Universidade Federal de Goiás
Instituto de Matemática e Estatística
Programa de Mestrado Profissional em Matemática em Rede Nacional

Transformações lineares no plano e aplicações

LEONARDO BERNARDES NOGUEIRA

Goiânia
2013

sistema de bibliotecas ufg

UFG

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR ELETRONICAMENTE OS TRABALHOS DE CONCLUSÃO DE CURSO NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: **Trabalho de Conclusão de Curso de Mestrado Profissional**

2. Identificação do Trabalho

Autor (a):	Leonardo Bernardes Nogueira				
E-mail:	leonardo.b.nogueira@gmail.com				
Seu e-mail pode ser disponibilizado na página? [x]Sim [ ] Não					
Vínculo empregatício do autor					
Agência de fomento:	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior		Sigla:	CAPES	
País: Brasil	UF: DF	CNPJ:			
Título:	Transformações lineares no plano e aplicações				
Palavras-chave:	Álgebra linear, teorema espectral, secções cônicas				
Título em outra língua:	Linear transformations on the plane and applications				
Palavras-chave em outra língua:	Linear Algebra, Spectral Theorem, Conic Section				
Área de concentração:	Matemática do Ensino Básico				
Data defesa: (15/03/2013)					
Programa de Pós-Graduação:	Mestrado Profissional em Matemática em Rede Nacional				
Orientador (a):	Dr. Maurílio Marcio Melo				
E-mail:	melo@mat.ufg.br				
Co-orientador(a):*					
E-mail:					

*Necessita do CPF quando não constar no SisPG

3. Informações de acesso ao documento:

Concorda com a liberação total do documento [x] SIM [ ] NÃO[1]

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF ou DOC do trabalho de conclusão de curso.

O sistema da Biblioteca Digital de Teses e Dissertações garante aos autores, que os arquivos contendo eletronicamente as teses, dissertações ou trabalhos de conclusão de curso, antes de sua disponibilização, receberão procedimentos de segurança, criptografia (para não permitir cópia e extração de conteúdo, permitindo apenas impressão fraca) usando o padrão do Acrobat.

Assinatura do (a) autor (a)

Data: 15 / 03 / 2013

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

LEONARDO BERNARDES NOGUEIRA

Transformações lineares no plano e aplicações

Trabalho de Conclusão de Curso apresentado ao Programa de Pós–Graduação do Instituto de Matemática e Estatística da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em matemática

Área de concentração: Matemática de Ensino Básico.

Orientador: Prof. Dr. Maurílio Márcio Melo

Goiânia
2013

Dados Internacionais de Catalogação na Publicação (CIP)
GPT/BC/UFG

Nogueira, Leonardo Bernardes.
N778t Transformações lineares no plano e aplicações [manuscrito] / Leonardo Bernardes Nogueira. – 2013.
62 f. : figs.

Orientador: Prof. Dr. Maurílio Márcio Melo.
Dissertação (Mestrado) – Universidade Federal de Goiás, Instituto de Matemática e Estatística, 2013.
Bibliografia.
Inclui lista de figuras
.
1. Álgebra linear. 2. Teorema espectral. 3. Seções Cônicas.
I. Título.

CDU: 512.64

Leonardo Bernardes Nogueira

Transformações Lineares no Plano e Aplicações

Trabalho de Conclusão de Curso defendido no Programa de Mestrado Profissional em Matemática em Rede Nacional – PROFMAT/UFG, do Instituto de Matemática e Estatística da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em Matemática, área de concentração Matemática do Ensino Básico, aprovado no dia 15 de março de 2013, pela Banca Examinadora constituída pelos professores:

Prof. Dr. Maurílio Márcio Melo
Instituto de Matemática e Estatística-UFG
Presidente da Banca

Prof. Dr. Venício Veloso Borges
Membro/PUC/GO

Prof. Dr. João Carlos da Rocha Medrado
Instituto de Matemática e Estatística-UFG
Membro/UFG

Leonardo Bernardes Nogueira

Licenciado em Matemática e especialista em Educação Matemática pela UnB. Professor da Secretaria de Educação do Distrito Federal.

À todos os meus alunos que já foram e aos que estão por vir, é por eles que quero ser um professor melhor.

Agradecimentos

Primeiro às pessoas que amo, por sempre estarem presente em minha vida, não apenas nos momentos felizes, mas principalmente nos mais difíceis. À Universidade Federal de Goiás (UFG) e ao PROFMAT, que proporcionam não só a mim, mas a vários estudantes, a oportunidade de sermos pessoas melhores. Ao professor Maurilio Márcio Melo, pelo apoio e incentivo para a conclusão deste trabalho. Aos demais professores que contribuíram com a formação de profissionais melhores. Aos meus pais, Guilherme Rocha Nogueira e Joana D'Arc Arantes Bernardes Nogueira e namorada, Halinna Dornelles Wawruk, pelo incentivo, paciência, carinho e apoio que me foram dados ao decorrer de todo o curso. À CAPES pelo suporte financeiro. Aos meus colegas do Mestrado, pelo agradável convívio, amizade e ajuda. Muito obrigado a todos!

Se vi mais longe foi por estar de pé sobre ombros de gigantes [...]

Isaac Newton,
Carta para Robert Hooke (15 de fevereiro de 1676).

Resumo

Nogueira, Leonardo Bernardes. **Transformações lineares no plano e aplicações**. Goiânia, 2013. 63p. Trabalho de conclusão de curso. Instituto de Matemática e Estatística, Universidade Federal de Goiás.

Este trabalho inicia-se com um breve embasamento histórico sobre o desenvolvimento de espaços vetoriais e transformações lineares. Em seguida, apresenta conceitos fundamentais básicos, que formam uma linguagem mínima necessária para falar sobre Álgebra Linear, com enfoque maior nos operadores lineares do plano $\mathbb{R}^2$. Através de exemplos, explora-se um vasto conjunto de transformações no plano a fim de mostrar outras aplicações de matrizes no ensino médio e prepara o terreno para a apresentação do Teorema Espectral para operadores auto-adjuntos de $\mathbb{R}^2$. Este Teorema diz que para todo operador auto-adjunto $T : \mathrm{E} \rightarrow \mathrm{E}$, num espaço vetorial de dimensão finita, munido de produto interno, existe uma base ortonormal $\{u_1, \ldots, u_n\} \subset \mathrm{E}$ formada por autovetores de T. O trabalho culmina com aplicações sobre o estudo das secções cônicas, formas quadráticas e equações do segundo grau em x e y, no qual o Teorema Espectral se traduz como Teorema dos Eixos Principais, embora essa nomenclatura não seja usada nesse trabalho (para um estudo mais aprofundado neste tema ver [3], [4], [5], [7]). Retomando assim um estudo feito por Joseph Louis Lagrange em "Recherche d'Arithmétique", entre 1773 e 1775, no qual estudou a propriedade de números que são a soma de dois quadrados. Assim, foi levado a estudar os efeitos das transformações lineares com coeficientes inteiros numa forma quadrática de duas variáveis.

Palavras–chave

Álgebra Linear, Teorema Espectral, Secções Cônicas

Abstract

Nogueira, Leonardo Bernardes. **Linear transformations on the plane and applications**. Goiânia, 2013. 63p. Completion of course work. Instituto de Matemática e Estatística, Universidade Federal de Goiás.

This paper begins with a brief history about the development of vector spaces and linear transformations, then presents fundamental concepts for the study of Linear Algebra, with greater focus on linear operators in the $\mathbb{R}^2$ space. Through examples it explores a wide range of operators in $\mathbb{R}^2$ in order to show other applications of matrices in high school and prepares the ground for the presentation a version of Spectral Theorem for self-adjoint operators in $\mathbb{R}^2$, which says that for every operator self-adjoint $T : \mathrm{E} \rightarrow \mathrm{E}$ in finite dimensional vector space with inner product, exists an orthonormal basis $\{u_1, \ldots, u_n\} \subset \mathrm{E}$ formed by eigenvectors of T, and culminates with their applications on the study of conic sections, quadratic forms and equations of second degree in x and y; on the study of operators associated to quadratic forms, a version of Spectral Theorem could be called as The Main Axis Theorem albeit this nomenclature is not used in this paper. Thereby summarizing a study made by Lagrange in "Recherche d'arithmétique ", between 1773 and 1775, which he studied the property of numbers that are the sum of two squares. Thus he was led to study the effects of linear transformation with integer coefficients in a quadratic form in two variables.

Keywords

Linear Algebra, Spectral Theorem, Conic Section

Sumário

Lista de Figuras

CAPÍTULO 1

Introdução

Aspectos Históricos

A área de álgebra linear ficou adormecida na matemática, não apresentando nada de substancial até a metade do século XVIII. Um assunto relevante cujas questões levaram ao desenvolvimento da teoria de sistemas lineares que, por sua vez, levaram ao desenvolvimento da teoria de espaços vetoriais é o estudo das curvas algébricas.

Em 1770, no entanto, o matemático Euler conseguiu caracterizar as transformações ortogonais para n = 2 e 3, quando estudava quadrados de números similares aos quadrados mágicos. Devido ao seu raciocínio puramente algébrico, ele também conseguiu generalizar as soluções para qualquer valor de n, não se restringindo somente a 3. Um processo semelhante não poderia ocorrer na geometria, pois, nesse caso, era preciso imaginar um espaço com dimensões maiores que 3.

Partindo dos estudos feitos por Euler, Joseph Louis Lagrange publicou o "Recherche d?Arithmétique", entre 1773 e 1775, no qual estudou a propriedade de números que são a soma de dois quadrados. Assim, foi levado a estudar os efeitos das transformações lineares com coeficientes inteiros numa forma quadrática de duas variáveis. A partir desse estudo, ele instituiu que o discriminante da nova forma quadrática é o produto do antigo discriminante pelo quadrado de uma quantidade (determinante da transformação linear).

Johann Carl Friedrich Gauss, por sua vez, também estudou a questão com duas e três variáveis. Ele apresentou uma notação similar a da matriz que caracteriza a transformação linear. Além disso, Gauss estabeleceu a fórmula e uma notação simbólica para a composição de duas transformações lineares e também para o produto, o que marca um passo fundamental em direção ao conceito de matriz.

Hermann Günter Grassmann publicou em 1844 a primeira versão de "Lineale Ausdehnungslehre". Nesse trabalho discutiu e obteve uma boa parte dos resultados elementares da teoria atual de espaços vetoriais e de álgebra linear, além de ter conseguido algo bem próximo de uma formalização axiomática, mas devido a sua forma obscura de apresentação, seus resultados não influenciaram seus contemporâneos e a maior parte de seus resultados foi redescoberto independentemente de seu trabalho.

Em 1846, Arthur Cayley publicou o tratado "Sur Quelques Résultats de Géometrié de Position". Esse tratado surgiu como um passo decisivo na direção de generalizar os espaços de dimensão maior que três, pois nesse trabalho ele mostrou que se podem obter resultados em geometria tridimensional trabalhando-se com espaços de dimensão maior que três. Esse resultado poderia ter sido obtido por Möbius, mas ele adotou uma postura comum à sua época e descartou essa possibilidade.

Giuseppe Peano (1858-1952) publicou em 1888 sua própria leitura do "Ausdehnungslehre", com o titulo de "Calcolo Geométrico"no qual escreveu uma definição axiomática do que ele chamou de "sistema linear", que foi considerada a primeira definição axiomática de um espaço vetorial, mas a teoria de espaços vetoriais não foi desenvolvida antes de 1920.

Situação Atual

A análise de livros didáticos de Matemática constitui um parâmetro indicador do estado atual em que se encontra o ensino da Álgebra Linear. Especificamente no ensino médio, pode-se constatar pela leitura de [6] que o conteúdo de Álgebra linear é apresentado simplesmente para a resolução de sistemas de equações, onde a única aplicação de matrizes se restringe apenas na resolução de sistemas.

Este trabalho mostra outros aspectos da aplicabilidade de conceitos de Álgebra Linear, como de espaços vetoriais, subespaços, bases, transformações lineares, produto interno e formas quadráticas, tendo sempre o foco no plano $\mathbb{R}^2$, ou seja em espaços de dimensão 2, visando um aprofundamento dos conceitos e mantendo a simplicidade para a aplicabilidade no ensino médio. Em outras palavras este trabalho mostra uma outra área de atuação para Álgebra Linear, que é retratada nos livros de ensino médio apenas pelo uso de matrizes e sistemas lineares, mostrando sua aplicabilidade no estudo das secções cônicas (parábolas, hipérboles e elipses), que são abordadas ao final do ensino médio.

Apresentação dos Capítulos

Dividimos o conteúdo deste trabalho em 9 Capítulos: No Capítulo 1, que se desenvolve no momento, contêm alguns aspectos históricos que motivaram o desenvolvimento do tema e uma abordagem sobre a situação atual do ensino de Álgebra linear no ensino médio. Os Capítulos 2 ao 5 desenvolvem os conceitos fundamentais e as proposições básicas, constituídas de vários exemplos geométricos para que se tenha a assimilação dos conceitos mais abstratos fundamentados nas ideias intuitivas de geometria. O Capítulo

6, aparentemente muito abstrato, contêm o conceito de isomorfismo, que confere sentido preciso à afirmação de que dois espaços vetoriais de mesma dimensão são algebricamente indistinguíveis, o que dá uma amplitude maior ao estudo aparentemente restrito ao plano, que tem dimensão 2. Os Capítulos 7 e 8 são uma preparação para a conclusão que aparece no Capítulo 9 com o estudo das cônicas, como sendo uma área bastante próspera para aplicações de matrizes e Álgebra Linear. A importância do Capítulo 8, não reside apenas como preparação para o Capítulo 9, nele se encontra uma demonstração do teorema Espectral para operadores auto-adjunto de $\mathbb{R}^2$, ponto alto do trabalho (para um estudo mais aprofundado ver [3], [4], [5], [7]).

CAPÍTULO 2

Espaços Vetoriais

A noção de espaço vetorial é a base do estudo que faremos, é o terreno onde desenvolve toda a Álgebra Linear. Este Capítulo apresenta os axiomas de espaços vetorial, deduz suas consequências mais imediatas e exibe os exemplos mais importantes dessa noção.

Um espaço vetorial E é um conjunto, cujos elementos são chamados vetores, no qual estão definidas duas operações: a *adição*, que a cada par de vetores $u, v \in$ E faz corresponder um novo vetor $u + v \in$ E, chamado a *soma* de u e v, e a *multiplicação* por um número real, que a cada número $\alpha \in \mathbb{R}$ e cada vetor $v \in$ E faz corresponder um vetor $\alpha \cdot v$, ou αv , chamado o *produto* de α por v. Essas operações devem satisfazer, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v, w \in$ E, as condições abaixo, chamadas os *axiomas* de espaço vetorial:

comutativa: $u + v = v + u$;

associativa: $w + (u + v) = (w + u) + v$ e $(\alpha\beta)v = \alpha(\beta v)$;

vetor nulo: existe um vetor $0 \in$ E, chamado *vetor nulo*, ou *vetor zero*, tal que $v + 0 = 0 + v = v$ para todo $v \in$ E

inverso aditivo: para cada vetor $v \in$ E existe um vetor $-v \in$ E, chamado o *inverso aditivo*, ou *simétrico* de v, tal que $-v + v = v + (-v) = 0$;

distributividade: $(\alpha + \beta)v = \alpha v + \beta v$ e $\alpha(u + v) = \alpha u + \alpha v$;

multiplicação por 1: $1 \cdot v = v$.

