $X = (x_1, \dots, x_n) \neq \vec{0}$, então $x_j \neq 0$, para algum j . Então

$$X \cdot X = x_1^2 + \dots + x_n^2 \geq x_j^2 > 0.$$



Produto Interno

O conceito de produto escalar pode ser estendido não somente ao \mathbb{R}^n , mas também a espaços vetoriais abstratos.

Definição 2.2. Seja \mathbb{V} um espaço vetorial. Um **produto escalar** ou **interno** em \mathbb{V} é uma função que associa a cada par ordenado de vetores V e W em \mathbb{V} um escalar denotado por $\langle V, W \rangle$ satisfazendo os seguintes axiomas:

- (a) Para todos os $V, W, U \in \mathbb{V}$, $\langle V + U, W \rangle = \langle V, W \rangle + \langle U, W \rangle$;
- (b) Para todos os $V, W \in \mathbb{V}$ e todo escalar α , $\langle \alpha V, W \rangle = \alpha \langle V, W \rangle$;
- (c) Para todos os $V, W \in \mathbb{V}$, $\overline{\langle V, W \rangle} = \langle W, V \rangle$, em que a barra significa o conjugado, ou seja, se $\alpha = a + ib \in \mathbb{C}$, então $\bar{\alpha} = a - ib$. Assim se $\langle V, W \rangle$ é um número real, então $\langle V, W \rangle = \langle W, V \rangle$.
- (d) Para todo $V \in \mathbb{V}$, $V \neq \vec{0}$, $\langle V, V \rangle > 0$.

Se está definido um produto interno em \mathbb{V} , dizemos que \mathbb{V} é um **espaço vetorial com produto interno**.

Observe que se o conjunto de escalares é o conjunto dos números reais, então o axioma (c) é equivalente a dizer que para todos $V, W \in \mathbb{V}$, $\langle V, W \rangle = \langle W, V \rangle$.

Exemplo 2.2. Segue da **Proposição 2.1** que $\langle V, W \rangle = V \cdot W$, para todos $V, W \in \mathbb{R}^n$ é um produto interno.

Exemplo 2.3. Seja $\mathbb{V} = \mathbb{C}^n$ o espaço vetorial sobre \mathbb{C} das n -uplas de números complexos. Vamos mostrar que definindo

$$\langle X, Y \rangle = \sum_{k=1}^n x_k \bar{y}_k, \quad \text{para todos } X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in \mathbb{C}^n$$

temos um produto interno.

Por exemplo, se $X = (1, i)$ e $Y = (1 + i, 2)$, então

$$\langle X, Y \rangle = 1 \overline{(1 + i)} + i \bar{2} = 1 - i + 2i = 1 + i.$$

Sejam $X, Y, Z \in \mathbb{C}^n$ e $\alpha \in \mathbb{C}$.

$$\begin{aligned} \text{(a)} \quad \langle X + Z, Y \rangle &= \sum_{k=1}^n (x_k + z_k) \bar{y}_k = \sum_{k=1}^n (x_k \bar{y}_k + z_k \bar{y}_k) = \sum_{k=1}^n x_k \bar{y}_k + \sum_{k=1}^n z_k \bar{y}_k = \\ &= \langle X, Y \rangle + \langle Z, Y \rangle; \end{aligned}$$

$$\text{(b)} \quad \langle \alpha X, Y \rangle = \sum_{k=1}^n \alpha x_k \bar{y}_k = \alpha \sum_{k=1}^n x_k \bar{y}_k = \alpha \langle X, Y \rangle;$$

$$\text{(c)} \quad \overline{\langle X, Y \rangle} = \overline{\sum_{k=1}^n x_k \bar{y}_k} = \sum_{k=1}^n \overline{x_k \bar{y}_k} = \sum_{k=1}^n y_k \bar{x}_k = \langle Y, X \rangle;$$

(d) Seja $X \neq \bar{0}$, $\langle X, X \rangle = \sum_{k=1}^n x_k \bar{x}_k = \sum_{k=1}^n |x_k|^2 > 0$.

Exemplo 2.4. Seja $\mathbb{V} = \mathcal{C}^0[a, b]$ o espaço vetorial das funções reais contínuas $f : [a, b] \rightarrow \mathbb{R}$. Vamos mostrar que definindo

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt, \quad \text{para todas as funções } f, g \in \mathcal{C}^0[a, b]$$

temos um produto interno em \mathbb{V} .

Por exemplo, se $f(t) = t, g(t) = e^t \in \mathcal{C}^0[0, 1]$, então

$$\langle f, g \rangle = \int_0^1 te^t dt = te^t \Big|_0^1 - \int_0^1 e^t dt = 1.$$

Sejam $f, g, h \in \mathcal{C}^0[a, b]$ e α um escalar.

(a) $\langle f + g, h \rangle = \int_a^b (f(t) + g(t))h(t)dt = \int_a^b f(t)h(t)dt + \int_a^b g(t)h(t)dt = \langle f, h \rangle + \langle g, h \rangle$.

(b) $\langle \alpha f, g \rangle = \int_a^b \alpha f(t)g(t)dt = \alpha \int_a^b f(t)g(t)dt = \alpha \langle f, g \rangle$.

(c) $\overline{\langle f, g \rangle} = \overline{\int_a^b f(t)g(t)dt} = \int_a^b \overline{f(t)g(t)}dt = \int_a^b \overline{f(t)}\overline{g(t)}dt = \int_a^b g(t)\overline{f(t)}dt = \langle g, f \rangle$, pois as funções são reais e o conjugado de um número real é igual a ele mesmo.

(d) Se $f \neq \bar{0}$, então, como f é contínua, existe um subintervalo de $[a, b]$, onde f^2 é limitada inferiormente por um número maior do que zero. Assim,

$$\langle f, f \rangle = \int_a^b (f(t))^2 dt > 0.$$

Exemplo 2.5. Seja $\mathbb{V} = \mathcal{C}^0([a, b], \mathbb{C})$ o espaço vetorial sobre \mathbb{C} das funções contínuas $f : [a, b] \rightarrow \mathbb{C}$, definidas em $[a, b]$ e tomando valores complexos. Vamos mostrar que definindo

$$\langle f, g \rangle = \int_a^b f(t) \overline{g(t)} dt, \quad \text{para todas as funções } f, g \in \mathcal{C}^0([a, b], \mathbb{C})$$

temos um produto interno em \mathbb{V} .

Por exemplo, se $f(t) = t, g(t) = e^{it} = \cos t + i \sin t \in \mathcal{C}^0([0, 2\pi], \mathbb{C})$, então

$$\langle f, g \rangle = \int_0^{2\pi} t e^{it} dt = \int_0^{2\pi} t e^{-it} dt = - \int_0^{-2\pi} s e^s ds = -s e^s \Big|_0^{-2\pi} + \int_0^{-2\pi} e^s ds = 2\pi i$$

Sejam $f, g, h \in \mathcal{C}^0([a, b], \mathbb{C})$ e α um escalar.

$$(a) \quad \langle f + g, h \rangle = \int_a^b (f(t) + g(t)) \overline{h(t)} dt = \int_a^b f(t) \overline{h(t)} dt + \int_a^b g(t) \overline{h(t)} dt = \langle f, h \rangle + \langle g, h \rangle.$$

$$(b) \quad \langle \alpha f, g \rangle = \int_a^b \alpha f(t) \overline{g(t)} dt = \alpha \int_a^b f(t) \overline{g(t)} dt = \alpha \langle f, g \rangle.$$

$$(c) \quad \overline{\langle f, g \rangle} = \overline{\int_a^b f(t) \overline{g(t)} dt} = \int_a^b \overline{f(t) \overline{g(t)}} dt = \int_a^b g(t) \overline{f(t)} dt = \langle g, f \rangle.$$

(d) Se $f \neq \bar{0}$, então, como f é contínua, existe um subintervalo de $[a, b]$, onde $|f|^2$ é limitada inferiormente por um número maior do que zero. Assim,

$$\langle f, f \rangle = \int_a^b |f(t)|^2 dt > 0.$$

A partir dos axiomas de produto interno podemos provar outras propriedades.

Proposição 2.2. *Seja \mathbb{V} um espaço com produto interno. São válidas as seguintes propriedades:*

- (a) *Para todos os $V, W, U \in \mathbb{V}$, $\langle V, W + U \rangle = \langle V, W \rangle + \langle V, U \rangle$;*
- (b) *Para todos os $V, W \in \mathbb{V}$ e todo escalar α , $\langle V, \alpha W \rangle = \bar{\alpha} \langle V, W \rangle$;*
- (c) *$\langle V, V \rangle = 0$ se, e somente se, $V = \bar{0}$;*

Demonstração. Sejam $V, W, U \in \mathbb{V}$ e α um escalar.

- (a) $\langle V, W + U \rangle = \overline{\langle W + U, V \rangle} = \overline{\langle W, V \rangle + \langle U, V \rangle} = \langle V, W \rangle + \langle V, U \rangle$;
- (b) $\langle V, \alpha W \rangle = \overline{\langle \alpha W, V \rangle} = \bar{\alpha} \overline{\langle W, V \rangle} = \bar{\alpha} \langle V, W \rangle$;
- (c) Se $V \neq \bar{0}$, então pela definição de produto interno, $\langle V, V \rangle > 0$. Se $V = \bar{0}$, então $\langle \bar{0}, \bar{0} \rangle = \langle \alpha \bar{0}, \bar{0} \rangle = \alpha \langle \bar{0}, \bar{0} \rangle$, para todo escalar α . O que implica que $\langle \bar{0}, \bar{0} \rangle = 0$.

■

2.1.2 Norma

Assim como o produto interno pode ser estendido ao \mathbb{R}^n e a espaços vetoriais quaisquer, a noção de norma ou comprimento de um vetor pode ser estendida ao \mathbb{R}^n e a espaços vetoriais quaisquer. Vamos definir a norma de um vetor já em um espaço vetorial qualquer.

Definição 2.3. Seja \mathbb{V} um espaço vetorial com produto interno. Para todo vetor $V \in \mathbb{V}$, definimos a **norma** de V denotada por $\|V\|$ como sendo

$$\|V\| = \sqrt{\langle V, V \rangle}.$$

Exemplo 2.6. Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ com o produto interno

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt, \quad \text{para todas as funções } f, g \in \mathcal{C}^0[-1, 1].$$

Sejam $f(t) = 1$, $g(t) = t$ e $h(t) = \cos \pi t$. Então

(a) $\|f\|^2 = \langle f, f \rangle = \int_{-1}^1 1dt = 2$. Assim, $\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{2}$.

(b) $\|g\|^2 = \langle g, g \rangle = \int_{-1}^1 t^2 dt = \frac{t^3}{3} \Big|_{-1}^1 = 2/3$. Assim, $\|g\| = \sqrt{\langle g, g \rangle} = \sqrt{2/3}$.

(c) $\|h\|^2 = \langle h, h \rangle = \int_{-1}^1 \cos^2 \pi t dt = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2 s ds = \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 + \cos 2s) ds = \frac{1}{2\pi} (s \Big|_{-\pi}^{\pi} + \frac{1}{2} \text{sen } 2s \Big|_{-\pi}^{\pi}) = 1$. Assim, $\|h\| = \sqrt{\langle h, h \rangle} = 1$.

Proposição 2.3. Seja \mathbb{V} um espaço vetorial com produto interno.

(a) Para todo $V \in \mathbb{V}$, $\|V\| \geq 0$ e $\|V\| = 0$ se, e somente se, $V = \vec{0}$;

- (b) Para todo vetor $V \in \mathbb{V}$ e para todo escalar α , $\|\alpha V\| = |\alpha| \|V\|$;
- (c) Para todos os vetores $V, W \in \mathbb{V}$, $|\langle V, W \rangle| \leq \|V\| \|W\|$ (Desigualdade de Cauchy-Schwarz);
- (d) Para todos os vetores $V, W \in \mathbb{V}$, $\|V + W\| \leq \|V\| + \|W\|$ (Desigualdade triangular);
-

Demonstração.

(a) Decorre da [Definição 2.2 na página 102](#) (d) de produto interno e da [Proposição 2.2 na página 106](#) (c).

(b) $\|\alpha V\| = \sqrt{\langle \alpha V, \alpha V \rangle} = \sqrt{|\alpha|^2 \langle V, V \rangle} = |\alpha| \sqrt{\langle V, V \rangle} = |\alpha| \|V\|.$

(c) A norma de $V - \lambda W$ é maior ou igual a zero, para qualquer escalar λ . Assim,

$$0 \leq \|V - \lambda W\|^2 = \langle V - \lambda W, V - \lambda W \rangle = \|V\|^2 - \lambda \langle W, V \rangle - \bar{\lambda} \langle V, W \rangle + |\lambda|^2 \|W\|^2.$$

Tomando

$$\lambda = \frac{\langle V, W \rangle}{\|W\|^2}$$

ficamos com

$$0 \leq \|V\|^2 - \frac{\langle V, W \rangle}{\|W\|^2} \langle W, V \rangle - \frac{\langle W, V \rangle}{\|W\|^2} \langle V, W \rangle + \frac{|\langle V, W \rangle|^2}{\|W\|^4} \|W\|^2 = \|V\|^2 - \frac{|\langle V, W \rangle|^2}{\|W\|^2}$$

Logo, $|\langle V, W \rangle| \leq \|V\| \|W\|.$

(d) Pelo item anterior temos que

$$\begin{aligned}\|V + W\|^2 &= \langle V + W, V + W \rangle = \langle V, V \rangle + \langle V, W \rangle + \langle W, V \rangle + \langle W, W \rangle \\ &= \|V\|^2 + 2\Re \langle V, W \rangle + \|W\|^2 \\ &\leq \|V\|^2 + 2|\langle V, W \rangle| + \|W\|^2 \\ &\leq \|V\|^2 + 2\|V\| \|W\| + \|W\|^2 \\ &\leq (\|V\| + \|W\|)^2;\end{aligned}$$

Tomando a raiz quadrada, segue-se o resultado. ■

Podemos, agora, estender o conceito de ângulo entre vetores para elementos de um espaço vetorial *real* com produto interno. Definimos o **ângulo entre dois vetores não nulos** V e W como sendo o número real θ entre 0 e π tal que

$$\cos \theta = \frac{\langle V, W \rangle}{\|V\| \|W\|}.$$

Segue da desigualdade de Cauchy-Schwarz ([Proposição 2.3 \(c\)](#)) que o ângulo θ está bem definido, pois

$$|\langle V, W \rangle| \leq \|V\| \|W\|$$

implica que

$$-1 \leq \frac{\langle V, W \rangle}{\|V\| \|W\|} \leq 1.$$

Exemplo 2.7. Seja $\mathbb{V} = C^0[-1,1]$ o conjunto das funções contínuas do intervalo $[-1,1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Vamos calcular o ângulo entre t^3 e t .

$$\cos \theta = \frac{\langle t^3, t \rangle}{\|t^3\| \|t\|} = \frac{\int_{-1}^1 t^4 dt}{\left(\int_{-1}^1 t^6 dt\right)^{1/2} \left(\int_{-1}^1 t^2 dt\right)^{1/2}} = \frac{2/5}{\sqrt{2/7} \sqrt{2/3}} = \frac{\sqrt{21}}{5}.$$

Assim o ângulo entre t^3 e t é

$$\theta = \arccos \frac{\sqrt{21}}{5} \approx 23^\circ.$$

2.1.3 Ortogonalidade

Vamos, agora, estender a espaços com produto interno o conceito de ortogonalidade.

Definição 2.4. Seja \mathbb{V} um espaço vetorial com produto interno.

- (a) Dizemos que um subconjunto não vazio \mathcal{X} de \mathbb{V} é **ortogonal** se para todo par V e W de elementos distintos de \mathcal{X} , $\langle V, W \rangle = 0$. Neste caso dizemos que os elementos de \mathcal{X} são **vetores ortogonais**.

- (b) Dizemos que um subconjunto não vazio \mathcal{X} de \mathbb{V} é **ortonormal** se \mathcal{X} é um conjunto ortogonal consistindo de vetores **unitários**, isto é, vetores cuja norma é igual a 1.
-

Exemplo 2.8. Considere o \mathbb{R}^3 com o produto interno usual

$$\langle (v_1, v_2, v_3), (w_1, w_2, w_3) \rangle = v_1 w_1 + v_2 w_2 + v_3 w_3.$$

O conjunto

$$\{W_1 = (1, 1, 1), W_2 = (-1, 1, 0), W_3 = (-1, -1, 2)\}$$

de \mathbb{R}^3 é ortogonal (verifique!).

$$\|W_1\|^2 = \langle W_1, W_1 \rangle = 1^2 + 1^2 + 1^2 = 3,$$

$$\|W_2\|^2 = \langle W_2, W_2 \rangle = (-1)^2 + 1^2 + 0^2 = 2,$$

$$\|W_3\|^2 = \langle W_3, W_3 \rangle = (-1)^2 + (-1)^2 + 2^2 = 6.$$

“Dividindo” cada vetor pela sua norma obtemos

$$U_1 = \frac{1}{\|W_1\|} W_1 = \frac{1}{\sqrt{3}} (1, 1, 1) = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right),$$

$$U_2 = \frac{1}{\|W_2\|} W_2 = \frac{1}{\sqrt{2}} (-1, 1, 0) = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right),$$

$$U_3 = \frac{1}{\|W_3\|} W_3 = \frac{1}{\sqrt{6}} (-1, -1, 2) = \left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right)$$

que formam um conjunto ortonormal.

Exemplo 2.9. Seja L um número real maior que zero. Seja $\mathbb{V} = C^0[-L, L]$ o conjunto das funções contínuas do intervalo $[-L, L]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-L}^L f(t)g(t)dt.$$

Vamos mostrar que o conjunto

$$\left\{ 1, \cos \frac{\pi t}{L}, \sin \frac{\pi t}{L}, \cos \frac{2\pi t}{L}, \sin \frac{2\pi t}{L}, \dots, \cos \frac{n\pi t}{L}, \sin \frac{n\pi t}{L}, \dots \right\}$$

é ortogonal. Como as funções do conjunto, exceto a primeira, são funções cujas primitivas são periódicas de período igual a $2L/n$, então a integral de $-L$ a L destas funções é igual a zero e portanto elas são ortogonais à função constante 1.

$$\begin{aligned} \left\langle \cos \frac{n\pi t}{L}, \sin \frac{m\pi t}{L} \right\rangle &= \int_{-L}^L \cos \frac{n\pi t}{L} \sin \frac{m\pi t}{L} dt = \frac{L}{\pi} \int_{-\pi}^{\pi} \cos ns \sin ms ds \\ &= \frac{L}{2\pi} \int_{-\pi}^{\pi} [\sin(m+n)s + \sin(m-n)s] ds = 0 \end{aligned}$$

Para $m \neq n$ temos que

$$\begin{aligned} \left\langle \cos \frac{n\pi t}{L}, \cos \frac{m\pi t}{L} \right\rangle &= \int_{-L}^L \cos \frac{n\pi t}{L} \cos \frac{m\pi t}{L} dt = \frac{L}{\pi} \int_{-\pi}^{\pi} \cos ns \cos ms ds \\ &= \frac{L}{2\pi} \int_{-\pi}^{\pi} [\cos(m+n)s + \cos(m-n)s] ds \\ &= \frac{L}{2\pi(m+n)} \sin(m+n)s \Big|_{-\pi}^{\pi} + \frac{L}{2\pi(m-n)} \sin(m-n)s \Big|_{-\pi}^{\pi} = 0, \\ \left\langle \sin \frac{n\pi t}{L}, \sin \frac{m\pi t}{L} \right\rangle &= \int_{-L}^L \sin \frac{n\pi t}{L} \sin \frac{m\pi t}{L} dt = \frac{L}{\pi} \int_{-\pi}^{\pi} \sin ns \sin ms ds \\ &= \frac{L}{2\pi} \int_{-\pi}^{\pi} [-\cos(m+n)s + \cos(m-n)s] ds = 0 \end{aligned}$$

Vamos calcular as normas dos elementos do conjunto.

$$\begin{aligned} \|1\|^2 &= \langle 1, 1 \rangle = \int_{-L}^L dt = 2L \\ \left\| \cos \frac{n\pi t}{L} \right\|^2 &= \left\langle \cos \frac{n\pi t}{L}, \cos \frac{n\pi t}{L} \right\rangle = \int_{-L}^L \cos^2 \frac{n\pi t}{L} dt \\ &= \frac{L}{\pi} \int_{-\pi}^{\pi} \cos^2 ns ds = \frac{L}{2\pi} \int_{-\pi}^{\pi} [1 + \cos 2ns] ds = L \\ \left\| \sin \frac{n\pi t}{L} \right\|^2 &= \left\langle \sin \frac{n\pi t}{L}, \sin \frac{n\pi t}{L} \right\rangle = \int_{-L}^L \sin^2 \frac{n\pi t}{L} dt \\ &= \frac{L}{\pi} \int_{-\pi}^{\pi} \sin^2 ns ds = \frac{L}{2\pi} \int_{-\pi}^{\pi} [1 - \cos 2ns] ds = L \end{aligned}$$

Assim o conjunto

$$\left\{ \frac{1}{\sqrt{2L}}, \frac{1}{\sqrt{L}} \cos \frac{\pi t}{L}, \frac{1}{\sqrt{L}} \sin \frac{\pi t}{L}, \frac{1}{\sqrt{L}} \cos \frac{2\pi t}{L}, \frac{1}{\sqrt{L}} \sin \frac{2\pi t}{L}, \dots, \frac{1}{\sqrt{L}} \cos \frac{n\pi t}{L}, \frac{1}{\sqrt{L}} \sin \frac{n\pi t}{L}, \dots \right\}$$

é ortonormal.

Exemplo 2.10. Seja L um número real maior que zero. Seja $\mathbb{V} = \mathcal{C}^0[0, L]$ o conjunto das funções contínuas do intervalo $[0, L]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_0^L f(t)g(t)dt.$$

Vamos mostrar que os conjuntos

$$\left\{ 1, \cos \frac{\pi t}{L}, \cos \frac{2\pi t}{L}, \dots, \cos \frac{n\pi t}{L}, \dots \right\} \quad \text{e} \quad \left\{ \sin \frac{\pi t}{L}, \sin \frac{2\pi t}{L}, \dots, \sin \frac{n\pi t}{L}, \dots \right\}$$

são ortogonais.

$$\left\langle 1, \cos \frac{n\pi t}{L} \right\rangle = \int_0^L \cos \frac{n\pi t}{L} dt = \frac{L}{\pi} \int_0^\pi \cos ns ds = \frac{L}{n\pi} \text{sen } ns \Big|_0^\pi = 0$$

Para $m \neq n$ temos que

$$\begin{aligned} \left\langle \cos \frac{n\pi t}{L}, \cos \frac{m\pi t}{L} \right\rangle &= \int_0^L \cos \frac{n\pi t}{L} \cos \frac{m\pi t}{L} dt = \frac{L}{\pi} \int_0^\pi \cos ns \cos ms ds \\ &= \frac{L}{2\pi} \int_0^\pi [\cos(m+n)s + \cos(m-n)s] ds \\ &= \frac{L}{2\pi(m+n)} \text{sen}(m+n)s \Big|_0^\pi + \frac{L}{2\pi(m-n)} \text{sen}(m-n)s \Big|_0^\pi = 0, \\ \left\langle \text{sen} \frac{n\pi t}{L}, \text{sen} \frac{m\pi t}{L} \right\rangle &= \int_0^L \text{sen} \frac{n\pi t}{L} \text{sen} \frac{m\pi t}{L} dt = \frac{L}{\pi} \int_0^\pi \text{sen } ns \text{sen } ms ds \\ &= \frac{L}{2\pi} \int_0^\pi [-\cos(m+n)s + \cos(m-n)s] ds = 0 \end{aligned}$$

Vamos calcular as normas dos elementos dos conjuntos.

$$\begin{aligned} \|1\|^2 &= \langle 1, 1 \rangle = \int_0^L dt = L \\ \left\| \cos \frac{n\pi t}{L} \right\|^2 &= \left\langle \cos \frac{n\pi t}{L}, \cos \frac{n\pi t}{L} \right\rangle = \int_0^L \cos^2 \frac{n\pi t}{L} dt \\ &= \frac{L}{\pi} \int_0^\pi \cos^2 ns ds = \frac{L}{2\pi} \int_0^\pi [1 + \cos 2ns] ds = L/2 \\ \left\| \text{sen} \frac{n\pi t}{L} \right\|^2 &= \left\langle \text{sen} \frac{n\pi t}{L}, \text{sen} \frac{n\pi t}{L} \right\rangle = \int_0^L \text{sen}^2 \frac{n\pi t}{L} dt \\ &= \frac{L}{\pi} \int_0^\pi \text{sen}^2 ns ds = \frac{L}{2\pi} \int_0^\pi [1 - \cos 2ns] ds = L/2 \end{aligned}$$

Assim os conjuntos

$$\left\{ \frac{1}{\sqrt{L}}, \sqrt{\frac{2}{L}} \cos \frac{\pi t}{L}, \sqrt{\frac{2}{L}} \cos \frac{2\pi t}{L}, \dots, \sqrt{\frac{2}{L}} \cos \frac{n\pi t}{L}, \dots \right\} \text{ e}$$

$$\left\{ \sqrt{\frac{2}{L}} \operatorname{sen} \frac{\pi t}{L}, \sqrt{\frac{2}{L}} \operatorname{sen} \frac{2\pi t}{L}, \dots, \sqrt{\frac{2}{L}} \operatorname{sen} \frac{n\pi t}{L}, \dots \right\}$$

são ortonormais.

Exemplo 2.11. Seja L um número real maior que zero. Seja $\mathbb{V} = C^0([-L, L], \mathbb{C})$ o espaço vetorial das funções contínuas do intervalo $[-L, L]$ com valores complexos munido do produto interno

$$\langle f, g \rangle = \int_{-L}^L f(t) \overline{g(t)} dt.$$

Vamos mostrar que o conjunto

$$\left\{ e^{\frac{in\pi t}{L}} = \cos \frac{n\pi t}{L} + i \operatorname{sen} \frac{n\pi t}{L} \mid n \in \mathbb{Z} \right\}$$

é ortogonal.

$$\begin{aligned} \left\langle e^{\frac{in\pi t}{L}}, e^{\frac{im\pi t}{L}} \right\rangle &= \int_{-L}^L e^{\frac{in\pi t}{L}} \overline{e^{\frac{im\pi t}{L}}} dt = \frac{L}{\pi} \int_{-\pi}^{\pi} e^{ins} \overline{e^{ims}} ds = \frac{L}{\pi} \int_{-\pi}^{\pi} e^{i(n-m)s} ds \\ &= \frac{L}{\pi i(n-m)} e^{i(n-m)s} \Big|_{-\pi}^{\pi} = 0, \quad \text{se } n \neq m. \end{aligned}$$

Vamos calcular as normas dos elementos do conjunto.

$$\|e^{\frac{in\pi t}{L}}\|^2 = \left\langle e^{\frac{in\pi t}{L}}, e^{\frac{in\pi t}{L}} \right\rangle = \int_{-L}^L e^{\frac{in\pi t}{L}} \overline{e^{\frac{in\pi t}{L}}} dt = \frac{L}{\pi} \int_{-\pi}^{\pi} e^{ins} \overline{e^{ins}} ds = \frac{L}{\pi} \int_{-\pi}^{\pi} 1 ds = 2L.$$

Assim o conjunto

$$\left\{ \frac{1}{\sqrt{2L}} e^{\frac{in\pi t}{L}} = \frac{1}{\sqrt{2L}} \cos \frac{n\pi t}{L} + i \frac{1}{\sqrt{2L}} \operatorname{sen} \frac{n\pi t}{L} \mid n \in \mathbb{Z} \right\}$$

é ortonormal

2.1.4 Projeção Ortogonal

Em um espaço com produto interno podemos definir a **projeção ortogonal** de um vetor V sobre um vetor não nulo W por

$$\operatorname{proj}_W V = \left(\frac{\langle V, W \rangle}{\|W\|^2} \right) W.$$

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 2.4. *Seja \mathbb{V} um espaço vetorial com produto interno. Seja $W \in \mathbb{V}$ um vetor não nulo. Então, $V - \operatorname{proj}_W V$ é ortogonal a W , para todo vetor $V \in \mathbb{V}$.*

Demonstração. Precisamos calcular o produto escalar de $V - \text{proj}_W V$ com W :

$$\langle V - \text{proj}_W V, W \rangle = \langle V, W \rangle - \langle \text{proj}_W V, W \rangle = \langle V, W \rangle - \left(\frac{\langle V, W \rangle}{\|W\|^2} \right) \langle W, W \rangle = 0,$$

pois por definição $\|W\| = \sqrt{\langle W, W \rangle}$. Portanto, $V - \text{proj}_W V$ é ortogonal a W . ■

Observe que se W é um vetor não nulo, então para todo vetor V ,

$$V = (V - \text{proj}_W V) + \text{proj}_W V.$$

Assim, pela **Proposição 2.4**, V pode ser escrito como uma soma de dois vetores ortogonais:

$$V - \text{proj}_W V \quad \text{e} \quad \text{proj}_W V,$$

sendo o segundo vetor um múltiplo escalar de W .

Exemplo 2.12. Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Vamos determinar a projeção de t^3 em t .

$$(\text{proj}_t t^3)(t) = \frac{\langle t^3, t \rangle}{\|t\|^2} t = \frac{\langle t^3, t \rangle}{\langle t, t \rangle} t = \frac{\int_{-1}^1 t^4 dt}{\int_{-1}^1 t^2 dt} t = \frac{2/5}{2/3} t = \frac{3}{5} t$$

Podemos escrever t^3 como uma soma de duas funções ortogonais:

$$t^3 = \left(t^3 - \frac{3}{5} t\right) + \frac{3}{5} t,$$



Figura 2.1: **Projeção ortogonal do vetor V sobre o vetor W**

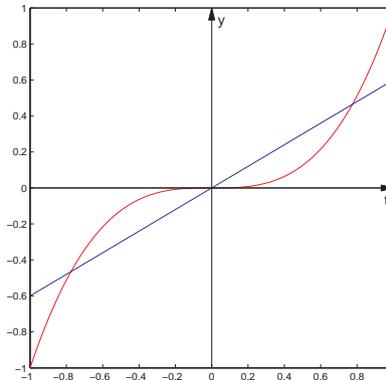


Figura 2.2: t^3 e a projeção de t^3 em t , $(\text{proj}_t t^3)(t) = \frac{3}{5}t$

sendo que $\frac{3}{5}t$ é um múltiplo escalar de t .

Proposição 2.5 (Teorema de Pitágoras). *Seja \mathbb{V} um espaço com produto interno. Sejam $V, W \in \mathbb{V}$. Se V e W são ortogonais ($\langle V, W \rangle = 0$), então*

$$\|V + W\|^2 = \|V\|^2 + \|W\|^2.$$

Demonstração. Sejam $V, W \in \mathbb{V}$.

$$\|V + W\|^2 = \langle V + W, V + W \rangle = \langle V, V \rangle + \langle V, W \rangle + \langle W, V \rangle + \langle W, W \rangle = \|V\|^2 + \|W\|^2,$$

pois $\langle V, W \rangle = \langle W, V \rangle = 0$. ■

Exemplo 2.13. Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

As funções $f(t) = 1$ e $g(t) = t$ são ortogonais, pois

$$\langle f, g \rangle = \langle 1, t \rangle = \int_{-1}^1 t dt = 0$$

Vamos verificar o Teorema de Pitágoras para $f(t) = 1$ e $g(t) = t$.

$$\|f\|^2 + \|g\|^2 = \|1\|^2 + \|t\|^2 = \langle 1, 1 \rangle + \langle t, t \rangle = \int_{-1}^1 dt + \int_{-1}^1 t^2 dt = 2 + \frac{2}{3} = \frac{8}{3}$$

$$\|f + g\|^2 = \|1 + t\|^2 = \langle 1 + t, 1 + t \rangle = \int_{-1}^1 (1 + t)^2 dt = (1/3)(1 + t)^3 \Big|_{-1}^1 = \frac{8}{3}$$

Proposição 2.6. *Seja \mathbb{V} um espaço vetorial com produto interno. Seja $W \in \mathbb{V}$ um vetor não nulo. Então a projeção de V em W é o múltiplo escalar de W que é mais “próximo” de V no sentido de que é a solução do problema*

$$\min_{X=\alpha W} \|V - X\|.$$

Demonstração. Seja X um múltiplo escalar qualquer de W . Pelo Teorema de Pitágoras ([Proposição 2.5](#)) temos que

$$\|V - X\|^2 = \|(V - \text{proj}_W V) + (\text{proj}_W V - X)\|^2 = \|V - \text{proj}_W V\|^2 + \|\text{proj}_W V - X\|^2, \quad (2.1)$$

pois $V - \text{proj}_W V$ é ortogonal a $\text{proj}_W V - X$ que é um múltiplo escalar de W .

Variando X como múltiplo escalar de W , vemos de (2.1) que o mínimo de $\|V - X\|^2$ ocorre somente para $X = \text{proj}_W V$, já que $\|\text{proj}_W V - V\|^2$ permanece fixo em (2.1) quando variamos X como múltiplo escalar de W . Portanto,

$$\min_{X=\alpha W} \|V - X\| = \|V - \text{proj}_W V\|.$$

■

Exemplo 2.14. Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

No [Exemplo 2.12 na página 117](#) vimos que

$$\text{proj}_t t^3 = \frac{3}{5} t$$

Assim, $\frac{3}{5} t$ é o múltiplo escalar de t que está mais “próximo” de t^3 no sentido de que $\|t^3 - \alpha t\| = \left(\int_{-1}^1 (t^3 - \alpha t)^2 dt \right)^{1/2}$ é mínimo.

Proposição 2.7. *Seja \mathbb{V} um espaço vetorial com produto interno. Seja \mathcal{X} um subconjunto de \mathbb{V} de vetores ortogonais não nulos.*

- (a) *Então o conjunto \mathcal{X} é linearmente independente (L.I.).*
- (b) *Se V pertence ao espaço de todas as combinações lineares de elementos de \mathcal{X} , ou seja, se $V \in [\mathcal{X}]$ (espaço gerado por \mathcal{X}), então existem $W_1, \dots, W_n \in \mathcal{X}$ tais que*

$$V = \sum_{i=1}^n \frac{\langle V, W_i \rangle}{\|W_i\|^2} W_i.$$

Demonstração. (a) Sejam V_1, \dots, V_k vetores quaisquer de \mathcal{X} . Considere a equação vetorial

$$x_1 V_1 + \dots + x_k V_k = \vec{0}. \quad (2.2)$$

Fazendo o produto escalar de ambos os membros de (2.2) com $V_i, i = 1, \dots, k$ e aplicando as propriedades do produto escalar, obtemos

$$x_1 \langle V_1, V_i \rangle + \dots + x_i \langle V_i, V_i \rangle + \dots + x_k \langle V_k, V_i \rangle = 0. \quad (2.3)$$

Mas, $\langle V_i, V_j \rangle = 0$, se $i \neq j$. Assim, de (2.3) obtemos que

$$x_i \|V_i\|^2 = 0.$$

Mas, como $V_i \neq \vec{0}$, então $\|V_i\| \neq 0$ e assim $x_i = 0$, para $i = 1, \dots, k$. Portanto o conjunto \mathcal{X} é L.I.

(b) Seja $V \in [\mathcal{X}]$. Então existem vetores $W_1, \dots, W_n \in \mathcal{X}$ e escalares $\alpha_1, \dots, \alpha_n$ tais que

$$V = \sum_{i=1}^n \alpha_i W_i. \quad (2.4)$$

Fazendo o produto escalar de V com W_j , para $j = 1, \dots, n$, obtemos que

$$\langle V, W_j \rangle = \left\langle \sum_{i=1}^n \alpha_i W_i, W_j \right\rangle = \sum_{i=1}^n \alpha_i \langle W_i, W_j \rangle = \alpha_j \|W_j\|^2.$$

Assim,

$$\alpha_j = \frac{\langle V, W_j \rangle}{\|W_j\|^2}, \quad \text{para } j = 1, \dots, n.$$

Substituindo-se este valor de α_j em (2.4) obtemos o resultado. ■

Exemplo 2.15. Vimos no Exemplo 2.8 na página 111 que o conjunto

$$\{W_1 = (1, 1, 1), W_2 = (-1, 1, 0), W_3 = (-1, -1, 2)\} \text{ de } \mathbb{R}^3$$

é ortogonal com relação ao produto interno usual

$$\langle (v_1, v_2, v_3), (w_1, w_2, w_3) \rangle = v_1 w_1 + v_2 w_2 + v_3 w_3.$$

Segue da Proposição que o conjunto $\{W_1, W_2, W_3\}$ é L.I. Como a dimensão do \mathbb{R}^3 é igual a 3, então este conjunto é uma base do \mathbb{R}^3 (Teorema 1.12 na página 75). Podemos assim escrever qualquer vetor $V = (x, y, z) \in \mathbb{R}^3$ facilmente como combinação linear de W_1, W_2 e W_3 . Por exemplo, se $V = (1, 2, 0)$, então

$$\langle (1, 2, 0), W_1 \rangle = 3, \quad \langle (1, 2, 0), W_2 \rangle = -1, \quad \langle (1, 2, 0), W_3 \rangle = -3.$$

$$\|W_1\|^2 = \langle W_1, W_1 \rangle = 3, \quad \|W_2\|^2 = \langle W_2, W_2 \rangle = 2, \quad \|W_3\|^2 = \langle W_3, W_3 \rangle = 6.$$

Assim,

$$\begin{aligned} V = (1, 2, 0) &= \frac{\langle (1, 2, 0), W_1 \rangle}{\|W_1\|^2} W_1 + \frac{\langle (1, 2, 0), W_2 \rangle}{\|W_2\|^2} W_2 + \frac{\langle (1, 2, 0), W_3 \rangle}{\|W_3\|^2} W_3 \\ &= 1 W_1 - \frac{1}{2} W_2 - \frac{1}{2} W_3 \end{aligned}$$

2.1.5 Coeficientes de Fourier

Definição 2.5. Seja \mathcal{X} um **subconjunto ortogonal** de um espaço vetorial com produto interno \mathbb{V} . Para todo $V \in \mathbb{V}$ chamamos de **coeficientes de Fourier** de V em relação a \mathcal{X} os escalares

$$\frac{\langle V, U_\lambda \rangle}{\|U_\lambda\|^2}, \quad \text{para } U_\lambda \in \mathcal{X}.$$

Exemplo 2.16. Seja L um número real maior que zero. Seja $\mathbb{V} = C^0([-1, 1], \mathbb{C})$ o espaço vetorial das funções contínuas do intervalo $[-1, 1]$ com valores complexos munido do produto interno

$$\langle f, g \rangle = \int_{-1}^1 f(t) \overline{g(t)} dt.$$

Vamos calcular os coeficientes de Fourier da função $f(t) = e^t$ em relação ao conjunto ortonormal

$$\mathcal{C} = \{e^{in\pi t} = \cos n\pi t + i \operatorname{sen} n\pi t \mid n \in \mathbb{Z}\}.$$

Os coeficientes de Fourier da função $f(t) = e^t$ são dados por

$$\begin{aligned} \frac{\langle e^t, 1 \rangle}{\|1\|^2} &= \frac{1}{2} \int_{-1}^1 e^t dt = \frac{e - e^{-1}}{2}, \\ \frac{\langle e^t, e^{in\pi t} \rangle}{\|e^{in\pi t}\|^2} &= \frac{1}{2} \int_{-1}^1 e^t \overline{e^{in\pi t}} dt = \frac{1}{2} \int_{-1}^1 e^{(1-in\pi)t} dt = \frac{1}{2} \frac{1}{(1-in\pi)} e^{(1-in\pi)t} \Big|_{-1}^1 \\ &= \frac{1}{2} \frac{1}{(1-in\pi)} \left(e^{(1-in\pi)} - e^{-(1-in\pi)} \right) = \frac{1}{2} \frac{(-1)^n (e - e^{-1})(1 + n\pi i)}{1 + n^2 \pi^2}, \quad \text{para } n \neq 0. \end{aligned}$$

Exemplo 2.17. Existe uma relação entre os coeficientes de Fourier de uma função $f \in C^0[-L, L]$ em relação ao conjunto

$$\mathcal{B} = \left\{ 1, \cos \frac{\pi t}{L}, \operatorname{sen} \frac{\pi t}{L}, \cos \frac{2\pi t}{L}, \operatorname{sen} \frac{2\pi t}{L}, \dots, \cos \frac{n\pi t}{L}, \operatorname{sen} \frac{n\pi t}{L}, \dots \right\}$$

do [Exemplo 2.9 na página 112](#) e os coeficientes de Fourier da função f em relação ao conjunto

$$\mathcal{C} = \left\{ e^{\frac{in\pi t}{L}} = \cos \frac{n\pi t}{L} + i \operatorname{sen} \frac{n\pi t}{L} \mid n \in \mathbb{Z} \right\}$$

do Exemplo 2.11 na página 115.

Sejam $c_n = \frac{\langle f, e^{\frac{i n \pi t}{L}} \rangle}{\|e^{\frac{i n \pi t}{L}}\|^2}$ os coeficientes de f em relação a \mathcal{B} ,
 $a_0 = \frac{\langle f, 1 \rangle}{\|1\|^2}$, $a_n = \frac{\langle f, \cos n \pi t \rangle}{\|\cos n \pi t\|^2}$ e $b_n = \frac{\langle f, \sin n \pi t \rangle}{\|\sin n \pi t\|^2}$ os coeficientes de Fourier de f em relação a \mathcal{C} . No Exemplo 2.9 na página 112 vimos que

$$\|1\|^2 = 2L, \quad \|\cos \frac{n \pi t}{L}\|^2 = L \quad \text{e} \quad \|\sin \frac{n \pi t}{L}\|^2 = L$$

e no Exemplo 2.11 na página 115 vimos que $\|e^{\frac{i n \pi t}{L}}\|^2 = 2L$. Assim,

$$c_0 = \frac{\langle f, 1 \rangle}{\|1\|^2} = a_0 \quad (2.5)$$

$$\begin{aligned} c_n &= \frac{\langle f, e^{\frac{i n \pi t}{L}} \rangle}{\|e^{\frac{i n \pi t}{L}}\|^2} = \frac{1}{2L} \langle f, \cos n \pi t + i \sin n \pi t \rangle \\ &= \frac{1}{2L} \langle f, \cos n \pi t \rangle - i \frac{1}{2L} \langle f, \sin n \pi t \rangle \\ &= \frac{1}{2} \left(\frac{\langle f, \cos n \pi t \rangle}{\|\cos \frac{n \pi t}{L}\|^2} - i \frac{\langle f, \sin n \pi t \rangle}{\|\sin \frac{n \pi t}{L}\|^2} \right) = \frac{1}{2} (a_n - i b_n), \quad n \neq 0 \end{aligned} \quad (2.6)$$

$$c_{-n} = \frac{1}{2} (a_n + i b_n) = \bar{c}_n, \quad n > 0. \quad (2.7)$$

Assim, de (2.5) e (2.6) segue-se que

$$a_0 = c_0, \quad a_n = 2\Re(c_n) \quad \text{e} \quad b_n = -2\Im(c_n), \quad \text{para } n > 0. \quad (2.8)$$

Ou seja, $a_0 = c_0$, a_n é igual a 2 vezes a parte real de c_n e b_n é igual a -2 vezes a parte imaginária de c_n , para $n > 0$.

Exemplo 2.18. Segue das relações (2.8) e do Exemplo 2.16 que os coeficientes de Fourier de $f(t) = e^t \in \mathcal{C}^0[-1, 1]$ em relação ao conjunto

$$\{1, \cos \pi t, \text{sen } \pi t, \cos 2\pi t, \text{sen } 2\pi t, \dots, \cos n\pi t, \text{sen } n\pi t, \dots\}$$

são dados por

$$a_0 = \langle e^t, 1 \rangle = c_0 = \frac{1}{2}(e - e^{-1}),$$

$$a_n = \langle e^t, \cos n\pi t \rangle = 2\Re(c_n) = \frac{(-1)^n(e - e^{-1})}{1 + n^2\pi^2}$$

$$b_n = \langle e^t, \text{sen } n\pi t \rangle = -2\Im(c_n) = \frac{(-1)^{n+1}(e - e^{-1})n\pi}{1 + n^2\pi^2}$$

Exercícios Numéricos (respostas na página 427)

- 2.1.1. Sejam $X = (1, 1, -2)$ e $Y = (a, -1, 2)$. Para quais valores de a , X e Y são ortogonais em relação ao produto interno usual de \mathbb{R}^3 ?
- 2.1.2. 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 $\{X, Y\}$ é ortonormal em relação ao produto interno usual de \mathbb{R}^3 ?

2.1.3. Considere o conjunto

$$\mathcal{B} = \{W_1 = (1, -1, 1), W_2 = (1, 1, 0), W_3 = (-1, 1, 2)\}.$$

- (a) Mostre que o conjunto \mathcal{B} é ortogonal em relação ao produto interno usual de \mathbb{R}^3 . Este conjunto é uma base de \mathbb{R}^3 ? Por que?
- (b) Determine um conjunto ortonormal de vetores $\mathcal{C} = \{U_1, U_2, U_3\}$ formado por múltiplos escalares dos vetores do conjunto \mathcal{B} .
- (c) Escreva o vetor $V = (2, 1, 0)$ como combinação linear de W_1, W_2 e W_3 .
- (d) Escreva o vetor $V = (x, y, z)$ de \mathbb{R}^3 como combinação linear de W_1, W_2 e W_3 .
- 2.1.4. Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

- (a) Mostre que o conjunto $\{1, t, t^2, \dots, t^n, \dots\}$ não é ortogonal.
- (b) Mostre que $f_1(t) = t$ e $f_2(t) = t^2$ são ortogonais e verifique a validade do Teorema de Pitágoras para $f_1(t) = t$ e $f_2(t) = t^2$.
- (c) Calcule a projeção de t^5 em t . Faça um esboço dos gráficos de t^5 e $\text{proj}_t t^5$ no mesmo sistema de coordenadas.
- (d) Escreva t^5 como uma soma de dois polinômios ortogonais $p_1(t)$ e $p_2(t)$ sendo $p_1(t)$ um múltiplo escalar de t .

(e) Determine o ângulo θ entre t e t^5 .

2.1.5. Considere o subconjunto $\mathcal{B} = \{1, t, t^2 - \frac{1}{3}\}$ do espaço $\mathbb{V} = \mathcal{C}^0[-1, 1]$ das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} .

(a) Mostre que o conjunto \mathcal{B} é ortogonal em relação ao produto interno

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Este conjunto é uma base do subespaço \mathcal{P}_2 dos polinômios de grau menor ou igual a 2? Por que?

(b) Determine um conjunto ortonormal de funções $\mathcal{C} = \{q_0, q_1, q_2\}$ formado por múltiplos escalares dos vetores do conjunto \mathcal{B} .

(c) Escreva o polinômio $1 + t + t^2$ como combinação linear dos elementos de \mathcal{B} .

(d) Escreva o polinômio $a + bt + ct^2$ como combinação linear dos elementos de \mathcal{B} .

Exercícios usando o MATLAB[®]

Comandos do MATLAB[®]:

```
>> syms t diz ao MATLAB® que a variável t é uma variável simbólica.
```

```
>> f=expr define uma função através da expr que deve ser uma expressão na variável simbólica t definida anteriormente.
```

Comandos do pacote GAAL:

>> `proj(g, f, a, b)` calcula a projeção de $f(t)$ em $g(t)$ com relação ao produto interno

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt.$$

>> `plotfproj(f, prj, a, b)` desenha as funções f e $prj(k)$, para k variando de 1 até o tamanho do vetor `prj`, no intervalo $[a, b]$.

2.1.6. Seja $\mathbb{V} = \mathcal{C}^0([-1, 1], \mathbb{C})$ com o produto interno

$$\langle f, g \rangle = \int_{-1}^1 f(t)\overline{g(t)}dt.$$

Seja $\mathcal{C}_3 = \{e^{int} \mid n = 0, 1, 2, 3\}$. Calcule os coeficientes de Fourier da função $f(t) = e^t - \alpha e^{2t}$, ($\alpha = \frac{e - e^{-1}}{e^2 - e^{-2}}$) em relação a \mathcal{C}_3 .

2.1.7. Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ com o produto interno

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Seja $\mathcal{B}_3 = \{1, \cos \pi t, \sin \pi t, \cos 2\pi t, \sin 2\pi t, \cos 3\pi t, \sin 3\pi t\}$. Calcule os coeficientes de Fourier da função $f(t) = e^t - \alpha e^{2t}$, ($\alpha = \frac{e - e^{-1}}{e^2 - e^{-2}}$) em relação a \mathcal{B} .

Exercícios Teóricos

2.1.8. Mostre que se $\langle X, V \rangle = \langle X, W \rangle$ para todo vetor X , então $V = W$.

2.1.9. Mostre que se V é ortogonal a W_1, \dots, W_k , então V é ortogonal a qualquer combinação linear de W_1, \dots, W_k .

2.1.10. Mostre que o conjunto de todos os vetores de \mathbb{V} (espaço com produto interno) ortogonais a um dado vetor V , $\{X \in \mathbb{V} \mid \langle X, V \rangle = 0\}$ é um subespaço de \mathbb{V} .

2.1.11. Sejam V e W vetores de um espaço com produto interno \mathbb{V} . Prove as **identidades polares**:

(a) $\langle V, W \rangle = \frac{1}{4} (\|V + W\|^2 - \|V - W\|^2)$, se o conjunto de escalares é o conjunto dos números reais.

(b) $\langle V, W \rangle = \frac{1}{4} (\|V + W\|^2 - \|V - W\|^2 + i\|V + iW\|^2 - i\|V - iW\|^2)$, se o conjunto de escalares é o conjunto dos números complexos.

(Sugestão: desenvolva os segundos membros das igualdades acima observando que $\|V + W\|^2 = \langle V + W, V + W \rangle$ e $\|V - W\|^2 = \langle V - W, V - W \rangle$)

2.1.12. Prove a **lei do paralelogramo** para vetores V e W de um espaço com produto interno \mathbb{V} :

$$\|V + W\|^2 + \|V - W\|^2 = 2 (\|V\|^2 + \|W\|^2).$$

(Sugestão: desenvolva o primeiro membro da igualdade acima observando que $\|V + W\|^2 = \langle V + W, V + W \rangle$ e $\|V - W\|^2 = \langle V - W, V - W \rangle$)

2.1.13. Seja $\{U_1, \dots, U_n\}$ uma base ortonormal de \mathbb{R}^n . Se $A = [U_1 \dots U_n]$ é uma matriz $n \times n$ cujas colunas são os vetores U_1, \dots, U_n , então A é invertível e $A^{-1} = A^t$. (Sugestão: mostre que $A^t A = I_n$.)

2.1.14. (Identidade de Parseval) Seja \mathcal{X} um subconjunto ortonormal de um espaço com produto interno \mathbb{V} . Se $V, W \in [\mathcal{X}]$, então existem vetores $U_1, \dots, U_n \in \mathcal{X}$ tais que

$$\langle V, W \rangle = \sum_{k=1}^n \langle V, U_k \rangle \langle U_k, W \rangle.$$

(Sugestão: use a [Proposição 2.7 na página 122.](#))

2.2 Bases Ortonormais e Subespaços Ortogonais

2.2.1 Bases Ortonormais

Seja \mathbb{V} um espaço vetorial com produto interno. Dizemos que \mathcal{B} é uma **base ortogonal**, se \mathcal{B} é uma base de \mathbb{V} que é ortogonal. Dizemos que \mathcal{B} é uma **base ortonormal**, se \mathcal{B} é uma base de \mathbb{V} que é ortonormal.

Exemplo 2.19. A **base canônica do \mathbb{R}^n** , que é formada pelos vetores $E_1 = (1, 0, \dots, 0), E_2 = (0, 1, 0, \dots, 0), \dots, E_n = (0, \dots, 0, 1)$ é uma base ortonormal do \mathbb{R}^n com relação ao produto interno usual

$$\langle (v_1, \dots, v_n), (w_1, \dots, w_n) \rangle = v_1 w_1 + \dots + v_n w_n.$$

Vamos mostrar a seguir, que a partir de uma base qualquer de um espaço com produto interno podemos encontrar uma base ortonormal com a propriedade de que o primeiro vetor da nova base seja um múltiplo escalar do primeiro vetor da base antiga. Nas [Figuras 2.4 e 2.5](#) vemos como isto é possível no caso em que o espaço é o \mathbb{R}^3 . Antes precisamos de um resultado que é uma generalização da [Proposição 2.4 na página 116](#).

Proposição 2.8. *Sejam W_1, W_2, \dots, W_n vetores ortogonais não nulos de \mathbb{V} , então para qualquer vetor V , o vetor $V - \text{proj}_{W_1} V - \dots - \text{proj}_{W_n} V$ é ortogonal a W_k , para $k = 1, \dots, n$.*

Demonstração. Vamos calcular o produto interno de $V - \text{proj}_{W_1} V - \dots - \text{proj}_{W_n} V$ com W_j , para $j = 1, \dots, n$, e ver que dá zero.

$$\left\langle V - \sum_{k=1}^n \text{proj}_{W_k} V, W_j \right\rangle = \langle V, W_j \rangle - \sum_{k=1}^n \frac{\langle V, W_k \rangle}{\|W_k\|^2} \langle W_k, W_j \rangle = \langle V, W_j \rangle - \frac{\langle V, W_j \rangle}{\|W_j\|^2} \langle W_j, W_j \rangle = 0,$$

pois $\langle W_k, W_j \rangle = 0$, se $j \neq k$.



Teorema 2.9. *Seja $\{V_1, \dots, V_n\}$ uma base de um espaço com produto interno \mathbb{V} . Então, existe uma base $\{U_1, \dots, U_n\}$ de \mathbb{V} que é ortonormal e tal que $U_1 = \left(\frac{1}{\|V_1\|}\right) V_1$.*

Demonstração. Usaremos o chamado **processo de ortogonalização de Gram-Schmidt** para construir uma base ortogonal. Depois “dividiremos” cada vetor da base encontrada pela sua norma, de forma a obtermos vetores ortogonais de norma igual a um.

(a) Sejam

$$\begin{aligned} W_1 &= V_1, \\ W_2 &= V_2 - \text{proj}_{W_1} V_2, \\ W_3 &= V_3 - \text{proj}_{W_1} V_3 - \text{proj}_{W_2} V_3, \\ &\dots \\ W_n &= V_n - \text{proj}_{W_1} V_n - \text{proj}_{W_2} V_n - \dots - \text{proj}_{W_{n-1}} V_n. \end{aligned}$$

Pela **Proposição 2.4 na página 116**, segue-se que W_2 é ortogonal a W_1 e $W_2 \neq \vec{0}$, 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, \dots, W_{n-1} seja uma base ortogonal do subespaço gerado por V_1, \dots, V_{n-1} segue-se da **Proposição 2.8** que W_n é ortogonal a W_1, \dots, W_{n-1} . $W_n \neq \vec{0}$, pois caso contrário, V_n pertenceria ao subespaço $[W_1, \dots, W_{n-1}] = [V_1, \dots, V_{n-1}]$. Como W_1, \dots, W_n são ortogonais não nulos, pela **Proposição 2.7 na página 122**, eles são L.I. e portanto formam uma base de \mathbb{V} .

(b) Sejam, agora

$$U_1 = \left(\frac{1}{\|W_1\|} \right) W_1, \quad U_2 = \left(\frac{1}{\|W_2\|} \right) W_2, \quad \dots, \quad U_n = \left(\frac{1}{\|W_n\|} \right) W_n.$$

Assim, $\{U_1, \dots, U_n\}$ é uma base ortonormal para o espaço \mathbb{V} .



Exemplo 2.20. Seja $\mathbb{V} = \mathbb{R}^3$ com o produto interno usual

$$\langle (v_1, v_2, v_3), (w_1, w_2, w_3) \rangle = v_1 w_1 + v_2 w_2 + v_3 w_3.$$

Considere a base formada pelos vetores $V_1 = (1, 1, 1)$, $V_2 = (0, 0, 1)$ e $V_3 = (1, 0, 0)$. Vamos encontrar uma base ortonormal para \mathbb{V} cujo primeiro vetor seja múltiplo escalar de V_1 . Sejam

$$W_1 = V_1 = (1, 1, 1)$$

$$W_2 = V_2 - \text{proj}_{W_1} V_2 = V_2 - \left(\frac{\langle V_2, W_1 \rangle}{\|W_1\|^2} \right) W_1 = (0, 0, 1) - \frac{1}{3}(1, 1, 1) = \left(-\frac{1}{3}, -\frac{1}{3}, \frac{2}{3}\right)$$

$$W_3 = V_3 - \text{proj}_{W_1} V_3 - \text{proj}_{W_2} V_3$$

$$= V_3 - \left(\frac{\langle V_3, W_1 \rangle}{\|W_1\|^2} \right) W_1 - \left(\frac{\langle V_3, W_2 \rangle}{\|W_2\|^2} \right) W_2$$

$$= (1, 0, 0) - \frac{1}{3}(1, 1, 1) - \frac{-1/3}{2/3} \left(-\frac{1}{3}, -\frac{1}{3}, \frac{2}{3}\right) = \left(\frac{1}{2}, -\frac{1}{2}, 0\right)$$

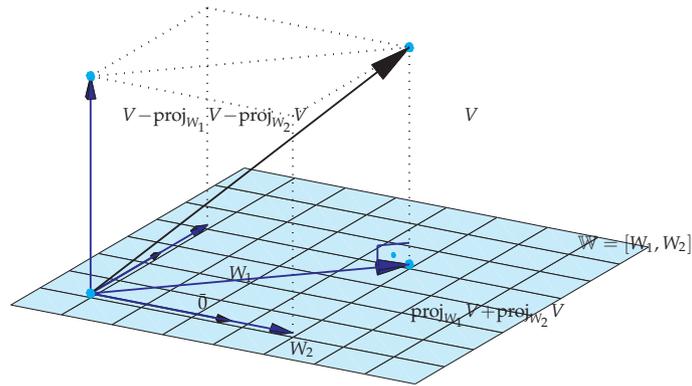


Figura 2.3: $V - \text{proj}_{W_1} V - \text{proj}_{W_2} V$ é ortogonal a W_1 e a W_2

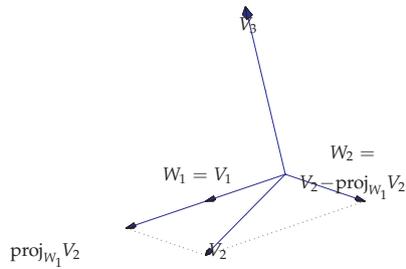


Figura 2.4: $W_1 = V_1$ e $W_2 = V_2 - \text{proj}_{W_1} V_2$

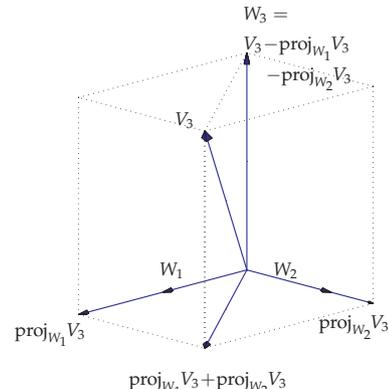


Figura 2.5: $W_3 = V_3 - \text{proj}_{W_1} V_3 - \text{proj}_{W_2} V_3$

Como $\|W_1\| = \sqrt{3}$, $\|W_2\| = \sqrt{2/3}$, $\|W_3\| = 1/\sqrt{2}$, temos que

$$\begin{aligned}u_1 &= \frac{1}{\|W_1\|} W_1 = \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right), \\u_2 &= \frac{1}{\|W_2\|} W_2 = \left(-\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right) \\u_3 &= \frac{1}{\|W_3\|} W_3 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0\right).\end{aligned}$$

Exemplo 2.21. Seja $\mathbb{V} = \mathcal{C}^0[0, 4]$ o conjunto das funções contínuas do intervalo $[0, 4]$ em \mathbb{R} com o produto interno definido por $\langle f, g \rangle = \int_0^4 f(t)g(t)dt$. Sabemos que $\{1, t, t^2, \dots, t^n\}$ é uma base de \mathcal{P}_n , o espaço dos polinômios de grau menor ou igual a n , para $n = 0, 1, 2, \dots$. Para cada n , \mathcal{P}_n pode ser visto como um subespaço de \mathbb{V} .

Vamos determinar uma base ortonormal para \mathcal{P}_2 . Para isto vamos aplicar o processo

de ortogonalização de Gram-Schmidt à base $\{1, t, t^2\}$.

$$p_0(t) = 1,$$

$$\begin{aligned} p_1(t) &= t - \left(\text{proj}_{p_0} t\right)(t) = t - \frac{\langle t, p_0 \rangle}{\|p_0\|^2} p_0(t) = t - \frac{\int_0^4 t p_0(t) dt}{\int_0^4 (p_0(t))^2 dt} p_0(t) \\ &= t - \frac{\int_0^4 t dt}{\int_0^4 dt} = t - \frac{8}{4} = t - 2 \end{aligned}$$

$$\begin{aligned} p_2(t) &= t^2 - \left(\text{proj}_{p_0} t^2\right)(t) - \left(\text{proj}_{p_1} t^2\right)(t) = t^2 - \frac{\langle t^2, p_0 \rangle}{\|p_0\|^2} p_0(t) - \frac{\langle t^2, p_1 \rangle}{\|p_1\|^2} p_1(t) \\ &= t^2 - \frac{\int_0^4 t^2 p_0(t) dt}{\int_0^4 (p_0(t))^2 dt} p_0(t) - \frac{\int_0^4 t^2 p_1(t) dt}{\int_0^4 (p_1(t))^2 dt} p_1(t) \\ &= t^2 - \frac{\int_0^4 t^2 dt}{\int_0^4 dt} - \frac{\int_0^4 t^2 (t-2) dt}{\int_0^4 (t-2)^2 dt} (t-2) \\ &= t^2 - \frac{64/3}{4} - \frac{64/3}{16/3} (t-2) = t^2 - 16/3 - 4(t-2) = t^2 - 4t + 8/3 = (t-2)^2 - 4/3 \end{aligned}$$

Vamos calcular as normas dos polinômios encontrados

$$\|p_0\|^2 = \langle p_0, p_0 \rangle = \int_0^4 (p_0(t))^2 dt = \int_0^4 dt = 4$$

$$\|p_1\|^2 = \langle p_1, p_1 \rangle = \int_0^4 (p_1(t))^2 dt = \int_0^4 (t-2)^2 dt = (1/3)(t-2)^3 \Big|_0^4 = 16/3$$

$$\begin{aligned} \|p_2\|^2 &= \langle p_2, p_2 \rangle = \int_0^4 (p_2(t))^2 dt = \int_0^4 [(t-2)^2 - 4/3]^2 dt \\ &= (1/5)(t-2)^5 \Big|_0^4 - (8/9)(t-2)^3 \Big|_0^4 + (16/9)t \Big|_0^4 = 256/45 \end{aligned}$$

Assim os polinômios

$$\begin{aligned}q_0(t) &= \frac{1}{\|p_0\|} p_0(t) = \frac{1}{2} \\q_1(t) &= \frac{1}{\|p_1\|} p_1(t) = \frac{\sqrt{3}}{4} (t - 2) \\q_2(t) &= \frac{1}{\|p_2\|} p_2(t) = \frac{3\sqrt{5}}{16} (t^2 - 4t + 8/3)\end{aligned}$$

são tais que $\{q_0(t) = 1/2, q_1(t) = (\sqrt{3}/4)(t - 2), q_2(t) = (3\sqrt{5}/16)(t^2 - 4t + 8/3)\}$ é uma base ortonormal de \mathcal{P}_2 .

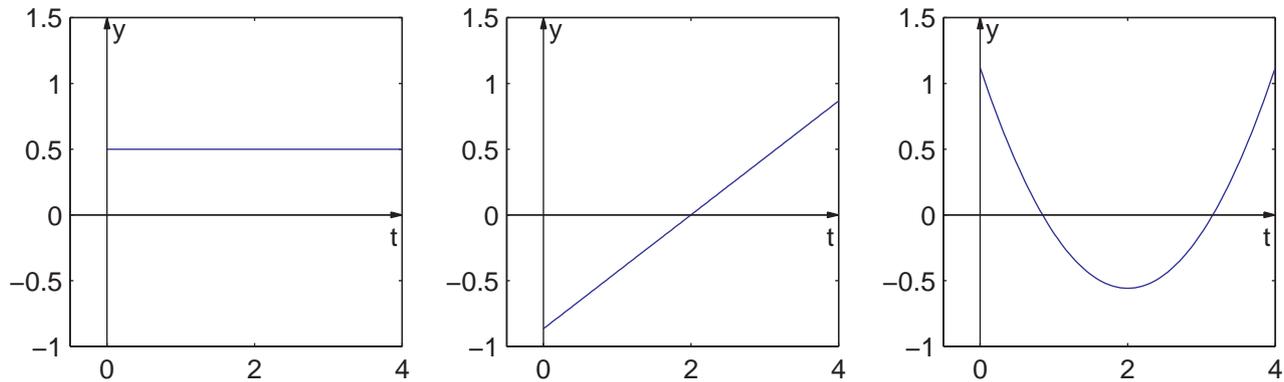


Figura 2.6: Polinômios ortonormais em relação ao produto interno $\langle f, g \rangle = \int_0^4 f(t)g(t)dt$ obtidos aplicando-se o processo de ortogonalização de Gram-Schmidt aos polinômios $1, t, t^2$

2.2.2 Aplicação: Polinômios de Legendre

Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Vamos aplicar o processo de ortogonalização de Gram-Schmidt à base $\mathcal{B}_n = \{1, t, \dots, t^n\}$ do subespaço das funções polinomiais de grau menor ou igual a n , que podemos identificar com \mathcal{P}_n .

$$p_0(t) = 1,$$

$$\begin{aligned} p_1(t) &= t - \left(\text{proj}_{p_0} t\right)(t) = t - \frac{\langle t, p_0 \rangle}{\|p_0\|^2} p_0(t) = t - \frac{\int_{-1}^1 t p_0(t) dt}{\int_{-1}^1 (p_0(t))^2 dt} p_0(t) \\ &= t - \frac{\int_{-1}^1 t dt}{\int_{-1}^1 dt} = t - 0 = t \end{aligned}$$

$$\begin{aligned} p_2(t) &= t^2 - \left(\text{proj}_{p_0} t^2\right)(t) - \left(\text{proj}_{p_1} t^2\right)(t) \\ &= t^2 - \frac{\langle t^2, p_0 \rangle}{\|p_0\|^2} p_0(t) - \frac{\langle t^2, p_1 \rangle}{\|p_1\|^2} p_1(t) \\ &= t^2 - \frac{\int_{-1}^1 t^2 p_0(t) dt}{\int_{-1}^1 (p_0(t))^2 dt} p_0(t) - \frac{\int_{-1}^1 t^2 p_1(t) dt}{\int_{-1}^1 (p_1(t))^2 dt} p_1(t) \\ &= t^2 - \frac{\int_{-1}^1 t^2 dt}{\int_{-1}^1 dt} - \frac{\int_{-1}^1 t^2 t dt}{\int_{-1}^1 t^2 dt} t = t^2 - \frac{2/3}{2} - 0 = t^2 - \frac{1}{3} \end{aligned}$$

Se continuarmos desta forma não encontraremos uma fórmula para todo $p_n(t)$. Vamos encontrar uma tal fórmula de outra maneira. Para isso defina

$$q_n(t) = \frac{d^n}{dt^n}(t^2 - 1)^n. \quad (2.9)$$

(a) O conjunto $\{q_n \mid n = 0, 1, \dots\}$ é ortogonal. Com efeito, se $m < n$, então

$$\langle q_n, q_m \rangle = \left\langle q_n, \sum_{k=0}^m \alpha_k t^k \right\rangle = \sum_{k=0}^m \alpha_k \langle q_n, t^k \rangle = \sum_{k=0}^m \alpha_k \int_{-1}^1 q_n(t) t^k dt = \sum_{k=0}^m \alpha_k 0 = 0,$$

pois se, $w_n(t) = (t^2 - 1)^n$, então

$$\frac{d^k w_n}{dt^k} \text{ se anula em } t = \pm 1, \text{ para } k = 0, 1, \dots, n-1.$$

(b) Como $\dim(\mathcal{P}_n) = n + 1$, então, pela [Proposição 2.7 na página 122](#), q_0, \dots, q_n formam uma base de \mathcal{P}_n . Assim, se $p \in \mathcal{P}_{n-1}$, então

$$\langle p, q_n \rangle = \left\langle \sum_{k=0}^{n-1} \alpha_k q_k, q_n \right\rangle = \sum_{k=0}^{n-1} \alpha_k \langle q_k, q_n \rangle = 0. \quad (2.10)$$

Como, também, p_0, \dots, p_n formam uma base ortogonal de \mathcal{P}_n , então pela [Proposição 2.7 na página 122](#) temos que

$$q_n = \sum_{k=1}^n \frac{\langle q_n, p_k \rangle}{\|p_k\|^2} p_k = \frac{\langle q_n, p_n \rangle}{\|p_n\|^2} p_n.$$

Ou seja, q_n é um múltiplo escalar de p_n . Comparando os coeficientes dos termos de grau n em p_n e q_n concluímos que

$$p_n = \frac{n!}{(2n)!} q_n.$$

Portanto os polinômios p_n que são obtidos aplicando-se o processo de ortogonalização de Gram-Schmidt aos polinômios $\{1, t, t^2, \dots, t^n, \dots\}$ são dados por

$$p_n(t) = \frac{n!}{(2n)!} \frac{d^n}{dt^n} (t^2 - 1)^n. \quad (2.11)$$

Esta é conhecida como **fórmula de Rodrigues**.

Vamos, agora, calcular $\|p_n\|$. Seja $w_n(t) = (t^2 - 1)^n$. Então $\frac{d^k w_n}{dt^k}$ se anula em $t = \pm 1$, para $k = 0, 1, \dots, n - 1$. Por isso, integrando-se por partes várias vezes temos que

$$\int_{-1}^1 \frac{d^n w_n}{dt^n} \frac{d^n w_n}{dt^n} dt = (2n)! \int_{-1}^1 (1 - t^2)^n dt = (2n)! \int_{-1}^1 (1 - t)^n (1 + t)^n dt = \frac{(n!)^2}{2n + 1} 2^{2n+1}.$$

Usando este resultado e (2.11) obtemos

$$\langle p_n, p_n \rangle = \left(\frac{n!}{(2n)!} \right)^2 \int_{-1}^1 \frac{d^n w_n}{dt^n} \frac{d^n w_n}{dt^n} dt = \left(\frac{n!}{(2n)!} \right)^2 \frac{(n!)^2}{2n + 1} 2^{2n+1} = \frac{2^{2n+1} (n!)^4}{(2n + 1) [(2n)!]^2}$$

Assim,

$$\|p_n\| = \sqrt{\langle p_n, p_n \rangle} = \frac{\sqrt{2} 2^n (n!)^2}{\sqrt{2n + 1} (2n)!}. \quad (2.12)$$

Dividindo p_n por $\frac{2^n (n!)^2}{(2n)!}$, obtemos um polinômio com norma mais favorável

$$P_n(t) = \frac{(2n)!}{2^n (n!)^2} p_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} (t^2 - 1)^n \quad (2.13)$$

que possui norma dada por

$$\|P_n\| = \sqrt{\frac{2}{2n + 1}}. \quad (2.14)$$

Os polinômios P_n , para $n = 0, 1, 2, \dots$ são chamados **polinômios de Legendre** e aparecem também no estudo de certas equações diferenciais.

$p_0(t) = 1$	$P_0(t) = 1$
$p_1(t) = t$	$P_1(t) = t$
$p_2(t) = t^2 - \frac{1}{3}$	$P_2(t) = \frac{3}{2}t^2 - \frac{1}{2}$
$p_3(t) = t^3 - \frac{3}{5}t$	$P_3(t) = \frac{5}{2}t^3 - \frac{3}{2}t$
$p_4(t) = t^4 - \frac{6}{7}t^2 + \frac{3}{35}$	$P_4(t) = \frac{35}{8}t^4 - \frac{15}{4}t^2 + \frac{3}{8}$
$p_5(t) = t^5 - \frac{10}{9}t^3 + \frac{5}{21}t$	$P_5(t) = \frac{63}{8}t^5 - \frac{35}{4}t^3 + \frac{15}{8}t$
$p_6(t) = t^6 - \frac{15}{11}t^4 + \frac{5}{11}t^2 - \frac{5}{231}$	$P_6(t) = \frac{231}{16}t^6 - \frac{315}{16}t^4 + \frac{105}{16}t^2 - \frac{5}{16}$

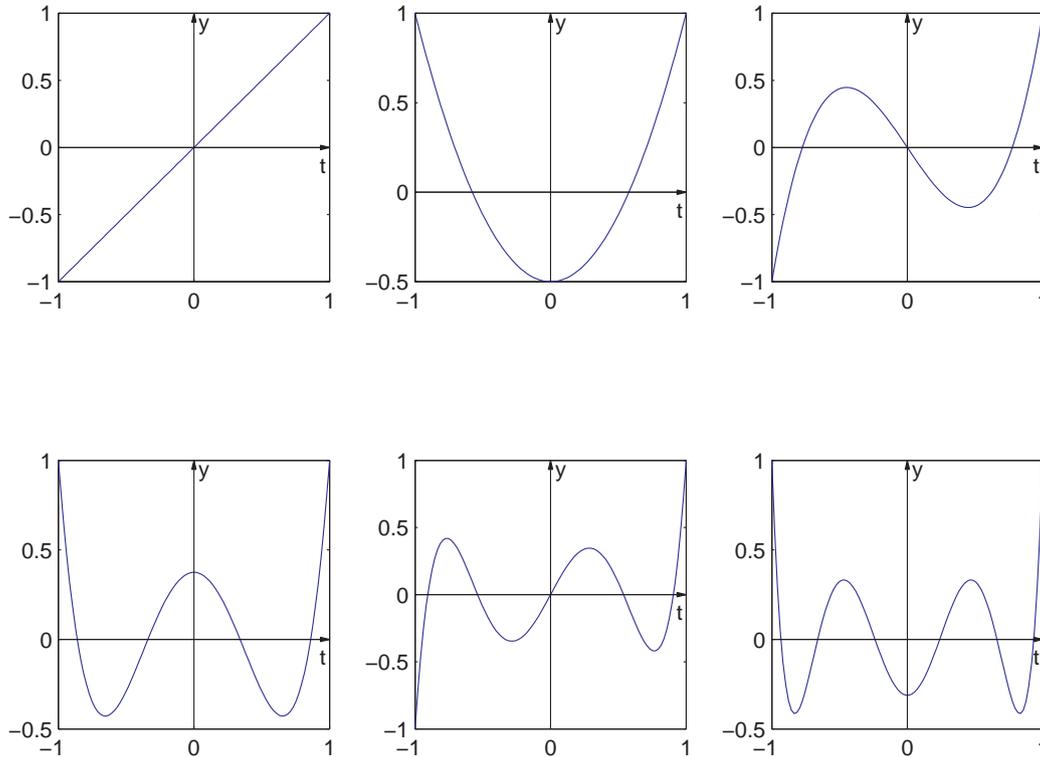


Figura 2.7: Polinômios de Legendre $P_n(t)$, para $n = 1, \dots, 6$

2.2.3 Complemento Ortogonal

Se V é um vetor não nulo de \mathbb{R}^3 , o conjunto dos vetores que são ortogonais a V , é um plano que passa pela origem e tem V como vetor normal. Neste caso dizemos que o plano é o subespaço ortogonal a $\{V\}$.

Definição 2.6. Seja \mathcal{S} um subconjunto não vazio de um espaço com produto interno \mathbb{V} . O **complemento ortogonal de \mathcal{S}** , denotado por \mathcal{S}^\perp , é o conjunto de todos os vetores de \mathbb{V} que são ortogonais a todo vetor de \mathcal{S} . Ou seja,

$$\mathcal{S}^\perp = \{X \in \mathbb{V} \mid \langle X, Y \rangle = 0 \text{ para todo } Y \in \mathcal{S}\}.$$

O conjunto \mathcal{S}^\perp é um subespaço, mesmo quando \mathcal{S} não o é.

Proposição 2.10. *Seja \mathcal{S} um subconjunto não vazio de um espaço vetorial com produto interno. Então, o conjunto \mathcal{S}^\perp é um subespaço.*

Demonstração. Vamos verificar as propriedades (0) e (0') do Teorema 1.3 na página 14.

(0) Sejam X_1 e X_2 vetores de \mathcal{S}^\perp . Então,

$$\langle X_1 + X_2, Y \rangle = \langle X_1, Y \rangle + \langle X_2, Y \rangle = 0 + 0 = 0, \quad \text{para todo } Y \in \mathcal{S}.$$

(0') Seja $X \in \mathcal{S}^\perp$ e α um escalar. Então,

$$\langle \alpha X, Y \rangle = \alpha \langle X, Y \rangle = \alpha 0 = 0, \quad \text{para todo } Y \in \mathcal{S}.$$

■

Exemplo 2.22. Se $\mathcal{S} = \{\vec{0}\} \subset \mathbb{V}$, então $\mathcal{S}^\perp = \mathbb{V}$. Se $\mathcal{S} = \mathbb{V}$, então $\mathcal{S}^\perp = \{\vec{0}\}$.