Observação: O mesmo símbolo 0 representa o vetor nulo e o número zero.

Exemplo 2.1 Para todo número natural n, o símbolo $\mathbb{R}^n$, representa o *espaço vetorial euclidiano* **n**-*dimensional*. Os elementos de $\mathbb{R}^n$ são as listas ordenadas $u = (u_1, \dots, u_n), v = (\beta_1, \dots, \beta_n)$ de números reais.

Por definição, a igualdade vetorial $u = v$ significa as n igualdades numéricas $u_1 = v_1, \dots, u_n = v_n$.

Os números $u_1, \dots, u_n$ são chamados as *coordenadas* do vetor u. As operações do espaço vetorial $\mathbb{R}^n$ são definidas por

$$u + v \quad = \quad (u_1 + v_1, \dots, u_n + v_n) \tag{2-1}$$

$$\alpha \cdot u \quad = \quad (\alpha u_1, \ldots, \alpha u_n) \tag{2-2}$$

O *vetor nulo* é, por definição, aquele cujas coordenadas são todas iguais a zero: $0 = (0,0,\ldots,0)$.

O inverso aditivo de $u = (u_1, \ldots, u_n)$ é $-u = (-u_1, \ldots, -u_n)$. Temos então que estas definições fazem de $\mathbb{R}^n$ um espaço vetorial. Com efeito temos que mostrar que todos os axiomas de espaço vetorial valem para $\mathbb{R}^n$, vide demonstração em [4]. Para $n = 1$, tem-se $\mathbb{R}^1 = \mathbb{R} =$ reta numérica, $\mathbb{R}^2$ é o plano euclidiano e $\mathbb{R}^3$ é o espaço euclidiano tridimensional da nossa experiência cotidiana.

Para ajudar a compreensão, os vetores de $\mathbb{R}^2$ e $\mathbb{R}^3$ podem ser representados por flechas com origem no mesmo ponto O. A soma $u + v$ é a flecha que liga a origem O ao vértice que lhe é oposto no paralelogramo que tem u e v como lados (Veja Figura).

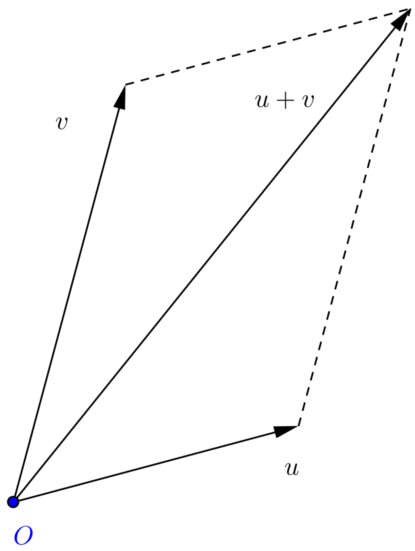


Figura 2.1: *Soma de vetores.*

Por sua vez, o produto αu é a flecha colinear a u, de comprimento α vezes o comprimento de u, com o mesmo sentido de u se $\alpha > 0$ e com sentido oposto se $\alpha < 0$.

Valem num espaço vetorial, como consequências dos axiomas, as regras operacionais habitualmente usadas nas manipulações numéricas. Vejamos algumas delas:

1. Se $w + u = w + v$ então $u = v$. Em particular, $w + u = w$ implica $u = 0$ e $w + u = 0$ implica $u = -w$. Com efeito, da igualdade $w + u = w + v$ segue-se que

   $$u = 0 + u = (-w + w) + u = -w + (w + u) = -w + (w + v) = (-w + w) + v = 0 + v = v$$

   Em particular, $w + u = w$ implica $w + u = w + 0$, logo $u = 0$. E se $w + u = 0$ então $w + u = w + (-w)$ logo $u = -w$.

2. Dados $0 \in \mathbb{R}$ e $v \in E$ tem-se $0 \cdot v = 0 \in E$. Analogamente, dados $\alpha \in \mathbb{R}$ e $0 \in E$, vale $\alpha \cdot 0 = 0$.
   Com efeito, $v + 0 \cdot v = 1 \cdot v + 0 \cdot v = (1+0) \cdot v = 1 \cdot v = v$, logo $0 \cdot v = 0$ como vimos acima. De modo análogo, como $\alpha \cdot 0 + \alpha \cdot 0 = \alpha \cdot (0+0) = \alpha \cdot 0$, segue de 1. que $\alpha \cdot 0 = 0$.
3. Se $\alpha \neq 0$ e $v \neq 0$ então $\alpha \cdot v \neq 0$.
   Com efeito, se fosse $\alpha \cdot v = 0$ então $v = 1 \cdot v = (\alpha^{-1} \cdot \alpha)v = \alpha^{-1} \cdot (\alpha v) = \alpha^{-1} \cdot 0 = 0$.
4. $(-1) \cdot v = -v$.
   Com efeito,
   $v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = [1 + (-1)] \cdot v = 0 \cdot v = 0$,
   logo $(-1)v = -v$, pela regra 1.

No que se segue, escreveremos $u - v$ para significar $u + (-v)$. Evidentemente,

$$u - v = w \Leftrightarrow u = v + w.$$

Exemplo 2.2 Sejam $u = (a,b)$ e $v = (c,d)$ vetores de $\mathbb{R}^2$ com $u \neq 0$, isto é $a \neq 0$ ou $b \neq 0$. A fim de que v seja múltiplo de u, isto é, $v = \alpha u$ para algum $\alpha \in \mathbb{R}$, é necessário e suficiente que se tenha $ad - bc = 0$. A necessidade é imediata pois $v = \alpha u$ significa $c = \alpha a$ e $d = \alpha b$. multiplicando a primeira destas igualdades por b e a segunda por a obtemos $bc = \alpha ab$ e $ad = \alpha ab$, logo $ad = bc$, ou seja, $ad - bc = 0$. Reciprocamente, se $ad = bc$ então, supondo $a \neq 0$, obtemos $d = \left(\frac{c}{a}\right) b$. Além disso, é claro que $c = \left(\frac{c}{a}\right) a$. logo, pondo $\alpha = \frac{c}{a}$, vem $d = \alpha b$ e $c = \alpha a$, isto é $v = \alpha u$. se for $b \neq 0$, tomaremos $\alpha = \frac{d}{b}$ para ter $v = \alpha u$.

CAPÍTULO 3

Subespaços

Um subespaço vetorial do espaço vetorial E *é um subconjunto* F $\subset$ E *que, relativamente às operações de* E*, é ainda um espaço vetorial. Os subespaços vetoriais constituem uma rica fonte de exemplos de espaços vetoriais.*

Seja E um espaço vetorial. Um *subespaço vetorial* (ou simplesmente um *subespaço*) de E é um subconjunto F $\subset$ E com as seguintes propriedades:

1. $0 \in \mathrm{F}$;
2. Se $u, v \in \mathrm{F}$ então $u + v \in \mathrm{F}$
3. se $v \in \mathrm{F}$ então, para todo $\alpha \in \mathbb{R}, \alpha v \in \mathrm{F}$

Para mostrar que um subespaço vetorial é um espaço vetorial, temos que mostrar que um subconjunto F de um espaço vetorial E, com as propriedades acima é uma espaço vetorial, ou seja, satisfaz os axiomas de espaço vetorial, vide demonstração em [4].

Segue-se que se u e v pertencem ao subespaço F e α, β são números reais quaisquer então $\alpha u + \beta v \in \mathrm{F}$, mais geralmente, dados $v_1, \ldots, v_m \in \mathrm{F}$ e $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$, tem-se $v = \alpha_1 v_1 + \ldots + \alpha_m v_m \in \mathrm{F}$.

O conjunto $\{0\}$, com o único elemento 0, e o espaço inteiro E são exemplos triviais de subespaços de E. Todo espaço vetorial é, em si mesmo, um subespaço.

Exemplo 3.1 Seja $v \in \mathrm{E}$ um vetor não-nulo. O conjunto $\mathrm{F} = \{\alpha v; \alpha \in \mathbb{R}\}$ de todos os múltiplos de v é um subespaço vetorial de E, chamado de reta que passa pela origem na direção de v.

Exemplo 3.2 Sejam $a_1, \ldots, a_n$ números reais. O conjunto H de todos os vetores $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$ tais que

$$a_1 x_1 + \ldots + a_n x_n = 0$$

é um subespaço vetorial de $\mathbb{R}^n$. No caso desinteressante em que $a_1 = \ldots = a_n = 0$, o subespaço H é todo o $\mathbb{R}^n$. Se, ao contrário, pelo menos um dos a_i é $\neq 0, H$ chama-se uma hiperplano de $\mathbb{R}^n$ que passa pela origem.

Exemplo 3.3 Sejam E um espaço vetorial e L um conjunto de índices. Se, para cada $\lambda \in \mathrm{L}, \mathrm{F}_\lambda$ é um subespaço vetorial de E, então a interseção

$$\mathrm{F} = \bigcap_{\lambda \in \mathrm{L}} \mathrm{F}_\lambda$$

é ainda um subespaço vetorial de E. Segue-se então do Exemplo 3.2 que o conjunto dos vetores $v = (x_1, \ldots, x_n) \in \mathbb{R}$ cujas coordenadas satisfazem as m condições abaixo

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

é um subespaço vetorial de $\mathbb{R}^n$, que é a interseção $\mathrm{F} = \mathrm{F}_1 \cap \ldots \cap \mathrm{F}_m$ dos hiperplanos F_i definidos, segundo o Exemplo 3.2, por cada uma das equações acima.

Seja X um subconjunto do espaço vetorial E. O *subespaço vetorial de* E *gerado por* X é, por definição, o conjunto do todas as combinações lineares

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m$$

de vetores $v_1, \ldots, v_m \in \mathrm{X}$.

Pode-se mostrar que o conjunto de todas as combinações lineares que se podem formar com vetores retirados dos conjunto X é, de fato, um subespaço vetorial, que indicaremos pelo símbolo $G(\mathrm{X})$.

O subespaço $G(\mathrm{X})$, gerado pelo subconjunto $\mathrm{X} \subset \mathrm{E}$, contém o conjunto X e, além disso, é o menor subespaço de E que contém X. Noutras palavras, se F é um subespaço vetorial de E e $\mathrm{X} \subset \mathrm{F}$, então $G(\mathrm{X}) \subset \mathrm{F}$. Evidentemente, se X já é um subespaço vetorial, então $G(\mathrm{X}) = \mathrm{X}$. Quando o subespaço $G(\mathrm{X})$ coincide com E, diz-se que X é um *conjunto de geradores de* E.

Explicitamente: um conjunto X é um conjunto de geradores do espaço vetorial E quando todo vetor $w \in \mathrm{E}$ pode exprimir-se como combinação linear

$$w = \alpha_1 v_1 + \ldots + \alpha_m v_m$$

de vetores $v_1, \ldots, v_m$ pertencentes a X

Exemplo 3.4 Se $v \in \mathrm{E}$ é um vetor não-nulo, o subespaço gerado por v é a reta que passa pela origem e contém v.

Exemplo 3.5 Sejam $u = (a,b)$ e $v = (c,d)$ vetores de $\mathbb{R}^2$ tais que nenhum deles é múltiplo do outro, então $u \neq 0, v \neq 0$ e, pelo Exemplo 2.2, afirmamos que $X = \{u,v\}$ é um conjunto de geradores de $\mathbb{R}^2$, ou seja, que qualquer vetor $w = (r,s) \in \mathbb{R}^2$ pode exprimir-se como uma combinação linear $w = xu + yv$. De fato esta igualdade vetorial em $\mathbb{R}^2$ equivale às duas igualdades numéricas

$$\begin{aligned} ax + cy &= r \\ bx + dy &= s. \end{aligned}$$

Como $ad - bc \neq 0$, o sistema de equações acima possui uma solução (x,y), logo existem $x,y \in \mathbb{R}$, tais que $xu + yv = w$. Esta mesma conclusão pode também ser obtida geometricamente, conforme mostra a figura 3.1. A partir da ponta de w, traçam-se paralelas às retas que contêm u e v, determinando assim os múltiplos xu, yv, que somados dão w.

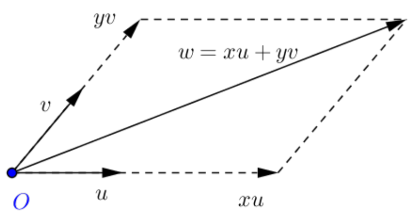


Figura 3.1: *Combinação linear de vetores em* $\mathbb{R}^2$

Exemplo 3.6 Os chamados *vetores canônicos.*

$$\begin{aligned} e_1 &= (1,0,0,\ldots,0), \\ e_2 &= (0,1,0,\ldots.0), \\ &\vdots \\ e_n &= (0,0,\ldots,0,1) \end{aligned}$$

constituem um conjunto de geradores do espaço $\mathbb{R}^n$. Com efeito, dado $v = (\alpha_1,\ldots,\alpha_n) \in \mathbb{R}^n$, tem-se $v = \alpha e_1 + \ldots + \alpha e_n$.

Resulta do Exemplo 3.5, que os únicos subespaços vetoriais de $\mathbb{R}^2$ são $\{0\}$, as retas que passam pela origem e o próprio $\mathbb{R}^2$. Com efeito, seja $F \subset \mathbb{R}^2$ um subespaço vetorial. Se F contém apenas o vetor nulo, então $F = \{0\}$. Se F contém algum vetor $u \neq 0$ então há duas possibilidades: ou todos os demais vetores de F são múltiplos de u, ou então F contém, além de u, um outro vetor v que não é múltiplo de u. Neste caso, F contém todas as combinações lineares $xu + yv$, logo $F = \mathbb{R}^2$, pelo Exemplo 3.5.

Exemplo 3.7 O sistema Linear de m equações a n incógnitas

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

possui uma solução $(x_1, \ldots, x_n)$ se, e somente se, o vetor $b = (b_1, \ldots, b_n)$ é combinação linear dos vetores-coluna

$$\begin{aligned} v_1 &= (a_{11}, a_{21}, \ldots, a_{m1}) \\ &\vdots \\ v_n &= (a_{1n}, a_{2n}, \ldots, a_{mn}) \end{aligned}$$

da matriz $A = [a_{ij}]$. Com efeito, estas equações significam que

$$b = x_1v_1 + x_2v_2 + \ldots + x_nv_n.$$

Em particular, se os vetores-colunas $v_1, \ldots, v_n$ gerarem $\mathbb{R}^n$, o sistema possui solução, seja qual for o segundo membro b.

Sejam F_1 e F_2 subespaços vetoriais de E. O subespaço vetorial de E gerado pela reunião $\mathrm{F}_1 \cup \mathrm{F}_2$ é, como se vê facilmente, o conjunto de todas as somas $v_1 + v_2$, onde $v_1 \in \mathrm{F}_1$ e $v_2 \in \mathrm{F}_2$. Ele é representado pelo símbolo $\mathrm{F}_1 + \mathrm{F}_2$.

Mais geralmente, dados os subconjuntos $X, Y \subset \mathrm{E}$, indica-se com $X + Y$ o conjunto cujos elementos são as somas $u + v$, onde $u \in X$ e $v \in V$. Quando $X = \{u\}$ reduz-se a um único elemento u, escreve-se $u + Y$ em vez de $\{u\} + Y$. Diz-se então que $u + Y$ resulta de Y pela translação de u.