Exemplo 2.23. Seja $\mathcal{S} = \{V\} \subset \mathbb{R}^n$, em que $V = (a_1, \dots, a_n) \neq \vec{0}$. Então, $\langle V, X \rangle = a_1 x_1 + \dots + a_n x_n$ e assim

$$\mathcal{S}^\perp = \{X = (x_1, \dots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \dots + a_n x_n = 0\},$$

que é um hiperplano que passa pela origem.

Exemplo 2.24. Seja $\mathcal{S} = \{V_1, \dots, V_m\} \subset \mathbb{R}^n$, onde $V_1 = (a_{11}, \dots, a_{1n}), \dots, V_m = (a_{m1}, \dots, a_{mn})$. Então o complementar ortogonal de \mathcal{S} é o espaço solução do sistema linear homogêneo

$$\mathcal{S}^\perp = \{X \in \mathbb{R}^n \mid AX = \vec{0}\}, \quad \text{onde } A = (a_{ij})_{m \times n}.$$

Se $\mathcal{S} = \mathbb{W}$ é um subespaço de dimensão finita, além de \mathbb{W}^\perp ser um subespaço, é válido o seguinte resultado.

Proposição 2.11. *Sejam \mathbb{W} um subespaço de dimensão finita de um espaço vetorial com produto interno \mathbb{V} e \mathbb{W}^\perp o seu complemento ortogonal. Então:*

$$\mathbb{V} = \mathbb{W} \oplus \mathbb{W}^\perp \quad \text{e} \quad (\mathbb{W}^\perp)^\perp = \mathbb{W}$$

Demonstração. Vamos, em primeiro lugar, mostrar que $\mathbb{V} = \mathbb{W} \oplus \mathbb{W}^\perp$. Para isso, vamos mostrar que todo vetor de \mathbb{V} se escreve como soma de um elemento de \mathbb{W} com um de \mathbb{W}^\perp e que $\mathbb{W} \cap \mathbb{W}^\perp = \{\vec{0}\}$. Seja W_1, \dots, W_n uma base ortogonal de \mathbb{W} . Seja V um vetor qualquer de \mathbb{V} .

Defina $W = \text{proj}_{W_1} V + \dots + \text{proj}_{W_n} V$. Pela [Proposição 2.8 na página 132](#), o vetor $U = V - W$ é ortogonal a W_k , para $k = 1, \dots, n$. Logo, U é ortogonal a todo vetor de \mathbb{W} e portanto $U \in \mathbb{W}^\perp$. Assim, $V = W + U$, com $W \in \mathbb{W}$ e $U \in \mathbb{W}^\perp$.

Além disso, se $V \in \mathbb{W} \cap \mathbb{W}^\perp$, então $\langle V, V \rangle = \|V\|^2 = 0$. Portanto $V = \vec{0}$ e $\mathbb{W} \cap \mathbb{W}^\perp = \{\vec{0}\}$.

Vamos mostrar, agora, que $(\mathbb{W}^\perp)^\perp = \mathbb{W}$. 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 $V \in (\mathbb{W}^\perp)^\perp$. Então, $V = W + U$, onde $W \in \mathbb{W}$ e $U \in \mathbb{W}^\perp$. Assim,

$$0 = \langle V, U \rangle = \langle W + U, U \rangle = \langle W, U \rangle + \langle U, U \rangle = \langle U, U \rangle = \|U\|^2.$$

Consequentemente, $U = \vec{0}$. Assim, $V \in \mathbb{W}$ e $(\mathbb{W}^\perp)^\perp \subseteq \mathbb{W}$. Portanto, $(\mathbb{W}^\perp)^\perp = \mathbb{W}$. ■

Seja \mathbb{W} um subespaço de *dimensão finita* de um espaço vetorial com produto interno \mathbb{V} . Dado um vetor $V \in \mathbb{V}$, em virtude da **Proposição 2.11** existe uma única decomposição

$$V = V_1 + V_2, \quad \text{com } V_1 \in \mathbb{W} \text{ e } V_2 \in \mathbb{W}^\perp.$$

O vetor V_1 é chamado **projeção ortogonal de V no subespaço \mathbb{W}** e é denotado por $\text{proj}_{\mathbb{W}}V$, ou seja,

$$\text{proj}_{\mathbb{W}}V = V_1.$$

Se $\{W_1, \dots, W_n\}$ é uma *base ortogonal* de \mathbb{W} , então decorre da demonstração da **Proposição 2.11** que

$$\text{proj}_{\mathbb{W}}V = \text{proj}_{W_1}V + \dots + \text{proj}_{W_n}V,$$

ou

$$\text{proj}_{\mathbb{W}}V = \frac{\langle V, W_1 \rangle}{\|W_1\|^2} W_1 + \dots + \frac{\langle V, W_n \rangle}{\|W_n\|^2} W_n.$$

Lembramos que $\frac{\langle V, W_1 \rangle}{\|W_1\|^2}, \dots, \frac{\langle V, W_n \rangle}{\|W_n\|^2}$ são os coeficientes de Fourier de V em relação ao conjunto $\{W_1, \dots, W_n\}$.

Exemplo 2.25. Seja $\mathbb{V} = \mathcal{C}^0[0, 4]$ o conjunto das funções contínuas do intervalo $[0, 4]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_0^4 f(t)g(t)dt.$$

Vamos calcular as projeções da função $f(t) = \sqrt{t} \in \mathbb{V}$ nos subespaços \mathcal{P}_0 , \mathcal{P}_1 e \mathcal{P}_2 , dos polinômios de grau menor ou igual 0, 1 e 2, respectivamente. Do **Exemplo 2.21 na página 137** temos que

$\{q_0(t) = 1/2\}$ é base ortonormal de \mathcal{P}_0 ,

$\{q_0(t) = 1/2, q_1(t) = (\sqrt{3}/4)(t - 2)\}$ é base ortogonal de \mathcal{P}_1 e

$\{q_0(t) = 1/2, q_1(t) = (\sqrt{3}/4)(t - 2), q_2(t) = (3\sqrt{5}/16)(t^2 - 4t + 8/3)\}$ é base ortogonal de \mathcal{P}_2 . Vamos calcular as projeções de \sqrt{t} nos polinômios q_0, q_1 e q_2 .

$$\begin{aligned} (\text{proj}_{q_0} \sqrt{t})(t) &= \frac{\langle \sqrt{t}, q_0 \rangle}{\|q_0\|^2} q_0(t) = \left(\int_0^4 \sqrt{t} q_0(t) dt \right) q_0(t) = (1/4) \int_0^4 \sqrt{t} dt = (1/6) t^{3/2} \Big|_0^4 = \frac{4}{3} \\ (\text{proj}_{q_1} \sqrt{t})(t) &= \frac{\langle \sqrt{t}, q_1 \rangle}{\|q_1\|^2} q_1(t) = \left(\int_0^4 \sqrt{t} q_1(t) dt \right) q_1(t) = (3/16) \left(\int_0^4 (t-2) \sqrt{t} dt \right) (t-2) \\ &= (3/16) \left((2/5) t^{5/2} \Big|_0^4 - 2(2/3) t^{3/2} \Big|_0^4 \right) (t-2) = (3/16)(32/15) (t-2) = \frac{2}{5} (t-2) \\ (\text{proj}_{q_2} \sqrt{t})(t) &= \frac{\langle \sqrt{t}, q_2 \rangle}{\|q_2\|^2} q_2(t) = \left(\int_0^4 \sqrt{t} q_2(t) dt \right) q_2(t) \\ &= (45/256) \left(\int_0^4 \sqrt{t} (t^2 - 4t + 8/3) dt \right) (t^2 - 4t + 8/3) \\ &= (45/256) \left((2/7) t^{7/2} \Big|_0^4 - 4(2/5) t^{5/2} \Big|_0^4 + (8/3)(2/3) t^{3/2} \Big|_0^4 \right) (t^2 - 4t + 8/3) \\ &= (45/256) \left(2^8/7 - 2^8/5 + 2^7/9 \right) (t^2 - 4t + 8/3) = -\frac{1}{14} (t^2 - 4t + 8/3) \end{aligned}$$

Assim as projeções nos subespaços $\mathcal{P}_0, \mathcal{P}_1$ e \mathcal{P}_2 são dadas por

$$\begin{aligned} (\text{proj}_{\mathcal{P}_0} \sqrt{t})(t) &= (\text{proj}_{q_0} \sqrt{t})(t) = \frac{4}{3} \\ (\text{proj}_{\mathcal{P}_1} \sqrt{t})(t) &= (\text{proj}_{q_0} \sqrt{t})(t) + (\text{proj}_{q_1} \sqrt{t})(t) = \frac{4}{3} + \frac{2}{5} (t-2) = \frac{2}{5} t + \frac{8}{15} \\ (\text{proj}_{\mathcal{P}_2} \sqrt{t})(t) &= (\text{proj}_{q_0} \sqrt{t})(t) + (\text{proj}_{q_1} \sqrt{t})(t) + (\text{proj}_{q_2} \sqrt{t})(t) \\ &= \frac{4}{3} + \frac{2}{5} (t-2) - \frac{1}{14} (t^2 - 4t + 8/3) = -\frac{1}{14} t^2 + \frac{24}{35} t + \frac{12}{35}. \end{aligned}$$

t	$\frac{2}{5}t + \frac{8}{15}$	$-\frac{1}{14}t^2 + \frac{24}{35}t + \frac{12}{35}$	\sqrt{t}
1.5000	1.1333	1.2107	1.2247
2.0000	1.3333	1.4286	1.4142
2.5000	1.5333	1.6107	1.5811
3.0000	1.7333	1.7571	1.7321
3.5000	1.9333	1.8679	1.8708

Exemplo 2.26. Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Para cada $n > 0$, seja \mathbb{W}_n o subespaço gerado (conjunto das combinações lineares) pelas funções

$$1, \cos \pi t, \sin \pi t, \dots, \cos n\pi t, \sin n\pi t.$$

Vamos calcular a projeção da função $f(t) = e^t$ no subespaço \mathbb{W}_n . Já mostramos no **Exemplo 2.9 na página 112** que $\mathcal{B} = \{1, \cos \pi t, \sin \pi t, \dots, \cos n\pi t, \sin n\pi t, \dots\}$ é

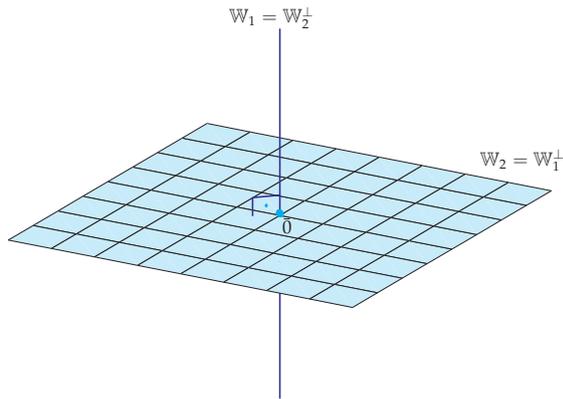


Figura 2.8: Complementos ortogonais

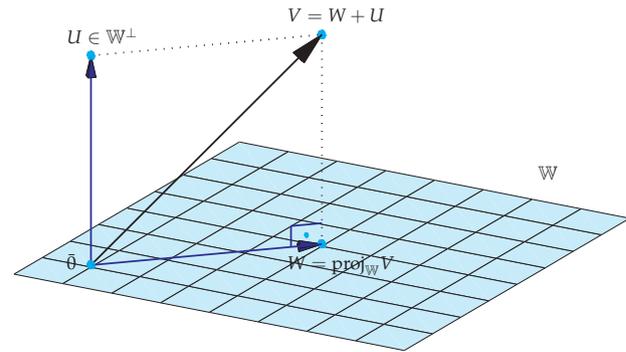


Figura 2.9: Decomposição de um ponto $X = Y + Z$, com $Y \in W, Z \in W^\perp$

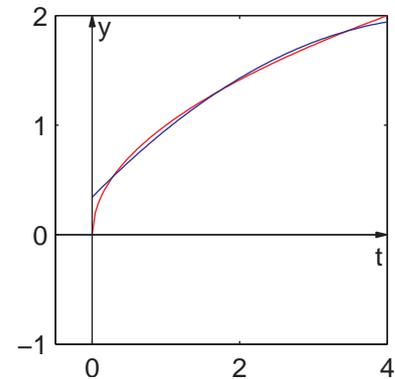
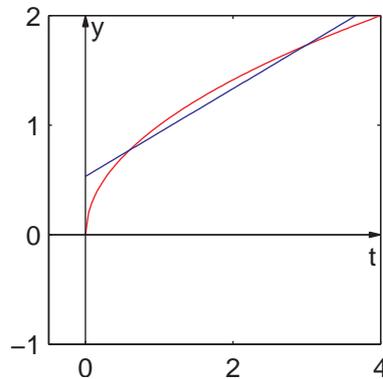
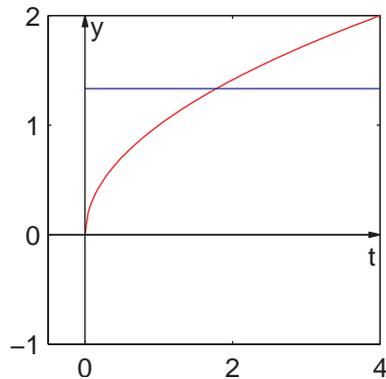


Figura 2.10: Projeções da função $f(t) = \sqrt{t}$ nos subespaços \mathcal{P}_n dos polinômios de grau menor ou igual a n , para $n = 0, 1, 2$.

um conjunto ortogonal. Assim,

$$\begin{aligned} \left(\text{proj}_{\mathbb{W}_n} e^t\right)(t) &= \frac{\langle e^t, 1 \rangle}{\|1\|^2} + \frac{\langle e^t, \cos \pi t \rangle}{\|\cos \pi t\|^2} \cos \pi t + \frac{\langle e^t, \text{sen } \pi t \rangle}{\|\text{sen } \pi t\|^2} \text{sen } \pi t + \dots \\ &+ \frac{\langle e^t, \cos n\pi t \rangle}{\|\cos n\pi t\|^2} \cos n\pi t + \frac{\langle e^t, \text{sen } n\pi t \rangle}{\|\text{sen } n\pi t\|^2} \text{sen } n\pi t. \end{aligned}$$

Assim,

$$\left(\text{proj}_{\mathbb{W}_n} e^t\right)(t) = a_0 + \sum_{k=1}^n a_k \cos k\pi t + \sum_{k=1}^n b_k \text{sen } k\pi t,$$

Em que a_0, a_1, \dots, a_n e b_1, \dots, b_n são os coeficientes de Fourier de e^t em relação ao conjunto \mathcal{B} , que já calculamos no **Exemplo 2.18** na página 127.

$$\begin{aligned} a_0 &= \frac{\langle e^t, 1 \rangle}{\|1\|^2} = c_0 = \frac{1}{2}(e - e^{-1}), \\ a_k &= \frac{\langle e^t, \cos k\pi t \rangle}{\|\cos k\pi t\|^2} = 2\Re(c_k) = \frac{(-1)^k(e - e^{-1})}{1 + k^2\pi^2} \\ b_k &= \frac{\langle e^t, \text{sen } k\pi t \rangle}{\|\text{sen } k\pi t\|^2} = -2\Im(c_k) = \frac{(-1)^{k+1}(e - e^{-1})k\pi}{1 + k^2\pi^2} \end{aligned}$$

Exemplo 2.27. Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Seja \mathcal{P}_n o espaço dos polinômios de grau menor ou igual a n , para $n = 0, 1, 2, \dots$. Vamos calcular a projeção da função $f(t) = e^t \in \mathbb{V}$ em relação ao subespaço \mathcal{P}_n .

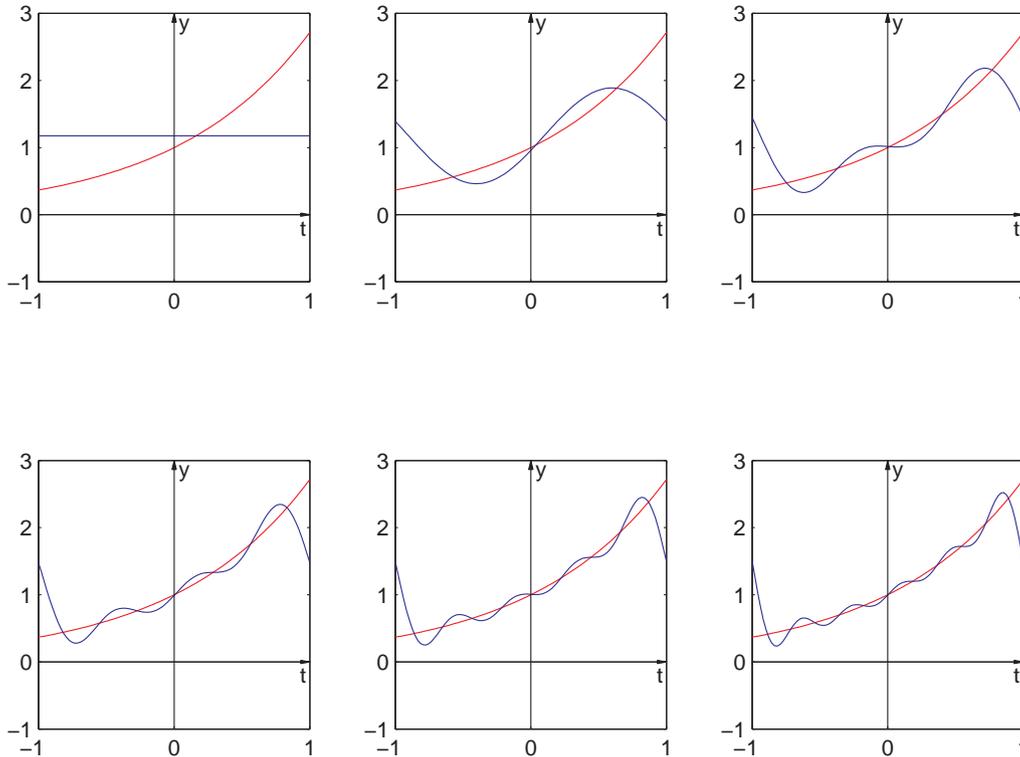


Figura 2.11: Projeções da função $f(t) = e^t$ nos subespaços W_n das combinações lineares das funções $1, \cos \pi t, \sin \pi t, \dots, \cos n\pi t, \sin n\pi t$, para $n = 0, \dots, 5$.

Para isso precisamos de uma base ortogonal de \mathcal{P}_n . Já mostramos na [Subseção 2.2.2 na página 141](#) que os polinômios de Legendre P_n formam uma tal base. Assim,

$$\text{proj}_{\mathcal{P}_n} f = \text{proj}_{P_0} f + \dots + \text{proj}_{P_n} f = \frac{\langle f, P_0 \rangle}{\|P_0\|^2} P_0 + \dots + \frac{\langle f, P_n \rangle}{\|P_n\|^2} P_n.$$

De (2.13) e (2.14) na página 143 temos que

$$\begin{aligned} \|P_n\|^2 &= \frac{2}{2n+1} \\ \langle f, P_n \rangle &= \frac{1}{2^n n!} \int_{-1}^1 e^t \frac{d^n}{dx^n} (t^2 - 1)^n dt = \frac{(-1)^n}{2^n n!} \int_{-1}^1 e^t (t^2 - 1)^n dt \\ &= \frac{(-1)^n}{2^n n!} \sum_{k=0}^n \binom{n}{k} (-1)^k \int_{-1}^1 t^{2k} e^t dt \\ &= \frac{(-1)^n}{2^n n!} \sum_{k=0}^n \binom{n}{k} (-1)^k \left(\sum_{j=0}^{2k} (-1)^j \binom{2k}{j} j! t^{2k-j} e^t \Big|_{-1}^1 \right) \\ &= \frac{(-1)^n}{2^n n!} \sum_{k=0}^n \sum_{j=1}^{2k} (-1)^{k+j} \binom{n}{k} \binom{2k}{j} j! t^{2k-j} e^t \Big|_{-1}^1 \\ &= \frac{(-1)^n}{2^n n!} \sum_{k=0}^n \sum_{j=1}^{2k} (-1)^{k+j} \binom{n}{k} \binom{2k}{j} j! (e - (-1)^j e^{-1}) \\ &= \frac{(-1)^n}{2^n n!} \left[\left(\sum_{k=0}^n \sum_{j=1}^{2k} (-1)^{k+j} \binom{n}{k} \binom{2k}{j} j! \right) e - \left(\sum_{k=0}^n \sum_{j=1}^{2k} (-1)^k \binom{n}{k} \binom{2k}{j} j! \right) e^{-1} \right] \\ &= \frac{(-1)^n}{2^n n!} (a_n e - b_n e^{-1}) \end{aligned}$$

em que

$$a_n = \sum_{k=0}^n \sum_{j=1}^{2k} (-1)^{k+j} \binom{n}{k} \binom{2k}{j} j!$$

$$b_n = \sum_{k=0}^n \sum_{j=1}^{2k} (-1)^k \binom{n}{k} \binom{2k}{j} j!$$

Assim a projeção da função $f(t) = e^t$ no subespaço \mathcal{P}_n é dada por

$$\begin{aligned} (\text{proj}_{\mathcal{P}_n} e^t)(t) &= \sum_{m=0}^n (\text{proj}_{P_m} e^t)(t) = \sum_{m=0}^n \frac{\langle e^t, P_m \rangle}{\|P_m\|^2} P_m(t) \\ &= \sum_{m=0}^n \frac{(-1)^m (2m+1)}{2^{m-1} m!} (a_m e - b_m e^{-1}) P_m(t) \end{aligned}$$

2.2.4 Distância de um Ponto a um Subespaço

No \mathbb{R}^3 a distância de um ponto P_0 a um plano $\pi : ax + by + cz + d = 0$, é definida como a menor distância entre P_0 e um ponto P do plano. Escrevemos:

$$\text{dist}(P_0, \pi) = \min_{P \in \pi} \|P_0 - P\|.$$

Esta distância é atingida para um ponto Q tal que $P_0 - Q$ é perpendicular ao plano. Assim,

$$\text{dist}(P_0, \pi) = \min_{P \in \pi} \|P_0 - P\| = \|P_0 - Q\|,$$

onde $Q \in \pi$ é tal que $P_0 - Q$ é perpendicular π .

A mesma coisa ocorre com subespaços **de dimensão finita** de um espaço vetorial com produto interno \mathbb{V} . Sejam \mathbb{W} um subespaço de \mathbb{V} e $V \in \mathbb{V}$. Definimos a **distância de V a \mathbb{W}** como

$$\text{dist}(V, \mathbb{W}) = \min_{X \in \mathbb{W}} \|V - X\|.$$

Teorema 2.12. *Seja \mathbb{V} um espaço vetorial com produto interno. Seja \mathbb{W} um subespaço de dimensão finita de \mathbb{V} . Para cada $V \in \mathbb{V}$*

$$\text{dist}(V, \mathbb{W}) = \min_{X \in \mathbb{W}} \|V - X\| = \|V - \text{proj}_{\mathbb{W}} V\| = \sqrt{\|V\|^2 - \|\text{proj}_{\mathbb{W}} V\|^2}.$$

Demonstração. Pela [Proposição 2.11](#), $\mathbb{V} = \mathbb{W} \oplus \mathbb{W}^\perp$. Seja $V \in \mathbb{V}$ fixo. Então, existe uma única decomposição de V como

$$V = W + U,$$

onde $W = \text{proj}_{\mathbb{W}} V \in \mathbb{W}$ e $U = V - \text{proj}_{\mathbb{W}} V \in \mathbb{W}^\perp$. Seja X um vetor qualquer de \mathbb{W} . Pelo Teorema de Pitágoras ([Proposição 2.5 na página 119](#)) temos que

$$\|V - X\|^2 = \|(V - W) + (W - X)\|^2 = \|V - W\|^2 + \|W - X\|^2, \quad (2.15)$$

pois $V - W = U \in \mathbb{W}^\perp$ e $W - X \in \mathbb{W}$.

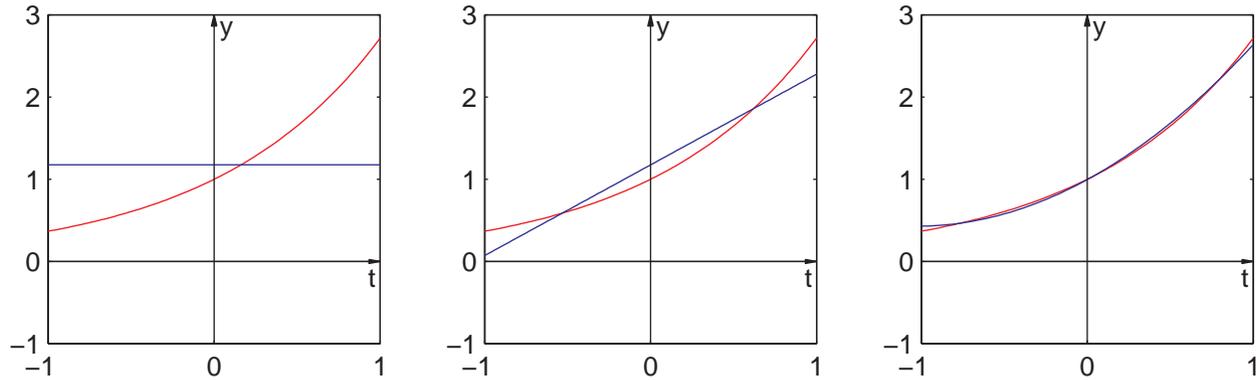


Figura 2.12: Projeções da função $f(t) = e^t$ nos subespaços \mathcal{P}_n dos polinômios de grau menor ou igual a n , para $n = 0, 1, 2$.

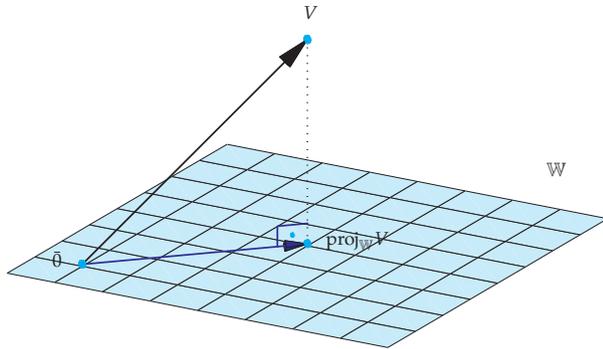


Figura 2.13: Ponto em um subespaço \mathbb{W} mais próximo do ponto V

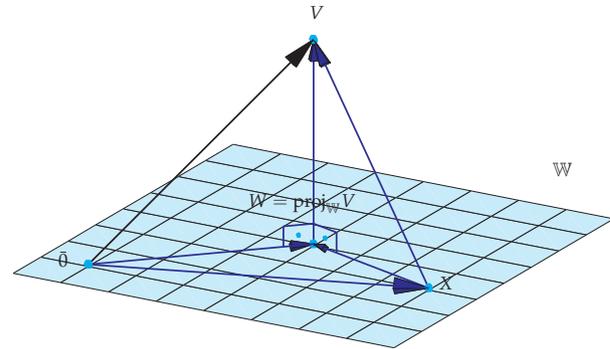


Figura 2.14: Distância de V a um ponto X qualquer do subespaço \mathbb{W}

Variando X em \mathbb{W} , vemos de (2.15) que o mínimo de $\|V - X\|^2$ ocorre somente para $X = W$, já que $\|W - V\|^2$ permanece fixo em (2.15) quando variamos X em \mathbb{W} . Portanto,

$$\min_{X \in \mathbb{W}} \|V - X\| = \|V - \text{proj}_{\mathbb{W}} V\| = \sqrt{\|V\|^2 - \|\text{proj}_{\mathbb{W}} V\|^2}.$$

A última igualdade segue-se novamente do Teorema de Pitágoras (Proposição 2.5 na página 119). ■

Exemplo 2.28. Seja $\mathbb{V} = C^0[0, 4]$ o conjunto das funções contínuas do intervalo $[0, 4]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_0^4 f(t)g(t)dt.$$

Vamos calcular a distância da função $f(t) = \sqrt{t} \in \mathbb{V}$ aos subespaços \mathcal{P}_0 , \mathcal{P}_1 e \mathcal{P}_2 , dos polinômios de grau menor ou igual 0, 1 e 2, respectivamente.

$$\text{dist}(f, \mathcal{P}_n) = \min_{g \in \mathcal{P}_n} \|f - g\| = \|f - \text{proj}_{\mathcal{P}_n} f\| = \sqrt{\|f\|^2 - \|\text{proj}_{\mathcal{P}_n} f\|^2}.$$

$$\|\sqrt{t}\|^2 = \langle \sqrt{t}, \sqrt{t} \rangle = \int_0^4 t dt = 8$$

Vimos no Exemplo 2.25 na página 149 que

$$\begin{aligned} (\text{proj}_{\mathcal{P}_0} \sqrt{t})(t) &= \frac{4}{3} \\ (\text{proj}_{\mathcal{P}_1} \sqrt{t})(t) &= \frac{2}{5}t + \frac{8}{15} \\ (\text{proj}_{\mathcal{P}_2} \sqrt{t})(t) &= -\frac{1}{14}t^2 + \frac{24}{35}t + \frac{12}{35} = -\frac{1}{14} \left(\left(t - \frac{24}{5}\right)^2 - \frac{696}{25} \right). \end{aligned}$$

Vamos calcular o quadrado das normas das projeções

$$\|\text{proj}_{\mathcal{P}_0} \sqrt{t}\|^2 = \frac{16}{9} \int_0^4 dt = \frac{64}{9}$$

$$\|\text{proj}_{\mathcal{P}_1} \sqrt{t}\|^2 = \int_0^4 \left(\frac{2}{5}t + \frac{8}{15} \right)^2 dt = \frac{1722}{225}$$

$$\|\text{proj}_{\mathcal{P}_2} \sqrt{t}\|^2 = \frac{1}{14^2} \int_0^4 \left(\left(t - \frac{24}{5} \right)^2 - \frac{696}{25} \right)^2 dt = \frac{9792}{1225}.$$

Assim as distâncias de \sqrt{t} aos subespaços \mathcal{P}_0 , \mathcal{P}_1 e \mathcal{P}_2 são dadas por

$$\text{dist}(\sqrt{t}, \mathcal{P}_0) = \sqrt{\|\sqrt{t}\|^2 - \|\text{proj}_{\mathcal{P}_0} \sqrt{t}\|^2} = \sqrt{8 - \frac{64}{9}} = \frac{2\sqrt{2}}{3}.$$

$$\text{dist}(\sqrt{t}, \mathcal{P}_1) = \sqrt{\|\sqrt{t}\|^2 - \|\text{proj}_{\mathcal{P}_1} \sqrt{t}\|^2} = \sqrt{8 - \frac{1722}{225}} = \frac{2\sqrt{2}}{15}.$$

$$\text{dist}(\sqrt{t}, \mathcal{P}_2) = \sqrt{\|\sqrt{t}\|^2 - \|\text{proj}_{\mathcal{P}_2} \sqrt{t}\|^2} = \sqrt{8 - \frac{9792}{1225}} = \frac{2\sqrt{2}}{35}.$$

Exemplo 2.29. Seja $f(t) = e^t \in \mathcal{C}^0[-1, 1]$. Para cada $n > 0$, seja

$$\mathbb{W}_n = [1, \cos \pi t, \text{sen } \pi t, \dots, \cos n\pi t, \text{sen } n\pi t].$$

Vamos calcular $\text{dist}(f, \mathbb{W}_n)$. Já mostramos no [Exemplo 2.9 na página 112](#) que

$$\mathcal{B} = \{1, \cos t, \text{sen } t, \dots, \cos nt, \text{sen } nt, \dots\}$$

é um conjunto ortogonal.

$$\begin{aligned}
 (\text{dist}(f, \mathbb{W}_n))^2 &= \|f\|^2 - \|\text{proj}_1 f\|^2 - \sum_{k=0}^n (\|\text{proj}_{\cos k\pi t} f\|^2 + \|\text{proj}_{\text{sen } k\pi t} f\|^2) \\
 &= \|f\|^2 - \frac{\langle f, 1 \rangle^2}{\|1\|^2} - \sum_{k=0}^n \left(\frac{\langle f, \cos k\pi t \rangle^2}{\|\cos k\pi t\|^2} + \frac{\langle f, \text{sen } k\pi t \rangle^2}{\|\text{sen } k\pi t\|^2} \right) \\
 &= \|f\|^2 - 2a_0^2 - \sum_{k=0}^n (a_k^2 + b_k^2),
 \end{aligned}$$

em que $a_0, \dots, a_n, b_1, \dots, b_n$ são os coeficiente de Fourier de f em relação ao conjunto \mathcal{B} , que já calculamos no [Exemplo 2.18 na página 127](#):

$$\begin{aligned}
 a_0 &= c_0 = \frac{1}{2}(e - e^{-1}), \\
 a_n &= 2\Re(c_n) = \frac{(-1)^n(e - e^{-1})}{1 + n^2\pi^2} \\
 b_n &= -2\Im(c_n) = \frac{(-1)^{n+1}(e - e^{-1})n\pi}{1 + n^2\pi^2}
 \end{aligned}$$

Como

$$\|f\|^2 = \langle f, f \rangle = \int_{-1}^1 e^{2t} dt = \frac{e^2 - e^{-2}}{2},$$

então

$$\begin{aligned}
 (\text{dist}(f, \mathbb{W}_n))^2 &= \frac{e^2 - e^{-2}}{2} - \frac{1}{2}(e - e^{-1})^2 - \sum_{k=1}^n \left(\frac{(e - e^{-1})^2}{(1 + k^2\pi^2)^2} + \frac{k^2\pi^2(e - e^{-1})^2}{(1 + k^2\pi^2)^2} \right) \\
 &= \frac{e^2 - e^{-2}}{2} - (e - e^{-1})^2 \left(\frac{1}{2} + \sum_{k=1}^n \frac{1}{1 + k^2\pi^2} \right)
 \end{aligned}$$

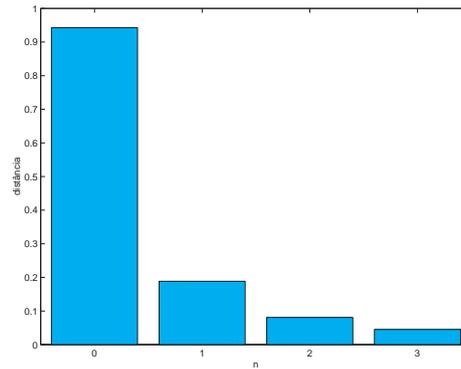


Figura 2.15: Distância de $f(t) = \sqrt{t}$ aos subespaços dos polinômios de grau menor ou igual a 0, 1, 2, 3.

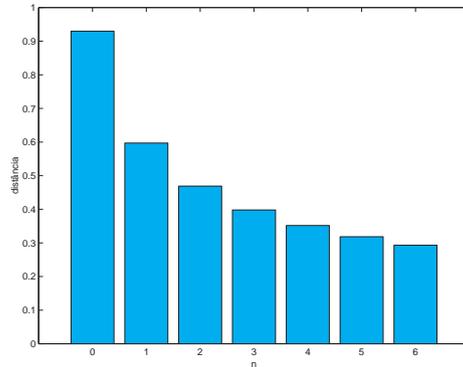


Figura 2.16: Distância de $f(t) = e^t$ ao subespaço gerado por $1/\sqrt{2}, \cos \pi t, \sin \pi t, \dots, \cos n\pi t, \sin n\pi t$

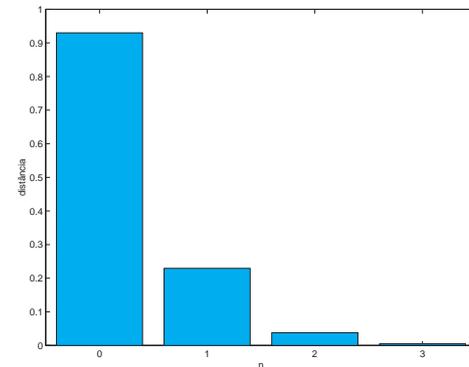


Figura 2.17: Distância de $f(t) = e^t$ ao subespaço \mathcal{P}_n

Exemplo 2.30. Seja $f(t) = e^t \in C^0[-1, 1]$. Seja \mathcal{P}_n o espaço dos polinômios de grau menor ou igual a n , para $n = 0, 1, 2, \dots$. Vamos calcular $\text{dist}(f, \mathcal{P}_n)$. Para isso precisamos de uma base ortogonal para \mathcal{P}_n . Já mostramos na [Subseção 2.2.2 na página 141](#) que os polinômios de Legendre P_0, \dots, P_n formam uma tal base. Assim,

$$(\text{dist}(f, \mathcal{P}_n))^2 = \|f\|^2 - \sum_{m=0}^n \|\text{proj}_{P_m} f\|^2 = \|f\|^2 - \sum_{m=0}^n \left(\frac{\langle f, P_m \rangle}{\|P_m\|} \right)^2.$$

De (2.14) na página 143 temos que

$$\|P_m\|^2 = \frac{2}{2m+1}$$

No [Exemplo 2.27 na página 153](#) mostramos que

$$\langle f, P_m \rangle = \frac{(-1)^m}{2^m m!} (a_m e - b_m e^{-1})$$

em que

$$a_m = \sum_{k=0}^m \sum_{j=1}^{2k} (-1)^{k+j} \binom{m}{k} \binom{2k}{j} j!$$

$$b_m = \sum_{k=0}^m \sum_{j=1}^{2k} (-1)^k \binom{m}{k} \binom{2k}{j} j!$$

Como

$$\|f\|^2 = \langle f, f \rangle = \int_{-1}^1 e^{2t} dt = \frac{e^2 - e^{-2}}{2}.$$

Então,

$$(\text{dist}(f, \mathcal{P}_n))^2 = \frac{e^2 - e^{-2}}{2} - \sum_{m=0}^n \frac{(2m+1)}{2^{2m-1} (m!)^2} (a_m e - b_m e^{-1})^2$$

2.2.5 Aplicação: Séries de Fourier

Convergência

Definição 2.7. Uma sequência de vetores $\{f_m\} = \{f_0, f_1, f_2, \dots, f_m, \dots\}$ de um espaço vetorial de funções com produto interno \mathbb{V} converge para um vetor f de \mathbb{V} se

$$\lim_{m \rightarrow \infty} \|f_m - f\| = 0.$$

Neste caso escrevemos $\lim_{m \rightarrow \infty} f_m = f$.

Proposição 2.13. Se uma sequência de vetores $\{f_m\}$ de um espaço vetorial de funções com produto interno \mathbb{V} converge para um vetor f de \mathbb{V} , então este vetor é único.

Demonstração. Vamos supor que $\lim_{m \rightarrow \infty} f_m = f$ e $\lim_{m \rightarrow \infty} f_m = g$, então pela desigualdade triangular temos que

$$\|f - g\| \leq \|f - f_m\| + \|g - f_m\|.$$

Passando ao limite obtemos que $\|f - g\| = 0$ o que implica que $f = g$. ■

Proposição 2.14. Se uma sequência de vetores $\{f_m\}$ de um espaço vetorial de funções com produto interno \mathbb{V} converge para um vetor f de \mathbb{V} , então para todo vetor g de \mathbb{V} a sequência de números reais $\{\langle f_m, g \rangle\}$ converge para $\langle f, g \rangle$. Ou seja, se $\lim_{m \rightarrow \infty} f_m = f$, então

$$\lim_{m \rightarrow \infty} \langle f_m, g \rangle = \left\langle \lim_{m \rightarrow \infty} f_m, g \right\rangle.$$

Demonstração. Seja $f = \lim_{m \rightarrow \infty} f_m$. Pela desigualdade de Cauchy-Schwarz, temos que

$$|\langle f_m, g \rangle - \langle f, g \rangle| = |\langle f_m - f, g \rangle| \leq \|f_m - f\| \|g\|.$$

Passando ao limite obtemos que $\lim_{m \rightarrow \infty} |\langle f_m, g \rangle - \langle f, g \rangle| = 0$. O que implica que $\lim_{m \rightarrow \infty} \langle f_m, g \rangle = \langle f, g \rangle$. ■

Definição 2.8. Uma **série de vetores** $\sum_{m=0}^{\infty} f_m$ de um espaço vetorial de funções com produto interno \mathbb{V} converge para um vetor f de \mathbb{V} se o limite da sequência das somas parciais converge para f , ou seja,

$$\lim_{m \rightarrow \infty} \sum_{n=0}^m f_n = f.$$

O seguinte resultado é uma consequência imediata da [Proposição 2.13](#).

Corolário 2.15. Se uma série de vetores $\sum_{m=0}^{\infty} f_m$ de um espaço vetorial de funções com produto interno \mathbb{V} converge para um vetor f de \mathbb{V} , então, para todo vetor g de \mathbb{V} ,

$$\sum_{m=0}^{\infty} \langle f_m, g \rangle = \left\langle \sum_{m=0}^{\infty} f_m, g \right\rangle.$$

O próximo resultado é uma generalização da [Proposição 2.7](#) na página 122.

Proposição 2.16. *Seja \mathbb{V} um espaço vetorial de funções com produto interno. Seja $\{g_0, g_1, g_2, \dots, g_n, \dots\}$ um subconjunto de \mathbb{V} de vetores ortogonais não nulos. Se $f = \sum_{m=0}^{\infty} c_m g_m$, então*

$$c_m = \frac{\langle f, g_m \rangle}{\|g_m\|^2}, \quad \text{para } m = 0, 1, 2, \dots$$

Demonstração. Seja $f = \sum_{m=0}^{\infty} c_m g_m$. Fazendo o produto escalar de f com g_n , para $n = 0, 1, 2, \dots$, obtemos que

$$\langle f, g_n \rangle = \left\langle \sum_{m=0}^{\infty} c_m g_m, g_n \right\rangle = \sum_{m=0}^{\infty} c_m \langle g_m, g_n \rangle = c_n \|g_n\|^2,$$

pois como os vetores g_m são ortogonais $\langle g_m, g_n \rangle = 0$, se $m \neq n$. Assim,

$$c_n = \frac{\langle f, g_n \rangle}{\|g_n\|^2}, \quad \text{para } n = 0, 1, 2, \dots$$

■

Séries de Fourier

Exemplo 2.31. Seja L um número real maior que zero. Seja $\mathbb{V} = \mathcal{C}^0[0, L]$ o conjunto das funções contínuas do intervalo $[0, L]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_0^L f(t)g(t)dt.$$

Já mostramos no [Exemplo 2.10 na página 113](#) que o conjunto

$$\left\{1, \cos \frac{\pi t}{L}, \cos \frac{2\pi t}{L}, \dots, \cos \frac{n\pi t}{L}, \dots\right\}$$

é ortogonal. Também calculamos as normas dos seus elementos.

$$\begin{aligned} \langle 1, 1 \rangle &= \int_0^L dt = L \\ \left\langle \cos \frac{n\pi t}{L}, \cos \frac{n\pi t}{L} \right\rangle &= \int_0^L \cos^2 \frac{n\pi t}{L} dt = \frac{L}{\pi} \int_0^\pi \cos^2 ns ds = \frac{L}{2\pi} \int_0^\pi [1 + \cos 2ns] ds = L/2 \end{aligned}$$

Assim, pela [Proposição 2.16](#), para toda função $f \in \mathcal{C}^0[0, L]$ que possa ser escrita como a série

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos \frac{m\pi t}{L},$$

teremos que os coeficientes da série serão dados por

$$a_m = \frac{\langle f, \cos \frac{m\pi t}{L} \rangle}{\|\cos \frac{m\pi t}{L}\|^2} = \frac{2}{L} \int_0^L f(t) \cos \frac{m\pi t}{L} dt, \quad \text{para } m = 0, 1, 2, \dots$$

Exemplo 2.32. Seja L um número real maior que zero. Seja $\mathbb{V} = \mathcal{C}^0[0, L]$ o conjunto das funções contínuas do intervalo $[0, L]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_0^L f(t)g(t)dt.$$

Já mostramos no [Exemplo 2.10 na página 113](#) que o conjunto

$$\left\{ \text{sen } \frac{\pi t}{L}, \text{sen } \frac{2\pi t}{L}, \dots, \text{sen } \frac{n\pi t}{L}, \dots \right\}$$

é ortogonal. Também calculamos as normas dos seus elementos.

$$\left\langle \text{sen } \frac{n\pi t}{L}, \text{sen } \frac{n\pi t}{L} \right\rangle = \int_0^L \text{sen}^2 \frac{n\pi t}{L} dt = \frac{L}{\pi} \int_0^\pi \text{sen}^2 ns ds = \frac{L}{2\pi} \int_0^\pi [1 - \cos 2ns] ds = L/2$$

Assim, pela [Proposição 2.16](#), para toda função $f \in \mathcal{C}^0[0, L]$ que possa ser escrita como a série

$$f(t) = \sum_{m=1}^{\infty} b_m \text{sen } \frac{m\pi t}{L},$$

teremos que os coeficientes da série serão dados por

$$b_m = \frac{\langle f, \text{sen } \frac{m\pi t}{L} \rangle}{\|\text{sen } \frac{m\pi t}{L}\|^2} = \frac{2}{L} \int_0^L f(t) \text{sen } \frac{m\pi t}{L} dt, \quad \text{para } m = 1, 2, \dots$$

Exemplo 2.33. Seja L um número real maior que zero. Seja $\mathbb{V} = \mathcal{CP}^0[-L, L]$ o conjunto das funções **seccionalmente contínuas** ou **contínuas por partes** que são funções que são contínuas nos pontos de $[-L, L]$ excetuando possivelmente em um número finito de pontos nos quais os limites laterais existem. Além disso duas

funções do conjunto são consideradas iguais se elas coincidem em todos os pontos onde ambas são contínuas. Vamos considerar $\mathbb{V} = \mathcal{C}\mathcal{P}^0[-L, L]$ com o produto interno definido por

$$\langle f, g \rangle = \int_{-L}^L f(t)g(t)dt.$$

Já mostramos no [Exemplo 2.9 na página 112](#) que o conjunto

$$\left\{1, \cos \frac{\pi t}{L}, \sin \frac{\pi t}{L}, \cos \frac{2\pi t}{L}, \sin \frac{2\pi t}{L}, \dots, \cos \frac{n\pi t}{L}, \sin \frac{n\pi t}{L}, \dots\right\}$$

é ortogonal. Também calculamos as normas dos seus elementos.

$$\begin{aligned} \langle 1, 1 \rangle &= \int_{-L}^L dt = 2L \\ \left\langle \cos \frac{n\pi t}{L}, \cos \frac{n\pi t}{L} \right\rangle &= \int_{-L}^L \cos^2 \frac{n\pi t}{L} dt = \frac{L}{\pi} \int_{-\pi}^{\pi} \cos^2 ns ds = \frac{L}{2\pi} \int_{-\pi}^{\pi} [1 + \cos 2ns] ds = L \\ \left\langle \sin \frac{n\pi t}{L}, \sin \frac{n\pi t}{L} \right\rangle &= \int_{-L}^L \sin^2 \frac{n\pi t}{L} dt = \frac{L}{\pi} \int_{-\pi}^{\pi} \sin^2 ns ds = \frac{L}{2\pi} \int_{-\pi}^{\pi} [1 - \cos 2ns] ds = L \end{aligned}$$

Assim, pela [Proposição 2.16](#), para toda função $f \in \mathcal{C}^0[-L, L]$ que possa ser escrita como a série

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos \frac{m\pi t}{L} + \sum_{m=1}^{\infty} b_m \sin \frac{m\pi t}{L},$$

teremos que os coeficientes da série serão dados por

$$a_m = \frac{\langle f, \cos \frac{m\pi t}{L} \rangle}{\|\cos \frac{m\pi t}{L}\|^2} = \frac{1}{L} \int_{-L}^L f(t) \cos \frac{m\pi t}{L} dt, \quad \text{para } m = 0, 1, 2, \dots \quad (2.16)$$

$$b_m = \frac{\langle f, \sin \frac{m\pi t}{L} \rangle}{\|\sin \frac{m\pi t}{L}\|^2} = \frac{1}{L} \int_{-L}^L f(t) \sin \frac{m\pi t}{L} dt, \quad \text{para } m = 1, 2, \dots \quad (2.17)$$

As séries dadas no [Exemplo 2.33](#) são chamadas de **séries de Fourier**, as do [Exemplo 2.31](#) de **séries de Fourier de cossenos** e as do [Exemplo 2.32](#) de **séries de Fourier de senos**. Elas aparecem no estudo de certas equações diferenciais.

Observe que as somas parciais da série de Fourier de uma função f são projeções da função f no subespaço

$$\mathbb{W}_n = [1, \cos \pi t, \text{sen } \pi t, \dots, \cos n\pi t, \text{sen } n\pi t],$$

ou seja,

$$\frac{a_0}{2} + \sum_{m=1}^n a_m \cos \frac{m\pi t}{L} + \sum_{m=1}^n b_m \text{sen } \frac{m\pi t}{L} = \left(\text{proj}_{\mathbb{W}_n} f \right) (t).$$

Na [Proposição 2.16 na página 167](#) fizemos a suposição de que a série $\sum_{m=0}^{\infty} c_m g_m$ convergia para a função f . Vamos considerar o problema inverso. Dada uma função $f \in \mathcal{C}^0[-L, L]$ podemos calcular os coeficientes a_m e b_m usando [\(2.16\)](#) e [\(2.17\)](#) e nos perguntar se a série obtida converge ou não. O teorema seguinte, cuja demonstração pode ser encontrada por exemplo em [\[15\]](#), afirma que para toda função $f \in \mathcal{C}^0[-L, L]$ a série de Fourier de f converge.

Teorema 2.17. *Seja L um número real maior que zero. Para toda função f pertencente ao espaço das funções contínuas por partes, $\mathcal{C}^0[-L, L]$, a série de Fourier de f*

$$\frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos \frac{m\pi t}{L} + \sum_{m=1}^{\infty} b_m \text{sen } \frac{m\pi t}{L},$$

em que

$$a_m = \frac{1}{L} \int_{-L}^L f(t) \cos \frac{m\pi t}{L} dt \quad \text{para } m = 0, 1, 2, \dots$$

$$b_m = \frac{1}{L} \int_{-L}^L f(t) \operatorname{sen} \frac{m\pi t}{L} dt, \quad \text{para } m = 1, 2, \dots$$

converge para f na norma $\|f\| = \left(\int_{-L}^L (f(t))^2 dt \right)^{\frac{1}{2}}$.

Se uma função $f \in \mathcal{C}^0[-L, L]$ é par, isto é, $f(-t) = f(t)$, para todo $t \in [-L, L]$, e pode ser escrita como a série

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos \frac{m\pi t}{L} + \sum_{m=1}^{\infty} b_m \operatorname{sen} \frac{m\pi t}{L},$$

então os coeficientes obtidos no [Exemplo 2.33](#) são dados por:

$$a_m = \frac{1}{L} \int_{-L}^L f(t) \cos \frac{m\pi t}{L} dt = \frac{2}{L} \int_0^L f(t) \cos \frac{m\pi t}{L} dt, \quad \text{para } m = 0, 1, 2, \dots$$

$$b_m = \frac{1}{L} \int_{-L}^L f(t) \operatorname{sen} \frac{m\pi t}{L} dt = 0 \quad \text{para } m = 1, 2, \dots$$

ou seja, os coeficientes b_m são iguais a zero e os a_m são iguais aos dados no [Exemplo 2.31](#).

Analogamente, se uma função $f \in \mathcal{C}^0[-L, L]$ é ímpar, isto é, $f(-t) = -f(t)$, para todo $t \in [-L, L]$, e pode ser escrita como a série

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos \frac{m\pi t}{L} + \sum_{m=1}^{\infty} b_m \operatorname{sen} \frac{m\pi t}{L},$$

então os coeficientes obtidos no **Exemplo 2.33** são dados por:

$$a_m = \frac{1}{L} \int_{-L}^L f(t) \cos \frac{m\pi t}{L} dt = 0 \quad \text{para } m = 0, 1, 2, \dots$$

$$b_m = \frac{1}{L} \int_{-L}^L f(t) \operatorname{sen} \frac{m\pi t}{L} dt = \frac{2}{L} \int_0^L f(t) \operatorname{sen} \frac{m\pi t}{L} dt, \quad \text{para } m = 1, 2, \dots$$

ou seja, os coeficientes a_m são iguais a zero e os b_m são iguais aos dados no **Exemplo 2.32**.

Para as funções f que são contínuas por partes em $[0, L]$ podemos prolongá-las de forma que elas se tornem par ou ímpar no intervalo $[-L, L]$ (verifique!). Assim, segue-se da observação que fizemos anteriormente, que as séries de Fourier de cossenos e de senos de f são séries de Fourier dos prolongamentos par e ímpar de f , respectivamente. Este raciocínio estende o resultado anterior para séries de Fourier de senos e de cossenos.

Corolário 2.18. *Seja L um número real maior que zero. Para toda função f pertencente ao espaço das funções contínuas por partes, $\mathcal{CP}^0[0, L]$, as séries de Fourier de cossenos de f*

$$\frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos \frac{m\pi t}{L},$$

e de Fourier de senos de f

$$\sum_{m=1}^{\infty} b_m \operatorname{sen} \frac{m\pi t}{L},$$

em que

$$a_m = \frac{2}{L} \int_0^L f(t) \cos \frac{m\pi t}{L} dt \quad \text{para } m = 0, 1, 2, \dots$$

$$b_m = \frac{2}{L} \int_0^L f(t) \operatorname{sen} \frac{m\pi t}{L} dt, \quad \text{para } m = 1, 2, \dots$$

convergem para f na norma $\|f\| = \left(\int_0^L (f(t))^2 dt \right)^{\frac{1}{2}}$.

Exemplo 2.34. Seja L um número real maior que zero. Considere a função $f_{cd}^{(0)} : [0, L] \rightarrow \mathbb{R}$ dada por

$$f_{cd}^{(0)}(t) = \begin{cases} 1, & \text{se } cL \leq t \leq dL, \\ 0, & \text{caso contrário,} \end{cases} \quad \text{para } c \text{ e } d \text{ fixos satisfazendo } 0 \leq c < d \leq 1.$$

Vamos calcular as séries de Fourier de senos e de cossenos de $f_{cd}^{(0)}$. Para a série de cossenos temos que

$$a_0 = \frac{2}{L} \int_{cL}^{dL} f(t) dt = \frac{2}{L} \int_{cL}^{dL} dt = 2(d - c),$$

$$a_m = \frac{2}{L} \int_{cL}^{dL} f(t) \cos \frac{m\pi t}{L} dt = \frac{2}{L} \int_{cL}^{dL} \cos \frac{m\pi t}{L} dt = \frac{2}{m\pi} \operatorname{sen} s \Big|_{m\pi c}^{m\pi d}, \quad \text{para } m = 1, 2, \dots$$

Assim a série de Fourier de cossenos de f é

$$f_{cd}^{(0)}(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos \frac{m\pi t}{L} = (d - c) + \frac{2}{\pi} \sum_{m=1}^{\infty} \frac{\operatorname{sen} m\pi d - \operatorname{sen} m\pi c}{m} \cos \frac{m\pi t}{L}.$$

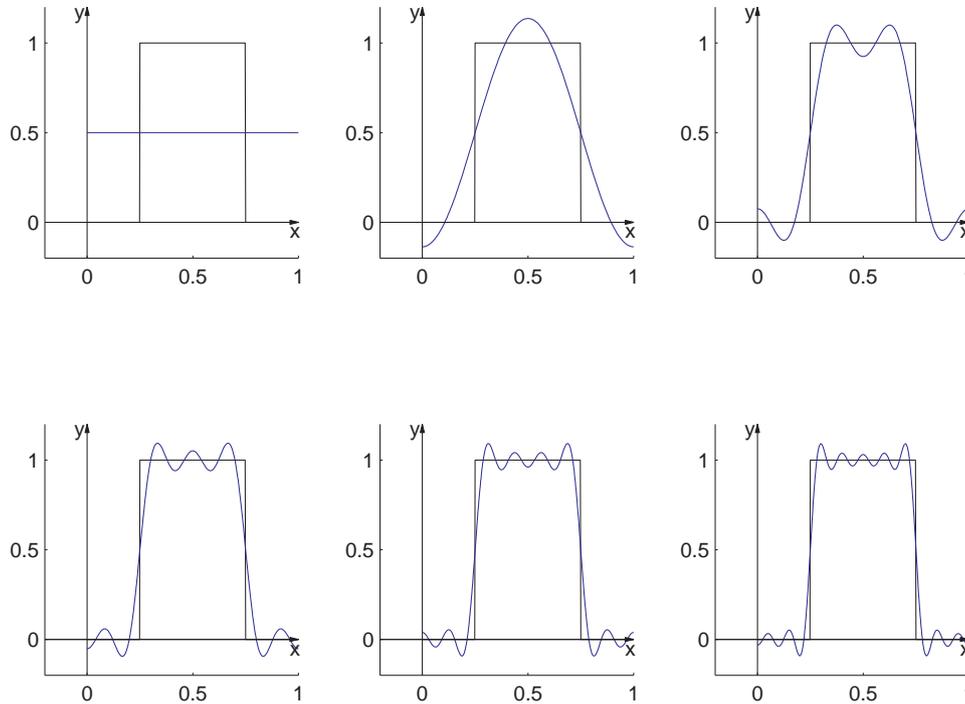


Figura 2.18: A função $f : [0, 1] \rightarrow \mathbb{R}$ definida por $f(t) = 1$, se $t \in [1/4, 3/4]$ e $f(t) = 0$, caso contrário e as somas parciais da série de Fourier de cossenos de f , para $n = 0, 2, 6, 10, 14, 18$

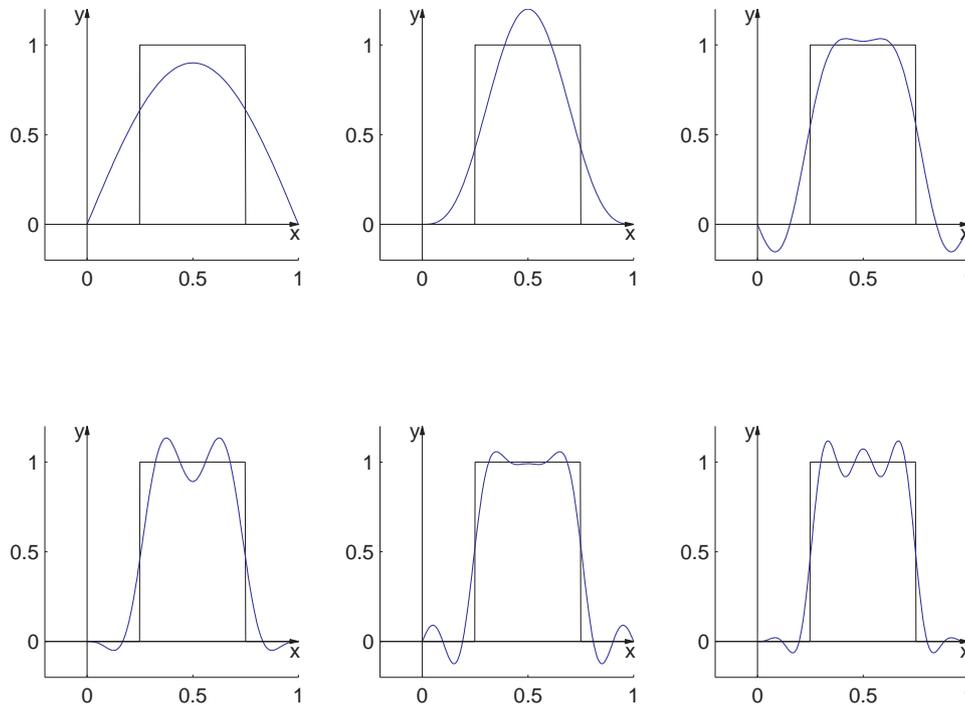


Figura 2.19: A função $f : [0,1] \rightarrow \mathbb{R}$ definida por $f(t) = 1$, se $t \in [1/4, 3/4]$ e $f(t) = 0$, caso contrário e as somas parciais da série de Fourier de senos de f , para $n = 1, \dots, 6$

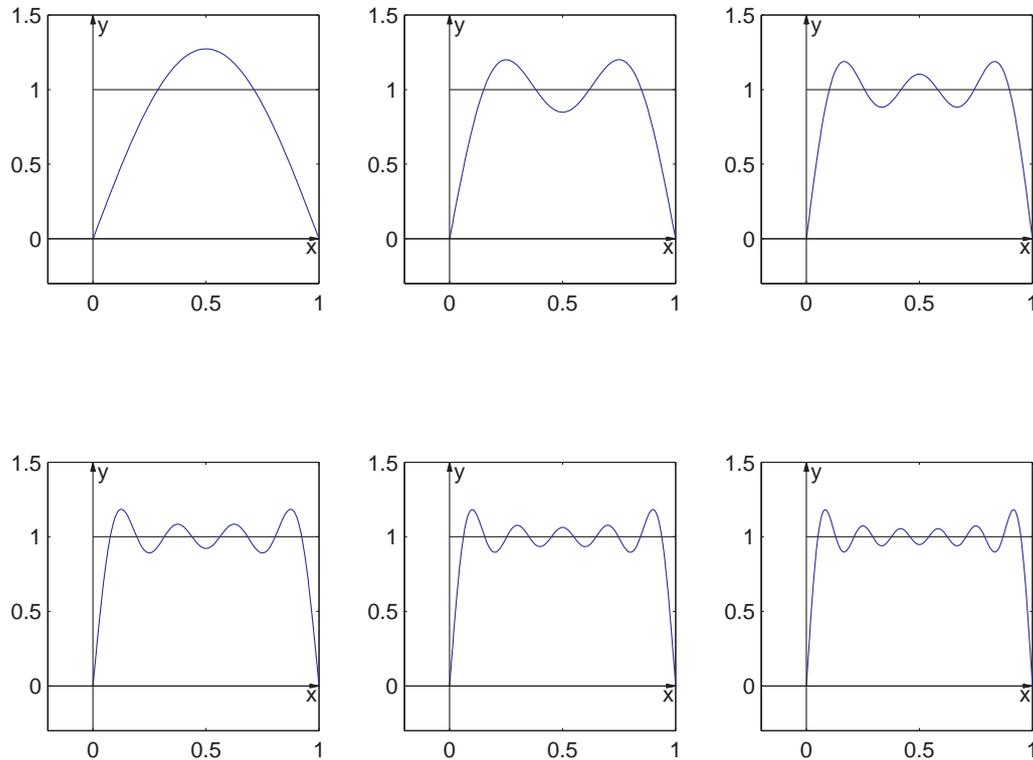


Figura 2.20: A função $f(t) = 1$ em $[0, 1]$ e as somas parciais da série de Fourier de senos de f , para $n = 1, 3, 5, 7, 9, 11$

Observe que a série de Fourier de cossenos da função constante igual a 1, $f_{01}^{(0)}$, tem somente o primeiro termo diferente de zero que é igual a 1.

Para a série de senos temos que para $m = 1, 2, \dots$,

$$b_m = \frac{2}{L} \int_{cL}^{dL} f(t) \operatorname{sen} \frac{m\pi t}{L} dt = \frac{2}{L} \int_{cL}^{dL} \operatorname{sen} \frac{m\pi t}{L} dt = -\frac{2}{m\pi} \cos s \Big|_{m\pi c}^{m\pi d}$$

Assim, a série de Fourier de senos de $f_{cd}^{(0)}$ é dada por

$$f_{cd}^{(0)}(t) = \sum_{m=1}^{\infty} b_m \operatorname{sen} \frac{m\pi t}{L} = \frac{2}{\pi} \sum_{m=1}^{\infty} \frac{\cos m\pi c - \cos m\pi d}{m} \operatorname{sen} \frac{m\pi t}{L}$$

Observe que para a função constante igual a 1, $f_{01}^{(0)}$ os termos de índice par são iguais a zero e neste caso a série de senos de $f_{01}^{(0)}$ é dada por

$$f_{01}^{(0)}(t) = \frac{4}{\pi} \sum_{m=1}^{\infty} \frac{1}{2m-1} \operatorname{sen} \frac{(2m-1)\pi t}{L}$$

Exemplo 2.35. Considere a função $f_{cd}^{(1)} : [0, L] \rightarrow \mathbb{R}$ dada por

$$f_{cd}^{(1)}(t) = \begin{cases} t, & \text{se } cL \leq t \leq dL, \\ 0, & \text{caso contrário,} \end{cases} \quad \text{para } c \text{ e } d \text{ fixos satisfazendo } 0 \leq c < d \leq 1.$$

Vamos calcular as séries de Fourier de senos e de cossenos de $f_{cd}^{(1)}$. Para a série de

cossenos temos que

$$\begin{aligned} a_0 &= \frac{2}{L} \int_{cL}^{dL} f(t) dt = \frac{2}{L} \int_{cL}^{dL} t dt = L(d^2 - c^2) \\ a_m &= \frac{2}{L} \int_{cL}^{dL} f(t) \cos \frac{m\pi t}{L} dt = \frac{2}{L} \int_{cL}^{dL} t \cos \frac{m\pi t}{L} dt = \frac{2L}{m^2\pi^2} \int_{m\pi c}^{m\pi d} s \cos s ds \\ &= \frac{2L}{m^2\pi^2} (s \operatorname{sen} s + \cos s) \Big|_{m\pi c}^{m\pi d} \end{aligned}$$

Assim a série de Fourier de cossenos de f é

$$f_{cd}^{(1)}(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos \frac{m\pi t}{L} = \frac{L(d^2 - c^2)}{2} + \frac{2L}{\pi^2} \sum_{m=1}^{\infty} \frac{(s \operatorname{sen} s + \cos s) \Big|_{m\pi c}^{m\pi d}}{m^2} \cos \frac{m\pi t}{L}$$

Observe que para a função $f_{cd}^{(1)}(t) = t$, para $0 \leq t \leq 1$, $f_{01}^{(1)}$, os termos de índice par são iguais a zero e neste caso a série de cossenos de $f_{01}^{(1)}$ é dada por

$$f_{01}^{(1)}(t) = \frac{L}{2} - \frac{4L}{\pi^2} \sum_{m=1}^{\infty} \frac{1}{(2m-1)^2} \cos \frac{(2m-1)\pi t}{L},$$

Para a série de senos temos que para $m = 1, 2, \dots$,

$$\begin{aligned} b_m &= \frac{2}{L} \int_{cL}^{dL} f(t) \operatorname{sen} \frac{m\pi t}{L} dt = \frac{2}{L} \int_{cL}^{dL} t \operatorname{sen} \frac{m\pi t}{L} dt = \frac{2L}{m^2\pi^2} \int_{m\pi c}^{m\pi d} s \operatorname{sen} s ds \\ &= \frac{2L}{m^2\pi^2} (-s \cos s + \operatorname{sen} s) \Big|_{m\pi c}^{m\pi d} \end{aligned}$$

Assim, a série de Fourier de senos de $f_{cd}^{(1)}$ é dada por

$$f_{cd}^{(1)}(t) = \sum_{m=1}^{\infty} b_m \operatorname{sen} \frac{m\pi t}{L} = \frac{2L}{\pi^2} \sum_{m=1}^{\infty} \frac{(-s \cos s + \operatorname{sen} s) \Big|_{m\pi c}^{m\pi d}}{m} \operatorname{sen} \frac{m\pi t}{L}$$

Observe que para a função $f(t) = t$, para $0 \leq t \leq 1$, $f_{01}^{(1)}$, temos que

$$b_m = \frac{2L}{m\pi} (-\cos m\pi) = \frac{(-1)^{m+1}2L}{m\pi}$$

e neste caso a série de cossenos de $f_{01}^{(1)}$ é dada por

$$f_{01}^{(1)}(t) = \sum_{m=1}^{\infty} b_m \operatorname{sen} \frac{m\pi t}{L} = \frac{2L}{\pi} \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \operatorname{sen} \frac{m\pi t}{L}$$

Com os coeficientes das funções destes dois exemplos podemos determinar as séries de Fourier de várias funções que são combinações lineares delas. Isto por que os coeficientes das séries dependem linearmente das funções, ou seja,

$$a_m(\alpha f + \beta g) = \alpha a_m(f) + \beta a_m(g) \quad \text{e} \quad b_m(\alpha f + \beta g) = \alpha b_m(f) + \beta b_m(g).$$

Por exemplo, a função

$$f(t) = \begin{cases} t, & \text{se } 0 \leq t \leq L/2 \\ L - t, & \text{se } L/2 < t \leq L \end{cases}$$

pode ser escrita como

$$f = f_{01/2}^{(1)} + Lf_{1/21}^{(0)} - f_{1/21}^{(1)}.$$

Assim os coeficientes a_m e b_m podem ser calculados como

$$\begin{aligned} a_m(f) &= a_m(f_{01/2}^{(1)}) + La_m(f_{1/21}^{(0)}) - a_m(f_{1/21}^{(1)}) \\ b_m(f) &= b_m(f_{01/2}^{(1)}) + Lb_m(f_{1/21}^{(0)}) - b_m(f_{1/21}^{(1)}) \end{aligned}$$

Coefficientes das Séries de Fourier de Senos e Cossenos de Funções Elementares

$f : [0, L] \rightarrow \mathbb{R}$	$a_m = \frac{2}{L} \int_0^L f(t) \cos \frac{m\pi t}{L} dt$	$b_m = \frac{2}{L} \int_0^L f(t) \sin \frac{m\pi t}{L} dt$
$f_{cd}^{(0)}(t) = \begin{cases} 1, & \text{se } cL \leq t \leq dL \\ 0, & \text{caso contrário} \end{cases}$	$a_0 = 2(d - c)$ $a_m = \frac{2}{m\pi} \sin s \Big _{m\pi c}^{m\pi d}$	$b_m = -\frac{2}{m\pi} \cos s \Big _{m\pi c}^{m\pi d}$
$f_{cd}^{(1)}(t) = \begin{cases} t, & \text{se } cL \leq t \leq dL \\ 0, & \text{caso contrário} \end{cases}$	$a_0 = L(d^2 - c^2)$ $a_m = \frac{2L}{m^2\pi^2} (s \sin s + \cos s) \Big _{m\pi c}^{m\pi d}$	$b_m = \frac{2L}{m^2\pi^2} (-s \cos s + \sin s) \Big _{m\pi c}^{m\pi d}$

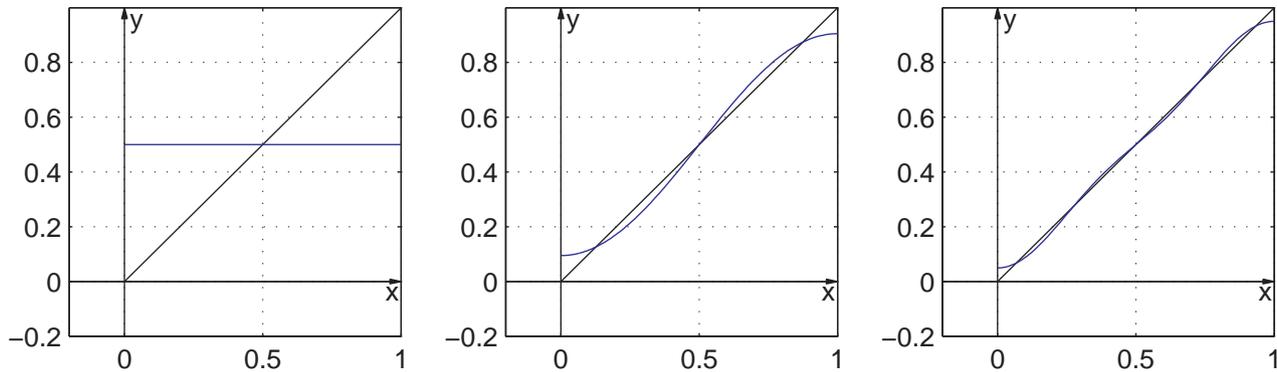


Figura 2.21: A função $f(t) = t$ em $[0, 1]$ e somas parciais da série de Fourier de cossenos para $n = 0, 1, 3$

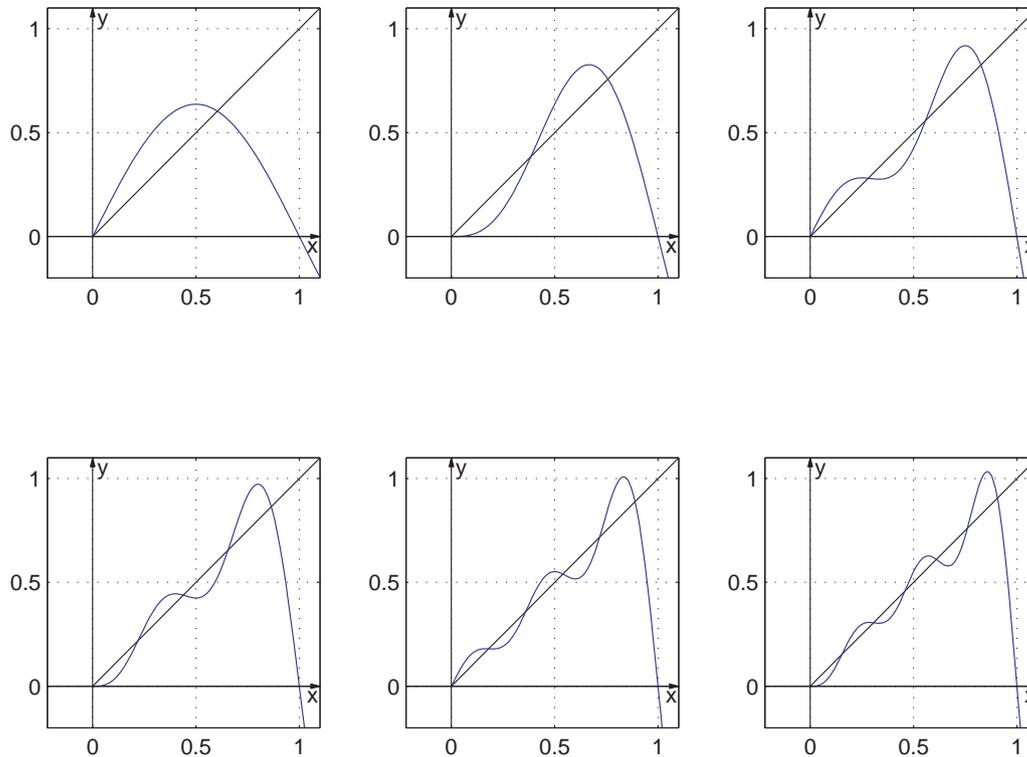


Figura 2.22: A função $f(t) = t$ em $[0, 1]$ e as somas parciais da série de Fourier de senos de f , para $n = 1, \dots, 6$

Exercícios Numéricos (respostas na página 432)

- 2.2.1.** Use o processo de ortogonalização de Gram-Schmidt para encontrar uma base ortonormal para o subespaço de \mathbb{R}^4 que tem como base $\{(1, 1, -1, 0), (0, 2, 0, 1), (-1, 0, 0, 1)\}$.
- 2.2.2.** 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 .
- 2.2.3.** Encontre uma base ortonormal para o subespaço de \mathbb{R}^3 que consiste de todos os vetores (a, b, c) tais que $a + b + c = 0$.
- 2.2.4.** Encontre uma base ortonormal para o subespaço do \mathbb{R}^4 que consiste de todos os vetores (a, b, c, d) tais que $a - b - 2c + d = 0$.
- 2.2.5.** Encontre uma base ortonormal para o espaço solução do sistema homogêneo

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

- 2.2.6.** 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 .
- 2.2.7.** Ache as equações dos planos em \mathbb{R}^3 ortogonais ao vetor $(2, 2, 2)$, que distam $\sqrt{3}$ do ponto $(1, 1, 1)$. Estes planos são subespaços de \mathbb{R}^3 ? Caso afirmativo, encontre uma base para eles.
- 2.2.8.** 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 .
- 2.2.9.** Seja $\mathbb{V} = C^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

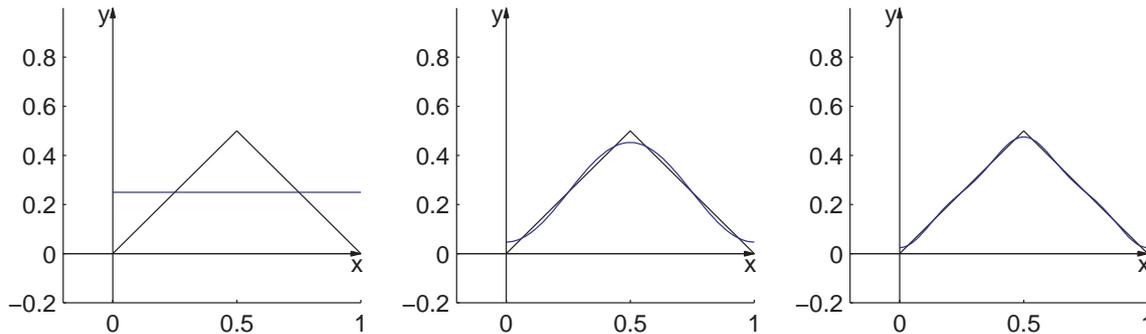


Figura 2.23: A função $f : [0,1] \rightarrow \mathbb{R}$, dada por $f(t) = t$ se $t \in [0,1/2]$ e $f(t) = 1 - t$ se $t \in [1/2,1]$ e somas parciais da série de Fourier de cossenos para $n = 0, 2, 6$

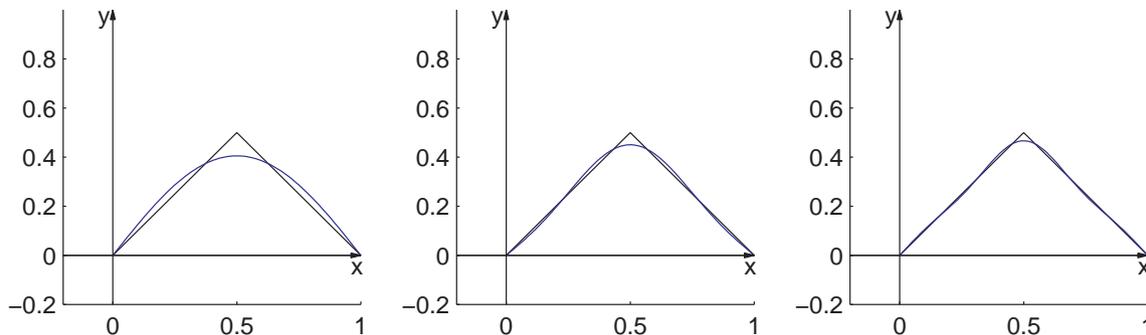


Figura 2.24: A função $f : [0,1] \rightarrow \mathbb{R}$, dada por $f(t) = t$ se $t \in [0,1/2]$ e $f(t) = 1 - t$ se $t \in [1/2,1]$ e somas parciais da série de Fourier de senos para $n = 1, 3, 5$

- (a) Sabendo-se que $\{p_0(t) = 1\}$ é base ortogonal de \mathcal{P}_0 , $\{p_0(t) = 1, p_1(t) = t\}$ é base ortogonal de \mathcal{P}_1 e $\{p_0(t) = 1, p_1(t) = t, p_2(t) = t^2 - 1/3\}$ é base ortogonal de \mathcal{P}_2 , calcule as projeções de $\sqrt{t+1}$ nos polinômios p_0, p_1 e p_2 .
- (b) Calcule as projeções da função $f(t) = \sqrt{t+1} \in \mathbb{V}$ nos subespaços $\mathcal{P}_0, \mathcal{P}_1$ e \mathcal{P}_2 , dos polinômios de grau menor ou igual 0, 1 e 2, respectivamente.

2.2.10. Seja $\mathbb{V} = \mathcal{C}^0[0, 4]$ o conjunto das funções contínuas do intervalo $[0, 4]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_0^4 f(t)g(t)dt.$$

- (a) Determine uma base ortogonal para os subespaços

$$\mathbb{W}_2 = \{p \in \mathcal{P}_2 \mid p(0) = 0\} \quad \text{e} \quad \mathbb{W}_3 = \{p \in \mathcal{P}_3 \mid p(0) = 0\}$$

em que \mathcal{P}_n é o espaço das funções polinomiais de grau menor ou igual n .

- (b) Calcule a projeção da função $f(t) = \sqrt{t} \in \mathbb{V}$ nos subespaços \mathbb{W}_2 e \mathbb{W}_3 .

Seja $\mathbb{V} = \mathcal{C}^0[0, 1]$ o conjunto das funções contínuas do intervalo $[0, 1]$ em \mathbb{R} com o produto interno definido por $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$. Determine as projeções das funções dadas em relação aos subespaços

$$\mathbb{W}_n = [1, \cos \pi t, \dots, \cos n\pi t] \quad \text{e} \quad \mathbb{V}_n = [\sin \pi t, \dots, \sin n\pi t].$$

2.2.11. $f(t) = \begin{cases} 0, & \text{se } 0 \leq t < 1/2, \\ 1, & \text{se } 1/2 \leq t \leq 1, \end{cases}$

2.2.12. $f(t) = \begin{cases} 1, & \text{se } 1/4 \leq t < 3/4, \\ 0, & \text{caso contrário,} \end{cases}$

2.2.13. $f(t) = \begin{cases} 0, & \text{se } 0 \leq t < 1/2, \\ t, & \text{se } 1/2 \leq t < 1, \end{cases}$

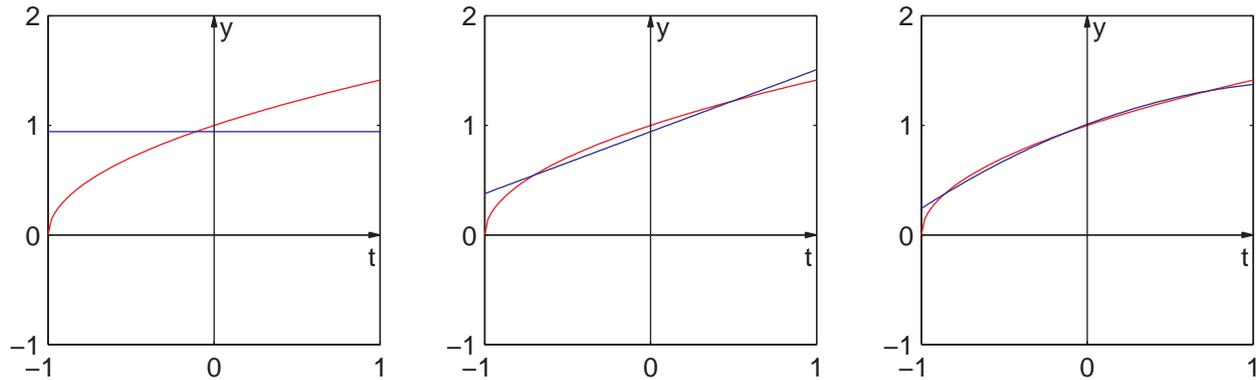


Figura 2.25: Projeções da função $f(t) = \sqrt{t+1}$ nos subespaços \mathcal{P}_n dos polinômios de grau menor ou igual a n , para $n = 0, 1, 2$.

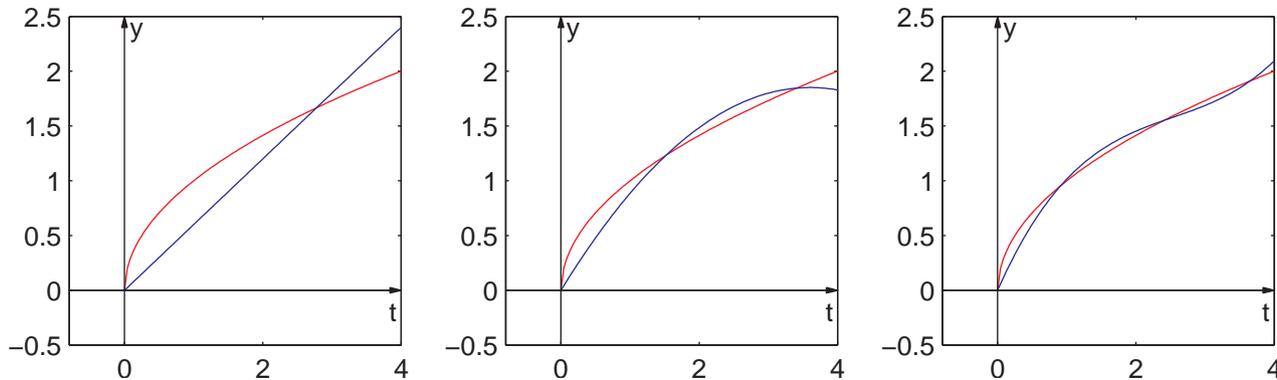


Figura 2.26: Projeções de $f(t) = \sqrt{t}$ nos subespaços $\mathbb{W}_2 = \{p \in \mathcal{P}_2 \mid p(0) = 0\}$ e $\mathbb{W}_3 = \{p \in \mathcal{P}_3 \mid p(0) = 0\}$

$$2.2.14. f(t) = \begin{cases} t, & \text{se } 0 \leq t < 1/2 \\ 1-t, & \text{se } 1/2 \leq t \leq 1 \end{cases}$$

$$2.2.15. f(t) = \begin{cases} t, & \text{se } 0 \leq t < 1/4 \\ 1/4, & \text{se } 1/4 \leq t < 3/4 \\ 1-t, & \text{se } 3/4 < t \leq 1 \end{cases}$$

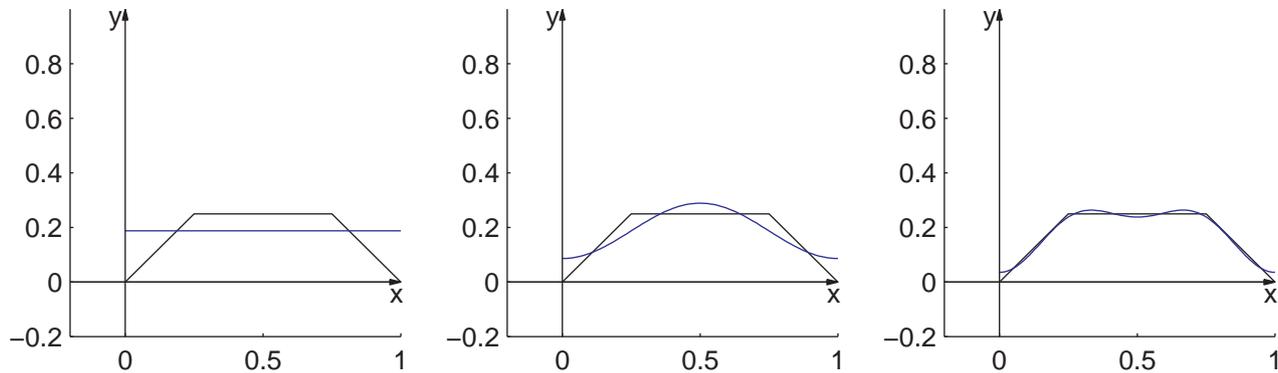


Figura 2.27: A função $f : [0,1] \rightarrow \mathbb{R}$ definida por $f(t) = t$, se $t \in [0,1/4]$, $f(t) = 1/4$, se $t \in [1/4,3/4]$ e $f(t) = 1 - t$, se $t \in [3/4,1]$ e somas parciais da série de Fourier de cossenos para $n = 0,1,2$

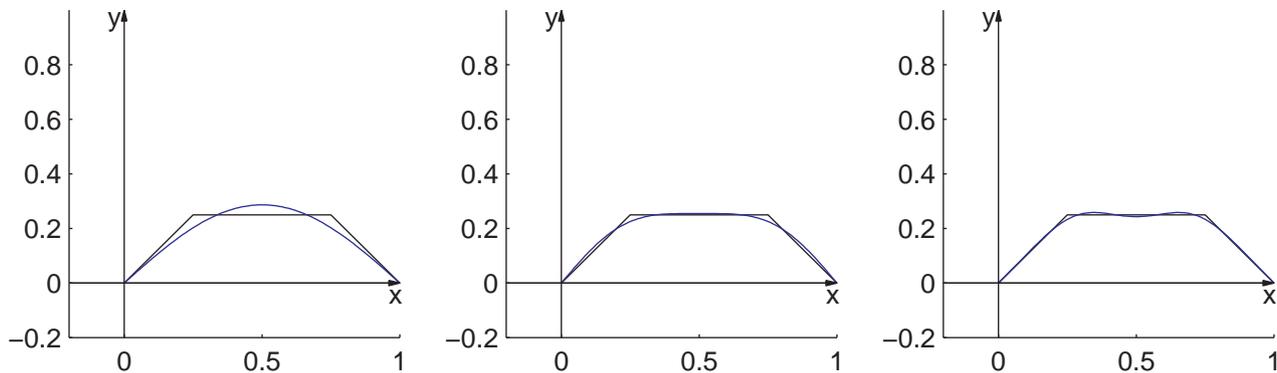


Figura 2.28: A função $f : [0,1] \rightarrow \mathbb{R}$ definida por $f(t) = t$, se $t \in [0,1/4]$, $f(t) = 1/4$, se $t \in [1/4,3/4]$ e $f(t) = 1 - t$, se $t \in [3/4,1]$ e somas parciais da série de Fourier de senos para $n = 1,3,5$

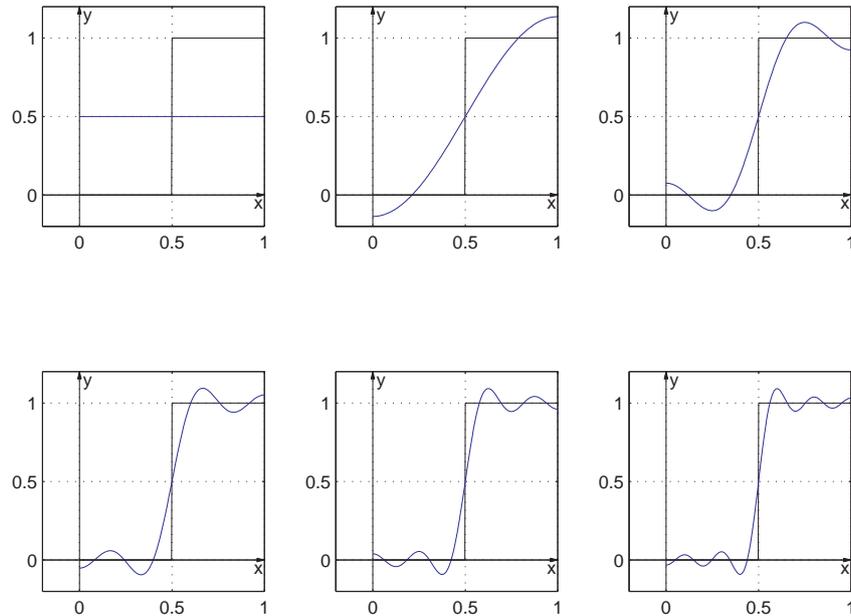


Figura 2.29: A função $f : [0, 1] \rightarrow \mathbb{R}$ definida por $f(t) = 1$, se $t \in [1/2, 1]$ e $f(t) = 0$, caso contrário e as somas parciais da série de Fourier de cossenos de f , para $n = 0, 1, 3, 5, 7, 9$

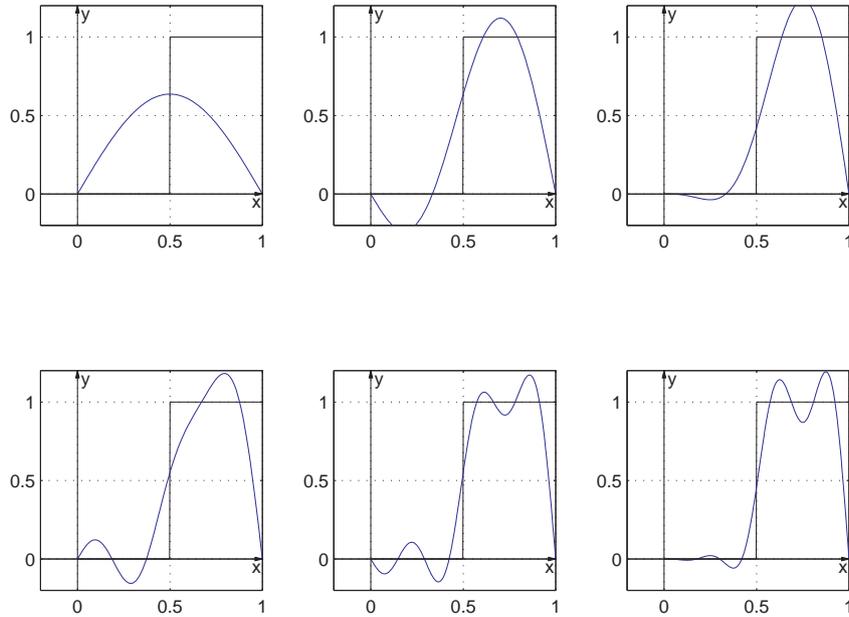


Figura 2.30: A função $f : [0, 1] \rightarrow \mathbb{R}$ definida por $f(t) = 1$, se $t \in [1/2, 1]$ e $f(t) = 0$, caso contrário e as somas parciais da série de Fourier de senos de f , para $n = 1, 2, 3, 5, 6, 7$

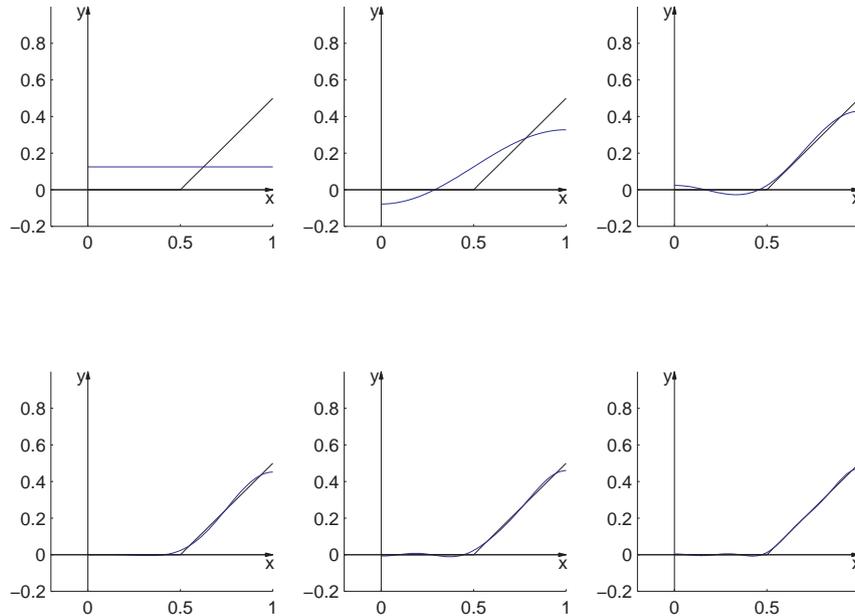


Figura 2.31: A função $f : [0, 1] \rightarrow \mathbb{R}$ definida por $f(t) = t - 1/2$, se $t \in [1/2, 1]$ e $f(t) = 0$, caso contrário e as somas parciais da série de Fourier de cossenos de f , para $n = 0, 1, 2, 3, 5, 6$

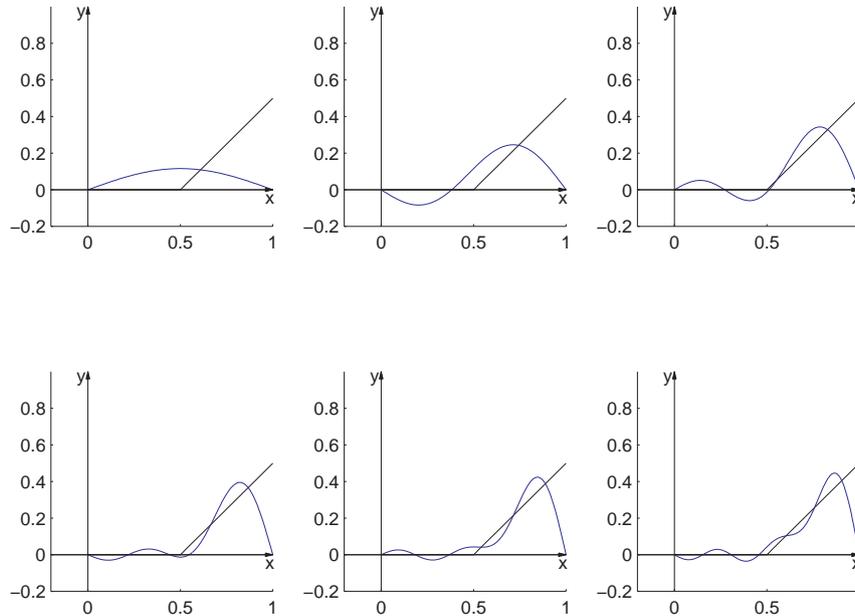


Figura 2.32: A função $f : [0, 1] \rightarrow \mathbb{R}$ definida por $f(t) = t - 1/2$, se $t \in [1/2, 1]$ e $f(t) = 0$, caso contrário e as somas parciais da série de Fourier de senos de f , para $n = 1, 2, 3, 4, 5, 6$

Exercícios usando o MATLAB[®]

Comandos do MATLAB[®]:

>> $V(i)=[]$ elimina a componente i do vetor V .

>> `syms t` diz ao MATLAB[®] que a variável t é uma variável simbólica.

>> `f=expr` define uma função através da `expr` que deve ser uma expressão na variável simbólica t definida anteriormente.

>> `prj=sum(p(1:k))` soma as componentes de 1 a k do vetor $p=[p_1; \dots; p_n]$.

Comandos do pacote GAAL:

>> `proj(g,f,a,b)` calcula a projeção de $f(t)$ em $g(t)$ com relação ao produto interno

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt.$$

>> `p=plendre(n,a,b)` calcula os n primeiros polinômios

$$Q_n(t) = P_n \left(\frac{2t - a - b}{b - a} \right),$$

em que $P_n(t)$ é o polinômio de Legendre de grau n . Assim, são válidas

$$\langle Q_n, Q_m \rangle = 0, \text{ se } m \neq n \quad \text{e} \quad \|Q_n\|^2 = \frac{2}{2n+1}$$

em relação ao produto interno

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt.$$

>> `p=plendren(n)` faz o mesmo que >> `p=plendren(n,-1,1)`.

>> `plotfproj(f,prj,a,b)` desenha as funções f e $prj(k)$, para k variando de 1 até o tamanho do vetor prj , no intervalo $[a,b]$.

>> `h=heaviside(t)` é a função que vale 0 se $t < 0$ e vale 1 se $t \geq 0$.

2.2.16. Seja $\mathbb{V} = C^0[-1,1]$ com o produto interno

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Considere a função $f(t) = e^t - \alpha e^{2t}$, ($\alpha = \frac{e-e^{-1}}{e^2-e^{-2}}$). Seja

$$\mathbb{W}_n = [1, \cos \pi t, \sin \pi t, \dots, \cos n\pi t, \sin n\pi t].$$

- Calcule $\text{proj}_{\mathbb{W}_n} f$, para $n = 0, 1, 2$. Faça os gráficos de f e das projeções.
- Seja \mathcal{P}_n o espaço dos polinômios de grau menor ou igual a n , para $n = 0, 1, 2, \dots$. Calcule $\text{proj}_{\mathcal{P}_n} f$, para $n = 0, 1, 2$. Faça os gráficos de f e das projeções.
- Discuta se é melhor projetar f em \mathbb{W}_n ou em \mathcal{P}_n .

Exercícios Teóricos

2.2.17. Mostre que se W_1, \dots, W_k são vetores não nulos ortogonais e $X = \alpha_1 W_1 + \dots + \alpha_k W_k$, então

$$X = \text{proj}_{W_1} X + \dots + \text{proj}_{W_k} X.$$

2.2.18. Sejam V_1, \dots, V_k vetores linearmente dependentes. Mostre que, aplicando-se o processo de ortogonalização de Gram-Schmidt aos vetores V_1, \dots, V_k , se obtém um vetor W_i que é nulo, para algum $i = 1, \dots, k$. (Sugestão: Seja V_i o primeiro vetor tal que $V_i \in [V_1, \dots, V_{i-1}] = [W_1, \dots, W_{i-1}]$ e use o exercício anterior.)

2.2.19. Seja \mathbb{W} um subespaço de um espaço vetorial com produto interno \mathbb{V} gerado pelos vetores V_1, \dots, V_k . Mostre que $V \in \mathbb{W}^\perp$ se, e somente se, V é ortogonal a V_i , para $i = 1, \dots, k$.

2.2.20. Sejam \mathbb{W}_1 e \mathbb{W}_2 subespaços de um espaço vetorial de dimensão finita com produto interno \mathbb{V} . Mostre que $(\mathbb{W}_1 + \mathbb{W}_2)^\perp = \mathbb{W}_1^\perp \cap \mathbb{W}_2^\perp$ e que $(\mathbb{W}_1 \cap \mathbb{W}_2)^\perp = \mathbb{W}_1^\perp + \mathbb{W}_2^\perp$.

2.2.21. Sejam \mathcal{S} e \mathcal{S}_0 subconjuntos de um espaço vetorial com produto interno \mathbb{V} . Mostre que $\mathcal{S}_0 \subseteq \mathcal{S}$ implica que $\mathcal{S}^\perp \subseteq \mathcal{S}_0^\perp$.

2.2.22. (Desigualdade de Bessel) Seja $\{U_1, \dots, U_m\}$ uma base ortonormal de um subespaço \mathbb{W} de um espaço vetorial com produto interno \mathbb{V} . Mostre que $\|V\|^2 \geq \sum_{j=1}^m |\langle V, U_j \rangle|^2, \forall V \in \mathbb{V}$.

2.2.23. Seja $\mathbb{V} = \mathcal{C}^0[-1, 1]$ o conjunto das funções contínuas do intervalo $[-1, 1]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

Prove a seguinte fórmula de recorrência para os polinômios $p_n(t)$ que são obtidos aplicando-se o processo de ortogonalização de Gram-Schmidt aos polinômios $\{1, t, t^2, \dots, t^n\}$:

$$p_{n+1}(t) = tp_n(t) - \frac{n^2}{(2n+1)(2n-1)}p_{n-1}(t).$$

Para isso siga os seguintes passos:

(a) Pela **Proposição 2.7 na página 122**, $tp_n(t)$ pode ser escrito em termos de p_0, \dots, p_{n+1} .

$$tp_n(t) = \sum_{k=0}^{n+1} \alpha_{nk} p_k(t), \quad (2.18)$$

$$\text{onde } \alpha_{nk} = \frac{\langle tp_n(t), p_k(t) \rangle}{\|p_k\|^2} = \frac{\langle p_n(t), tp_k(t) \rangle}{\|p_k\|^2}.$$

Os coeficientes $\alpha_{nk} = 0$, para $k < n - 1$, pois neste caso $tp_k(t) \in \mathcal{P}_{n-1}$ e da mesma forma que em (2.10) p_n é ortogonal a todo polinômio de \mathcal{P}_{n-1} . O coeficiente $\alpha_{nn} = 0$, pois $t(p_n(t))^2$ é uma função ímpar. O coeficiente $\alpha_{n(n+1)} = 1$, pois os coeficientes de grau $n + 1$ nos dois membros de (2.18) são iguais a 1. Mostre que

$$\alpha_{n(n-1)} = \frac{\langle p_n(t), tp_{n-1}(t) \rangle}{\|p_{n-1}\|^2} = \frac{\alpha_{(n-1)n} \|p_n\|^2}{\|p_{n-1}\|^2} = \frac{\|p_n\|^2}{\|p_{n-1}\|^2} = \frac{n^2}{(2n+1)(2n-1)} \quad (2.19)$$

(b) Substituindo-se os valores de α_{nk} encontrados no item anterior, em (2.18), obtenha a fórmula de recorrência que dá o polinômio p_{n+1} em função de p_n e p_{n-1}

2.2.24. Seguindo os mesmos passos do exercício anterior mostre a seguinte fórmula de recorrência para os polinômios de Legendre:

$$P_{n+1}(t) = \frac{2n+1}{n+1} tP_n(t) - \frac{n}{n+1} P_{n-1}(t)$$

2.2.25. Seja $\mathbb{V} = C^0[a, b]$ o conjunto das funções contínuas do intervalo $[a, b]$ em \mathbb{R} com o produto interno definido por

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt.$$

(a) Mostre que os polinômios $Q_n(t) = P_n\left(\frac{2t-a-b}{b-a}\right)$ são ortogonais e que $\|Q_n\| = \sqrt{\frac{2}{2n+1}}$, em que $P_n(t)$ são os polinômios de Legendre, para $n = 0, 1, 2, \dots$

(b) Mostre que o conjunto

$$\left\{1, \cos \frac{2\pi t}{b-a}, \operatorname{sen} \frac{2\pi t}{b-a}, \dots, \cos \frac{2n\pi t}{b-a}, \operatorname{sen} \frac{2n\pi t}{b-a}, \dots\right\}$$

é ortogonal.

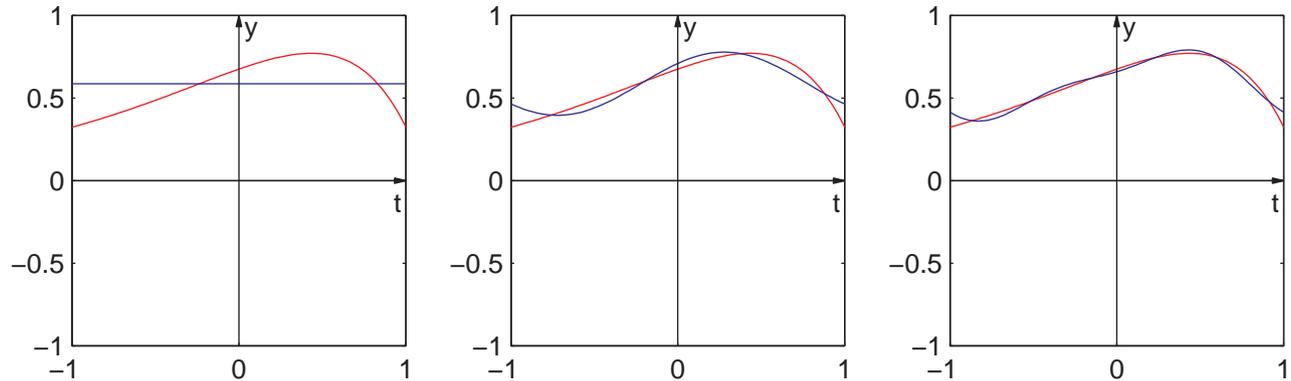


Figura 2.33: Projeções da função $f(t) = e^t - \alpha e^{2t}$, ($\alpha = \frac{e - e^{-1}}{e^2 - e^{-2}}$) nos subespaços \mathbb{W}_n das combinações lineares das funções $1, \cos \pi t, \sin \pi t, \dots, \cos n\pi t, \sin n\pi t$, para $n = 0, 1, 2$.

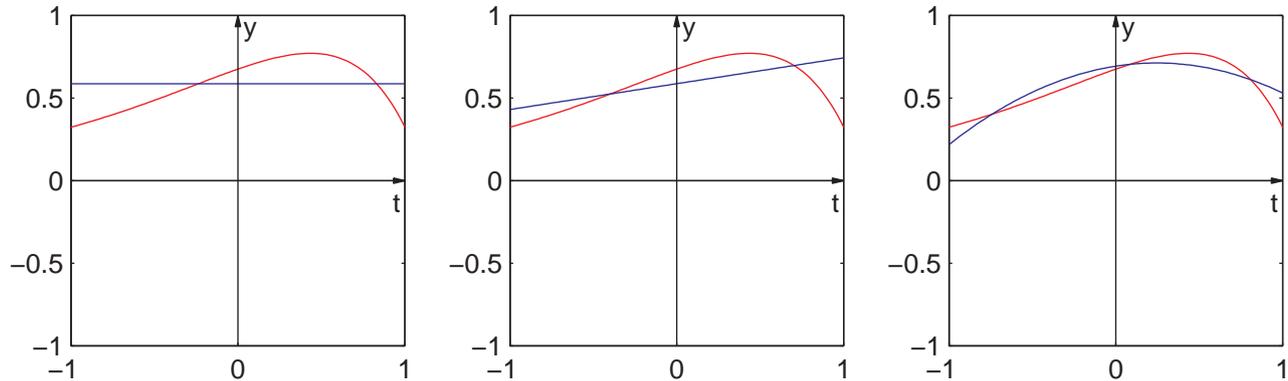


Figura 2.34: Projeções da função $f(t) = e^t - \alpha e^{2t}$, ($\alpha = \frac{e - e^{-1}}{e^2 - e^{-2}}$) nos subespaços \mathcal{P}_n dos polinômios de grau menor ou igual a n , para $n = 0, 1, 2$.

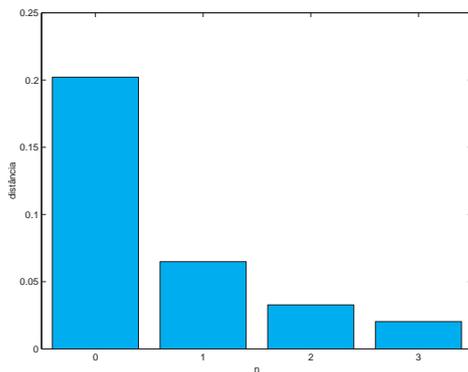


Figura 2.35: Distância de $f(t) = e^t - \alpha e^{2t}$, $(\alpha = \frac{e-e^{-1}}{e^2-e^{-2}})$ ao subespaço gerado por $1/\sqrt{2}, \cos \pi t, \sin \pi t, \dots, \cos n\pi t, \sin n\pi t$

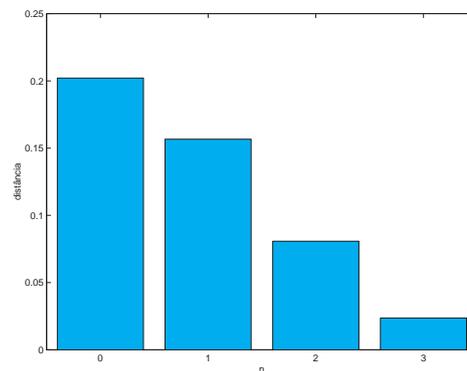


Figura 2.36: Distância de $f(t) = e^t - \alpha e^{2t}$, $(\alpha = \frac{e-e^{-1}}{e^2-e^{-2}})$ ao subespaço \mathcal{P}_n

Capítulo 3

Transformações Lineares

3.1 Definição, Exemplos e Propriedades

3.1.1 Definição e Exemplos

Lembramos que uma **função f de um conjunto A em um conjunto B** , $f : A \rightarrow B$, 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 3.1. Sejam \mathbb{V} e \mathbb{W} espaços vetoriais. Uma função $T : \mathbb{V} \rightarrow \mathbb{W}$ é uma **transformação linear** se

$$\boxed{T(\alpha X) = \alpha T(X)} \quad \text{e} \quad \boxed{T(X + Y) = T(X) + T(Y)}, \quad (3.1)$$

para todos $X, Y \in \mathbb{V}$ e todos os escalares α .

Exemplo 3.1. Sejam \mathbb{V} e \mathbb{W} espaços vetoriais. A função O , que leva todo vetor de \mathbb{V} no vetor nulo de \mathbb{W} , ou seja, $O(X) = \vec{0}$, para todo $X \in \mathbb{V}$, é claramente uma transformação linear e é chamada a **transformação linear nula**. Também a **transformação identidade**, I , de \mathbb{V} em \mathbb{V} que leva todo vetor de \mathbb{V} nele mesmo, ou seja, $I(X) = X$, para todo $X \in \mathbb{V}$ é uma transformação linear.

Exemplo 3.2. Sejam $P_x, P_y : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ as funções que levam todo vetor nas suas projeções nos eixos x e y , respectivamente, ou seja,

$$P_x(x, y) = (x, 0) \text{ e } P_y(x, y) = (0, y), \text{ para todo par } (x, y) \in \mathbb{R}^2.$$

Deixamos para o leitor a verificação de que P_x e P_y são transformações lineares.

Exemplo 3.3. Sejam $R_x, R_y : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 3.4. 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, $P_r : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por

$$P_r(x, y) = \text{proj}_{(a,b)}(x, y) = \frac{(a, b) \cdot (x, y)}{\|(a, b)\|^2} (a, b).$$

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 3.6**.

Exemplo 3.5. 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

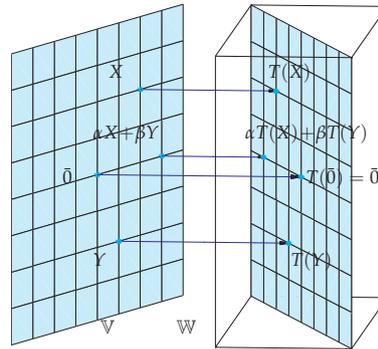


Figura 3.1: Transformação Linear $T : V \rightarrow W$

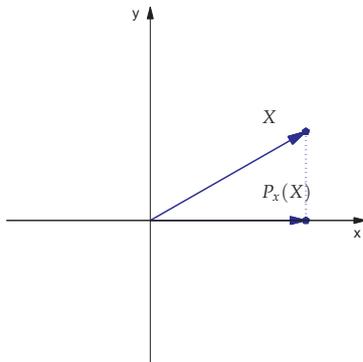


Figura 3.2: Projeção no eixo x

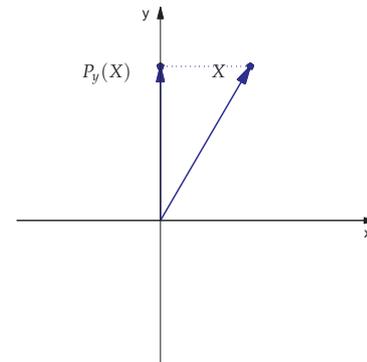


Figura 3.3: Projeção no eixo y

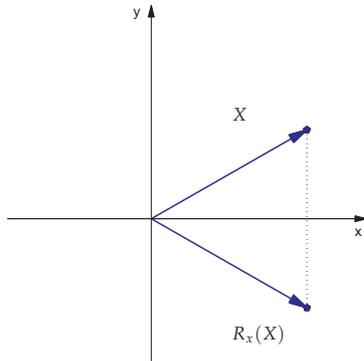


Figura 3.4: Reflexão em relação ao eixo x

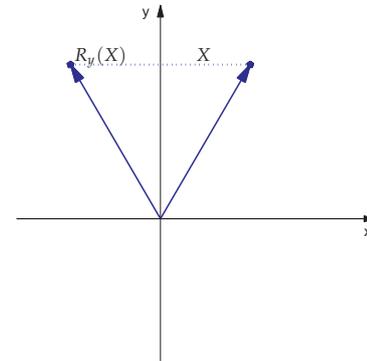


Figura 3.5: Reflexão em relação ao eixo y

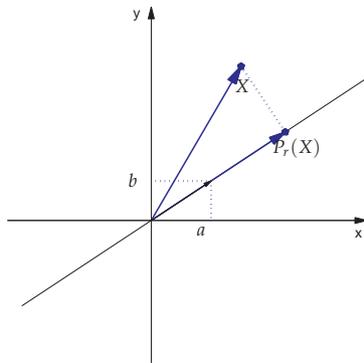


Figura 3.6: Projeção na reta r

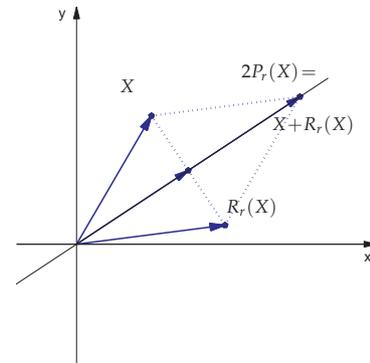


Figura 3.7: Reflexão em relação a reta r

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 3.6**.

Os quatro últimos exemplos são um caso particular do que é apresentado no próximo exemplo.

Exemplo 3.6. Considere a transformação $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, dada por

$$T(X) = T \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{bmatrix}, \quad \text{para todo } X \in \mathbb{R}^n,$$

que pode ser escrita como

$$T(X) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A X,$$

em que $A = (a_{ij})_{m \times n}$. Segue das propriedades da álgebra matricial, 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

$$P_r \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a^2}{a^2+b^2} & \frac{ab}{a^2+b^2} \\ \frac{ab}{a^2+b^2} & \frac{b^2}{a^2+b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad R_r \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a^2-b^2}{a^2+b^2} & \frac{2ab}{a^2+b^2} \\ \frac{2ab}{a^2+b^2} & \frac{b^2-a^2}{a^2+b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Exemplo 3.7. Defina $D : \mathcal{C}^1[a, b] \rightarrow \mathcal{C}^0[a, b]$ por

$$D(f) = f' \quad (\text{a derivada de } f), \text{ para toda função } f \in \mathcal{C}^1[a, b].$$

D é uma transformação linear pois

$$D(f + g) = (f + g)' = f' + g' = D(f) + D(g),$$

$$D(\alpha f) = (\alpha f)' = \alpha f' = \alpha D(f).$$

para todas as funções $f, g \in \mathcal{C}^1[a, b]$ e todos os escalares α .

Exemplo 3.8. Seja $T : \mathcal{C}^0[a, b] \rightarrow \mathcal{C}^1[a, b]$ definida por

$$T(f)(x) = \int_a^x f(t)dt \quad \text{para } a \leq x \leq b, \text{ para toda função } f \in \mathcal{C}^0[a, b].$$

T é uma transformação linear pois

$$\begin{aligned} T(f+g)(x) &= \int_a^x (f+g)(t)dt = \int_a^x (f(t) + g(t))dt \\ &= \int_a^x f(t)dt + \int_a^x g(t)dt \\ &= T(f)(x) + T(g)(x) = (T(f) + T(g))(x), \end{aligned}$$

$$\begin{aligned} T(\alpha f)(x) &= \int_a^x (\alpha f)(t)dt = \int_a^x (\alpha f(t))dt \\ &= \alpha \int_a^x f(t)dt = \alpha(T(f)(x)) = (\alpha T(f))(x), \end{aligned}$$

para $a \leq x \leq b$, todas as funções $f, g \in \mathcal{C}^0[a, b]$ e todos os escalares α .

3.1.2 Propriedades

Segue da [Definição 3.1](#) que toda transformação linear $T : \mathbb{V} \rightarrow \mathbb{W}$ leva o vetor nulo de \mathbb{V} no vetor nulo de \mathbb{W} . Pois, se X é um vetor qualquer de \mathbb{V} , então

$$T(\vec{0}) = T(0X) = 0T(X) = \vec{0}.$$

Segue também da [Definição 3.1](#) que uma função $T : \mathbb{V} \rightarrow \mathbb{W}$ é uma transformação linear se, e somente se,

$$T(\alpha X + \beta Y) = \alpha T(X) + \beta T(Y),$$

para todos os vetores $X, Y \in \mathbb{V}$ 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 $X, Y \in \mathbb{V}$ 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.

Sejam $T : \mathbb{V} \rightarrow \mathbb{W}$ e $S : \mathbb{V} \rightarrow \mathbb{W}$ duas transformações lineares. Vamos mostrar que a soma delas também é uma transformação linear.

$$\begin{aligned} (T + S)(\alpha X + \beta Y) &= T(\alpha X + \beta Y) + S(\alpha X + \beta Y) \\ &= \alpha T(X) + \beta T(Y) + \alpha S(X) + \beta S(Y) \\ &= \alpha(T(X) + S(X)) + \beta(T(Y) + S(Y)) \\ &= \alpha(T + S)(X) + \beta(T + S)(Y) \end{aligned}$$

É fácil ver que para qualquer escalar α , a transformação αT é também uma transformação linear. Assim, o conjunto das transformações lineares de \mathbb{V} em \mathbb{W} , denotado por $\mathcal{L}(\mathbb{V}; \mathbb{W})$, é um subespaço do espaço das funções de \mathbb{V} em \mathbb{W} , $\mathcal{F}(\mathbb{V}; \mathbb{W})$ ([Exercício 1.1.7 na página 12](#)). Portanto, o conjunto $\mathcal{L}(\mathbb{V}; \mathbb{W})$ é um espaço vetorial.

Quando \mathbb{W} é o conjunto de escalares (\mathbb{R} ou \mathbb{C}), que vamos denotar por \mathbb{K} , o conjunto $\mathcal{L}(\mathbb{V}; \mathbb{K})$ é também denotado por \mathbb{V}^* e é chamado o **espaço dual de \mathbb{V}** . Os elementos de \mathbb{V}^* são chamados **funcionais lineares de \mathbb{V}** .

Se $T : \mathbb{V} \rightarrow \mathbb{W}$ é uma transformação linear e \mathcal{B} é uma base de \mathbb{V} , então todo vetor $V \in \mathbb{V}$ pode ser escrito como combinação linear de vetores de \mathcal{B} , ou seja, existem vetores $V_1, \dots, V_n \in \mathcal{B}$ e escalares $\alpha_1, \dots, \alpha_n$ tais que

$$V = \alpha_1 V_1 + \dots + \alpha_n V_n = \sum_{i=1}^n \alpha_i V_i.$$

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 $U : \mathbb{V} \rightarrow \mathbb{W}$ é outra transformação linear tal que $U(V_\lambda) = T(V_\lambda)$ para todo $V_\lambda \in \mathcal{B}$, então aplicando-se o raciocínio acima, temos que

$$U(V) = T(V), \quad \text{para todo } V \in \mathbb{V}.$$

Ou seja, $U = T$. Isto prova o seguinte teorema.

Teorema 3.1. *Sejam \mathbb{V} e \mathbb{W} espaços vetoriais. Uma transformação linear $T : \mathbb{V} \rightarrow \mathbb{W}$ é totalmente caracterizada pelos seus valores em uma base de \mathbb{V} . Ou seja, se \mathcal{B} é uma base de \mathbb{V} e uma função T está definida para valores em \mathcal{B} ,*

$$T(V_\lambda) = W_\lambda, \quad \text{para todo } V_\lambda \in \mathcal{B}.$$

Então, existe um única transformação linear definida em todo espaço \mathbb{V} , $T : \mathbb{V} \rightarrow \mathbb{W}$, tal que $T(V_\lambda) = W_\lambda$, para todo $V_\lambda \in \mathcal{B}$.

Exemplo 3.9. Seja $R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a transformação linear definida na base canônica por (Figura 3.9)

$$\begin{aligned} R_\theta(E_1) &= \cos \theta E_1 + \text{sen } \theta E_2 = (\cos \theta, \text{sen } \theta) \\ R_\theta(E_2) &= -\text{sen } \theta E_1 + \cos \theta E_2 = (-\text{sen } \theta, \cos \theta). \end{aligned}$$

Assim, como $(x, y) = x(1, 0) + y(0, 1) = xE_1 + yE_2$, então

$$\begin{aligned} R_\theta \begin{bmatrix} x \\ y \end{bmatrix} &= xR_\theta(E_1) + yR_\theta(E_2) = x \begin{bmatrix} \cos \theta \\ \text{sen } \theta \end{bmatrix} + y \begin{bmatrix} -\text{sen } \theta \\ \cos \theta \end{bmatrix} \\ &= \begin{bmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \end{aligned}$$

Segue da linearidade, que R_θ faz uma rotação de um ângulo θ em todo vetor $X = (x, y) \in \mathbb{R}^2$. Pois, se escrevemos $X = (x, y) = (r \cos \alpha, r \text{sen } \alpha)$, então

$$\begin{aligned} R_\theta(X) &= R_\theta(r \cos \alpha, r \text{sen } \alpha) = rR_\theta(\cos \alpha, \text{sen } \alpha) \\ &= r(\cos \alpha \cos \theta - \text{sen } \alpha \text{sen } \theta, \cos \alpha \text{sen } \theta + \text{sen } \alpha \cos \theta) \\ &= r(\cos(\alpha + \theta), \text{sen }(\alpha + \theta)) = (r \cos(\alpha + \theta), r \text{sen }(\alpha + \theta)). \end{aligned}$$

Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma transformação linear tal que

$$T \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \dots, \quad T \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Sejam $X = (x_1, \dots, x_n)$ um vetor qualquer do \mathbb{R}^n e

$$E_1 = (1, 0, \dots, 0), E_2 = (0, 1, 0, \dots, 0), \dots, E_n = (0, \dots, 0, 1).$$

Como $X = x_1 E_1 + \dots + x_n E_n$, então

$$\begin{aligned} T(X) &= x_1 T(E_1) + \dots + x_n T(E_n) = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = AX, \end{aligned}$$

em que as colunas de A são $T(E_1), \dots, T(E_n)$, ou seja, $A = [T(E_1) \dots T(E_n)]$, com $T(E_i)$, para $i = 1, \dots, n$ escritos como matrizes colunas. Isto prova o seguinte resultado.

Proposição 3.2. *Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma transformação linear. Então, T é dada por*

$$T(X) = AX, \quad \text{para todo } X \in \mathbb{R}^n,$$

em que $A = (a_{ij})_{m \times n} = [T(E_1) \dots T(E_n)]$, com $T(E_i)$, para $i = 1, \dots, n$, escritos como matrizes colunas e $E_1 = (1, 0, \dots, 0)$, $E_2 = (0, 1, 0, \dots, 0)$, \dots , $E_n = (0, \dots, 0, 1)$. A matriz A é chamada **matriz da transformação T** (em relação às bases canônicas).

Exemplo 3.10. As 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

$$[P_x(E_1) \quad P_x(E_2)] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad [P_y(E_1) \quad P_y(E_2)] = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

as matrizes das reflexões em relação aos eixos x e y ,

$$[R_x(E_1) \quad R_x(E_2)] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad [R_y(E_1) \quad R_y(E_2)] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$

as matrizes da projeção ortogonal e da reflexão em relação a uma reta

$$r : (x, y) = t(a, b), \text{ como}$$

$$[P_r(E_1) \quad P_r(E_2)] = \begin{bmatrix} \frac{a^2}{a^2+b^2} & \frac{ab}{a^2+b^2} \\ \frac{ab}{a^2+b^2} & \frac{b^2}{a^2+b^2} \end{bmatrix} \quad [R_r(E_1) \quad R_r(E_2)] = \begin{bmatrix} \frac{a^2-b^2}{a^2+b^2} & \frac{2ab}{a^2+b^2} \\ \frac{2ab}{a^2+b^2} & \frac{b^2-a^2}{a^2+b^2} \end{bmatrix}$$

e a matriz da rotação de um ângulo θ ,

$$[R_\theta(E_1) \quad R_\theta(E_2)] = \begin{bmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{bmatrix}.$$

3.1.3 Aplicação: Matriz Jacobiana

Dizemos que uma função $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é **diferenciável** em um ponto $X_0 \in \mathbb{R}^n$ se existe uma transformação linear $T_{X_0} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que

$$\lim_{H \rightarrow \vec{0}} \frac{1}{\|H\|} (T_{X_0}(H) - F(X_0 + H) + F(X_0)) = \vec{0}. \quad (3.2)$$

Este limite é uma generalização do limite de funções reais, significa que para todo $\epsilon > 0$, existe um $\delta > 0$, tal que $\|H\| < \delta$ implica que $\frac{\|T_{X_0}(H) - F(X_0 + H) + F(X_0)\|}{\|H\|} < \epsilon$. No caso em que $m = n = 1$, a transformação

linear é $T_{X_0}(H) = F'(X_0)H$, para todo $H \in \mathbb{R}$. Vamos supor que a função F seja diferenciável em X_0 e vamos determinar a matriz da transformação linear T_{X_0} em relação às bases canônicas de \mathbb{R}^n , \mathcal{B} , e de \mathbb{R}^m , \mathcal{B}' . Para isto, vamos calcular o limite (3.2) segundo o caminho $H = hE_j$, para $j = 1, \dots, n$, $h > 0$, em que $E_1 = (1, 0, \dots, 0)$, $E_2 = (0, 1, 0, \dots, 0)$, \dots , $E_n = (0, \dots, 0, 1)$ são os vetores da base canônica de \mathbb{R}^n , \mathcal{B} .

$$\lim_{h \rightarrow 0^+} \frac{1}{h} (T_{X_0}(hE_j) - F(X_0 + hE_j) + F(X_0)) = \vec{0}.$$

Como T_{X_0} é uma transformação linear, obtemos

$$T_{X_0}(E_j) = \lim_{h \rightarrow 0^+} \frac{1}{h} (F(X_0 + hE_j) - F(X_0)).$$

Calculando-se o limite (3.2) para o caminho $H = hE_j$, para $h < 0$, obtemos

$$T_{X_0}(E_j) = \lim_{h \rightarrow 0^-} \frac{1}{h} (F(X_0 + hE_j) - F(X_0)).$$

Assim, a função $F(X) = (F_1(X), \dots, F_m(X))$ é tal que $F_i(X)$ tem derivadas parciais em X_0 , para $i = 1, \dots, m$ e

$$T_{X_0}(E_j) = \begin{bmatrix} \frac{\partial F_1}{\partial x_j}(X_0) \\ \vdots \\ \frac{\partial F_m}{\partial x_j}(X_0) \end{bmatrix}.$$

Portanto a matriz da transformação linear T_{X_0} em relação às bases canônicas de \mathbb{R}^n , \mathcal{B} , e de \mathbb{R}^m , \mathcal{B}' , é

$$[T_{X_0}]_{\mathcal{B}'}^{\mathcal{B}} = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(X_0) & \dots & \frac{\partial F_1}{\partial x_n}(X_0) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(X_0) & \dots & \frac{\partial F_m}{\partial x_n}(X_0) \end{bmatrix}.$$

Esta matriz é chamada **matriz jacobiana** da função F em X_0 .

Exercícios Numéricos (respostas na página 440)

- 3.1.1. Seja $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ uma transformação linear para a qual sabemos que $T(1, 1) = (2, -3)$ e $T(0, 1) = (1, 2)$.
- Determine $T(3, -2)$;
 - Determine $T(a, b)$.
- 3.1.2. Determine a transformação $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ tal que $T(1, 1) = (3, 2, 1)$ e $T(0, -2) = (0, 1, 0)$. Encontre $T(1, 0)$ e $T(0, 1)$.
- 3.1.3. Determine expressões para as transformações lineares $P_{xy}, P_{yz}, P_{xz} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, que são projeções nos planos xy, yz e xz , respectivamente.
- 3.1.4. Considere o plano $\pi : x + 2y + 3z = 0$. Encontre expressões para as seguintes transformações lineares:
- A projeção ortogonal no plano π , $P_\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.
 - A reflexão em relação ao plano π , $R_\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.
- 3.1.5. 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.
- 3.1.6. Considere a reta $r : (x, y, z) = t(1, 1, 1)$.
- Seja $P_r : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ a projeção ortogonal na reta r . Encontre uma expressão para $P_r(x, y, z)$.
 - Seja $R_r : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ a reflexão em relação à reta r . Encontre uma expressão para $R_r(x, y, z)$.
- 3.1.7. 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

- 3.1.8. Considere o plano $\pi : ax + by + cz = 0$. Encontre expressões para as seguintes transformações lineares:
- A projeção ortogonal no plano π , $P_\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

(b) A reflexão em relação ao plano π , $R_\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

3.1.9. Determine expressões para as transformações lineares $R_{\theta,x}, R_{\theta,y}, R_{\theta,z} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, que são rotações de θ em relação aos eixos x, y e z , respectivamente.

3.1.10. 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 , $P_r : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

(b) A reflexão em relação à reta r , $R_r : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

3.1.11. Seja \mathbb{V} um espaço vetorial de dimensão finita. Seja $\mathcal{B} = \{V_1, \dots, V_n\}$ uma base de \mathbb{V} . Mostre que $\mathcal{B}^* = \{f_1, \dots, f_n\} \subset \mathbb{V}^*$ definido por $f_i(V_j) = \delta_{ij}$, para $i, j = 1, \dots, n$, em que δ_{ij} é o **delta de Kronecker** ($\delta_{ij} = 0$, se $i \neq j$ e $\delta_{ii} = 1$) é uma base de \mathbb{V}^* . Esta base é chamada **base dual de \mathcal{B}** .

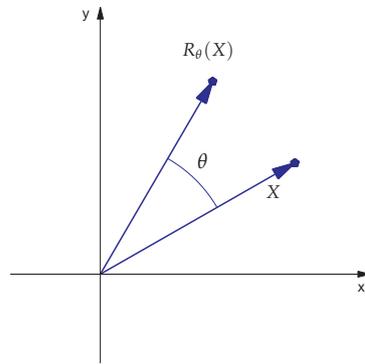


Figura 3.8: Transformação rotação de um ângulo θ

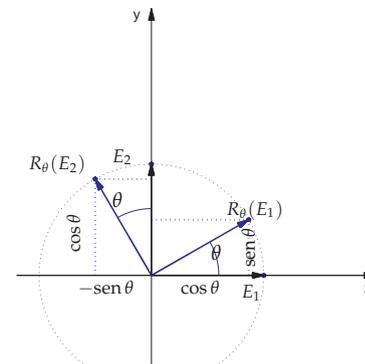


Figura 3.9: Transformação rotação sobre os vetores E_1 e E_2

3.2 A Imagem e o Núcleo

Definição 3.2. Sejam \mathbb{V} e \mathbb{W} espaços vetoriais. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear.

(a) O **núcleo** de T é definido pelo conjunto

$$\mathcal{N}(T) = \{X \in \mathbb{V} \mid T(X) = \bar{0}\}.$$

(b) A **imagem** de T é definida pelo conjunto

$$\mathcal{I}(T) = \{Y \in \mathbb{W} \mid Y = T(X), \text{ para algum } X \in \mathbb{V}\}$$

Exemplo 3.11. Sejam $O : \mathbb{V} \rightarrow \mathbb{W}$ a transformação linear nula e $I : \mathbb{V} \rightarrow \mathbb{V}$ a transformação identidade. Então $\mathcal{N}(O) = \mathbb{V}$, $\mathcal{I}(O) = \{\bar{0}\}$, $\mathcal{N}(I) = \{\bar{0}\}$ e $\mathcal{I}(I) = \mathbb{V}$.

Teorema 3.3. Sejam \mathbb{V} e \mathbb{W} espaços vetoriais. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. O núcleo, $\mathcal{N}(T)$, e a imagem, $\mathcal{I}(T)$ são subespaços de \mathbb{V} e de \mathbb{W} , respectivamente.

Demonstração. Para mostrar que um conjunto é um subespaço precisamos mostrar as propriedades (0) e (0') do **Teorema 1.3 na página 14**. Vamos mostrar em primeiro lugar que o núcleo de T é um subespaço.

(0) 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, $X_1 + X_2 \in \mathcal{N}(T)$;

(0') 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.

(0) Se $Y_1, Y_2 \in \mathcal{I}(T)$, então existem $X_1, X_2 \in \mathbb{V}$ 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, $Y_1 + Y_2 \in \mathcal{I}(T)$.

(0') Se $Y \in \mathcal{I}(T)$ e α é um escalar, então existe $X \in \mathbb{V}$ 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 3.12. A imagem de um funcional linear não nulo $f : \mathbb{V} \rightarrow \mathbb{R}$ é \mathbb{R} , pois $\{\bar{0}\}$ e \mathbb{R} são os únicos subespaços do espaço vetorial \mathbb{R} .

Definição 3.3. Sejam \mathbb{V} e \mathbb{W} espaços vetoriais. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. A dimensão do núcleo de T é chamada de **nulidade** de T e a dimensão da imagem de T é chamada **posto** de T .

Proposição 3.4. *Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Se $\{V_1, \dots, V_n\}$ é uma base de \mathbb{V} , então a imagem de T é gerada por $T(V_1), \dots, T(V_n)$.*

Demonstração. Seja $W \in \mathcal{I}(T)$. Então, existe $V \in \mathbb{V}$ tal que $T(V) = W$. Como $\{V_1, \dots, V_n\}$ é base de \mathbb{V} , existem escalares $\alpha_1, \dots, \alpha_n$ tais que $V = \alpha_1 V_1 + \dots + \alpha_n V_n$. 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, $T(V_1), \dots, T(V_n)$ geram $\mathcal{I}(T)$. ■

Exemplo 3.13. Vamos considerar as projeções nos eixos x e y (Figuras 3.2 e 3.3 na página 206)

$$P_x \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad P_y \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

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 3.14. Vamos considerar as reflexões em relação aos eixos x e y (Figuras 3.4 e 3.5 na página 207)

$$R_x \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad R_y \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

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 3.15. Consideremos a projeção ortogonal e a reflexão em relação a uma reta $r : (x, y) = t(a, b)$ (Figuras 3.6 e 3.7 na página 207)

$$P_r \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a^2}{a^2+b^2} & \frac{ab}{a^2+b^2} \\ \frac{ab}{a^2+b^2} & \frac{b^2}{a^2+b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad R_r \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a^2-b^2}{a^2+b^2} & \frac{2ab}{a^2+b^2} \\ \frac{2ab}{a^2+b^2} & \frac{b^2-a^2}{a^2+b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

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 3.16. Geometricamente vemos que a rotação de um ângulo θ (Figura 3.8 na página 220)

$$R_\theta \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

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 $-\theta$.

Exemplo 3.17. Considere a transformação $D : \mathcal{C}^1[a, b] \rightarrow \mathcal{C}^0[a, b]$ definida por

$$D(f) = f' \quad (\text{a derivada de } f), \text{ para toda função } f \in \mathcal{C}^1[a, b].$$

Vamos mostrar que a imagem de D é igual à $\mathcal{C}^0[a, b]$. Seja $g \in \mathcal{C}^0[a, b]$. A função $f(x) = \int_a^x g(t)dt$ está bem definida, pois g é contínua. Além disso,

$$\begin{aligned} D(f)(x) &= \frac{d}{dx} \left(\int_a^x g(t)dt \right) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\int_a^{x+h} g(t)dt - \int_a^x g(t)dt \right) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\int_x^{x+h} g(t)dt \right) \\ &= \lim_{\xi \rightarrow x} g(\xi) = g(x) \end{aligned}$$

pele Teorema do Valor Médio para integrais. O núcleo de D é igual ao conjunto das funções constantes em $[a, b]$.

Exemplo 3.18. Seja $T : \mathcal{C}^0[a, b] \rightarrow \mathcal{C}^1[a, b]$ definida por

$$T(f)(x) = \int_a^x f(t)dt \quad \text{para } a \leq x \leq b, \text{ para toda função } f \in \mathcal{C}^0[a, b].$$

A imagem de T é igual ao conjunto das funções $g \in \mathcal{C}^1[a, b]$ tais que $g(a) = 0$. Pois para toda função $g \in \mathcal{C}^1[a, b]$, com $g(a) = 0$, a função $f = g'$ é tal que

$$T(f)(x) = T(g')(x) = \int_a^x g'(t)dt = g(x) - g(a) = g(x),$$

pele Teorema Fundamental do Cálculo. O núcleo de T é formado somente pela função nula. Pois, $T(f) = \bar{0}$ implica que $\int_a^x f(t)dt = 0$, para todo $x \in [a, b]$. Assim, $\frac{d}{dx} \left(\int_a^x f(t)dt \right) = f(x) = 0$. Portanto, f é identicamente nula em $[a, b]$.

Exemplo 3.19. Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma transformação linear. Pela [Proposição 3.2 na página 215](#) a transformação T é dada por

$$T(X) = AX, \quad \text{para todo } X \in \mathbb{R}^n,$$

onde $A = (a_{ij})_{m \times n} = [T(E_1) \dots T(E_n)]$, com $T(E_i)$, para $i = 1, \dots, n$ escritos como matrizes colunas e $E_1 = (1, 0, \dots, 0)$, $E_2 = (0, 1, 0, \dots, 0)$, \dots , $E_n = (0, \dots, 0, 1)$. 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 , chamado **espaço coluna de A** .

3.2.1 Espaço Linha e Espaço Coluna de uma Matriz

Definição 3.4. 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 3.20. 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 3.5. *Seja A uma matriz $m \times n$. O espaço linha e o espaço coluna de A possuem a mesma dimensão.*

Demonstração. Seja R a forma escalonada reduzida da matriz A .

- (a) Vamos mostrar que o espaço linha de A é igual ao espaço linha de R .
A matriz R é obtida de A aplicando-se uma sequência de operações elementares às linhas de A . Assim, toda linha de R é uma combinação linear das linhas de A . Então, o espaço linha de R está contido no espaço linha de A . O mesmo argumento mostra que o espaço linha de A está contido no espaço linha de R . Portanto, eles são iguais.
- (b) Vamos mostrar que a dimensão do espaço coluna de A é igual a dimensão do espaço coluna de R . A dimensão do espaço coluna de uma matriz é igual ao número máximo de colunas L.I. As colunas de A , são L.I. se, e somente se, o sistema $AX = \vec{0}$ tem somente a solução trivial. Mas, a solução deste sistema é a mesma do sistema $RX = \vec{0}$. Assim, as colunas de A são L.I. se, e somente se, as colunas de R são L.I. Analogamente, r colunas de A são L.I. se, e somente se, as colunas correspondentes de R são L.I.

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.



Posto de uma Matriz

Definição 3.5. Seja A uma matriz $m \times n$. 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 .

Exemplo 3.21. 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 .

3.2.2 Injetividade e Sobrejetividade

Dizemos que uma função $f : A \rightarrow B$ 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 3.4** o seguinte resultado.

Corolário 3.6. *Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Seja $\{V_1, \dots, V_n\}$ base de \mathbb{V} . T é sobrejetiva se, e somente se, $T(V_1), \dots, T(V_n)$ geram \mathbb{W} .*

Exemplo 3.22. Todo funcional linear não nulo $f : \mathbb{V} \rightarrow \mathbb{R}$ é sobrejetivo, pois $\{\bar{0}\}$ e \mathbb{R} são os únicos subespaços do espaço vetorial \mathbb{R} .

Exemplo 3.23. 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 3.15 e 3.16 na página 224**).

Exemplo 3.24. A transformação $D : \mathcal{C}^1[a, b] \rightarrow \mathcal{C}^0[a, b]$ definida por

$$D(f) = f' \quad (\text{a derivada de } f), \text{ para toda função } f \in \mathcal{C}^1[a, b]$$

é sobrejetiva (**Exemplo 3.17 na página 224**).

Exemplo 3.25. A transformação $T : C^0[a, b] \rightarrow C^1[a, b]$ definida por

$$T(f)(x) = \int_a^x f(t)dt \quad \text{para } a \leq x \leq b, \text{ para toda função } f \in C^0[a, b]$$

não é sobrejetiva, pois a imagem de T é o conjunto das funções $g \in C^1[a, b]$ tais que $g(a) = 0$ (**Exemplo 3.18 na página 225**).

Dizemos que uma função $f : A \rightarrow B$ é **injetiva**, se $f(x) = f(y)$ implica que $x = y$.

Teorema 3.7. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ 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 3.26. 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 3.15 e 3.16 na página 224**).

Exemplo 3.27. A transformação $D : \mathcal{C}^1[a, b] \rightarrow \mathcal{C}^0[a, b]$ definida por

$$D(f) = f' \quad (\text{a derivada de } f), \text{ para toda função } f \in \mathcal{C}^1[a, b]$$

não é injetiva, pois $\mathcal{N}(D)$ é o conjunto das funções constantes (Exemplo 3.17 na página 224).

Exemplo 3.28. Seja $T : \mathcal{C}^0[a, b] \rightarrow \mathcal{C}^1[a, b]$ definida por

$$T(f)(x) = \int_a^x f(t) dt \quad \text{para } a \leq x \leq b, \text{ para toda função } f \in \mathcal{C}^0[a, b]$$

é injetiva, pois $\mathcal{N}(T) = \{\bar{0}\}$ (Exemplo 3.18 na página 225).

Teorema 3.8 (da Dimensão do Núcleo e da Imagem). *Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Se \mathbb{V} tem dimensão finita, então a soma da dimensão do núcleo de T com a dimensão da imagem de T é igual a dimensão de \mathbb{V} , ou seja,*

$$\dim(\mathcal{N}(T)) + \dim(\mathcal{I}(T)) = \dim(\mathbb{V}).$$

Demonstração. Vamos supor que $1 \leq \dim(\mathcal{N}(T)) < \dim(\mathbb{V})$. Sejam V_1, \dots, V_p vetores de \mathbb{V} , que formam uma base para o núcleo de T . Vamos estendê-la a uma base de \mathbb{V} . Sejam V_{p+1}, \dots, V_n vetores de \mathbb{V} tais que $V_1, \dots, V_p, V_{p+1}, \dots, V_n$ formam uma base de \mathbb{V} . Vamos mostrar que $T(V_{p+1}), \dots, T(V_n)$ formam uma base da imagem de T . Para isso, precisamos mostrar que eles geram a imagem de T e que são L.I.

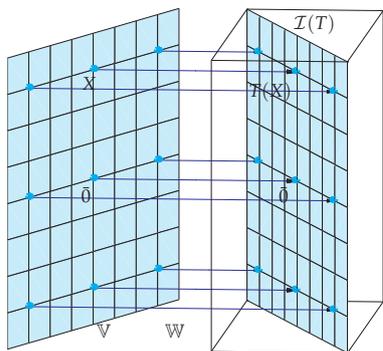


Figura 3.10: $\mathcal{N}(T) = \{\vec{0}\}$

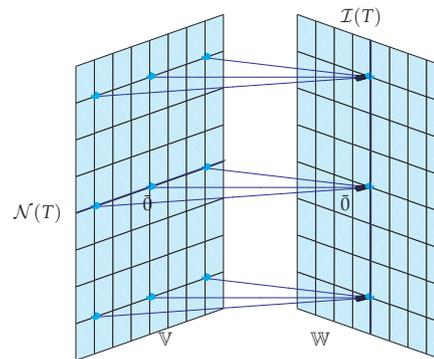


Figura 3.11: $\mathcal{N}(T) \neq \{\vec{0}\}$

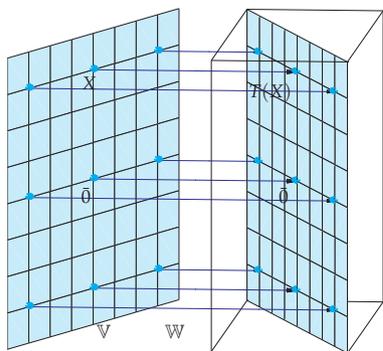


Figura 3.12: Transformação linear injetiva
($\mathcal{N}(T) = \{\vec{0}\}$)

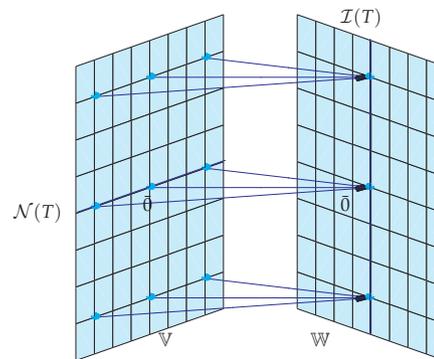


Figura 3.13: Transformação linear não injetiva
($\mathcal{N}(T) \neq \{\vec{0}\}$)

Pela [Proposição 3.4 na página 223](#), $T(V_{p+1}), \dots, T(V_n)$ geram a imagem de T , pois $T(V_1) = \dots = T(V_p) = \bar{0}$. Vamos mostrar que $T(V_{p+1}), \dots, T(V_n)$ são linearmente independentes. Considere a combinação linear nula:

$$x_{p+1}T(V_{p+1}) + \dots + x_nT(V_n) = \bar{0}.$$

Pela linearidade de T segue-se que

$$T(x_{p+1}V_{p+1} + \dots + x_nV_n) = \bar{0}.$$

Mas isto implica que $x_{p+1}V_{p+1} + \dots + x_nV_n \in \mathcal{N}(T)$. Assim, existem escalares y_1, \dots, y_p tais que $x_{p+1}V_{p+1} + \dots + x_nV_n = y_1V_1 + \dots + y_pV_p$. De onde segue-se que

$$y_1V_1 + \dots + y_pV_p - x_{p+1}V_{p+1} - \dots - x_nV_n = \bar{0}.$$

Como V_1, \dots, V_n é base de \mathbb{V} , então $y_1 = \dots = y_p = x_{p+1} = \dots = x_n = 0$, ou seja, $T(V_{p+1}), \dots, T(V_n)$ são L.I.

Portanto, a dimensão da imagem de T é igual a diferença entre a dimensão de \mathbb{V} e a dimensão do núcleo de A , de onde segue-se o resultado. Deixamos como exercício para o leitor os casos em que $\dim(\mathcal{N}(T)) = 0$ e $\dim(\mathcal{N}(T)) = \dim(\mathbb{V})$. ■

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 3.8 da Dimensão do Núcleo e da Imagem](#).

Corolário 3.9. *Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear entre espaços de dimensão finita. Suponha que $\dim(\mathbb{V}) = \dim(\mathbb{W})$.*

- (a) Se T é sobrejetiva, então T é injetiva.
(b) Se T é injetiva, então T é sobrejetiva.
-

Se uma transformação linear $T : \mathbb{V} \rightarrow \mathbb{W}$ é injetiva e sobrejetiva, então T é chamada **isomorfismo**. Neste caso, dizemos que os espaços vetoriais \mathbb{V} e \mathbb{W} são **isomorfos**. Se existe um isomorfismo que leva o espaço vetorial \mathbb{V} de dimensão finita em \mathbb{W} , então, pelo **Teorema 3.8 da Dimensão do Núcleo e da Imagem**, $\dim(\mathbb{V}) = \dim(\mathbb{W})$. A recíproca também é verdadeira como mostra o próximo resultado.

Teorema 3.10. *Sejam \mathbb{V} e \mathbb{W} espaços vetoriais de dimensão finita. \mathbb{V} é isomorfo a \mathbb{W} se, e somente se,*

$$\dim(\mathbb{V}) = \dim(\mathbb{W}).$$

Demonstração. Depois da observação feita acima, basta provarmos que se $\dim(\mathbb{V}) = \dim(\mathbb{W})$, então \mathbb{V} e \mathbb{W} são isomorfos. Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_n\}$ bases de \mathbb{V} e \mathbb{W} , respectivamente. Pelo **Teorema 3.1 na página 213**, existe uma transformação linear $T : \mathbb{V} \rightarrow \mathbb{W}$, tal que $T(V_i) = W_i$, para $i = 1, \dots, n$. Pela **Proposição 3.4 na página 223**, a imagem de T é gerada pelos vetores $T(V_1), \dots, T(V_n)$ e assim

$$\mathcal{I}(T) = [T(V_1), \dots, T(V_n)] = [W_1, \dots, W_n] = \mathbb{W}.$$

Ou seja, T é sobrejetiva. Pelo **Teorema 3.8 da Dimensão do Núcleo e da Imagem**, T é também injetiva e é portanto um isomorfismo. ■

Deste teorema segue-se que todos os espaços vetoriais de dimensão n são isomorfos ao \mathbb{R}^n . Segue também deste teorema que para termos uma transformação linear entre dois espaços vetoriais de dimensão finita, \mathbb{V} e \mathbb{W} , injetiva e não sobrejetiva ou sobrejetiva e não injetiva temos que ter $\dim(\mathbb{V}) \neq \dim(\mathbb{W})$.

Corolário 3.11. *Seja \mathbb{V} um espaço vetorial. \mathbb{V} é isomorfo ao \mathbb{R}^n se, e somente se, $\dim(\mathbb{V}) = n$.*

Exemplo 3.29. O espaço \mathcal{P}_n é isomorfo a \mathbb{R}^{n+1} e o espaço \mathcal{M}_{mn} é isomorfo a \mathbb{R}^{nm} .

Exercícios Numéricos (respostas na página 441)

- 3.2.1. 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 ?
- 3.2.2. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ uma transformação linear dada por $T(x, y, z) = (z, x - y, -z)$.
- Encontre uma base para o núcleo de T .
 - Encontre uma base para a imagem de T .
 - Descreva geometricamente o núcleo e a imagem de T .
- 3.2.3. Discuta como o posto de A varia com t .
- $A = \begin{bmatrix} 1 & 1 & t \\ 1 & t & 1 \\ t & 1 & 1 \end{bmatrix}$
 - $A = \begin{bmatrix} t & 3 & -1 \\ 3 & 6 & -2 \\ -1 & -3 & -t \end{bmatrix}$
- 3.2.4. Seja $D : \mathcal{P}_3 \rightarrow \mathcal{P}_3$ o operador de derivação $D(p) = p'$. Determine uma base para o núcleo de D e para a sua imagem.
- 3.2.5. Seja $T : \mathbb{V} \rightarrow \mathbb{R}^5$ uma transformação linear.
- Se T é sobrejetiva e $\dim(\mathcal{N}(T)) = 2$, qual a dimensão de \mathbb{V} ?
 - Se T é sobrejetiva e injetiva, qual a dimensão de \mathbb{V} ?
- 3.2.6. Dê exemplos de transformações lineares $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ tais que
- $\mathcal{N}(T) = \{(x, y, z) \in \mathbb{R}^3 \mid z = -x\}$,
 - $\mathcal{I}(T) = \{(x, y, z) \in \mathbb{R}^3 \mid x = y\}$.
- 3.2.7. Dê um exemplo de uma transformação linear $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tal que $\mathcal{N}(T) = \mathcal{I}(T)$.
- 3.2.8. Defina $T : \mathcal{P}_n \rightarrow \mathcal{P}_n$ por $T(p) = p + p'$. Mostre que T é um isomorfismo.

Exercícios Teóricos

- 3.2.9. Sejam $X = [x_1 \dots x_m]^t$ e $Y = [y_1 \dots y_n]^t$ matrizes $m \times 1$ e $n \times 1$, respectivamente. Seja $A = XY^t$. Mostre que se $X \neq \mathbf{0}$, então $\{X\}$ é uma base para o espaço coluna de A . Qual o posto de A ?
- 3.2.10. Mostre que se A é uma matriz, $m \times n$, de posto igual a 1, então existem matrizes $X = [x_1 \dots x_m]^t$ e $Y = [y_1 \dots y_n]^t$, $m \times 1$ e $n \times 1$, respectivamente, tais que $A = XY^t$. (Sugestão: Tome X tal que $\{X\}$ é uma base para o espaço coluna de A .)
- 3.2.11. Seja A uma matriz $n \times n$ de posto 1. Mostre que
- (a) $A^2 = (\text{tr}(A))A$.
 - (b) $A^n = (\text{tr}(A))^{n-1}A$, para $n = 2, 3, \dots$
- (Sugestão: use o exercício anterior)
- 3.2.12. 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.
- 3.2.13. 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) $\text{posto}(C) \leq \min(\text{posto}(A), \text{posto}(B))$.
 - (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.