Quando os subespaços $\mathrm{F}_1, \mathrm{F}_2 \subset \mathrm{E}$ têm em comum apenas o elemento $\{0\}$, escreve-se $\mathrm{F}_1 \oplus \mathrm{F}_2$ em vez de $\mathrm{F}_1 + \mathrm{F}_2$ e diz-se que $F = \mathrm{F}_1 \oplus \mathrm{F}_2$ é *soma direta de* F_1 *e* F_2.

Teorema 3.1 *Sejam* $\mathrm{F}, \mathrm{F}_1, \mathrm{F}_2$ *subespaços vetoriais de* E*, com* $\mathrm{F}_1 \subset \mathrm{F}$ *e* $\mathrm{F}_2 \subset \mathrm{F}$*. As seguintes afirmações são equivalentes:*

1. $\mathrm{F} = \mathrm{F}_1 \oplus \mathrm{F}_2$
2. *todo elemento* $w \in \mathrm{F}$ *se escreve, de modo único, como a soma* $w = v_1 + v_2$ *onde* $v_1 \in \mathrm{F}_1$ *e* $v_2 \in \mathrm{F}_2$

Prova. Provaremos que $(1) \Rightarrow (2)$. Para isso, suponhamos que $\mathrm{F}_1 \cap \mathrm{F}_2 = \{0\}$ e que se tenha $u_1 + u_2 = v_1 + v_2$, com $u_1, v_1 \in \mathrm{F}_1$ e $u_2, v_2 \in \mathrm{F}_2$. Então $u_1 - v_1 = v_2 - u_2$. Como $u_1 - v_1 \in \mathrm{F}_1$ e $v_2 - u_2 \in \mathrm{F}_2$, segue-se que $u_1 - v_1$ e $v_2 - u_2$ pertencem ambos a F_1 e a F_2. Mas $\mathrm{F}_1 \cap \mathrm{F}_2 = \{0\}$. Logo $u_1 - v_1 = v_2 - u_2 = 0$, ou seja, $u_1 = v_1$ e $u_2 = v_2$. Para provar que $(2) \Rightarrow (1)$, seja $v \in \mathrm{F}_1 \cap \mathrm{F}_2$. Então $0 + v = v + 0$ com $0, v \in \mathrm{F}_1$ e $v, 0 \in \mathrm{F}_2$. Pela hipótese

(2), isto implica $0 = v$ e portanto $F_1 \cap F_2 = \{0\}$. $\square$

Exemplo 3.8 Em $\mathbb{R}^2$, sejam F_1 o subespaço gerado pelo vetor $e_1 = (1,0)$ e F_2 o subespaço gerado pelo vetor $e_2 = (0,1)$. Então F_1 é a reta que passa pela origem na direção de e_1, ou seja, é o conjunto de vetores da forma $(\alpha_1, 0)$, enquanto que F_2 tem a forma $(0, \alpha_2)$. É claro que $\mathbb{R}^2 = F_1 \oplus F_2$.

A noção de subespaço vetorial abrange as retas, planos e seus análogos multidimensionais apenas nos casos em que esses conjuntos contém a origem. Para incluir retas, planos, etc. que não passam pela origem, tem-se a noção de variedade afim, que discutiremos agora.

Seja E um espaço vetorial. Se $x, y \in E$ e $x \neq y$, a reta que une os pontos x, y é por definição o conjunto

$$r = \{(1-t)x + ty; t \in \mathbb{R}\}$$

Pondo $v = y - x$, podemos ver que $r = \{x + tv; t \in \mathbb{R}\}$.

Um subconjunto $V \subset E$ chama-se uma *variedade afim* quando a reta que une dois pontos quaisquer de V está contida em V. Assim, $V \subset E$ é uma variedade afim se, e somente se, cumpre a seguinte condição:

$$x, y \in V, t \in \mathbb{R} \Rightarrow (1-t)x + ty \in V$$

Exemplo 3.9 Um exemplo óbvio de variedade afim é um subespaço vetorial. Ao contrário dos subespaços vetoriais, que nunca são vazios, pois devem conter o zero, a definição acima é formulada de tal modo que o conjunto vazio a cumpre, logo é uma variedade afim.

Teorema 3.2 *Seja* V *uma variedade afim não-nula no espaço vetorial* E*. Existe um único subespaço vetorial* $F \subset E$ *tal que, para todo* $x \in V$ *tem-se*

$$V = x + F = \{x + v; v \in F\}.$$

Vide demonstração em [4].

Exemplo 3.10 Vimos no exemplo 3.7 que o conjunto V das soluções de um sistema linear de m equações e n incógnitas é uma variedade afim. Supondo $V \neq \varnothing$, tomemos $x_0 \in V$ e chamemos de F o subespaço vetorial de $\mathbb{R}^n$ formado pelas soluções do sistema homogêneo

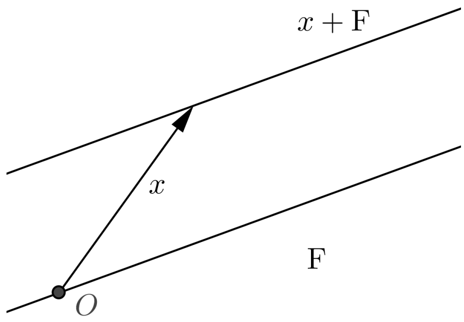


Figura 3.2: *Variedade Afim, translação por um vetor.*

correspondente (descrito no Exemplo 3.3). Tem-se $\mathrm{V} = x_0 + \mathrm{F}$. Diz-se então que "todas as soluções do sistema se obtêm somando uma solução particular com a solução geral do sistema homogêneo associado".

CAPÍTULO 4

Bases

Os espaços vetoriais de dimensão finita possuem um estrutura algébrica extremamente simples, evidenciada pelas ideias de base e dimensão. Uma vez fixada uma base num espaço vetorial de dimensão n, seus elementos são meramente combinações lineares dos n vetores básicos, com coeficientes univocamente determinados.

Seja E um espaço vetorial. Diz-se que um conjunto $X \subset E$ é *linearmente independente* (abreviadamente, L.I.) quando nenhum vetor $v \in X$ é combinação linear de outros elementos de X. Para evitar ambiguidade, no caso em que $X = \{v\}$ consta de um único elemento v, diz-se que X é L.I., por definição, quando $v \neq 0$. Quando X é L.I., diz-se também que os seus elementos são *linearmente independentes*.

Quando o conjunto X é L.I. seus elementos são todos $\neq 0$, pois o vetor nulo é combinação linear de quaisquer outros: $0 = 0 \cdot v_1 + \ldots + 0 \cdot v_m$ (se não há "outros", $X = \{v\}, v \neq 0$).

Um critério extremamente útil para verificar a independência linear de um conjunto é dado pelo teorema abaixo.

Teorema 4.1 *Seja* X *um conjunto L.I. no espaço vetorial* E*. Se* $\alpha_1 v_1 + \ldots + \alpha_m v_m = 0$ *com* $v_1, \ldots, v_m \in X$*, então* $\alpha_1 = \ldots = \alpha_m = 0$*. Reciprocamente, se a única combinação linear nula de vetores de* X *é aquela cujos coeficientes são todos iguais a zero, então* X *é um conjunto L.I..*

Pode-se obter uma demonstração do teorema acima em [4], mas para um abordagem mais ampla do assunto vide [5] e [7].

Corolário 4.1.1 *Se* $v = \alpha_1 v_1 + \ldots + \alpha_m v_m = \beta_1 v_1 + \ldots + \beta_m v_m$ *e os vetores* $v_1, \ldots, v_m$ *são L.I. então* $\alpha_1 = \beta_1, \ldots, \alpha_m = \beta_m$*.*

Com efeito, tem-se neste caso $(\alpha_1 - \beta_1)v_1 + \ldots + (\alpha_m - \beta_m)v_m = 0$ *logo* $(\alpha_1 - \beta_1) = \ldots = (\alpha_m - \beta_m) = 0$.$\square$

Exemplo 4.1 Os vetores canônicos $e_1 = (1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$ em $\mathbb{R}^n$ são L.I.. Com efeito, $\alpha_1 e_1 + \ldots + \alpha_n e_n = 0$ significa $(\alpha_1, \ldots, \alpha_n) = 0$, logo $\alpha_1 = \ldots = \alpha = 0$.

Teorema 4.2 *Sejam $v_1, \ldots, v_m$ vetores não-nulos do espaço vetorial* E. *Se nenhum deles é combinação linear dos anteriores então o conjunto* $X = \{v_1, \ldots, v_m\}$ *é L.I..*

Prova. Suponhamos, por absurdo, que uma combinação linear dos vetores dados, com coeficientes não todos nulos, fosse igual a zero. Se $\alpha_r v_r$ fosse a última parcela não-nula dessa combinação, teríamos então

$$\alpha_1 v_1 + \ldots + \alpha_r v_r = 0 = 0$$

com $\alpha_r \neq 0$. Daí viria $v_r = -\frac{\alpha_1}{\alpha_r} v_1 - \ldots - \frac{\alpha_{r-1}}{\alpha_r} v_{r-1}$, logo v_r seria combinação linear dos elementos anteriores a ele na lista $v_1, \ldots, v_m$.(Observe que $r > 1$ pois $v_1 \neq 0$) □

Observação: Vale um resultado análogo, com "subsequentes"em vez de "anteriores"no enunciado.

Um conjunto $X \subset E$ diz-se *linearmente dependente* (abreviadamente, L.D.) quando não é L.I.

Isto significa que algum dos vetores $v \in X$ é combinação linear de outros elementos de X, ou então $X = \{0\}$. A fim de que X seja L.D. é necessário e suficiente que exista uma combinação linear nula $\alpha_1 v_1 + \ldots + \alpha_m v_m = 0$ de vetores $v_1, \ldots, v_m \in X$ com algum coeficiente $\alpha_i \neq 0$. Se $X \subset Y$ e X é L.D., então Y também é L.D.. Se $0 \in X$ então o conjunto X é L.D..

Exemplo 4.2 Os vetores $u = (1,2,3), v = (4,5,6)$ e $w = (7,8,9)$ em $\mathbb{R}^3$ são L.D. pois $w = 2v - u$.

Exemplo 4.3 Quando os vetores $v_1, \ldots, v_m$ são L.D., isto não significa que qualquer um deles seja combinação linear dos demais. Por exemplo, se $u = (1,2), v = (3,4), w = (4,8)$ então $\{u, v, w\} \subset \mathbb{R}^2$ é um conjunto L.D. pois $w = 4u + 0 \cdot v$ porém v não é combinação linear de u e w.

Uma base de um espaço vetorial E é um conjunto $\mathcal{B} \subset E$ linearmente independente que gera E. Isto significa que todo vetor $v \in E$ se exprime, de modo único, como combinação linear $v = \alpha_1 v_1 + \ldots + \alpha_m v_m$ de elementos $v_1, \ldots, v_m$ da base $\mathcal{B}$ se $\mathcal{B} = \{v_1, \ldots, v_m\}$ é uma base de E e $v = \alpha_1 v_1 + \ldots + \alpha_m v_m$, então os números $\alpha_1, \ldots, \alpha_m$ chamam-se as *coordenadas* do vetor v na base $\mathcal{B}$.

Teorema 4.3 *Se os vetores $v_1, \ldots, v_m$ geram o espaço vetorial* E *então qualquer conjunto com mais de m vetores é **L.D.**.*

Corolário 4.3.1 *Se os vetores $v_1, \ldots, v_m$ geram o espaço vetorial* E *e os vetores $u_1, \ldots, u_n$ são **L.I.**, então $n \leqslant m$.*

Corolário 4.3.2 *Se o espaço vetorial* E *admite uma base $\mathcal{B} = \{u_1, \ldots, u_n\}$ com n elementos, qualquer outra base de* E *possui também n elementos.*

Diz-se que o espaço vetorial E tem *dimensão finita* quando admite uma base $\mathcal{B} = \{v_1, \ldots, v_n\}$ com um números finito n de elementos. Este número, que é o mesmo para todas as bases de E, chama-se a *dimensão* do espaço vetorial E: n= dim E. Por extensão, diz-se que o espaço vetorial $\mathrm{E} = \{0\}$ tem dimensão zero.

Corolário 4.3.3 *Se a dimensão de* E *é n, um conjunto com n vetores gera* E *se, e somente se, é L.I..*

Para uma demonstração do teorema acima e de seus corolários vide [2], [4] e [5].

Teorema 4.4 *Seja* E *um espaço vetorial de dimensão finita n. Então:*

(a) Todo conjunto X *de geradores de* E *contém uma base,*
(b) Todo conjunto L.I. $\{v_1, \ldots, v_m\} \subset \mathrm{E}$ está contido numa base,
(c) Todo subespaço vetorial $\mathrm{F} \subset \mathrm{E}$ tem dimensão finita, a qual é $\leqslant n$,
(d) Se a dimensão do subespaço $\mathrm{F} \subset \mathrm{E}$ é igual a n, então $\mathrm{F} = \mathrm{E}$.

Vide demonstração do teorema acima em [2] e [4].

Exemplo 4.4 O espaço vetorial $\mathrm{M}(m \times n)$, das matrizes $m \times n$, tem dimensão finita, igual a $m \cdot n$. uma base para $\mathrm{M}(m \times n)$ é formada pelas matrizes $\mathbf{e}_{ij}$, cujo ij-ésimo elemento (na interseção da i-ésima linha com a j-ésima coluna) é igual a 1 e os demais elementos são iguais a zero.

Diz-se que a variedade afim $\mathrm{V} \subset \mathrm{E}$ tem *dimensão* r quando $\mathrm{V} = x + \mathrm{F}$, onde o subespaço vetorial $\mathrm{F} \subset \mathrm{E}$ tem dimensão r.

CAPÍTULO 5

Transformações Lineares

Sejam E, F espaços vetoriais. Uma *transformação linear* $T : \mathrm{E} \to \mathrm{F}$ é uma correspondência que associa a cada vetor $v \in \mathrm{E}$ um vetor $T(v) \in \mathrm{F}$ de modo que valham, para quaisquer $u, v \in \mathrm{E}$ e $\alpha \in \mathbb{R}$ as relações

$$T(u+v) = T(u) + T(v) \qquad \text{e} \qquad T(\alpha \cdot v) = \alpha \cdot T(v).$$

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

Se $T : \mathrm{E} \to \mathrm{F}$ é uma transformação então $T(0) = 0$. Com efeito, $T(0) = T(0+0) = T(0) + T(0)$. Além disso, dados $u, v \in \mathrm{E}$ e $\alpha, \beta \in \mathbb{R}$, tem-se $T(\alpha u + \beta v) = T(\alpha u) + T(\beta v) = \alpha \cdot T(u) + \beta \cdot T(v)$. Mais geralmente, dados $v_1, \ldots, v_m$ em E e $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$, vale a igualdade

$$T(\alpha_1 v_1 + \ldots + \alpha_m v_m) = \alpha \cdot T(v_1) + \ldots + \alpha_m \cdot T(v_m).$$

Daí resultam $T(-v) = -T(v)$ e $T(u-v) = T(u) - T(v)$.