- 3.2.14.** 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. (Sugestão: Se $A = [A_1 \dots A_n]$, então $PA = [PA_1 \dots PA_n]$. Mostre que A_1, \dots, A_n são L.I. se, e somente se, PA_1, \dots, PA_n são L.I.)
- 3.2.15.** Seja A uma matriz $m \times n$. Mostre que
- 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.
 - O posto de A é igual ao maior inteiro positivo r tal que alguma submatriz $r \times r$ de A possui determinante não nulo.
- 3.2.16.** Determine o núcleo e a imagem do operador linear definido no [Exercício 3.1.9 na página 219](#).
- 3.2.17.** Considere o plano $\pi : ax + by + cz = 0$.
- Determine o núcleo e a imagem da projeção ortogonal no plano π , P_π . Responda se P_π é sobrejetiva e se é injetiva.
 - 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.
- 3.2.18.** Considere a reta $r : (x, y, z) = t(a, b, c)$.
- Determine o núcleo e a imagem da projeção ortogonal na reta r , P_r . Responda se P_r é sobrejetiva e se é injetiva.
 - 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.
- 3.2.19.** Seja $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ um funcional linear.
- Mostre que existem escalares a, b, c tais que $f(x, y, z) = ax + by + cz$.
 - Descreva geometricamente todas as possibilidades para o núcleo de f .

- 3.2.20. Prove o Teorema da dimensão do Núcleo e da Imagem nos casos em que $\dim(\mathcal{N}(T)) = 0$ e $\dim(\mathcal{N}(T)) = \dim(\mathbb{V})$.
- 3.2.21. Sejam $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear e $\mathcal{B} = \{V_1, \dots, V_n\}$ um conjunto de vetores de \mathbb{V} . Mostre que se $\mathcal{C} = \{T(V_1), \dots, T(V_n)\}$ é L.I., então \mathcal{B} também o é.
- 3.2.22. Seja \mathbb{V} um espaço vetorial de dimensão finita. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Mostre que T é injetiva se, e somente se, $\dim(\mathcal{I}(T)) = \dim(\mathbb{V})$.
- 3.2.23. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ 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.
- 3.2.24. Sejam $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Suponha que $\dim(\mathbb{V}) = \dim(\mathbb{W})$. Mostre que T é injetiva se, e somente se, a imagem por T de uma base de \mathbb{V} é uma base de \mathbb{W} .
- 3.2.25. Seja λ um escalar. Considere o operador diferencial $D - \lambda I : \mathcal{C}^\infty(\mathbb{R}) \rightarrow \mathcal{C}^\infty(\mathbb{R})$, dado por $(D - \lambda I)(f) = f' - \lambda f$, para todo $f \in \mathcal{C}^\infty(\mathbb{R})$.
- (a) Mostre que o núcleo de $D - \lambda I$ tem dimensão 1. (Sugestão: Mostre que $\{e^{\lambda t}\}$ é uma base para o núcleo de $D - \lambda I$ que é a solução da equação diferencial $y' - \lambda y = 0$. Veja [Exemplo 1.19 na página 22](#))
- (b) Mostre que $D - \lambda I$ é sobrejetivo. (Sugestão: Mostre que a equação diferencial $y' - \lambda y = f$ tem solução para todo $f \in \mathcal{C}^\infty(\mathbb{R})$. Sejam $w(t) = f(t)e^{-\lambda t}$ e $W(t) = \int_0^t w(s)ds$, Mostre que $y(t) = W(t)e^{\lambda t}$ é solução.)
- 3.2.26. Mostre que se $V \in \mathbb{V}$ é tal que $f(V) = 0$, para todo funcional linear de \mathbb{V} , então $V = \vec{0}$. (Sugestão: se $V \neq \vec{0}$, complete a uma base de \mathbb{V} defina um funcional não nulo que é diferente de zero em V .)
- 3.2.27. Definimos o **espaço bidual** de \mathbb{V} por $\mathbb{V}^{**} = \mathcal{L}(\mathbb{V}^*; \mathbb{R})$, onde $\mathbb{V}^* = \mathcal{L}(\mathbb{V}; \mathbb{R})$ é o espaço dual de \mathbb{V} . Seja \mathbb{V} um espaço de dimensão finita. Considere a função $\zeta : \mathbb{V} \rightarrow \mathbb{V}^{**}$ que associa a cada vetor $V \in \mathbb{V}$ o elemento $\zeta(V) = V^{**} \in \mathbb{V}^{**}$, tal que $V^{**}(f) = f(V)$, para todo $f \in \mathbb{V}^*$. Mostre que ζ é um isomorfismo. (Sugestão: mostre que ζ é linear e que $\mathcal{N}(\zeta) = \{\vec{0}\}$.)
- 3.2.28. Seja \mathbb{V} um espaço vetorial de dimensão finita. Dada uma base $\mathcal{F} = \{f_1, \dots, f_n\}$ do espaço dual \mathbb{V}^* ,

mostre que existe uma base $\mathcal{B} = \{V_1, \dots, V_n\}$ de \mathbb{V} tal que $\mathcal{B}^* = \mathcal{F}$, onde $\mathcal{B}^* = \{g_1, \dots, g_n\}$ é a base dual de \mathcal{B} , ou seja, definida por $g_i(V_j) = \delta_{ij}$, onde δ_{ij} é o delta de Kronecker. (Sugestão: use o isomorfismo do exercício anterior.)

3.2.29. Seja f um funcional linear não nulo de um espaço vetorial \mathbb{V} .

(a) Mostre que existe um vetor $U \in \mathbb{V}$ tal que $f(U) = 1$.

(b) Mostre que $\mathbb{V} = \mathcal{N}(f) \oplus [U]$. (Sugestão: resolva a equação $V = W + xU$, com $W \in \mathcal{N}(f)$.)

3.3 Composição de Transformações Lineares

Sejam $T : \mathbb{V} \rightarrow \mathbb{U}$ e $S : \mathbb{U} \rightarrow \mathbb{W}$ transformações lineares. A **composição de S com T** , denotada por ST é a função de \mathbb{V} em \mathbb{W} definida por

$$(ST)(X) = S(T(X)), \quad \text{para todo } X \in \mathbb{V}.$$

Proposição 3.12. *Se $T : \mathbb{V} \rightarrow \mathbb{U}$ e $S : \mathbb{U} \rightarrow \mathbb{W}$ são transformações lineares, então a composição $ST : \mathbb{V} \rightarrow \mathbb{W}$ é uma transformação linear.*

Demonstração. Sejam $X_1, X_2 \in \mathbb{V}$ e α, β escalares.

$$\begin{aligned}(ST)(\alpha X_1 + \beta X_2) &= S(T(\alpha X_1 + \beta X_2)) = S(\alpha T(X_1) + \beta T(X_2)) \\ &= \alpha S(T(X_1)) + \beta S(T(X_2)) = \alpha(ST)(X_1) + \beta(ST)(X_2)\end{aligned}$$

■

Podemos, agora, provar outras propriedades da álgebra das transformações lineares.

Proposição 3.13. *Sejam S, T e U transformações lineares com domínios e contra-domínios apropriados. Então,*

(a) $S(TU) = (ST)U$;

$$(b) S(T + U) = ST + SU \text{ e } (S + T)U = SU + TU$$

$$(c) \alpha(TU) = (\alpha T)U = T(\alpha U);$$

(d) $IT = TI = T$, onde I é a transformação identidade.

Demonstração. Seja X pertencente ao domínio de U .

$$(a) S(TU)(X) = S((TU)(X)) = S(T(U(X))) = ST((U(X)));$$

(b) Pela linearidade de S temos que

$$\begin{aligned} S(T + U)(X) &= S((T + U)(X)) = S(T(X) + U(X)) = S(T(X)) + S(U(X)) \\ &= ST(X) + SU(X) = (ST + SU)(X); \end{aligned}$$

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

$$(c) \alpha(ST)(X) = \alpha S(T(X)) = (\alpha S)(T(X)) = (\alpha S)T(X) \text{ e} \\ \alpha(ST)(X) = \alpha S(T(X)) = S(\alpha T(X)) = S(\alpha T)(X).$$

(d) A demonstração deste item é simples e deixamos como exercício.



3.3.1 Matriz de uma Transformação Linear

Definição 3.6. Seja $\mathcal{B} = \{V_1, \dots, V_n\}$ uma base de um espaço vetorial \mathbb{V} . Todo vetor $V \in \mathbb{V}$ se escreve de maneira única como combinação linear de V_1, \dots, V_n (**Teorema 1.9 na página 70**), ou seja, existem escalares $\alpha_1, \dots, \alpha_n$ tais que $V = \alpha_1 V_1 + \dots + \alpha_n V_n$. Definimos o **vetor de coordenadas de V em relação à base (ordenada) $\mathcal{B} = \{V_1, \dots, V_n\}$** , 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} . Assim, $[V_1]_{\mathcal{B}} = E_1, [V_2]_{\mathcal{B}} = E_2, \dots, [V_n]_{\mathcal{B}} = E_n$, em que E_1, \dots, E_n são os vetores da base canônica do \mathbb{R}^n . Pois, $V_i = 0V_1 + \dots + 1V_i + \dots + 0V_n$, para $i = 1, \dots, n$.

Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_m\}$ bases dos espaços vetoriais \mathbb{V} e \mathbb{W} , respectivamente. Seja $T: \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Seja

$$[V]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad [T(V_1)]_{\mathcal{C}} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad [T(V_2)]_{\mathcal{C}} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \dots, \quad [T(V_n)]_{\mathcal{C}} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Então,

$$\begin{aligned} T(V) &= T(x_1V_1 + \dots + x_nV_n) = x_1T(V_1) + \dots + x_nT(V_n) \\ &= x_1(a_{11}W_1 + \dots + a_{m1}W_m) + \dots + x_n(a_{1n}W_1 + \dots + a_{mn}W_m) \\ &= (x_1a_{11} + \dots + x_na_{1n})W_1 + \dots + (x_1a_{m1} + \dots + x_na_{mn})W_m. \end{aligned}$$

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,

$$[T(V)]_{\mathcal{C}} = \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A[V]_{\mathcal{B}},$$

em que $A = [[T(V_1)]_C \ \dots \ [T(V_n)]_C]$. 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,

$$[T]_{\mathcal{B}}^{\mathcal{C}} = [[T(V_1)]_C \ \dots \ [T(V_n)]_C].$$

Isto prova o seguinte resultado, que é uma generalização da **Proposição 3.2** na página 215.

Teorema 3.14. *Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_m\}$ bases dos espaços vetoriais \mathbb{V} e \mathbb{W} , respectivamente. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Então, a matriz $m \times n$*

$$[T]_{\mathcal{B}}^{\mathcal{C}} = [[T(V_1)]_C \ \dots \ [T(V_n)]_C],$$

é tal que

$$[T(V)]_C = [T]_{\mathcal{B}}^{\mathcal{C}} [V]_{\mathcal{B}}, \quad \text{para todo vetor } V \in \mathbb{V}.$$

Aqui $[V]_{\mathcal{B}}$ é o vetor de coordenadas de V em relação à base \mathcal{B} , $[T(V)]_C$ é o vetor de coordenadas de $T(V)$ em relação à base \mathcal{C} e a matriz $[T]_{\mathcal{B}}^{\mathcal{C}} = [[T(V_1)]_C \ \dots \ [T(V_n)]_C]$ é matriz da transformação linear T em relação às bases \mathcal{B} e \mathcal{C} .

Exemplo 3.30. Seja $D : \mathcal{P}_3 \rightarrow \mathcal{P}_2$ a transformação que associa a cada polinômio $p(t)$ a sua derivada $p'(t)$. Sejam $\mathcal{B} = \{1, t, t^2, t^3\}$ e $\mathcal{C} = \{1, t, t^2\}$. Vamos determinar a matriz da transformação linear D em relação a estas bases. Para isto, vamos escrever

o resultado da aplicação de D em cada elemento de \mathcal{B} como combinação linear dos elementos de \mathcal{C} .

$$\begin{aligned} D(1) &= 0 = 0(1) + 0(t) + 0(t^2) \\ D(t) &= 1 = 1(1) + 0(t) + 0(t^2) \\ D(t^2) &= 2t = 0(1) + 2(t) + 0(t^2) \\ D(t^3) &= 3t^2 = 0(1) + 0(t) + 3(t^2) \end{aligned}$$

Assim,

$$[D]_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} [D(1)]_{\mathcal{C}} & [D(t)]_{\mathcal{C}} & [D(t^2)]_{\mathcal{C}} & [D(t^3)]_{\mathcal{C}} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Vamos, agora, usar a matriz obtida para calcular $D(p)$ para um polinômio $p(t) = a_0 + a_1x + a_2x^2 + a_3x^3$. Claramente $[p]_{\mathcal{B}} = [a_0 \ a_1 \ a_2 \ a_3]^t$. Assim,

$$[D(p)]_{\mathcal{C}} = [D]_{\mathcal{B}}^{\mathcal{C}}[p]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ 3a_3 \end{bmatrix}.$$

Portanto, $D(p) = a_1(1) + 2a_2(x) + 3a_3(x^2)$, como já sabíamos.

Quando a transformação linear é a transformação identidade $I_{\mathbb{V}} : \mathbb{V} \rightarrow \mathbb{V}$, definida por $I_{\mathbb{V}}(X) = X$, para todo $X \in \mathbb{V}$, então aplicando o resultado anterior ([Teorema 3.14](#)) a esta transformação, obtemos uma relação entre os vetores de coordenadas de um vetor X em relação à duas bases.

$$[X]_{\mathcal{C}} = [I_{\mathbb{V}}(X)]_{\mathcal{C}} = [I_{\mathbb{V}}]_{\mathcal{B}}^{\mathcal{C}}[X]_{\mathcal{B}} = P[X]_{\mathcal{B}},$$

em que $P = [[V_1]_{\mathcal{C}} \ \dots \ [V_n]_{\mathcal{C}}]$ é chamada **matriz mudança de base de \mathcal{B} para \mathcal{C}** .

Corolário 3.15. Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_n\}$ bases de um mesmo espaço vetorial \mathbb{V} . Então, Para todo $X \in \mathbb{V}$,

$$[X]_{\mathcal{C}} = [I_{\mathbb{V}}]_{\mathcal{B}}^{\mathcal{C}} [X]_{\mathcal{B}},$$

em que $P = [I_{\mathbb{V}}]_{\mathcal{B}}^{\mathcal{C}} = [[V_1]_{\mathcal{C}} \dots [V_n]_{\mathcal{C}}]$ é a matriz da transformação linear identidade $I_{\mathbb{V}} : \mathbb{V} \rightarrow \mathbb{V}$ em relação às bases \mathcal{C} e \mathcal{B} .

Exemplo 3.31. 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 [24]), 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

$$[I_{\mathbb{R}^2}]_{\mathcal{B}}^{\mathcal{C}} = [[E_1]_{\mathcal{C}} \ [E_2]_{\mathcal{C}}] = [A^{-1}E_1 \ A^{-1}E_2] = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}.$$

Portanto

$$[(x, y)]_{\mathcal{C}} = [I_{\mathbb{R}^2}]_{\mathcal{B}}^{\mathcal{C}} [(x, y)]_{\mathcal{B}} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (x+y)/2 \\ (x-y)/2 \end{bmatrix}.$$

Teorema 3.16. *Sejam $T : \mathbb{V} \rightarrow \mathbb{U}$ e $S : \mathbb{U} \rightarrow \mathbb{W}$ transformações lineares. Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$, $\mathcal{C} = \{U_1, \dots, U_p\}$ e $\mathcal{D} = \{W_1, \dots, W_m\}$ bases de \mathbb{V} , \mathbb{U} e \mathbb{W} 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, \dots, W_m são L.I., então $c_{ij} = \sum_{k=1}^p a_{ik}b_{kj}$, ou seja, $C = AB$, como queríamos provar. ■

Exemplo 3.32. 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

$$\begin{aligned} U_1 &= U = (a, b, c) \\ U_2 &= \left(-\frac{b}{\sqrt{a^2 + b^2}}, \frac{a}{\sqrt{a^2 + b^2}}, 0\right) \\ U_3 &= \left(-\frac{ac}{\sqrt{a^2 + b^2}}, -\frac{bc}{\sqrt{a^2 + b^2}}, \sqrt{a^2 + b^2}\right) \end{aligned}$$

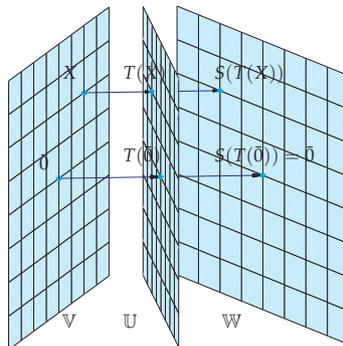


Figura 3.14: Composição das Transformações Lineares $T : V \rightarrow U$ e $S : U \rightarrow W$

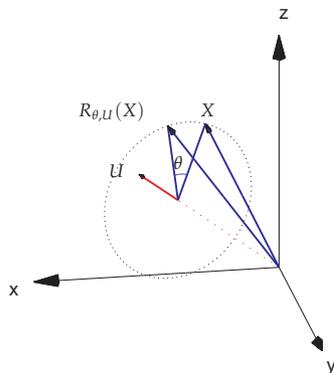


Figura 3.15: Rotação de um ângulo θ em torno de um eixo determinado pelo vetor U

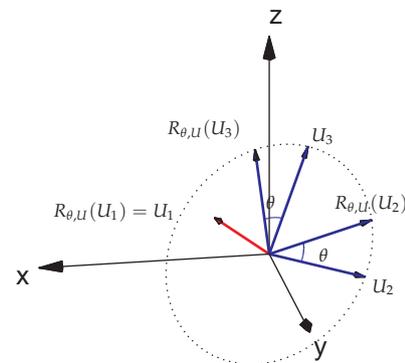


Figura 3.16: $R_{\theta, U}(U_1)$, $R_{\theta, U}(U_2)$ e $R_{\theta, U}(U_3)$

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

$$R_\theta(U_1) = U_1 = (a, b, c)$$

$$R_\theta(U_2) = \cos \theta U_2 + \sin \theta U_3 = \left(\frac{-b \cos \theta - ac \sin \theta}{\sqrt{a^2 + b^2}}, \frac{a \cos \theta - bc \sin \theta}{\sqrt{a^2 + b^2}}, \sqrt{a^2 + b^2} \sin \theta \right)$$

$$R_\theta(U_3) = -\sin \theta U_2 + \cos \theta U_3 = \left(\frac{b \sin \theta - ac \cos \theta}{\sqrt{a^2 + b^2}}, \frac{-a \sin \theta - bc \cos \theta}{\sqrt{a^2 + b^2}}, \sqrt{a^2 + b^2} \cos \theta \right).$$

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 $R_{\theta, U} = R_{\theta, U} I_{\mathbb{R}^3}$ e assim

$$[R_{\theta, U}]_{\mathcal{B}}^{\mathcal{B}} = [R_{\theta, U} I_{\mathbb{R}^3}]_{\mathcal{B}}^{\mathcal{B}} = [R_{\theta, U}]_{\mathcal{C}}^{\mathcal{B}} [I_{\mathbb{R}^3}]_{\mathcal{B}}^{\mathcal{C}}$$

Agora,

$$[R_{\theta, U}]_{\mathcal{C}}^{\mathcal{B}} = [[R_{\theta, U}(U_1)]_{\mathcal{B}} \ [R_{\theta, U}(U_2)]_{\mathcal{B}} \ [R_{\theta, U}(U_3)]_{\mathcal{B}}] = [R_{\theta, U}(U_1) \ R_{\theta, U}(U_2) \ R_{\theta, U}(U_3)],$$

e

$$[I_{\mathbb{R}^3}]_{\mathcal{B}}^{\mathcal{C}} = [[E_1]_{\mathcal{C}} \ [E_2]_{\mathcal{C}} \ [E_3]_{\mathcal{C}}] = \begin{bmatrix} \langle E_1, U_1 \rangle & \langle E_2, U_1 \rangle & \langle E_3, U_1 \rangle \\ \langle E_1, U_2 \rangle & \langle E_2, U_2 \rangle & \langle E_3, U_2 \rangle \\ \langle E_1, U_3 \rangle & \langle E_2, U_3 \rangle & \langle E_3, U_3 \rangle \end{bmatrix} = [U_1 \ U_2 \ U_3]^t,$$

pois como as bases \mathcal{B} e \mathcal{C} são ortonormais, pela [Proposição 2.7 na página 122](#),

$$E_j = \sum_{i=1}^3 \langle E_j, U_i \rangle U_i \text{ e } U_j = \sum_{i=1}^3 \langle U_j, E_i \rangle E_i = (\langle U_j, E_1 \rangle, \langle U_j, E_2 \rangle, \langle U_j, E_3 \rangle).$$

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,

$$[R_{\theta,U}(U_1) \ R_{\theta,U}(U_2) \ R_{\theta,U}(U_3)] [U_1 \ U_2 \ U_3]^t =$$

$$\begin{bmatrix} a & \frac{-b \cos \theta - ac \sin \theta}{\sqrt{a^2 + b^2}} & \frac{b \sin \theta - ac \cos \theta}{\sqrt{a^2 + b^2}} \\ b & \frac{a \cos \theta - bc \sin \theta}{\sqrt{a^2 + b^2}} & \frac{-a \sin \theta - bc \cos \theta}{\sqrt{a^2 + b^2}} \\ c & \sqrt{a^2 + b^2} \sin \theta & \sqrt{a^2 + b^2} \cos \theta \end{bmatrix} \begin{bmatrix} a & b & c \\ -\frac{b}{\sqrt{a^2 + b^2}} & \frac{a}{\sqrt{a^2 + b^2}} & 0 \\ -\frac{ac}{\sqrt{a^2 + b^2}} & -\frac{bc}{\sqrt{a^2 + b^2}} & \sqrt{a^2 + b^2} \end{bmatrix} =$$

$$\begin{bmatrix} a^2(1 - \cos \theta) + \cos \theta & ab(1 - \cos \theta) - c \sin \theta & ac(1 - \cos \theta) + b \sin \theta \\ ab(1 - \cos \theta) + c \sin \theta & b^2(1 - \cos \theta) + \cos \theta & bc(1 - \cos \theta) - a \sin \theta \\ ac(1 - \cos \theta) - b \sin \theta & bc(1 - \cos \theta) + a \sin \theta & c^2(1 - \cos \theta) + \cos \theta \end{bmatrix}$$

que é a matriz de $R_{\theta,U}$ em relação à base canônica. Finalmente,

$$R_{\theta,U} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a^2(1 - \cos \theta) + \cos \theta & ab(1 - \cos \theta) - c \sin \theta & ac(1 - \cos \theta) + b \sin \theta \\ ab(1 - \cos \theta) + c \sin \theta & b^2(1 - \cos \theta) + \cos \theta & bc(1 - \cos \theta) - a \sin \theta \\ ac(1 - \cos \theta) - b \sin \theta & bc(1 - \cos \theta) + a \sin \theta & c^2(1 - \cos \theta) + \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

3.3.2 Invertibilidade

Dizemos que uma transformação linear $T : \mathbb{V} \rightarrow \mathbb{W}$ é **invertível** se, existe uma função $U : \mathbb{W} \rightarrow \mathbb{V}$ tal que $TU = I_{\mathbb{W}}$ e $UT = I_{\mathbb{V}}$. A função U é única (verifique!) e denotada por T^{-1} .

Proposição 3.17. *Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear invertível. Então, $T^{-1} : \mathbb{W} \rightarrow \mathbb{V}$ é também uma transformação linear.*

Demonstração. Sejam $Y_1, Y_2 \in \mathbb{W}$ 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 $f : A \rightarrow B$ é invertível se, e somente se, é injetiva e sobrejetiva.

Teorema 3.18. Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_m\}$ bases dos espaços vetoriais \mathbb{V} e \mathbb{W} , respectivamente. Uma transformação linear $T : \mathbb{V} \rightarrow \mathbb{W}$ é invertível se, e somente se, $[T]_{\mathcal{B}}^{\mathcal{C}}$ é invertível ($\det([T]_{\mathcal{B}}^{\mathcal{C}}) \neq 0$). 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 3.8 na página 231](#), que $n = \dim(\mathbb{V}) = \dim(\mathbb{W}) = m$. Além disso, existe uma transformação linear, T^{-1} , tal que $TT^{-1} = I_{\mathbb{W}}$ e $T^{-1}T = I_{\mathbb{V}}$. Assim,

$$I_n = [I_{\mathbb{V}}]_{\mathcal{B}}^{\mathcal{B}} = [T^{-1}T]_{\mathcal{B}}^{\mathcal{B}} = [T^{-1}]_{\mathcal{C}}^{\mathcal{B}} [T]_{\mathcal{B}}^{\mathcal{C}}.$$

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 $\dim(\mathbb{V}) = n = m = \dim(\mathbb{W})$. 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 3.7 na página 230**) e como $\dim(\mathbb{V}) = \dim(\mathbb{W})$, então pelo **Corolário 3.9 na página 233** segue-se que T é também sobrejetiva e portanto invertível. ■

Quando a transformação linear é a transformação identidade $I_{\mathbb{V}} : \mathbb{V} \rightarrow \mathbb{V}$, definida por $I_{\mathbb{V}}(X) = X$, para todo $X \in \mathbb{V}$, então aplicando o resultado anterior (**Teorema 3.18**) a esta transformação, obtemos o seguinte resultado.

Corolário 3.19. *Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_n\}$ bases de um mesmo espaço vetorial \mathbb{V} . A matriz mudança de base $P = [I_{\mathbb{V}}]_{\mathcal{B}}^{\mathcal{C}}$ é invertível e*

$$P^{-1} = ([I_{\mathbb{V}}]_{\mathcal{B}}^{\mathcal{C}})^{-1} = [I_{\mathbb{V}}]_{\mathcal{C}}^{\mathcal{B}}$$

Exemplo 3.33. Seja $T : \mathcal{P}_2 \rightarrow \mathcal{P}_2$ definida por $T(p) = p + p' + p''$, para todo $p \in \mathcal{P}_2$, onde p' e p'' denotam a primeira e a segunda derivada de p , respectivamente. Vamos verificar se T é invertível. Seja $\mathcal{B} = \{1, x, x^2\}$. 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} .

$$\begin{aligned} T(1) &= 1 &= 1(1) + 0(x) + 0(x^2) \\ T(x) &= x + 1 &= 1(1) + 1(x) + 0(x^2) \\ T(x^2) &= x^2 + 2x + 2 &= 2(1) + 2(x) + 1(x^2) \end{aligned}$$

Assim, a matriz de T em relação a \mathcal{B} é

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} [T(1)]_{\mathcal{B}} & [T(x)]_{\mathcal{B}} & [T(x^2)]_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Esta matriz é invertível e assim pelo [Teorema 3.18](#) 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} . Seja $p(x) = a_0 + a_1x + a_2x^2$ um polinômio qualquer de \mathcal{P}_2 . Então, $[p]_{\mathcal{B}} = [a_0 \ a_1 \ a_2]^t$ e pelo [Teorema 3.14 na página 245](#), temos que

$$[T^{-1}(p)]_{\mathcal{B}} = [T^{-1}]_{\mathcal{B}}^{\mathcal{B}} [p]_{\mathcal{B}} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_0 - a_1 \\ a_1 - 2a_2 \\ a_2 \end{bmatrix}.$$

Portanto, $T^{-1}(p) = (a_0 - a_1) + (a_1 - 2a_2)x + a_2x^2$.

3.3.3 Semelhança

Corolário 3.20. *Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_n\}$ bases de um espaço vetorial \mathbb{V} . Se $T : \mathbb{V} \rightarrow \mathbb{V}$ é uma transformação linear, então*

$$[T]_{\mathcal{C}}^{\mathcal{C}} = P^{-1} [T]_{\mathcal{B}}^{\mathcal{B}} P.$$

Demonstração. Pelo Teorema 3.16 na página 248 temos que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = [I_{\mathbb{V}}]_{\mathcal{B}}^{\mathcal{C}} [T]_{\mathcal{B}}^{\mathcal{B}} [I_{\mathbb{V}}]_{\mathcal{C}}^{\mathcal{B}}.$$

Mas pelo Corolário 3.19 na página 255 a matriz $P = [I_{\mathbb{V}}]_{\mathcal{C}}^{\mathcal{B}}$ é invertível e $P^{-1} = [I_{\mathbb{V}}]_{\mathcal{B}}^{\mathcal{C}}$. De onde segue-se o resultado. ■

Uma transformação linear de um espaço vetorial \mathbb{V} 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 3.20 pode ser estabelecido da seguinte forma:

Se $T : \mathbb{V} \rightarrow \mathbb{V}$ é uma transformação linear, $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_n\}$ são bases de \mathbb{V} , 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 $\text{tr}(AB) = \text{tr}(BA)$ (ver por exemplo [24]), então

$$\text{tr}(P^{-1}AP) = \text{tr}(A).$$

Assim, em virtude do [Corolário 3.20](#), se \mathbb{V} é um espaço vetorial de dimensão finita, podemos definir o **traço de um operador linear** $T : \mathbb{V} \rightarrow \mathbb{V}$ como sendo

$$\text{tr}(T) = \text{tr}([T]_{\mathcal{B}}^{\mathcal{B}}),$$

onde \mathcal{B} é uma base de \mathbb{V} .

De forma análoga, como $\det(AB) = \det(A)\det(B) = \det(BA)$ (ver por exemplo [\[24\]](#)), então

$$\det(P^{-1}AP) = \det(A).$$

Assim, em virtude do [Corolário 3.20](#), se \mathbb{V} é um espaço vetorial de dimensão finita, podemos definir o **determinante de um operador linear** $T : \mathbb{V} \rightarrow \mathbb{V}$ como sendo

$$\det(T) = \det([T]_{\mathcal{B}}^{\mathcal{B}}),$$

onde \mathcal{B} é uma base de \mathbb{V} qualquer. Agora podemos dizer que um operador é invertível se, e somente se, o seu determinante é diferente de zero ([Teorema 3.18 na página 254](#)).

Exemplo 3.34. Vamos obter uma expressão para a reflexão na reta $r : y = 2x$, $R_r : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, usando o [Corolário 3.20](#). 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, $\mathcal{C} = \{V_1 = (1, 2), V_2 = (-2, 1)\}$.

$$\begin{aligned} R_r(V_1) &= R_r(1, 2) = (1, 2) = 1V_1 + 0V_2 \\ R_r(V_2) &= R_r(-2, 1) = (2, -1) = 0V_1 - 1V_2. \end{aligned}$$

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 3.20](#), 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 3.20](#), cuja demonstração é inteiramente análoga e deixamos como exercício para o leitor.

Corolário 3.21. *Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{B}' = \{V'_1, \dots, V'_n\}$ bases de \mathbb{V} e $\mathcal{C} = \{W_1, \dots, W_m\}$ e $\mathcal{C}' = \{W'_1, \dots, W'_m\}$ bases de \mathbb{W} . 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} .

3.3.4 Aplicação: Equações Diferenciais Lineares

Vamos fazer uma aplicação ao estudo das equações diferenciais lineares homogêneas com coeficientes constantes, ou seja, equações da forma

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0,$$

onde $a_0, \dots, a_{n-1} \in \mathbb{R}$.

Proposição 3.22. *Sejam $S, T : \mathbb{V} \rightarrow \mathbb{V}$ operadores lineares, tais que $\mathcal{N}(S)$ e $\mathcal{N}(T)$ têm dimensão finita.*

(a) *Então, $\dim(\mathcal{N}(ST)) \leq \dim(\mathcal{N}(S)) + \dim(\mathcal{N}(T))$.*

(b) *Se T é sobrejetivo, então $\dim(\mathcal{N}(ST)) = \dim(\mathcal{N}(S)) + \dim(\mathcal{N}(T))$.*

Demonstração. Seja $\{V_1, \dots, V_k\}$ base do núcleo de T . Seja $\{W_1, \dots, W_l\}$ base de $\mathcal{N}(S) \cap \mathcal{I}(T)$. Estenda-a a uma base do núcleo de S , $\{W_1, \dots, W_r\}$, $r \geq l$. Sejam U_1, \dots, U_l tais que $T(U_1) = W_1, \dots, T(U_l) = W_l$.

(a) Vamos mostrar que $V_1, \dots, V_k, U_1, \dots, U_l$ geram o núcleo de ST .

Seja $V \in \mathcal{N}(ST)$. Então, $T(V) \in \mathcal{N}(S) \cap \mathcal{I}(T)$. Assim, existem escalares $\alpha_1, \dots, \alpha_l$ tais que

$$T(V) = \alpha_1 W_1 + \dots + \alpha_l W_l.$$

Como $W_1 = T(U_1), \dots, W_l = T(U_l)$, então

$$T(V - \alpha_1 U_1 - \dots - \alpha_l U_l) = \vec{0}.$$

Logo, $V - \alpha_1 U_1 - \dots - \alpha_l U_l \in \mathcal{N}(T)$ o que implica que existem escalares β_1, \dots, β_k tais que

$$V - \alpha_1 U_1 - \dots - \alpha_l U_l = \beta_1 V_1 + \dots + \beta_k V_k.$$

Portanto, $V_1, \dots, V_k, U_1, \dots, U_l$ geram o núcleo de ST . De onde segue-se o resultado.

- (b) Se T é sobrejetivo, então $\mathcal{N}(S) \cap \mathcal{I}(T) = \mathcal{N}(S)$ e assim $l = r$. Vamos mostrar que $V_1, \dots, V_k, U_1, \dots, U_l$ são L.I. Considere a combinação linear nula

$$x_1 V_1 + \dots + x_k V_k + y_1 U_1 + \dots + y_l U_l = \vec{0}. \quad (3.3)$$

Como $V_1, \dots, V_k \in \mathcal{N}(T)$, então aplicando-se T ficamos com

$$y_1 W_1 + \dots + y_l W_l = \vec{0}.$$

Como W_1, \dots, W_l são L.I., então $y_1 = \dots = y_l = 0$. Substituindo-se os valores de y_1, \dots, y_l em (3.3) e usando o fato de que V_1, \dots, V_k também são L.I., obtemos que $x_1 = \dots = x_k = 0$ e $V_1, \dots, V_k, U_1, \dots, U_l$ são L.I. De onde segue-se o resultado.



Corolário 3.23. *Seja $D : \mathcal{C}^\infty(\mathbb{R}) \rightarrow \mathcal{C}^\infty(\mathbb{R})$ o operador de derivação, $D(f) = f'$, para todo $f \in \mathcal{C}^\infty(\mathbb{R})$. Seja $p(D) = (D - \lambda_1 I) \dots (D - \lambda_n I)$. Então, $\dim(\mathcal{N}(p(D))) = n$.*

Demonstração. Pelo Exercício 2.25 na página 239 $\dim(\mathcal{N}(D - \lambda_n I)) = 1$. Podemos escrever

$$p(D) = q(D)(D - \lambda_n I),$$

onde $q(D)$ é de ordem $n - 1$. Vamos supor que o resultado seja verdadeiro para $n - 1$, ou seja, que $\dim(\mathcal{N}(q(D))) = n - 1$.

Como, pelo Exercício 2.25 na página 239 $(D - \lambda_n I)$ é sobrejetivo, então pela Proposição 3.22 segue-se que

$$\dim(\mathcal{N}(p(D))) = \dim(\mathcal{N}(q(D))) + \dim(\mathcal{N}(D - \lambda_n I)) = (n - 1) + 1 = n.$$

■

Corolário 3.24. O espaço solução de uma equação diferencial linear homogênea com coeficientes constantes é um subespaço de $\mathcal{C}^\infty(\mathbb{R})$ de dimensão n .

Demonstração. Seja a equação diferencial linear homogênea com coeficientes constantes dada por

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0, \quad (3.4)$$

onde $a_0, \dots, a_{n-1} \in \mathbb{R}$. Considere o polinômio auxiliar

$$p(t) = t^n + a_{n-1}t^{n-1} + \dots + a_0.$$

Seja $p(D) : \mathcal{C}^\infty \rightarrow \mathcal{C}^\infty$ o operador definido por

$$p(D)(y) = y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y.$$

O conjunto solução da equação diferencial (3.4) é igual ao núcleo do operador $p(D)$.

Vamos mostrar que $\dim(\mathcal{N}(p(D))) = n$.

Pelo Teorema Fundamental da Álgebra o polinômio $p(t) = t^n + a_{n-1}t^{n-1} + \dots + a_0$ é fatorável da forma $p(t) = (t - \lambda_1) \dots (t - \lambda_n)$, em que $\lambda_1, \dots, \lambda_n$ são raízes reais ou complexas de $p(t)$.

- (a) Se as raízes são reais, então a solução da equação diferencial (3.4) é o núcleo de $p(D) = (D - \lambda_1 I) \dots (D - \lambda_n I)$ que pelo [Corolário 3.23](#) tem dimensão n .
- (b) Se $p(t)$ tiver raízes complexas, então pelo mesmo raciocínio empregado no item anterior, o conjunto solução da equação (3.4) tem dimensão n em $C^\infty(\mathbb{R}; \mathbb{C})$. Agora, as raízes complexas aparecem aos pares, sendo uma o conjugado da outra. $e^{(\alpha+i\beta)t}$ pertence ao núcleo de $(D - (\alpha + i\beta)I)$, $e^{(\alpha-i\beta)t}$ pertence ao núcleo de $(D - (\alpha - i\beta)I)$ e $e^{\alpha t} \cos \beta t$ e $e^{\alpha t} \sin \beta t$ são funções linearmente independentes que pertencem ao núcleo de $(D - (\alpha + i\beta)I)(D - (\alpha - i\beta)I)$. Assim, o núcleo do operador formado por cada par de raízes complexas tem dimensão 2 em $C^\infty(\mathbb{R})$. Portanto, também neste caso a dimensão do espaço solução de (3.4) é igual a n .



Exercícios Numéricos (respostas na página 443)

- 3.3.1. Seja $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 $[T]_{\mathcal{C}}^{\mathcal{C}}$.
- 3.3.2. Seja $D : \mathcal{P}_1 \rightarrow \mathcal{P}_1$ definida por $D(p) = p'$, a derivada de p , para todo $p \in \mathcal{P}_1$. Seja $\mathcal{C} = \{1 + t, 1 - t\}$. Encontre $[D]_{\mathcal{C}}^{\mathcal{C}}$.
- 3.3.3. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ 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 $\{V_1, V_2, V_3\}$.
- 3.3.4. Seja $T : \mathcal{P}_2 \rightarrow \mathcal{P}_2$ a transformação linear dada por $T(p)(t) = tp'(t) + p''(t)$.
- (a) Encontre a matriz A que representa T com relação a $\mathcal{B} = \{1, t, t^2\}$.
- (b) Encontre a matriz B que representa T com relação a $\mathcal{C} = \{1, t, 1 + t^2\}$.
- (c) Encontre a matriz P tal que $B = P^{-1}AP$.
- (d) Se $p(t) = a_0 + a_1t + a_2(1 + t^2)$, calcule $T^n(p)$.
- 3.3.5. 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 $P_r : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ a projeção ortogonal na reta r . Encontre $[P_r]_{\mathcal{C}}^{\mathcal{C}}$ e $[P_r]_{\mathcal{B}}^{\mathcal{B}}$.
- (b) Seja $R_r : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ a reflexão em relação à reta r . Encontre $[R_r]_{\mathcal{C}}^{\mathcal{C}}$ e $[R_r]_{\mathcal{B}}^{\mathcal{B}}$.

- 3.3.6. 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 $R_{\pi/2, r} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ 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 $[R_{\pi/2, r}]_{\mathcal{B}}^{\mathcal{B}}$.
- 3.3.7. Para cada uma das transformações lineares T verifique se T é invertível e calcule a inversa, T^{-1} , se ela existe.
- (a) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida por $T(x, y, z) = (x + 2y + z, y + 2z, z)$.
 - (b) $T : \mathcal{P}_2 \rightarrow \mathcal{P}_2$ definida por $T(p) = p - 2p' - p''$.
 - (c) $T : \mathbb{R}^3 \rightarrow \mathcal{P}_2$ definida por $T(a, b, c) = (a + b + c) + (a + 2b + c)t + (a + 2c)t^2$.
 - (d) $T : \mathcal{P}_2 \rightarrow \mathbb{R}^3$ definida por $T(p) = (p(-1), p(0), p(1))$.

Exercícios Teóricos

- 3.3.8. Sejam \mathbb{V} e \mathbb{W} espaços vetoriais de dimensão finita. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Mostre que o posto de T é igual ao posto da matriz $[T]_{\mathcal{B}}^{\mathcal{C}}$, onde \mathcal{B} e \mathcal{C} bases de \mathbb{V} e \mathbb{W} , respectivamente. (Sugestão: mostre que nulidade de T é igual a nulidade da matriz $[T]_{\mathcal{B}}^{\mathcal{C}}$.)
- 3.3.9. Determine a matriz do operador $D : \mathcal{P}_n \rightarrow \mathcal{P}_n$ definido por $D(p) = p'$, para todo $p \in \mathcal{P}_n$, em que p' é a derivada de p , em relação à base $\{1, t, t^2, \dots, t^n\}$.

- 3.3.10. Seja $Y = (a, b, c) \in \mathbb{R}^3$. Determine a matriz do operador linear $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, 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.

- 3.3.11. Seja c uma constante diferente de zero. Uma transformação linear $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $T(x, y) = (x, y + cx)$, é chamada **cisalhamento ao longo do eixo y** .

- Mostre que a matriz elementar que corresponde a trocar duas linhas é a matriz de uma reflexão em relação à reta $y = x$.
- 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.
- 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, por exemplo, em [24] as matrizes elementares 2×2 e compare com as matrizes das transformações definidas acima.)

- 3.3.12.** Mostre que se $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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 3.3.11](#).)
- 3.3.13.** Seja A uma matriz triangular superior $n \times n$ com todos os elementos da diagonal iguais a zero. Mostre que $A^n = \tilde{0}$. (Sugestão: considere o operador $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ cuja matriz na base canônica é igual a A e determine a matriz de T^n .)
- 3.3.14.** Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{B}' = \{V'_1, \dots, V'_n\}$ bases de \mathbb{V} e $\mathcal{C} = \{W_1, \dots, W_m\}$ e $\mathcal{C}' = \{W'_1, \dots, W'_m\}$ bases de \mathbb{W} . 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 3.20 na página 256](#).)

- 3.3.15.** Seja $\mathcal{B} = \{V_1, \dots, V_n\}$ uma base de um espaço vetorial \mathbb{V} . Seja P uma matriz $n \times n$ invertível. Mostre que $\mathcal{C} = \{W_1, \dots, W_n\}$ é uma base de \mathbb{V} , onde $W_j = \sum_{i=1}^n p_{ij}V_i$, para $j = 1, \dots, n$. Assim, P é a matriz mudança de base de \mathcal{B} para \mathcal{C} .
- 3.3.16.** Sejam $T : \mathbb{V} \rightarrow \mathbb{U}$ e $S : \mathbb{U} \rightarrow \mathbb{W}$ transformações lineares. Suponha que os espaços vetoriais \mathbb{V}, \mathbb{U} e \mathbb{W} são de dimensão finita. Sejam A e B matrizes tais que o produto AB esteja definido. Mostre que
- $\text{posto}(ST) \leq \text{posto}(S)$. (Sugestão: mostre que $\mathcal{I}(ST) \subseteq \mathcal{I}(S)$.)
 - $\text{posto}(AB) \leq \text{posto}(A)$. (Sugestão: use o item anterior.)
 - $\text{posto}(AB) \leq \text{posto}(B)$. (Sugestão: use o fato de que o posto de uma matriz é igual ao de sua transposta.)
 - $\text{posto}(ST) \leq \text{posto}(T)$. (Sugestão: use o item anterior.)

- (e) Se S é injetiva, então $\text{posto}(ST) = \text{posto}(T)$.
- (f) Se T é sobrejetiva, então $\text{posto}(ST) = \text{posto}(S)$.
- (g) Dê um exemplo em que $\text{posto}(T) = \text{posto}(S) > \text{posto}(ST)$.

3.3.17. Sejam \mathbb{W}_1 e \mathbb{W}_2 subespaços de um espaço vetorial \mathbb{V} tais que $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2$. Defina o operador linear $P : \mathbb{V} \rightarrow \mathbb{V}$, **projecção em \mathbb{W}_1 paralelo a \mathbb{W}_2** da seguinte forma: todo vetor $V \in \mathbb{V}$ se escreve de maneira única como soma $V = V_1 + V_2$, com $V_1 \in \mathbb{W}_1$ e $V_2 \in \mathbb{W}_2$. Define-se então $P(V) = V_1$. Mostre que $\mathcal{I}(P) = \mathbb{W}_1$, $\mathcal{N}(P) = \mathbb{W}_2$ e $P^2 = P$.

3.3.18. Mostre que se $P : \mathbb{V} \rightarrow \mathbb{V}$ é um operador linear tal que $P^2 = P$ (**operador idempotente**), então P é uma projecção. (Sugestão: Sejam $\mathbb{W}_1 = \mathcal{I}(P)$ e $\mathbb{W}_2 = \mathcal{N}(P)$. Para todo $V \in \mathbb{V}$, $V = (V - P(V)) + P(V)$. Mostre que $(V - P(V)) \in \mathcal{N}(P)$ e que se $W \in \mathcal{I}(P)$, então $P(W) = W$.)

3.3.19. Suponha que um espaço vetorial, \mathbb{V} , admita uma decomposição em **soma direta**

$$\mathbb{V} = \mathbb{W}_1 \oplus \dots \oplus \mathbb{W}_k, \text{ isto é,}$$

- (i) $\mathbb{V} = \mathbb{W}_1 + \dots + \mathbb{W}_k$ e
- (ii) $\mathbb{W}_i \cap (\mathbb{W}_1 + \dots + \mathbb{W}_{i-1} + \mathbb{W}_{i+1} + \dots + \mathbb{W}_k) = \{\vec{0}\}$, para $i = 1, \dots, k$.

Para $i = 1, \dots, k$, seja $P_i : \mathbb{V} \rightarrow \mathbb{V}$ definida da seguinte forma. Seja $V \in \mathbb{V}$. V se escreve de maneira única como $V = V_1 + \dots + V_k$, com $V_1 \in \mathbb{W}_1, \dots, V_k \in \mathbb{W}_k$. Então, $P_i(V) = V_i$. Mostre que

- (a) $P_i^2 = P_i$, para $i = 1, \dots, k$.
- (b) $P_1 + \dots + P_k = I$.
- (c) $P_i P_j = O$, a transformação linear nula, se $i \neq j$.

3.3.20. Sejam $P_1, \dots, P_k : \mathbb{V} \rightarrow \mathbb{V}$ operadores lineares tais que

- (i) $P_1 + \dots + P_k = I$.
- (ii) $P_i P_j = O$, a transformação linear nula, se $i \neq j$.

Mostre que

- (a) $P_i^2 = P_i$, para $i = 1, \dots, k$.
- (b) $\mathbb{V} = \mathbb{W}_1 \oplus \dots \oplus \mathbb{W}_k$, onde $\mathbb{W}_i = \mathcal{I}(P_i)$, para $i = 1, \dots, k$.
- (c) Se $V = V_1 + \dots + V_k$, com $V_1 \in \mathbb{W}_1, \dots, V_k \in \mathbb{W}_k$. Então, $P_i(V) = V_i$, para $i = 1, \dots, k$.

3.3.21. Sejam $\mathbb{W}_1, \dots, \mathbb{W}_k$ subespaços de um espaço vetorial \mathbb{V} .

- (a) Seja $\mathbb{V} = \mathbb{W}_1 \oplus \dots \oplus \mathbb{W}_k$. Se \mathcal{B}_i é uma base de \mathbb{W}_i , para $i = 1, \dots, k$, mostre que $\mathcal{B}_1 \cap \dots \cap \mathcal{B}_k = \emptyset$ e $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base de \mathbb{V} .
- (b) Reciprocamente, se \mathcal{B}_i é base de \mathbb{W}_i , para $i = 1, \dots, k$, com $\mathcal{B}_1 \cap \dots \cap \mathcal{B}_k = \emptyset$ e $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ base de \mathbb{V} , mostre que $\mathbb{V} = \mathbb{W}_1 \oplus \dots \oplus \mathbb{W}_k$.

3.3.22. Seja $P : \mathbb{V} \rightarrow \mathbb{V}$ um operador idempotente ($P^2 = P$) de um espaço de dimensão finita. Mostre que

- (a) Existe uma base \mathcal{B} tal que a matriz de P em relação a \mathcal{B} , $A = (a_{ij})_{n \times n}$, é tal que $a_{11} = \dots = a_{rr} = 1$ e os demais $a_{ij} = 0$, onde $r = \text{posto}(P)$.
- (b) O traço de P é igual ao seu posto.

3.3.23. Um operador $T : \mathbb{V} \rightarrow \mathbb{V}$ 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 $T(V) = \bar{0}$.

3.3.24. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear. Seja $\{V_1, \dots, V_p\}$ uma base do núcleo de T . Mostre que o conjunto de todos os vetores $X \in \mathbb{V}$ que são solução da equação $T(X) = W$, para $W \in \mathcal{I}(T)$ é o conjunto dos vetores da forma $X_0 + \alpha_1 V_1 + \dots + \alpha_p V_p$, em que X_0 é uma solução particular de $T(X) = W$.

3.4 A Adjunta

Proposição 3.25. *Seja \mathbb{V} um espaço vetorial de **dimensão finita** com produto interno. Seja f um funcional linear sobre \mathbb{V} , ou seja, uma transformação linear de \mathbb{V} no conjunto de escalares (\mathbb{R} ou \mathbb{C}). Então existe um único vetor $W \in \mathbb{V}$ tal que*

$$f(V) = \langle V, W \rangle, \quad \text{para todo } V \in \mathbb{V}.$$

Demonstração. Seja $g_W : \mathbb{V} \rightarrow \mathbb{R}$ (ou \mathbb{C}) a função definida por $g_W(V) = \langle V, W \rangle$, para todo $V \in \mathbb{V}$. Segue das propriedades do produto interno que g_W é um funcional linear. Seja $\mathcal{B} = \{U_1, \dots, U_n\}$ uma base ortonormal de \mathbb{V} . Então, pela **Proposição 2.7** na página 122,

$$W = \langle W, U_1 \rangle U_1 + \dots + \langle W, U_n \rangle U_n.$$

Mas, $f(U_k) = g_W(U_k)$ se, e somente se, $f(U_k) = \langle U_k, W \rangle$ ou $\langle W, U_k \rangle = \overline{f(U_k)}$.

Assim, seja $W = \sum_{k=1}^n \overline{f(U_k)} U_k$. Então,

$$g_W(U_j) = \langle U_j, W \rangle = \left\langle U_j, \sum_{k=1}^n \overline{f(U_k)} U_k \right\rangle = \sum_{k=1}^n f(U_k) \langle U_j, U_k \rangle = f(U_j).$$

Ou seja, f e g_W são iguais nos elementos da base \mathcal{B} . Logo, pelo **Teorema 3.1** na página 213, temos que $f(V) = g_W(V) = \langle V, W \rangle$, para todo $V \in \mathbb{V}$. ■

Exemplo 3.35. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ um funcional linear. Então, existe uma matriz $A = [a_1 \dots a_n]$, $1 \times n$, tal que

$$f(X) = f \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = AX = [a_1 \dots a_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Neste caso, o vetor

$$W = f(E_1)E_1 + \dots + f(E_n)E_n = a_1E_1 + \dots + a_nE_n = (a_1, \dots, a_n)$$

é tal que $f(X) = \langle X, W \rangle$.

Definição 3.7. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear entre espaços vetoriais com produto interno. Uma função $T^* : \mathbb{W} \rightarrow \mathbb{V}$ é chamada a **adjunta** de T se,

$$\langle T(V), W \rangle = \langle V, T^*(W) \rangle, \quad \text{para todos } V \in \mathbb{V} \text{ e } W \in \mathbb{W}. \quad (3.5)$$

Simbolicamente, a equação (3.5) está dizendo que temos que “colocar” $*$ em T , quando “passamos” T para o outro lado do produto interno.

Proposição 3.26. *Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear entre espaços vetoriais com produto interno. Se T tem uma adjunta, $T^* : \mathbb{W} \rightarrow \mathbb{V}$, então ela é uma transformação linear e é a única que satisfaz (3.5).*

Demonstração. Vamos mostrar em primeiro lugar que se T^* é adjunta de T , então ela é uma transformação linear. Sejam $W_1, W_2 \in \mathbb{W}$ e α, β escalares. Seja $V \in \mathbb{V}$ qualquer.

$$\begin{aligned} \langle V, T^*(\alpha W_1 + \beta W_2) \rangle &= \langle T(V), \alpha W_1 + \beta W_2 \rangle \\ &= \bar{\alpha} \langle T(V), W_1 \rangle + \bar{\beta} \langle T(V), W_2 \rangle \\ &= \bar{\alpha} \langle V, T^*(W_1) \rangle + \bar{\beta} \langle V, T^*(W_2) \rangle \\ &= \langle V, \alpha T^*(W_1) + \beta T^*(W_2) \rangle \end{aligned}$$

De onde segue-se que

$$\langle V, T^*(\alpha W_1 + \beta W_2) - (\alpha T^*(W_1) + \beta T^*(W_2)) \rangle = 0, \quad \text{para todo vetor } V \in \mathbb{V},$$

em particular para $V = T^*(\alpha W_1 + \beta W_2) - (\alpha T^*(W_1) + \beta T^*(W_2))$. O que implica que $V = \bar{0}$. Portanto, $T^*(\alpha W_1 + \beta W_2) = \alpha T^*(W_1) + \beta T^*(W_2)$, como queríamos provar.

Seja $U : \mathbb{W} \rightarrow \mathbb{V}$ uma transformação linear tal que

$$\langle T(V), W \rangle = \langle V, T^*(W) \rangle = \langle V, U(W) \rangle, \quad \text{para todos } V \in \mathbb{V} \text{ e } W \in \mathbb{W}.$$

Então, $\langle V, T^*(W) - U(W) \rangle = 0$, para todos $V \in \mathbb{V}$ e $W \in \mathbb{W}$. Isto implica, tomando $V = T^*(W) - U(W)$, que $T^*(W) - U(W) = \vec{0}$, para todo $W \in \mathbb{W}$. De onde segue-se que $T^* = U$.

■

Teorema 3.27. *Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear entre espaços de **dimensão finita** com produto interno. Então T possui adjunta T^* , ou seja, existe uma única transformação linear*

$$T^* : \mathbb{W} \rightarrow \mathbb{V}$$

que satisfaz

$$\langle T(V), W \rangle = \langle V, T^*(W) \rangle, \quad \text{para todos } V \in \mathbb{V} \text{ e } W \in \mathbb{W}.$$

Demonstração. Seja $W \in \mathbb{W}$. Seja $g_W : \mathbb{V} \rightarrow \mathbb{R}$ (ou \mathbb{C}) a função definida por $g_W(V) = \langle T(V), W \rangle$, para todo $V \in \mathbb{V}$. Segue das propriedades do produto interno e da linearidade de T que g_W é um funcional linear.

Pela **Proposição 3.25** existe um único $V' \in \mathbb{V}$ (que depende de W) tal que

$$g_W(V) = \langle T(V), W \rangle = \langle V, V' \rangle, \quad \text{para todo } V \in \mathbb{V}.$$

Seja $T^* : \mathbb{W} \rightarrow \mathbb{V}$ a função definida por $T^*(W) = V'$, para todo $W \in \mathbb{W}$. Já mostramos que T^* é única e é uma transformação linear.

■

Para determinar a adjunta de uma transformação linear T só precisamos da matriz de T em relação a bases ortonormais.

Proposição 3.28. Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_m\}$ bases *ortonormais* de \mathbb{V} e \mathbb{W} , respectivamente. Se $T : \mathbb{V} \rightarrow \mathbb{W}$ é uma transformação linear, então

$$[T]_{\mathcal{B}}^{\mathcal{C}} = A, \quad \text{onde } [A]_{ij} = a_{ij} = \langle T(V_j), W_i \rangle, \text{ para } i = 1, \dots, m \text{ e } j = 1, \dots, n.$$

Demonstração. Seja $A = (a_{ij})_{m \times n}$ a matriz definida por $a_{ij} = \langle T(V_j), W_i \rangle$. Pela [Proposição 2.7 na página 122](#), para o vetor $T(V_j)$, temos que

$$T(V_j) = \sum_{i=1}^m \langle T(V_j), W_i \rangle W_i.$$

Então, $[T(V_j)]_{\mathcal{C}} = \begin{bmatrix} \langle T(V_j), W_1 \rangle \\ \vdots \\ \langle T(V_j), W_m \rangle \end{bmatrix}$. Como $[T]_{\mathcal{B}}^{\mathcal{C}} = [[T(V_1)]_{\mathcal{C}} \dots [T(V_n)]_{\mathcal{C}}]$, então

$$[T]_{\mathcal{B}}^{\mathcal{C}} = A.$$

■

Proposição 3.29. Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_m\}$ bases *ortonormais* de \mathbb{V} e \mathbb{W} , respectivamente. Se $T : \mathbb{V} \rightarrow \mathbb{W}$ é uma transformação linear e $A = [T]_{\mathcal{B}}^{\mathcal{C}}$, então

$$[T^*]_{\mathcal{C}}^{\mathcal{B}} = B, \quad \text{onde } [B]_{ij} = b_{ij} = \bar{a}_{ji}, \text{ para } i = 1, \dots, n \text{ e } j = 1, \dots, m.$$

Demonstração. Pela [Proposição 3.28](#) a matriz $B = [T^*]_{\mathcal{C}}^{\mathcal{C}}$ é dada por

$$b_{ij} = \langle T^*(W_j), V_i \rangle = \overline{\langle V_i, T^*(W_j) \rangle} = \overline{\langle T(V_i), W_j \rangle} = \bar{a}_{ji}$$

■

Exemplo 3.36. Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma transformação linear. Então existe uma matriz A , $m \times n$, tal que

$$T(X) = AX, \quad \text{para todo } X \in \mathbb{R}^n.$$

Como a base canônica do \mathbb{R}^n é uma base ortonormal, então a adjunta $T^* : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é dada por

$$T^*(Y) = A^t Y, \quad \text{para todo } Y \in \mathbb{R}^m.$$

Podemos, agora, provar mais algumas propriedades da álgebra das transformações lineares.

Proposição 3.30. *Sejam S e T transformações lineares com domínios e contra-domínios apropriados. Então,*

$$(a) (S + T)^* = S^* + T^*; \quad (b) (\alpha T)^* = \bar{\alpha} T^*; \quad (c) (ST)^* = T^* S^*; \quad (d) T^{**} = T; \quad (e) I^* = I.$$

Demonstração. Para provar estas propriedades vamos usar o fato de que duas transformações lineares $T_1, T_2 : \mathbb{V} \rightarrow \mathbb{W}$ são iguais se $\langle W, T_1(V) \rangle = \langle W, T_2(V) \rangle$, para todos $V \in \mathbb{V}$ e $W \in \mathbb{W}$.

(a)

$$\begin{aligned} \langle V, (S + T)^*(W) \rangle &= \langle (S + T)(V), W \rangle = \langle S(V), W \rangle + \langle T(V), W \rangle = \\ &= \langle V, S^*(W) \rangle + \langle V, T^*(W) \rangle = \langle V, S^*(W) + T^*(W) \rangle = \langle V, (S^* + T^*)(W) \rangle \end{aligned}$$

(b)

$$\langle V, (\alpha T)^*(W) \rangle = \langle (\alpha T)(V), W \rangle = \alpha \langle T(V), W \rangle = \alpha \langle V, T^*(W) \rangle = \langle V, (\bar{\alpha} T^*)(W) \rangle.$$

(c)

$$\langle V, (ST)^*(W) \rangle = \langle (ST)(V), W \rangle = \langle S(T(V)), W \rangle = \langle T(V), S^*(W) \rangle = \langle V, (T^* S^*)(W) \rangle.$$

(d)

$$\langle T^{**}(V), W \rangle = \overline{\langle W, T^{**}(V) \rangle} = \overline{\langle T^*(W), V \rangle} = \langle V, T^*(W) \rangle = \langle T(V), W \rangle.$$

(e) Este item é simples e deixamos como exercício para o leitor.



Teorema 3.31. *Sejam $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear entre espaços vetoriais com produto interno de dimensão finita. Então*

$$(a) \mathcal{N}(T) = \mathcal{I}(T^*)^\perp e \mathbb{V} = \mathcal{N}(T) \oplus \mathcal{I}(T^*).$$

$$(b) \mathcal{N}(T^*) = \mathcal{I}(T)^\perp e \mathbb{W} = \mathcal{N}(T^*) \oplus \mathcal{I}(T).$$

Demonstração. (a) Um vetor $V \in \mathcal{N}(T)$ se, e somente se, $\langle T(V), W \rangle = 0$, para todo $W \in \mathbb{W}$. Mas, $\langle T(V), W \rangle = \langle V, T^*(W) \rangle = 0$, para todo $W \in \mathbb{W}$ se, e somente se, $V \in \mathcal{I}(T^*)^\perp$. Portanto, $\mathcal{N}(T) = \mathcal{I}(T^*)^\perp e \mathbb{V} = \mathcal{N}(T) \oplus \mathcal{I}(T^*)$, pela **Proposição 2.11 na página 148**.

(b) Basta aplicar o item anterior a T^* e usar o fato de que $T^{**} = T$.



3.4.1 Aplicação: Problema de Quadrados Mínimos

Muitos problemas, quando modelados, levam a sistemas lineares $AX = B$, 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,

$$\min \|AX - B\|^2.$$

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.

Teorema 3.32. *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 equações normais.

Demonstração. Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ a transformação linear definida por $T(X) = AX$, para todo $X \in \mathbb{R}^n$. O problema de quadrados mínimos

$$\min \|AX - B\|^2$$

pode ser escrito como

$$\min_{Y \in \mathcal{I}(T)} \|Y - B\|^2 \quad \text{e} \quad Y = A X.$$

Um vetor $X \in \mathbb{R}^n$ é solução deste problema se, e somente se, $B - Y = B - A X \in \mathcal{I}(T)^\perp$ (Teorema 2.12 na página 157). Mas, pelo Teorema 3.31, $\mathcal{I}(T)^\perp = \mathcal{N}(T^*)$. Assim, X é solução do problema de quadrados mínimos se, e somente se,

$$A^t(B - AX) = \vec{0}.$$

Ou seja, a solução do problema de quadrados mínimos é a solução do sistema linear

$$A^t A X = A^t B.$$

■

Exemplo 3.37. 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 $\sum_{i=1}^4 (y_i - ax_i - b)^2$ seja mínimo. Substituindo-se estes pontos na equação da reta obtemos o seguinte sistema

$$\begin{cases} -3a + b = 6 \\ + b = 4 \\ + b = 0 \\ 2a + b = 2 \end{cases}$$

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 nor-

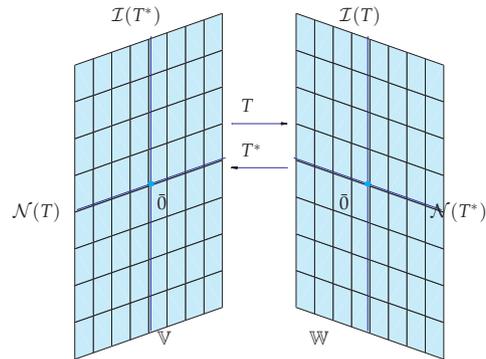


Figura 3.17: Subespaços $\mathcal{N}(T)$, $\mathcal{I}(T^*)$, $\mathcal{I}(T)$ e $\mathcal{N}(T^*)$

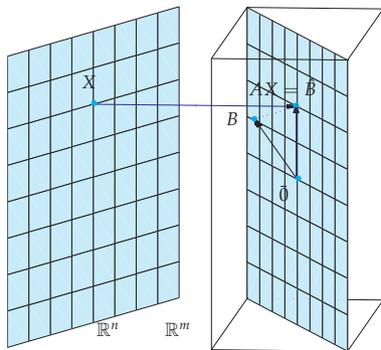


Figura 3.18: A solução de quadrados mínimos

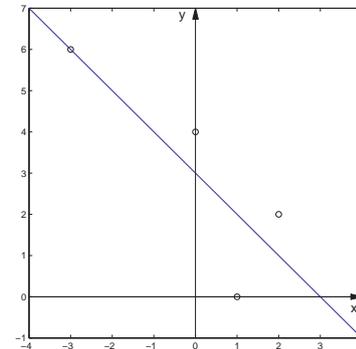


Figura 3.19: Retas que “melhor” se ajusta a quatro pontos

mais $A^tAX = A^tB$. 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 3.38. 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

$$\begin{cases} 4a - 2b + c = 0 \\ a - b + c = 2 \\ a + b + c = 2 \\ 4a + 2b + c = 10 \end{cases}$$

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 3.39. 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

$$\begin{cases} -2a & & + c & = & 4 \\ & + 2b & + c & = & 4 \\ & a & - 3b & + c & = 10 \\ & 3a & + b & + c & = 10 \end{cases}$$

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 $A^t A X = A^t B$. Aqui,

$$A^t A = \begin{bmatrix} 14 & 0 & 2 \\ 0 & 14 & 0 \\ 2 & 0 & 4 \end{bmatrix} \quad \text{e} \quad A^t B = \begin{bmatrix} 44 \\ 20 \\ 14 \end{bmatrix}$$

Escalonando a matriz aumentada $[A^t A | A^t B]$ obtemos que a solução de quadrados mínimos é $X = [18/13 \quad -6/7 \quad 82/13]^t$, 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 $r = \sqrt{\frac{57724}{8281}} \approx 2,6$.

Exemplo 3.40. Vamos, agora, acrescentar o par $(1/2, 3/2)$ ao conjunto de dados do Exemplo 1.70 na página 85 obtendo o seguinte conjunto de dados

x	-1	0	1/2	1	2	3
y	1	4	3/2	1	3	0

Como no Exemplo 1.70 na página 85, dividindo-se o intervalo $[-1, 3]$ em dois subintervalos e usando a base $\{q_1, q_2, q_3, q_4, q_5\}$, o problema de encontrar um spline

$$f(x) = c_1q_1(x) + c_2q_2(x) + c_3q_3(x) + c_4q_4(x) + c_5q_5(x)$$

que melhor se ajusta ao conjunto de pontos $(-1, 1)$, $(0, 4)$, $(1/2, 3/2)$, $(1, 1)$, $(2, 3)$, $(3, 0)$ no sentido de quadrados mínimos toma a forma

$$\min \sum_{i=1}^6 (c_1q_1(x_i) + c_2q_2(x_i) + c_3q_3(x_i) + c_4q_4(x_i) + c_5q_5(x_i) - f(x_i))^2$$

ou

$$\min \|AX - B\|,$$

em que a matriz A é definida por $a_{ij} = q_j(x_i)$, B por $b_j = y_j$ e X por $x_j = c_j$, para $i = 1, \dots, 6$, $j = 1, \dots, 5$. Neste caso

$$A = \begin{bmatrix} \frac{1}{4} & 1 & \frac{1}{4} & 0 & 0 \\ \frac{1}{32} & \frac{23}{32} & \frac{23}{32} & \frac{1}{32} & 0 \\ \frac{1}{256} & \frac{121}{256} & \frac{235}{256} & \frac{27}{256} & 0 \\ 0 & \frac{1}{4} & 1 & \frac{1}{4} & 0 \\ 0 & \frac{1}{32} & \frac{23}{32} & \frac{23}{32} & \frac{1}{32} \\ 0 & 0 & \frac{1}{4} & 1 & \frac{1}{4} \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 4 \\ 3/2 \\ 1 \\ 3 \\ 0 \end{bmatrix}, \quad X = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix},$$

Os coeficientes c_j obtidos resolvendo o problema de quadrados mínimos são

$$c_1 = -34.4039, \quad c_2 = 10.7147, \quad c_3 = -4.3737, \quad c_4 = 9.6455, \quad c_5 = -34.2351$$

Exercícios Numéricos (respostas na página 448)

3.4.1. Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma transformação linear definida por $T(X) = AX$, para todo $X \in \mathbb{R}^n$. Determine uma base para cada um dos seguintes subespaços $\mathcal{I}(T^*)$, $\mathcal{N}(T)$, $\mathcal{I}(T)$ e $\mathcal{N}(T^*)$, para cada uma das seguintes matrizes.

$$(a) \begin{bmatrix} 4 & -2 \\ 1 & 3 \\ 2 & 1 \\ 3 & 4 \end{bmatrix}$$

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

3.4.2. 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 .

3.4.3. 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 .

3.4.4. Encontre a equação da parábola que melhor se ajusta aos pontos dados no sentido de quadrados mínimos, ou seja, tal que $\sum_{i=1}^4 (y_i - ax_i^2 - bx_i - c)^2$ seja mínimo:

$$(a) P_1 = (-2, 1), P_2 = (-1, 2), P_3 = (1, 0) \text{ e } P_4 = (2, 7).$$

$$(b) P_1 = (-2, 1), P_2 = (-1, 3), P_3 = (1, 3) \text{ e } P_4 = (2, 11).$$

3.4.5. Encontre a equação do círculo que melhor se ajusta aos pontos dados 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:

$$(a) P_1 = (-2, 0), P_2 = (0, 1), P_3 = (1, -2) \text{ e } P_4 = (2, 1).$$

$$(b) P_1 = (-2, 1), P_2 = (-1, -2), P_3 = (0, 1) \text{ e } P_4 = (2, 0).$$

3.4.6. Encontre a solução de quadrados mínimos dos seguintes sistemas:

$$(a) \begin{cases} x + 2y = 3 \\ 2x + 4y = 2 \\ -x - 2y = 1 \end{cases} \quad (b) \begin{cases} -x + y = 10 \\ 2x + y = 5 \\ x - 2y = 20 \end{cases} \quad (c) \begin{cases} x + y + z = 4 \\ -x + y + z = 2 \\ -y + z = 1 \\ x + z = 2 \end{cases}$$

Exercícios usando o MATLAB[®]

Comandos do MATLAB[®]:

>> `x=linspace(a,b)` cria um vetor contendo 100 valores igualmente espaçados entre a e b .

>> `plot(x,f(x))` desenha a função $f(x)$ ligando os pontos $(x_i, f(x_i))$.

Comandos do pacote GAAL:

>> `qk=spline1(k,x,nbp,a,b)` calcula o spline q_k em x para um intervalo $[a, b]$ dividido em $nbp-1$ subintervalos.

>> `A=spline1(X,nbp,a,b)` cria a matriz $a_{ij} = q_j(X_i)$ para um intervalo $[a, b]$ dividido em $nbp-1$ subintervalos.

>> `f=spline1(C,x,nbp,a,b)` calcula a soma $C_k q_k(x)$ com $k=1:nbp+2$

>> `po([X,Y])` desenha os pontos $(X(i), Y(i))$.

- 3.4.7. (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])`, onde a, b, c e d são os coeficientes já encontrados. Desenhe os eixos coordenados com o comando `eixos`.

- 3.4.8. (a) Use o comando $P = \text{randi}(6, 2)$, 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 $\sum (x_i^2 - ax_i y_i - by_i^2 - cx_i - dy_i - e)^2$ seja mínimo. As matrizes $M = \text{matvand}(P, 2)$, $B = -M(:, 1)$ 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])`, onde a, b, c, d e e são os coeficientes encontrados no item anterior. Desenhe os eixos coordenados com o comando `eixos`.

Exercícios Teóricos

- 3.4.9. Sejam $T, S : \mathbb{V} \rightarrow \mathbb{W}$ transformações lineares entre espaços vetoriais com produto interno. Mostre que se

$$\langle T(V), W \rangle = \langle S(V), W \rangle, \quad \text{para todos } V \in \mathbb{V} \text{ e } W \in \mathbb{W}, \text{ então } T = S.$$

- 3.4.10. Seja \mathbb{W} um subespaço de um espaço vetorial com produto interno \mathbb{V} , gerado pelos vetores V_1, \dots, V_k . Mostre que $V \in \mathbb{W}^\perp$ se, e somente se, V é ortogonal a V_i , para $i = 1, \dots, k$.

- 3.4.11. Seja $T : \mathbb{V} \rightarrow \mathbb{W}$ uma transformação linear entre espaços vetoriais de dimensão finita com produto interno.

- (a) Mostre que $\mathcal{N}(T^*T) = \mathcal{N}(T)$;
(Sugestão: se $T^*T(V) = \vec{0}$, então $T(V) \in \mathcal{I}(T) \cap \mathcal{N}(T^*)$.)
- (b) Mostre que se T é injetiva, então T^*T é invertível.
- (c) Mostre que $\mathcal{I}(T^*) = \mathcal{I}(T^*T)$.
- (d) Mostre que se T é sobrejetiva, então TT^* é invertível.

3.4.12. Sejam A uma matriz $m \times n$ com colunas linearmente independentes e B uma matriz $m \times 1$. Mostre que neste caso, a matriz $A^t A$ é invertível e que vale a seguinte fórmula para a solução do problema de quadrados mínimos, $\min \|AX - B\|^2$,

$$X = (A^t A)^{-1} A^t B.$$

3.4.13. Sejam $\mathcal{B} = \{V_1, \dots, V_n\}$ e $\mathcal{C} = \{W_1, \dots, W_n\}$ duas bases ortonormais de um espaço vetorial com produto interno \mathbb{V} real.

(a) Mostre que a matriz mudança de base de \mathcal{B} para \mathcal{C} é

$$P = (p_{ij})_{n \times n}, \quad \text{em que } p_{ij} = \langle V_j, W_i \rangle, \text{ para } i, j = 1, \dots, n.$$

(b) Mostre que $\langle V_i, V_j \rangle = \sum_{k=1}^n \langle V_i, W_k \rangle \langle V_j, W_k \rangle$.

(c) Mostre que P é uma matriz ortogonal, ou seja, que $P^{-1} = P^t$. (Sugestão: mostre que $P^t P = I_n$.)

3.4.14. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço com produto interno.

(a) Mostre que se \mathbb{W} é um subespaço de \mathbb{V} tal que $T(W) \in \mathbb{W}$, para todo $W \in \mathbb{W}$, então $T^*(V) \in \mathbb{W}^\perp$, para todo $V \in \mathbb{W}^\perp$. (Sugestão: tome $V \in \mathbb{W}^\perp$ e mostre que $\langle W, T^*(V) \rangle = 0$, para todo $W \in \mathbb{W}$.)

(b) Mostre que se $TT^* = T^*T$, então $\|T(V)\| = \|T^*(V)\|$, para todo vetor $V \in \mathbb{V}$. (Sugestão: use o fato de que $\|T(V)\|^2 = \langle T(V), T(V) \rangle$.)

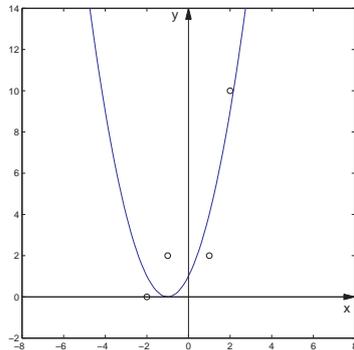


Figura 3.20: Parábola que “melhor” se ajusta a quatro pontos

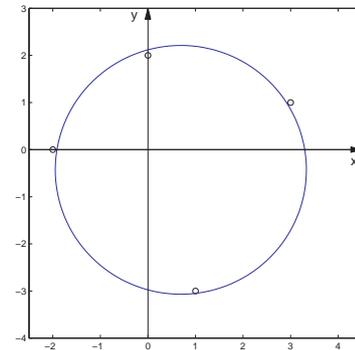


Figura 3.21: Círculo que “melhor” se ajusta a quatro pontos

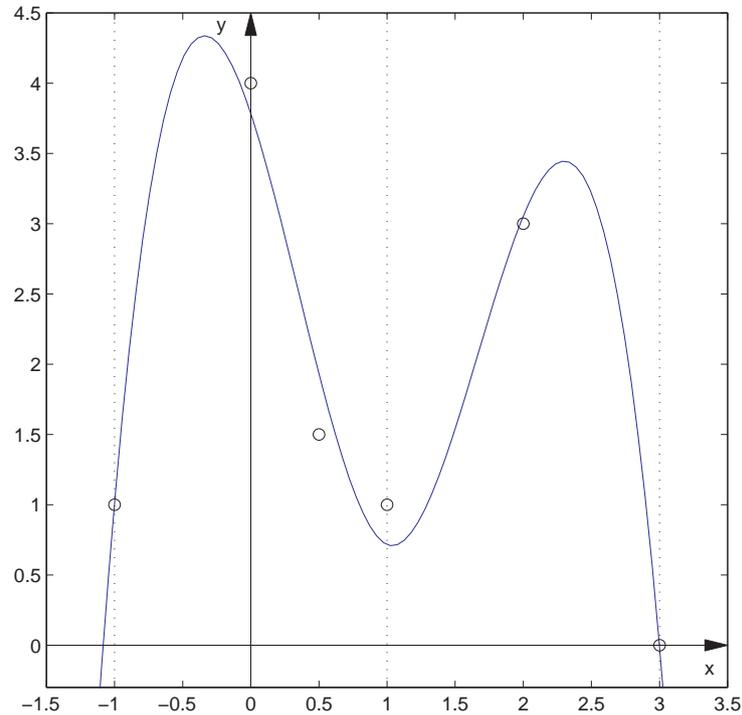


Figura 3.22: Ajuste dos dados do Exemplo 3.40 por splines dividindo-se o intervalo $[-1, 3]$ em dois subintervalos

Teste do Capítulo

1. Seja $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ uma aplicação linear tal que $F(E_1) = E_1 + E_2$, $F(E_2) = E_1$ e $F(E_3) = E_2$, em que $\{E_1, E_2, E_3\}$ é a base canônica do \mathbb{R}^3 .
 - (a) Ache $F(x, y, z)$, onde $(x, y, z) \in \mathbb{R}^3$.
 - (b) Ache uma base para $\mathcal{N}(F)$, uma base para a $\mathcal{I}(F)$ e suas dimensões.
 - (c) Diga se F é injetiva, sobrejetiva e isomorfismo.
 2. Dê um exemplo de um isomorfismo $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ diferente da identidade do \mathbb{R}^3 . Prove que é isomorfismo.
 3. Dê um exemplo de uma transformação linear T sobrejetiva e não injetiva. Prove que é sobrejetiva e não injetiva.
 4. Considere o plano $\pi : x + y + z = 0$. 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 $U_1 = 1/\sqrt{3}(1, 1, 1)$, $U_2 = 1/\sqrt{2}(-1, 1, 0)$ e $U_3 = 1/\sqrt{6}(-1, -1, 2)$.
 - (a) Seja $P_\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ a projeção ortogonal no plano π . Encontre $[P_\pi]_{\mathcal{C}}^{\mathcal{C}}$ e $[P_\pi]_{\mathcal{B}}^{\mathcal{B}}$.
 - (b) Seja $R_\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ a reflexão em relação ao plano π . Encontre $[R_\pi]_{\mathcal{C}}^{\mathcal{C}}$ e $[R_\pi]_{\mathcal{B}}^{\mathcal{B}}$.
-

Capítulo 4

Diagonalização

4.1 Diagonalização de Operadores

4.1.1 Motivação: Sistemas de Equações Diferenciais Lineares

Vamos considerar o problema de encontrar as funções que dão a evolução das populações de duas espécies, S_1 e S_2 , convivendo em um mesmo ecossistema no tempo $t > 0$. Vamos denotar as populações das espécies S_1 e S_2 em um instante t por $x_1(t)$ e $x_2(t)$, respectivamente.