A *soma* de duas transformações lineares $T, G : \mathrm{E} \to \mathrm{F}$ e o *produto* de uma transformação linear $T : \mathrm{E} \to \mathrm{F}$ por um número $\alpha \in \mathbb{R}$ são as transformações lineares $T + G : \mathrm{E} \to \mathrm{F}$ e $\alpha T : \mathrm{E} \to \mathrm{F}$, definidas respectivamente por $(T+G)v = T(v) + G(v)$ e $(\alpha T)v = \alpha \cdot T(v)$, para todo $v \in \mathrm{E}$. O símbolo 0 indica a transformação linear nula $0 : \mathrm{E} \to \mathrm{F}$ definida por $0 \cdot v = 0$ e, definindo $-T : \mathrm{E} \to \mathrm{F}$ por $(-T)(v) = -A(v)$, vê-se que $(-T) + T = T + (-T) = 0$.

Dadas as transformações lineares $R : \mathrm{E} \to \mathrm{F}$, $S : \mathrm{F} \to \mathrm{G}$, onde o domínio de S coincide com a imagem de R, define-se o *produto* $SR : \mathrm{E} \to \mathrm{G}$ pondo, para cada $v \in \mathrm{E}, (SR)(v) = S(R(v))$.

Vemos que SR é uma transformação linear. Observamos também que SR nada mais é do que a composta $S \circ R$ das funções S e R. Segue-se então dos princípios gerais que se $T : \mathrm{G} \to \mathrm{H}$ é outra transformação linear, valem as propriedades

Associatividade: $(TS)R = T(SR)$

De fato, $(TS)R = (TS) \circ R$, portanto $(TS) \circ R(v) = (TS)(R(v)) = T(S(R(v))) = T(SR(v)) = T \circ SR(v)$ para todo v, logo $(TS)R = T(SR)$.

A linearidade tampouco é necessária para mostrar que, dadas $R\mathrm{E} \to \mathrm{F}$ e $S, T : \mathrm{F} \to \mathrm{G}$, tem-se a

Distributividade à esquerda: $(S+T)R = SR + TR$, que decorre simplesmente da definição de $S+T$.

Usando a linearidade de $T : \mathrm{F} \to \mathrm{G}$, vemos que, dadas, $R, S : \mathrm{E} \to \mathrm{F}$, vale a propriedade

Distributividade à direita: $T(R+S) = TR + TS$.

Com efeito, para todo $v \in \mathrm{E}$, tem-se

$$[T(R+S)](v) = C[(R+S)(v)] = C[(R(v) + S(v)] = T(R(v)) + T(S(v)) =$$
$$= (TR)(v) + (TS)(v) = (TR + TS)(v)$$

Homogeneidade: $S(\alpha R) = \alpha(SR)$, válida para $\alpha \in \mathbb{R}$, $R : \mathrm{E} \to \mathrm{F}$ e $S : \mathrm{F} \to \mathrm{G}$ quaisquer.

As transformações lineares $T : \mathrm{E} \to \mathrm{E}$ do espaço vetorial E em si mesmo são chamados *operadores lineares* em E. Um operador linear especial é o operador *identidade* $I : \mathrm{E} \to \mathrm{E}$ definido por $I(v) = v$ para todo $v \in \mathrm{E}$.

Uma transformação linear $T : \mathrm{E} \to \mathrm{F}$ é um tipo particular de função que tem o espaço vetorial E como domínio e o espaço F como contra-domínio. Em geral, para se definir uma função $f : \mathrm{X} \to \mathrm{Y}$ é necessário especificar o valor de $f(x)$ para cada elemento x no seu domínio X. O que torna as transformações lineares tão manejáveis é que para se conhecer uma transformação T, basta que se saibam os valores $T(v)$ que T assume nos vetores $v \in \mathcal{B}$, onde $\mathcal{B}$ é uma base de E. Isto é particularmente útil quando E tem dimensão finita. Neste caso, um número finito de valores $T(v_1), \ldots, T(v_n)$ (onde $\{v_1, \ldots, v_n\} \subset \mathrm{E}$ é uma base) atribuídos arbitrariamente, definem inteiramente uma transformação linear $T : \mathrm{E} \to \mathrm{F}$. Mais precisamente, vale o seguinte Teorema:

Teorema 5.1 *Sejam* E, F *espaços vetoriais e* $\mathcal{B}$ *uma base de* E. *A cada vetor* $u \in \mathcal{B}$, *façamos corresponder (de maneira arbitrária) um vetor* $u' \in \mathrm{F}$. *Então existe uma única transformação linear* $T : \mathrm{E} \to \mathrm{F}$ *tal que* $T(u) = u'$.

Prova. Todo vetor $v \in \mathrm{E}$ se exprime, de modo único, como uma combinação linear $v = \alpha_1 u_1 + \ldots + \alpha_m u_m$ de elementos $u_1, \ldots, u_m$ da base $\mathcal{B}$. Definimos $T : \mathrm{E} \to \mathrm{F}$ pondo

$$T(v) = \alpha_1 u'_1 + \ldots + \alpha_m u'_m$$

Dados $v, w \in \mathrm{E}$ temos

$$v = \alpha_1 u_1 + \ldots + \alpha_m u_m$$

e

$$w = \beta_1 u_1 + \ldots + \beta_m u_m$$

então

$$v + w = \sum_{i=1}^{m} (\alpha_i + \beta_i) u_i$$

logo

$$T(v+w) = \sum (\alpha_i + \beta_i) u_i' = \sum \alpha_i u_i' + \sum \beta_i u_i' = T(v) + T(w).$$

De maneira análoga se vê que $T(\alpha v) = \alpha \cdot T(v)$, portanto $T : \mathrm{E} \to \mathrm{F}$, assim definida, é uma transformação linear, tal que $T(u) = u'$, para todo $u \in \mathcal{B}$. Quanto à unicidade, seja $G : \mathrm{E} \to \mathrm{F}$ outra transformação linear tal que $G(u) = u'$ para todo $u \in \mathcal{B}$. Então, para cada $v = \sum \alpha_i u_i \in \mathrm{E}$ tem-se

$$G(v) = G\left(\sum \alpha_i u_i\right) = \sum \alpha_i \cdot G(u_i) = \sum \alpha_i \cdot u_i' = T(v)$$

portanto $G = T$. Isto completa a demonstração. □

Em virtude do Teorema 5.1, se quisermos definir uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ basta escolher, para cada $j = 1, \ldots, n$, um vetor $v_j = (a_{1j}, a_{2j}, \ldots, a_{mj}) \in \mathbb{R}^m$ e dizer que $v_j = A \cdot e_j$ é a imagem do j-ésimo vetor da base canônica, $e_j = (0, \ldots, 1, \ldots, 0)$, pela transformação linear T. A partir daí, fica determinada a imagem $T(v)$ de qualquer vetor $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Com efeito, tem-se $v = x_1 e_1 + \ldots + x_n e_n$, logo

$$\begin{aligned} T(v) = T\left(\sum_{j=1}^{n} x_j e_j\right) &= \sum_{j=1}^{n} x_j T(e_j) = \sum_{j=1}^{n} (a_{1j} x_j, a_{2j} x_j, \ldots, a_{mj} x_j) \\ &= \left(\sum_{j=1}^{n} a_{1j} x_j, \sum_{j=1}^{n} a_{2j} x_j, \ldots, \sum_{j=1}^{n} a_{mj} x_j\right), \end{aligned}$$

ou seja,

$$T(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_m),$$

onde

$$\begin{aligned} y_1 &= a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \\ y_2 &= a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n \\ &\vdots \\ y_m &= a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n. \end{aligned} \tag{5-1}$$

Resumindo: uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ fica inteiramente determinada por uma matriz $A = [a_{ij}] \in \mathrm{M}(m \times n)$. Os vetores-colunas dessa matriz são as imagens $T(e_j)$ dos vetores da base canônica de $\mathbb{R}^n$. A imagem de $T(v)$ de um vetor arbitrário $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$ é o vetor $w = (y_1, \ldots, y_m) \in \mathbb{R}^m$ cujas coordenadas são dadas pelas equações 5-1 acima, nas quais ocorrem os vetores-linha da matriz A. Diz-se que A é a *matriz da transformação* de T relativa às bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$.

Dessa maneira temos que o espaço vetorial $\mathrm{M}(m \times n)$, das matrizes $m \times n$ munido das operações usuais de soma, multiplicação por escalar e produto, satisfazem todas as condições de transformações lineares, ou seja, para R e S, transformações lineares cujas matrizes associadas sejam A e B respectivamente, temos que as matriz de $R+S$, αR e $R \circ S$ (onde a soma, multiplicação por escalar e composição estejam definidas) são $A+B$, αA e $A \times B$ respectivamente (para um estudo mais aprofundado desses resultados vide [1], [2] e [8]). A partir de agora faremos uso abertamente de todas as propriedades matriciais, como determinantes, transposta, ortogonalidade, semelhança, adjunta e etc.

Exemplo 5.1 Seja $v = (x, y, z)$ um vetor genérico do $\mathbb{R}^3$. Então, a fórmula

$$T(v) = (2x - y + z, x + 3y - 2z) \tag{5-2}$$

Define uma transformação $T : \mathbb{R}^3 \to \mathbb{R}^2$, é fácil ver que T assim definida é linear. Por exemplo, a imagem do vetor canônico $e_1 = (1, 0, 0)$ é o vetor $w = (2, 1)$. Observe que, se

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 3 & -2 \end{pmatrix} \tag{5-3}$$

, então

$$Av^t = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 3 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y + z \\ x + 3y - 2z \end{pmatrix}. \tag{5-4}$$

Comparando 5-2 e 5-4, vemos que a transformação $T : \mathbb{R}^3 \to \mathbb{R}^2$ é dada por

$$T(v) = (Av^t)^t. \tag{5-5}$$

Observe que no exemplo acima, utilizamos a noção de matriz transposta, indicada pelo expoente t, isso se faz necessário uma vez que a imagem de uma transformação linear é um vetor linha e o produto da matriz A pelo vetor coluna v é um vetor coluna.

Exemplo 5.2 (Rotação de ângulo θ em torno da origem em $\mathbb{R}^2$) Trata-se do operador $R : \mathbb{R}^2 \to \mathbb{R}^2$, que leva cada vetor v no vetor $R(v)$ que dele resulta pela rotação de ângulo θ em torno da origem. A figura 5.1 deixa claro que $R(u+v) = R(u) + R(v)$. É bem mais claro ainda que $R(\alpha v) = \alpha R v$ para $v \in \mathbb{R}^2$ e $\alpha \in \mathbb{R}$, logo R é uma transformação linear.

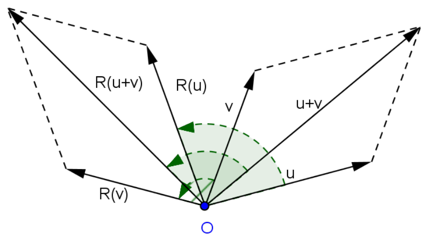


Figura 5.1: *Rotação de Vetores.*

Para um vetor $v = (x, y) \in \mathbb{R}^2$ arbitrário, seja $R(v) = (x', y')$. Sabemos que existe uma matriz A associada a transformação R, tal que $R(v) = (Av^t)^t$, temos que $A \in \mathrm{M}(2 \times 2)$. Queremos então determinar a matriz

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

onde $R(e_1) = (a, c)$ e $R(e_2) = (b, d)$, com $e_1 = (1, 0)$ e $e_2 = (0, 1)$.

Ora, pelas definições de seno e cosseno, o vetor unitário $R(e_1)$, que forma com e_1 um ângulo θ, tem coordenadas
$cos\theta$ e $\operatorname{sen}\theta$, ou seja, $R(e_1) = (\cos\theta, \operatorname{sen}\theta)$. Além disso, como e_2 forma com e_1 um ângulo reto, $R(e_2)$ também forma com $R(e_1)$ um ângulo reto. Logo $R(e_2) = (-\operatorname{sen}\theta, \cos\theta)$

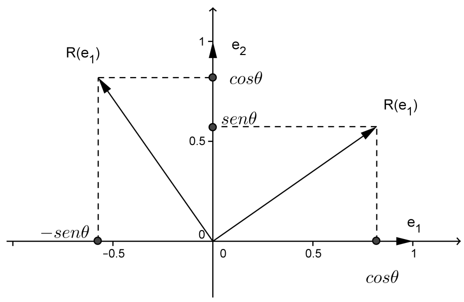


Figura 5.2: *Rotação de ângulo* θ.

Portanto, a rotação $R : \mathbb{R}^2 \to \mathbb{R}^2$ leva um vetor $v = (x, y)$ no vetor $R(v) = (Av^t)^t = (x', y')$, onde

$$A = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

portanto

$$Av^t = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\operatorname{sen}\theta \\ x\operatorname{sen}\theta + y\cos\theta \end{pmatrix}$$

obtemos assim:

$$x' = x\cos\theta - y\operatorname{sen}\theta$$
$$y' = x\operatorname{sen}\theta + y\cos\theta$$

A matriz de R relativa à base canônica de $\mathbb{R}^2$ é portanto a matriz A.

Exemplo 5.3 (Projeção ortogonal sobre uma reta) A reta $y = ax$ é o conjunto dos pontos $(x, ax) \in \mathbb{R}^2$, onde x varia em $\mathbb{R}$. Ela é o subespaço vetorial de $\mathbb{R}^2$ gerado por $(1, a)$. Consideremos o operador $P : \mathbb{R}^2 \to \mathbb{R}^2$ que faz corresponder a cada $v = (x, y) \in \mathbb{R}^2$ o vetor $P(v) = (x', ax')$, cuja extremidade é o pé da perpendicular baixada de v sobre a reta $y = ax$ como é mostrado na figura 5.3.

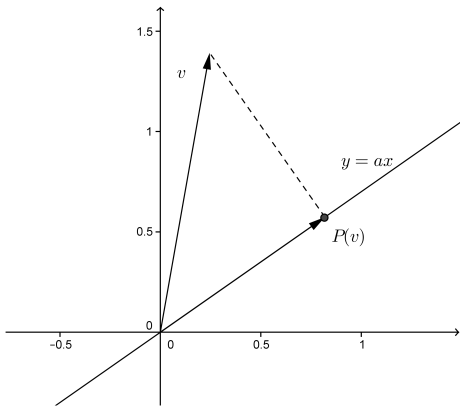


Figura 5.3: *Projeção ortogonal sobre uma reta.*

Queremos determinar x' em função de x e y, o que nos dará as coordenadas (x', ax') de $P(v)$ em função das coordenadas de v. No caso particular em que $a = 0$, a reta $y = ax$ é o eixo das abscissas e a projeção $P(v)$ é simplesmente a $(x, 0)$. As equações da projeção P sobre o eixo horizontal são portanto $x' = x$ e $y' = 0$. A matriz de P na base canônica de $\mathbb{R}^2$ é $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. No caso geral, a extremidade do vetor $P(v)$ é o vértice do ângulo reto num triângulo retângulo cujos demais vértices são a origem e a extremidade do vetor v. Pelo teorema de Pitágoras, temos

$$dist(v, 0)^2 = dist(P(v), 0)^2 + dist(v, P(v))^2,$$

ou seja,

$$x^2 + y^2 = (x')^2 + a^2(x')^2 + (x - x')^2 + (y - ax')^2.$$

Supondo $x' \neq 0$, desenvolvendo, simplificando e dividindo ambos os membros por x', obtemos $(1 + a^2)x' = x + ay$, donde

$$x' = \frac{x + ay}{1 + a^2}, \quad \text{ou seja} \quad x' = \frac{1}{1 + a^2}x + \frac{a}{1 + a^2}y.$$

O caso $x' = 0$ significa que $v(x, y)$ está sobre a perpendicular à reta $y = ax$ passando pela origem. Ora, a equação dessa perpendicular é $x + ay = 0$, logo a expressão $x' = \frac{(x+ay)}{(1+a^2)}$ fornece x' em função de x e y em quase todos os casos, falta analisarmos quando a reta se trata do eixo y, cuja equação nesse caso é $x + ay = 0$ com $a = 0$, assim a projeção $P(v)$ é simplesmente igual a $(0, y)$. As equações da projeção P sobre o eixo vertical são portanto $x' = 0, y' = y$. A matriz de P na base canônica de $\mathbb{R}^2$ é $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Vemos em particular, que a projeção $P : \mathbb{R}^2 \to \mathbb{R}^2$ é um operador linear, cujas matrizes na base canônica de $\mathbb{R}^2$ são:

$$\text{caso a reta seja } y = ax \quad \begin{pmatrix} \frac{1}{1+a^2} & \frac{a}{1+a^2} \\ \frac{a}{1+a^2} & \frac{a^2}{1+a^2} \end{pmatrix}; \quad \text{caso a reta seja o eixo } y \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Observe que a segunda matriz é o $\lim\limits_{a \to \infty}$ da primeira matriz, mas não vamos nos aprofundar nesse tema.

Exemplo 5.4 (Reflexão em torno de uma reta.) Seja $S : \mathbb{R}^2 \to \mathbb{R}^2$ a reflexão em torno da reta $y = ax$. Para todo $v = (x, y) \in \mathbb{R}^2$, a reta $y = ax$ é a bissetriz do ângulo entre v e $S(v)$ e é perpendicular à reta que liga v a $S(v)$. Seja $P : \mathbb{R}^2 \to \mathbb{R}^2$ a projeção ortogonal sobre a reta $y = ax$. A figura 5.4 mostra que, para todo $v \in \mathbb{R}^2$, tem-se $v + S(v) = 2P(v)$, ou seja, $I + S = 2P$, onde $I : \mathbb{R}^2 \to \mathbb{R}^2$ é o operador identidade. Daí vem $S = 2P - I$. Usando o exemplo anterior, concluímos que, para todo $v = (x, y)$, tem-se $S(v) = (x', y')$, onde

$$x' = \frac{1 - a^2}{1 + a^2}x + \frac{2a}{1 + a^2}y, \quad y' = \frac{2a}{1 + a^2}x - \frac{1 - a^2}{1 + a^2}y.$$

Obtemos assim a matriz A associada a transformação S é:

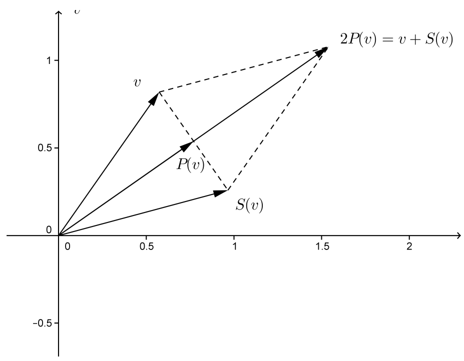


Figura 5.4: *Reflexão em torno de uma reta.*

$$A = \begin{pmatrix} \frac{1-a^2}{1+a^2} & \frac{2a}{1+a^2} \\ \frac{2a}{1+a^2} & -\frac{1-a}{1+a^2} \end{pmatrix}.$$

Observamos que a matriz A, a princípio, não contempla o caso em que a reta se trata do eixo y, isso se dá porque o parâmetro a na matriz A é o coeficiente angular da reta, ou seja, a $tan(\theta)$, onde θ é o ângulo que a reta faz com o eixo positivo x, de fato o eixo y é descrito pela equação $y = ax$ quando a tende ao infinito e portanto a matriz A retrata, de fato, a reflexão com relação a uma reta que passa pela origem.

Mostraremos agora uma outra maneira de se obter a reflexão sobre uma reta, só que nesse caso vamos enfatizar a argumentação sobre o ângulo θ que a reta faz com o eixo positivo x.

Exemplo 5.5 (Reflexão em torno de uma reta L que passa pela origem.) É uma transformação que leva cada ponto sobre sua imagem especular com relação a à reta L. A figura 5.5 ilustra a reflexão T com relação ao eixo x, para qual $T(x,y) = (x,-y)$. Obviamente,

$$T(e_1) = e_1 = (1,0) \quad \text{e} \quad T(e_2) = (0,-1)$$

de modo que a matriz de T é

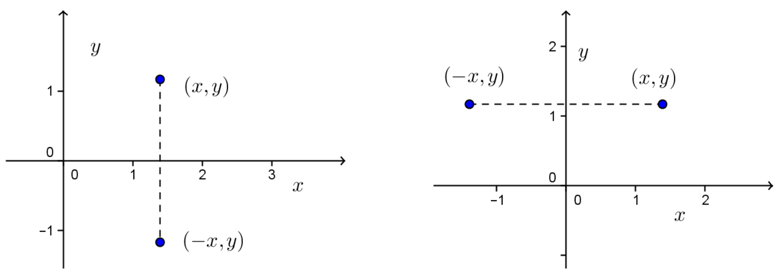
$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Como uma verificação, notamos que

$$T(x,y) = \left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right]^t = (x,-y).$$

De maneira análoga, obtemos que a matriz de reflexão com relação ao eixo y é

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



(a) *Reflexão sobre o eixo x.* (b) *Reflexão sobre o eixo y.*

Figura 5.5: *(a) e (b) representam as reflexão com relação ao eixo x e y respectivamente.*

Seja L a reta que passa através da origem do $\mathbb{R}^2$ e forma um ângulo θ com o eixo x positivo. Então, pode-se conseguir a reflexão T com relação a L através de uma rotação de um ângulo $-\theta$ para mover L sobre o eixo x, em seguida uma reflexão com relação ao eixo x e, finalmente, uma rotação de um ângulo θ para mover L de volta à sua posição original. Temos então três operadores lineares envolvidos R_θ, T_x e $R_{-\theta}$ que fazem respectivamente as movimentações citadas acima. Para cada um dos operadores citados temos uma matriz associada, logo a matriz A associada a transformação $R_\theta \circ T_x \circ R_{-\theta}$ é o produto das suas matrizes individuais, portanto

$$A = \begin{pmatrix} \cos(\theta) & -\mathrm{sen}\,(\theta) \\ \mathrm{sen}\,(\theta) & \cos(\theta) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos(-\theta) & -\mathrm{sen}\,(-\theta) \\ \mathrm{sen}\,(-\theta) & \cos(-\theta) \end{pmatrix}$$

$$A = \begin{pmatrix} \cos(\theta) & -\mathrm{sen}\,(\theta) \\ \mathrm{sen}\,(\theta) & \cos(\theta) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos(\theta) & \mathrm{sen}\,(\theta) \\ -\mathrm{sen}\,(\theta) & \cos(\theta) \end{pmatrix},$$

portanto

$$A = \begin{pmatrix} \cos^2(\theta) - \mathrm{sen}^2(\theta) & 2\mathrm{sen}\,(\theta)\cos(\theta) \\ 2\mathrm{sen}\,(\theta)\cos(\theta) & -(\cos^2(\theta) - \mathrm{sen}^2(\theta)) \end{pmatrix}.$$

Usando as igualdades trigonométricas

$$\begin{aligned} \cos(2\theta) &= \cos^2(\theta) - \mathrm{sen}^2(\theta) \\ \mathrm{sen}\,(2\theta) &= 2\mathrm{sen}\,(\theta)\cos(\theta), \end{aligned}$$

obtemos a matriz

$$A = \begin{pmatrix} \cos(2\theta) & \mathrm{sen}\,(2\theta) \\ \mathrm{sen}\,(2\theta) & -\cos(2\theta) \end{pmatrix}.$$

Para mais exemplos de operadores lineares vide [1], [2], [3], [7] e [8].

CAPÍTULO 6

Núcleo e Imagem

Nesta seção, será examinada com cuidado a possibilidade de uma transformação linear admitir ou não uma inversa. Veremos que isto está associado à existência e à unicidade da solução de um sistema de equações lineares. Será introduzido o conceito de isomorfismo, que dará um sentido preciso à afirmação de que dois espaços vetoriais de mesma dimensão são algebricamente indistinguíveis. Tudo começa com o núcleo e a imagem de uma transformação.

À toda transformação linear $T : \mathrm{E} \rightarrow \mathrm{F}$ estão associados dois subespaços vetoriais indispensáveis para estudar o comportamento de T: o núcleo de T, que é o subespaço de E, e a imagem de T, que é um subespaço de F, vide demonstração desse fato em [14].

A *imagem* de T é o subconjunto $Im(T) \subset \mathrm{F}$, formado por todos os vetores $w = T(v) \in \mathrm{E}$ que são imagens de elementos de E pela transformação T.

A noção de imagem tem sentido seja qual for a função $T : \mathrm{E} \rightarrow \mathrm{F}$, seja linear ou não. Quando T é linear, então $Im(T)$ é um subespaço vetorial de F, como se vê facilmente.

Se $Im(T) = F$, dizemos que a transformação T é sobrejetiva. Isto significa que, para qualquer $w \in \mathrm{F}$ dado, pode-se achar $v \in \mathrm{E}$ tal que $T(v) = w$.

Seja $\mathrm{X} \subset \mathrm{E}$ um conjunto de geradores do espaço vetorial E. A imagem da transformação linear $T : \mathrm{E} \rightarrow \mathrm{F}$ é o subespaço vetorial de F gerado pelos vetores $T(v), v \in \mathrm{X}$. Em particular, T é sobrejetiva se, e somente se, transforma X num conjunto de geradores de F. Se $v_1, \ldots, v_n$ geram E os vetores $T(v_1), \ldots, T(v_n)$ geram $Im(T)$. Segue-se que a dimensão de $Im(T)$ é menor do que ou igual à dimensão do domínio de T.

Uma transformação linear $R : \mathrm{F} \rightarrow \mathrm{E}$ chama-se uma *inversa à direita* da transformação $T : \mathrm{E} \rightarrow \mathrm{F}$ quando se tem $T(R(w)) = w$ para todo $w \in \mathrm{F}$.

Teorema 6.1 *A fim de que uma transformação linear* $T : \mathrm{E} \rightarrow \mathrm{F}$*, entre espaços vetoriais de dimensão finita, possua uma inversa à direita R é necessário e suficiente que T seja sobrejetiva.*

Prova. Se T admite uma inversa à direita $R : \mathrm{F} \rightarrow \mathrm{E}$ então para todo $w \in \mathrm{F}$ tem-se $T(R(w)) = w$, logo $w = T(v)$, onde $v = R(w)$, e T é sobrejetiva. Suponhamos, em seguida, que T seja sobrejetiva. A fim de definir uma transformação linear $R : \mathrm{F} \rightarrow \mathrm{E}$ com

$T(R(w)) = w$, $\forall w \in \mathrm{F}$, tomamos uma base $\mathcal{B} = \{w_1, \ldots, w_m\} \subset \mathrm{F}$. Como T é sobrejetiva, podemos escolher vetores $v_1, \ldots, v_m \in \mathrm{E}$ tais que $T(v_1) = w_1, \ldots, T(v_m) = w_m$. Pelo Teorema 5.1, existe uma transformação linear $R : \mathrm{F} \to \mathrm{E}$ tal que $R(w_1) = v_1, \ldots, R(w_m) = v_m$. Afirmamos que, para todo $w \in \mathrm{F}$, tem-se $T(R(w)) = w$, de fato, sendo $\mathcal{B}$ uma base, podemos escrever $w = \beta_1 w_1 + \ldots + \beta_m w_m$, portanto

$$\begin{aligned} T(R(w)) &= T(\beta_1 R(w_1) + \ldots + \beta_m R(w_m)) \\ &= T(\beta_1 v_1 + \ldots + \beta_m v_m) \\ &= \beta_1 T(v_1) + \ldots + \beta_m T(v_m) \\ &= \beta_1 w_1 + \ldots + \beta_m w_m = w. \end{aligned}$$

□

O *núcleo* da transformação linear $T : \mathrm{E} \to \mathrm{F}$ é o conjunto dos vetores $v \in \mathrm{E}$ tais que $T(v) = 0$. Usaremos a notação $\mathcal{N}(T)$ para representar o núcleo de T. É fácil ver que $\mathcal{N}(T)$ é um subespaço vetorial de E. Uma transformação linear $T : \mathrm{E} \to \mathrm{F}$ chama-se *injetiva* quando $v \neq v'$ em $\mathrm{E} \Rightarrow T(v) \neq T(v')$ em F. Equivalentemente: $T(v) = T(v') \Rightarrow v = v'$. Esta noção tem sentido para qualquer função $T : \mathrm{E} \to \mathrm{F}$, seja ela linear ou não. No caso linear, porém, o teorema abaixo simplifica a verificação da injetividade.

Teorema 6.2 *A fim de que uma transformação linear $T : \mathrm{E} \to \mathrm{F}$ seja injetiva é necessário e suficiente que seu núcleo $\mathcal{N}(T)$ contenha apenas o vetor nulo.*

Prova. Seja T injetiva. Então $v \in \mathcal{N}(T) = \{0\}$. Reciprocamente, seja $\mathcal{N}(T) = \{0\}$. Então $T(v) = v' \Rightarrow T(v - v') = T(v) - T(v') = 0 \Rightarrow v - v' = 0 \Rightarrow v - v' \in \mathcal{N}(T) \Rightarrow v - v' = 0 \Rightarrow v = v'$. □

Teorema 6.3 *Uma transformação linear é injetiva se, e somente se, leva vetores* **L.I.** *em vetores* **L.I.**.

Prova. Seja $T : \mathrm{E} \to \mathrm{F}$ uma transformação linear injetiva. Se os vetores $v_1, \ldots, v_n \in \mathrm{E}$ são L.I., vamos provar que $T(v_1), \ldots, T(v_n)$ são L.I. em F. Com efeito, se $\alpha_1 T(v_1) + \ldots + \alpha_n T(v_n) = 0$, então $T(\alpha_1 v_1 + \ldots + \alpha_n v_n) = 0$, logo $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ pois T é injetiva. Como $v_1, \ldots, v_n$ são L.I., segue-se que $\alpha_1 = \ldots = \alpha_n = 0$, portanto $T(v_1), \ldots, T(v_n)$ são L.I.. Reciprocamente se a transformação linear $T : \mathrm{E} \to \mathrm{F}$ leva vetores L.I. em vetores L.I., então $v \neq 0$ em $\mathrm{E} \Rightarrow \{v\}$ L.I. $\Rightarrow$ $\{T(v)\}$ L.I. $\Rightarrow T(v) \neq 0$, portanto $\mathcal{N}(T) = \{0\}$ e T é injetiva. □

Segue-se deste teorema que se E tem dimensão finita n e $T : \mathrm{E} \to \mathrm{F}$ é uma transformação linear injetiva, então $\dim\mathrm{F} \geqslant n$. Assim, por exemplo, não existe uma transformação linear injetiva de $\mathbb{R}^3$ em $\mathbb{R}^2$.