Inicialmente vamos supor que a taxa de crescimento da população de uma espécie não depende do que ocorre com a outra espécie e que esta taxa é proporcional a sua população existente (ou equivalentemente que a taxa de crescimento relativa é constante). Ou seja, vamos supor que

$$\begin{aligned}x_1'(t) &= \frac{dx_1}{dt}(t) = ax_1(t) \\x_2'(t) &= \frac{dx_2}{dt}(t) = dx_2(t)\end{aligned}$$

em que $a, d \in \mathbb{R}$. Temos aqui um sistema de equações diferenciais, ou seja, um sistema de equações que envolvem derivadas das funções que são incógnitas. Neste caso as duas equações são desacopladas, isto é, podem ser resolvidas independentemente. A solução do sistema é $x_1(t) = x_1(0)e^{at}$ e $x_2(t) = x_2(0)e^{dt}$, para $t \geq 0$ (Exemplo 1.19 na página 22).

Vamos supor, agora, que as duas populações interagem de forma que a taxa de crescimento da população de uma espécie depende de forma linear não somente da sua população existente, mas também da população existente da outra espécie. Ou seja, vamos supor que

$$\begin{aligned}x_1'(t) &= ax_1(t) + bx_2(t) \\x_2'(t) &= cx_1(t) + dx_2(t)\end{aligned}$$

Por exemplo, se os indivíduos de uma espécie competem com os da outra por alimento ($a, d > 0$ e $b, c < 0$), ou os indivíduos da espécie S_1 são predadores dos da outra ($a, b, d > 0$ e $c < 0$). Neste caso a solução de uma equação depende da outra. Podemos escrever este sistema na forma de uma equação diferencial matricial

$$X'(t) = AX(t), \tag{4.1}$$

em que $X'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix}$, $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ e $X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$. Vamos supor que

existam matrizes P e D tais que

$$A = PDP^{-1}, \quad (4.2)$$

em que $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. Substituindo-se (4.2) em (4.1) obtemos

$$X'(t) = PDP^{-1}X(t).$$

Multiplicando-se à esquerda por P^{-1} , obtemos

$$P^{-1}X'(t) = DP^{-1}X(t).$$

Fazendo a mudança de variável $Y(t) = P^{-1}X(t)$, obtemos

$$Y'(t) = DY(t),$$

que pode ser escrito na forma

$$\begin{aligned} y_1'(t) &= \lambda_1 y_1(t) \\ y_2'(t) &= \lambda_2 y_2(t) \end{aligned}$$

Estas equações estão desacopladas e têm soluções dadas por $y_1(t) = y_1(0)e^{\lambda_1 t}$ e $y_2(t) = y_2(0)e^{\lambda_2 t}$. Assim a solução da equação (4.1) é

$$X(t) = PY(t) = P \begin{bmatrix} y_1(0)e^{\lambda_1 t} \\ y_2(0)e^{\lambda_2 t} \end{bmatrix} = P \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} Y(0) = P \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} P^{-1}X(0).$$

Observe que o que possibilitou a resolução do sistema de equações foi a hipótese feita de que a matriz A pode ser escrita como $A = PDP^{-1}$, em que D é uma matriz diagonal.

Vamos descobrir como podemos determinar matrizes P e D , quando elas existem, tais que $A = PDP^{-1}$, ou equivalentemente, $D = P^{-1}AP$, com D sendo uma matriz diagonal. Chamamos **diagonalização** ao processo de encontrar as matrizes P e D .

4.1.2 Operadores e Matrizes Diagonalizáveis

Definição 4.1. 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

$$B = P^{-1}AP$$

A relação de semelhança satisfaz as seguintes propriedades:

- toda matriz quadrada é semelhante a si mesma;
- se uma matriz A é semelhante a B , então B é semelhante a A e
- se A é semelhante a B e B é semelhante a C , então A é semelhante a C .

Deixamos como exercício a verificação destas propriedades.

Definição 4.2. (a) Dizemos que uma matriz A , $n \times n$, é **diagonalizável**, se ela é semelhante a uma matriz diagonal. Ou seja, se existem matrizes Q e D tais que $A = Q^{-1}DQ$, em que D é uma matriz diagonal.

(b) Dizemos que um operador $T : \mathbb{V} \rightarrow \mathbb{V}$ de um espaço de dimensão finita \mathbb{V} é **diagonalizável**, se existe uma base \mathcal{C} de \mathbb{V} tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é uma matriz diagonal.

Exemplo 4.1. Toda matriz diagonal

$$A = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

é diagonalizável, pois

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

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 4.1. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador de um espaço vetorial \mathbb{V} de dimensão finita. Seja \mathcal{B} uma base de \mathbb{V} . T é um operador diagonalizável se, e somente se, $[T]_{\mathcal{B}}^{\mathcal{B}}$ é uma matriz diagonalizável.*

Demonstração. Se um operador linear $T : \mathbb{V} \rightarrow \mathbb{V}$ é diagonalizável, então existe uma base \mathcal{C} de \mathbb{V} tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é diagonal. Então, pelo [Corolário 3.20 na página 256](#) 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 $\mathcal{B} = \{V_1, \dots, V_n\}$, então existe uma matriz P tal que

$$D = P^{-1}AP$$

é uma matriz diagonal. Seja $\mathcal{C} = \{W_1, \dots, W_n\}$, em que

$$W_j = \sum_{i=1}^n p_{ij}V_i, \text{ para } j = 1, \dots, n.$$

\mathcal{C} é uma base de \mathbb{V} ([Exercício 3.3.15 na página 267](#)) e

$$[T]_{\mathcal{C}}^{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}^{\mathcal{B}}P = D.$$

Portanto, T é um operador diagonalizável. ■

4.1.3 Autovalores e Autovetores

Se um operador T é diagonalizável, então existe uma base $\mathcal{C} = \{V_1, \dots, V_n\}$ tal que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}.$$

O que implica que

$$T(V_j) = \sum_{i=1}^n d_{ij}V_i = \lambda_j V_j, \quad \text{para } j = 1, \dots, n.$$

Isto motiva a seguinte definição.

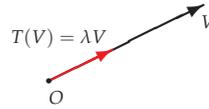
Definição 4.3. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear. Um escalar λ é chamado **autovalor** de T , se existe um vetor *não nulo* $V \in \mathbb{V}$, tal que

$$T(V) = \lambda V. \tag{4.3}$$

Um vetor *não nulo* que satisfaça (4.3), é chamado de **autovetor** de T .



$$\lambda > 1$$



$$0 < \lambda < 1$$



$$\lambda < 0$$

Observe que a equação (4.3) pode ser escrita como

$$T(V) = \lambda I(V)$$

ou

$$(T - \lambda I)(V) = \vec{0}. \quad (4.4)$$

Assim, os autovetores são os vetores não nulos que pertencem ao núcleo de $T - \lambda I$ e os autovalores são os escalares λ tais que

$$\mathcal{N}(T - \lambda I) \neq \{\vec{0}\}.$$

Seja $\mathcal{B} = \{V_1, \dots, V_n\}$ uma base de \mathbb{V} .

Escrevendo a equação (4.4) na base \mathcal{B} obtemos

$$(A - \lambda I_n)X = \vec{0}.$$

em que $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ e $X = [V]_{\mathcal{B}}$. E esta equação tem solução não trivial se, e somente se, $\det(A - \lambda I_n) = 0$ (ver por exemplo [24]).

Com a definição de determinante de um operador (Página 256), podemos dizer que existe um vetor **não nulo** $V \in \mathbb{V}$ que satisfaz (4.4) se, e somente se,

$$\det(T - \lambda I) = 0.$$

Assim temos um método para encontrar os autovalores e os autovetores de um operador T .

Proposição 4.2. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear de um espaço de dimensão finita.*

(a) *Os autovalores de T são as raízes do polinômio*

$$p(\lambda) = \det(T - \lambda I_n) \quad (4.5)$$

que estão no conjunto de escalares.

(b) *Para cada autovalor λ , os autovetores associados a λ são os vetores não nulos do núcleo de $T - \lambda I$,*

$$\mathcal{N}(T - \lambda I).$$

Definição 4.4. *Seja T um operador em um espaço de dimensão finita. O polinômio*

$$p(\lambda) = \det(T - \lambda I) \quad (4.6)$$

*é chamado **polinômio característico de T** .*

Assim, para determinarmos os autovalores de um operador T em um espaço de dimensão finita precisamos determinar as raízes, que estão no conjunto de escalares, do seu polinômio característico, que tem a forma

$$p(\lambda) = (-1)^n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0.$$

Um resultado sobre polinômios que muitas vezes é útil, é o que diz que se a_0, a_1, \dots, 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

$$p(\lambda) = -\lambda^3 + 6\lambda^2 - 11\lambda + 6,$$

então as possíveis raízes racionais são $\pm 1, \pm 2, \pm 3$ e ± 6 . Substituindo estes valores em $p(\lambda)$, vemos que $p(1) = 0$, ou seja, 1 é uma raiz de $p(\lambda)$. Finalmente, dividindo $p(\lambda)$ por $\lambda - 1$, obtemos que

$$p(\lambda) = (\lambda - 1)(-\lambda^2 + 5\lambda - 6).$$

Como as raízes de $-\lambda^2 + 5\lambda - 6$ são 2 e 3, então as raízes de $p(\lambda)$, são 1, 2 e 3.

Exemplo 4.2. Vamos determinar os autovalores e autovetores do operador $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por $T(X) = AX$, para todo $X \in \mathbb{R}^3$, em que

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

Para este operador o polinômio característico é

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I_3) = \det \begin{bmatrix} 2 - \lambda & 0 & 1 \\ 2 & 1 - \lambda & 2 \\ 1 & 0 & 2 - \lambda \end{bmatrix} \\ &= (1 - \lambda) \det \begin{bmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{bmatrix} \\ &= (1 - \lambda)(\lambda^2 - 4\lambda + 3) = (1 - \lambda)^2(3 - \lambda). \end{aligned}$$

Como os autovalores de A são as raízes reais de $p(\lambda)$, temos que os autovalores de A são $\lambda_1 = 1$ e $\lambda_2 = 3$.

Agora, vamos determinar os autovetores associados aos autovalores $\lambda_1 = 1$ e $\lambda_2 = 3$. Para isto vamos resolver os sistemas $(A - \lambda_1 I_3)X = \vec{0}$ e $(A - \lambda_2 I_3)X = \vec{0}$. O sistema

$$(A - \lambda_1 I_3)X = \vec{0} \quad \text{é}$$

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

cuja solução geral é

$$\mathbb{W}_1 = \{(-\alpha, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\}.$$

que é o conjunto de todos os autovetores associados a $\lambda_1 = 1$ acrescentado o vetor nulo. Enquanto o sistema

$$(A - \lambda_2 I_3)X = \vec{0} \quad \text{é}$$

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

cuja solução geral é

$$\mathbb{W}_2 = \{(\alpha, 2\alpha, \alpha) \mid \alpha \in \mathbb{R}\},$$

que é o conjunto de todos os autovetores associados a $\lambda_2 = 3$ acrescentado o vetor nulo.

Exemplo 4.3. Vamos determinar os autovalores e autovetores do operador $T : \mathcal{P}_2 \rightarrow \mathcal{P}_2$ definido por $T(p)(x) = p(x) + (3x + 2)p'(x)$, para todo $p \in \mathcal{P}_2$. Seja $\mathcal{B} = \{1, x, x^2\}$, então

$$[T]_{\mathcal{B}}^{\mathcal{B}} = A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 4 & 4 \\ 0 & 0 & 7 \end{bmatrix}.$$

Para este operador o polinômio característico é

$$\begin{aligned} p(\lambda) &= \det(T - \lambda I) = \det(A - \lambda I_3) = \det \begin{bmatrix} 1 - \lambda & 2 & 0 \\ 0 & 4 - \lambda & 4 \\ 0 & 0 & 7 - \lambda \end{bmatrix} \\ &= (1 - \lambda)(4 - \lambda)(7 - \lambda) \end{aligned}$$

Portanto os autovalores de T são $\lambda_1 = 1, \lambda_2 = 4$ e $\lambda_3 = 7$.

Agora, vamos determinar os autovetores associados aos autovalores λ_1, λ_2 e λ_3 . Para isto vamos resolver os sistemas

$$(A - \lambda_1 I_3)X = \vec{0}, (A - \lambda_2 I_3)X = \vec{0} \text{ e } (A - \lambda_3 I_3)X = \vec{0}.$$

Como

$$(A - \lambda_1 I_3)X = \vec{0}$$

é

$$\begin{bmatrix} 0 & 2 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

então a solução geral do sistema $(A - \lambda_1 I_3)X = \vec{0}$ é $\{(\alpha, 0, 0) \mid \alpha \in \mathbb{R}\}$. Portanto, o conjunto de todos os autovetores associados a $\lambda_1 = 1$ acrescentado o vetor nulo é

$$\mathbb{W}_1 = \{p(x) = \alpha(1) \mid \alpha \in \mathbb{R}\}.$$

O sistema

$$(A - \lambda_2 I_3)X = \vec{0}$$

é

$$\begin{bmatrix} -3 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Assim a solução geral do sistema $(A - \lambda_2 I_3)X = \vec{0}$ é $\{\alpha(2, 3, 0) \mid \alpha \in \mathbb{R}\}$. Portanto, o conjunto de todos os autovetores associados a $\lambda_2 = 4$ acrescentado o vetor nulo é

$$\mathbb{W}_2 = \{p(x) = \alpha(2(1) + 3(x)) \mid \alpha \in \mathbb{R}\}.$$

O sistema

$$(A - \lambda_3 I_3)X = \vec{0}$$

é

$$\begin{bmatrix} -6 & 2 & 0 \\ 0 & -3 & 4 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Assim a solução geral do sistema $(A - \lambda_3 I_3)X = \vec{0}$ é $\{\alpha(4, 12, 9) \mid \alpha \in \mathbb{R}\}$. Portanto, o conjunto de todos os autovetores associados a $\lambda_2 = 7$ acrescentado o vetor nulo é

$$\mathbb{W}_3 = \{p(x) = \alpha(4(1) + 12(x) + 9(x^2)) \mid \alpha \in \mathbb{R}\}.$$

Nos exemplos anteriores, para cada autovalor encontramos todos os autovetores associados a ele. Para cada autovalor λ , o conjunto dos autovetores associados a ele acrescentado o vetor nulo é o núcleo de $T - \lambda I$, $\mathcal{N}(T - \lambda I)$, que é um subespaço. Este subespaço recebe o nome de **autoespaço associado ao autovalor** λ .

Já vimos que se um operador T é diagonalizável, então os vetores da base em relação a qual a matriz de T é diagonal são autovetores associados a autovalores. Portanto, estes autovetores são L.I. Vamos mostrar, a seguir, que esta é uma condição necessária e suficiente para que um operador seja diagonalizável.

Teorema 4.3. *Um operador $T : \mathbb{V} \rightarrow \mathbb{V}$ de um espaço vetorial \mathbb{V} de dimensão n é **diagonalizável** se, e somente se, ele possui n autovetores linearmente independentes. Ou seja, T é diagonalizável se, e somente se, o espaço \mathbb{V} tem uma base formada de autovetores de T .*

Demonstração. Vamos primeiro provar que se T é diagonalizável, então ele possui n autovetores L.I. Se um operador T é diagonalizável, então existe uma base $\mathcal{C} = \{V_1, \dots, V_n\}$ tal que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}.$$

O que implica que

$$T(V_j) = \sum_{i=1}^n d_{ij}V_i = \lambda_j V_j, \quad \text{para } j = 1, \dots, n.$$

Como \mathcal{C} é uma base, então os autovetores V_1, \dots, V_n são L.I.

Suponha, agora, que V_1, \dots, V_n são n autovetores linearmente independentes associados a $\lambda_1, \dots, \lambda_n$, respectivamente. Então, $\mathcal{C} = \{V_1, \dots, V_n\}$ é uma base e

$$[T]_{\mathcal{C}}^{\mathcal{C}} = D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}.$$

Ou seja, T é diagonalizável. ■

Segue do teorema anterior que se queremos diagonalizar um operador devemos procurar o maior número possível de autovetores linearmente independentes. Para cada autovalor λ , uma base do autoespaço associado a λ é um conjunto com o maior número possível de autovetores L.I. associados a λ . Vamos mostrar que ao juntarmos as bases dos autoespaços o novo conjunto formado continua linearmente independente.

Vamos antes provar um resultado sobre os autoespaços.

Proposição 4.4. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear.*

- (a) *Para todo autovalor λ , o autoespaço associado a λ , $\mathbb{W}_\lambda = \mathcal{N}(T - \lambda I)$, é T -invariante, isto é, $T(\mathbb{W}_\lambda) \subseteq \mathbb{W}_\lambda$.*
- (b) *A interseção de dois autoespaços associados a autovalores diferentes é igual ao subespaço nulo. Ou seja, se $\lambda \neq \mu$ são autovalores de T , então $\mathcal{N}(T - \lambda I) \cap \mathcal{N}(T - \mu I) = \{\vec{0}\}$*

Demonstração. (a) Se $V \in \mathbb{W}_\lambda = \mathcal{N}(T - \lambda I)$, então $T(V) = \lambda V \in \mathbb{W}_\lambda$, pois \mathbb{W}_λ é um subespaço.

- (b) Se $V \in \mathcal{N}(T - \lambda I) \cap \mathcal{N}(T - \mu I)$, então $T(V) = \lambda V = \mu V$. O que implica que $(\lambda - \mu)V = \vec{0}$. Assim, se $\lambda \neq \mu$, então $V = \vec{0}$.



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 4.5. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear de um espaço de dimensão finita \mathbb{V} . Sejam $\lambda_1, \dots, \lambda_l$ autovalores distintos de T . Se para cada $i = 1, \dots, l$, $\{V_1^{(i)}, \dots, V_{k_i}^{(i)}\}$ é um conjunto de autovetores L.I. associados a λ_i , então $\{V_1^{(1)}, \dots, V_{k_1}^{(1)}, \dots, V_1^{(l)}, \dots, V_{k_l}^{(l)}\}$ é um conjunto L.I.*

Demonstração. Vamos, em primeiro lugar, mostrar que se

$$V_1 + \dots + V_l = \bar{0}, \quad (4.7)$$

com $V_i \in \mathcal{N}(T - \lambda_i I)$, então $V_i = \bar{0}$, para $i = 1, \dots, l$. Vamos supor, por indução, que

$$V_1 + \dots + V_{l-1} = \bar{0},$$

com $V_i \in \mathcal{N}(T - \lambda_i I)$, implica que $V_i = \bar{0}$, para $i = 1, \dots, l-1$. Aplicando $(T - \lambda_l I)$ a ambos os membros da equação acima obtemos

$$(T - \lambda_l I)V_1 + \dots + (T - \lambda_l I)V_{l-1} = \bar{0}.$$

Pela **Proposição 4.4** e pela hipótese de indução temos que $V_i = \bar{0}$, para $i = 1, \dots, l-1$. Substituindo-se $V_1 = \dots = V_{l-1} = \bar{0}$ na equação (4.7) obtemos que também $V_l = 0$.

Precisamos mostrar que a única solução da equação

$$x_1^{(1)}V_1^{(1)} + \dots + x_{k_1}^{(1)}V_{k_1}^{(1)} + \dots + x_1^{(l)}V_1^{(l)} + \dots + x_{k_l}^{(l)}V_{k_l}^{(l)} = \bar{0} \quad (4.8)$$

é a solução trivial. Como $V_i = x_1^{(i)}V_1^{(i)} + \dots + x_{k_i}^{(i)}V_{k_i}^{(i)} \in \mathcal{N}(T - \lambda_i I)$, pelo que mostramos acima, a equação (4.8) implica que

$$V_i = x_1^{(i)}V_1^{(i)} + \dots + x_{k_i}^{(i)}V_{k_i}^{(i)} = \bar{0},$$

para $i = 1, \dots, l$. Como para cada $i = 1, \dots, l$, $V_1^{(i)}, \dots, V_{k_i}^{(i)}$ são L.I., então $x_1^{(i)} = \dots = x_{k_i}^{(i)} = 0$. O que prova que todos os autovetores juntos são L.I. ■

Exemplo 4.4. Considere o operador $T : \mathcal{P}_2 \rightarrow \mathcal{P}_2$ definido por
 $T(p)(x) = p(x) + (3x + 2)p'(x)$, para todo $p \in \mathcal{P}_2$.

Seja $\mathcal{B} = \{1, x, x^2\}$, então

$$[T]_{\mathcal{B}}^{\mathcal{B}} = A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 4 & 4 \\ 0 & 0 & 7 \end{bmatrix}.$$

já vimos no **Exemplo 4.3 na página 302** que seu polinômio característico é $p(\lambda) = (1 - \lambda)(4 - \lambda)(7 - \lambda)$, os seus autovalores são $\lambda_1 = 1, \lambda_2 = 4$ e $\lambda_3 = 7$ e os autoespaços correspondentes são

$$\mathbb{W}_1 = \{p(x) = \alpha(1) \mid \alpha \in \mathbb{R}\},$$

$$\mathbb{W}_2 = \{p(x) = \alpha(2(1) + 3(x)) \mid \alpha \in \mathbb{R}\} \quad \text{e}$$

$$\mathbb{W}_3 = \{p(x) = \alpha(4(1) + 12(x) + 9(x^2)) \mid \alpha \in \mathbb{R}\},$$

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

$$\{p_1(x) = 1\}$$

é uma base para \mathbb{W}_1 . Assim, não podemos ter um número maior de autovetores L.I. associados a $\lambda_1 = 1$ (**Teorema 1.10 na página 72**).

Para \mathbb{W}_2 , temos que

$$\{p_2(x) = 2 + 3x\}$$

é uma base para \mathbb{W}_2 . Assim, não podemos ter um número maior de autovetores L.I. associados a $\lambda_2 = 4$.

Para \mathbb{W}_3 , temos que

$$\{p_3(x) = 4 + 12x + 9x^2\}$$

é uma base para \mathbb{W}_3 . Assim, não podemos ter um número maior de autovetores L.I. associados a $\lambda_3 = 7$.

Como p_1 é um autovetor L.I. associados a λ_1 , p_2 é um autovetor L.I. associados a λ_2 e p_3 é um autovetor L.I. associado a λ_3 , então pela **Proposição 4.5 na página 306** os autovetores juntos p_1, p_2 e p_3 são L.I. e portanto formam uma base de \mathcal{P}_2 . Seja $\mathcal{C} = \{p_1, p_2, p_3\}$. Então,

$$D = [T]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 7 \end{bmatrix}, P = [[p_1]_{\mathcal{B}} \ [p_2]_{\mathcal{B}} \ [p_3]_{\mathcal{B}}] = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 12 \\ 0 & 0 & 9 \end{bmatrix}$$

são tais que

$$D = P^{-1}AP.$$

Exemplo 4.5. Considere o operador $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por $T(X) = AX$, para todo $X \in \mathbb{R}^3$, em que

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

Já vimos no **Exemplo 4.2 na página 300** que o seu polinômio característico é

$$p(\lambda) = \det(A - \lambda I_3) = (1 - \lambda)^2(3 - \lambda),$$

que os seus autovalores são $\lambda_1 = 1$ e $\lambda_2 = 3$ e que os autoespaços correspondentes são

$$\mathbb{W}_1 = \{(-\alpha, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$$

e

$$\mathbb{W}_2 = \{(\alpha, 2\alpha, \alpha) \mid \alpha \in \mathbb{R}\},$$

respectivamente.

Para $\lambda_1 = 1$, temos que $\{W_1 = (-1, 0, 1), W_2 = (0, 1, 0)\}$ é uma base de \mathbb{W}_1 . Assim, não podemos ter um número maior de autovetores L.I. associados a λ_1 . De forma análoga para $\lambda_2 = 3$, $\{W_3 = (1, 2, 1)\}$ é um conjunto com o maior número possível de autovetores L.I. associados a λ_2 . Assim as matrizes

$$P = [W_1 \ W_2 \ W_3] = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

são tais que

$$D = P^{-1}AP.$$

Exemplo 4.6. Considere o operador $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definido por $T(X) = AX$, para todo $X \in \mathbb{R}^2$, em que

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

O seu polinômio característico é $p(\lambda) = \det(A - \lambda I_2) = \lambda^2$, assim T possui um único autovalor: $\lambda_1 = 0$. O autoespaço correspondente a $\lambda_1 = 0$ é

$$\mathbb{W}_1 = \{(\alpha, 0) \mid \alpha \in \mathbb{R}\}.$$

Assim, para $\lambda_1 = 1$, temos que $\{V_1 = (1, 0)\}$ é uma base de \mathbb{W}_1 . Assim, não podemos ter mais autovetores L.I. associados a λ_1 e como só temos um autovalor não podemos ter dois autovetores L.I. Portanto, pelo **Teorema 4.3 na página 304**, o operador T **não** é diagonalizável, ou seja, não existem matrizes P e D tais que $D = P^{-1}AP$.

4.1.4 Subespaços Invariantes e o Teorema de Cayley-Hamilton

Definição 4.5. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear. Um subespaço \mathbb{W} de \mathbb{V} é T -invariante, se $T(\mathbb{W}) \subseteq \mathbb{W}$, ou seja, se $T(W) \in \mathbb{W}$, para todo $W \in \mathbb{W}$.

Exemplo 4.7. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear. A imagem de T , $\mathcal{I}(T)$, o núcleo de T , $\mathcal{N}(T)$ e para cada autovalor λ , o autoespaço associado a λ , $\mathbb{W}_\lambda = \mathcal{N}(T - \lambda I)$ são exemplos de subespaços T -invariantes.

Proposição 4.6. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear de um espaço de dimensão finita \mathbb{V} . Seja \mathbb{W} um subespaço T -invariante. Seja $T_{\mathbb{W}} : \mathbb{W} \rightarrow \mathbb{W}$, a **restrição de T a \mathbb{W}** , definida por $T_{\mathbb{W}}(V) = T(V)$, para todo $V \in \mathbb{W}$. Então o polinômio característico de $T_{\mathbb{W}}$ é um fator do polinômio característico de T .*

Demonstração. Sejam $\mathcal{B}_1 = \{V_1, \dots, V_k\}$ uma base de \mathbb{W} e $\mathcal{B} = \{V_1, \dots, V_k, V_{k+1}, \dots, V_n\}$ uma extensão a uma base de \mathbb{V} . Sejam $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ e $B = [T_{\mathbb{W}}]_{\mathcal{B}_1}^{\mathcal{B}_1}$. Então,

$$A = \begin{bmatrix} B & C \\ \bar{0} & D \end{bmatrix},$$

em que C e D são matrizes de tamanhos apropriados. Sejam $p_T(\lambda)$ e $p_{T_{\mathbb{W}}}(\lambda)$ os polinômios característicos de T e de $T_{\mathbb{W}}$, respectivamente. Então,

$$p_T(\lambda) = \det \begin{bmatrix} B - \lambda I_k & C \\ \bar{0} & D - \lambda I_{n-k} \end{bmatrix} = \det(B - \lambda I_k) \det(D - \lambda I_{n-k}) = p_{T_{\mathbb{W}}}(\lambda) \det(D - \lambda I_{n-k}).$$

■

A **multiplicidade de uma raiz** λ de um polinômio $p(t)$ é o maior inteiro positivo k tal que $(t - \lambda)^k$ é um fator de $p(t)$. Se o polinômio característico de um operador $T : \mathbb{V} \rightarrow \mathbb{V}$ é da forma $p_T(t) = (t - \lambda)^m r(t)$, em que $r(t)$ não tem λ como raiz, então a multiplicidade do autovalor λ no polinômio característico de T é igual a m e é chamada de **multiplicidade algébrica de λ** .

Corolário 4.7. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear de um espaço de dimensão finita \mathbb{V} . Seja λ um autovalor de T que tem multiplicidade algébrica m . Então*

$$\dim(\mathcal{N}(T - \lambda I)) \leq m,$$

ou seja, a dimensão do autoespaço associado a λ é menor ou igual a multiplicidade algébrica.

Demonstração. Sejam $\mathbb{W}_\lambda = \mathcal{N}(T - \lambda I)$ e $p = \dim(\mathbb{W}_\lambda)$. Seja \mathcal{B}_1 uma base de \mathbb{W}_λ . A matriz da restrição de T a \mathbb{W}_λ , $T_{\mathbb{W}_\lambda}$, em relação a \mathcal{B}_1 é λI_p . Assim, o polinômio característico $T_{\mathbb{W}_\lambda}$, é

$$p_{T_{\mathbb{W}_\lambda}}(t) = (\lambda - t)^p.$$

O que implica, pela **Proposição 4.6**, que o polinômio característico de T é da forma

$$p_T(t) = (\lambda - t)^p q(t),$$

em que $q(t)$ é um polinômio. Como por outro lado o polinômio característico de T é da forma

$$p_T(t) = (t - \lambda)^m r(t),$$

em que $r(t)$ não tem λ como raiz, então $\dim(\mathbb{W}_\lambda) = p \leq m$. ■

Para cada autovalor λ , a dimensão do autoespaço associado a λ , $\mathcal{N}(T - \lambda I)$, é chamada **multiplicidade geométrica de λ** . Do **Corolário 4.7** segue-se que a multiplicidade geométrica é menor ou igual a multiplicidade algébrica. O próximo resultado diz que é necessário que a multiplicidade geométrica seja igual a multiplicidade algébrica de cada autovalor de um operador para que ele seja diagonalizável.

Corolário 4.8. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear de um espaço de dimensão finita \mathbb{V} . Suponha que o polinômio característico de T seja da forma $p_T(t) = (-1)^n (t - \lambda_1)^{m_1} \dots (t - \lambda_k)^{m_k}$, com $\lambda_1, \dots, \lambda_k$ distintos. T é diagonalizável se, e somente se,*

$$\dim(\mathcal{N}(T - \lambda_i I)) = m_i, \quad \text{para } i = 1, \dots, k.$$

Ou seja, T é diagonalizável se, e somente se, a multiplicidade algébrica de λ_i é igual a sua multiplicidade geométrica, para $i = 1, \dots, k$.

Demonstração. Seja $n = \dim(\mathbb{V})$. Para $i = 1, \dots, k$, sejam $\mathbb{W}_i = \mathcal{N}(T - \lambda_i I)$ e $p_i = \dim(\mathbb{W}_i)$.

Vamos supor que T seja diagonalizável. Seja \mathcal{C} a base de \mathbb{V} consistindo de autovetores de T . Para $i = 1, \dots, k$, sejam $\mathcal{C}_i = \mathcal{C} \cap \mathbb{W}_i$ e n_i o número de vetores de \mathcal{C}_i . Como, pelo [Corolário 4.7](#), $p_i \leq m_i$, então

$$n = \sum_{i=1}^k n_i \leq \sum_{i=1}^k p_i \leq \sum_{i=1}^k m_i = n.$$

Logo, $p_i = \dim(\mathbb{W}_i) = m_i$.

Vamos supor, agora, que $p_i = \dim(\mathbb{W}_i) = m_i$, para $i = 1, \dots, k$. Seja \mathcal{C}_i uma base de \mathbb{W}_i , para $i = 1, \dots, k$. Pela [Proposição 4.5 na página 306](#), $\mathcal{C} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_k$ é um conjunto L.I. e como ele tem $m_1 + \dots + m_k = n$ elementos, então \mathcal{C} é uma base de \mathbb{V} consistindo de autovetores de T . O que prova que T é diagonalizável. ■

Se $p(t) = a_0 + a_1t + \dots + a_nt^n$ é um polinômio e T é um operador linear, definimos o operador linear

$$p(T) = a_0I + a_1T + \dots + a_nT^n.$$

Teorema 4.9 (Cayley-Hamilton). *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço de dimensão finita. Seja $p_T(t)$ o polinômio característico de T . Então, $p_T(T) = O$, a transformação linear nula.*

Demonstração. Vamos mostrar que $p_T(T)(V) = \vec{0}$, para todo $V \in \mathbb{V}$. Seja V um vetor não nulo de \mathbb{V} . Seja k o menor inteiro tal que $V, T(V), T^2(V), \dots, T^k(V)$ são L.D. Então, existem escalares a_0, \dots, a_{k-1} tais que

$$T^k(V) = a_0V + a_1T(V) + \dots + a_{k-1}T^{k-1}(V). \quad (4.9)$$

Seja $\mathbb{W} = [V, T(V), \dots, T^{k-1}(V)]$. \mathbb{W} é um subespaço T -invariante (verifique!) e se $\mathcal{B} = \{V, T(V), \dots, T^{k-1}(V)\}$, então

$$[T_{\mathbb{W}}]_{\mathcal{B}}^{\mathcal{B}} = B = \begin{bmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{k-1} \end{bmatrix}$$

e o polinômio característico de $T_{\mathbb{W}}$ é $p_{T_{\mathbb{W}}}(t) = (-1)^{k+1}(a_0 + a_1t + \dots + a_{k-1}t^{k-1} - t^k)$ (verifique, fazendo o desenvolvendo do determinante de $B - \lambda I_k$ em termos da 1ª linha).

Pela **Proposição 4.6** na página 312, existe um polinômio $q(t)$ tal que

$$p_T(t) = q(t)p_{T_{\mathbb{W}}}(t).$$

Logo,

$$p_T(T)(V) = q(T)p_{T_{\mathbb{W}}}(T)(V) = q(T)(-1)^{k+1}(a_0I + a_1T + \dots + a_{k-1}T^{k-1} - T^k)(V) = q(T)(\bar{0}) = \bar{0},$$

para $V \neq \bar{0}$, por (4.9). Como também $p_T(T)(\bar{0}) = \bar{0}$, então $p_T(T) = O$. ■

4.1.5 Aplicação: Cálculo das Potências de uma Matriz

Exemplo 4.8. Seja $A = \begin{bmatrix} 0 & -1 & 3 \\ 0 & 2 & 0 \\ 0 & -1 & 3 \end{bmatrix}$. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por $T(X) = AX$,

para todo $X \in \mathbb{R}^3$. O seu polinômio característico é $p(t) = -t(t-2)(t-3)$. Portanto, os autovalores de T são $\lambda_1 = 0, \lambda_2 = 2$ e $\lambda_3 = 3$.

$$A^2 = AA = \begin{bmatrix} 0 & -5 & 9 \\ 0 & 4 & 0 \\ 0 & -5 & 9 \end{bmatrix}.$$

Vamos calcular A^k , para $k \geq 3$. Existem polinômios $q_k(t)$ e $r_k(t)$ tais que

$$t^k = q_k(t)p(t) + r_k(t), \quad \text{em que o grau de } r_k(t) \text{ é menor que 3.} \quad (4.10)$$

Pelo Teorema de Cayley-Hamilton (Exercício 31 na página 328), $p(A) = \bar{0}$ e assim substituindo-se t por A em (4.10) obtemos

$$A^k = q_k(A)p(A) + r_k(A) = r_k(A).$$

Para determinar $r_k(t)$ vamos usar o fato de que os autovalores de T também anulam $p(t)$. Sejam $a, b, c \in \mathbb{R}$ tais que $r_k(t) = at^2 + bt + c$. Substituindo-se os autovalores de A , $t = 0, 2, 3$ em (4.10) obtemos o sistema

$$\begin{cases} 0 & = & & + & c \\ 2^k & = & 4a & + & 2b & + & c \\ 3^k & = & 9a & + & 3b & + & c \end{cases}$$

cuja solução é $a = 3^{k-1} - 2^{k-1}$, $b = 3 \cdot 2^{k-1} - 2 \cdot 3^{k-1}$ e $c = 0$. Assim, para $k \geq 3$, temos que

$$\begin{aligned} A^k &= r_k(A) = aA^2 + bA + cI_3 \\ &= (3^{k-1} - 2^{k-1})A^2 + (3 \cdot 2^{k-1} - 2 \cdot 3^{k-1})A \\ &= \begin{bmatrix} 0 & 2^k - 3^k & 3^k \\ 0 & 2^k & 0 \\ 0 & 2^k - 3^k & 3^k \end{bmatrix}. \end{aligned}$$

Exemplo 4.9. Seja $A = \begin{bmatrix} 4 & -1 & -4 \\ 1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por

$$T(X) = AX, \text{ para todo } X \in \mathbb{R}^3.$$

O seu polinômio característico é $p(t) = -t(t-3)^2$. Portanto, os autovalores de T são $\lambda_1 = 0$ e $\lambda_2 = 3$.

$$A^2 = \begin{bmatrix} 15 & -6 & -15 \\ 6 & 3 & -6 \\ 0 & 0 & 0 \end{bmatrix}.$$

Vamos calcular A^k , para $k \geq 3$. Existem polinômios $q_k(t)$ e $r_k(t)$ tais que

$$t^k = q_k(t)p(t) + r_k(t), \quad \text{em que o grau de } r_k(t) \text{ é menor que } 3. \quad (4.11)$$

Usando o fato de que os autovalores de T também anulam $p(t)$ temos apenas duas equações, pois um dos autovalores é uma raiz dupla de $p(t)$. Sejam $a, b, c \in \mathbb{R}$ tais que $r_k(t) = at^2 + bt + c$. Substituindo-se os autovalores de A , $t = 0$ e 3 em (4.11), obtemos

$$\begin{cases} 0 & = & c \\ 3^k & = & 9a + 3b + c \end{cases}$$

Agora, derivando-se (4.11) obtemos

$$kt^{k-1} = q'_k(t)p(t) + q_k(t)p'(t) + r'_k(t).$$

Usando o fato de que quando um polinômio tem uma raiz dupla ela é também raiz da sua derivada, substituindo-se $t = 3$ em (4.12) e acrescentando às equações anteriores obtemos o sistema

$$\begin{cases} 0 & = & c \\ 3^k & = & 9a + 3b + c \\ k3^{k-1} & = & 6a + b \end{cases}$$

cuja solução é $a = (k-1)3^{k-2}$, $b = (2-k)3^{k-1}$ e $c = 0$. Assim, para $k \geq 3$, pelo Teorema de Cayley-Hamilton, temos que

$$\begin{aligned} A^k &= q(A)p(A) + r_k(A) \\ &= r_k(A) = aA^2 + bA + cI_3 \\ &= (k-1)3^{k-2}A^2 + (2-k)3^{k-1}A \\ &= \begin{bmatrix} (k+3)3^{k-1} & -k3^{k-1} & -(k+3)3^{k-1} \\ k3^{k-1} & (3-k)3^{k-1} & -k3^{k-1} \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Exemplo 4.10. Seja $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por $T(X) = AX$,

para todo $X \in \mathbb{R}^3$. O seu polinômio característico é

$$p(t) = -t(t^2 + 1) = -t(t - i)(t + i).$$

Portanto, os autovalores de $T_{\mathbb{C}} : \mathbb{C}^n \rightarrow \mathbb{C}^n$, definido por $T_{\mathbb{C}}(X) = AX$, para todo $X \in \mathbb{C}^3$ são $\lambda_1 = 0, \lambda_2 = i$ e $\lambda_3 = -i$.

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

Vamos calcular A^k , para $k \geq 3$. Existem polinômios $q_k(t)$ e $r_k(t)$ tais que

$$t^k = q_k(t)p(t) + r_k(t), \quad \text{em que o grau de } r_k(t) \text{ é menor que 3.} \quad (4.12)$$

Para determinar $r_k(t)$ vamos usar o fato de que os autovalores de $T_{\mathbb{C}}$ também anulam $p(t)$. Seja $a, b, c \in \mathbb{R}$ tais que $r_k(t) = at^2 + bt + c$. Substituindo-se $t = 0, i, -i$ em (4.12) obtemos o sistema

$$\begin{cases} 0 & = & & + & c \\ i^k & = & -a & + & ib & + & c \\ (-i)^k & = & -a & - & ib & + & c \end{cases}$$

cuja solução é $a = -\cos(\frac{k\pi}{2}), b = \sin(\frac{k\pi}{2})$ e $c = 0$. Assim, para $k \geq 3$, pelo Teorema de Cayley-Hamilton, temos que

$$\begin{aligned} A^k &= q(A)p(A) - \cos\left(\frac{k\pi}{2}\right)A^2 + \sin\left(\frac{k\pi}{2}\right)A \\ &= -\cos\left(\frac{k\pi}{2}\right)A^2 + \sin\left(\frac{k\pi}{2}\right)A \\ &= \begin{bmatrix} 0 & 0 & 0 \\ \sin\left(\frac{k\pi}{2}\right) & \cos\left(\frac{k\pi}{2}\right) & -\sin\left(\frac{k\pi}{2}\right) \\ -\cos\left(\frac{k\pi}{2}\right) & \sin\left(\frac{k\pi}{2}\right) & \cos\left(\frac{k\pi}{2}\right) \end{bmatrix}. \end{aligned}$$

Exercícios Numéricos (respostas na página 455)

4.1.1. Ache o polinômio característico, os autovalores e os autovetores do operador $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, definido por $T(X) = AX$, para todo $X \in \mathbb{R}^n$, para cada matriz:

$$(a) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

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

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

$$(b) \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$$

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

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

4.1.2. Ache bases para os auto-espacos associados a cada autovalor do operador $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, definido por $T(X) = AX$, para todo $X \in \mathbb{R}^n$, para cada matriz:

$$(a) \begin{bmatrix} 2 & 0 & 0 \\ 3 & -1 & 0 \\ 0 & 4 & 3 \end{bmatrix}$$

$$(c) \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

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

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

4.1.3. Verifique quais das matrizes são diagonalizáveis:

$$(a) \begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix}$$

$$(c) \begin{bmatrix} 1 & 1 & -2 \\ 4 & 0 & 4 \\ 1 & -1 & 4 \end{bmatrix}$$

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

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

4.1.4. Ache para cada matriz A , se possível, uma matriz não-singular P tal que $P^{-1}AP$ seja diagonal:

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

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

$$(b) \begin{bmatrix} 4 & 2 & 3 \\ 2 & 1 & 2 \\ -1 & -2 & 0 \end{bmatrix}$$

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

4.1.5. Determine para cada matriz A as potências A^k , para $k = 1, 2, 3, \dots$

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

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

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

$$(d) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}$$

Exercícios usando o MATLAB[®]

>> `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 a_{11} , a_{12} , \dots , a_{mn} e a armazena numa variável A ;

>> `A=[A1,...,An]` cria uma matriz A formada pelas matrizes, definidas anteriormente, A_1 , \dots , A_n colocadas uma ao lado da outra;

>> `solve(expr)` determina a solução da equação $\text{expr}=0$. 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 .

- 4.1.6.** Defina as matrizes $B=\text{sym}(\text{randi}(2))$ e $A=[B-B', \text{zeros}(2,1); \text{zeros}(1,2), \text{randi}]$. A matriz A é diagonalizável? Por que?
- 4.1.7.** Defina as matrizes $L=[\text{eye}(2), \text{zeros}(2,1); \text{randi}(1,2), 0]$ e $A=\text{sym}(L*L')$. 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 $\text{inv}(P)*A*P=D$, se possível. Verifique o resultado. Use o comando $[P,D]=\text{eig}(A)$ e compare com as matrizes que você encontrou.
- 4.1.8.** Defina $a=\text{randi}, b=\text{randi}$ e $A=\text{sym}([2*a, a-b, a-b; 0, a+b, b-a; 0, b-a, a+b])$. 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 $\text{inv}(P)*A*P=D$, se possível. Verifique o resultado. Use o comando $[P,D]=\text{eig}(A)$ e compare com as matrizes que você encontrou.
- 4.1.9.** Defina $a=\text{randi}, b=\text{randi}$ e $A=\text{sym}([a, 0, b; 2*b, a-b, 2*b; b, 0, a])$. 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 $\text{inv}(P)*A*P=D$, se possível. Verifique o resultado. Use o comando $[P,D]=\text{eig}(A)$ e compare com as matrizes que você encontrou.
- 4.1.10.** Defina $a=\text{randi}$ (se obtiver $a=0$, repita até que a seja diferente de zero) e $b=\text{randi}$. Determinar os autovalores e autovetores do operador $T: \mathcal{P}_2 \rightarrow \mathcal{P}_2$ definido por $T(p)(x) = p(x) + (ax + b)p'(x)$, para todo $p \in \mathcal{P}_2$. Determine, se possível, uma base \mathcal{C} de \mathcal{P}_2 tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é uma matriz diagonal.
- 4.1.11.** Repita o exercício anterior, mas com $a=0$.

4.1.12. Use o MATLAB[®] para resolver os Exercícios Numéricos

Exercícios Teóricos

4.1.13. 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 .

4.1.14. Seja λ um autovalor de T com autovetor associado V . Demonstre que λ^k é um autovalor de $T^k = T \dots T$ associado a V , em que k é um inteiro positivo.

4.1.15. Um operador T é chamado **nilpotente** se $T^k = O$, a transformação linear nula, para algum inteiro positivo k . Demonstre que se T é nilpotente, então o único autovalor de T é 0. (Sugestão: use o exercício anterior)

4.1.16. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial de dimensão finita \mathbb{V} .

- (a) Mostre que o determinante de T é o produto de todas as raízes do seu polinômio característico, incluindo as multiplicidades; (Sugestão: $\det(T - tI) = p_T(t) = (-1)^n(t - \lambda_1) \dots (t - \lambda_n)$)
- (b) Mostre que T é singular se, e somente se, 0 for um autovalor de T .

4.1.17. Seja λ um autovalor de um operador invertível T com autovetor associado V . Mostre que $1/\lambda$ é um autovalor de T^{-1} com autovetor associado V .

4.1.18. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear. Mostre que se \mathbb{W} é um subespaço T -invariante, então \mathbb{W} é

- (a) T^k -invariante, para todo inteiro positivo k e
- (b) $p(T)$ -invariante, para todo polinômio $p(t)$.

4.1.19. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear. Sejam $\lambda_1, \dots, \lambda_k$ seus autovalores distintos. Seja $\mathbb{W}_i = \mathcal{N}(T - \lambda_i I)$, para $i = 1, \dots, k$.

(a) Mostre que T é diagonalizável se, e somente se,

$$\mathbb{V} = \mathbb{W}_1 \oplus \cdots \oplus \mathbb{W}_k.$$

(Sugestão: $V_1 + \dots + V_k = \bar{0}$, com $V_i \in \mathcal{N}(T - \lambda_i I)$, para $i = 1, \dots, k$ implica que $V_1 = \dots = V_k = \bar{0}$, pois caso contrário seria uma contradição com a [Proposição 4.5 na página 306](#))

(b) Suponha que T seja diagonalizável. Para $i = 1, \dots, k$, seja $P_i : \mathbb{V} \rightarrow \mathbb{V}$ definida da seguinte forma. Seja $V \in \mathbb{V}$. V se escreve de maneira única como $V = V_1 + \dots + V_k$, com $V_1 \in \mathbb{W}_1, \dots, V_k \in \mathbb{W}_k$. Então, $P_i(V) = V_i$. Mostre que

i. $P_i^2 = P_i$, para $i = 1, \dots, k$.

ii. $P_1 + \dots + P_k = I$.

iii. $P_i P_j = O$, a transformação linear nula, se $i \neq j$.

iv. $T = \lambda_1 P_1 + \dots + \lambda_k P_k$.

(Sugestão: use o [Exercício 3.3.19 na página 268](#).)

4.1.20. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear. Um polinômio com o coeficiente do termo de mais alto grau igual a 1 é chamado **polinômio mônico**. O polinômio mônico de menor grau tal que $p(T) = O$, a transformação linear nula, é chamado **polinômio mínimo** de T . Mostre que se T é diagonalizável e $\lambda_1, \dots, \lambda_k$ são os seus autovalores distintos, então o polinômio mínimo de T é $m(t) = (t - \lambda_1) \dots (t - \lambda_k)$. (Sugestão: Se $V_i \in \mathcal{N}(T - \lambda_i I)$, então $m(T)(V_i) = m(\lambda_i)V_i = \bar{0}$. Escreva $V = V_1 + \dots + V_k$, com $V_i \in \mathcal{N}(T - \lambda_i I)$ e $m(T) = (T - \lambda_1 I) \dots (T - \lambda_k I)$.)

4.1.21. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial de dimensão n . Mostre que

(a) Para $k \geq n$, $T^k = r_k(T)$, em que $r_k(t)$ é um polinômio de grau menor que n , que depende de k . (Sugestão: use o fato de que existem polinômios $q(t)$ e $r_k(t)$ tais que $t^k = q(t)p_T(t) + r_k(t)$ e o [Teorema de Cayley-Hamilton na página 315](#).)

(b) Se T é diagonalizável e seu polinômio característico tem m raízes distintas, então para $k \geq m$, $T^k = r_k(T)$, em que $r_k(t)$ é um polinômio de grau menor que m , que depende de k . (Sugestão:

use o fato de que existem polinômios $q(t)$ e $r_k(t)$ tais que $t^k = q(t)m_T(t) + r_k(t)$, em que $m_T(t)$ é o polinômio mínimo e o [Exercício 4.1.20](#).)

(c) A dimensão do espaço gerado por $I, T, T^2, \dots, T^k, \dots$ é menor ou igual a n .

4.1.22. Seja $p(t) = (-1)^n(t^n + a_{n-1}t^{n-1} + \dots + a_0)$ o polinômio característico do operador linear $T : \mathbb{V} \rightarrow \mathbb{V}$. Mostre que

(a) T é invertível se, e somente se, $a_0 \neq 0$.

(b) Se T é invertível, então

$$T^{-1} = \frac{-1}{a_0} [T^{n-1} + a_{n-1}T^{n-2} + \dots + a_1I].$$

(Sugestão: use o [Teorema de Cayley-Hamilton na página 315](#).)

4.1.23. Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ um operador linear definido por $T(X) = AX$, para todo $X \in \mathbb{R}^n$, em que

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{n-1} \end{bmatrix},$$

para $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$. Mostre que o polinômio característico de T é

$$p_T^{(n)}(t) = (-1)^{n+1}(a_0 + a_1t + \dots + a_{n-1}t^{n-1} - t^n).$$

(Sugestão: faça o desenvolvendo do determinante de $A - \lambda I_n$ em termos da 1ª linha)

4.1.24. Seja $p(t) = a_0 + a_1t + \dots + a_{n-1}t^{n-1} + t^n$ um polinômio mônico de grau n . Mostre que a matriz $n \times n$

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}$$

é tal que o polinômio característico do operador linear $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definido por $T(X) = AX$, para todo $X \in \mathbb{R}^n$ é $p_T(t) = (-1)^n p(t)$. Esta matriz é chamada **matriz companheira do polinômio** $p(t)$.

4.1.25. Seja \mathbb{V} um espaço vetorial real. Seja $\mathbb{V}_{\mathbb{C}}$ a **complexificação de \mathbb{V}** , ou seja, é o conjunto das expressões $V + iW$, em que $V, W \in \mathbb{V}$ e $i = \sqrt{-1}$. Em $\mathbb{V}_{\mathbb{C}}$, $V + iW = V' + iW'$ significa $V = V'$ e $W = W'$. A soma em $\mathbb{V}_{\mathbb{C}}$ é definida por $(V_1 + iW_1) + (V_2 + iW_2) = (V_1 + V_2) + i(W_1 + W_2)$. A multiplicação por escalar complexo é definido por $(\alpha + i\beta)(V + iW) = (\alpha V - \beta W) + i(\alpha W + \beta V)$. Para todo operador linear $T : \mathbb{V} \rightarrow \mathbb{V}$, define-se a **complexificação de T** , $T_{\mathbb{C}} : \mathbb{V}_{\mathbb{C}} \rightarrow \mathbb{V}_{\mathbb{C}}$ por $T_{\mathbb{C}}(V + iW) = T(V) + iT(W)$, para todos $V, W \in \mathbb{V}$. Mostre que

- (a) Toda base de \mathbb{V} é uma base de $\mathbb{V}_{\mathbb{C}}$.
- (b) Se $\mathcal{B} = \{V_1, \dots, V_n\}$ é uma base de \mathbb{V} e $T : \mathbb{V} \rightarrow \mathbb{V}$ é um operador linear de \mathbb{V} , então $[T]_{\mathcal{B}}^{\mathcal{B}} = [T_{\mathbb{C}}]_{\mathcal{B}}^{\mathcal{B}}$. Portanto, os polinômios característicos de T e de $T_{\mathbb{C}}$ são iguais.
- (c) Se $V + iW$ é um autovetor de $T_{\mathbb{C}}$ associado a $\lambda = \alpha + i\beta$, com $\beta \neq 0$, então $\mathbb{W} = [V, W]$ é um subespaço T -invariante.
- (d) Se $V + iW$ é um autovetor de $T_{\mathbb{C}}$ associado a $\lambda = \alpha + i\beta$, com $\beta \neq 0$, então $V - iW$ é também um autovetor de $T_{\mathbb{C}}$ associado a $\bar{\lambda} = \alpha - i\beta$ e que $\mathcal{B}_{\lambda} = \{W, V\}$ é uma base (ordenada) de um subespaço \mathbb{W} de \mathbb{V} , T -invariante, tal que a matriz da restrição $T_{\mathbb{W}}$ de T a \mathbb{W} é

$$[T_{\mathbb{W}}]_{\mathcal{B}_{\lambda}}^{\mathcal{B}_{\lambda}} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}.$$

4.1.26. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial real de dimensão finita. Então, existe um subespaço \mathbb{W} de \mathbb{V} , T -invariante, com $1 \leq \dim(\mathbb{W}) \leq 2$.

4.1.27. Um operador linear $T : \mathbb{V} \rightarrow \mathbb{V}$ em um espaço vetorial real de dimensão n é chamado **semi-simples** se a complexificação $T_{\mathbb{C}} : \mathbb{V}_{\mathbb{C}} \rightarrow \mathbb{V}_{\mathbb{C}}$ é diagonalizável. Mostre que se T é um operador semi-simples, então

existe uma base \mathcal{C} de \mathbb{V} tal que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & D_k \end{bmatrix}, \quad \text{em que } D_i = [\lambda_i] \text{ ou } D_i = \begin{bmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{bmatrix}.$$

4.1.28. Considere o sistema de equações diferenciais lineares

$$\begin{cases} \frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \vdots \\ \frac{dx_n}{dt} = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{cases}$$

Ele pode ser escrito na forma de uma equação diferencial vetorial $\frac{dX}{dt} = AX$, em que $A = (a_{ij})_{n \times n}$. Suponha que a matriz A seja diagonalizável, ou seja, suponha que exista uma matriz P tal que

$$P^{-1}AP = D,$$

em que $D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}$. Mostre que a solução do sistema é

$$X(t) = P \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & e^{\lambda_n t} \end{bmatrix} P^{-1}X(0), \quad \text{para } t \geq 0.$$

- 4.1.29.** Sejam $T, S : \mathbb{V} \rightarrow \mathbb{V}$ com \mathbb{V} de dimensão finita. Mostre que TS e ST 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 TS , então SV é autovetor de ST .)
- 4.1.30.** Seja $T : \mathbb{V} \rightarrow \mathbb{V}$. Suponha que \mathbb{V} tenha dimensão finita. Mostre que o traço de T é igual a soma das raízes do seu polinômio característico, incluindo as multiplicidades. (Sugestão: use o fato de que $\text{tr}(AB) = \text{tr}(BA)$.)
- 4.1.31.** (Teorema de Cayley-Hamilton para matrizes) Se $p(t) = a_0 + a_1t + \dots + a_nt^n$ é um polinômio e A é uma matriz $n \times n$, definimos a matriz

$$p(A) = a_0I_n + a_1A + \dots + a_nA^n.$$

Seja $p(t)$ o polinômio característico da matriz A , então $p(A) = \bar{0}$.

4.2 Operadores Auto-adjuntos e Normais

Vamos supor que para um operador linear $T : \mathbb{V} \rightarrow \mathbb{V}$ em um espaço vetorial com produto interno de dimensão finita, exista uma base ortonormal \mathcal{B} tal que $[T]_{\mathcal{B}}^{\mathcal{B}}$ seja diagonal. Então, $[T^*]_{\mathcal{B}}^{\mathcal{B}}$ também é diagonal. Assim, $TT^* = T^*T$, pois matrizes diagonais comutam. Isto motiva a seguinte definição.

Definição 4.6. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial de dimensão finita com produto interno. Dizemos que T é **operador normal** se existe T^* e $TT^* = T^*T$.

São válidas as seguintes propriedades para um operador normal.

Proposição 4.10. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador normal em um espaço vetorial com produto interno. São válidas as seguintes propriedades:*

- (a) $\|T(V)\| = \|T^*(V)\|$, para todo vetor $V \in \mathbb{V}$.
- (b) Se V é tal que $T(V) = \lambda V$, então $T^*(V) = \bar{\lambda}V$.
- (c) Os autovetores de T associados a autovalores diferentes são ortogonais.

Demonstração. (a) Seja $V \in \mathbb{V}$.

$$\|T(V)\|^2 = \langle T(V), T(V) \rangle = \langle V, T^*T(V) \rangle = \langle V, TT^*(V) \rangle = \langle T^*(V), T^*(V) \rangle = \|T^*(V)\|^2.$$

(b) Como T é normal, então $T - \lambda I$ também é normal (verifique!). Assim, pelo item anterior

$$\|T^*(V) - \bar{\lambda}V\| = \|(T - \lambda I)^*(V)\| = \|(T - \lambda I)(V)\| = \|T(V) - \lambda V\| = 0.$$

Logo, $T^*(V) = \bar{\lambda}V$.

(c) Sejam V_1 e V_2 autovetores de T associados aos autovalores λ_1 e λ_2 , respectivamente, com $\lambda_1 \neq \lambda_2$. Então, pelo item anterior, temos que

$$\lambda_1 \langle V_1, V_2 \rangle = \langle \lambda_1 V_1, V_2 \rangle = \langle T(V_1), V_2 \rangle = \langle V_1, T^*(V_2) \rangle = \langle V_1, \bar{\lambda}_2 V_2 \rangle = \lambda_2 \langle V_1, V_2 \rangle.$$

Assim,

$$(\lambda_1 - \lambda_2) \langle V_1, V_2 \rangle = 0.$$

Como, $\lambda_1 \neq \lambda_2$, então $\langle V_1, V_2 \rangle = 0$.

■

Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador normal. Para encontrarmos uma base ortonormal \mathcal{B} tal que $[T]_{\mathcal{B}}^{\mathcal{B}}$ seja diagonal, precisamos encontrar, para cada autovalor, autovetores ortonormais associados a eles, já que autovetores associados a autovalores diferentes já são ortogonais. 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 4.11. Considere o operador $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por $T(X) = AX$, para todo $X \in \mathbb{R}^3$, em que

$$A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}.$$

Este operador é normal, pois $T^* = T$. O seu polinômio característico é

$$p(\lambda) = \det(A - \lambda I_3) = (\lambda - 2)^2(8 - \lambda)$$

Portanto os autovalores de T são $\lambda_1 = 2$ e $\lambda_2 = 8$.

Os autovetores associados aos autovalores $\lambda_1 = 2$ e $\lambda_2 = 8$ são as soluções de $(A - \lambda_1 I_3)X = \vec{0}$ e $(A - \lambda_2 I_3)X = \vec{0}$ 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 usar o processo de ortogonalização de Gram-Schmidt aos vetores V_1 e V_2 .

$$W_1 = V_1 = (-1, 0, 1); \quad W_2 = V_2 - \text{proj}_{W_1} V_2 = (-1/2, 1, -1/2)$$

$$U_1 = \left(\frac{1}{\|W_1\|} \right) W_1 = (-1/\sqrt{2}, 0, 1/\sqrt{2})$$

$$U_2 = \left(\frac{1}{\|W_2\|} \right) W_2 = (-1/\sqrt{6}, 2/\sqrt{6}, -1/\sqrt{6})$$

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

$$U_3 = \left(\frac{1}{\|V_3\|} \right) V_3 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$$

forma uma base ortonormal para \mathbb{W}_2 .

Como o operador T é normal, autovetores associados a autovalores diferentes são ortogonais. Portanto, U_1 , U_2 e U_3 são ortonormais e assim a base $\mathcal{C} = \{U_1, U_2, U_3\}$ é tal que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

Vamos mostrar que usando o procedimento usado no exemplo anterior, para operadores normais, cujas raízes do polinômio característico pertencem ao conjunto de escalares, sempre encontramos uma base ortonormal em que a matriz do operador é diagonal. Para isso, precisamos provar o seguinte resultado.

Proposição 4.11. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial com produto interno. Se \mathbb{W} é um subespaço T -invariante, então \mathbb{W}^\perp é T^* -invariante.*

Demonstração. Seja $V \in \mathbb{W}^\perp$. Vamos mostrar que $T^*(V)$ também pertence a \mathbb{W}^\perp . Seja $W \in \mathbb{W}$ qualquer. Temos que

$$\langle W, T^*(V) \rangle = \langle T(W), V \rangle = 0,$$

pois, como \mathbb{W} é T -invariante, então $T(W) \in \mathbb{W}$. Logo, $T^*(V) \in \mathbb{W}^\perp$ e portanto \mathbb{W}^\perp é T^* -invariante. ■

Teorema 4.12. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial com produto interno de dimensão finita tal que as raízes do seu polinômio característico pertencem ao conjunto de escalares. Existe uma base ortonormal de \mathbb{V} , \mathcal{C} , tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é diagonal se, e somente se, o operador T é normal.*

Demonstração. Vamos supor que T seja normal. Seja $n = \dim(\mathbb{V})$. O resultado é claramente verdadeiro para $n = 1$.

Suponha que o resultado seja verdadeiro para operadores lineares em espaços de dimensão $n - 1$, cujas raízes do polinômio característico estejam no conjunto de escalares.

Seja λ_1 uma raiz do polinômio característico de T , $p_T(t)$. Seja V_1 um autovetor unitário de T associado a λ_1 . Seja $\mathbb{W}_1 = [V_1]$. \mathbb{W}_1 é claramente T -invariante. Pela **Proposição 4.10 na página 329**, V_1 é também autovetor de T^* . Logo, \mathbb{W}_1 é também T^* -invariante. Como $T^{**} = T$, então pela **Proposição 4.11**, $\mathbb{W}_2 = \mathbb{W}_1^\perp$ é T -invariante e T^* -invariante.

Seja $T_{\mathbb{W}_2}$ a restrição de T a \mathbb{W}_2 . A dimensão de \mathbb{W}_2 é $n - 1$, pois pela **Proposição 2.11 na página 148** $\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_1^\perp$. O polinômio característico de $T_{\mathbb{W}_2}$, pela **Proposição 4.6 na página 312** é um fator de $p_T(t)$.

Como $(T_{\mathbb{W}_2})^*$ é igual a restrição de T^* a \mathbb{W}_2 , então $T_{\mathbb{W}_2}$ é um operador normal em um espaço de dimensão $n - 1$, cujas raízes do seu polinômio característico estão no conjunto de escalares. Como estamos assumindo que o resultado é verdadeiro neste caso, então existe uma base ortonormal de \mathbb{W}_2 , \mathcal{C}_2 , tal que $[T_{\mathbb{W}_2}]_{\mathcal{C}_2}^{\mathcal{C}_2}$ é diagonal. Claramente, $\mathcal{C} = \{V_1\} \cup \mathcal{C}_2$ é uma base ortonormal de \mathbb{V} tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é diagonal.

Por outro lado, se existe uma base ortonormal, tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é diagonal. Então, $[T^*]_{\mathcal{C}}^{\mathcal{C}}$ também é diagonal. Assim, $TT^* = T^*T$, pois matrizes diagonais comutam. ■

Se o espaço vetorial for complexo, isto é, se o conjunto de escalares é o conjunto dos números complexos, então todas as raízes de um polinômio pertencem ao conjunto de escalares e portanto todo operador normal é diagonalizável. Entretanto, se o espaço vetorial é real, isto é, se o conjunto de escalares é o conjunto de números reais,

então pode acontecer do polinômio característico de um operador normal ter raízes complexas que portanto não pertencem ao conjunto de escalares. Neste caso ele não será diagonalizável. Se o operador for tal que $T^* = T$, então isto não acontece.

Definição 4.7. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial com produto interno. Dizemos que T é **operador auto-adjunto (ou hermitiano)** se, $T^* = T$.

Proposição 4.13. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador auto-adjunto em um espaço vetorial com produto interno de dimensão finita. São válidas as seguintes propriedades:*

- (a) *Se o espaço vetorial \mathbb{V} é complexo, então os autovalores de T são reais.*
 - (b) *As raízes do polinômio característico de T são reais.*
 - (c) *Os autovetores de T associados a autovalores diferentes são ortogonais.*
-

Demonstração.

- (a) Seja $V \in \mathbb{V}$, $V \neq \bar{0}$ tal que $T(V) = \lambda V$. Então, pela **Proposição 4.10(b) na página 329**, temos que

$$\lambda V = T(V) = T^*(V) = \bar{\lambda}V.$$

Logo, $\lambda = \bar{\lambda}$, ou seja, λ é real.

- (b) Sejam \mathcal{B} uma base ortonormal de \mathbb{V} e $A = [T]_{\mathcal{B}}^{\mathcal{B}}$. Defina o operador $S : \mathbb{C}^n \rightarrow \mathbb{C}^n$ por $S(X) = AX$, para todo $X \in \mathbb{C}^n$. O operador S é auto-adjunto e tem o mesmo polinômio característico de T . Assim, as raízes do polinômio característico de T são os autovalores de S , que são reais.
- (c) Este item é um caso particular da [Proposição 4.10\(c\) na página 329](#).



Teorema 4.14. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial **real** com produto interno de dimensão finita. Existe uma base ortonormal de \mathbb{V} , \mathcal{C} , tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é diagonal se, e somente se, o operador T é auto-adjunto.*

Demonstração. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador auto-adjunto em um espaço vetorial real. Como pela [Proposição 4.13](#) as raízes do polinômio característico de T são reais, então elas são autovalores de T . Segue, então, do [Teorema 4.12 na página 333](#), que existe uma base ortonormal de \mathbb{V} , \mathcal{C} , tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é diagonal.

Por outro lado, suponha que um operador linear $T : \mathbb{V} \rightarrow \mathbb{V}$ em um espaço vetorial real é tal que existe uma base ortonormal de \mathbb{V} , \mathcal{C} , tal que $D = [T]_{\mathcal{C}}^{\mathcal{C}}$ é diagonal. Então, os elementos da diagonal de D são autovalores de T , que são reais e portanto $[T^*]_{\mathcal{C}}^{\mathcal{C}} = [T]_{\mathcal{C}}^{\mathcal{C}}$, ou seja, T é auto-adjunto.



Segue do [Teorema 4.14](#), que todo operador auto-adjunto em um espaço vetorial real é diagonalizável. Como consequência, toda matriz simétrica com entradas reais é diagonalizável. Mais ainda.

Teorema 4.15. Uma matriz A , $n \times n$ com entradas reais é simétrica se, e somente se, existe uma matriz P ortogonal (isto é, $P^{-1} = P^t$) e uma matriz diagonal D tal que

$$D = P^t A P.$$

Assim, se A é simétrica com entradas reais, então ela é diagonalizável.

Demonstração. Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ um operador definido por $T(X) = AX$, para todo $X \in \mathbb{R}^n$. Seja $\mathcal{B} = \{E_1, \dots, E_n\}$ a base canônica de \mathbb{R}^n . Pelo Teorema 4.14 existe uma base ortonormal \mathcal{C} de \mathbb{R}^n tal que $[T]_{\mathcal{C}}^{\mathcal{C}} = D$ é diagonal. Seja P a matriz mudança de base de \mathcal{B} para \mathcal{C} . Pelo Exercício 4.1.13c na página 287 a matriz P é ortogonal, ou seja, satisfaz $P^{-1} = P^t$. Assim, pelo Corolário 3.20 na página 256, temos que

$$D = [T]_{\mathcal{C}}^{\mathcal{C}} = P^{-1} A P = P^t A P.$$

Por outro lado se existem matrizes P simétrica e D diagonal tais que $D = P^t A P$, então $A = P D P^t$ e $A^t = (P D P^t)^t = P D^t P^t = P D P^t = A$, ou seja, A é simétrica. ■

Exemplo 4.12. Considere a matriz

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

Esta matriz é a matriz do operador do Exemplo 4.11 na página 331. Assim, a matriz

$$P = [U_1 U_2 U_3] = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

satisfaz $D = P^tAP$, em que

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}.$$

Já vimos que mesmo se o polinômio característico de um operador tenha suas raízes no conjunto de escalares, se ele não for normal, não existe uma base ortonormal, \mathcal{C} , tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ seja diagonal. Entretanto, sempre existe uma base \mathcal{C} , tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ seja triangular superior.

Teorema 4.16 (Schur). *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador em um espaço vetorial com produto interno de dimensão finita. Suponha que as raízes do polinômio característico de T estejam no conjunto de escalares. Então, existe uma base ortonormal \mathcal{C} de \mathbb{V} tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é uma matriz triangular superior.*

Demonstração. Seja $n = \dim(\mathbb{V})$. O resultado é claramente verdadeiro para $n = 1$.

Suponha que o resultado seja verdadeiro para operadores lineares em espaços de dimensão $n - 1$, cujas raízes do polinômio característico estejam no conjunto de escalares.

Seja λ_1 uma raiz do polinômio característico de T , $p_T(t)$. Então, $\bar{\lambda}_1$ é raiz do polinômio característico de T^* (verifique!) e autovalor de T^* , pela [Proposição 4.2\(b\) na página 299](#), pois se λ_1 é real, então $\bar{\lambda}_1 = \lambda_1$ e se λ_1 não é real, então o conjunto de escalares é \mathbb{C} e $\bar{\lambda}_1$ pertence ao conjunto de escalares. Seja V_1 um autovetor unitário

de T^* associado a $\bar{\lambda}_1$. Seja $\mathbb{W}_1 = [V_1]$. Como \mathbb{W}_1 é T^* -invariante e $T^{**} = T$, então pela [Proposição 4.11 na página 333](#), $\mathbb{W}_2 = \mathbb{W}^\perp$ é T -invariante. Seja $T_{\mathbb{W}_2}$ a restrição de T a $\mathbb{W}_2 = \mathbb{W}^\perp$. A dimensão de \mathbb{W}_2 é $n - 1$ (pois pela [Proposição 2.11 na página 148](#) $\mathbb{V} = \mathbb{W} \oplus \mathbb{W}^\perp$) e o seu polinômio característico, pela [Proposição 4.6 na página 312](#), é um fator de $p_T(t)$.

Assim, $T_{\mathbb{W}_2}$ é um operador linear em um espaço de dimensão $n - 1$, cujas raízes do seu polinômio característico estão no conjunto de escalares. Como estamos assumindo que o resultado é verdadeiro neste caso, então existe uma base ortonormal de \mathbb{W}_2 , \mathcal{C}_2 , tal que $[T_{\mathbb{W}_2}]_{\mathcal{C}_2}^{\mathcal{C}_2}$ é triangular superior. Claramente, $\mathcal{C} = \mathcal{C}_2 \cup \{V_1\}$ é uma base ortonormal de \mathbb{V} tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ é uma matriz triangular superior. ■

Exercícios Numéricos (respostas na página 468)

4.2.1. Para cada matriz dada A , ache uma matriz ortogonal P tal que $P^t A P$ seja diagonal:

$$(a) \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

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

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

$$(g) \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{bmatrix}$$

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

$$(d) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{bmatrix}$$

$$(f) \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

$$(h) \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Exercícios Teóricos

4.2.2. (a) Mostre que $X = (x, y)$ é ortogonal a $V = (a, b) \neq \vec{0}$ com $\|X\| = \|V\|$ se, e somente se, $X = (-b, a)$ ou $X = (b, -a)$.

(b) 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.)

4.2.3. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador normal em um espaço com produto interno. Mostre que

(a) $\mathcal{N}(T) = \mathcal{N}(T^*)$. (Sugestão: use a [Proposição 4.10 na página 329](#).)

(b) Se \mathbb{V} é de dimensão finita, então $\mathcal{I}(T) = \mathcal{I}(T^*)$. (Sugestão: use o [Teorema 3.31 na página 277](#).)

4.2.4. (Teorema Espectral) Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial de dimensão finita com produto interno. Suponha que T seja auto-adjunto se \mathbb{V} for um espaço vetorial real e que T seja normal se \mathbb{V} for um espaço vetorial complexo. Sejam $\lambda_1, \dots, \lambda_k$ autovalores distintos de T . Sejam $\mathbb{W}_i = \mathcal{N}(T - \lambda_i I)$ e $P_i : \mathbb{V} \rightarrow \mathbb{V}$ a projeção ortogonal em \mathbb{W}_i , para $i = 1, \dots, k$. Mostre que

(a) $\mathbb{V} = \mathbb{W}_1 \oplus \dots \oplus \mathbb{W}_k$.

(b) $\mathbb{W}_i^\perp = \mathbb{W}_1 + \dots + \mathbb{W}_{i-1} + \mathbb{W}_{i+1} + \dots + \mathbb{W}_k$.

(c) $P_i P_j = O$, a transformação linear nula, se $i \neq j$.

(d) $I = P_1 + \dots + P_k$, chamada **resolução da identidade induzida por T** .

(e) $T = \lambda_1 P_1 + \dots + \lambda_k P_k$, chamada **decomposição espectral do operador T** .

(Sugestão: use o [Exercício 4.1.19 na página 323](#).)

4.2.5. Seja \mathbb{V} um espaço vetorial com produto interno de dimensão finita. Seja \mathbb{W} um subespaço de \mathbb{V} . Mostre que a projeção ortogonal em \mathbb{W} , $P : \mathbb{V} \rightarrow \mathbb{V}$, $P(V) = \text{proj}_{\mathbb{W}} V$, é um operador auto-adjunto. (Sugestão: mostre que $\langle P(V), W \rangle = \langle V, P(W) \rangle$, decompondo $V = V_1 + V_2$, $W = W_1 + W_2$, com $V_1, W_1 \in \mathbb{W}$ e $V_2, W_2 \in \mathbb{W}^\perp$.)

4.2.6. Seja \mathbb{V} um espaço vetorial complexo. Mostre que um operador $T : \mathbb{V} \rightarrow \mathbb{V}$ é normal se, e somente se, existe um polinômio $p(t)$ tal que $T^* = p(T)$. (Sugestão: Mostre que se $T = \lambda_1 P_1 + \dots + \lambda_k P_k$ é a decomposição espectral de T e $p(t)$ é um polinômio, então $p(T) = p(\lambda_1) P_1 + \dots + p(\lambda_k) P_k$. Use o fato de que existe um polinômio $p(t)$ tal que $p(\lambda_i) = \bar{\lambda}_i$, para $i = 1, \dots, k$.)

4.2.7. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador normal em um espaço vetorial complexo. Seja \mathbb{W} um subespaço de \mathbb{V} . Mostre que

(a) Se \mathbb{W} é T -invariante, então \mathbb{W} é também T^* -invariante. (Sugestão: use o exercício anterior e o [Exercício 4.1.18 na página 323](#).)

(b) Se \mathbb{W} é T -invariante, então \mathbb{W}^\perp também é T -invariante.

4.2.8. Dizemos que um operador auto-adjunto $T : \mathbb{V} \rightarrow \mathbb{V}$, em um espaço com produto interno de dimensão finita é **(definido) positivo** se $\langle T(V), V \rangle > 0$, para todo $V \in \mathbb{V}, V \neq \vec{0}$. Mostre que são equivalentes as seguintes afirmações:

(a) O operador T é definido positivo.

(b) O operador T é auto-adjunto e todos os autovalores de T são positivos.

(c) Existe um operador auto-adjunto e definido positivo S tal que $T = S^2$. O operador S é chamado **raiz quadrada** de T .

(Sugestão: Mostre que (a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a). Na parte (b) \Rightarrow (c) defina S na base em que a matriz de T é diagonal.)