Teorema 6.4 *Seja $T : \mathrm{E} \to \mathrm{F}$ uma transformação linear. Para todo $b \in Im(T)$, o conjunto $\mathrm{V} = \{x \in \mathrm{E}; T(x) = b\}$, formado pelas soluções do sistema linear $T(x) = b$, é uma variedade afim em E, paralela ao $\mathcal{N}(T)$.*

Observe que o sistema $T(x) = b$ é o usualmente conhecido sistema matricial $AX = B$, das aulas do ensino médio, onde A é a matriz associada à transformação T, enquanto que X e B são os vetores colunas dos vetores x e b, ou seja, $X^t = x$ e $B^t = b$.

Prova. Fixemos $x_0 \in \mathrm{V}$, isto é, com $T(x_0) = b$. Afirmamos que $\mathrm{V} = x_0 + \mathcal{N}(T)$. Com efeito, $v \in \mathcal{N}(T) \Rightarrow T(x_0 + v) = T(x_0) + T(v) = b + 0 = b \Rightarrow x_0 + x \in \mathrm{V}$. Logo $x_0 + \mathcal{N}(T) \subset \mathrm{V}$. Reciprocamente,

$$\begin{gathered} x \in \mathrm{V} \Rightarrow x = x_0 + (x - x_0) = x_0 + v \\ \Downarrow \\ b = T(x) = T(x_0 + v) = t(x_0) + t(v) = b + T(v) \\ \Downarrow \\ b = b + T(v) \Rightarrow T(v) = 0 \Rightarrow x = x_0 + v \in x_0 + \mathcal{N}(T). \end{gathered}$$

Logo $\mathrm{V} \subset x_0 + \mathcal{N}(T)$. □

Teorema 6.5 *Sejam E e F espaços vetoriais de dimensão finita e T transformação linear, então $T : \mathrm{E} \to \mathrm{F}$ possui inversa à esquerda se, e somente se, é injetiva.*

Prova. Seja $R : \mathrm{F} \to \mathrm{E}$ inversa à esquerda de T. Então $T(u) = T(v) \Rightarrow u = R(T(u)) = R(T(v)) = v$, logo T é injetiva. Reciprocamente, suponha que T seja injetiva. A fim de obter uma inversa à esquerda R para T, tomemos $\{v_1, \ldots, v_n\} \subset \mathrm{E}$, uma base. Pelo Teorema 6.3, os vetores $T(v_1), \ldots, T(v_n) \in \mathrm{F}$ são L.I., logo podemos achar vetores $w_1, \ldots, w_k \in \mathrm{F}$ tais que

$$\{T(v_1, \ldots, T(v_n), w_1, \ldots, w_k)\} \subset \mathrm{F}$$

seja uma base. Pelo Teorema 5.1, a fim de se definir a transformação linear $R : \mathrm{F} \to \mathrm{E}$, basta especificar seus valores nos elementos desta base. Poremos $R(T(v_1)) = v_1, \ldots, R(T(v_n)) = v_n, b(w_1) = 0, \ldots, b(w_k) = 0$. Dado qualquer $v \in \mathrm{E}$, tem-se $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$, logo

$$R(T(v)) \quad = \quad R(\alpha_1 T(v_1) + \ldots + \alpha_n T(v_n))$$

$$\begin{aligned} &= \alpha_1 R(T(v_1) + \ldots + \alpha_n R(T(v_n) \\ &= \alpha_1 v_1 + \ldots + \alpha_n v_n, \end{aligned}$$

portanto R é uma inversa à esquerda de T. $\square$

Uma transformação $T : \mathrm{E} \to \mathrm{F}$ chama-se invertível quando existe $R : \mathrm{F} \to \mathrm{R}$, transformação linear, tal que $RT = I_{\mathrm{E}}$ e $TR = I_{\mathrm{F}}$, ou seja, quando R é, ao mesmo tempo, inversa à esquerda e à direita de T, em outras palavras quando a matriz A associada a T for quadrada, ou seja, $A \in \mathrm{M}(n \times n)$, e o determinante de A for diferente de zero.

Neste caso, dizemos que R é a *inversa* de T e escrevemos $R = T^{-1}$, o que significa que a matriz B associada a R é a inversa da matriz A associada a T.

A fim de que a transformação linear T, seja invertível, é necessário e suficiente que ela seja injetiva e sobrejetiva. Diz-se, então que T é uma *bijeção linear* entre E e F ou, mais apropriadamente que $T : \mathrm{E} \to \mathrm{F}$ é um isomorfismo e que os espaços vetoriais E e F são isomorfos.

Se $T : \mathrm{E} \to \mathrm{F}$ e $R : \mathrm{F} \to \mathrm{G}$ são isomorfismos, então $T^{-1} : \mathrm{F} \to \mathrm{E}$ e $RT : \mathrm{E} \to \mathrm{G}$ também são isomorfismos. Tem-se $(RT)^{-1} = T^{-1}R^{-1}$ e para $\alpha \neq 0$, $(\alpha T)^{-1} = \frac{1}{\alpha} \cdot T^{-1}$. Observe que as mesmas relações valem para as suas matrizes associadas.

Um isomorfismo $T : \mathrm{E} \to \mathrm{F}$ entre espaços vetoriais transforma toda base de E numa base de F. Reciprocamente, se uma transformação linear $T : \mathrm{E} \to \mathrm{F}$ leva alguma base de E numa base F então T é um isomorfismo.

Do que foi dito acima resulta, em particular, que dois espaços vetoriais de dimensão finita isomorfos têm a mesma dimensão. A recíproca é verdadeira, como veremos agora.

Com efeito, seja E um espaço vetorial de dimensão finita n. Fixando uma base $\{v_1, \ldots, v_n\} \subset \mathrm{E}$, podemos definir uma transformação linear $T : \mathbb{R}^n \to \mathrm{E}$ pondo, para cada $v = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, $T(v) = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Temos que $T(e_1) = v_1, \ldots, T(e_n) = v_n$. Assim, T transforma a base canônica $\{e_1, \ldots, e_n\}$ na base $\{v_1, \ldots, v_n\} \subset \mathrm{E}$, logo é um isomorfismo entre $\mathbb{R}^n$ e E.

Noutras palavras, todo espaço de dimensão n é isomorfo a $\mathbb{R}^n$.

Como vimos $T^{-1} : \mathrm{E} \to \mathbb{R}^n$ e o produto $RT^{-1} : \mathrm{E} \to \mathrm{F}$ de T^{-1} por outro isomorfismo $R : \mathbb{R}^n \to \mathrm{F}$ são isomorfismos, segue-se que dois espaços vetoriais E e F, ambos de dimensão n, são isomorfos.

Isso justifica que a restrição desse trabalho ao espaço $\mathbb{R}^2$ é apenas aparente, uma vez que todos os espaços de dimensão 2 são isomorfos a $\mathbb{R}^2$.

Teorema 6.6 (Teorema do Núcleo e da Imagem) *Sejam* E, F *espaços vetoriais de dimensão finita. Para toda transformação linear* $T : \mathrm{E} \to \mathrm{F}$ *tem-se que* $\dim \mathrm{E} = \dim \mathcal{N}(T) + \dim Im(T)$.

Prova. O teorema resulta imediatamente da seguinte afirmação mais precisa, que provaremos a seguir: se $\{T(u_1),\ldots,T(u_p)\}$ é uma base de $Im(T)$ e $\{v_1,\ldots,v_q\}$ é uma base de $\mathcal{N}(T)$ então $\{u_1,\ldots,u_p,v_1,\ldots,v_q\}$ é uma base de E.

Com efeito, em primeiro lugar, se tivermos

$$\alpha_1 u_1+\ldots+\alpha_p+\beta v_1+\ldots+\beta v_q=0, \tag{6-1}$$

então, aplicando a transformação T a ambos os membros desta igualdade e lembrando que $v_1,\ldots,v_q$ pertencem ao núcleo de T, obtemos a igualdade

$$\alpha_1 T(u_1)+\ldots+\alpha_p T(u_p)=0.$$

Como os vetores $T(u_1),\ldots,T(u_p)$ são L.I., resulta daí que $\alpha_1=\ldots=\alpha_p=0$. Portanto a igualdade 6-1 se reduz a igualdade

$$\beta_1 v_1+\ldots+\beta_q v_q=0.$$

Como $v_1,\ldots,v_q$ são L.I., concluímos que $\beta_1=\ldots=\beta_q=0$. Isto mostra que os vetores $u_1,\ldots,u_p,v_1,\ldots,v_q$ são L.I..

Em seguida, consideramos um vetor arbitrário $w\in$ E. Como $T(w)\in Im(T)$, podemos escrever

$$T(w)=\alpha_1 T(u_1)+\ldots+\alpha_p T(u_p),$$

pois $\{T(u_1),\ldots,T(u_p)\}$ é uma base da imagem de T. A igualdade acima pode ser reescrita como

$$T[w-(\alpha_1 u_1+\ldots+\alpha_p u_p)]=0.$$

Assim, o vetor $w-(\alpha_1 u_1+\ldots+\alpha_p u_p)$ pertence ao núcleo de T, logo pode ser expresso como combinação linear dos elementos da base $\{v_1,\ldots,v_q\}$. Temos então

$$w-(\alpha_1 u_1+\ldots+\alpha_p u_p)=\beta_1 v_1+\ldots+\beta_q v_q,$$

ou seja, $w=\alpha_1 u_1+\ldots+\alpha_p u_p+\beta_1 v_1+\ldots+\beta_q v_q$. isto mostra que os vetores $\{u_1,\ldots,u_p,v_1,\ldots,v_q\}$ geram E e portanto constituem uma base. □

Corolário 6.6.1 *Sejam* E, F *espaços vetoriais de mesma dimensão finita* n. *Uma transformação linear T : E $\to$ F é injetiva se, e somente se, é sobrejetiva e portanto é um isomorfismo.*

Com efeito, temos $n=\dim\mathcal{N}(T)+\dim Im(T)$. logo $\mathcal{N}(T)=\{0\}$ se, e somente se $\dim Im(T)=n$, ou seja, $Im(T)=$ F.

Teorema 6.7 *Se uma transformação linear* $T : \mathrm{E} \to \mathrm{F}$ *tem uma inversa à esquerda* $R : \mathrm{F} \to \mathrm{E}$ *e uma inversa à direita* $S : \mathrm{F} \to \mathrm{E}$*, então* $R = S$ *e* T *é um isomorfismo, com* $T^{-1} = R = S$*.*

Prova. Tem-se $RT = I_{\mathrm{E}}$ e $TS = I_{\mathrm{F}}$. Portanto $R = RI_{\mathrm{F}} = R(TS) = (RT)S = I_{\mathrm{E}}S = S$. □

Corolário 6.7.1 *Seja* $\dim \mathrm{E} = \dim \mathrm{F}$*. Se as transformações lineares* $T : \mathrm{E} \to \mathrm{F}$*,* $R : \mathrm{F} \to \mathrm{E}$ *são tais que* $RT = I_{\mathrm{E}}$*, então* $TR = I_{\mathrm{F}}$ *e* $R = T^{-1}$*.*

Com efeito, $RT = I_{\mathrm{E}} \Rightarrow T$ injetiva $\Rightarrow T$ sobrejetiva $\Rightarrow TS = I_{\mathrm{F}}$ para algum $S \Rightarrow S = R \Rightarrow TR = I_{\mathrm{F}}$.

Exemplo 6.1 (Projeção sobre o eixo das abscissas) *Considere o operador linear* $T : \mathbb{R}^2 \to \mathbb{R}^2$*, tal que* $T(x,y) = (0,y)$*, nestas condições* $\mathcal{N}(T)$ *é o subespaço gerado pelo vetor* e_1 *e* $Im(T)$ *é o subespaço gerado pelo vetor* e_1*, então como no exemplo 3.8 temos que* $\mathbb{R}^2 = \mathcal{N}(T) \oplus Im(T)$

Exemplo 6.2 *Para todo operador linear* $T : \mathrm{E} \to \mathrm{E}$ *num espaço vetorial de dimensão finita vale a relação* $\dim \mathrm{E} = \dim \mathcal{N}(T) + \dim Im(T)$*. Isto porém não implica que se tenha sempre* $\mathrm{E} = \mathcal{N}(T) \oplus Im(T)$*. Por exemplo, se* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *é definido por* $T(x,y) = (x-y, x-y)$ *então, tomando* $w = (1,1)$*, temos* $w = T(v)$*, com* $v = (2,1)$ *e* $T(w) = 0$*, logo* $\mathcal{N}(T) \cap Im(T)$ *contém o vetor não nulo* w

CAPÍTULO 7

Produto Interno

Os axiomas de espaço vetorial não são suficientes para abordar certas noções geométricas como ângulo, perpendicularismo, comprimento, distância, etc. Isto se torna possível com a introdução de um produto interno.

Um *produto interno* num espaço vetorial E é um funcional bilinear simétrico e positivo em E. Mais precisamente, um produto interno é uma função $E \times E \to \mathbb{R}$, que associa a cada par de vetores $u, v \in E$ um número real $\langle u, v \rangle$, chamado de produto interno de u por v, de modo que sejam válidas as seguintes propriedades, para quaisquer $u, u', v, v' \in E$ e $\alpha \in \mathbb{R}$:

Bilinearidade: $\langle u + u', v \rangle = \langle u, v \rangle + \langle u', v \rangle$, $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$, $\langle u, v + v' \rangle = \langle u, v \rangle + \langle u, v' \rangle$, $\langle u, \alpha v \rangle = \alpha \langle u, v \rangle$;

Comutatividade (simetria): $\langle u, v \rangle = \langle v, u \rangle$;

Positividade: $\langle u, u \rangle > 0$ se $u \neq 0$.

Como $\langle 0, v \rangle = \langle 0 + 0, v \rangle = \langle 0, v \rangle + \langle 0, v \rangle$, segue-se que $\langle 0, v \rangle = \langle v, 0 \rangle = 0$ para todo $v \in E$.

Resulta da positividade que se $\langle u, v \rangle = 0$, para todo $v \in E$, então $u = 0$. Com efeito, se $u \neq 0$ teríamos $\langle u, v \rangle \neq 0$ pelo menos quando $v = u$.

Segue-se desta observação que se $u, u' \in E$ são vetores tais que $\langle u, v \rangle = \langle u', v \rangle$ para todo $v \in E$, então $u = u'$. Com efeito, isto implica que $\langle u - u' \rangle = 0$ para todo $v \in E$, logo $u - u' = 0$ e $u = u'$.

O número não negativo $|u| = \sqrt{\langle u, u \rangle}$ chama-se a *norma* ou o *comprimento* do vetor u. Com esta notação, tem-se $|u|^2 = \langle u, u \rangle$ e a igualdade

$$\langle u + v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle,$$

lê-se: $|u + v|^2 = |u|^2 + |v|^2 + 2\langle u, v \rangle$.

Quando $|u| = 1$ diz-se que $u \in E$ é um *vetor unitário*. Todo vetor $u \neq 0$ se escreve como $u = |u| \cdot u'$, onde u' é um vetor unitário. Basta pôr $u' = |u|^{-1} \cdot u$.

Exemplo 7.1 No espaço euclidiano $\mathbb{R}^n$, *produto interno* canônico dos vetores $u = (\alpha_1, \ldots, \alpha_n)$ e $v = (\beta_1, \ldots, \beta_n)$ é definido por $\langle u, v\rangle = \alpha_1\beta_1 + \ldots + \alpha_n\beta_n$. Este é o produto interno que consideraremos em $\mathbb{R}^n$.

Exemplo 7.2 Consideremos $\mathbb{R}^2$ como modelo aritmético do plano euclidiano, no qual se introduziu um sistema de coordenadas cartesianas. Dados $u = (\alpha_1, \alpha_2)$ e $v = (\beta_1, \beta_2)$, os números

$$|u| = \sqrt{\alpha_1^2 + \alpha_2^2}$$

e

$$|v| = \sqrt{\beta_1^2 + \beta_2^2}$$

medem realmente os comprimentos das flechas que representam esses vetores. Suponhamos $u \neq 0, v \neq 0$ e chamemos de θ o ângulo formado por essas flechas. Afirmamos que o produto interno $\langle u, v\rangle = \alpha_1\beta_1 + \alpha_2\beta_2$ acima definido é igual a $|u||v|\cos\theta$. Isto será provado em três passos: $1^{\underline{o}}$) Se os vetores u e v são perpendiculares, então $\langle u, v\rangle = 0 = |u||v|\cos 90^o$. Com efeito, por um lado,

$$|u+v|^2 = \langle u+v, u+v\rangle = |u|^2 + |v|^2 + 2\langle u, v\rangle$$

e por outro lado, pelo Teorema de Pitágoras,

$$|u+v|^2 = |u|^2 + |v|^2,$$

logo $\langle u, v\rangle = 0$. $2^{\underline{o}}$) Se $|u| = |v| = 1$, então $\langle u, v\rangle = \cos\theta$. Com efeito, tomando o vetor

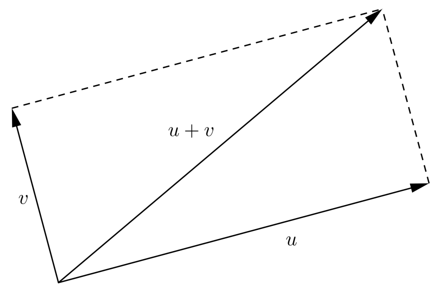


Figura 7.1: *Teorema de Pitágoras em sua forma vetorial.*

unitário u^* perpendicular a u temos, pela definição de seno e cosseno, $v = \cos\theta \cdot u + \operatorname{sen}\theta \cdot u^*$. Tomando o produto interno de ambos os membros desta igualdade por u, vem $\langle u, v\rangle = \cos\theta \cdot \langle u, u\rangle + \operatorname{sen}\,\theta \cdot \langle u, u^*\rangle$. Como $\langle u, u\rangle = 1$ e $\langle u, u^*\rangle = 0$, pelo primeiro passo, temos $\langle u, v\rangle = \cos\theta$. $3^{\underline{o}}$) Caso geral: pomos $u = |u| \cdot u'$ e $v = |v| \cdot v'$, onde u' e v' são vetores unitários na mesma direção e sentido de u e v respectivamente. Então $\langle u, v\rangle = |u||v|\langle u', v'\rangle = |u||v|\cos\theta$. Vemos, em particular, que o vetores u,v formam um

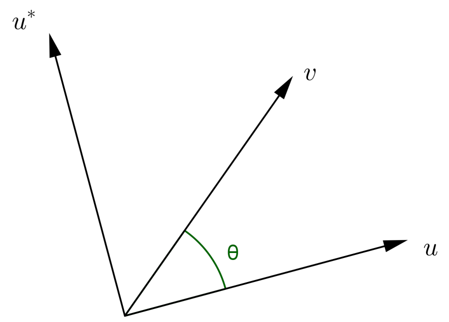


Figura 7.2: *Decomposição do vetor v em uma base ortogonal.*

ângulo agudo quando $\langle u, v\rangle > 0$, um ângulo obtuso quando $\langle u, v\rangle < 0$ e um ângulo reto quando $\langle u, v\rangle = 0$.

Seja E um espaço com produto interno. Dois vetores $u, v \in$ E chamam-se *ortogonais* (ou *perpendiculares*) quando $\langle u, v\rangle = 0$. Escreve-se, então $u \perp v$. Em particular, 0 é ortogonal a qualquer vetor de E. Um conjunto X $\subset$ E diz-se ortogonal quando todos os vetores de X são ortogonais dois a dois. Se, além disso, todos os vetores de X são unitários, então X chama-se um conjunto *ortonormal*. Portanto, o conjunto X $\subset$ E é ortonormal se, e somente se, dados $u, v \in$ X tem-se $\langle u, v\rangle = 0$ se $u \neq v$ e $\langle u, v\rangle = 1$ se $v = u$. Uma *base ortonormal* é uma base de E que é um conjunto ortonormal.

Teorema 7.1 *Num espaço vetorial* E *com produto interno, todo conjunto ortonormal* X *de vetores não nulo é* **L.I.**.

Prova. Sejam $v_1, \ldots, v_n \in$ X. Temos $\langle v_i, v_j\rangle = 0$ se $i \neq j$. Se $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ é uma combinação linear nula desses vetores então, para cada $i = 1, 2, \ldots, n$, tomamos os produto interno de ambos os membros desta igualdade por v_i e temos

$$\alpha_1 \langle v_1, v_i\rangle + \ldots + \alpha_n \langle v_n, v_i\rangle = 0.$$

Logo $\alpha_i \langle v_i, v_i\rangle = \alpha_i |v_i|^2 = 0$, pois todos os produtos internos $\langle v_j, v_i\rangle$, com $j \neq i$, são nulos em virtude da ortogonalidade de X. Além disso, como os vetores pertencentes ao conjunto X são todos não nulos, resulta de $\alpha_i |v_i|^2 = 0$ que $\alpha_i = 0$. Assim, os coeficientes da combinação linear $\sum \alpha_i v_i = 0$ são todos iguais a zero e os vetores do conjunto X são, portanto, linearmente independentes. $\square$

Exemplo 7.3 A base canônica $\{e_1, \ldots, e_n\} \subset \mathbb{R}^n$ é ortonormal: tem-se $\langle e_i, e_j\rangle = 0$ se $i \neq j$ e $\langle e_i, e_j\rangle = 1$ se $i = j$. No plano $\mathbb{R}^2$ os vetores $u = (1, 1)$ e $v = (-1, 1)$ são ortonormais. Pondo

$$u' = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) \quad \text{e} \quad v' = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$$

o conjunto $\{u', v'\} \subset \mathbb{R}^2$ é uma base ortonormal.

Num espaço vetorial E com produto interno, seja u um vetor unitário. Dado qualquer $v \in$ E, o vetor $\langle u, v\rangle \cdot u$ chama-se a *projeção ortogonal de v sobre o eixo que contém u*. A justificativa para esta denominação está no fato de que, escrevendo $w = v - \langle u, v\rangle u$, tem-se $v = \langle u, v\rangle u + w$, onde w é *perpendicular a u*. Com efeito, tomando o produto interno de u por ambos os membros da igualdade $w = v - \langle u, v\rangle u$ tem-se

$$\langle u, w\rangle = \langle u, v\rangle - \langle u, v\rangle\langle u, u\rangle = \langle u, v\rangle - \langle u, v\rangle = 0,$$

pois $\langle u, u\rangle = 1$.

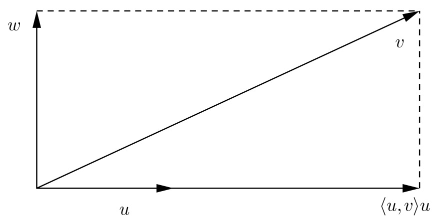


Figura 7.3: *Projeção ortogonal.*

Quando se tem apenas $u \neq 0$, o eixo que contém u é o mesmo que contém o vetor unitário $u' = \frac{u}{|u|}(= |u|^{-1} \cdot u)$. A projeção ortogonal de v sobre este eixo é, portanto, igual a $\langle u', v\rangle u'$, ou seja, $\frac{\langle u,v\rangle}{\langle u,u\rangle} \cdot u$. Usaremos a notação

$$pr_u(v) = \frac{\langle u, v\rangle}{\langle u, u\rangle} \cdot u$$

para indicar a projeção ortogonal do vetor v sobre o eixo que contém o vetor não-nulo u.

Se $z = pr_u(v)$, tem-se $v = z + w$, com $w \perp z$. Pelo Teorema de Pitágoras, $|v|^2 = |z|^2 + |w|^2$. Em particular vemos que $|z| \leqslant |v|$, isto é, o comprimento da projeção $pr_u(v)$ é menor do que ou igual ao comprimento v.

Ora, a norma de vetor $pr_u(v)$ é igual a $\frac{|\langle u,v\rangle|}{|u|}$. Segue-se então que, para quaisquer $u, v \in$ E tem-se $\frac{|\langle u,v\rangle|}{|u|} \leqslant |v|$, ou seja

$$|\langle u, v\rangle| \leqslant |u| \cdot |v| \qquad (\textit{desigualdade de Schwarz}).$$

A rigor, o argumento acima prova a desigualdade de Schwarz apenas no caso em $u \neq 0$, mas ela é óbvia no caso em que u, v é múltiplo um do outro. Isto resulta do raciocínio acima pois, no Teorema de Pitágoras $|v|^2 = |z|^2 + |w|^2$, dizer $|v| = |z|$ significa que $w = 0$, isto é, que v é múltiplo do u.

CAPÍTULO 8

Subespaços Invariantes, Operadores Auto-Adjuntos, Um Caso Particular do Teorema Espectral

A Adjunta de uma transformação Linear $T : \mathrm{E} \to \mathrm{F}$ é uma transformação linear $T^* : \mathrm{F} \to \mathrm{E}$ tal que, para $v \in \mathrm{E}$ e $w \in \mathrm{F}$ quaisquer se tenha:

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle.$$

A transposta de uma matriz $A = [a_{ij}] \in \mathrm{M}(m \times n)$ é a matriz $A^t = [a_{ji}] \in \mathrm{M}(n \times m)$ que tem como linhas as colunas de A e como colunas as linhas de A, na mesma ordem.

Teorema 8.1 *Sejam $\mathcal{U} = \{u_1, \ldots, u_n\} \subset \mathrm{E}$ e $\mathcal{V} = \{v_1, \ldots, v_m\} \subset \mathrm{F}$ bases ortonormais. Se $A = [aij] \in \mathrm{M}(m \times n)$ é a matriz da transformação linear $T : \mathrm{E} \to \mathrm{F}$ nas bases $\mathcal{U}, \mathcal{V}$ então a matriz da adjunta $T^* : \mathrm{F} \to \mathrm{E}$ nas bases $\mathcal{V}, \mathcal{U}$ é a transposta $A^t = [a_{ji}] \in \mathrm{M}(n \times m)$ de A.*

Prova. Por definição de um transformação linear, temos

$$T(u_j) = \sum_{i=1}^{m} a_{ij} v_i \quad (j = 1, \ldots, n)$$

e

$$T^*(v_i) = \sum_{r=1}^{n} b_{ri} u_r,$$

onde $B = [b_{ri}] \in \mathrm{M}(n \times m)$ é a matriz de T^* nas bases $\mathcal{V}, \mathcal{U}$, a ser determinada. Como ambas as bases são ortonormais, temos, para cada $i = 1, \ldots, m$ e cada $j = 1, \ldots, n$:

$$b_{ji} = \langle u_j, T^*(v_i) \rangle = \langle T(u_j), v_i \rangle = a_{ij},$$

portanto, $B = A^t$, transposta de A. $\square$

É apresentada a seguir uma lista de propriedades operacionais da adjunta de um transformação linear, as quais se traduzem em propriedades de transposta de matriz. A validez dessas propriedades decorre da observação de que duas transformações lineares $T,R : \mathrm{E} \to \mathrm{F}$ são iguais quando se tem $\langle T(u), v\rangle = \langle R(u), v\rangle$ para quaisquer $u \in \mathrm{E}$ e $v \in \mathrm{F}$

$$\begin{aligned} I^* &= I & \quad & (I_n^t = I_n) \\ (T+R)^* &= T^* + R^* & & (A+B)^t = A^t + B^t \\ (\alpha T)^* &= \alpha t^* & & (\alpha A)^t = \alpha A^t \\ (RT)^* &= T^* R^* & & (BA)^t = A^t B^t \\ T^{**} &= T & & (A^t)^t = A \end{aligned}$$

onde A e B são as matrizes associadas às transformações lineares T e R respectivamente.

Diz-se que um subespaço vetorial $\mathrm{F} \subset \mathrm{E}$ é *invariante* pelo operador linear $T : \mathrm{E} \to \mathrm{E}$ quando $T(\mathrm{F}) \subset \mathrm{F}$, isto é, quando a imagem de $T(v)$ de qualquer vetor $v \in \mathrm{F}$ é ainda um vetor de F.

Exemplo 8.1 Os subespaços $\{0\}$ e E são invariantes por qualquer operador linear T : $\mathrm{E} \to \mathrm{E}$. O núcleo $\mathcal{N}(T)$ e a imagem $Im(T)$ são também exemplos óbvios de subespaços invariantes. Um subespaço F de dimensão 1 (reta passando pela origem) é invariante por T se, e somente se, existe um número λ tal que $T(v) = \lambda v$ para todo $v \in \mathrm{F}$. [Com efeito, fixando um vetor $u \neq 0$ em F, todos os demais elementos de F são da forma $\alpha u,\ \alpha \in \mathbb{R}$. Como $T(u) \in \mathrm{F}$, tem-se $T(u) = \lambda u$. Para qualquer outro $v \in \mathrm{F}$, vale $v = \alpha u$ logo $T(v) = \alpha T(u) = \alpha\lambda u = \lambda\alpha u$, logo $T(v) = \lambda v$, como o mesmo λ.]

Um vetor $v \neq 0$ em E chama-se um *autovetor* do operador linear $T : \mathrm{E} \to \mathrm{E}$ quando existe $\lambda \in \mathbb{R}$ tal que

$$T(v) = \lambda v.$$

O número $\lambda \in \mathbb{R}$, por sua vez, chama-se *autovalor* do operador linear T quando existe um vetor não-nulo $v \in \mathrm{E}$ tal que $T(v) = \lambda v$. Diz-se então que o autovalor λ corresponde ao autovetor v e , vice-versa, que o autovetor v também corresponde ao autovalor λ. Então, para todo $w = \alpha v$, tem-se $T(w) = \lambda w$.

Achar um autovetor (ou, o que é equivalente, um autovalor) do operador T é, portanto, o mesmo que achar um subespaço de dimensão 1 invariante por T.

Analogamente, diz-se que o número real λ é um *autovalor* da *matriz* $A \in \mathrm{M}(n \times n)$ quando λ é um autovalor do operador $T : \mathbb{R}^n \to \mathbb{R}^n$, cuja matriz na base canônica é A. Isto significa que existe um vetor $x \neq 0$ em $\mathbb{R}^n$ tal que $T(x) = \lambda x$ ou, o que é o mesmo, uma matriz não-nula $\mathbf{x} \in \mathrm{M}(n \times 1)$ tal que $A\mathbf{x} = \lambda\mathbf{x}$.

Exemplo 8.2 Uma rotação $R : \mathbb{R}^2 \to \mathbb{R}^2$ em torno da origem, de ângulo diferente de 0°ou 180°, não admite outros subespaços invariantes além de $\{0\}$ e $\mathbb{R}^2$.

Exemplo 8.3 O operador $T : \mathbb{R}^2 \to \mathbb{R}^2$, definido por $T(x,y) = (x+\alpha y, y)$, chama-se *cisalhamento*. Se $\alpha \neq 0$, os únicos subespaços invariantes por T são $\{0\}, \mathbb{R}^2$ e o eixo das abscissas. Com efeito, qualquer outro subespaço de $\mathbb{R}^2$ é uma reta F, formada pelos múltiplos $kv = (ka, kb)$ de um vetor $v = (a,b)$, com $b \neq 0$. Se $k \neq 0$ tem-se $kv \in$ F mas $T(kv) = (ka + \alpha kb, kb) = kv + (\alpha kb, 0) \notin$ F logo F não é invariante por T.