4.2.9. Um operador $T : \mathbb{V} \rightarrow \mathbb{V}$ em um espaço com produto interno tal que $\|T(V)\| = \|V\|$, para todo $V \in \mathbb{V}$ é chamado **operador unitário**, se o espaço vetorial \mathbb{V} é complexo e **operador ortogonal** se o espaço vetorial \mathbb{V} é real. Mostre que T é unitário (ortogonal) se, e somente se, $\langle T(V), T(W) \rangle = \langle V, W \rangle$, para todos $V, W \in \mathbb{V}$. (Sugestão: use as identidades polares, [Exercício 2.1.11 na página 131](#).)

4.2.10. Mostre que um operador T é unitário (ortogonal) se, e somente se, T^* existe e $T^*T = TT^* = I$. (Sugestão: mostre que se T é unitário, então T é invertível e $T^* = T^{-1}$.)

4.2.11. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador em um espaço com produto interno de dimensão finita. Mostre que as afirmações seguintes são equivalentes:

(a) O operador T é unitário (ortogonal).

(b) O operador T leva base ortonormal em base ortonormal.

(c) Existe uma base ortonormal $\mathcal{B} = \{V_1, \dots, V_n\}$, tal que $\{T(V_1), \dots, T(V_n)\}$ é uma base ortonormal.

(Sugestão: Mostre que (a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a). Na parte (c) \Rightarrow (a), mostre que $\langle T(V), T(W) \rangle = \langle V, W \rangle$, para todos $V, W \in \mathbb{V}$.)

- 4.2.12.** Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear invertível em um espaço com produto interno de dimensão finita. Mostre que existe um operador auto-adjunto definido positivo P e um operador unitário (ortogonal) U , tal que $T = PU$. Esta decomposição é única chamada de **decomposição polar de T** . (Sugestão: Sejam $P = (TT^*)^{1/2}$ e $U = P^{-1}T$. Mostre que $UU^* = I$.)
- 4.2.13.** Uma matriz $A, n \times n$, é chamada (**definida**) **positiva** se o operador $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definido por $T(X) = AX$, para todo $X \in \mathbb{R}^n$ é um operador definido positivo. Para $k = 1, \dots, n$, seja A_k a submatriz obtida de uma matriz $A, n \times n$, 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 simétrica e definida positiva, então
- A é não singular;
 - $\det(A) > 0$;
 - as submatrizes principais A_1, \dots, A_n são todas definidas positivas. (Sugestão: considere vetores X_k tais que os últimos $n - k$ elementos são nulos e observe que $\langle T(X), X \rangle = X^t A X$.)
- 4.2.14.** Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear normal em um espaço vetorial real de dimensão n . Mostre que existe uma base ortonormal \mathcal{C} de \mathbb{V} tal que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & D_k \end{bmatrix}, \quad \text{em que } D_i = [\lambda_i] \text{ ou } D_i = \begin{bmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{bmatrix}.$$

(Sugestão: use o [Exercício 4.1.27 na página 326](#).)

- 4.2.15.** Seja \mathbb{V} um espaço vetorial de dimensão finita. Mostre que um operador linear $T : \mathbb{V} \rightarrow \mathbb{V}$ é unitário se, e somente se, ele é normal e $|\lambda| = 1$ para todo autovalor λ de T .
- 4.2.16.** Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear ortogonal em um espaço vetorial real de dimensão n . Mostre que

existe uma base ortonormal \mathcal{C} de \mathbb{V} tal que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & D_k \end{bmatrix}, \quad \text{em que } D_i = [\pm 1] \text{ ou } D_i = \begin{bmatrix} \cos \theta_i & -\text{sen } \theta_i \\ \text{sen } \theta_i & \cos \theta_i \end{bmatrix}.$$

4.2.17. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial de dimensão n . Mostre que se $T^k = O$, a transformação linear nula, para algum $k > n$, então $T^n = O$. (Sugestão: Se A é uma matriz triangular superior, então $[A^j]_{ii} = (a_{ii})^j$ e assim o polinômio característico de T é $p_T(\lambda) = (-1)^n \lambda^n$. Use o [Teorema de Schur na página 338](#) e o de [Cayley-Hamilton na página 315](#).)

4.3 Aplicação na Identificação de Cônicas

Uma **equação quadrática** nas variáveis x e y tem a forma

$$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

onde 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, hipérbolas 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 4.1 na página 348](#).

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 4.13. Considere a cônica C cuja equação é

$$5x^2 - 4xy + 8y^2 - 36 = 0.$$

Esta equação pode ser escrita como

$$X^t A X - 36 = 0, \tag{4.13}$$

onde

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}.$$

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

$$(A - 4I_2)X = \vec{0}$$

ou

$$\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

cujas soluções são

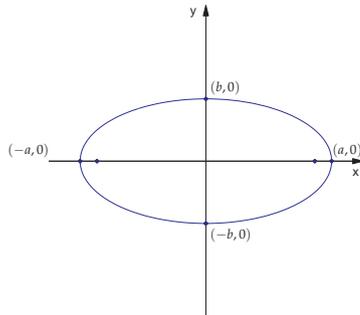
$$\mathbb{V}_1 = \{(2\alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

Assim, $V_1 = (2, 1)$ é uma base para \mathbb{V}_1 , pois gera \mathbb{V}_1 e é L.I. E $W_1 = \frac{V_1}{\|V_1\|} = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$ é uma base ortonormal para \mathbb{V}_1 .

Os autovetores associados a $\lambda_2 = 9$ são as soluções não nulas do sistema

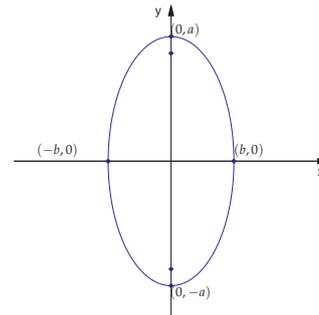
$$(A - 9I_2)X = \vec{0}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, a > b$$

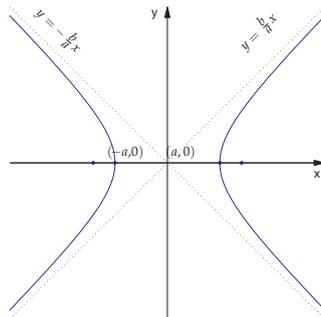


Elipse

$$\frac{y^2}{a^2} + \frac{x^2}{b^2} = 1, a > b$$

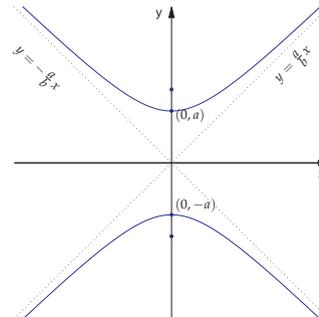


$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$



Hipérbole

$$\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1$$



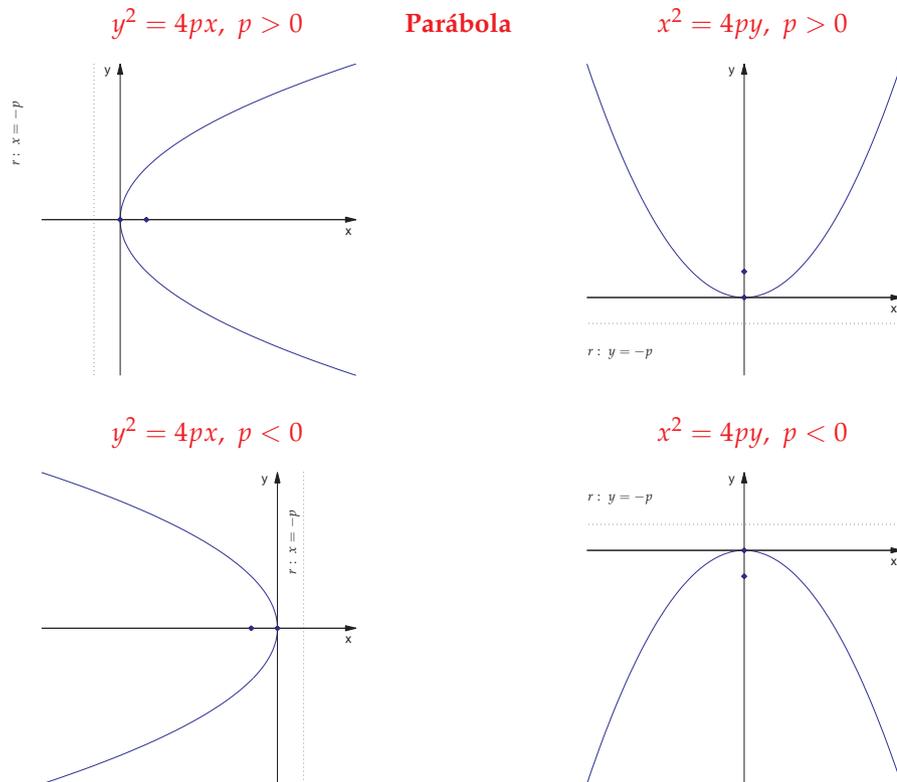


Figura 4.1: Cônicas não degeneradas com equações na forma padrão

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

$$W_2 = \frac{V_2}{\|V_2\|} = \left(\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$$

é uma base ortonormal para \mathbb{V}_2 . Portanto,

$$D = P^t A P$$

onde,

$$D = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}, \text{ e } P = [W_1 \ W_2] = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}.$$

Vamos fazer a mudança de variáveis $X = P X'$, onde $X' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ na equação (4.13).

Substituindo $X = P X'$ na equação (4.13), 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

$$\frac{x'^2}{9} + \frac{y'^2}{4} = 1 \tag{4.14}$$

que é a equação de uma elipse cujo esboço é mostrado na [Figura 4.2](#). 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 (primeira coluna de P) e o eixo y' passa pela origem, é paralelo e possui o mesmo sentido de W_2 (segunda coluna de P). Depois, a partir da equação (4.14), verificamos na [Figura 4.1 na página 348](#) a forma da curva em relação aos eixos x' e y' .

Exemplo 4.14. 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

$$X^t AX + KX + 4 = 0, \quad (4.15)$$

onde

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix} \text{ e } K = \begin{bmatrix} \frac{20}{\sqrt{5}} & -\frac{80}{\sqrt{5}} \end{bmatrix}.$$

A matriz A é a mesma do exemplo anterior. Assim, temos que

$$D = P^t AP$$

onde,

$$D = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}, \text{ e } P = [W_1 \ W_2] = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}.$$

Vamos fazer a mudança de variáveis $X = PX'$, onde $X' = \begin{bmatrix} x' \\ y' \end{bmatrix}$.

Substituindo $X = PX'$ na equação (4.15), obtemos

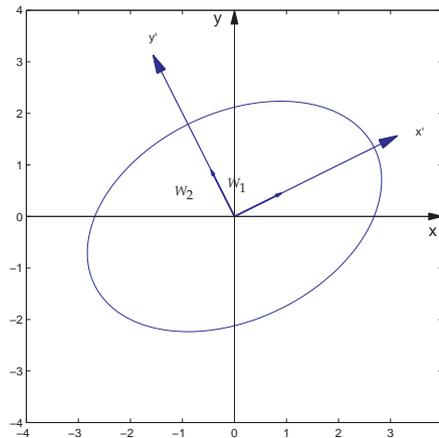


Figura 4.2: Elipse do Exemplo 4.13

$$X'^t(P^t AP)X' + KPX' + 4 = 0$$

ou

$$X'^t DX' + KPX' + 4 = 0,$$

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

$$x'' = x' - 1 \quad \text{e} \quad (4.16)$$

$$y'' = y' - 2 \quad (4.17)$$

obtemos

$$4x''^2 + 9y''^2 - 36 = 0$$

ou

$$\frac{x''^2}{9} + \frac{y''^2}{4} = 1 \quad (4.18)$$

que é a equação de uma elipse cujo esboço é mostrado na [Figura 4.3](#). 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 (a primeira coluna de P). O eixo y'' tem a direção e o sentido do vetor W_2 (a segunda

coluna de P). O eixo x'' tem equação $y'' = 0$. Usando a equação (4.16) obtemos $y' = 2$. O eixo y'' tem equação $x'' = 0$. Usando a equação (4.17) obtemos $x' = 1$. Depois, a partir da equação (4.18), verificamos na Figura 4.1 na página 348 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 4.17. *Considere a equação*

$$ax^2 + bxy + cy^2 + dx + ey + f = 0, \quad (4.19)$$

com $a, b, c, d, e, f \in \mathbb{R}$, sendo $b \neq 0$ e a e c não simultaneamente nulos. Então existe um sistema de coordenadas (x', y') , onde a equação (4.19) tem a forma

$$\lambda_1 x'^2 + \lambda_2 y'^2 + d'x' + e'y' + f = 0,$$

onde λ_1, λ_2 são os autovalores de

$$A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}.$$

Mais ainda,

$$X' = QX,$$

onde $X' = \begin{bmatrix} x' \\ y' \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \end{bmatrix}$ e Q é uma matriz ortogonal ($Q^{-1} = Q^t$).

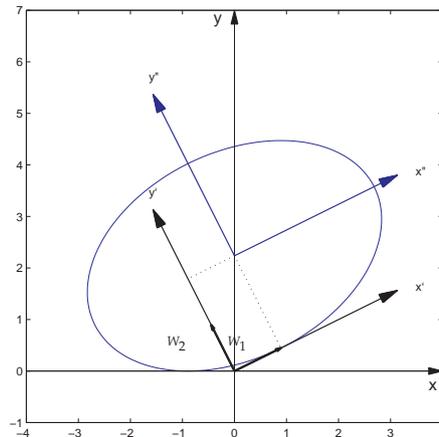


Figura 4.3: Elipse do Exemplo 4.14

Exercícios Numéricos (respostas na página 474)

Identificar a cônica, achar a equação no último sistema de coordenadas utilizado e fazer um esboço do gráfico.

4.3.1. $9x^2 - 4xy + 6y^2 = 30$;

4.3.2. $3x^2 - 8xy - 12y^2 + 81 = 0$;

4.3.3. $2x^2 - 4xy - y^2 = -24$;

4.3.4. $21x^2 + 6xy + 13y^2 - 132 = 0$;

4.3.5. $4x^2 - 20xy + 25y^2 - 15x - 6y = 0$;

4.3.6. $9x^2 + y^2 + 6xy - 10\sqrt{10}x + 10\sqrt{10}y + 90 = 0$;

4.3.7. $5x^2 + 5y^2 - 6xy - 30\sqrt{2}x + 18\sqrt{2}y + 82 = 0$;

4.3.8. $5x^2 + 12xy - 12\sqrt{13}x = 36$;

4.3.9. $6x^2 + 9y^2 - 4xy - 4\sqrt{5}x - 18\sqrt{5}y = 5$;

4.3.10. $x^2 - y^2 + 2\sqrt{3}xy + 6x = 0$;

4.3.11. $8x^2 + 8y^2 - 16xy + 33\sqrt{2}x - 31\sqrt{2}y + 70 = 0$;

4.3.12. $x^2 - 6xy - 7y^2 + 10x + 2y + 9 = 0$;

Exercícios usando o MATLAB[®]

Comandos do pacote GAAL:

>> [P,D]=diagonal(A) diagonaliza a matriz A, de forma que $AP=PD$, onde 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.
- >> `ellipse(a,b)` desenha a elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.
- >> `ellipse(a,b, [U1 U2])` desenha a elipse $\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1$, onde x' e y' são as coordenadas em relação à base ortonormal `U1` e `U2`.
- >> `ellipse(a,b, [U1 U2], X0)` desenha a elipse $\frac{x''^2}{a^2} + \frac{y''^2}{b^2} = 1$, onde 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 $\frac{x'^2}{a^2} - \frac{y'^2}{b^2} = 1$, onde x' e y' são as coordenadas em relação à base ortonormal `U1` e `U2`.
- >> `hiperbx(a,b, [U1 U2], X0)` desenha a hipérbole $\frac{x''^2}{a^2} - \frac{y''^2}{b^2} = 1$, onde 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 $\frac{y'^2}{a^2} - \frac{x'^2}{b^2} = 1$, onde x' e y' são as coordenadas em relação à base ortonormal `U1` e `U2`.
- >> `hiperby(a,b, [U1 U2], X0)` desenha a hipérbole $\frac{y''^2}{a^2} - \frac{x''^2}{b^2} = 1$, onde 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 $y'^2 = 4px'$, onde x' e y' são as coordenadas em relação à base ortonormal `U1` e `U2`.
- >> `parabx(p, [U1 U2], X0)` desenha a parábola $y''^2 = 4px''$, onde 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 $x'^2 = 4py'$, onde x' e y' são as coordenadas em relação à base ortonormal U1 e U2.

>> paraby(p, [U1 U2], X0) desenha a parábola $x''^2 = 4py''$, onde x'' e y'' são as coordenadas em relação ao sistema de coordenadas determinado pela base ortonormal U1 e U2 e por X0.

4.3.13. Use o MATLAB[®] para resolver os **Exercícios Numéricos**

Exercícios Teóricos

4.3.14. Demonstre o [Teorema 4.17](#) na página 353.

4.3.15. 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.

- Mostre que $\lambda\mu = ac - b^2/4$.
- Mostre que se $\lambda\mu > 0$, então C é uma elipse, um ponto ou o conjunto vazio.
- Mostre que se $\lambda\mu < 0$, então C é uma hipérbole, ou um par de retas concorrentes.
- Mostre que se $\lambda\mu = 0$, então temos duas possibilidades:
 - Se $\lambda \neq 0$ ou $\mu \neq 0$, então C é uma parábola, um par de retas paralelas, uma reta ou o conjunto vazio.
 - Se $\lambda = \mu = 0$, então C é uma reta.

4.4 Forma Canônica de Jordan

Já vimos ([Exemplo 4.6 na página 310](#)) que nem todo operador $T : \mathbb{V} \rightarrow \mathbb{V}$ em um espaço de dimensão finita é diagonalizável, ou seja, nem sempre existe uma base \mathcal{C} de \mathbb{V} 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 4.8. 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 $j = 1, \dots, m$. J_{λ_j} é chamado **bloco de Jordan**.

Exemplo 4.15. 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.

4.4.1 Autoespaço Generalizado

Vamos supor que exista uma base $\mathcal{C} = \{V_1^{(1)}, \dots, V_{m_1}^{(1)}, \dots, V_1^{(k)}, \dots, V_{m_k}^{(k)}\}$ de um espaço vetorial \mathbb{V} de dimensão $m_1 + \dots + m_k = n$ tal que a matriz do operador linear $T : \mathbb{V} \rightarrow \mathbb{V}$, esteja na forma canônica de Jordan

$$[T]_{\mathcal{C}}^{\mathcal{C}} = J = \begin{bmatrix} J_{\lambda_1} & \bar{0} & \dots & \bar{0} \\ \bar{0} & J_{\lambda_2} & \dots & \bar{0} \\ \vdots & & \ddots & \vdots \\ \bar{0} & \dots & \bar{0} & J_{\lambda_m} \end{bmatrix} \quad \text{em que} \quad J_{\lambda_j} = \begin{bmatrix} \lambda_j & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda_j & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_j & 0 \\ 0 & 0 & 0 & \dots & 1 & \lambda_j \end{bmatrix}_{m_j \times m_j}.$$

Assim,

$$\begin{aligned} T(V_i^{(j)}) &= \lambda_j V_i^{(j)} + V_{i+1}^{(j)} \quad \text{ou} \quad (T - \lambda_j I)(V_i^{(j)}) = V_{i+1}^{(j)} \quad \text{e} \quad (4.20) \\ T(V_{m_j}^{(j)}) &= \lambda_j V_{m_j}^{(j)} \end{aligned}$$

para $i = 1, \dots, m_j - 1$ e $j = 1, \dots, k$. Ou seja,

$$\begin{aligned} (T - \lambda_j I)(V_{m_j}^{(j)}) &= \bar{0}, \\ (T - \lambda_j I)^2(V_{m_j-1}^{(j)}) &= (T - \lambda_j I)(V_{m_j}^{(j)}) = \bar{0}, \\ &\vdots \\ (T - \lambda_j I)^{m_j}(V_1^{(j)}) &= (T - \lambda_j I)^{m_j-1}(V_2^{(j)}) = \bar{0}. \end{aligned} \quad (4.21)$$

Portanto, os vetores da base \mathcal{C} , $V_i^{(j)}$, e os elementos da diagonal de J , λ_j , satisfazem as equações

$$(T - \lambda I)^l(V) = \bar{0},$$

para $l = 1, \dots, m_j$. Isto motiva a seguinte definição.

Definição 4.9. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear. Um vetor $V \in \mathbb{V}$ 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 4.10. Seja λ um autovalor de um operador $T : \mathbb{V} \rightarrow \mathbb{V}$. O **autoespaço generalizado** associado a λ , é o subconjunto definido por

$$\mathbb{W}_\lambda = \{V \in \mathbb{V} \mid (T - \lambda I)^k(V) = \bar{0}, \text{ para algum inteiro } k > 0\}.$$

Observe que \mathbb{W}_λ consiste dos autovetores generalizados acrescentado o vetor nulo e contém o autoespaço associado a λ . Além disso \mathbb{W}_λ é um subespaço (verifique!). Assim, temos a seguinte sequência de subespaços encaixados:

$$\mathcal{N}(T - \lambda I_n) \subseteq \mathcal{N}(T - \lambda I_n)^2 \subseteq \cdots \subseteq \mathbb{W}_\lambda \subseteq \mathbb{V}.$$

Exemplo 4.16. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ 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,

$$\mathbb{W}_3 = \mathcal{N}(A - 3I_3)^2 = \mathbb{R}^3.$$

Já mostramos na [Proposição 4.4 na página 306](#) que um vetor não pode ser autovetor associado a dois autovalores distintos e que os autoespaços são T -invariantes. O próximo resultado, generaliza isto para autovetores e autoespaços generalizados, ou seja, não existe autovetor generalizado associado a autovalores distintos e os autoespaços generalizados são T -invariantes.

Proposição 4.18. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço de dimensão finita.*

- (a) *Se λ é um autovalor de T , então \mathbb{W}_λ é um subespaço T -invariante.*
 - (b) *A interseção de dois autoespaços generalizados associados a autovalores diferentes é igual ao subespaço nulo, ou seja, se $\mu \neq \lambda$ são autovalores de T , então $\mathbb{W}_\lambda \cap \mathbb{W}_\mu = \{0\}$.*
-

Demonstração. (a) Para mostrar que \mathbb{W}_λ é T -invariante, considere $V \in \mathbb{W}_\lambda$. Seja p tal que $(T - \lambda I)^p(V) = \bar{0}$. Então:

$$(T - \lambda I)^p T(V) = T(T - \lambda I)^p(V) = T\bar{0} = \bar{0}.$$

Portanto, $T(V) \in \mathbb{W}_\lambda$.

(b) Sejam $\mu \neq \lambda$ autovalores de T . Inicialmente vamos mostrar que

$$\mathbb{W}_\lambda \cap \mathcal{N}(T - \mu I) = \{\bar{0}\}. \quad (4.22)$$

Vamos supor que exista um vetor $V \neq \bar{0}$ em $\mathbb{W}_\lambda \cap \mathcal{N}(T - \mu I)$. Seja p o menor inteiro tal que $(T - \lambda I)^p(V) = \bar{0}$. Então, $W = (T - \lambda I)^{p-1}(V)$ é um autovetor de T associado a λ e é também um autovetor associado a μ , pois

$$(T - \mu I)(W) = (T - \mu I)(T - \lambda I)^{p-1}(V) = (T - \lambda I)^{p-1}(T - \mu I)(V) = (T - \lambda I)^{p-1}\bar{0} = \bar{0}.$$

Isto está em contradição com a [Proposição 4.4\(a\) na página 306](#). Logo, V tem que ser igual ao vetor nulo.

Vamos supor, agora, que exista um vetor $V \neq \bar{0}$ em $\mathbb{W}_\lambda \cap \mathbb{W}_\mu$. Seja p o menor inteiro tal que $(T - \mu I)^p(V) = \bar{0}$. Então, $W = (T - \mu I)^{p-1}(V)$ é um autovetor de T associado a μ e é também um autovetor generalizado associado a λ , pois

$$(T - \lambda I)^k(T - \mu I)^{p-1}(V) = (T - \mu I)^{p-1}(T - \lambda I)^k(V) = (T - \mu I)^{p-1}\bar{0} = \bar{0},$$

para algum inteiro positivo k . Isto está em contradição com [\(4.22\)](#). Logo, V tem que ser igual ao vetor nulo. ■

No [Corolário 4.7 na página 313](#) provamos que a dimensão do autoespaço é menor ou igual a multiplicidade de λ no polinômio característico. Na próxima proposição generalizamos este resultado para autoespaços generalizados. Este resultado é útil na determinação dos autoespaços generalizados.

Proposição 4.19. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em que as raízes do polinômio característico são autovalores de T . Seja λ é um autovalor de T tal que a multiplicidade de λ no polinômio característico de T é igual a m , então*

- (a) $\dim(\mathbb{W}_\lambda) \leq m$,
- (b) $\mathbb{W}_\lambda = \mathcal{N}(T - \lambda I)^m$.

Demonstração. Seja p a dimensão de \mathbb{W}_λ .

- (a) Como pela **Proposição 4.18**, \mathbb{W}_λ é T -invariante, então pela **Proposição 4.6 na página 312** o polinômio característico de $T_{\mathbb{W}_\lambda}$ é um fator do polinômio característico de T ,

$$p_T(t) = p_{T_{\mathbb{W}_\lambda}}(t)q(t).$$

Vamos mostrar que o polinômio característico de $T_{\mathbb{W}_\lambda}$ é da forma

$$p_{T_{\mathbb{W}_\lambda}}(t) = (-1)^p(t - \lambda)^p.$$

Para isto vamos mostrar que λ é o único autovalor de $T_{\mathbb{W}_\lambda}$.

Sejam $\mu \neq \lambda$ e $W \in \mathbb{W}_\lambda$, tal que $(T_{\mathbb{W}_\lambda} - \mu I)W = \bar{0}$. Então, $W \in \mathbb{W}_\lambda \cap \mathbb{W}_\mu = \{\bar{0}\}$, pela **Proposição 4.18**. Portanto, $p_T(t) = (-1)^p(t - \lambda)^p q(t)$ e $p = \dim(\mathbb{W}_\lambda) \leq m$.

- (b) Claramente $\mathcal{N}(T - \lambda I)^m \subseteq \mathbb{W}_\lambda$. Vimos que \mathbb{W}_λ é T -invariante e na demonstração do item anterior que $p_{T_{\mathbb{W}_\lambda}} = (-1)^p(t - \lambda)^p$, com $p \leq m$. O que implica, pelo **Teorema de Cayley-Hamilton na página 315**, que

$$(T_{\mathbb{W}_\lambda} - \lambda I)^p = O,$$

a transformação linear nula. Logo, para todo $W \in \mathbb{W}_\lambda$ temos que

$$(T - \lambda I)^m(W) = (T - \lambda I)^{m-p}(T - \lambda I)^p(W) = (T - \lambda I)^{m-p}(\bar{0}) = \bar{0}.$$

O que mostra que $\mathbb{W}_\lambda \subseteq \mathcal{N}(T - \lambda I)^m$. Portanto, $\mathbb{W}_\lambda = \mathcal{N}(T - \lambda I)^m$.



Exemplo 4.17. Seja $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ 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, pela **Proposição 4.19**, $\mathbb{W}_3 = \mathcal{N}(T - 3I_4) = \{(\alpha, 0, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. 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, $\mathcal{N}(T - 2I_4) = \{(\beta, \alpha, \alpha, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$. 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, pela **Proposição 4.19**, $\mathbb{W}_2 = \mathcal{N}(T - 2I_4)^2 = \{(\gamma, \alpha + \beta, 2\beta, 2\alpha) \mid \alpha, \beta, \gamma \in \mathbb{R}\}$.

Por (4.21) na página 360, temos que se existe uma base \mathcal{C} de \mathbb{V} tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ está na forma canônica de Jordan, então a base \mathcal{C} é formada de autovetores generalizados. Assim, para encontrarmos a forma canônica de Jordan precisamos encontrar o maior número possível de autovetores generalizados linearmente independentes.

O próximo resultado garante que um conjunto formado por grupos de autovetores generalizados L.I., associados a autovalores diferentes, é L.I. Este resultado é uma generalização da [Proposição 4.5 na página 306](#).

Proposição 4.20. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial \mathbb{V} de dimensão finita. Sejam $\lambda_1, \dots, \lambda_l$ autovalores distintos de T . Se para cada $i = 1, \dots, l$, $\{V_1^{(i)}, \dots, V_{k_i}^{(i)}\}$ é um conjunto de autovetores generalizados L.I. associados a λ_i , então $\{V_1^{(1)}, \dots, V_{k_1}^{(1)}, \dots, V_1^{(l)}, \dots, V_{k_l}^{(l)}\}$ é um conjunto L.I.*

Demonstração. Seja p um inteiro positivo suficientemente grande. Vamos, em primeiro lugar, mostrar que se

$$V_1 + \cdots + V_l = \bar{0}, \quad (4.23)$$

com $V_i \in \mathcal{N}(T - \lambda_i I)^p$, então $V_i = \bar{0}$, para $i = 1, \dots, l$. Vamos supor, por indução, que

$$V_1 + \cdots + V_{l-1} = \bar{0},$$

com $V_i \in \mathcal{N}(T - \lambda_i I)^p$, implica que $V_i = \bar{0}$, para $i = 1, \dots, l-1$. Aplicando-se $(T - \lambda_l I)^p$ a ambos os membros da equação (4.23) obtemos

$$(T - \lambda_l I)^p V_1 + \cdots + (T - \lambda_l I)^p V_{l-1} = \bar{0}.$$

Pela **Proposição 4.18** e pela hipótese de indução temos que

$$V_i = \bar{0}, \text{ para } i = 1, \dots, l-1.$$

Substituindo-se $V_1 = \cdots = V_{l-1} = \bar{0}$ na equação (4.23) obtemos que também $V_l = 0$.

Precisamos mostrar que a única solução da equação

$$x_1^{(1)} V_1^{(1)} + \cdots + x_{k_1}^{(1)} V_{k_1}^{(1)} + \cdots + x_1^{(l)} V_1^{(l)} + \cdots + x_{k_l}^{(l)} V_{k_l}^{(l)} = \bar{0} \quad (4.24)$$

é a solução trivial. Como $V_i = x_1^{(i)} V_1^{(i)} + \cdots + x_{k_i}^{(i)} V_{k_i}^{(i)} \in \mathcal{N}(T - \lambda_i I)^p$, pelo que mostramos acima, a equação (4.24) implica que

$$V_i = x_1^{(i)} V_1^{(i)} + \cdots + x_{k_i}^{(i)} V_{k_i}^{(i)} = \bar{0},$$

para $i = 1, \dots, l$. Como para cada $i = 1, \dots, l$, $V_1^{(i)}, \dots, V_{k_i}^{(i)}$ são L.I., então

$$x_1^{(i)} = \cdots = x_{k_i}^{(i)} = 0.$$

O que prova que todos os autovetores generalizados juntos são L.I. ■

Exemplo 4.18. Seja $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ o operador definido por $T(X) = AX$, para todo $X \in \mathbb{R}^4$, em que

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

Já vimos no **Exemplo 4.17** que os autoespaços generalizados de T são

$$\mathbb{W}_3 = \mathcal{N}(A - 3I_4) = \{(\alpha, 0, 0, \alpha) \mid \alpha \in \mathbb{R}\} \quad \text{e}$$

$$\mathbb{W}_2 = \mathcal{N}(A - 2I_4)^2 = \{(\gamma, \alpha + \beta, 2\beta, 2\alpha) \mid \alpha, \beta, \gamma \in \mathbb{R}\}.$$

O vetor $V_1 = (1, 0, 0, 1)$ forma uma base de \mathbb{W}_3 e os vetores

$$V_2 = (0, 1, 0, 2), V_3 = (0, 1, 2, 0) \text{ e } V_4 = (1, 0, 0, 0)$$

formam uma base de \mathbb{W}_2 . Portanto, pela **Proposição 4.20**, segue-se que V_1, V_2, V_3, V_4 são L.I. e assim formam uma base de \mathbb{R}^4 .

Vamos mostrar que se as raízes do polinômio característico de T são autovalores, então sempre existe uma base de \mathbb{V} formada de autovetores generalizados.

Teorema 4.21. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador tal que seu polinômio característico é da forma

$$p(t) = (-1)^n (t - \lambda_1)^{n_1} \dots (t - \lambda_k)^{n_k}, \text{ com } \lambda_1, \dots, \lambda_k \text{ distintos.}$$

Sejam $\mathcal{B}_1, \dots, \mathcal{B}_k$ bases dos autoespaços generalizados \mathbb{W}_i associados a λ_i , para $i = 1, \dots, k$. Então,

- (a) $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ é uma base de \mathbb{V} ;
- (b) $\dim(\mathbb{W}_i) = n_i$, para $i = 1, \dots, k$.

Demonstração. Vamos mostrar que $\mathbb{V} = \mathbb{W}_1 \oplus \dots \oplus \mathbb{W}_k$. Os dois itens são então decorrência do [Exercício 3.3.21 na página 269](#) e da [Proposição 4.19 na página 365](#).

Vamos provar para o caso em que $k = 2$. Sejam

$$p_1(t) = (t - \lambda_1)^{n_1} \text{ e } p_2(t) = (t - \lambda_2)^{n_2}.$$

Como o máximo divisor comum de $p_1(t)$ e $p_2(t)$ é igual a 1, então existem polinômios $m_1(t)$ e $m_2(t)$ tais que

$$1 = m_1(t)p_1(t) + m_2(t)p_2(t).$$

Este resultado pode ser encontrado em qualquer livro de Álgebra Abstrata como por exemplo [\[5\]](#). Portanto, para todo vetor $V \in \mathbb{V}$

$$V = I(V) = m_1(T)p_1(T)(V) + m_2(T)p_2(T)(V).$$

Mas, $m_1(T)p_1(T)(V)$ pertence a $\mathbb{W}_2 = \mathcal{N}(p_2(T))$ e $m_2(T)p_2(T)(V)$ pertence a $\mathbb{W}_1 = \mathcal{N}(p_1(T))$, pois pelo Teorema de Cayley-Hamilton:

$$p_1(T)(m_2(T)p_2(T)(V)) = m_2(T)(p_1(T)p_2(T))(V) = m_2(T)((-1)^n p(T)(V)) = \vec{0}$$

e

$$p_2(T)(m_1(T)p_1(T)(V)) = m_1(T)(p_1(T)p_2(T))(V) = m_1(T)((-1)^n p(T)(V)) = \bar{0}.$$

Assim,

$$\mathbb{V} = \mathbb{W}_1 + \mathbb{W}_2.$$

Como pela [Proposição 4.18 na página 363](#), $\mathbb{W}_1 \cap \mathbb{W}_2 = \{\bar{0}\}$, então

$$\mathbb{V} = \mathbb{W}_1 \oplus \mathbb{W}_2.$$

■

Observação. Segue do [Teorema 4.21](#) que, para determinarmos o autoespaço generalizado \mathbb{W}_λ , associado a um autovalor λ de um operador $T : \mathbb{V} \rightarrow \mathbb{V}$, devemos determinar

$$\mathcal{N}(T - \lambda I), \mathcal{N}(T - \lambda I)^2, \dots, \mathcal{N}(T - \lambda I)^q,$$

em que q é o menor inteiro positivo tal que $\dim(\mathcal{N}(T - \lambda I)^q)$ é igual a multiplicidade de λ no polinômio característico de T . Então, $\mathbb{W}_\lambda = \mathcal{N}(T - \lambda I)^q$.

4.4.2 Ciclos de Autovetores Generalizados

Por (4.21), os vetores que formam a base \mathcal{C} tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ está na forma canônica de Jordan são autovetores generalizados. Mas além disso, por (4.20), segue-se que os vetores da base \mathcal{C} que correspondem a um bloco de Jordan formam um ciclo no sentido que se segue.

Definição 4.11. Seja V um autovetor generalizado associado ao autovalor λ de um operador $T : \mathbb{V} \rightarrow \mathbb{V}$. Seja p o menor inteiro tal que $(T - \lambda I)^p(V) = \vec{0}$. O conjunto

$$\mathcal{C}_V = \{V, (T - \lambda I)(V), \dots, (T - \lambda I)^{p-1}(V)\}$$

é chamado um **ciclo de autovetores generalizados associados a λ** . Os vetores $(T - \lambda I)^{p-1}(V)$ e V são chamados **vetor final** e **vetor inicial** do ciclo, respectivamente. Dizemos também que o **comprimento do ciclo** é p .

Observe que o vetor final de um ciclo de autovetores generalizados de um operador T é o único autovetor de T no ciclo. Assim, o maior número possível de ciclos de autovetores generalizados associados a um autovalor λ , com os vetores finais formando um conjunto L.I., é igual à dimensão do núcleo de $T - \lambda I$.

Exemplo 4.19. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ o operador definido por $T(X) = AX$, para todo $X \in \mathbb{R}^3$, em que

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

O polinômio característico de T é $p(t) = -(t-3)^3$. Assim o único autovalor de T é $\lambda = 3$. A forma escalonada reduzida de

$$A - 3I_3 = \begin{bmatrix} 1 & 1 & -1 \\ -1 & -2 & 3 \\ -1 & -1 & 1 \end{bmatrix} \quad \text{é} \quad \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, $\mathcal{N}(T - 3I) = \{(-\alpha, 2\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. A forma escalonada reduzida de

$$(A - 3I_3)^2 = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 0 & -2 \\ -1 & 0 & -1 \end{bmatrix} \quad \text{é} \quad \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, $\mathcal{N}(T - 3I)^2 = \{(-\alpha, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$. Agora, $(A - 3I_3)^3 = \bar{0}$. Assim, $\mathcal{N}(A - 3I_3)^3 = \mathbb{W}_3 = \mathbb{R}^3$.

Os ciclos de \mathbb{W}_3 têm comprimento no máximo igual a 3. Para obtermos um ciclo de maior comprimento possível em \mathbb{W}_3 , devemos tomar um vetor em \mathbb{W}_3 que não pertença ao $\mathcal{N}(T - 3I)^2$. Para determinar um tal vetor calculamos $(T - 3I)^2$ aplicado em vetores de uma base de $\mathbb{W}_3 = \mathbb{R}^3$. Ou seja,

$$(A - 3I_3)^2 [E_1 \ E_2 \ E_3] = (A - 3I_3)^2 I_3 = (A - 3I_3)^2.$$

Daí concluímos que os vetores $E_1 = (1, 0, 0)$ e $E_3 = (0, 0, 1)$ não pertencem a $\mathcal{N}(T - 3I)^2$. Assim,

$$\mathcal{C}_{E_1} = \{E_1, (A - 3I_3)E_1 = (1, -1, -1), (A - 3I_3)^2 E_1 = (1, -2, -1)\}$$

é um ciclo de autovetores generalizados associados a $\lambda_1 = 3$ de maior comprimento possível.

	Base de		
	$\mathcal{N}(T - 3I)^3$		
		$\mathcal{N}(T - 3I)^2$	
			$\mathcal{N}(T - 3I)$
E_1 \cap $\mathcal{N}(T - 3I)^3 \setminus \mathcal{N}(T - 3I)^2$	E_1	$(T - 3I)E_1$	$(T - 3I)^2E_1$

Como queremos encontrar o maior número possível de autovetores generalizados L.I. que compõem ciclos, o resultado seguinte garante que basta que os autovetores que estão nos ciclos (os vetores finais) sejam L.I. para que o conjunto completo seja L.I.

Teorema 4.22. *Seja λ um autovalor de um operador $T : \mathbb{V} \rightarrow \mathbb{V}$. Sejam C_1, \dots, C_q ciclos de autovetores generalizados de T associados a λ tais que os vetores finais dos ciclos são linearmente independentes. Então, os C_i 's são disjuntos (isto é, $C_i \cap C_j = \emptyset$, para $i \neq j$) e a união $\mathcal{C} = \bigcup_{i=1}^q C_i$ é linearmente independente.*

Demonstração. Deixamos como exercício a verificação de que os C_i 's são disjuntos.

Vamos inicialmente supor que tenhamos apenas dois ciclos cada um com comprimento igual a 2 com dois vetores cada,

$\mathcal{C}_1 = \{V_1, (T - \lambda I)(V_1)\}$, $\mathcal{C}_2 = \{V_2, (T - \lambda I)(V_2)\}$ e que $(T - \lambda I)(V_1)$ e $(T - \lambda I)(V_2)$ são L.I. Considere a seguinte combinação linear nula:

$$x_{10}V_1 + x_{20}V_2 + x_{11}(T - \lambda I)(V_1) + x_{21}(T - \lambda I)(V_2) = \vec{0}. \quad (4.25)$$

Aplicando-se $T - \lambda I$ em (4.25), obtemos:

$$x_{10}(T - \lambda I)(V_1) + x_{20}(T - \lambda I)(V_2) = \vec{0}.$$

O que implica que, $x_{10} = x_{20} = 0$. Substituindo-se estes valores em (4.25) obtemos que $\mathcal{C}_1 \cup \mathcal{C}_2$ é L.I.

Seja p o número de vetores de \mathcal{C} . Se $p = 1$ o resultado é claramente verdadeiro. Suponhamos que $p > 1$ e que o resultado seja verdadeiro se a união dos ciclos contiver menos de p vetores. Vamos mostrar que ele é verdadeiro se o número de vetores de \mathcal{C} é p . Considere a seguinte combinação linear nula:

$$x_{10}V_1 + \cdots + x_{q0}V_q + \sum_{i=1}^q \sum_{j=1}^{p_i-1} x_{ij}(T - \lambda I)^j(V_i) = \vec{0}, \quad (4.26)$$

em que $\mathcal{C}_i = \{V_i, (T - \lambda_i)V_i, \dots, (T - \lambda_i)^{p_i-1}V_i\}$. Aplicando-se $T - \lambda I$ em (4.26), obtemos:

$$\sum_{i=1}^q \sum_{j=1}^{p_i-1} x_{i(j-1)}(T - \lambda I)^j(V_i) = \vec{0},$$

que é uma combinação linear nula de vetores de $\mathcal{C}' = \bigcup_{i=1}^q \mathcal{C}'_i$, em que \mathcal{C}'_i é o ciclo obtido de \mathcal{C}_i retirando-se o vetor inicial, portanto tem o mesmo vetor final de \mathcal{C}_i . \mathcal{C}' tem menos de p vetores. Como estamos supondo o resultado verdadeiro neste caso, temos que, $x_{ij} = 0$, para $i = 1, \dots, q$ e $j = 0, \dots, p_i - 2$. Substituindo-se estes valores em (4.26), obtemos que $\mathcal{C} = \bigcup_{i=1}^q \mathcal{C}_i$ é linearmente independente. ■

Observação. Para encontrar uma base \mathcal{C} tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ esteja na forma de Jordan, devemos, para cada autovalor λ , encontrar um número de ciclos igual a dimensão do autoespaço determinado por λ . Sendo os ciclos de comprimento o maior possível e com vetores finais (os autovetores de cada ciclo) L.I. Para isso, para cada autovalor λ com multiplicidade p no polinômio característico de T :

- (a) Devemos determinar bases para $\mathcal{N}(T - \lambda I), \mathcal{N}(T - \lambda I)^2, \dots, \mathcal{N}(T - \lambda I)^q$, em que q é o menor inteiro positivo tal que $\dim(\mathcal{N}(T - \lambda I)^q) = p$.
 - (b) Os vetores iniciais dos maiores ciclos são os vetores que estão numa base de $\mathbb{W}_{\lambda} = \mathcal{N}(T - \lambda I)^q$ e que **não** pertencem a $\mathcal{N}(T - \lambda I)^{q-1}$. Para isso, podemos tomar vetores que pertencem a uma base de $\mathcal{N}(T - \lambda I)^q$ e que não são combinação linear de uma base de $\mathcal{N}(T - \lambda I)^{q-1}$ ou tais que $(T - \lambda I)^{q-1}$ aplicado neles seja diferente do vetor nulo.
 - (c) O número de ciclos deve ser igual a dimensão do autoespaço associado a λ . O número de vetores da união de todos os ciclos deve ser igual a dimensão de \mathbb{W}_{λ} . Para os ciclos de comprimento $j = q - 1, \dots, 1$, se existirem, os vetores iniciais são vetores que pertencem a uma base de $\mathcal{N}(T - \lambda I)^j$ tais que $(T - \lambda I)^{j-1}$ aplicado neles são linearmente independentes com os autovetores (últimos vetores) dos ciclos já determinados.
-

Base de			
$\mathcal{N}(T - \lambda I)^q$			
⋮			
$\mathcal{N}(T - \lambda I)^2$			
$\mathcal{N}(T - \lambda I)$			
$\begin{matrix} V_1 \\ \cap \\ \mathcal{N}(T - \lambda I)^q \setminus \mathcal{N}(T - \lambda I)^{q-1} \\ \vdots \\ V_k \\ \cap \\ \mathcal{N}(T - \lambda I)^{q_k} \setminus \mathcal{N}(T - \lambda I)^{q_k-1} \\ \vdots \end{matrix}$	$\begin{matrix} V_1, (T - \lambda I)V_1, (T - \lambda I)^2V_1, \dots \\ \vdots \\ V_k, (T - \lambda I)V_k, \dots \\ \vdots \end{matrix}$	$\begin{matrix} (T - \lambda I)^{q-2}V_1 \\ \vdots \\ (T - \lambda I)^{q_k-2}V_k \\ \vdots \end{matrix}$	$\begin{matrix} (T - \lambda I)^{q-1}V_1 \\ \vdots \\ (T - \lambda I)^{q_k-1}V_k \\ \vdots \end{matrix}$

Exemplo 4.20. Seja $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ o operador definido por $T(X) = AX$, para todo $X \in \mathbb{R}^4$, em que

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

Já vimos no [Exemplo 4.18 na página 369](#), que os autovalores de T são $\lambda_1 = 3$ e $\lambda_2 = 2$ e que

$$\begin{aligned} \mathcal{N}(T - 3I) &= \{(\alpha, 0, 0, \alpha) \mid \alpha \in \mathbb{R}\} = \mathbb{W}_3, \\ \mathcal{N}(T - 2I) &= \{(\beta, \alpha, \alpha, \alpha) \mid \alpha, \beta \in \mathbb{R}\} \text{ e} \\ \mathcal{N}(T - 2I)^2 &= \{(\gamma, \alpha + \beta, 2\beta, 2\alpha) \mid \alpha, \beta, \gamma \in \mathbb{R}\} = \mathbb{W}_2. \end{aligned}$$

Todos os ciclos em \mathbb{W}_3 têm comprimento igual a 1. Seja $V_1 = (1, 0, 0, 1)$. Este vetor forma uma base de \mathbb{W}_3 e será o primeiro vetor da base \mathcal{C} . Os ciclos de \mathbb{W}_2 têm comprimento no máximo igual a 2. Para obtermos um ciclo de maior comprimento possível em \mathbb{W}_2 , devemos tomar um vetor em \mathbb{W}_2 que não pertença a $\mathcal{N}(T - 2I)$.

$\{W_1 = (1, 0, 0, 0), W_2 = (0, 1, 1, 1)\}$ é uma base de $\mathcal{N}(T - 2I)$.

$\{W_3 = (0, 1, 0, 2), W_4 = (0, 1, 2, 0), W_5 = (1, 0, 0, 0)\}$ é uma base de $\mathbb{W}_2 = \mathcal{N}(T - 2I)^2$.

Vamos descobrir entre os vetores W_3, W_4, W_5 , quais **não** pertencem a $\mathcal{N}(T - 2I)$. Para isso vamos calcular $(A - 2I_3)[W_3 \ W_4 \ W_5]$:

$$(A - 2I_3)[W_3 \ W_4 \ W_5] = (A - 2I_3) \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \end{bmatrix}.$$

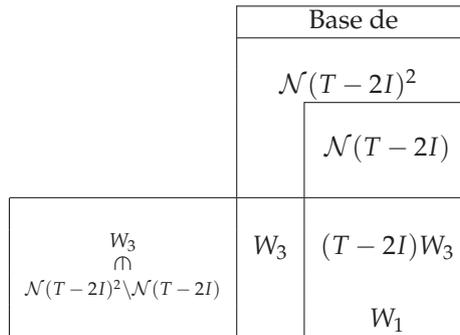
Portanto, W_3 e W_4 pertencem a \mathbb{W}_2 , mas não pertencem a $\mathcal{N}(T - 2I)$.

Sejam $V_2 = W_3 = (0, 1, 0, 2)$, $V_3 = (A - 2I_4)V_2 = (1, 1, 1, 1)$ o segundo e o terceiro vetor de \mathcal{C} , respectivamente. Precisamos ainda um vetor que pertença a $\mathcal{N}(T - 2I)$ e que seja linearmente independente com o vetor $V_3 = (T - 2I)(V_2) = (1, 1, 1, 1)$. É fácil ver que W_1 e W_2 **não** são combinação linear de V_3 . Seja $V_4 = W_1 = (1, 0, 0, 0)$ o quarto vetor da base \mathcal{C} .

Assim, se

$$P = [V_1 \ V_2 \ V_3 \ V_4] = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 \end{bmatrix}, \quad \text{então} \quad P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

que está na forma canônica de Jordan.



Exemplo 4.21. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ o operador definido por $T(X) = AX$, para todo $X \in \mathbb{R}^3$, em que

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

Já vimos no Exemplo 4.16 na página 362 que

$$\mathcal{N}(T - 3I) = \{(-\beta, 2\beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\} \text{ e } \mathcal{N}(T - 3I)^2 = \mathbb{R}^3 = \mathbb{W}_3.$$

Precisamos encontrar um vetor que pertença ao $\mathcal{N}(T - 3I)^2 = \mathbb{R}^3$, mas não pertença ao $\mathcal{N}(T - 3I)$ que é o subespaço que tem como base $\{V_1 = (0, 0, 1), V_2 = (-1, 2, 0)\}$. Para isso vamos descobrir quais vetores da base canônica **não** pertencem a $\mathcal{N}(T - 3I)$.

$$(A - 3I_3)[E_1 \ E_2 \ E_3] = (A - 3I_3)I_3 = A - 3I_3 = \begin{bmatrix} 2 & 1 & 0 \\ -4 & -2 & 0 \\ -2 & -1 & 0 \end{bmatrix}$$

Daí concluímos que $E_1 = (1, 0, 0)$ e $E_2 = (0, 1, 0)$ não pertencem a $\mathcal{N}(T - 3I)$. Consideremos então o ciclo iniciado com

$$W_1 = E_1. \quad W_2 = (A - 3I_3)E_1 = (2, -4, 2).$$

Vamos tomar para W_3 um autovetor de T que não é múltiplo escalar de W_2 . Por exemplo, $W_3 = V_1$. Assim, se

$$P = [W_1 \ W_2 \ W_3] = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -4 & 0 \\ 0 & 2 & 1 \end{bmatrix}, \quad \text{então} \quad P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

que está na forma canônica de Jordan.

	Base de	
	$\mathcal{N}(T - 3I)^2$	
	$\mathcal{N}(T - 3I)$	
E_1 \cap $\mathcal{N}(T - 3I)^2 \setminus \mathcal{N}(T - 3I)$	E_1	$(T - 3I)E_1$ V_1

O próximo teorema, garante a existência de uma base em relação a qual a matriz do operador está na forma canônica de Jordan, se as raízes do polinômio característico são autovalores do operador.

Teorema 4.23. *Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador tal que seu polinômio característico é da forma*

$$p(t) = (-1)^n (t - \lambda_1)^{n_1} \dots (t - \lambda_k)^{n_k}, \text{ com } \lambda_1, \dots, \lambda_k \text{ distintos.}$$

Então,

- (a) *Para $j = 1, \dots, k$, o autoespaço generalizado associado a λ_j tem uma base consistindo de uma união de ciclos de autovetores generalizados associados a λ_j .*
 - (b) *Existe uma base \mathcal{C} de \mathbb{V} formada por uma união de ciclos de autovetores generalizados tal que $[T]_{\mathcal{C}}^{\mathcal{C}}$ está na forma canônica de Jordan.*
-

Demonstração.

- (a) Vamos mostrar que se λ é um autovalor de T , então existe uma base de \mathbb{W}_λ consistindo de uma união de ciclos de autovetores generalizados. Seja $U : \mathbb{W}_\lambda \rightarrow \mathbb{W}_\lambda$, definido por

$$U(W) = (T - \lambda I)(W), \quad \text{para todo } W \in \mathbb{W}_\lambda.$$

Observe que W é um autovetor generalizado de T associado a λ se, e somente se, W é um autovetor generalizado de U associado a zero. Assim, basta provarmos o resultado para U .

Seja $p = \dim(\mathbb{W}_\lambda)$. Vamos supor que o resultado seja verdadeiro se a dimensão do autoespaço generalizado for menor que p . $\mathcal{I}(U)$ é o autoespaço generalizado da restrição de U a $\mathcal{I}(U)$ associado a zero. Como $\mathcal{N}(U)$ é não trivial, então $\dim(\mathcal{I}(U)) < p$. Como estamos supondo o resultado verdadeiro neste caso, existem $\mathcal{C}'_1, \dots, \mathcal{C}'_k$ ciclos de autovetores generalizados da restrição de U a $\mathcal{I}(U)$

tais que $\mathcal{C}' = \bigcup_{i=1}^k \mathcal{C}'_i$ é uma base de $\mathcal{I}(U)$.

Para cada $i = 1, \dots, k$, $\mathcal{C}'_i = \{W_i, \dots, U^{n_i}(W_i)\}$, para $W_i \in \mathcal{I}(U)$. Logo, existem $V_1, \dots, V_k \in \mathbb{W}_\lambda$ tais que $\mathcal{C}'_i = \{U(V_i), \dots, U^{n_i+1}(V_i)\}$, para $i = 1, \dots, k$.

Para cada $i = 1, \dots, k$, seja $\mathcal{C}_i = \{V_i\} \cup \mathcal{C}'_i$. Sejam V_{k+1}, \dots, V_r tais que

$$\{U^{m_1+1}(V_1), \dots, U^{m_k+1}(V_k), V_{k+1}, \dots, V_r\} \quad \text{é uma base de } \mathcal{N}(U).$$

Seja $\mathcal{C} = \bigcup_{i=1}^k \mathcal{C}_i \cup \{V_{k+1}\} \cup \dots \cup \{V_r\}$. \mathcal{C} é uma união de ciclos L.I. pela

Proposição 4.22 na página 374 e como \mathcal{C} tem

$$\dim(\mathcal{I}(U)) + \dim(\mathcal{N}(U)) = \dim(\mathbb{W}_\lambda)$$

vetores, então \mathcal{C} é uma base de \mathbb{W}_λ formada por ciclos de autovetores generalizados de U e portanto de T .

- (b) Sejam $\mathcal{C}_j = \{V_j, \dots, (T - \lambda_j I)^{m_j-1}(V_j)\}$ ciclos de autovetores generalizados tais que $\mathcal{C} = \cup \mathcal{C}_j$ seja uma base \mathbb{V} , que existe pelo item anterior e pelo **Teorema 4.21** na página 370. Então,

$$\begin{aligned}T(T - \lambda_j I)^i(V_j) &= (T - \lambda_j I)^{i+1}(V_j) + \lambda_j(T - \lambda_j I)^i(V_j). \quad \text{para } i = 0, \dots, m_j - 2 \text{ e} \\T(T - \lambda_j I)^{m_j-1}(V_j) &= (T - \lambda_j I)(T - \lambda_j I)^{m_j-1}(V_j) + \lambda_j(T - \lambda_j I)^{m_j-1}(V_j) \\&= \lambda_j(T - \lambda_j I)^{m_j-1}(V_j),\end{aligned}$$

Logo, a matriz $[T]_{\mathcal{C}}^{\mathcal{C}}$ está na forma canônica de Jordan.



4.4.3 Aplicação: Funções de Matrizes e Sistemas de Equações Diferenciais Lineares

Vamos estender funções $f \in \mathcal{C}^\infty(I)$ ao conjunto de todas as matrizes quadradas cujas raízes do polinômio característico sejam reais e estejam no intervalo I . Vamos estender inicialmente para blocos de Jordan.

Se

$$J_\lambda = \begin{bmatrix} \lambda & 0 & \cdots & 0 & 0 \\ 1 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 0 \\ 0 & 0 & \cdots & 1 & \lambda \end{bmatrix}_{m \times m},$$

então definimos $f(J_\lambda)$ por

$$f(J_\lambda) = \begin{bmatrix} f(\lambda) & 0 & \cdots & 0 & 0 \\ f'(\lambda) & f(\lambda) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{f^{(m-2)}(\lambda)}{(m-2)!} & \frac{f^{(m-3)}(\lambda)}{(m-3)!} & \cdots & f(\lambda) & 0 \\ \frac{f^{(m-1)}(\lambda)}{(m-1)!} & \frac{f^{(m-2)}(\lambda)}{(m-2)!} & \cdots & f'(\lambda) & f(\lambda) \end{bmatrix}.$$

Para $k > 0$, se $f(x) = x^k$, então $f(J_\lambda) = J_\lambda^k$ (Exercício 4.4.7 na página 396). Além disso, para quaisquer funções $f, g \in \mathcal{C}^\infty(I)$, e escalar α , $(\alpha f)(J_\lambda) = \alpha f(J_\lambda)$, $(f + g)(J_\lambda) = f(J_\lambda) + g(J_\lambda)$, $(fg)(J_\lambda) = f(J_\lambda)g(J_\lambda)$ (verifique!).

Se a matriz A é semelhante a uma matriz na forma canônica de Jordan, $A = PJP^{-1}$,

em que

$$J = \begin{bmatrix} J_{\lambda_1} & \bar{0} & \cdots & \bar{0} \\ \bar{0} & J_{\lambda_2} & \cdots & \bar{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{0} & \bar{0} & \cdots & J_{\lambda_k} \end{bmatrix},$$

então definimos $f(A)$ por

$$f(A) = P \begin{bmatrix} f(J_{\lambda_1}) & \bar{0} & \cdots & \bar{0} \\ \bar{0} & f(J_{\lambda_2}) & \cdots & \bar{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{0} & \bar{0} & \cdots & f(J_{\lambda_k}) \end{bmatrix} P^{-1}.$$

Para $k > 0$, se $f(x) = x^k$, então $f(A) = A^k$. Além disso, para quaisquer funções $f, g \in C^\infty(I)$ e escalar α , $(f + g)(A) = f(A) + g(A)$, $(f g)(A) = f(A)g(A)$ (verifique!).

Sistemas de Equações Diferenciais Lineares - caso em que os autovalores são reais

Considere o sistema de equações diferenciais lineares

$$\begin{cases} \frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ \frac{dx_n}{dt} = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{cases}$$

Ele pode ser escrito na forma de uma equação diferencial vetorial

$$\frac{dX}{dt} = AX,$$

em que $A = (a_{ij})_{n \times n}$. Suponha que a matriz A seja semelhante a uma matriz na forma canônica de Jordan, ou seja, suponha que exista uma matriz P tal que $P^{-1}AP = J$, em que

$$J = \begin{bmatrix} J_{\lambda_1} & \bar{0} & \cdots & \bar{0} \\ \bar{0} & J_{\lambda_2} & \cdots & \bar{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{0} & \bar{0} & \cdots & J_{\lambda_k} \end{bmatrix} \quad \text{e} \quad J_{\lambda_i} = \begin{bmatrix} \lambda_i & 0 & \cdots & 0 & 0 \\ 1 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 0 \\ 0 & 0 & \cdots & 1 & \lambda_i \end{bmatrix}.$$

Vamos mostrar que $X(t) = e^{tA}X(0)$ é a solução do sistema. Para isso, vamos mostrar que $\frac{d}{dt}e^{tA} = e^{tA}A$ e que $X(t) = e^{tA}X(0)$ é a única solução do sistema que tem valor $X(0)$ em $t = 0$.

$$e^{tA} = P \begin{bmatrix} e^{tJ_{\lambda_1}} & \bar{0} & \cdots & \bar{0} \\ \bar{0} & e^{tJ_{\lambda_2}} & \cdots & \bar{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{0} & \bar{0} & \cdots & e^{tJ_{\lambda_k}} \end{bmatrix} P^{-1},$$

em que

$$e^{tJ_{\lambda_i}} = \begin{bmatrix} e^{t\lambda_i} & 0 & \cdots & 0 & 0 \\ te^{t\lambda_i} & e^{t\lambda_i} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{t^{m-2}}{(m-2)!}e^{t\lambda_i} & \frac{t^{m-3}}{(m-3)!}e^{t\lambda_i} & \cdots & e^{t\lambda_i} & 0 \\ \frac{t^{m-1}}{(m-1)!}e^{t\lambda_i} & \frac{t^{m-2}}{(m-2)!}e^{t\lambda_i} & \cdots & te^{t\lambda_i} & e^{t\lambda_i} \end{bmatrix}.$$

$$\begin{aligned}
\frac{d}{dt}e^{tJ_{\lambda_i}} &= \begin{bmatrix} \lambda_i e^{t\lambda_i} & 0 & \cdots & 0 & 0 \\ (\lambda_i t + 1)e^{t\lambda_i} & \lambda_i e^{t\lambda_i} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \left(\frac{\lambda_i t^{m-2}}{(m-2)!} + \frac{t^{m-3}}{(m-3)!}\right)e^{t\lambda_i} & \left(\frac{\lambda_i t^{m-3}}{(m-3)!} + \frac{t^{m-4}}{(m-4)!}\right)e^{t\lambda_i} & \cdots & \lambda_i e^{t\lambda_i} & 0 \\ \left(\frac{\lambda_i t^{m-1}}{(m-1)!} + \frac{t^{m-2}}{(m-2)!}\right)e^{t\lambda_i} & \left(\lambda_i \frac{t^{m-2}}{(m-2)!} + \frac{t^{m-3}}{(m-3)!}\right)e^{t\lambda_i} & \cdots & te^{t\lambda_i} & \lambda_i e^{t\lambda_i} \end{bmatrix} \\
&= \begin{bmatrix} e^{t\lambda_i} & 0 & \cdots & 0 & 0 \\ te^{t\lambda_i} & e^{t\lambda_i} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{t^{m-2}}{(m-2)!}e^{t\lambda_i} & \frac{t^{m-3}}{(m-3)!}e^{t\lambda_i} & \cdots & e^{t\lambda_i} & 0 \\ \frac{t^{m-1}}{(m-1)!}e^{t\lambda_i} & \frac{t^{m-2}}{(m-2)!}e^{t\lambda_i} & \cdots & te^{t\lambda_i} & e^{t\lambda_i} \end{bmatrix} \begin{bmatrix} \lambda_i & 0 & \cdots & 0 & 0 \\ 1 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 0 \\ 0 & 0 & \cdots & 1 & \lambda_i \end{bmatrix} \\
&= e^{tJ_{\lambda_i}} J_{\lambda_i}.
\end{aligned}$$

Assim,

$$\begin{aligned}
 \frac{d}{dt}e^{tA} &= P \begin{bmatrix} \frac{d}{dt}e^{tJ_{\lambda_1}} & \bar{0} & \cdots & \bar{0} \\ \bar{0} & \frac{d}{dt}e^{tJ_{\lambda_2}} & \cdots & \bar{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{0} & \bar{0} & \cdots & \frac{d}{dt}e^{tJ_{\lambda_k}} \end{bmatrix} P^{-1} \\
 &= P \begin{bmatrix} e^{tJ_{\lambda_1}} J_{\lambda_1} & \bar{0} & \cdots & \bar{0} \\ \bar{0} & e^{tJ_{\lambda_2}} J_{\lambda_2} & \cdots & \bar{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{0} & \bar{0} & \cdots & e^{tJ_{\lambda_k}} J_{\lambda_k} \end{bmatrix} P^{-1} \\
 &= P \begin{bmatrix} e^{tJ_{\lambda_1}} & \bar{0} & \cdots & \bar{0} \\ \bar{0} & e^{tJ_{\lambda_2}} & \cdots & \bar{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{0} & \bar{0} & \cdots & e^{tJ_{\lambda_k}} \end{bmatrix} \begin{bmatrix} J_{\lambda_1} & \bar{0} & \cdots & \bar{0} \\ \bar{0} & J_{\lambda_2} & \cdots & \bar{0} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{0} & \bar{0} & \cdots & J_{\lambda_k} \end{bmatrix} P^{-1} \\
 &= e^{tA} A
 \end{aligned}$$

Seja $X(t)$ uma solução do sistema tal que em $t = 0$ vale $X(0)$. Então, $Z(t) = e^{-tA}X(t)$ é tal que

$$\frac{d}{dt}Z(t) = \frac{d}{dt}e^{-tA}X(t) + e^{-tA}\frac{d}{dt}X(t) = -e^{-tA}AX(t) + e^{-tA}AX(t) = \bar{0}.$$

Aqui usamos o fato de que $\frac{d}{dt}(Y(t)X(t)) = \frac{dY}{dt}(t)X(t) + Y(t)\frac{dX}{dt}(t)$ (verifique!).

Logo, $Z(t) = Z(0) = e^{-0A}X(0) = X(0)$, para todo $t \in \mathbb{R}$ e

$$X(t) = e^{tA}e^{-tA}X(t) = e^{tA}Z(t) = e^{tA}X(0),$$

pois $e^{tA}e^{-tA} = I_n$ (verifique!). Portanto, $X(t) = e^{tA}X(0)$ é a única solução do sistema de equações diferenciais lineares

$$\frac{dX}{dt} = AX,$$

com o valor pré-determinado em $t = 0$.

Exemplo 4.22. Considere o sistema de equações diferenciais lineares

$$\begin{cases} \frac{dx}{dt} = 4x - 3y - z \\ \frac{dy}{dt} = 0 \\ \frac{dz}{dt} = x - 3y + 2z \end{cases}$$

Ele pode ser escrito na forma de uma equação diferencial vetorial

$$\frac{dX}{dt} = AX,$$

em que $A = \begin{bmatrix} 4 & -3 & -1 \\ 0 & 0 & 0 \\ 1 & -3 & 2 \end{bmatrix}$. Seja $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definido por $T(X) = AX$, para

todo $X \in \mathbb{R}^3$. O seu polinômio característico é $p(\lambda) = -\lambda(\lambda - 3)^2$. Portanto, os autovalores de T são $\lambda_1 = 0$ e $\lambda_2 = 3$.

Vamos usar o Teorema de Cayley-Hamilton para calcular e^{tA} . Seja $r_t(\lambda)$ um polinômio de grau menor que 3 tal que $e^{t\lambda} = r_t(\lambda)$, para $\lambda = 0$ e 3 e que $\frac{d}{d\lambda}e^{t\lambda} = r'_t(\lambda)$,

para $\lambda = 3$. Seja $q_t : \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$q_t(\lambda) = \begin{cases} \frac{e^{t\lambda} - r_t(\lambda)}{p(\lambda)} & \text{se } \lambda \neq 0, 3, \\ \lim_{h \rightarrow 1} \frac{e^{th} - r_t(h)}{p(h)} = \lim_{h \rightarrow 1} \frac{te^{th} - r'_t(h)}{p'(h)} & \text{se } \lambda = 0, \\ \lim_{h \rightarrow 3} \frac{e^{th} - r_t(h)}{p(h)} = \lim_{h \rightarrow 3} \frac{(1+t)e^{th} - r''_t(h)}{p''(h)} & \text{se } \lambda = 3. \end{cases}$$

As funções $q_t(\lambda)$ e $r_t(\lambda)$ pertencem a $C^\infty(\mathbb{R})$ e são tais que

$$e^{t\lambda} = q_t(\lambda)p(\lambda) + r_t(\lambda), \quad \text{em que } r_t(\lambda) \text{ é um polinômio de grau menor que 3.} \quad (4.27)$$

Usando o fato de que os autovalores de T também anulam $p(\lambda)$ temos duas equações, pois um dos autovalores é uma raiz dupla de $p(\lambda)$. Mas, quando um polinômio tem uma raiz dupla ela é também raiz da sua derivada. Sejam $a, b, c \in \mathbb{R}$ tais que $r_t(x) = ax^2 + bx + c$. Substituindo-se $\lambda = 0$ e 3 em (4.27), derivando-se (4.27) e substituindo-se $\lambda = 3$ obtemos o sistema

$$\begin{cases} 1 & = & c \\ e^{3t} & = & 9a + 3b + c \\ te^{3t} & = & 6a + b \end{cases}$$

cuja solução é $a = \frac{1+(3t-1)e^{3t}}{9}$, $b = \frac{-2+(-3t_2)e^{3t}}{3}$ e $c = 1$. Assim, pelo Teorema de Cayley-Hamilton, temos que

$$\begin{aligned} e^{tA} &= q(A)p(A) + \frac{1+(3t-1)e^{3t}}{9}A^2 + \frac{-2+(-3t_2)e^{3t}}{3}A + I_3 \\ &= \frac{1+(3t-1)e^{3t}}{9}A^2 + \frac{-2+(-3t_2)e^{3t}}{3}A + I_3 \\ &= \begin{bmatrix} (1+t)e^{3t} & 1-e^{3t} & -te^{3t} \\ 0 & 1 & 0 \\ te^{3t} & 1-e^{3t} & (1-t)e^{3t} \end{bmatrix}. \end{aligned}$$

Assim, a solução do sistema para $x(0), y(0)$ e $z(0)$ dados é

$$\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} (1+t)e^{3t} & 1 - e^{3t} & -te^{3t} \\ 0 & 1 & 0 \\ te^{3t} & 1 - e^{3t} & (1-t)e^{3t} \end{bmatrix} \begin{bmatrix} x(0) \\ y(0) \\ z(0) \end{bmatrix}$$

Sistemas de Equações Diferenciais Lineares - caso em que existem autovalores complexos

Vamos resolver inicialmente o caso em que a matriz A é 2×2 . Se o polinômio característico $p(\lambda) = \det(A - \lambda I_2)$ tem uma raiz complexa, então ele tem duas raízes complexas conjugadas: $\lambda_1 = \alpha + i\beta$ e $\lambda_2 = \alpha - i\beta$ e pelo [Exercício 4.1.27 na página 326](#), existe uma matriz P tal que

$$A = PD_{\alpha,\beta}P^{-1}, \quad \text{em que} \quad D_{\alpha,\beta} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}.$$

Definimos

$$e^{tD_{\alpha,\beta}} = e^{t\alpha} \begin{bmatrix} \cos t\beta & -\text{sen } t\beta \\ \text{sen } t\beta & \cos t\beta \end{bmatrix} \quad \text{e} \quad e^{tA} = Pe^{tD_{\alpha,\beta}}P^{-1}.$$

Vamos mostrar que $\frac{d}{dt}e^{tA} = e^{tA}A$.

$$\begin{aligned} \frac{d}{dt}e^{tD_{\alpha,\beta}} &= \alpha e^{t\alpha} \begin{bmatrix} \cos t\beta & -\text{sen } t\beta \\ \text{sen } t\beta & \cos t\beta \end{bmatrix} + \beta e^{t\alpha} \begin{bmatrix} -\text{sen } t\beta & -\cos t\beta \\ \cos t\beta & -\text{sen } t\beta \end{bmatrix} \\ &= e^{t\alpha} \begin{bmatrix} \alpha \cos t\beta - \beta \text{sen } t\beta & -\beta \cos t\beta - \alpha \text{sen } t\beta \\ \alpha \text{sen } t\beta + \beta \cos t\beta & -\beta \text{sen } t\beta + \alpha \cos t\beta \end{bmatrix} \\ &= e^{t\alpha} \begin{bmatrix} \cos t\beta & -\text{sen } t\beta \\ \text{sen } t\beta & \cos t\beta \end{bmatrix} \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \\ &= e^{tD_{\alpha,\beta}}D_{\alpha,\beta} \end{aligned}$$

Assim,

$$\frac{d}{dt}e^{tA} = P \frac{d}{dt}e^{tD_{\alpha,\beta}}P^{-1} = Pe^{tD_{\alpha,\beta}}D_{\alpha,\beta}P^{-1} = e^{tA}A.$$

Para o caso geral, vamos usar o [Exercício 4.4.6 na página 395](#), ou seja, existe uma matriz P tal que

$$PAP^{-1} = J = \begin{bmatrix} J_{\lambda_1} & \bar{0} & \cdots & \bar{0} & \bar{0} \\ \bar{0} & J_{\lambda_2} & \cdots & \bar{0} & \bar{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \bar{0} & \bar{0} & \cdots & J_{\alpha_{k-1}\beta_{k-1}} & \bar{0} \\ \bar{0} & \bar{0} & \cdots & \bar{0} & J_{\alpha_k\beta_k} \end{bmatrix}, \text{ em que}$$

$$J_{\lambda_j} = \begin{bmatrix} \lambda_j & 0 & \cdots & 0 & 0 \\ 1 & \lambda_j & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_j & 0 \\ 0 & 0 & \cdots & 1 & \lambda_j \end{bmatrix},$$

$$J_{\alpha_i\beta_i} = \begin{bmatrix} D_{\alpha_i\beta_i} & \bar{0} & \cdots & \bar{0} & \bar{0} \\ I_2 & D_{\alpha_i\beta_i} & \cdots & \bar{0} & \bar{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \bar{0} & \bar{0} & \cdots & D_{\alpha_i\beta_i} & \bar{0} \\ \bar{0} & \bar{0} & \cdots & I_2 & D_{\alpha_i\beta_i} \end{bmatrix} \text{ e } D_{\alpha_i\beta_i} = \begin{bmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{bmatrix}.$$

Definindo

$$e^{tJ_{\alpha_i\beta_i}} = \begin{bmatrix} e^{tD_{\alpha_i\beta_i}} & 0 & \cdots & 0 & 0 \\ te^{tD_{\alpha_i\beta_i}} & e^{tD_{\alpha_i\beta_i}} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{t^{m-2}}{(m-2)!}e^{tD_{\alpha_i\beta_i}} & \frac{t^{m-3}}{(m-3)!}e^{tD_{\alpha_i\beta_i}} & \cdots & e^{tD_{\alpha_i\beta_i}} & 0 \\ \frac{t^{m-1}}{(m-1)!}e^{tD_{\alpha_i\beta_i}} & \frac{t^{m-2}}{(m-2)!}e^{tD_{\alpha_i\beta_i}} & \cdots & te^{tD_{\alpha_i\beta_i}} & e^{tD_{\alpha_i\beta_i}} \end{bmatrix}$$

podemos definir

$$e^{tA} = P \begin{bmatrix} e^{tJ_{\lambda_1}} & \bar{0} & \cdots & \bar{0} & \bar{0} \\ \bar{0} & e^{tJ_{\lambda_2}} & \cdots & \bar{0} & \bar{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \bar{0} & \bar{0} & \cdots & e^{tJ_{\alpha_{k-1}\beta_{k-1}}} & \bar{0} \\ \bar{0} & \bar{0} & \cdots & \bar{0} & e^{tJ_{\alpha_k\beta_k}} \end{bmatrix} P^{-1}.$$

Exercícios Numéricos (respostas na página 492)

4.4.1. Quais das seguintes matrizes estão na forma canônica de Jordan.

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

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

4.4.2. Determine, se possível, uma matriz P , tal que $P^{-1}AP = J$ esteja na forma canônica 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 Teóricos

4.4.3. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador tal que seu polinômio característico é da forma

$$p(\lambda) = (-1)^n (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_k)^{n_k},$$

com $\lambda_1, \dots, \lambda_k$ distintos. Seja $\mathbb{W} = \mathcal{I}(T - \lambda_k)^{m_k}$. Mostre que:

- (a) \mathbb{W} é T -invariante.
- (b) $\mathbb{W} \cap \mathcal{N}(T - \lambda_i I)^{m_i}$ é invariante por $(T - \lambda_k I)^{m_k}$.
- (c) A restrição de $(T - \lambda_k I)^{m_k}$ a $\mathbb{W} \cap \mathcal{N}(T - \lambda_i I)^{m_i}$ é injetiva e portanto sobrejetiva.

4.4.4. Sejam $\mathcal{C}_1, \dots, \mathcal{C}_k$ ciclos de autovetores generalizados associados a um autovalor λ . Mostre que se os vetores finais dos ciclos são distintos, então os ciclos são disjuntos, isto é, $\mathcal{C}_i \cap \mathcal{C}_j = \emptyset$, para $i \neq j$.

4.4.5. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador tal que seu polinômio característico é da forma

$$p(\lambda) = (-1)^n (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_k)^{n_k},$$

com $\lambda_1, \dots, \lambda_k$ distintos. Mostre que:

- (a) T é diagonalizável se, e somente se, o autoespaço generalizado \mathbb{W}_i associado a λ_i é igual ao autoespaço $\mathcal{N}(T - \lambda_i I)$ associado a λ_i , para $i = 1, \dots, k$.
- (b) T é diagonalizável se, e somente se, $\mathcal{N}(T - \lambda_i I) = \mathcal{N}(T - \lambda_i I)^2$, para $i = 1, \dots, k$.

4.4.6. Seja $T : \mathbb{V} \rightarrow \mathbb{V}$ um operador linear em um espaço vetorial real. Mostre que existe uma base \mathcal{C} de \mathbb{V} tal que

$$[T]_{\mathcal{C}}^{\mathcal{C}} = J = \begin{bmatrix} J_{\lambda_1} & \bar{0} & \cdots & \bar{0} & \bar{0} \\ \bar{0} & J_{\lambda_2} & \cdots & \bar{0} & \bar{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \bar{0} & \bar{0} & \cdots & J_{\alpha_{k-1}\beta_{k-1}} & \bar{0} \\ \bar{0} & \bar{0} & \cdots & \bar{0} & J_{\alpha_k\beta_k} \end{bmatrix}, \quad \text{em que}$$

$$J_{\lambda_j} = \begin{bmatrix} \lambda_j & 0 & \cdots & 0 & 0 \\ 1 & \lambda_j & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_j & 0 \\ 0 & 0 & \cdots & 1 & \lambda_j \end{bmatrix},$$

$$J_{\alpha_i \beta_i} = \begin{bmatrix} D_{\alpha_i \beta_i} & \bar{0} & \cdots & \bar{0} & \bar{0} \\ I_2 & D_{\alpha_i \beta_i} & \cdots & \bar{0} & \bar{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \bar{0} & \bar{0} & \cdots & D_{\alpha_i \beta_i} & \bar{0} \\ \bar{0} & \bar{0} & \cdots & I_2 & D_{\alpha_i \beta_i} \end{bmatrix} \text{ e } D_{\alpha_i \beta_i} = \begin{bmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{bmatrix}.$$

(Sugestão: considere a complexificação de \mathbb{V} e do operador T , [Exercício 4.1.25 na página 326.](#))

4.4.7. Seja $J = \begin{bmatrix} \lambda & 0 & \cdots & 0 & 0 \\ 1 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 0 \\ 0 & 0 & \cdots & 1 & \lambda \end{bmatrix}$ um bloco de Jordan $m \times m$ correspondente a λ . Seja $N = J - \lambda I_m$.

Mostre que:

(a) $N^m = O$, a transformação linear nula e para $1 \leq k < m$, $[N^k]_{ij} = \begin{cases} 1 & \text{se } i - j = k \\ 0 & \text{caso contrário} \end{cases}$

(b) $J^k = \begin{cases} \sum_{j=0}^k \binom{k}{j} \lambda^{k-j} N^j & \text{para } 0 < k < m. \\ \sum_{j=0}^{m-1} \binom{k}{j} \lambda^{k-j} N^j & \text{para } k \geq m. \end{cases}$

(c) Para todo inteiro $k \geq m$,

$$J^k = \begin{bmatrix} \lambda^k & 0 & \cdots & 0 & 0 \\ k\lambda^{k-1} & \lambda^k & \cdots & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ \binom{k}{m-2}\lambda^{k-m+2} & & & \lambda^k & 0 \\ \binom{k}{m-1}\lambda^{k-m+1} & \binom{k}{m-2}\lambda^{k-m+2} & \cdots & k\lambda^{k-1} & \lambda^k \end{bmatrix}$$

Respostas dos Exercícios

1.1. Definição e Exemplos (página 12)

1.1.1. 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 $X = \frac{5}{4}U - \frac{1}{6}V$.

1.1.2. 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 $Y = \frac{1}{4}U + \frac{1}{2}V$.

1.1.3. 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

$$\begin{cases} a & & b & = & 1 \\ a & + & b & = & 2 \\ a & + & 3b & = & 7 \end{cases}$$

que tem solução $a = 1$ e $b = 2$.

1.1.4. 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 $3 = -5a + 5a + 2b = 0$, ou seja, basta que $a = -3/5$ e $b = 3/2$.

```

1.1.5. >> 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 .

- 1.1.6. (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. $1x = x^1 = x$.

- 1.1.7. (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 $x \in \mathcal{X}$;

1.2. Subespaços (página 39)

- 1.2.1. (a) Não é subespaço pois $(1, 1, 1)$ pertence ao conjunto, mas $2(1, 1, 1) = (2, 2, 2)$ não pertence;
- (b) É um subespaço pois é um plano que passa pela origem;
- (c) Não é um subespaço pois $(1, 1, 1)$ pertence ao conjunto, mas $-1(1, 1, 1) = (-1, -1, -1)$ não pertence;
- (d) Não é subespaço pois $(1, 0, 0)$ e $(0, -1, 0)$ pertencem ao conjunto, mas $(1, 0, 0) + (0, -1, 0) = (1, -1, 0)$ não pertence;

- (e) É um subespaço pois é uma reta que passa pela origem (eixo y);
- (f) É um subespaço pois é um plano que passa pela origem;
- (g) Não é um subespaço pois é um plano que não passa pela origem.
- (h) Não é um subespaço pois $(1, 1, 2)$ e $(-1, -1, 2)$ pertencem ao conjunto, mas $(1, 1, 2) + (-1, -1, 2) = (0, 0, 4)$ não pertencem.

1.2.2. (a) Não é subespaço pois $(2, 0, 0, 0)$ pertence ao conjunto, mas $(0, 0, 0, 0) = 0(2, 0, 0, 0)$ não pertence;

(b) É um subespaço. Pois,

(0) Se $X_1 = (x_1, y_1, z_1, w_1)$ e $X_2 = (x_2, y_2, z_2, w_2)$ pertencem ao conjunto, então $z_1 = x_1 = 2y_1$, $z_2 = x_2 = 2y_2$ e $w_1 = x_1 - 3y_1$, $w_2 = x_2 - 3y_2$. O que implica que $X_1 + X_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2, w_1 + w_2)$ também pertence ao conjunto já que $z_1 + z_2 = x_1 + x_2 = 2(y_1 + y_2)$ e $w_1 + w_2 = (x_1 + x_2) - 3(y_1 + y_2)$.

(0') Se $X = (x, y, z, w)$ pertence ao conjunto, então $z = x = 2y$ e $w = x - 3y$. O que implica que $\alpha X = (\alpha x, \alpha y, \alpha z, \alpha w)$ também pertence, pois $\alpha z = \alpha x = 2(\alpha y)$ e $\alpha w = (\alpha x) - 3(\alpha y)$.

(c) É um subespaço. Pois,

(0) Se $X_1 = (x_1, y_1, z_1, w_1)$ e $X_2 = (x_2, y_2, z_2, w_2)$ pertencem ao conjunto, então $x_1 = y_1 = 0$ e $x_2 = y_2 = 0$. O que implica que $X_1 + X_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2, w_1 + w_2)$ também pertence ao conjunto já que $x_1 + x_2 = y_1 + y_2 = 0$.

(0') Se $X = (x, y, z, w)$ pertence ao conjunto, então $x = y = 0$. O que implica que $\alpha X = (\alpha x, \alpha y, \alpha z, \alpha w)$ também pertence, pois $\alpha x = \alpha y = 0$.

(d) É um subespaço. Pois,

(0) Se $X_1 = (x_1, y_1, z_1, w_1)$ e $X_2 = (x_2, y_2, z_2, w_2)$ pertencem ao conjunto, então $x_1 = 0$, $x_2 = 0$, $y_1 = -w_1$ e $y_2 = -w_2$. O que implica que $X_1 + X_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2, w_1 + w_2)$ também pertence ao conjunto já que $x_1 + x_2 = 0$ e $y_1 + y_2 = -(w_1 + w_2)$.

(0') Se $X = (x, y, z, w)$ pertence ao conjunto, então $x = 0$ e $y = -w$. O que implica que $\alpha X = (\alpha x, \alpha y, \alpha z, \alpha w)$ também pertence, pois $\alpha x = 0$ e $\alpha y = -(\alpha w)$.

1.2.3. (a) É um subespaço de $C^0[-1, 1]$ pois se $f(-1) = f(1)$ e $g(-1) = g(1)$, então $(f + g)(-1) = f(-1) + g(-1) = f(1) + g(1) = (f + g)(1)$ e $(\alpha f)(-1) = \alpha f(-1) = \alpha f(1) = (\alpha f)(1)$. Portanto, é um espaço vetorial.

(b) Não é um subespaço de $C^0[0, 1]$ pois $f(x) = x$ é uma função não decrescente, mas $((-1)f)(x) = -x$ não é uma função não decrescente. Portanto, não é um espaço vetorial.

(c) Não é um subespaço de $C^0[-1, 1]$ pois se $f(x) = 1 + x$ e $g(x) = 1 - x$, então $(f + g)(x) = 1$ não pertence ao conjunto. Portanto, não é um espaço vetorial.

(d) É um subespaço de $C^0[-1, 1]$ pois se $f(-1) = f(1) = 0$ e $g(-1) = g(1) = 0$, então $(f + g)(-1) = f(-1) + g(-1) = 0 = f(1) + g(1) = (f + g)(1)$ e $(\alpha f)(-1) = \alpha f(-1) = 0 = \alpha f(1) = (\alpha f)(1)$. Portanto, é um espaço vetorial.

(e) Não é um subespaço do espaço dos polinômios pois o polinômio nulo $\bar{0} = 0(t^3)$ não tem grau igual a 3. Portanto, não é um espaço vetorial.

1.2.4. (a) É um subespaço do espaço das matrizes \mathcal{M}_{mn} pois se B_1 e B_2 são tais que $AB_1 = B_1A$ e $AB_2 = B_2A$, então $A(B_1 + B_2) = AB_1 + AB_2 = B_1A + B_2A = (B_1 + B_2)A$ e se B é tal que $AB = BA$, então $A(\alpha B) = \alpha AB = \alpha BA = (\alpha B)A$. Portanto é um espaço vetorial.

(b) Não é um subespaço do espaço das matrizes \mathcal{M}_{mn} , pois se B é tal que $AB \neq BA$, então $\bar{0} = 0B$ é tal que $AB = BA = \bar{0}$. Portanto não é um espaço vetorial.

(c) É um subespaço do espaço das matrizes \mathcal{M}_{mn} pois se B_1 e B_2 são tais que $B_1A = \bar{0}$ e $B_2A = \bar{0}$, então $(B_1 + B_2)A = B_1A + B_2A = \bar{0} + \bar{0} = \bar{0}$ e se B é tal que $BA = \bar{0}$, então $(\alpha B)A = \alpha(BA) = \alpha\bar{0} = \bar{0}$. Portanto é um espaço vetorial.

1.2.5. Sejam $f, g \in \mathcal{F}(\mathcal{X}; \mathbb{R})$ tais que $f(x_0) = g(x_0) = 0$.

$$(0) (f + g)(x_0) = f(x_0) + g(x_0) = 0;$$

$$(0') (\alpha f)(x_0) = \alpha f(x_0) = \alpha 0 = 0.$$

1.2.6. Seja $V = (a, b, c)$ um vetor qualquer de \mathbb{R}^3 . Vamos mostrar que existem $V_1 \in \mathbb{W}_1$ e $V_2 \in \mathbb{W}_2$ tais que $V = V_1 + V_2$. Um elemento qualquer de \mathbb{W}_1 é da forma (z, z, z) enquanto um elemento qualquer de \mathbb{W}_2 é da forma $(x, y, 0)$. Assim, precisamos encontrar x, y e z tais que

$$(a, b, c) = (z, z, z) + (x, y, 0) = (x + z, y + z, z)$$

que é equivalente ao sistema

$$\begin{cases} x + z = a \\ y + z = b \\ z = c \end{cases}$$

que tem solução única $x = a - c$, $y = b - c$ e $z = c$. Portanto, $\mathbb{R}^3 = \mathbb{W}_1 \oplus \mathbb{W}_2$ pois todo elemento de \mathbb{R}^3 se escreve de maneira única como a soma de um elemento de \mathbb{W}_1 e um elemento de \mathbb{W}_2 .

1.2.7. (a)

```
>> syms a b c d
>> A=[1,0,0,1;0,1,0,0;1,1,1,1;...
1,1,1,0;a,b,c,d].'
```

$$\begin{bmatrix} 1 & 0 & 1 & 1 & a \\ 0 & 1 & 1 & 1 & b \\ 0 & 0 & 1 & 1 & c \\ 1 & 0 & 1 & 0 & d \end{bmatrix}$$

```
>> escalona(A);
[ 1 0 0 0 a - c ]
```

$$\begin{bmatrix} 0 & 1 & 0 & 0 & b - c \\ 0 & 0 & 1 & 0 & c + d - a \\ 0 & 0 & 0 & 1 & -d + a \end{bmatrix}$$

Portanto, os vetores da letra (a) geram o \mathbb{R}^4 .

```
(b) >> A=[1,2,1,0;1,1,-1,0;0,0,0,1;a,b,c,d].'
```

$$\begin{bmatrix} 1 & 1 & 0 & a \\ 2 & 1 & 0 & b \\ 1 & -1 & 0 & c \\ 0 & 0 & 1 & d \end{bmatrix}$$

```
>> escalona(A);
```

$$\begin{bmatrix} 1 & 0 & 0 & -a + b \\ 0 & 1 & 0 & -b + 2a \\ 0 & 0 & 1 & d \\ 0 & 0 & 0 & c + 3a - 2b \end{bmatrix}$$

Continua ? (s/n) n

Portanto, os vetores da letra (b) **não** geram o \mathbb{R}^4 .

```
(c) >> A=[6,4,-2,4;2,0,0,1;3,2,-1,2;...
5,6,-3,2;0,4,-2,-2;a,b,c,d].'
```

$$\begin{bmatrix} 6 & 2 & 3 & 5 & 0 & a \\ 4 & 0 & 2 & 6 & 4 & b \\ -2 & 0 & -1 & -3 & -2 & c \\ 4 & 1 & 2 & 2 & -2 & d \end{bmatrix}$$

```
>> escalona(A);
```

$$\begin{bmatrix} 1 & 0 & 1/2 & 3/2 & 1 & 1/4 b \\ 0 & 1 & 0 & -2 & -3 & -3/4 b + 1/2 a \\ 0 & 0 & 0 & -2 & -3 & d - 1/2 a - 1/4 b \\ 0 & 0 & 0 & 0 & 0 & c + 1/2 b \end{bmatrix}$$

Continua ? (s/n) n

Portanto, os vetores da letra (c) **não** geram o \mathbb{R}^4 .

```
(d) >> A=[1,1,0,0;1,2,-1,1;...
0,0,1,1;2,1,1,1;a,b,c,d].'
```

$$A = \begin{bmatrix} 1 & 1 & 0 & 2 & a \\ 1 & 2 & 0 & 1 & b \\ 0 & -1 & 1 & 1 & c \\ 0 & 1 & 1 & 1 & d \end{bmatrix}$$

```
>> escalona(A);
```

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -a + 2b - 3/2 d + 3/2 c \\ 0 & 1 & 0 & 0 & 1/2 d - 1/2 c \\ 0 & 0 & 1 & 0 & c + b - a \\ 0 & 0 & 0 & 1 & 1/2 d - b + a - 1/2 c \end{bmatrix}$$

Portanto, os vetores da letra (d) geram o \mathbb{R}^4 .

1.2.8. (a) `>> A=[1,0,1,0,0;1,2,3,1,0;...`
`2,1,3,1,0]`

$$\begin{array}{ccccc} 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 1 & 0 \\ 2 & 1 & 3 & 1 & 0 \end{array}$$

`>> R=escalona(A)`

$$\begin{array}{ccccc} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}$$

Encontramos a forma reduzida escalonada da matriz $[A \mid \vec{0}]$, que corresponde ao sistema

$$\begin{cases} x_1 + x_3 = 0 \\ x_2 + x_3 = 0 \\ x_4 = 0 \end{cases}$$

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]`

$$\begin{array}{ccccc} 1 & 1 & 2 & -1 & 0 \\ 2 & 3 & 6 & -2 & 0 \\ -2 & 1 & 2 & 2 & 0 \end{array}$$

`>> R=escalona(A)`

$$\begin{array}{ccccc} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}$$

Encontramos a forma reduzida escalonada da matriz $[A \mid \vec{0}]$, que corresponde ao sistema

$$\begin{cases} x_1 + x_2 - x_4 = 0 \\ x_2 + 2x_3 = 0 \end{cases}$$

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, $\mathcal{S} = \{V_1 = (1, 0, 0, 1), V_2 = (0, -2, 1, 0)\}$ gera \mathbb{W} .

1.2.9. O subespaço \mathbb{V} é um plano que passa pela origem, paralelo aos vetores $(-1, 2, 3)$ e $(1, 3, 4)$. O subespaço \mathbb{W} é um plano que passa pela origem, paralelo aos vetores $(1, 2, -1)$ e $(0, 1, 1)$.

```
>> V1=[-1,2,3]; V2=[1,3,4];
>> N1=pv(V1,V2)
N1 =   -1    7   -5
>> V3=[1,2,-1]; V4=[0,1,1];
>> N2=pv(V3,V4)
N2 =    3   -1    1
>> V=pv(N1,N2)
V =    2  -14  -20
```

A equação paramétrica da reta interseção dos dois subespaços é $(x, y, z) = t(2, -14, -20)$, para qualquer $t \in \mathbb{R}$.

1.2.10. Por um lado, $1 = 1 \sin^2 t + 1 \cos^2 t$, $\cos 2t = (-1) \sin^2 t + 1 \cos^2 t$, implica que $\alpha 1 + \beta \cos 2t = (\alpha - \beta) \sin^2 t + (\alpha + \beta) \cos^2 t$, ou seja, toda combinação linear de 1 e $\cos 2t$ também é combinação linear de $\sin^2 t$ e de $\cos^2 t$. Por outro lado, $\sin^2 t = \frac{1}{2}(1) - \frac{1}{2} \cos 2t$ e $\cos^2 t = \frac{1}{2}(1) + \frac{1}{2} \cos 2t$ implica que $\alpha \sin^2 t + \beta \cos^2 t = \frac{\alpha + \beta}{2}(1) + \frac{\beta - \alpha}{2} \cos 2t$.

1.2.11. (a) Podemos, por exemplo, tirar c em função de a e b na equação $3a - 5b + 2c = 0$ obtendo

$$c = \frac{1}{2}(5b - 3a).$$

Substituindo o valor de c em um ponto qualquer do plano obtemos que os vetores do plano são da forma

$$(a, b, c) = (a, b, \frac{1}{2}(5b - 3a)) = a(1, 0, -\frac{3}{2}) + b(0, 1, \frac{5}{2}).$$

Assim, os vetores $V_1 = (1, 0, -\frac{3}{2})$ e $V_2 = (0, 1, \frac{5}{2})$ geram o subespaço.

(b) Podemos, por exemplo, obter que $a_{12} = \frac{3}{2}a_{11}$. Substituindo o valor de a_{12} obtido em uma matriz qualquer do subespaço obtemos que as matrizes do subespaço são da forma

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & \frac{3}{2}a_{11} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} \begin{pmatrix} 1 & \frac{3}{2} \\ 0 & 0 \end{pmatrix} + a_{21} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + a_{22} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Assim as matrizes $M_1 = \begin{pmatrix} 1 & \frac{3}{2} \\ 0 & 0 \end{pmatrix}$, $M_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ e $M_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ geram o subespaço

- (c) Um polinômio $p(t) = at^3 + bt^2 + ct + d$ pertence ao subespaço se, e somente se, $p(2) = 0$, ou seja, se $a2^3 + b2^2 + c2 + d = 0$. Podemos, por exemplo, tirar o valor de d em função de a, b e c obtendo $d = -8a - 4b - 2c$. Substituindo este valor de d em um polinômio qualquer do subespaço obtemos

$$p(t) = at^3 + bt^2 + ct + d = at^3 + bt^2 + ct + (-8a - 4b - 2c) = a(t^3 - 8) + b(t^2 - 4) + c(t - 2).$$

Assim, os polinômios $t^3 - 8$, $t^2 - 4$ e $t - 2$ geram o subespaço.

- (d) Um polinômio $p(t) = at^3 + bt^2 + ct + d$ pertence ao subespaço se, e somente se, $p(2) = p(-1)$, ou seja, se $a2^3 + b2^2 + c2 + d = a(-1)^3 + b(-1)^2 + c(-1) + d$ ou ainda, $9a + 3b + 3c = 0$. Podemos, por exemplo, tirar o valor de c em função de a e b obtendo $c = -3a - b$. Substituindo este valor de c em um polinômio qualquer do subespaço obtemos

$$p(t) = at^3 + bt^2 + ct + d = at^3 + bt^2 + (-3a - b)t + d = a(t^3 - 3t) + b(t^2 - 1) + d(1).$$

Assim, os polinômios $t^3 - 3t$, $t^2 - 1$ e $t - 2$ geram o subespaço.

- 1.2.12. Seja $p(t) = \alpha_0 + \dots + \alpha_n t^n$ um polinômio qualquer. Como $p(t)$ tem grau n os polinômios de grau maior que n não entram na combinação linear. Vamos mostrar que a equação

$$p(t) = \alpha_0 + \dots + \alpha_n t^n = x_0(2) + x_1(t+1) + \dots + x_n(t^n + 1)$$

tem solução. Agrupando os termos de mesmo grau obtemos que a equação acima é equivalente ao sistema

$$\begin{cases} 2x_0 + x_1 + \dots + x_n = \alpha_0 \\ x_1 + \dots = \alpha_1 \\ \dots \\ x_n = \alpha_n \end{cases} \quad \text{que tem solução. Portanto o conjunto dado gera o espaço dos polinômios.}$$

- 1.2.13. Seja $p(t) = \alpha_0 + \dots + \alpha_n t^n$ um polinômio qualquer tal que $p(-t) = p(t)$. Então, $p(t) - p(-t) = 2a_1t + 2a_3t^3 + \dots + 2a_{2k+1}t^{2k+1} = 0$, para todo $t \in \mathbb{R}$. O que implica que todos os coeficientes de p com índice ímpar são iguais a zero. Assim, todo polinômio do subespaço é da forma $p(t) = a_0 + a_2t^2 + \dots + a_{2k}t^{2k}$, para algum $k \in \mathbb{N}$. Portanto $1, t^2, t^4, \dots, t^{2n}, \dots$ geram o subespaço dos polinômios pares.

1.3. Dependência Linear (página 61)

- 1.3.1. (a) >> v1=[1,1,2,1];v2=[1,0,0,2];
>> v3=[4,6,8,6];v4=[0,3,2,1];
>> A=[v1;v2;v3;v4;zeros(1,4)].
1 1 4 0 0
1 0 6 3 0
2 0 8 2 0
1 2 6 1 0
>> R=escalona(A)

```

1   0   0  -3   0
0   1   0  -1   0
0   0   1   1   0
0   0   0   0   0

```

Logo, o sistema $x_1(1, 1, 2, 1) + x_2(1, 0, 0, 2) + x_3(4, 6, 8, 6) + x_4(0, 3, 2, 1) = \vec{0}$ admite solução não trivial. Isto implica que os vetores da letra (a) são L.D.

```

(b) >> v1=[1,-2,3,-1];v2=[-2,4,-6,2];
>> A=[v1;v2;zeros(1,4)].'
      1   -2   0
     -2   4   0
      3  -6   0
     -1   2   0
>> R=escalona(A)
      1   -2   0
      0   0   0
      0   0   0
      0   0   0

```

Logo, o sistema $x_1(1, -2, 3, -1) + x_2(-2, 4, -6, 2) = \vec{0}$ admite solução não trivial. Isto implica que os vetores da letra (b) são L.D. Observe que o segundo vetor é -2 vezes o primeiro.

```

(c) >> v1=[1,1,1,1];v2=[2,3,1,2];
>> v3=[3,1,2,1];v4=[2,2,1,1];
>> A=[v1;v2;v3;v4;zeros(1,4)].'
      1   2   3   2   0
      1   3   1   2   0
      1   1   2   1   0
      1   2   1   1   0
>> R=escalona(A)
      1   0   0   0   0
      0   1   0   0   0
      0   0   1   0   0
      0   0   0   1   0

```

Logo, o sistema $x_1(1, 1, 1, 1) + x_2(2, 3, 1, 2) + x_3(3, 1, 2, 1) + x_4(2, 2, 1, 1) = \vec{0}$ só admite a solução trivial. Isto implica que os vetores da letra (c) são L.I.

```

(d) >> v1=[4,2,-1,3];v2=[6,5,-5,1];v3=[2,-1,3,5];
>> A=[v1;v2;v3;zeros(1,4)].'
      4   6   2   0
      2   5  -1   0
     -1  -5   3   0
      3   1   5   0
>> R=escalona(A)

```

$$\begin{array}{cccc} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Logo, o sistema $x_1(4, 2, -1, 3) + x_2(2, 3, 1, 2) + x_3(2, -1, 3, 5) = \vec{0}$ admite solução não trivial. Isto implica que os vetores da letra (d) são L.D.

```
1.3.2. >> 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=0','a')
ans = [ 2] [-2]
```

1.3.3. (a) $x_1(t^2 - 2t + 3) + x_2(2t^2 + t + 8) + x_3(t^2 + 8t + 7) = \vec{0}(t) = 0$, para todo $t \in \mathbb{R}$. Agrupando os termos correspondentes a mesma potência de t obtemos

$(x_1 + 2x_2 + x_3)t^2 + (-2x_1 + x_2 + 8x_3)t + (3x_1 + 8x_2 + 7x_3) = 0$, para todo $t \in \mathbb{R}$. Como o polinômio nulo tem todos os seus coeficientes iguais a zero, obtemos o sistema linear

$$\begin{cases} x_1 + 2x_2 + x_3 = 0 \\ -2x_1 + x_2 + 8x_3 = 0 \\ 3x_1 + 8x_2 + 7x_3 = 0 \end{cases}$$

```
>> A=[1,2,1,0;-2,1,8,0;3,8,7,0]
>> escalona(A)
[ 1, 2, 1, 0]
[-2, 1, 8, 0]
```

```
[ 3, 8, 7, 0]
eliminação 1:
(2)*linha 1 + linha 2 ==> linha 2
(-3)*linha 1 + linha 3 ==> linha 3
[ 1, 2, 1, 0]
[ 0, 5, 10, 0]
[ 0, 2, 4, 0]
eliminação 2:
(1/5)*linha 2 ==> linha 2
[ 1, 2, 1, 0]
[ 0, 1, 2, 0]
[ 0, 2, 4, 0]
(-2)*linha 2 + linha 1 ==> linha 1
(-2)*linha 2 + linha 3 ==> linha 3
[ 1, 0, -3, 0]
[ 0, 1, 2, 0]
[ 0, 0, 0, 0]
```

A solução do sistema é $x_3 = \alpha$, $x_2 = -2\alpha$ e $x_1 = 3\alpha$, para todo $\alpha \in \mathbb{R}$. Portanto o sistema e a equação vetorial inicial possuem solução não trivial o que implica que os três polinômios são L.D.

- (b) $x_1(t^2 - 1) + x_2(t + 1) + x_3(t + 2) = \vec{0}(t) = 0$, para todo $t \in \mathbb{R}$. Agrupando os termos correspondentes a mesma potência de t obtemos

$(x_1)t^2 + (x_2 + x_3)t + (-x_1 + x_2 + 2x_3) = 0$, para todo $t \in \mathbb{R}$. Como o polinômio nulo tem todos os seus coeficientes iguais a

zero, obtemos o sistema linear $\begin{cases} x_1 & & & = 0 \\ -x_1 & + & x_2 & + & x_3 & = 0 \\ & & -x_1 & + & x_2 & + & 2x_3 & = 0 \end{cases}$

```
>> A=[1,0,0,0;0,1,1,0;-1,1,2,0]
>> escalona(A)
[ 1, 0, 0, 0]
[ 0, 1, 1, 0]
[-1, 1, 2, 0]
eliminação 1:
(1)*linha 1 + linha 3 ==> linha 3
[ 1, 0, 0, 0]
[ 0, 1, 1, 0]
[ 0, 1, 2, 0]
eliminação 2:
(-1)*linha 2 + linha 3 ==> linha 3
[ 1, 0, 0, 0]
[ 0, 1, 1, 0]
[ 0, 0, 1, 0]
eliminação 3:
(-1)*linha 3 + linha 2 ==> linha 2
```

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

Assim o sistema e a equação vetorial inicial possuem somente a solução trivial $x_1 = x_2 = x_3 = 0$. Portanto, os polinômios são L.I.

1.3.4. (a) $W[f_1, f_2, f_3](t) = \det \begin{bmatrix} t & \cos t & \sin t \\ 1 & -\sin t & \cos t \\ 0 & -\cos t & -\sin t \end{bmatrix}$ e $W[f_1, f_2, f_3](\pi) = \det \begin{bmatrix} \pi & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = \pi \neq 0$. Portanto, as funções são L.I.

(b) $W[f_1, f_2, f_3](t) = \det \begin{bmatrix} \cos t & 1 & \sin^2 t/2 \\ -\sin t & 0 & \frac{1}{2}\sin t \\ -\cos t & 0 & \frac{1}{2}\cos t \end{bmatrix} = -\det \begin{bmatrix} -\sin t & \frac{1}{2}\sin t \\ -\cos t & \frac{1}{2}\cos t \end{bmatrix} = 0$ para todo $t \in [-\pi, \pi]$. Assim, as funções podem ser L.D. Vamos tentar descobrir uma relação entre as funções. Observamos que $\cos t = \cos(2\frac{t}{2}) = \cos^2 t/2 - \sin^2 t/2 = 1 - 2\sin^2 t/2$. Assim, a primeira função é combinação linear das outras duas. Portanto as funções são L.D.

(c) $W[f_1, f_2, f_3](t) = \det \begin{bmatrix} 1 & e^t + e^{-t} & e^t - e^{-t} \\ 0 & e^t - e^{-t} & e^t + e^{-t} \\ 0 & e^t + e^{-t} & e^t - e^{-t} \end{bmatrix} = \det \begin{bmatrix} e^t - e^{-t} & e^t + e^{-t} \\ e^t + e^{-t} & e^t - e^{-t} \end{bmatrix} = -4$. Assim as funções são L.I.

1.3.5. Sejam $y_1(t) = e^t \cos 3t$ e $y_2(t) = e^t \sin 3t$. $y_1'(t) = e^t(\cos 3t - 3 \sin 3t)$, $y_2'(t) = e^t(\sin 3t + 3 \cos 3t)$. $y_1''(t) = -2e^t(4 \cos 3t + 3 \sin 3t)$, $y_2''(t) = 2e^t(-4 \sin 3t + 3 \cos 3t)$. Assim, $y_1'' - 2y_1' + 10y_1 = e^t 0 = 0$ e $y_2'' - 2y_2' + 10y_2 = e^t 0 = 0$, ou seja, y_1 e y_2 são soluções da equação diferencial. Agora, $W[y_1, y_2](t) = \det \begin{bmatrix} e^t \cos 3t & e^t \sin 3t \\ e^t(\cos 3t - 3 \sin 3t) & e^t(\sin 3t + 3 \cos 3t) \end{bmatrix}$ e $W[y_1, y_2](0) = \det \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix} = 3 \neq 0$. Portanto as funções são soluções L.I. da equação diferencial dada.

1.3.6. (a) $x_1 Y_1 + x_2 Y_2 + x_3 Y_3 = x_1(X_1 + X_2) + x_2(X_1 + X_3) + x_3(X_2 + X_3) = (x_1 + x_2)X_1 + (x_1 + x_3)X_2 + (x_2 + x_3)X_3 = \vec{0}$. Como X_1, X_2 e X_3 são por hipótese L.I., os escalares que os estão multiplicando têm que ser iguais a zero. O que leva ao sistema

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_3 = 0 \\ x_2 + x_3 = 0 \end{cases}$$

```
>> 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 Y_1, Y_2 e Y_3 são L.I.

- (b) $x_1Y_1 + x_2Y_2 + x_3Y_3 = x_1X_1 + x_2(X_1 + X_3) + x_3(X_1 + X_2 + X_3) = (x_1 + x_2 + x_3)X_1 + x_3X_2 + (x_2 + x_3)X_3 = \vec{0}$ Como X_1, X_2 e X_3 são por hipótese L.I., os escalares que os estão multiplicando têm que ser iguais a zero. O que leva ao sistema
$$\begin{cases} x_1 + x_2 + x_3 = 0 \\ x_3 = 0 \\ x_2 + x_3 = 0 \end{cases}$$
 Assim, o sistema e a equação vetorial inicial têm somente a solução trivial $x_1 = x_2 = x_3 = 0$. Portanto os vetores Y_1, Y_2 e Y_3 são L.I.

- 1.3.7.** $x_1p_1(t) + \dots + x_np_n(t) = x_0(1 + t + t^2 + \dots + t^n) + x_1(t + t^2 + \dots + t^n) + x_2(t^2 + t^3 + \dots + t^n) + \dots + x_n(t^n) = x_0(1) + (x_0 + x_1)t + (x_0 + x_1 + x_2)t^2 + \dots + (x_0 + x_1 + \dots + x_n)t^n = \vec{0}(t) = 0$. Como todos os coeficientes do polinômio nulo têm que ser iguais a zero, então obtemos o seguinte sistema
$$\begin{cases} x_0 = 0 \\ x_0 + x_1 = 0 \\ \vdots \\ x_0 + x_1 + \dots + x_n = 0 \end{cases}$$
 Assim, o sistema e a equação vetorial inicial têm somente a solução trivial. Portanto, os polinômios são L.I.

- 1.3.8.** $x_1f_1(t) + \dots + x_kf_k(t) = e^{\lambda t}(x_1(t) + x_2(t^2) + \dots + x_k(t^k)) = \vec{0}(t) = 0$. Dividindo-se por $e^{\lambda t}$ e considerando que um polinômio nulo tem todos os seus coeficientes iguais a zero, obtemos que $x_1 = \dots = x_k = 0$, ou seja, as funções são L.I.

1.3.9.
$$W[f_1, \dots, f_n](t) = \det \begin{bmatrix} f_1(t) & f_2(t) & \dots & f_k(t) \\ f_1'(t) & f_2'(t) & \dots & f_k'(t) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(k-1)}(t) & f_2^{(k-1)}(t) & \dots & f_k^{(k-1)}(t) \end{bmatrix} = \det \begin{bmatrix} e^{\lambda_1 t} & e^{\lambda_2 t} & \dots & e^{\lambda_k t} \\ \lambda_1 e^{\lambda_1 t} & \lambda_2 e^{\lambda_2 t} & \dots & \lambda_k e^{\lambda_k t} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} e^{\lambda_1 t} & \lambda_2^{k-1} e^{\lambda_2 t} & \dots & \lambda_k^{k-1} e^{\lambda_k t} \end{bmatrix}$$

$$W[f_1, \dots, f_n](0) = \det \begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_k \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \dots & \lambda_k^{k-1} \end{bmatrix} = \prod_{i < j} (\lambda_i - \lambda_j) \neq 0$$

Portanto, as funções são L.I.

- 1.3.10.** Precisamos saber se a equação abaixo tem somente a solução trivial ou não. $x_1AX_1 + \dots + x_nAX_n = A(x_1X_1 + \dots + x_nX_n) = \vec{0}$. Como a matriz A é invertível, então $x_1X_1 + \dots + x_nX_n = \vec{0}$. Como os vetores X_1, \dots, X_n são L.I., então $x_1 = \dots = x_n = 0$. Portanto AX_1, \dots, AX_n são L.I.
- 1.3.11.** Vamos mostrar que o conjunto $\mathcal{X}_n = \{2, 1 + t, \dots, 1 + t^n\}$ é L.I. Para isso, vamos mostrar que a equação abaixo tem somente a solução trivial. $x_0(2) + x_1(t + 1) + \dots + x_n(t^n + 1) = \vec{0}(t) = 0$. Agrupando os termos de mesmo grau obtemos que a equação acima é equivalente ao sistema

$$\begin{cases} 2x_0 + x_1 + \dots + x_n = 0 \\ x_1 + \dots + x_n = 0 \\ \vdots \\ x_n = 0 \end{cases}$$

que tem somente a solução trivial. Portanto o conjunto \mathcal{X}_n é L.I. Como

qualquer subconjunto finito \mathcal{Y} , do conjunto $\{2, 1+t, \dots, 1+t^n, \dots\}$ está contido em algum conjunto \mathcal{X}_n , então \mathcal{Y} é L.I. Isto prova que o conjunto infinito $\{2, 1+t, \dots, 1+t^n, \dots\}$ é L.I.

1.3.12. (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].

1.3.13. >> A=randi(3,2)*randi(2,5,2)

A = -2 4 -2 -8 -8
-4 0 -4 -8 0
5 -3 5 13 6

>> escalona(A)

ans = 1 0 1 2 0
0 1 0 -1 -2
0 0 0 0 0

$A_3 = 1A_1 + 0A_2$, $A_4 = 2A_1 - A_2$, $A_5 = 0A_1 - 2A_2$. Observe que as relações que são válidas entre as colunas de A são válidas entre as colunas da forma escalonada reduzida de A.

1.4. Base e Dimensão de Subespaços (página 87)

1.4.1. Como a dimensão do \mathbb{R}^4 é 4, 4 vetores que são L.I. geram o espaço e formam portanto uma base (Teorema 1.12(a) na página 75). Para verificar se os vetores são L.I., precisamos saber se a equação vetorial,

$$x_1V_1 + x_2V_2 + x_3V_3 + x_4V_4 = \vec{0}$$

só possui a solução trivial. Esta equação é equivalente ao sistema linear, cuja matriz é formada pelos vetores V_1, V_2, V_3, V_4 escritos como colunas.

```
(a) >> v1=[1,0,0,1];v2=[0,1,0,0];v3=[1,1,1,1];
>> v4=[0,1,1,1];A=[v1;v2;v3;v4].';
>> escalona(A)
[ 1, 0, 1, 0]
[ 0, 1, 1, 1]
[ 0, 0, 1, 1]
[ 1, 0, 1, 1]

[ 1, 0, 0, 0]
[ 0, 1, 0, 0]
[ 0, 0, 1, 0]
[ 0, 0, 0, 1]
```

Assim a equação acima tem somente a solução trivial. Portanto os vetores são L.I. e formam uma base do \mathbb{R}^4 .

(b) Não é base pois 3 vetores não geram o \mathbb{R}^4 .

```
(c) >> v1=[0,0,1,1];v2=[-1,1,1,2];v3=[1,1,0,0];
>> v4=[2,1,2,1];A=[v1;v2;v3;v4].';
>> escalona(A)
[ 0, -1, 1, 2]
[ 0, 1, 1, 1]
[ 1, 1, 0, 2]
[ 1, 2, 0, 1]

[ 1, 0, 0, 0]
[ 0, 1, 0, 0]
[ 0, 0, 1, 0]
[ 0, 0, 0, 1]
```

Assim a equação acima tem somente a solução trivial. Portanto os vetores são L.I. e formam uma base do \mathbb{R}^4 .

- 1.4.2. (a) $(a, a, c) = (a, a, 0) + (0, 0, c) = a(1, 1, 0) + c(0, 0, 1)$. Logo, $\mathcal{S} = \{V_1 = (1, 1, 0), V_2 = (0, 0, 1)\}$ gera o subespaço. Além disso, \mathcal{S} é L.I., pois um vetor não é múltiplo escalar do outro. Portanto, \mathcal{S} é uma base do subespaço.
- (b) $(0, b, c) = (0, b, 0) + (0, 0, c) = b(0, 1, 0) + c(0, 0, 1)$ Logo, $\mathcal{S} = \{V_1 = (0, 1, 0), V_2 = (0, 0, 1)\}$ gera o subespaço. Além disso, \mathcal{S} é L.I., pois um vetor não é múltiplo escalar do outro. Portanto, \mathcal{S} é uma base do subespaço.

$$\begin{aligned}
 (c) \quad & (a - b, b + c, 2a - b + c) = \\
 & (a, 0, 2a) + (-b, b, -b) + (0, c, c) = \\
 & a(1, 0, 2) + b(-1, 1, -1) + c(0, 1, 1)
 \end{aligned}$$

Logo, $\mathcal{S} = \{V_1 = (1, 0, 2), V_2 = (-1, 1, -1), V_3 = (0, 1, 1)\}$ gera o subespaço.

Agora, para verificar se \mathcal{S} é L.I., precisamos saber se a equação

$$x(1, 0, 2) + y(-1, 1, -1) + z(0, 1, 1) = (0, 0, 0)$$

possui somente a solução trivial.

```
>> v1=[1,0,2];v2=[-1,1,-1];v3=[0,1,1];
>> A=[v1;v2;v3;zeros(1,3)].';
>> R=escalona(A)
     1     0     1     0
     0     1     1     0
     0     0     0     0
```

Uma solução particular da equação acima é:

$$z = 1, \quad y = -1, \quad x = -1.$$

Substituindo estes valores na equação vetorial acima:

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

ou ainda,

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

Assim, $\mathcal{S} = \{V_1 = (1, 0, 2), V_2 = (-1, 1, -1)\}$ é uma base para o subespaço (eles são L.I., pois um não é múltiplo escalar do outro).

1.4.3. (a)

$$\begin{aligned}
 & (a, b, c, a + b) = \\
 & (a, 0, 0, a) + (0, b, 0, b) + (0, 0, c, 0) = \\
 & a(1, 0, 0, 1) + b(0, 1, 0, 1) + c(0, 0, 1, 0)
 \end{aligned}$$

Logo, $\mathcal{S} = \{V_1 = (1, 0, 0, 1), V_2 = (0, 1, 0, 1), V_3 = (0, 0, 1, 0)\}$ gera o subespaço.

Agora, para verificar se \mathcal{S} é L.I., precisamos saber se a equação

$$x(1, 0, 0, 1) + y(0, 1, 0, 1) + z(0, 0, 1, 0) = (0, 0, 0, 0)$$

possui somente a solução trivial.

```
>> v1=[1,0,0,1];v2=[0,1,0,1];v3=[0,0,1,0];
>> A=[v1;v2;v3;zeros(1,4)].';
>> R=escalona(A)
    1    0    0    0
    0    1    0    0
    0    0    1    0
    0    0    0    0
```

Assim, $\mathcal{S} = \{V_1 = (1, 0, 0, 1), V_2 = (0, 1, 0, 1), V_3 = (0, 0, 1, 0)\}$ é uma base para o subespaço e a dimensão é 3.

(b)

$$\begin{aligned} (a, b, a - b, a + b) &= (a, 0, a, a) + (0, b, -b, b) \\ &= a(1, 0, 1, 1) + b(0, 1, -1, 1) \end{aligned}$$

Logo, $\mathcal{S} = \{V_1 = (1, 0, 1, 1), V_2 = (0, 1, -1, 1)\}$ gera o subespaço. Como um vetor não é múltiplo escalar do outro, eles são L.I.

Portanto, \mathcal{S} é uma base do subespaço e a dimensão do subespaço é 2.

(c)

$$\begin{aligned} (a + c, a - b, b + c, -a + b) &= \\ (a, a, 0, -a) + (0, -b, b, b) + (c, 0, c, 0) &= \\ a(1, 1, 0, -1) + b(0, -1, 1, 1) + c(1, 0, 1, 0) \end{aligned}$$

Logo, $\mathcal{S} = \{V_1 = (1, 1, 0, -1), V_2 = (0, -1, 1, 1), V_3 = (1, 0, 1, 0)\}$ gera o subespaço.

Agora, para verificar se são L.I., precisamos saber se a equação

$$x(1, 1, 0, -1) + y(0, -1, 1, 1) + z(1, 0, 1, 0) = (0, 0, 0, 0)$$

possui somente a solução trivial.

```
>> v1=[1,1,0,-1];v2=[0,-1,1,1];v3=[1,0,1,0];
>> A=[v1;v2;v3;zeros(1,4)].';
>> R=escalona(A)
    1    0    1    0
    0    1    1    0
    0    0    0    0
    0    0    0    0
```

Uma solução particular da equação acima é:

$$z = 1, \quad y = -1, \quad x = -1.$$

Substituindo estes valores na equação vetorial acima:

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

ou ainda,

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

Assim, $S = \{V_1 = (1, 1, 0, -1), V_2 = (0, -1, 1, 1)\}$ é uma base para o subespaço. E a dimensão do subespaço é 2.

- 1.4.4. Como a dimensão do \mathbb{R}^3 é 3, 3 vetores que são L.I. geram o espaço e formam portanto uma base. Para verificar se os vetores são L.I., precisamos saber se a equação vetorial,

$$x_1V_1 + x_2V_2 + x_3V_3 = \vec{0}$$

só possui a solução trivial. Esta equação é equivalente ao sistema linear, cuja matriz é formada pelos vetores V_1, V_2, V_3 escritos como colunas.

```
>> syms a
>> A=[a^2,0,1;0,a,2;1,0,1]
>> escalona(A)
[ a^2,  0,  1]
[  0,  a,  2]
[  1,  0,  1]
eliminação 1:
linha 3 <==> linha 1
[  1,  0,  1]
[  0,  a,  2]
[ a^2,  0,  1]
(-a^2)*linha 1 + linha 3 ==> linha 3
[  1,  0,  1]
[  0,  a,  2]
[  0,  0, 1-a^2]
eliminação 2:
(1/a)*linha 2 ==> linha 2
[  1,  0,  1]
[  0,  1, 2/a]
[  0,  0, 1-a^2]
```

Portanto, para $a \neq 0, 1, -1$ os vetores V_1, V_2 e V_3 formam uma base.

- 1.4.5. (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
```

$$\begin{cases} x_1 - x_3 = 0 \\ x_2 + 2x_3 = 0 \end{cases}$$

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, $\mathcal{S} = \{V = (1, -2, 1)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então \mathcal{S} 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
```

$$\begin{cases} x_1 - 3x_3 = 0 \\ x_2 + 3x_3 = 0 \\ x_4 = 0 \end{cases}$$

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, $\mathcal{S} = \{V = (3, -3, 1, 0)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então \mathcal{S} 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
```

$$\begin{cases} x_2 & & & = 0 \\ & x_3 & & = 0 \\ & & x_4 & = 0 \end{cases}$$

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, $\mathcal{S} = \{V = (1, 0, 0, 0)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então \mathcal{S} 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
```

$$\begin{cases} x_1 & - & 3x_3 & = & 0 \\ & x_2 & & = & 0 \end{cases}$$

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, $\mathcal{S} = \{V = (1, 0, 1)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então \mathcal{S} 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
```

$$\begin{cases} x_1 & - & 3x_3 & = & 0 \\ & x_2 & - & 2x_3 & = & 0 \end{cases}$$

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, $\mathcal{S} = \{V = (3, 2, 1)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então \mathcal{S} 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
```

$$\begin{cases} x_1 & - & x_3 & = & 0 \\ x_2 & - & 3x_3 & = & 0 \end{cases}$$

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, $\mathcal{S} = \{V = (1, 3, 1)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então \mathcal{S} é base para \mathbb{W} .

```
(d) >> 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]
>> solve(det(B))
ans = [ 1] [-1] [ 3] [ 2]
>> Bm1=subs(B,x,-1)
     2     2     3     4
     0     0     3     2
     0     0     4     3
     0     0     0     3
>> escalona([Bm1,zeros(4,1)])
     1     1     0     0     0
     0     0     1     0     0
     0     0     0     1     0
     0     0     0     0     0
```

$$\begin{cases} x_1 + x_2 - 3x_3 & = & 0 \\ & x_3 & = & 0 \\ & & x_4 & = & 0 \end{cases}$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(-\alpha, \alpha, 0, 0) \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de \mathbb{W} temos:

$$(-\alpha, \alpha, 0, 0) = \alpha(-1, 1, 0, 0).$$

Logo, $\mathcal{S} = \{V = (-1, 1, 0, 0)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então \mathcal{S} é base para \mathbb{W} .

```
>> B1=subs(B,x,1)
    0    2    3    4
    0   -2    3    2
    0    0    2    3
    0    0    0    1
>> escalona([B1,zeros(4,1)])
    0    1    0    0    0
    0    0    1    0    0
    0    0    0    1    0
    0    0    0    0    0
```

$$\begin{cases} x_2 & & & = 0 \\ & x_3 & & = 0 \\ & & x_4 & = 0 \end{cases}$$

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, $\mathcal{S} = \{V = (1, 0, 0, 0)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então \mathcal{S} é base para \mathbb{W} .

```
>> B2=subs(B,x,2)
   -1    2    3    4
    0   -3    3    2
    0    0    1    3
    0    0    0    0
>> escalona([B2,zeros(4,1)])
    1    0    0   29/3    0
    0    1    0    7/3    0
    0    0    1    3    0
    0    0    0    0    0
```

$$\begin{cases} x_1 & & + & (29/7)x_4 & = 0 \\ & x_2 & & + & (7/3)x_4 & = 0 \\ & & x_3 & + & 3x_4 & = 0 \end{cases}$$

Este sistema tem como solução geral

$$\mathbb{W} = \{(-29/3)\alpha, -(7/3)\alpha, -3\alpha, \alpha \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de \mathbb{W} temos:

$$\begin{aligned} &(-29/3)\alpha, -(7/3)\alpha, -3\alpha, \alpha) = \\ &\alpha(-29/3, -7/3, -3, 1). \end{aligned}$$

Logo, $S = \{V = (-29/3, -7/3, -3, 1)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então S é base para \mathbb{W} .

```
>> B3=subs(B,x,3)
-2    2    3    4
 0    -4   3    2
 0     0   0    3
 0     0   0   -1
>> escalona([B3,zeros(4,1)])
 1     0  -9/4   0  0
 0     1  -3/4   0  0
 0     0   0     1  0
 0     0   0     0  0
```

$$\begin{cases} x_1 & - & (9/4)x_3 & = & 0 \\ & x_2 & - & (3/4)x_3 & = & 0 \\ & & & x_4 & = & 0 \end{cases}$$

Este sistema tem como solução geral

$$\mathbb{W} = \{((9/4)\alpha, (3/4)\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}.$$

Agora, para qualquer elemento de \mathbb{W} temos:

$$((9/4)\alpha, (3/4)\alpha, \alpha, 0) = \alpha(9/4, 3/4, 1, 0).$$

Logo, $S = \{V = (9/4, 3/4, 1, 0)\}$ gera \mathbb{W} . Como um conjunto formado por um único vetor não nulo é sempre L.I., então S é base para \mathbb{W} .

1.4.6. (a)

```
>> v1=[2,1,3];v2=[3,-1,4];v3=[2,6,4];
>> A=[v1;v2;v3;zeros(1,3)].';
>> escalona(A)
[ 1, 0, 4, 0]
[ 0, 1, -2, 0]
[ 0, 0, 0, 0]
```

A equação $xV_1 + yV_2 + zV_3 = \vec{0}$ admite solução não trivial.

(b) 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 com vetor normal $N = V_1 \times V_2 = (7, 1, -5)$, ou seja, é o plano $7x + y - 5z = 0$.

1.4.7. (a) Não. O \mathbb{R}^3 é um subespaço de dimensão 3.

(b) V_3 deve ser um vetor que não seja combinação linear de V_1 e V_2 .

```
(c) >> v1=[1,1,1];v2=[3,-1,4];
>> syms a b c
>> A=[v1;v2;[a,b,c]].';
>> escalona(A)
[      1,      0, 4*a-3*c]
[      0,      1,      c-a]
[      0,      0, b-5*a+4*c]
```

Seja $V_3 = (a, b, c)$ tal que $b - 5a + 4c \neq 0$. Por exemplo, $V_3 = (0, 0, 1)$, é tal que V_1, V_2 e V_3 formam uma base de \mathbb{R}^3 .

1.4.8. Fazendo $z = \alpha$ e $y = \beta$, 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 $(x, y, z) = \alpha(-4, 0, 1) + \beta(-2, 1, 0) = \alpha V_1 + \beta V_2$, $\forall \alpha, \beta \in \mathbb{R}$, onde $V_1 = (-4, 0, 1)$ 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 conseguir isso é a seguinte. Pelo menos um dos vetores da base canônica não é combinação linear de V_1 e V_2 . Para descobrir qual, podemos escalonar a matriz cujas colunas são os vetores V_1, V_2, E_1, E_2, E_3 , ou seja,

```
>> V1=[-4;0;1];V2=[-2;1;0];
>> A=[V1,V2,eye(3)];
>> escalona(A)
[ -4, -2,  1,  0,  0]
[  0,  1,  0,  1,  0]
[  1,  0,  0,  0,  1]

[  1,  0,  0,  0,  1]
[  0,  1,  0,  1,  0]
[  0,  0,  1,  2,  4]
```

Assim, nenhum dos vetores da base canônica é combinação linear de V_1 e V_2 . Portanto podemos tomar como V_3 qualquer um dos vetores da base canônica que teremos $\{V_1, V_2, V_3\}$ sendo uma base de \mathbb{R}^3 .

1.4.9. Seja $p(t) = at^2 + bt + c$ um polinômio qualquer de \mathcal{P}_2 .

$x_1(1) + x_2(t - 1) + x_3(t^2 - 3t + 1) = (x_3)t^2 + (x_2 - 3x_3)t + (x_1 - x_2 + x_3)1 = at^2 + bt + c$. Como dois polinômios são iguais se,

e somente se, os coeficientes dos termos de mesmo grau são iguais, então $\begin{cases} x_3 = a \\ x_2 - 3x_3 = b \\ x_1 + x_2 - 3x_3 = c \end{cases}$ que tem solução $x_3 = a, x_2 = b + 3a$ e $x_1 = c + 3a - b - 3a = c - b$. Assim, o sistema e a equação vetorial inicial têm solução única, para todo polinômio $p(t) = at^2 + bt + c \in \mathcal{P}_2$. Portanto, pelo Teorema 1.9 na página 70, os polinômios $1, t - 1$ e $t^2 - 3t + 1$ formam uma base de \mathcal{P}_2 .

1.4.10. Se as funções fossem L.I., então o subespaço teria dimensão 3, mas $\cos 2t = \cos^2 t - \sin^2 t = 2\cos^2 t - 1$. Logo, a função $\cos 2t$ não é necessária para gerar o subespaço. As funções 1 e $\cos^2 t$ são claramente L.I. Assim elas formam uma base para o subespaço e portanto a dimensão do subespaço é igual a 2.

1.4.11. (a) Um polinômio $p(t) = at^3 + bt^2 + ct + d$ é tal que $p(0) = 0$ se, e somente se, $d = 0$. Assim, todo polinômio de \mathbb{W}_1 é da forma $p(t) = at^3 + bt^2 + ct = a(t^3) + b(t^2) + c(t)$. Assim, t, t^2, t^3 geram \mathbb{W}_1 e como claramente eles são L.I., então $\{t, t^2, t^3\}$ é uma base de \mathbb{W}_1 .

(b) Um polinômio $q(t) = at^3 + bt^2 + ct + d$ é tal que $q(1) = 0$ se, e somente se, $a + b + c + d = 0$ ou $d = -a - b - c$. Assim, todo polinômio de \mathbb{W}_1 é da forma $p(t) = at^3 + bt^2 + ct + (-a - b - c) = a(t^3 - 1) + b(t^2 - 1) + c(t - 1)$. Assim, $t - 1, t^2 - 1, t^3 - 1$ geram \mathbb{W}_2 e como claramente eles são L.I., então $\{t - 1, t^2 - 1, t^3 - 1\}$ é uma base de \mathbb{W}_2 .

(c) Um polinômio $p(t) = at^3 + bt^2 + ct + d$ é tal que $p(0) = 0$ e $p(1) = 0$ se, e somente se, $d = 0$ e $a + b + c + d = 0$, ou seja, se $d = 0$ e $c = -a - b$. Assim, todo polinômio de $\mathbb{W}_1 \cap \mathbb{W}_2$ é da forma $p(t) = at^3 + bt^2 + (-a - b)t = a(t^3 - t) + b(t^2 - t)$. Assim, $t^3 - t$ e $t^2 - t$ geram $\mathbb{W}_1 \cap \mathbb{W}_2$ e como claramente eles são L.I., então $\{t^3 - t, t^2 - t\}$ é uma base para $\mathbb{W}_1 \cap \mathbb{W}_2$.

1.4.12. (a) $a_{ji} = a_{ij}$, para $i = 1, \dots, n$ e $j = i, \dots, n$. $\{E_{ij} + E_{ji} \mid i = 1, \dots, n - 1, j = i + 1, \dots, n\} \cup \{E_{ii} \mid i = 1, \dots, n\}$ é uma base para o espaço das matrizes simétricas. Assim a dimensão deste espaço é igual a $n + (n - 1) + \dots + 1 = \frac{n(n+1)}{2}$.

(b) $a_{ji} = -a_{ij}$, para $i = 1, \dots, n$ e $j = i, \dots, n$. $\{E_{ij} - E_{ji} \mid i = 1, \dots, n - 1, j = i + 1, \dots, n\}$ é uma base para o espaço das matrizes anti-simétricas. Assim a dimensão deste espaço é igual a $(n - 1) + (n - 2) + \dots + 1 = \frac{n(n-1)}{2}$.

(c) $a_{nn} = -a_{11} - \dots - a_{(n-1)(n-1)}$. $\{E_{ij} \mid i, j = 1, \dots, n, i \neq j\} \cup \{E_{11} - E_{nn}, \dots, E_{(n-1)(n-1)} - E_{nn}\}$. Assim a dimensão deste espaço é igual a $n^2 - 1$.

(d) $a_{in} = a_{1i}$, para $i = 1, \dots, n$. $\{E_{ij} \mid i = 2, \dots, n, j = 1, \dots, n - 1\} \cup \{E_{1i} + E_{in} \mid i = 2, \dots, n - 1\} \cup \{E_{11} + E_{1n} + E_{nn}\}$. Assim, a dimensão do espaço é $(n - 1)^2 + (n - 2) + 1 = n^2 - n$.

(e) $a_{11} + \dots + a_{1n} = a_{1n} + \dots + a_{nn}$. $a_{nn} = -a_{11} - \dots - a_{1(n-1)} - a_{2n} - \dots - a_{(n-1)n}$. $\{E_{ij} \mid i = 2, \dots, n, j = 1, \dots, n - 1\} \cup \{E_{11} - E_{nn}, \dots, E_{1(n-1)} - E_{nn}, E_{1n}, E_{2n} - E_{nn}, \dots, E_{(n-1)n} - E_{nn}\}$. Assim, a dimensão do espaço é $n^2 - 1$.

1.4.13. Seja $p(t) = a_0 + \dots + a_n t^n$ um polinômio tal que $p(a) = 0$. Então $a_0 + a_1 a + \dots + a_n a^n = 0$, ou $a_0 = -a_1 a - \dots - a_n a^n$. Assim, $p(t) = (-a_1 a - \dots - a_n a^n) + a_1 t + \dots + a_n t^n = a_1(t - a) + a_2(t^2 - a^2) + \dots + a_n(t^n - a^n)$. Assim, $t - a, t^2 - a^2, \dots, t^n - a^n$ geram \mathbb{W} . Como eles são claramente L.I., então eles formam uma base de \mathbb{W} . A dimensão de \mathbb{W} é igual a n .

- 1.4.14. Não, pois eles podem ser L.D., por exemplo se os três vetores forem iguais a um vetor não nulo, a dimensão do espaço gerado por eles é igual a 1.
- 1.4.15. Não, pois se A for igual a matriz nula, então $AX_1 = \dots = AX_n = \vec{0}$. Supondo que a matriz A seja singular, então como a dimensão do \mathbb{R}^n é igual a n , basta mostrar que AX_1, \dots, AX_n são L.I.
 $x_1AX_1 + \dots + x_nAX_n = A(x_1X_1 + \dots + x_nX_n) = \vec{0}$. Como a matriz A é invertível, então $x_1X_1 + \dots + x_nX_n = \vec{0}$. Como os vetores X_1, \dots, X_n são L.I., então $x_1 = \dots = x_n = 0$. Portanto AX_1, \dots, AX_n são L.I. e formam uma base do \mathbb{R}^n .
- 1.4.16. Seja $\{V_1, V_2\}$ uma base de \mathbb{V} e $\{W_1, W_2\}$ uma base de \mathbb{W} . Um vetor V está na interseção se $V = x_1V_1 + x_2V_2 = y_1W_1 + y_2W_2$. O que implica que $x_1V_1 + x_2V_2 - y_1W_1 - y_2W_2 = \vec{0}$. Como espaço tem dimensão 3 estes quatro vetores são L.D., o que implica que existem escalares x_1, x_2, y_1, y_2 não todos nulos que satisfazem a equação acima. Ou seja, $\mathbb{V} \cap \mathbb{W} \neq \{\vec{0}\}$. O mesmo não seria verdade se o espaço tivesse dimensão igual a 4. Por exemplo, se $\mathbb{V} = [E_1, E_2]$ e $\mathbb{W} = [E_3, E_4]$ são tais que $\mathbb{V} \cap \mathbb{W} = \{\vec{0}\}$.
- 1.4.17. Já mostramos na solução do [Exercício 12 na página 406](#) que o conjunto gera o espaço dos polinômios e na solução do [Exercício 11 na página 411](#) que o conjunto é L.I. Portanto, o conjunto é uma base do espaço dos polinômios.
- 1.4.18. Já mostramos na solução do [Exercício 13 na página 406](#) que eles geram o espaço dos polinômios pares. Como eles são claramente L.I., então eles formam uma base do espaço dos polinômios pares.
- 1.4.19.

```
>> A=randi(4,3)*randi(3,5,2);
>> R=escalona(A)
[ 6, -2, 1, 8, 2]
[ 12, 6, -1, 8, 9]
[ 20, 12, 1, 15, 16]
[ 0, 8, 5, 1, 4]
R = [ 1, 0, 0, 0, 1, 1/2]
[ 0, 1, 0, -1/2, 1/2]
[ 0, 0, 1, 1, 0]
[ 0, 0, 0, 0, 0, 0]
```

O conjunto solução de $AX = \vec{0}$ é o mesmo de $RX = \vec{0}$. Assim, a mesma relação que é válida entre as colunas de R é válida entre as colunas de A . Portanto, as colunas de A que correspondem aos pivôs de R formam uma base para o subespaço gerado pelas colunas de A , pois as outras colunas são combinação linear destas.

- 1.4.20.

```
>> A=randi(4,2)
A = 2 1
     2 -4
     3 -1
     0 2
>> B=[A,eye(4)];
>> R=escalona(B)
[ 2, 1, 1, 0, 0, 0]
[ 2, -4, 0, 1, 0, 0]
[ 3, -1, 0, 0, 1, 0]
[ 0, 2, 0, 0, 0, 1]
```

```
R = [ 1, 0, 0, 0, 1/3, 1/6]
[ 0, 1, 0, 0, 0, 1/2]
[ 0, 0, 1, 0, -2/3, -5/6]
[ 0, 0, 0, 1, -2/3, 5/3]
```

As colunas de B que correspondem aos pivôs de R formam uma base para o subespaço gerado pelas colunas de B, pois as outras colunas são combinação linear destas.

```
1.4.21. (a) >> A=randi(4,3)*randi(3,5,2)
A = 5 -4 1 -5 -3
-9 5 -4 -3 1
5 -5 0 -7 -5
6 -3 3 11 0
>> escalona([A(:,1)+A(:,2),A])
[ 1, 5, -4, 1, -5, -3]
[ -4, -9, 5, -4, -3, 1]
[ 0, 5, -5, 0, -7, -5]
[ 3, 6, -3, 3, 11, 0]
ans =
[ 1, 0, 1, 1, 0, 2]
[ 0, 1, -1, 0, 0, -1]
[ 0, 0, 0, 0, 1, 0]
[ 0, 0, 0, 0, 0, 0]
```

A base do subespaço é formada por $V_1 = A_1 + A_2$, $V_2 = A_2 = (5, -9, 5, 6)$, $V_3 = A_4 = (-5, -3, -7, 11)$.

```
(b) >> B=A*randi(5,2)
B = -61 -17
42 -35
-78 -33
11 62
>> escalona([B,A])
[ -61, -17, 5, -4, 1, -5, -3]
[ 42, -35, -9, 5, -4, -3, 1]
[ -78, -33, 5, -5, 0, -7, -5]
[ 11, 62, 6, -3, 3, 11, 0]
ans =
[1, 0, 0, 1, 1, -2, 2]
[0, 1, 0, -1, -1, 9/4, -2]
[0, 0, 1, 8, 9, -71/4, 17]
[0, 0, 0, 0, 0, 0, 0]
```

A base do subespaço é formada por $V_1 = (-61, 42, -78, 11)$, $V_2 = (-17, -35, -33, 62)$, $V_3 = A_1 = (5, -9, 5, 6)$.

2.1. Produto Escalar e Norma (página 128)

```

2.1.1. >> syms a
>> x=[1,1,-2];y=[a,-1,2];
>> solve(pe(x,y))
ans = 5

2.1.2. >> 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]

2.1.3. (a) >> w1=[1,-1,1],w2=[1,1,0],w3=[-1,1,2]
>> pe(w1,w2),pe(w1,w3),pe(w2,w3)
ans = 0, 0, 0

```

O conjunto dado é uma base de \mathbb{R}^3 , pois sendo um conjunto ortogonal é L.I. (Proposição 2.7 na página 122). E um subconjunto de \mathbb{R}^3 L.I. com 3 vetores é uma base de \mathbb{R}^3 pois a dimensão \mathbb{R}^3 é igual a 3 (Teorema 1.12 na página 75).

```

(b) >> no(w1),no(w2),no(w3)
ans = 3^(1/2),2^(1/2),6^(1/2)
>> u1=w1/no(w1),u2=w2/no(w2),u3=w3/no(w3)

```

$$W_1 = (1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3}),$$

$$W_2 = (1/\sqrt{2}, 1/\sqrt{2}, 0),$$

$$W_3 = (-1/\sqrt{6}, 1/\sqrt{6}, \sqrt{2}/\sqrt{3}).$$

```

(c) >> v=[2,1,0];
>> pe(v,w1),pe(v,w2),pe(v,w3)
ans = 1, 3, -1
>> pe(w1,w1),pe(w2,w2),pe(w3,w3)
ans = 3, 2, 6

```

$$V = (2,1,0) = (1/3)W_1 + (3/2)W_2 + (-1/6)W_3.$$

```

(d) >> syms x y z
>> v=[x,y,z];
>> pe(v,w1),pe(v,w2),pe(v,w3)
ans = x-y+z, x+y, -x+y+2*z

```

$$(x,y,z) = (x-y+z)/3 W_1 + (x+y)/2 W_2 + (-x+y+2z)/6 W_3.$$

2.1.4. (a) As funções t e t^3 , por exemplo, não são ortogonais.

$$\langle t, t^3 \rangle = \int_{-1}^1 t^4 dt = 2/5 \neq 0$$

(b)

$$\langle f_1, f_2 \rangle = \langle t, t^2 \rangle = \int_{-1}^1 t^3 dt = 0$$

$$\begin{aligned} \|f_1\|^2 + \|f_2\|^2 &= \|t\|^2 + \|t^2\|^2 = \langle t, t \rangle + \langle t^2, t^2 \rangle \\ &= \int_{-1}^1 t^2 dt + \int_{-1}^1 t^4 dt = \frac{2}{3} + \frac{2}{5} \\ &= \frac{16}{15} \\ \|f_1 + f_2\|^2 &= \|t + t^2\|^2 = \langle t + t^2, t + t^2 \rangle \\ &= \int_{-1}^1 (t + t^2)^2 dt = (1/3)t^3 \Big|_{-1}^1 \\ &= 2(1/4)t^4 \Big|_{-1}^1 + (1/5)t^5 \Big|_{-1}^1 \\ &= \frac{2}{3} + 0 + \frac{2}{5} = \frac{16}{15} \end{aligned}$$

(c)

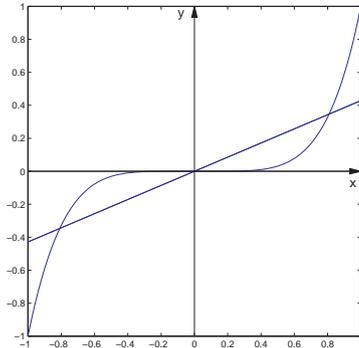
$$\text{proj}_t t^5 = \frac{\langle t^5, t \rangle}{\|t\|^2} t = \frac{\int_{-1}^1 t^6 dt}{\int_{-1}^1 t^2 dt} t = \frac{2/7}{2/3} t = \frac{3}{7} t$$

Geometricamente a projeção de t^5 em t dá o múltiplo escalar de t que está mais “próximo” de t^5 no sentido de que $\|t^5 - \alpha t\| = \left(\int_{-1}^1 (t^5 - \alpha t)^2 dt \right)^{1/2}$ é mínimo.

(d)

$$t^5 = \left(t^5 - \frac{3}{7} t \right) + \frac{3}{7} t,$$

sendo que $\frac{3}{7}t$ é um múltiplo escalar de t .



(e)

$$\begin{aligned} \cos \theta &= \frac{\langle t^5, t \rangle}{\|t^5\| \|t\|} \\ &= \frac{\int_{-1}^1 t^6 dt}{\left(\int_{-1}^1 t^{10} dt\right)^{1/2} \left(\int_{-1}^1 t^2 dt\right)^{1/2}} \\ &= \frac{2/7}{\sqrt{2/11} \sqrt{2/3}} = \frac{\sqrt{33}}{7}. \end{aligned}$$

Assim o ângulo entre t^5 e t é

$$\theta = \arccos \frac{\sqrt{33}}{7} \approx 35^\circ.$$

2.1.5. (a)

```
>> syms t
>> p0=sym(1);p1=t;p2=t^2-1/3;
>> int(p0*p1,-1,1),int(p0*p2,-1,1),int(p1*p2,-1,1)
ans = 0, 0, 0
```

O conjunto dado é uma base de \mathcal{P}_2 , pois sendo um conjunto ortogonal é L.I. (Proposição 2.7 na página 122). E um subconjunto de \mathcal{P}_2 L.I. com 3 vetores é uma base de \mathcal{P}_2 pois a dimensão \mathcal{P}_2 é igual a 3 (Teorema 1.12 na página 75).

```
(b) >> n0=(int(p0^2,-1,1))^(1/2)
      n0 = 2^(1/2)
      >> n1=(int(p1^2,-1,1))^(1/2)
      n1 = 1/3*6^(1/2)
      >> n2=(int(p2^2,-1,1))^(1/2)
      n2 = 1/45*8^(1/2)*45^(1/2)
      >> q0=p0/n0
      q0 = 1/2*2^(1/2)
      >> q1=p1/n1
      q1 = 1/2*t*2^(1/2)*3^(1/2)
      >> q2=p2/n2
      1/8*(t^2-1/3)*8^(1/2)*45^(1/2)
```

$$q_0 = \frac{\sqrt{2}}{2}, \quad q_1 = \frac{\sqrt{6}}{2} t, \quad q_2 = \frac{\sqrt{8}}{\sqrt{45}} \left(t^2 - \frac{1}{3}\right).$$

```
(c) >> f=1+t+t^2;
      >> int(f*p0,-1,1),int(f*p1,-1,1),int(f*p2,-1,1)
      ans = 8/3, 2/3, 8/45
      >> int(p0^2,-1,1),int(p1^2,-1,1),int(p2^2,-1,1)
      ans = 2, 2/3, 8/45
```

$$1 + t + t^2 = \left(\frac{4}{3}\right)1 + 1t + 1\left(t^2 - \frac{1}{3}\right).$$

```
(d) >> syms a b c
      >> p0=sym(1);p1=t;p2=t^2-1/3;
      >> f=a+b*t+c*t^2;
      >> int(f*p0,-1,1),int(f*p1,-1,1),int(f*p2,-1,1)
      ans = 2*a+2/3*c, 2/3*b, 8/45*c
```

$$a + bt + ct^2 = \left(a + \frac{1}{3}c\right)1 + bt + c\left(t^2 - \frac{1}{3}\right).$$

2.1.6.

$$\begin{aligned}
\langle e^t - \alpha e^{2t}, 1 \rangle &= \frac{1}{2} \int_{-1}^1 e^t dt - \frac{\alpha}{2} \int_{-1}^1 e^{2t} dt \\
&= \frac{1}{2} (e - e^{-1}) - \frac{\alpha}{4} (e^2 - e^{-2}) \\
&= \frac{1}{4} (e - e^{-1}), \\
\langle e^t - \alpha e^{2t}, e^{in\pi t} \rangle &= \frac{1}{2} \int_{-1}^1 e^t e^{in\pi t} dt - \frac{\alpha}{2} \int_{-1}^1 e^{2t} e^{in\pi t} dt \\
&= \frac{1}{2} \int_{-1}^1 e^{(1-in\pi)t} dt - \frac{\alpha}{2} \int_{-1}^1 e^{(2-in\pi)t} dt \\
&= \frac{1}{2(1-in\pi)} e^{(1-in\pi)t} \Big|_{-1}^1 \\
&\quad - \frac{\alpha}{2(2-in\pi)} e^{(2-in\pi)t} \Big|_{-1}^1 \\
&= \frac{1}{2(1-in\pi)} (e^{(1-in\pi)} - e^{-(1-in\pi)}) \\
&\quad - \frac{\alpha}{2(2-in\pi)} (e^{(2-in\pi)} - e^{-(2-in\pi)}) \\
&= \frac{(-1)^n (e - e^{-1})}{2} \left(\frac{1}{1-in\pi} - \frac{1}{2-in\pi} \right) \\
&= \frac{1}{2} \frac{(-1)^n (e - e^{-1}) (1 - n^2 \pi^2 + 3n\pi i)}{(1 + n^2 \pi^2)(4 + n^2 \pi^2)}, \\
&\quad \text{para } n \neq 0.
\end{aligned}$$

2.1.7.

$$\begin{aligned}
 a_0 &= \langle f, 1/\sqrt{2} \rangle = \sqrt{2}c_0 = \frac{\sqrt{2}}{4}(e - e^{-1}), \\
 a_n &= \langle f, \cos n\pi t \rangle = c_n + \bar{c}_n \\
 &= \frac{(-1)^n(e - e^{-1})(1 - n^2\pi^2)}{(1 + n^2\pi^2)(4 + n^2\pi^2)} \\
 b_n &= \langle f, \text{sen } n\pi t \rangle = \frac{\bar{c}_n - c_n}{i} \\
 &= \frac{(-1)^{n+1}(e - e^{-1})3n\pi}{(1 + n^2\pi^2)(4 + n^2\pi^2)}
 \end{aligned}$$

2.2. Bases Ortogonais e Subespaços Ortogonais (página 184)

2.2.1. >> 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)

$$\begin{aligned}
 u_1 &= \left[\frac{1}{3}\sqrt{3} \quad \frac{1}{3}\sqrt{3} \quad -\frac{1}{3}\sqrt{3} \quad 0 \right] \\
 u_2 &= \left[-\frac{2}{33}\sqrt{11}\sqrt{3} \quad \frac{4}{33}\sqrt{11}\sqrt{3} \quad \frac{2}{33}\sqrt{11}\sqrt{3} \quad \frac{1}{11}\sqrt{11}\sqrt{3} \right] \\
 u_3 &= \left[-\frac{2}{55}\sqrt{110} \quad -\frac{3}{110}\sqrt{110} \quad -\frac{7}{110}\sqrt{110} \quad \frac{3}{55}\sqrt{110} \right]
 \end{aligned}$$

2.2.2. >> 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)

$$\begin{aligned}
 u_1 &= \left[\frac{1}{3}\sqrt{3} \quad \frac{1}{3}\sqrt{3} \quad \frac{1}{3}\sqrt{3} \right] \\
 u_2 &= \left[-\frac{1}{3}\sqrt{2}\sqrt{3} \quad \frac{1}{6}\sqrt{2}\sqrt{3} \quad \frac{1}{6}\sqrt{2}\sqrt{3} \right] \\
 u_3 &= \left[0 \quad -\frac{1}{2}\sqrt{2} \quad \frac{1}{2}\sqrt{2} \right]
 \end{aligned}$$

2.2.3. 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/|w1|, u2=w2/|w2|
```

$$\begin{aligned} u_1 &= \left[-\frac{1}{2}\sqrt{2} \quad 0 \quad \frac{1}{2}\sqrt{2} \right] \\ u_2 &= \left[-\frac{1}{6}\sqrt{3}\sqrt{2} \quad \frac{1}{3}\sqrt{3}\sqrt{2} \quad -\frac{1}{6}\sqrt{3}\sqrt{2} \right] \end{aligned}$$

2.2.4. 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/|w1|, u2=w2/|w2|, u3=w3/|w3|
```

$$\begin{aligned} u_1 &= \left[-\frac{1}{2}\sqrt{2} \quad 0 \quad 0 \quad \frac{1}{2}\sqrt{2} \right] \\ u_2 &= \left[\frac{1}{3}\sqrt{3} \quad 0 \quad \frac{1}{3}\sqrt{3} \quad \frac{1}{3}\sqrt{3} \right] \\ u_3 &= \left[\frac{1}{42}\sqrt{42} \quad \frac{1}{7}\sqrt{42} \quad -\frac{1}{21}\sqrt{42} \quad \frac{1}{42}\sqrt{42} \right] \end{aligned}$$

2.2.5. >> 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

$$\mathbb{W} = \{(-3\alpha, 4\alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

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/ao(v)
```

$$u = \left[-\frac{3}{26} \sqrt{26} \quad \frac{2}{13} \sqrt{26} \quad \frac{1}{26} \sqrt{26} \right]$$

```
2.2.6. >> 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 $\vec{P_1P_2}$ 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,W2)
W2 = [ 7/25, 1, 1/25]
>> U1=W1/ao(W1), U2=W2/ao(W2)
```

$$U_1 = \left[1/10 \sqrt{2} \quad 0 \quad -\frac{7}{10} \sqrt{2} \right]$$

$$U_2 = \left[\frac{7}{45} \sqrt{3} \quad 5/9 \sqrt{3} \quad 1/45 \sqrt{3} \right]$$

Para completarmos a uma base ortonormal de \mathbb{R}^3 , basta acrescentarmos $U_3 = U_1 \times U_2$.

>> U3=pv(U1,U2)

$$U_3 = \left[\frac{7}{18} \sqrt{2}\sqrt{3} \quad -1/9 \sqrt{2}\sqrt{3} \quad 1/18 \sqrt{2}\sqrt{3} \right]$$

```
2.2.7. >> 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)
```

$$\text{expr2} = 1/6 |6 + d| \sqrt{3}$$

```
>> 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)
```

$$U_1 = \left[1/2 \sqrt{2} \quad 0 \quad -1/2 \sqrt{2} \right]$$

$$U_2 = \left[-1/6 \sqrt{3}\sqrt{2} \quad 1/3 \sqrt{3}\sqrt{2} \quad -1/6 \sqrt{3}\sqrt{2} \right].$$

2.2.8.

2.2.9. (a)

$$\begin{aligned}(\operatorname{proj}_{p_0} \sqrt{t+1})(t) &= \frac{\langle \sqrt{t+1}, p_0 \rangle}{\|p_0\|^2} p_0(t) = \\ &= \frac{(2/3)(t+1)^{\frac{3}{2}} \Big|_{-1}^1}{2} = \frac{2\sqrt{2}}{3} \\ (\operatorname{proj}_{p_1} \sqrt{t+1})(t) &= \frac{\langle \sqrt{t+1}, p_1 \rangle}{\|p_1\|^2} p_1(t) \\ &= \frac{2\sqrt{2}}{5} t \\ (\operatorname{proj}_{p_2} \sqrt{t+1})(t) &= \frac{\langle \sqrt{t+1}, p_2 \rangle}{\|p_2\|^2} p_2(t) \\ &= -\frac{\sqrt{2}}{7} \left(t^2 - \frac{1}{3}\right)\end{aligned}$$

(b)

$$\begin{aligned}
 \left(\text{proj}_{p_0} \sqrt{t+1}\right)(t) &= \left(\text{proj}_{p_0} \sqrt{t+1}\right)(t) \\
 &= \frac{2\sqrt{2}}{3} \\
 \left(\text{proj}_{p_1} \sqrt{t+1}\right)(t) &= \left(\text{proj}_{p_0} \sqrt{t+1}\right)(t) \\
 &\quad + \left(\text{proj}_{p_1} \sqrt{t+1}\right)(t) \\
 &= \frac{2\sqrt{2}}{3} + \frac{2\sqrt{2}}{5}t \\
 &= 2\sqrt{2}\left(\frac{1}{3} + \frac{1}{5}t\right) \\
 \left(\text{proj}_{p_2} \sqrt{t+1}\right)(t) &= \left(\text{proj}_{p_0} \sqrt{t+1}\right)(t) \\
 &\quad + \left(\text{proj}_{p_1} \sqrt{t+1}\right)(t) \\
 &\quad + \left(\text{proj}_{p_2} \sqrt{t+1}\right)(t) \\
 &= \frac{2\sqrt{2}}{3} + \frac{2\sqrt{2}}{5}t \\
 &\quad - \frac{\sqrt{2}}{7}\left(t^2 - \frac{1}{3}\right) \\
 &= \sqrt{2}\left(-\frac{1}{7}t^2 + \frac{2}{5}t + \frac{15}{21}\right).
 \end{aligned}$$

2.2.10. (a)

```
>> syms t
>> p1=t;p2=t^2;p3=t^3;
>> q1=p1;
>> q2=p2-proj(q1,p2,0,4)
q2 = t^2-3*t
>> q3=p3-proj(q1,p3,0,4)-proj(q2,p3,0,4)
q3 = t^3+32/5*t-16/3*t^2
```

$\{q_1(t) = t, q_2(t) = t^2 - 3t\}$ é uma base para \mathbb{W}_2 e $\{q_1(t) = t, q_2(t) = t^2 - 3t, q_3(t) = t^3 + \frac{32}{5}t - \frac{16}{3}t^2\}$ é uma base para \mathbb{W}_3 .

(b)

```
>> f=sqrt(t)
f = t^(1/2)
```

```

>> pr1=proj(q1,f,0,4)
pr1 = 3/5*t
>> pr2=proj(q2,f,0,4)
pr2 = -1/7*t^2+3/7*t
>> pr3=proj(q3,f,0,4)
pr3 = 1/16*t^3+2/5*t-1/3*t^2

```

$$\begin{aligned} (\text{proj}_{W_2} \sqrt{t})(t) &= \text{proj}_{q_1} \sqrt{t} + \text{proj}_{q_2} \sqrt{t} = \left(\frac{3}{5}t\right) + \left(-\frac{1}{7}t^2 + \frac{3}{7}t\right) = \frac{36}{35}t - \frac{1}{7}t^2. & (\text{proj}_{W_2} \sqrt{t})(t) &= \text{proj}_{q_1} \sqrt{t} + \text{proj}_{q_2} \sqrt{t} + \\ \text{proj}_{q_3} \sqrt{t} &= (\text{proj}_{W_2} \sqrt{t})(t) + \text{proj}_{q_3} \sqrt{t} = \left(\frac{36}{35}t - \frac{1}{7}t^2\right) + \left(\frac{1}{16}t^3 + \frac{2}{5}t - \frac{1}{3}t^2\right) = \frac{10}{7}t - \frac{10}{21}t^2 + \frac{1}{16}t^3. \end{aligned}$$

$$2.2.11. (\text{proj}_{W_n} f)(t) = \frac{1}{2} - \frac{2}{\pi} \sum_{m=1}^n \frac{\text{sen } \frac{m\pi}{2}}{m} \cos m\pi t.$$

$$(\text{proj}_{V_n} f)(t) = \frac{2}{\pi} \sum_{m=1}^n \frac{\cos \frac{m\pi}{2} - (-1)^m}{m} \text{sen } m\pi t$$

$$2.2.12. (\text{proj}_{W_n} f)(t) = \frac{1}{2} + \frac{2}{\pi} \sum_{m=1}^n \frac{\text{sen } \frac{3m\pi}{4} - \text{sen } \frac{m\pi}{4}}{m} \cos m\pi t.$$

$$(\text{proj}_{V_n} f)(t) = \frac{2}{\pi} \sum_{m=1}^n \frac{\cos \frac{m\pi}{4} - \cos \frac{3m\pi}{4}}{m} \text{sen } m\pi t$$

$$2.2.13. (\text{proj}_{W_n} f)(t) = \frac{3}{8} + \frac{2}{\pi^2} \sum_{m=1}^n \frac{\cos m\pi - \cos \frac{m\pi}{2} - \frac{m\pi}{2} \text{sen } \frac{m\pi}{2}}{m^2} \cos m\pi t.$$

$$\begin{aligned} (\text{proj}_{V_n} f)(t) &= \\ \frac{2}{\pi^2} \sum_{m=1}^n \frac{\frac{m\pi}{2} \cos \frac{m\pi}{2} - m\pi \cos m\pi - \text{sen } \frac{m\pi}{2}}{m^2} \text{sen } m\pi t \end{aligned}$$

$$2.2.14. (\text{proj}_{W_n} f)(t) = \frac{1}{4} + \frac{2}{\pi^2} \sum_{m=1}^n \frac{2 \cos \frac{m\pi}{2} - 1 - (-1)^m}{m^2} \cos m\pi t.$$

$$(\text{proj}_{V_n} f)(t) = \frac{4}{\pi^2} \sum_{m=1}^n \frac{\text{sen } \frac{m\pi}{2}}{m^2} \text{sen } m\pi t$$

$$2.2.15. (\text{proj}_{W_n} f)(t) =$$

$$\frac{3}{16} + \frac{2}{\pi^2} \sum_{m=1}^n \frac{\cos \frac{m\pi}{4} + \cos \frac{3m\pi}{4} - 1 - (-1)^m}{m^2} \cos m\pi t.$$

$$\left(\text{proj}_{V_n} f \right) (t) =$$

$$\frac{2}{\pi^2} \sum_{m=1}^n \frac{\text{sen} \frac{m\pi}{4} + \text{sen} \frac{3m\pi}{4}}{m^2} \text{sen} m\pi t$$

- 2.2.16. (a)

```
>> syms t a
>> fa=exp(t)-a*exp(2*t)
fa = exp(t)-a*exp(2*t)
>> pa=simplify(proj(sym(1),fa,-1,1))
pa=1/4*(2*exp(-1)-a-2*exp(-3)+a*exp(-4))*exp(2)
>> pa=simplify([pa;...
proj(cos(pi*t),fa,-1,1);...
proj(sin(pi*t),fa,-1,1)]);
>> pa=simplify([pa;...
proj(cos(2*pi*t),fa,-1,1);...
proj(sin(2*pi*t),fa,-1,1)]);
>> f=subs(fa,a,...
(exp(1)-exp(-1))/(exp(2)-exp(-2)));
>> p=subs(pa,a,...
(exp(1)-exp(-1))/(exp(2)-exp(-2)));
>> prj=[p(1);sum(p(1:3));sum(p(1:5))]
>> plotfproj(f,prj,-1,1)
```
- (b)

```
>> pa=proj(pl(1),fa,-1,1);
>> pa=simplify([pa;...
proj(pl(2),fa,-1,1);...
proj(pl(3),fa,-1,1)]);
>> p=subs(pa,a,...
(exp(1)-exp(-1))/(exp(2)-exp(-2)));
>> prj=[p(1);sum(p(1:2));sum(p(1:3))]
>> plotfproj(f,prj,-1,1)
```

3.1. Definição, Exemplos e Propriedades (página 218)

3.1.1. (a) `>> V1=[1;1];V2=[0;1]; A=[V1,V2];`
`>> escalona([A,[3;-2]])`

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

$$(3, -2) = 3V_1 - 5V_2. \text{ Assim, } T(3, -2) = 3T(V_1) - 5T(V_2) = 3(2, -3) - 5(1, 2) = (1, -19).$$

(b) `>> syms a b`
`>> escalona([A,[a;b]])`

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b-a \end{bmatrix}$$

$$(a, b) = aV_1 + (b - a)V_2. \text{ Assim, } T(a, b) = aT(V_1) + (b - a)T(V_2) = a(2, -3) + (b - a)(1, 2) = (a + b, -5a + 2b).$$

3.1.2. `>> V1=[1;1];V2=[0;-2]; A=[V1,V2];`
`>> escalona([A,[1;0],[0;1]])`

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1/2 & -1/2 \end{bmatrix}$$

$$(1, 0) = V_1 + 1/2V_2 \text{ e } (0, 1) = -1/2V_2. \text{ Assim, } T(1, 0) = T(V_1) + 1/2T(V_2) = (3, 2, 1) + 1/2(0, 1, 0) = (3, 5/2, 1) \text{ e } T(0, 1) = -1/2T(V_2) = -1/2(0, 1, 0) = (0, -1/2, 0).$$

3.1.3. $P_{xy}(x, y, z) = (x, y, 0)$, $P_{yz}(x, y, z) = (0, y, z)$ e $P_{xz}(x, y, z) = (x, 0, z)$.

3.1.4. (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) $R_\pi(x, y, z) = 2P_\pi(x, y, z) - (x, y, z) = (x, y, z) - 2P_{(1,2,3)}(x, y, z) = 1/7(6x - 2y - 3z, 3y - 2x - 6z, -2z - 3x - 6y)$.

3.1.5. $R_{\pi/3,x}(E_1) = E_1, R_{\pi/3,x}(E_2) = 1/2E_2 + \sqrt{3}/2E_3, R_{\pi/3,x}(E_3) = -\sqrt{3}/2E_2 + 1/2E_3$. 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, $R_{\pi/3,y} \begin{bmatrix} x \\ y \\ z \end{bmatrix} =$

$$\begin{bmatrix} 1/2 & 0 & \sqrt{3}/2 \\ 0 & 1 & 0 \\ -\sqrt{3}/2 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

$$R_{\pi/3,z}(E_3) = E_3, R_{\pi/3,z}(E_1) = 1/2E_1 + \sqrt{3}/2E_2, R_{\pi/3,z}(E_2) = -\sqrt{3}/2E_1 + 1/2E_2. \quad \text{Portanto, } R_{\pi/3,z} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

3.1.6. (a) $P_r(x, y, z) = \text{proj}_{(1,1,1)}(x, y, z) = \frac{\langle (x,y,z), (1,1,1) \rangle}{\|(1,1,1)\|^2} = \frac{x+y+z}{3}(1, 1, 1) = \left(\frac{x+y+z}{3}, \frac{x+y+z}{3}, \frac{x+y+z}{3}\right).$

(b) $R_r(x, y, z) = 2\text{proj}_{(1,1,0)}(x, y, z) - (x, y, z) = 2\left(\frac{x+y+z}{3}, \frac{x+y+z}{3}, \frac{x+y+z}{3}\right) - (x, y, z) = \left(\frac{-x+2y+2z}{3}, \frac{2x-y+2z}{3}, \frac{2x+2y-z}{3}\right).$

3.1.7. 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 $[R_r]_{\mathcal{B}} = \begin{bmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{bmatrix}$.

3.2. A Imagem e o Núcleo (página 236)

3.2.1. Esta transformação é a projeção no plano xy . Se a imagem da reta r é um ponto $P_0 = (x_0, y_0, 0)$, então a reta é perpendicular ao plano xy , ou seja, passa por P_0 e tem como vetor diretor um vetor perpendicular ao plano xy , por exemplo $V = (0, 0, 1)$. Assim, as equações paramétricas da reta r são da forma $(x, y, z) = (x_0, y_0, 0) + t(0, 0, 1)$ ou $x = x_0, y = y_0, z = t, \forall t \in \mathbb{R}$.

3.2.2. (a) Seja $\mathcal{B} = \{E_1, E_2, E_3\}$ a base canônica. $[T]_{\mathcal{B}}^{\mathcal{B}} = [T(E_1)T(E_2)T(E_3)] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

```
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 $[T]_{\mathcal{B}}$. 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 $(1, 0, -1)$.

3.2.3. (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.

3.2.4. O núcleo de D é o conjunto dos polinômios de grau menor ou igual a 3 cuja derivada é igual ao polinômio nulo. Ou seja, $\{p_1(t) = 1\}$ é uma base para $\mathcal{N}(D)$. Seja $p(t) = a_0 + a_1t + a_2t^2 + a_3t^3 \in \mathcal{P}_3$. Então, $D(p) = a_1 + 2a_2t + 3a_3t^2$, ou seja, $1, 2t, 3t^2$ geram a imagem de D e como eles são L.I., $\{1, 2t, 3t^2\}$ é base para $\mathcal{I}(D)$.

3.2.5. (a) $\dim(\mathbb{V}) = \dim(\mathcal{N}(T)) + \dim(\mathcal{I}(T)) = 2 + 5 = 7$.

(b) $\dim(\mathbb{V}) = \dim(\mathcal{N}(T)) + \dim(\mathcal{I}(T)) = 0 + 5 = 5$.

3.2.6. (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: $W = \vec{0}$, então $T(E_3) = \vec{0}$. A transformação T está totalmente caracterizada por seus valores em uma base do domínio.

3.2.7. Seja $\{V_1, V_2\}$ uma base de \mathbb{R}^2 . Defina $T(V_1) = \vec{0}$ e $T(V_2) = V_1$. Ex.: $T(1,0) = (0,0)$ e $T(0,1) = (1,0)$.

3.2.8. Basta mostrarmos que $\mathcal{N}(T) = \{\vec{0}\}$. Seja $p(t) = a_0 + a_1t + \dots + a_n t^n \in \mathcal{P}_n$. $T(p) = (a_0 + a_1) + (a_1 + 2a_2)t + \dots + (a_{n-1} + na_n)t^{n-1} + a_n t^n$. $T(p) = \vec{0}$ implica que $a_n = a_{n-1} = \dots = a_0 = 0$. Ou seja, $p = \vec{0}$ e $\mathcal{N}(T) = \{\vec{0}\}$. $\dim(\mathcal{I}(T)) = \dim(\mathcal{P}_n) - \mathcal{N}(T) = \dim(\mathcal{P}_n)$. Logo, T é injetiva e sobrejetiva e portanto um isomorfismo.

3.3. Composição de Transformações Lineares (página 264)

3.3.1. Seja $\mathcal{B} = \{E_1, E_2\}$ a base canônica de \mathbb{R}^2 . Como $P = [I]_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, $A = [T]_{\mathcal{B}}^{\mathcal{B}} = [T(E_1)T(E_2)] = \begin{bmatrix} 2 & -1 \\ 1 & -3 \end{bmatrix}$ e $P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$. Então, $[T]_{\mathcal{C}}^{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}[T]_{\mathcal{B}}^{\mathcal{B}}[I]_{\mathcal{C}}^{\mathcal{B}} = P^{-1}AP = \begin{bmatrix} 8 & 13 \\ -5 & -9 \end{bmatrix}$.

3.3.2. Seja $\mathcal{B} = \{1, t\}$. Então, $P = [I]_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $A = [D]_{\mathcal{B}}^{\mathcal{B}} = [[D(1)]_{\mathcal{B}}[D(t)]_{\mathcal{B}}] = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ e $P^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$. Portanto, $[D]_{\mathcal{C}}^{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}[D]_{\mathcal{B}}^{\mathcal{B}}[I]_{\mathcal{C}}^{\mathcal{B}} = P^{-1}AP = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix}$.

3.3.3. (a) $P = [I]_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 1 \end{bmatrix}$.

(b) $B = [T]_{\mathcal{C}}^{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}[T]_{\mathcal{B}}^{\mathcal{B}}[I]_{\mathcal{C}}^{\mathcal{B}} = P^{-1}AP$. 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

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

$$\text{Assim, } B = [X_1 X_2 X_3] = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & -1 \end{bmatrix}.$$

3.3.4. (a) $T(1) = 0 \cdot 1 + 0 \cdot t + 0 \cdot t^2$. $T(t) = 0 \cdot 1 + 1 \cdot t + 0 \cdot t^2$. $T(t^2) = 2 \cdot 1 + 0 \cdot t + 2 \cdot t^2$. Assim, $A = [T]_B^B = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

(b) $T(1+t^2) = T(t^2) = 0 \cdot 1 + 0 \cdot t + 2 \cdot (1+t^2)$. Assim, $B = [T]_C^C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

(c) $B = [T]_C^C = [I]_B^C [T]_B^B [I]_C^B$. Assim, $P = [I]_C^B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(d) Se $p(t) = a_0 + a_1 t + a_2 (1+t^2)$, então

$$[T(p)]_C = [T]_C^C [p]_C = B[p]_C.$$

$$[T^2(p)]_C = [T]_C^C [T(p)]_C = B^2[p]_C.$$

$$[T^n(p)]_C = [T]_C^C [T^{n-1}(p)]_C = B^n [p]_C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1^n & 0 \\ 0 & 0 & 2^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ a_1 \\ 2^n a_2 \end{bmatrix}$$

$$T^n(p) = a_1 t + (2^n a_2)(1+t^2).$$

3.3.5. (a) $P_r(U_1) = U_1 = 1U_1 + 0U_2 + 0U_3$,

$$P_r(U_2) = \vec{0} = 0U_1 + 0U_2 + 0U_3 \text{ e}$$

$$P_r(U_3) = \vec{0} = 0U_1 + 0U_2 + 0U_3.$$

$$\text{Assim, } [P_r]_C^C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$$\text{Seja } P = [I]_C^B = \begin{bmatrix} \sqrt{3}/3 & \sqrt{6}/3 & 0 \\ \sqrt{3}/3 & -\sqrt{6}/6 & \sqrt{2}/2 \\ \sqrt{3}/3 & -\sqrt{6}/3 & -\sqrt{2}/2 \end{bmatrix}.$$

$$P_r(E_1) = \text{proj}_{(1,1,1)} E_1 = \frac{\langle E_1, (1,1,1) \rangle}{\|(1,1,1)\|^2} (1,1,1) = (1/3, 1/3, 1/3) = 1/3E_1 + 1/3E_2 + 1/3E_3$$

$$P_r(E_2) = \text{proj}_{(1,1,1)} E_2 = \frac{\langle E_2, (1,1,1) \rangle}{\|(1,1,1)\|^2} (1,1,1) = (1/3, 1/3, 1/3) = 1/3E_1 + 1/3E_2 + 1/3E_3$$

$$P_r(E_3) = \text{proj}_{(1,1,1)} E_3 = \frac{\langle E_3, (1,1,1) \rangle}{\|(1,1,1)\|^2} (1, 1, 1) = (1/3, 1/3, 1/3) = 1/3E_1 + 1/3E_2 + 1/3E_3$$

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, $[R_r]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}.$

3.3.6. $R_{\pi/2,r}(U_1) = U_1 = 1U_1 + 0U_2 + 0U_3,$

$$R_{\pi/2,r}(U_2) = U_3 = 0U_1 + 0U_2 + 1U_3,$$

$$R_{\pi/2,r}(U_3) = -U_2 = 0U_1 - 1U_2 + 0U_3.$$

Assim, $[R_{\pi/2,r}]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$

$$E_1 = \langle E_1, U_1 \rangle U_1 + \langle E_1, U_2 \rangle U_2 + \langle E_1, U_3 \rangle U_3 = 1/\sqrt{2}U_1 + 1/\sqrt{2}U_3$$

$$R_{\pi/2,r}(E_1) = 1/\sqrt{2}R_{\pi/2,r}(U_1) + 1/\sqrt{2}R_{\pi/2,r}(U_3) = 1/\sqrt{2}U_1 - 1/\sqrt{2}U_2 = 1/\sqrt{2}(1/\sqrt{2}, 1/\sqrt{2}, 0) - 1/\sqrt{2}(0, 0, 1) = (1/2, 1/2, -1/\sqrt{2})$$

$$E_2 = \langle E_2, U_1 \rangle U_1 + \langle E_2, U_2 \rangle U_2 + \langle E_2, U_3 \rangle U_3 = 1/\sqrt{2}U_1 - 1/\sqrt{2}U_3$$

$$R_{\pi/2,r}(E_2) = 1/\sqrt{2}R_{\pi/2,r}(U_1) - 1/\sqrt{2}R_{\pi/2,r}(U_3) = 1/\sqrt{2}U_1 + 1/\sqrt{2}U_3 = 1/\sqrt{2}(1/\sqrt{2}, 1/\sqrt{2}, 0) + 1/\sqrt{2}(0, 0, 1) = (1/2, 1/2, 1/\sqrt{2})$$

$$E_3 = \langle E_3, U_1 \rangle U_1 + \langle E_3, U_2 \rangle U_2 + \langle E_3, U_3 \rangle U_3 = U_2$$

$$R_{\pi/2,r}(E_3) = R_{\pi/2,r}(U_2) = U_3 = (1/\sqrt{2}, -1/\sqrt{2}, 0)$$

$$\text{Assim, } [R_{\pi/2,r}]_B^B = \begin{bmatrix} 1/2 & 1/2 & 1/\sqrt{2} \\ 1/2 & 1/2 & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}.$$

3.3.7. (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 $\{1, t, t^2\}$ 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, $T^{-1}(a + bt + ct^2) = (a + 2b + 10c) + (b + 4c)t + ct^2$.

(c) A matriz de T em relação às bases $\{E_1, E_2, E_3\}$ e $\{1, t, t^2\}$ 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 $T^{-1}(a + bt + ct^2) = (4a - 2b - c, -a + b, -2a + b + c)$.

(d) A matriz de T em relação às bases $\{1, t, t^2\}$ e $\{E_1, E_2, E_3\}$ 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 $T^{-1}(a, b, c) = b + (c/2 - a/2)t + (a/2 - b + c/2)t^2$.

3.4. A Adjunta (página 284)

```

3.4.1. (a) >> A=[4,-2;1,3;2,1;3,4];
>> escalona(A)
[ 1, 0]
[ 0, 1]
[ 0, 0]
[ 0, 0]
>> escalona(A')
[ 1, 0, 5/14, 5/14]
[ 0, 1, 4/7, 11/7]

```

$\mathcal{N}(A) = \{\vec{0}\}$, $\mathcal{I}(A^t) = \mathbb{R}^2$. Portanto, $\mathcal{N}(A)$ não tem base e $\{(1,0), (0,1)\}$ é base de $\mathcal{I}(A^t)$. $\mathcal{N}(A^t) = \{((-5/14)\beta - (5/14)\alpha, (-4/7)\beta - (11/7)\alpha, \beta, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$. Portanto, $\{(-5, -22, 0, 14), (-5, -8, 14, 0)\}$ é base de $\mathcal{N}(A^t)$. $\{(1, 0, 5/14, 5/14), (0, 1, 4/7, 11/7)\}$ é base para $\mathcal{I}(A)$.

```

(b) >> A=[1,0,0,0;0,1,1,1;0,0,1,1;1,1,2,2];
>> escalona(A)
[ 1, 0, 0, 0]
[ 0, 1, 0, 0]

```

```
[ 0, 0, 1, 1]
[ 0, 0, 0, 0]
>> escalona(A')
[ 1, 0, 0, 1]
[ 0, 1, 0, 1]
[ 0, 0, 1, 1]
[ 0, 0, 0, 0]
```

$\mathcal{N}(A) = \{(0, 0, -\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. Portanto, $\{(0, 0, -1, 1)\}$ é base para $\mathcal{N}(A)$. $\{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 1)\}$ é base para $\mathcal{I}(A^t)$. $\mathcal{N}(A^t) = \{(-\alpha, -\alpha, -\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. Portanto, $\{(-1, -1, -1, 0)\}$ é base para $\mathcal{N}(A^t)$. $\{(1, 0, 0, 1), (0, 1, 0, 1), (0, 0, 1, 1)\}$ é base para $\mathcal{I}(A)$.

3.4.2. $\mathbb{W}^\perp = \{(x, y, z) \mid x - y + z = 0\}$. Assim, $X \in \mathbb{W}^\perp$ se, e somente se, $X = (\beta - \alpha, \beta, \alpha) = \alpha(-1, 0, 1) + \beta(1, 1, 0)$. Portanto, $W_1 = (-1, 0, 1)$ e $W_2 = (1, 1, 0)$ formam uma base de \mathbb{W}^\perp . \mathbb{W} é uma reta que passa pela origem e \mathbb{W}^\perp é um plano que passa pela origem, perpendicular a \mathbb{W} .

3.4.3. >> A=[1,0,-2,1;0,1,3,-2]
A=[1, 0, -2, 1]
[0, 1, 3, -2]

$X = (2\beta - \alpha, 2\alpha - 3\beta, \beta, \alpha) = \alpha(-1, 2, 0, 1) + \beta(2, -3, 1, 0), \forall \alpha, \beta \in \mathbb{R}$. Assim, $W_1 = (-1, 2, 0, 1)$ e $W_2 = (2, -3, 1, 0)$ formam uma base para \mathbb{W}^\perp .

3.4.4. (a) >> P=[-2,1;-1,2;1,0;2,7];
>> A=matvand(P(:,1),2),B=P(:,2)
A = 4 -2 1
1 -1 1
1 1 1
4 2 1
B = 1
2
0
7
>> escalona([A'*A,A'*B])
[34, 0, 10, 34]
[0, 10, 0, 10]
[10, 0, 4, 10]
[1, 0, 0, 1]
[0, 1, 0, 1]
[0, 0, 1, 0]

A equação da parábola é $y = x^2 + x$.

(b) >> P=[-2,1;-1,3;1,3;2,11];
>> A=matvand(P(:,1),2),B=P(:,2)

```

A = 4  -2  1
    1  -1  1
    1   1  1
    4   2  1
B = 1
    3
    3
    11
>> escalona([A'*A,A'*B])
[ 34,  0, 10, 54]
[  0, 10,  0, 20]
[ 10,  0,  4, 18]
[  1,  0,  0,  1]
[  0,  1,  0,  2]
[  0,  0,  1,  2]

```

A equação é $y = x^2 + 2x + 2$.

3.4.5. (a)

```
>> P=[-2,0;0,1;1,-2;2,1];
>> A=matvand(P,1),B=P(:,1).^2+P(:,2).^2
```

```

A = -2   0   1
     0   1   1
     1  -2   1
     2   1   1
B = 4
    1
    5
    5
>> escalona([A'*A,A'*B])
[  9,  0,  1,  7]
[  0,  6,  0, -4]
[  1,  0,  4, 15]
[   1,   0,   0,   13/35]
[   0,   1,   0,  -2/3]
[   0,   0,   1, 128/35]

```

A equação do círculo é $x^2 + y^2 - (13/35)x + (2/3)y = 128/35$.

(b)

```
>> P=[-2,1;-1,-2;0,1;2,0];
>> A=matvand(P,1),B=P(:,1).^2+P(:,2).^2
```

```

A = -2   1   1
     -1  -2   1
     0   1   1
     2   0   1
B = 5
    5

```

```

      1
      4
>> escalona([A'*A,A'*B])
[ 9, 0, -1, -7]
[ 0, 6, 0, -4]
[ -1, 0, 4, 15]
[ 1, 0, 0, 0, -13/35]
[ 0, 1, 0, 0, -2/3]
[ 0, 0, 1, 128/35]

```

A equação do círculo é $x^2 + y^2 + (13/35)x + (2/3)y = 128/35$.

3.4.6. (a) >> A=[1,2;2,4;-1,-2];B=[3;2;1];
 >> escalona([A'*A,A'*B])
 [6, 12, 6]
 [12, 24, 12]
 [1, 2, 1]
 [0, 0, 0]

$X = (1 - 2\alpha, \alpha) = (1, 0) + \alpha(-2, 1), \forall \alpha \in \mathbb{R}$.

(b) >> A=[-1,1;2,1;1,-2];B=[10;5;20];
 >> escalona([A'*A,A'*B])
 [6, -1, 20]
 [-1, 6, -25]
 [1, 0, 19/7]
 [0, 1, -26/7]

$X = (19/7, -26/7)$.

(c) >> A=[1,1,1;-1,1,1;0,-1,1;1,0,1];B=[4;2;1;2];
 >> escalona([A'*A,A'*B])
 [3, 0, 1, 4]
 [0, 3, 1, 5]
 [1, 1, 4, 9]
 [1, 0, 0, 11/15]
 [0, 1, 0, 16/15]
 [0, 0, 1, 9/5]

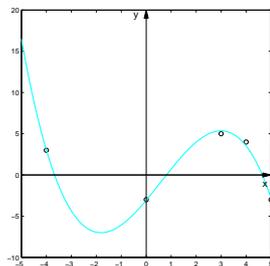
$X = (11/15, 16/15, 9/15)$.

3.4.7. >> P=randi(5,2)
 P = 3 5
 5 -3
 0 -3
 4 4
 -4 3

```

>> A=matvand(P(:,1),3), B=P(:,2)
A = 27    9    3    1
    125   25    5    1
         0    0    0    1
        64   16    4    1
       -64   16   -4    1
B = 5
    -3
    -3
     4
     3
>> R=escalona([A'*A,A'*B])
[ 24546, 3368, 1218, 152, -176]
[ 3368, 1218, 152, 66, 82]
[ 1218, 152, 66, 8, 4]
[ 152, 66, 8, 5, 6]
R = [ 1,0,0,0, -35077/157992]
     [ 0,1,0,0, 33866/85579]
     [ 0,0,1,0, 7430353/2053896]
     [ 0,0,0,1, -262092/85579]
>> a=R(1,5);b=R(2,5);c=R(3,5);d=R(4,5);
>> clf,po(P),syms x,plotf1(a*x^3+b*x^2+c*x+d,[-5,5])
>> eixos

```



3.4.8. >> P=randi(6,2)

```

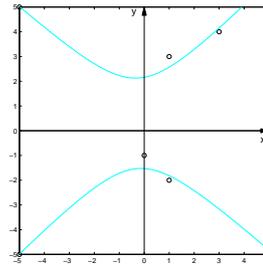
P = 0    -1
     1    -2
     3     4
    -5    -5
     1     3
    -5     5

```

```

>> M=matvand(P,2),B=-M(:,1),A=M(:,2:6)
M =  0    0    1    0   -1    1
     1   -2    4    1   -2    1
     9   12   16    3    4    1
    25   25   25   -5   -5    1
     1    3    9    1    3    1
    25  -25   25   -5    5    1
B =  0
     -1
     -9
    -25
     -1
    -25
A =  0    1    0   -1    1
     -2   4    1   -2    1
     12  16    3    4    1
     25  25   -5   -5    1
     3    9    1    3    1
    -25  25   -5    5    1
>> R=escalona([A'*A,A'*B])
[ 1407,   211,    37,  -189,   13,  -109]
[   211,  1604,  -189,   82,   80, -1407]
[    37,  -189,   61,   13,   -5,   221]
[  -189,   82,   13,   80,    4,  -37]
[    13,   80,   -5,    4,    6,  -61]
R = [1,0,0,0,0,  35943/287650]
     [0,1,0,0,0, -301491/287650]
     [0,0,1,0,0,  127343/287650]
     [0,0,0,1,0,  95187/143825]
     [0,0,0,0,1,  18123/5230]
>> a=R(1,6);b=R(2,6);c=R(3,6);
>> d=R(4,6);e=R(5,6);
>> clf,po(P),syms x y
>> plotci(x^2+a*x*y+b*y^2+c*x+d*y+e, [-5,5], [-5,5])
>> eixos

```



4.1. Diagonalização de Operadores (página 320)

4.1.1.

```
(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]

```

$$V_{-2} = \{(0, 0, \alpha) \mid \alpha \in \mathbb{R}\}$$

$$V_1 = \{(6\alpha, 3\alpha, 8\alpha) \mid \alpha \in \mathbb{R}\}$$

$$V_3 = \{(0, 5\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$$

(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]

```

$$V_1 = \{(-\alpha, \alpha, \alpha) \mid \alpha \in \mathbb{R}\}$$

$$V_2 = \{(\alpha, 0, 0) \mid \alpha \in \mathbb{R}\}$$

$$\mathbb{V}_4 = \{(7\alpha, -4\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$$

(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]

```

$$\mathbb{V}_{-1} = \{(-\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}, \mathbb{V}_2 = \{(-2\alpha, -3\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\} \text{ e } \mathbb{V}_4 = \{(8\alpha, 5\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$$

4.1.2. (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. ($xV_1 + yV_2 = \vec{0}$ 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.

4.1.3.

```

(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 4.5 na página 306). A matriz A é diagonalizável pois, é 2×2 e possui dois autovetores L.I. (Teorema 4.3 na página 304).

```

(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 4.3 na página 304).

```

(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 4.5 na página 306). A matriz A é diagonalizável pois, é 3×3 e possui três autovetores L.I. (Teorema 4.3 na página 304).

```
(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 4.5 na página 306). A matriz A é diagonalizável pois, é
3 x 3 e possui três autovetores L.I. (Teorema 4.3 na página 304).
```

4.1.4.

```
(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]
```

$V_1 = \{(\beta, -2\alpha, \alpha) \mid \alpha, \beta \in \mathbb{R}\}$. $\{(1,0,0), (0,-2,1)\}$ é base para V_1 , pois gera 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)

$V_3 = \{(\alpha, 0, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{(1,0,1)\}$ é base para V_3 , pois gera 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 4.3 na página 304**).

```
(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]
>> 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]
```

```
>> 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]
>> 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]
```

$V_0 = \{(-\alpha, 0, 3\alpha) \mid \alpha \in \mathbb{R}\}$. $\{(-1, 0, 3)\}$ é base para V_0 , pois gera V_0 ($(-\alpha, 0, 3\alpha) = \alpha(-1, 0, 3)$) e um vetor não nulo é L.I.

$V_2 = \{(2\alpha, \alpha, 0) \mid \alpha \in \mathbb{R}\}$. $\{(2, 1, 0)\}$ é base para V_2 , pois gera V_2 ($(2\alpha, \alpha, 0) = \alpha(2, 1, 0)$) e um vetor não nulo é L.I.

$V_3 = \{(\alpha, 0, 0) \mid \alpha \in \mathbb{R}\}$. $\{(1, 0, 0)\}$ é base para V_3 , pois gera V_3 ($(\alpha, 0, 0) = \alpha(1, 0, 0)$) e um vetor não nulo é L.I.

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

```
4.1.5. (a) >> A=[2,1,2;0,0,0;0,1,3]; syms t
>> p=det(A-t*eye(3))
p =
-(2-t)*t*(3-t)
>> B=matvand([0;2;3],2)
B = 0 0 1
    4 2 1
    9 3 1
>> R=escalona([B,[0;2^k;3^k]])
[ 0, 0, 1, 0]
[ 4, 2, 1, 2^k]
[ 9, 3, 1, 3^k]
```

```

R=[1,0,0,      -2^(-1+k)+3^(-1+k)]
  [0,1,0, -2*3^(-1+k)+3*2^(-1+k)]
  [0,0,1,      0]
>> simplify(R(1,4)*A^2+R(2,4)*A)
[      2^k, -2^(-1+k)+2*3^(-1+k),      -2^(1+k)+2*3^k]
[      0,      0,      0]
[      0,      3^(-1+k),      3^k]

```

$$A^k = \begin{bmatrix} 2^k & -2^{-1+k} + 2 \cdot 3^{-1+k} & -2^{1+k} + 2 \cdot 3^k \\ 0 & 0 & 0 \\ 0 & 3^{-1+k} & 3^k \end{bmatrix}$$

```

(b) >> A=[3,0,-3;1,3,-1;0,0,0]
A = 3      0      -3
    1      3      -1
    0      0      0
>> p=det(A-t*eye(3))
p = -(3-t)^2*t
>> B=matvand([0;3],2)
B = 0      0      1
    9      3      1
>> B=[B;[6,1,0]]
B = 0      0      1
    9      3      1
    6      1      0
>> R=escalone([B,[0;3^k;k*3^(k-1)]])
[      0,      0,      1,      0]
[      9,      3,      1,      3^k]
[      6,      1,      0,      k*3^(-1+k)]
R=[1,0,0,      -1/9*3^k+1/9*k*3^k]
  [0,1,0, -k*3^(-1+k)+2*3^(-1+k)]
  [0,0,1,      0]
>> Ak=simplify(R(1,4)*A^2+R(2,4)*A)
Ak =
[      3^k,      0,      -3^k]
[ k*3^(-1+k),      3^k, -k*3^(-1+k)]
[      0,      0,      0]

```

$$A^k = \begin{bmatrix} 3^k & 0 & -3^k \\ k \cdot 3^{-1+k} & 3^k & -k \cdot 3^{-1+k} \\ 0 & 0 & 0 \end{bmatrix}$$

```
(c) >> A=[0,-2,3;0,1,0;0,-2,3]
A = 0 -2 3
    0 1 0
    0 -2 3
>> p=det(A-t*eye(3))
p = -t*(1-t)*(3-t)
>> B=matvand([0;1;3],2)
B = 0 0 1
    1 1 1
    9 3 1
>> R=escalona([B,[0;1;3^k]])
[ 0, 0, 1, 0]
[ 1, 1, 1, 1]
[ 9, 3, 1, 3^k]
>> Ak=simplify(R(1,4)*A^2+R(2,4)*A)
Ak =
[ 0, 1-3^k, 3^k]
[ 0, 1, 0]
[ 0, 1-3^k, 3^k]
```

$$A^k = \begin{bmatrix} 0 & 1-3^k & 3^k \\ 0 & 1 & 0 \\ 0 & 1-3^k & 3^k \end{bmatrix}$$

```
(d) >> A=[0,0,0;0,0,-1;-1,1,0]
A = 0 0 0
    0 0 -1
   -1 1 0
>> p=det(A-t*eye(3))
p = -t*(t^2+1)
>> solve(p)
[ 0][ i][ -i]
>> B=matvand([0;i;-i],2)
B =[ 0, 0, 1]
   [-1, i, 1]
   [-1, -i, 1]
>> escalona([B,[0;i^k;(-i)^k]])
[ 0, 0, 1, 0]
[ -1, i, 1, i^k]
[ -1, -i, 1, (-i)^k]
R =
[1,0,0, -cos(1/2*k*pi)]
[0,1,0, sin(1/2*k*pi)]
[0,0,1, 0]
```

```
>> Ak=simplify(R(1,4)*A^2+R(2,4)*A)
Ak =
[      0,      0,      0]
[-cos(1/2*k*pi), cos(1/2*k*pi), -sin(1/2*k*pi)]
[-sin(1/2*k*pi), sin(1/2*k*pi), cos(1/2*k*pi)]
```

$$A^k = \begin{bmatrix} 0 & 0 & 0 \\ -\cos(1/2k\pi) & \cos(1/2k\pi) & -\sin(1/2k\pi) \\ -\sin(1/2k\pi) & \sin(1/2k\pi) & \cos(1/2k\pi) \end{bmatrix}$$

```
4.1.6. >> 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.

```
4.1.7. >> 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

$$V_0 = \{(-2\alpha, 2\alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

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 .

4.2. Operadores Auto-adjuntos e Normais (página 340)

4.2.1.

```
(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]
```

$\mathbb{V}_0 = \{(-\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_1 = (-1, 1)\}$ 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]
```

$\mathbb{V}_1 = \{(-\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_1 = (-1, 1)\}$ 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 .

```
>> B3=subs(B,x,3)
[-1, 1]
[ 1, -1]
>> escalona(B3)
[1, -1]
[0, 0]
```

$\mathbb{V}_3 = \{(\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_2 = (1, 1)\}$ 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]
>> 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]
>> 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]
>> 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 4.10 na página 329**). 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. ($xV_1 + yV_2 = \vec{0}$ se, e somente se, $(x, -y, y) = (0, 0, 0)$ ou $x = 0$ e $y = 0$). Sejam $W_1 = V_1$, $W_2 = V_2 - \text{proj}_{W_1} V_2 = V_2 - \vec{0} = V_2$. 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 4.10 na página 329](#)). 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]
>> 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]
```

```

>> 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]

>> 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}_2 , pois gera \mathbb{V}_2 ($(\alpha, \alpha, 0) = \alpha(1, 1, 0)$) 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}_2 .

Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais (Proposição 4.10 na página 329). 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}_1 ($(-\alpha - \beta, \alpha, \beta) = \alpha(-1, 1, 0) + \beta(-1, 0, 1)$) 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 4.10 na página 329). 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 $W_1 = V_1$, $W_2 = V_2 - \text{proj}_{W_1} V_2 = V_2 - \bar{0} = V_2$. 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}_3 ($(\alpha, \alpha, \beta, \beta) = \alpha(1, 1, 0, 0) + \beta(0, 0, 1, 1)$) 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 4.10 na página 329).

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]
>> 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]
```

```
>> 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]
>> 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]
```

$V_0 = \{(\alpha, \beta, 0, 0) \mid \alpha, \beta \in \mathbb{R}\}$. $\{V_1 = (1, 0, 0, 0), V_2 = (0, 1, 0, 0)\}$ é base para V_0 , pois gera 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 V_0 .

$V_1 = \{(0, 0, -\alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_3 = (0, 0, -1, 1)\}$ é base para V_1 , pois gera 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 V_1 .

$V_{-1} = \{(0, 0, \alpha, \alpha) \mid \alpha \in \mathbb{R}\}$. $\{V_4 = (0, 0, 1, 1)\}$ é base para V_{-1} , pois gera 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 V_{-1} . Como a matriz A é simétrica, autovetores associados a autovalores diferentes são ortogonais ([Proposição 4.10 na página 329](#)). Portanto,

$\{U_1, U_2, U_3, U_4\}$ é uma base ortonormal de autovetores de A .

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

4.3. Aplicação ao Estudo de Cônicas (página 355)

```
4.3.1. >> A=[9,-2;-2,6]; K=[-10,-20];
>> syms x y; X=[x;y];
>> expr=simplify(X.'*A*X+K*X-5)

9x2 - 4xy + 6y2 - 10x - 20y - 5
>> [P,D]=diagonal(A)

P = [ sqrt(5)/5  -2*sqrt(5)/5 ]
     [ 2*sqrt(5)/5  sqrt(5)/5 ]

D=[5, 0]
   [0,10]
>> syms x1 y1; X1=[x1;y1];
>> expr=subst(expr,X,P*X1)

5x12 + 10y12 - 10sqrt(5)x1 - 5
>> syms x2 y2; X2=[x2;y2]; X0=[5^(1/2);0];
>> expr=subst(expr,X1,X2+X0)

5x22 - 30 + 10y22
>> expr=expr/30

x22/6 + y22/3 - 1
>> ellipse(sqrt(6),sqrt(3),P,X0)
```

```

4.3.2. >> A=[3,-4;-4,-12];
>> K=[-30,-64];
>> expr=simplify(X.'*A*X+K*X)


$$3x^2 - 8xy - 12y^2 - 30x - 64y$$

>> [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)


$$-13x_1^2 + 4y_1^2 - 286\sqrt{17}x_1/17 + 56\sqrt{17}y_1/17$$

>> X0=[-286/(2*13*17^(1/2));-56/(2*4*17^(1/2))]
[-11*17^(1/2)/17]
[- 7*17^(1/2)/17]
>> expr=subst(expr,X1,X2+X0)

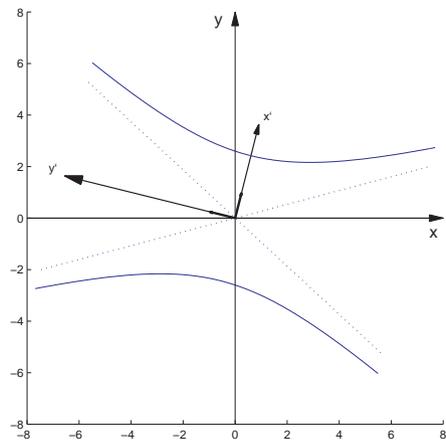
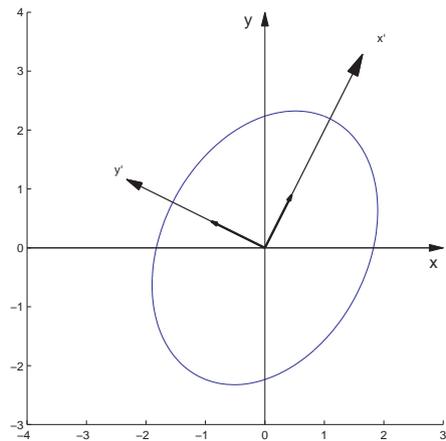

$$-13x_2^2 + 81 + 4y_2^2$$

>> expr=expr/81


$$-\frac{13}{81}x_2^2 + 1 + \frac{4}{81}y_2^2$$

>> hiperbx(9/sqrt(13),9/2,P,X0)

```



```
4.3.3. >> A=[2,-2;-2,-1];
>> K=[-4,-8];
>> expr=simplify(X.'*A*X+K*X+14)


$$2x^2 - 4xy - y^2 - 4x - 8y + 14$$


>> [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)


$$-2x_1^2 + 3y_1^2 - 4\sqrt{5}x_1 + 14$$


>> X0=[-5^(1/2);0];
>> expr=subst(expr,X1,X2+X0)


$$-2x_2^2 + 24 + 3y_2^2$$


>> expr=expr/24


$$-x_2^2/12 + y_2^2/8 + 1$$


>> hiperbx(sqrt(12),sqrt(8),P,X0)
```

```

4.3.4. >> A=[21,3;3,13];
>> K=[-114,34];
>> expr=simplify(X.'*A*X+K*X+73)

 $21x^2 + 6xy + 13y^2 - 114x + 34y + 73$ 

>> [P,D]=diagonal(A)


$$P = \begin{bmatrix} 3\sqrt{10}/10 & -1\sqrt{10}/10 \\ 1\sqrt{10}/10 & 3\sqrt{10}/10 \end{bmatrix}$$


D=[22, 0]
 [ 0,12]
>> expr=subst(expr,X,P*X1)

 $22x_1^2 + 12y_1^2 - \frac{154}{5}\sqrt{10}x_1 + \frac{108}{5}\sqrt{10}y_1 + 73$ 

>> X0=[154*10^(1/2)/(5*2*22);-108*10^(1/2)/(5*2*12)];
>> expr=subst(expr,X1,X2+X0)

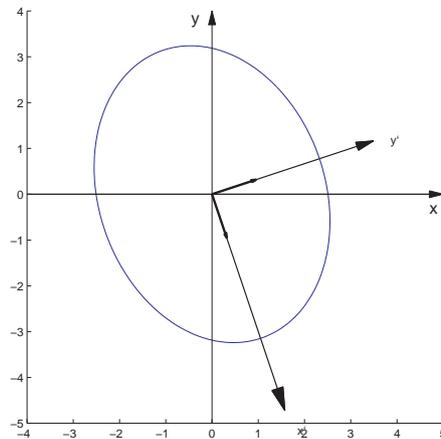
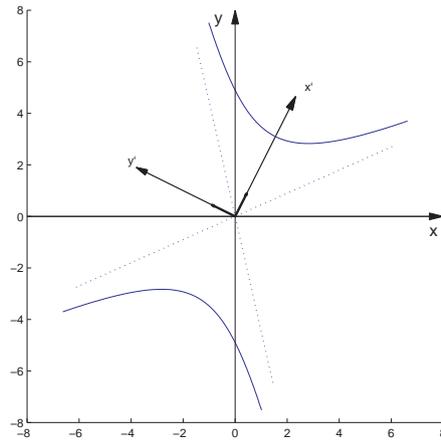
 $22x_2^2 - 132 + 12y_2^2$ 

>> expr=expr/132

 $x_2^2/6 + y_2^2/11 - 1$ 

>> ellipse(sqrt(6),sqrt(11),P,X0)

```



```
4.3.5. >> A=[4,-10;-10,25];
>> K=[-15,-6];
>> expr=simplify(X.'*A*X+K*X)
```

$$4x^2 - 20xy + 25y^2 - 15x - 6y$$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} \frac{5}{29}\sqrt{29} & -\frac{2}{29}\sqrt{29} \\ \frac{2}{29}\sqrt{29} & \frac{5}{29}\sqrt{29} \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 0 \\ 0 & 29 \end{bmatrix}$$

```
>> expr=subst(expr,X,P*X1)
```

$$29y_1^2 - 3\sqrt{29}x_1$$

```
>> expr=expr/29
```

$$y_1^2 - \frac{3}{29}\sqrt{29}x_1$$

```
>> parabx(3/(4*sqrt(29)),P)
```

```

4.3.6. >> A=[9,3;3,1]; K=[-10*10^(1/2),10*10^(1/2)];
>> expr=simplify(X.'*A*X+K*X+90)

9 x2 + 6 xy + y2 - 10 √10 x + 10 √10 y + 90
>> [P,D]=diagonal(A)

P = [ 3√10/10  -√10/10 ]
     [ √10/10   3√10/10 ]

D = [10, 0]
     [ 0, 0]
>> expr=subst(expr,X,P*X1)

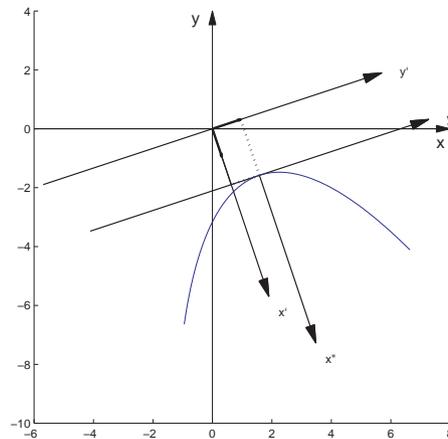
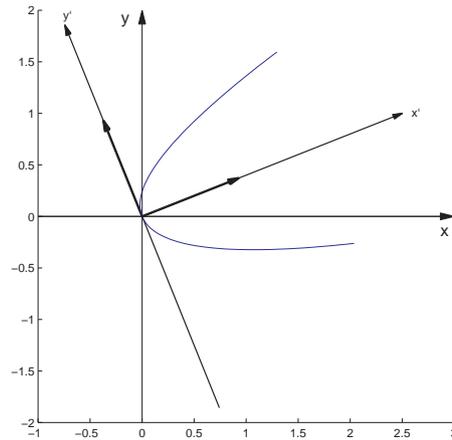
10 x12 - 20 x1 + 40 y1 + 90
>> X0=[154*10^(1/2)/(5*2*22);-108*10^(1/2)/(5*2*12)];
>> expr=subst(expr,x1,x2+1)

10 x22 + 80 + 40 y1
>> expr=subst(expr,y1,y2-2)

10 x22 + 40 y2
>> expr=expr/10

x22 + 4 y2
>> paraby(-1,P,[1;-2])

```



```
4.3.7. >> A=[5,-3;-3,5];
>> K=[-30*(2)^(1/2),18*(2)^(1/2)];
>> expr=simplify(X.'*A*X+K*X+82)


$$5x^2 - 6xy + 5y^2 - 30\sqrt{2}x + 18\sqrt{2}y + 82$$


>> [P,D]=diagonal(A)


$$P = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$


D = [2, 0]
     [0, 8]
>> expr=subst(expr,X,P*X1)


$$2x_1^2 + 8y_1^2 - 12x_1 + 48y_1 + 82$$


>> X0=[3;-3];
>> expr=subst(expr,X1,X2+X0)


$$2x_2^2 - 8 + 8y_2^2$$


>> expr=expr/8


$$x_2^2/4 - 1 + y_2^2$$


>> ellipse(2,1,P,X0)
```

```
4.3.8. >> A=[5,6;6,0];
>> K=[-12*(13)^(1/2),0];
>> expr=simplify(X.'*A*X+K*X-36)
```

$$5x^2 + 12xy - 12\sqrt{13}x - 36$$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} 2/\sqrt{13} & 3/\sqrt{13} \\ -3/\sqrt{13} & 2/\sqrt{13} \end{bmatrix}$$

$$D = \begin{bmatrix} -4 & 0 \\ 0 & 9 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 9 \end{bmatrix}$$

```
>> expr=subst(expr,X,P*X1)
```

$$-4x_1^2 + 9y_1^2 - 24x_1 - 36y_1 - 36$$

```
>> X0=[-3;2];
```

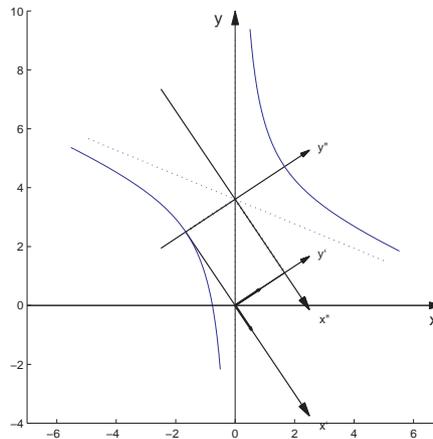
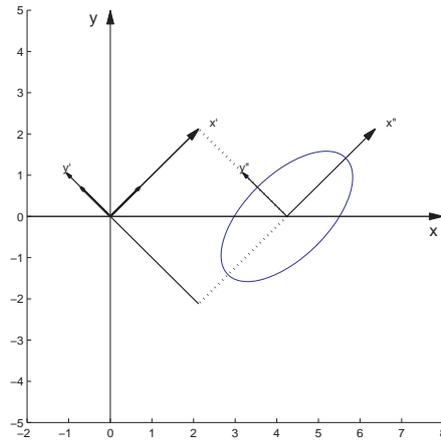
```
>> expr=subst(expr,X1,X2+X0)
```

$$-4x_2^2 - 36 + 9y_2^2$$

```
>> expr=expr/36
```

$$-x_2^2/9 - 1 + y_2^2/4$$

```
>> hiperby(2,3,P,X0)
```



```

4.3.9. >> A=[6,-2;-2,9];
>> K=[-4*5^(1/2),-18*5^(1/2)];
>> expr=simplify(X.'*A*X+K*X-5)


$$6x^2 - 4xy + 9y^2 - 4\sqrt{5}x - 18\sqrt{5}y - 5$$


>> [P,D]=diagonal(A)


$$P = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$


D = [5, 0]
     [0, 10]
>> expr=subst(expr,X,P*X1)


$$5x_1^2 + 10y_1^2 - 26x_1 - 32y_1 - 5$$


>> X0=[26/10;32/20];
>> expr=subst(expr,X1,X2+X0)


$$5x_2^2 - \frac{322}{5} + 10y_2^2$$


>> expr=expr*5/322


$$\frac{25}{322}x_2^2 - 1 + \frac{25}{161}y_2^2$$


>> ellipse(sqrt(322)/5,sqrt(161)/5,P,X0)

```

```

4.3.10. >> A=[1,3^(1/2);3^(1/2),-1];
>> K=[6,0];
>> expr=simplify(X.'*A*X+K*X)

 $x^2 + 2xy\sqrt{3} - y^2 + 6x$ 

>> [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)

 $2x_1^2 - 2y_1^2 + 3\sqrt{3}x_1 - 3y_1$ 

>> X0=[-3*3^(1/2)/4;-3/4];
>> expr=subst(expr,X1,X2+X0)

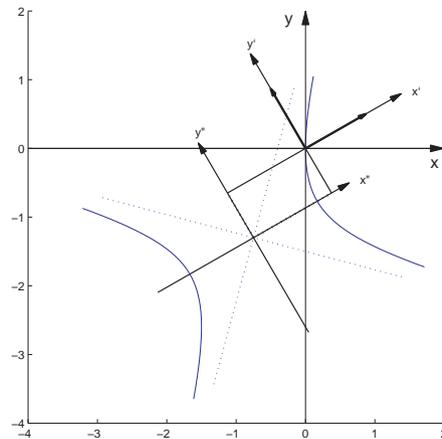
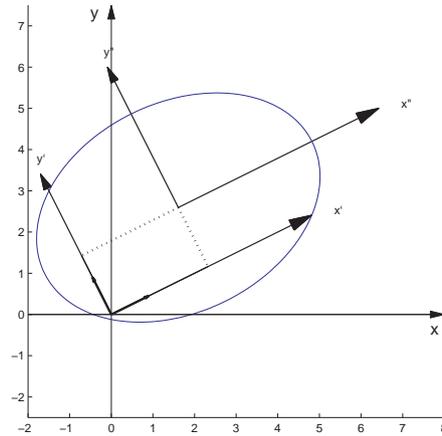
 $2x_2^2 - 9/4 - 2y_2^2$ 

>> 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)

```



```

4.3.11. >> A=[8,-8;-8,8];
>> K=[33*2^(1/2),-31*2^(1/2)];
>> expr=simplify(X.'*A*X+K*X+70)


$$8x^2 - 16xy + 8y^2 + 33\sqrt{2}x - 31\sqrt{2}y + 70$$


>> [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)


$$16y_1^2 + 2x_1 - 64y_1 + 70$$


>> expr=subst(expr,y1,y2+2)


$$16y_2^2 + 6 + 2x_1$$


>> expr=subst(expr,x1,x2-3)


$$16y_2^2 + 2x_2$$


>> expr=expr/16


$$y_2^2 + x_2/8$$


>> parabx(-1/32,P,[-3;2])

```

```
4.3.12. >> A=[1,-3;-3,-7];
>> K=[10,2];
>> expr=simplify(X.'*A*X+K*X+9)
```

$$x^2 - 6xy - 7y^2 + 10x + 2y + 9$$

```
>> [P,D]=diagonal(A)
```

$$P = \begin{bmatrix} 1/\sqrt{10} & -3/\sqrt{10} \\ 3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix}$$

$$D = \begin{bmatrix} -8 & 0 \\ 0 & 2 \end{bmatrix}$$

```
>> expr=subst(expr,X,P*X1)
```

$$-8x_1^2 + 2y_1^2 + \frac{8}{5}\sqrt{10}x_1 - \frac{14}{5}\sqrt{10}y_1 + 9$$

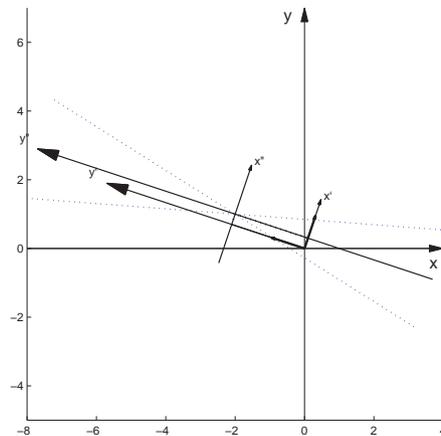
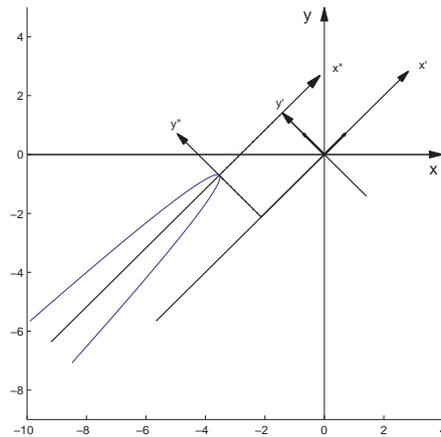
```
>> X0=[1/10^(1/2);7/10^(1/2)];
```

```
>> expr=subst(expr,X1,X2+X0)
```

$$-8x_2^2 + 2y_2^2$$

```
>> hiperby(4,1,P,X0,'d')
```

Esta é uma cônica degenerada. A equação representa as duas retas $y''^2 = 4x''^2$, ou $y'' = \pm 2x''$.



4.4. Forma Canônica de Jordan (página 394)

4.4.1. As matrizes dos itens (a) e (b) estão na forma canônica de Jordan.

4.4.2. (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
>> P1=[(A-2*eye(5))*N22(:,2),N22(:,2)]
P1 = 5 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
>> P2=[(A+eye(5))*Nm12(:,3),Nm12(:,3)]
P2 = 0 0
      0 0
      -1 0
        0 0
        0 1
>> P3=Nm11(:,2);
```

```
>> P=[P1,P2,P3]
P = 5    0    0    0    0
     0    1    0    0    0
     0    0   -1    0    0
     0    0    0    0    1
     0    0    0    1    0
>> inv(P)*A*P
ans =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
```

(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)
N23 = [ 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-2*eye(6))^3*N24
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
[ 1, 0, 0, 0, 0]
[ 1/3, 0, 0, 0, 0]
>> P1=[(A-2*eye(6))^3*N24(:,1),
(A-2*eye(6))^2*N24(:,1),
(A-2*eye(6))*N24(:,1),N24(:,1)]
P1 = [ 0, 0, 0, 1]
      [ 0, 0, 1, -5/3]
      [ 0, 0, -1, 0]
      [ 0, 1, -5/3, 0]
      [ 1, -5/3, -2/3, 0]
      [ 1/3, -2/3, 0, 0]

```

Como a dimensão de $\mathcal{N}(A - 2I_6)$ é igual a 2, a dimensão de K_2 é igual a 5 e P_1 tem 4 colunas, devemos encontrar apenas mais um ciclo de comprimento igual a 1.

```

>> P2=N21(:,2);
>> Nm11=null(A+eye(6))
Nm11 = [ 0]
        [ 0]
        [ 0]
        [ 0]
        [ 0]
        [ 1]
>> P=[P1,P2,Nm11]
P = [ 0, 0, 0, 1, 0, 0]
     [ 0, 0, 1, -5/3, 0, 0]
     [ 0, 0, -1, 0, -1, 0]
     [ 0, 1, -5/3, 0, 1, 0]
     [ 1, -5/3, -2/3, 0, 0, 0]
     [ 1/3, -2/3, 0, 0, 0, 1]
>> inv(P)*A*P
[ 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]

```

- (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+eye(6))^2*Nm13
[ 2, 1, 0, 2, 0]
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
>> P1=[(A+eye(6))^2*Nm13(:,1),
(A+eye(6))*Nm13(:,1),Nm13(:,1)]
[ 2, -1, 0]
[ 0, 3, 0]
[ 0, -2, 0]
[ 0, 1, 0]
[ 0, 0, 1]
[ 0, 0, 0]

```

Como a dimensão $\mathcal{N}(A + I_6)$ é igual a 2, a dimensão de K_{-1} é igual 5 e P_1 tem 3 colunas, devemos encontrar mais um ciclo de comprimento igual a 2. Para isso, vamos descobrir um vetor da base de $\mathcal{N}(A + I_6)^2$ tal que $(A + I_6)$ aplicado a ele não seja múltiplo escalar do autovetor do ciclo já determinado.

```

>> rref([(A+eye(6))^2*Nm13(:,1),(A+eye(6))*Nm12])
[ 1, -1/2, 0, 1/2, 0]

```

```

[ 0, 0, 0, 0, 1]
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0]
>> P2=[(A+eye(6))*Nm12(:,4),Nm12(:,4)]
[ 1, 0]
[ 1, 0]
[ -2, -2]
[ 1, 0]
[ 0, 1]
[ 0, 0]
>> Nm41=null(A+4*eye(6))
[ -2]
[ 0]
[ -1]
[ 1]
[ -1]
[ 1]
>> P=[P1,P2,Nm41]
[ 2, -1, 0, 1, 0, -2]
[ 0, 3, 0, 1, 0, 0]
[ 0, -2, 0, -2, -2, -1]
[ 0, 1, 0, 1, 0, 1]
[ 0, 0, 1, 0, 1, -1]
[ 0, 0, 0, 0, 0, 1]
>> inv(P)*A*P
[ -1, 1, 0, 0, 0, 0]
[ 0, -1, 1, 0, 0, 0]
[ 0, 0, -1, 0, 0, 0]
[ 0, 0, 0, -1, 1, 0]
[ 0, 0, 0, 0, -1, 0]
[ 0, 0, 0, 0, 0, -4]

```

(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, 0, 1, 0, 0, 0]
[ 1, 0, 1, 0, 3, -4]
[ 0, 0, 1, 0, 0, 0, 0]
[ 1, 0, 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-eye(8))^2*N13
[ -1, 0, 4, 0, -1, 1, -3]
[ 0, 0, 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 0, 0]
[ 0, 0, 0, 0, 0, 0, 0]
>> P1=[(A-eye(8))^2*N13(:,1),...
(A-eye(8))*N13(:,1),N13(:,1)]
[ -1, -1, 0]
[ 0, -1, 0]
[ 0, 1, 0]
[ 0, 1, 0]
[ 0, 0, 1]
[ 0, 0, 0]
[ 0, 0, 0]
[ 0, 0, 0]

```

O próximo bloco não pode ser iniciado por nenhuma coluna de $N13$, pois $(A - I_8)^2$ multiplicado pelas colunas de $N13$ são

L.D. com a primeira coluna de P1.

```
>> (A-eye(8))*N12
[ -1,  0,  0,  1,  0,  4]
[  0,  0,  0,  0,  0, -1]
[  1,  0,  0,  0,  1, -2]
[  1,  0,  0,  0,  1, -4]
[  0,  0,  0,  0,  0, -1]
[  0,  0,  0,  0,  0,  1]
[  0,  0,  0,  0,  0,  0]
[  0,  0,  0,  0,  0,  0]
>> P2=[(A-eye(8))*N12(:,1),N12(:,1)]
[ -1,  0]
[  0,  0]
[  1,  1]
[  1,  0]
[  0,  1]
[  0,  0]
[  0,  0]
[  0,  0]
```

Precisamos saber quais colunas de N12 geram ciclos que são L.I. com os ciclos P1 e P2. Para isso basta que os últimos vetores dos ciclos sejam L.I.

```
>> rref([P1(:,1),P2(:,1),(A-eye(8))*N12])
[  1,  0,  0,  0,  0, -1, -1,  0]
[  0,  1,  1,  0,  0,  0,  1,  0]
[  0,  0,  0,  0,  0,  0,  0,  1]
[  0,  0,  0,  0,  0,  0,  0,  0]
[  0,  0,  0,  0,  0,  0,  0,  0]
[  0,  0,  0,  0,  0,  0,  0,  0]
[  0,  0,  0,  0,  0,  0,  0,  0]
[  0,  0,  0,  0,  0,  0,  0,  0]
>> P3=[(A-eye(8))*N12(:,6),N12(:,6)]
[  4,  0]
[ -1,  0]
[ -2, -4]
[ -4,  0]
[ -1,  0]
[  1,  0]
[  0,  1]
[  0,  0]
>> N31=null(A-3*eye(8))
[  1]
[  0]
[  0]
```

```
[ 2]
[ 2]
[ 1]
[ 1]
[-1]
>> P=[P1,P2,P3,N31]
[-1, -1,  0, -1,  0,  4,  0,  1]
[ 0, -1,  0,  0,  0, -1,  0,  0]
[ 0,  1,  0,  1,  1, -2, -4,  0]
[ 0,  1,  0,  1,  0, -4,  0,  2]
[ 0,  0,  1,  0,  1, -1,  0,  2]
[ 0,  0,  0,  0,  0,  1,  0,  1]
[ 0,  0,  0,  0,  0,  0,  1,  1]
[ 0,  0,  0,  0,  0,  0,  0, -1]
>> inv(P)*A*P
[ 1, 1, 0, 0, 0, 0, 0, 0]
[ 0, 1, 1, 0, 0, 0, 0, 0]
[ 0, 0, 1, 0, 0, 0, 0, 0]
[ 0, 0, 0, 1, 1, 0, 0, 0]
[ 0, 0, 0, 0, 1, 0, 0, 0]
[ 0, 0, 0, 0, 0, 1, 1, 0]
[ 0, 0, 0, 0, 0, 0, 1, 0]
[ 0, 0, 0, 0, 0, 0, 0, 3]
```

Bibliografia

- [1] Mário Barone Júnior. *Álgebra Linear*. Instituto de Matemática e Estatística da USP, São Paulo, 3a. edição, 1988.
- [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. *Introdução à Álgebra*. IMPA, Rio de Janeiro, 1979.
- [6] Adilson Gonçalves e Rita M. L. de Souza. *Introdução à Álgebra Linear*. Edgard Blücher, Rio de Janeiro, 1977.

- [7] João Pitombeira de Carvalho. *Álgebra Linear - Introdução*. Livros Técnicos e Científicos Editora S.A., Rio de Janeiro, 2a. edição, 1977.
- [8] John B. Fraleigh e Raymond A. Beauregard. *Linear Algebra*. Addison Wesley, Reading, Massachusetts, 3a. edição, 1995.
- [9] M. W. Frazier. *An Introduction to Wavelets through Linear Algebra*. Springer Verlag, New York, 1999.
- [10] Stephen H. Friedberg, Arnold J. Insel, e Lawrence E. Spence. *Linear Algebra*. Prentice Hall, Upper Saddle River, New Jersey, 3a. edição, 1997.
- [11] G. H. Golub e C. F. Van Loan. *Matrix Computations*. Johns Hopkins U.P., Baltimore, 3a. edição, 1996.
- [12] Stanley I. Grossman. *Elementary Linear Algebra*. Saunders College Publishing, New York, 5a. edição, 1994.
- [13] Morris W. Hirsch e Stephen Smale. *Differential Equations, Dynamical Systems and Linear Algebra*. Academic Press, Inc., New York, 1974.
- [14] Kenneth Hoffman e Ray Kunze. *Álgebra Linear*. Livros Técnicos e Científicos Ed. S.A., Rio de Janeiro, 3a. edição, 1979.
- [15] Donald Kreider, Donald R. Ostberg, Robert C. Kuller, e Fred W. Perkins. *Introdução à Análise Linear*. Ao Livro Técnico S.A., Rio de Janeiro, 1972.
- [16] Serge Lang. *Linear Algebra*. Springer Verlag, New York, 3a. edição, 1987.
- [17] Steven Leon, Eugene Herman, e Richard Faulkenberry. *ATLAST Computer Exercises for Linear Algebra*. Prentice Hall, Upper Saddle River, New Jersey, 1996.
- [18] Steven J. Leon. *Álgebra Linear com Aplicações*. Livros Técnicos e Científicos Editora S.A., Rio de Janeiro, 5a. edição, 1998.
- [19] Emília Giraldes, Vitor H. Fernandes, e Maria P. M Smith. *Curso de Álgebra Linear e Geometria Analítica*. Mc Graw Hill, Lisboa, 1995.
- [20] Elon L. Lima. *Álgebra Linear*. IMPA, Rio de Janeiro, 2a. edição, 1996.

- [21] Seymour Lipschutz. *Álgebra Linear*. McGraw-Hill, São Paulo, 3a. edição, 1994.
- [22] Mathworks Inc. *Student Edition of MATLAB Version 5 for Windows*. Prentice Hall, Upper Saddle River, New Jersey, 1997.
- [23] Ben Noble e James W. Daniel. *Applied Linear Algebra*. Prentice Hall, Upper Saddle River, New Jersey, 3a. edição, 1988.
- [24] Reginaldo J. Santos. *Matrizes Vetores e Geometria Analítica*. Imprensa Universitária da UFMG, Belo Horizonte, 2004.
- [25] Georgi E. Shilov. *Linear Algebra*. Dover Publications Inc., New York, 1977.

Índice Alfabético

- Adjunta, [273](#)
- Ângulo
 - entre vetores, [109](#)
- Autoespaço, [303](#)
- Autoespaço generalizado, [362](#)
- Autovalore(s), [297](#)
- Autovetor generalizado, [361](#)
- Autovetore(s), [297](#)

- Base
 - canônica, [132](#)
 - canônica de, [67](#)
 - de espaço vetorial, [67](#)
 - dual, [219](#)
 - ortogonal, [132](#)
 - ortonormal, [132](#)
- Bloco de Jordan, [358](#)

- Ciclo de autovetores generalizados, [372](#)
- Coefficientes de Fourier, [124](#)
- Complexificação de um espaço vetorial, [326](#)
- Complexificação de um operador linear, [326](#)
- Cônicas
 - (não) degeneradas, [345](#)
 - identificação de, [345](#)
- Conjunto de geradores, [32](#)
- Conjunto imagem, [203](#)
- Conjunto linearmente (in)dependente, [44](#)
- Conjunto ortogonal, [110](#)
- Conjunto ortonormal, [110](#)
- Contradomínio, [203](#)

- Decomposição espectral, [341](#)
- Decomposição polar de um operador linear, [342](#)
- Delta de Kronecker, [219](#)

- Determinante
 de um operador linear, 258
 diagonal, 355
 Diagonalização
 de matrizes, 293
 Dimensão, 73
 Dimensão (in)finita, 73
 Distância
 de um ponto a um subespaço, 156
 Domínio, 203

 eig, 321
 elipse, 356
 Equação (equações)
 diferencial linear, 23
 diferencial linear homogênea, 24
 diferencial, 22
 normais, 278
 quadrática, 345
 Espaço (espaços)
 conjunto de geradores, 32
 finitamente gerado, 32
 bidual, 239
 coluna, 226
 das frações racionais, 41
 das transformações lineares, 211
 dual, 211
 gerado por um conjunto, 31
 linha, 226
 \mathbb{R}^n , 1
 solução, 26
 vetoriais isomorfos, 234
 vetorial, 5
 vetorial com produto interno, 102
 vetorial complexo, 334
 vetorial real, 334

 Forma canônica de Jordan, 358
 Formula de Rodrigues, 143
 Fração racional, 81
 Frações parciais, 81
 Função, 203
 contínua por partes, 169
 diferenciável, 216
 seccionalmente contínua, 169
 Funções de matrizes, 384
 Funcional linear, 211

 Grau de um polinômio, 20

 hiperbx, 356
 hiperby, 356
 Hiperplano, 19

 identidades polares, 131
 Identificação de cônicas, 345
 Imagem, 203, 221
 inv, 321
 Isomorfismo, 234

 Lei do paralelogramo, 131

 Matriz (matrizes)

- (definida) positiva, 343
 - coluna, 3
 - companheira, 325
 - da transformação linear, 215, 245
 - de rotação, 340
 - diagonalizável, 295
 - jacobiana, 216
 - linha, 3
 - mudança de base, 243, 246
 - ortogonal, 287, 337
 - posto de, 228
 - semelhantes, 257, 294
 - submatriz principal de, 343
- Multiplicidade
- algébrica, 312
 - de uma raiz, 312
 - geométrica, 313
- Múltiplo escalar, 3, 6, 7
- Norma, 106
- Norma de um vetor, 106
- Núcleo, 221
- Nulidade, 222
- numeric, 322
- Operador
- auto-adjunto, 329, 335
 - definido positivo, 342
 - diagonalizável, 295
 - idempotente, 268
 - linear, 257
 - nilpotente, 323
 - normal, 329
 - ortogonal, 342
 - positivo, 342
 - projeção, 268
 - raiz quadrada de, 342
 - semi-simples, 326
 - unitário, 342
- parabx, 356
- paraby, 356
- Polinômio característico, 299
- Polinômio mônico, 324
- Polinômio mínimo, 324
- Polinômios de Legendre, 143
- Posto, 222
- de uma matriz, 228
- Problema de quadrados mínimos, 278
- Processo de ortogonalização de Gram-Schmidt, 134
- Produto
- escalar ou interno, 100
- Projeção, 268
- Projeção ortogonal, 116
- Projeção ortogonal no subespaço, 149
- Raiz quadrada de um operador, 342
- randi, 322
- Resolução da identidade, 341
- Restrição de um operador, 312
- Seção cônica, 345

- Sistema de equações diferenciais lineares, 384
- Soma de subespaços, 26, 27
- Soma direta de subespaços, 26, 268
- Splines (cúbicos), 21, 37
- Subespaço(s), 14
- invariante, 311
 - soma de, 26
 - soma direta de, 26
- Submatriz principal, 343
- subs, 321
- subst, 355
- sym, 322
-
- Teorema de Cayley-Hamilton, 315, 328
- Teorema de Schur, 338
- Teorema espectral, 341
- Traço de um operador linear, 257
- Transformação linear, 204
- adjunta, 273
 - identidade, 205
 - injetiva, 230
 - invertível, 252
 - nula, 205
 - sobrejetiva, 229
-
- Vetor (vetores), 2, 6
- de coordenadas em relação a uma base, 243
 - diferença de, 3, 7
 - iguais, 2
 - inverso aditivo, 6
 - multiplicação por escalar, 3
 - múltiplo escalar, 3, 6, 7
 - norma de, 106
 - nulo, 3, 6
 - ortogonais, 110
 - produto escalar ou interno de, 100
 - simétrico, 3, 6
 - soma de, 3
 - unitário, 110
-
- Wronskiano, 58