Teorema 8.2 *A autovalores diferentes do mesmo operador correspondem autovetores linearmente independentes.*

Prova. Dado o operador linear $T : \mathrm{E} \to \mathrm{E}$, sejam $v_1, \ldots, v_m$ vetores não-nulos em E tais que $T(v_1) = \lambda_1 v_1, \ldots, T(v_n) = \lambda_m v_m$, onde os números reais $\lambda_1, \ldots, \lambda_m$ são dois a dois diferentes. Provaremos, por indução que esses são L.I.. A afirmação é óbvia quando $m = 1$. Suponha-a verdadeira para $m-1$ vetores, concluiremos daí sua validez para m. Dada a combinação linear nula

$$\alpha_1 v_1 + \ldots + \alpha_m v_m = 0, \tag{8-1}$$

aplicamos o operador T a ambos os membros desta igualdade, levando em conta que $T(v_i) = \lambda_i v_i$. Resulta então que

$$\lambda_1 \alpha_1 v_1 + \ldots + \lambda_m \alpha_m v_m = 0. \tag{8-2}$$

Multiplicamos a igualdade 8-1 por λ_m e subtraindo de 8-2 vem:

$$(\lambda_1 \lambda_2)\alpha_1 v_1 + \ldots + (\lambda_{m-1} - \lambda_m)\alpha_{m-1} v_{m-1} = 0.$$

Pela hipótese de indução, os vetores $v_1, \ldots, v_{m-1}$ são L.I.. Logo

$$(\lambda_1 \lambda_2)\alpha_1 = \ldots = (\lambda_{m-1} - \lambda_m)\alpha_{m-1} = 0.$$

Como os autovetores são todos diferentes, as $m-1$ diferenças nos parênteses acima são $\neq 0$, logo $\alpha_1 = \ldots = \alpha_{m-1} = 0$. Isto reduz a igualdade 8-1 a $\alpha_m v_m = 0$. Como $v_m \neq 0$, segue-se que $\alpha_m = 0$. Assim, a igualdade 8-1 só pode ocorrer quando todos os coeficientes α_i são nulos, o que prova o teorema. □

Corolário 8.2.1 *Sejam dim* $\mathrm{E} = n$. *Se um operador linear* $T : \mathrm{E} \to \mathrm{E}$ *possui n autovalores diferentes então existe uma base* $\{v_1, \ldots, v_n\} \subset \mathrm{E}$ *em relação à qual a matriz de T é diagonal (isto é, tem a forma* $[a_{ij}]$ *com* $a_{ij} = 0$ *se* $i \neq j$*).*

Com efeito, se $T(v_1) = \lambda_1 v_1, \ldots, T(v_n) = \lambda_n v_n$ com os v_i não nulos e os λ_i dois a dois distintos então $\{v_1, \ldots, v_n\}$ é, em virtude do Teorema 8.2, uma base de E. A matriz de T nesta base é

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

na qual os termos que não aparecem são iguais a zero.

A igualdade $T(v) = \lambda v$ equivale a $(T - \lambda I)v = 0$, logo v é um autovetor do operador $T : \mathrm{E} \to \mathrm{E}$ se, e somente se, o operador $T - \lambda I : \mathrm{E} \to \mathrm{E}$ não possui inversa, ou seja, se o determinante da sua matriz associada for igual a zero. Para a validação dos resultado de matrizes e um estudo mais aprofundado vide [3] e [7].

Exemplo 8.4 Um caso particular importante ocorre quando $\dim \mathrm{E} = 2$. Vimos no Exemplo 2.2 que se $\{u, v\} \subset \mathrm{E}$ é uma base então os vetores $\alpha u + \beta v$ e $\gamma u + \delta v$ são L.D. se, e somente se, $\alpha\delta - \beta\gamma = 0$. Dados o operador $T : \mathrm{E} \to \mathrm{E}$ e a base $\{u, v\} \subset \mathrm{E}$, sejam $T(u) = au + cv$ e $T(v) = bu + dv$. Noutras palavras, a matriz do operador T na base $\{u, v\}$ é

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

Então $(T - \lambda I)(u) = (a - \lambda)u + cv$ e $(T - \lambda I)(v) = bu + (d - \lambda)v$. A fim de que $T - \lambda I$ não possua inversa é necessário e suficiente que $\det(A - \lambda I) = 0$, ou seja, $(a - \lambda)(d - \lambda) - bc = 0$, ou ainda, que λ seja raiz do polinômio

$$p(\lambda) = \lambda^2 - (a + d)\lambda + ad - bc,$$

chamado *polinômio característico* do operador T.

Portanto, o número real λ é um autovalor do operador $T : \mathrm{E} \to \mathrm{E}$ onde $\dim \mathrm{E} = 2$, se, e somente se, é uma raiz do polinômio característico do operador T, o qual, por definição é $p(\lambda) = \lambda^2 - (a + d)\lambda + ad - bc$. Observe que essa definição se justifica pela teoria das matrizes onde raiz desse polinômio representa o valor para o qual o operador não é invertível. Os coeficientes de $p(\lambda)$ são tirados da matriz A em relação a uma base qualquer de E.

Observação. A matriz A do operador T muda quando se passa de uma base para outra. Mas o polinômio $p(\lambda)$ (isto é, as expressões $a + d$ e $ad - bc$, que são seus coeficientes) permanece sem alteração. Vide [5] e [7].

Exemplo 8.5 No caso da rotação $R : \mathbb{R}^2 \to \mathbb{R}^2$, $R(x, y) = (x\ cos\theta - y\ sen\theta,\ x\ sen\theta + y\ cos\theta)$, temos $a = cos\theta,\ b = -sen\theta,\ c = sen\theta,\ d = cos\theta$, logo o polinômio característico de R é

$$p(\lambda) = \lambda^2 - (2cos\theta)\lambda + 1.$$

Se $\theta \neq 0°$ e $\theta \neq 180°$, o trinômio $p(\lambda)$ não possui raiz real pois nesse caso seu discriminante $\Delta = 4(cos^2\theta - 1)$ é negativo. Consequentemente R só possui autovalor (reais) se $\theta = 0°$ ou $\theta = 180°$.

Exemplo 8.6 Definamos o operador $T : \mathbb{R}^2 \to \mathbb{R}^2$ pondo $T(x,y) = (4x+3y, x+2y)$. Seu polinômio característico é $p(\lambda) = \lambda^2 - 6\lambda + 5$, cujas raízes são $\lambda_1 = 1$ e $\lambda_2 = 5$. Estes números são autovalores de T. Existem, portanto, vetores não-nulos v_1 e v_2 em $\mathbb{R}^2$, tais que $T(v_1) = v_1$ e $T(v_2) = 5v_2$. Pelo Teorema 8.2, v_1 e v_2 formam uma base $\mathbb{R}^2$, em relação à qual a matriz do operador T tem a forma diagonal:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}.$$

A fim de determinar $v_1 = (x,y)$ e $v_2 = (r,s)$ exprimimos as igualdades $T(v_1) = v_1$ e $t(v_2) = 5v_2$ em termos de coordenadas, obtemos os sistemas lineares

$$\begin{cases} 4x+3y &= x, \\ x+2y &= y \end{cases} \qquad e \qquad \begin{cases} 4r+3s &= 5r \\ r+2s &= 5s \end{cases}.$$

Ambos os sistemas acima são indeterminados, e tinham que ser assim pois se v é autovetor de T, todo múltiplo αv também é. Tomando uma solução não-nula de cada um desses sistemas obtemos $v_1 = (1,-1), v_2 = (3,1)$ tais que $\{v_1, v_2\} \subset \mathbb{R}^2$ é uma base formada por autovetores de T.

Um operador linear $T : \mathrm{E} \to \mathrm{E}$, num espaço vetorial munido de produto interno, chama-se *auto-adjunto* quando $T = T^*$, ou seja, quando $\langle T(u), v\rangle = \langle u, T(v)\rangle$ para quaisquer $u, v \in \mathrm{E}$.

Teorema 8.3 *$T : \mathrm{E} \to \mathrm{E}$ é auto-adjunto se, e somente se, sua matriz $A = [a_{ij}]$ relativamente a uma (e portanto a qualquer) base ortonormal $\mathcal{U} = \{u_1, \ldots, u_n\} \subset \mathrm{E}$ é uma matriz simétrica.*

Prova. Sabemos que $\langle u_i, T((u_j))\rangle$=[i-ésima coordenada do vetor $T(u_j)$ na base $\mathcal{U}$]=[i-ésimo elemento da j-ésima coluna de A]=a_{ij}. Portanto a matriz A é simétrica se, e somente se, $\langle u_i, T(u_j)\rangle = \langle T(u_i), u_j\rangle$ para quaisquer $i, j = 1, \ldots, n$. Devido à linearidade de T e à bilinearidade do produto interno, isto equivale a dizer que $\langle u, T(v)\rangle = \langle T(u), v\rangle$ para quaisquer $u, v \in \mathrm{E}$, ou seja, que T é auto-adjunta. $\square$

Teorema 8.4 *Se $\lambda_1, \ldots, \lambda_m$ são autovalores dois a dois diferentes do operador auto-adjunto $T : \mathrm{E} \to \mathrm{E}$, os autovetores correspondentes $v_1, \ldots, v_m$ são dois a dois ortogonais.*

Prova. Para $i \neq j$ quaisquer:

$$\begin{aligned}(\lambda_i - \lambda_j)\langle v_i - v_j\rangle &= \langle \lambda_i v_i, v_j\rangle - \langle v_i, \lambda_j v_j\rangle = \langle T(v_i), v_j\rangle - \langle v_i, T(v_j)\rangle \\ &= \langle T(v_i), v_j\rangle - \langle T(v_i), v_j\rangle = 0 \text{ pois } T \text{ é auto-adjunto.}\end{aligned}$$

□

Observação: Se $T(v) = \lambda v$ então, para todo múltiplo $w = \alpha v$, tem-se ainda $T(w) = \lambda w$. Logo, na situação do Teorema 8.4, os vetores $v_1, \ldots, v_m$ podem ser tomados unitários, caso haja conveniência.

Um problema importante sobre operadores num espaço vetorial de dimensão finita é o de encontrar uma base em relação à qual a matriz desse operador seja a mais simples possível. Mostraremos que, se $T : \mathrm{E} \to \mathrm{E}$ e um operador auto-adjunto (associado a uma matriz simétrica, pois $T = T^* \Rightarrow A = A^t$) num espaço vetorial de dimensão finita com produto interno, existe uma base ortonormal em E, relativamente à qual a matriz de T é uma *matriz diagonal* $A = [a_{ij}]$, isto é, $a_{ij} = 0$ se $i \neq j$.

Quando se diz que a matriz do operador $T : \mathrm{E} \to \mathrm{E}$ na base $\{u_1, \ldots, u_n\} \subset \mathrm{E}$ é uma matriz diagonal, isto significa que, para todo $j = i, \ldots, n$, tem-se $T(u_j) = \lambda_j u_j$, ou seja, os vetores da base dada são todos eles autovetores de T.

Vamos mostrar um caso particular do Teorema Espectral para operadores auto-adjuntos em que o espaço tem dimensão 2.

Teorema 8.5 (Teorema Espectral para Operadores Auto-Adjuntos de $\mathbb{R}^2$) *Seja $T : \mathrm{E} \to \mathrm{E}$ um operador auto-adjunto num espaço vetorial de dimensão 2, munido de produto interno. Existe uma base ortonormal $\{u_1, u_2\} \subset \mathrm{E}$ formada por autovetores de T.*

Prova. Seja $\{v, w\} \subset \mathrm{E}$ uma base ortonormal arbitrária. Em virtude do Teorema 8.3, temos $T(v) = av + bw, T(w) = bv + cw$. Como vimos no Exemplo 8.4, os autovalores de T são as raízes reais do polinômio característico $p(\lambda) = \lambda^2 - (a+c)\lambda + ac - b^2$. O discriminante deste trinômio é $\Delta = (a+c)^2 - 4(ac - b^2) = (a-c)^2 + 4b^2 \geqslant 0$. Se $\Delta = 0$, então $b = 0, a = c$ e $T = aI$, logo todo vetor não-nulo em E é um autovetor. Se $\Delta > 0$, então o trinômio $p(\lambda)$ possui 2 raízes reais distintas λ_1, λ_2. Isto, como sabemos, quer dizer que os operadores $T - \lambda_1 I$ e $T - \lambda_2 I$ são ambos não invertíveis, logo existem vetores não-nulos (que podemos supor unitário) $u_1, u_2 \in \mathrm{E}$ tais que $T(u_1) = \lambda_1 u_1$ e $T(u_2) = \lambda_2 u_2$. Pelo Teorema 8.4 $\{u_1, u_2\} \subset \mathrm{E}$ é uma base ortonormal de autovetores de T. □

Para uma abordagem mais completa sobre o Teorema Espectral vide [3], [4], [5] e [7].

CAPÍTULO 9

Seções Cônicas e Formas Quadráticas

Aplicaremos agora todo nosso estudo de Álgebra Linear ao problema geométrico de se determinar o gráfico, no plano xy, de um equação da forma

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0 \tag{9-1}$$

onde os coeficientes $a, b, \ldots, f$ são constantes reais, com a, b e c não todos nulos. Tal equação é chamada de **equação de segundo grau em x e y.**

A razão pela qual escrevemos $2b$ em lugar de b para o coeficiente do termo xy é que a natureza do gráfico é determinada, em grande parte, pela **forma quadrática associada**

$$ax^2 + 2bxy + cy^2 = \langle T(x,y),(x,y)\rangle \tag{9-2}$$

em x e y, que corresponde ao operador auto-adjunto $T : \mathbb{R}^2 \to \mathbb{R}^2$ associado à matriz simétrica 2×2

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}. \tag{9-3}$$

Em geral, a forma quadrática q nas variáveis $x_1, x_2, \ldots, x_n$ que corresponde ao operador auto-adjunto $T : \mathbb{R}^n \to \mathbb{R}^n$ é a função $q : \mathbb{R}^n \to \mathbb{R}$ definida por

$$q(x) = \langle x, T(x)\rangle.$$

A chave para a análise da equação de segundo grau em 9-1 é o fato de que a matriz simétrica A associada a transformação T em 9-3 é ortogonalmente diagonalizável, ou seja, de acordo com o Teorema 8.5, existe uma base $\mathcal{B}$ ortonormal tal que A nessa base $\mathcal{B}$ é uma matriz diagonal. Usando esse fato, demonstraremos que, exceto por alguns casos degenerados, o gráfico de toda equação de segundo grau em x e y é uma seção cônica. A expressão **seção cônica** vem do fato destas serem as curvas em que um plano intercepta um cone. O cone utilizado é um cone circular reto com duas *folhas* que se estendem ao infinito em ambos os sentidos, como mostra a figura 9.1. Há três tipo de seções cônicas, conforme ilustrado na figura 9.1. Se o plano secante é paralelo a alguma geratiz do cone, então a curva de interseção é uma **parábola**. De outro modo, ou ela é uma única curva fechada - elipse -, ou é uma hipérbole com dois ramos.

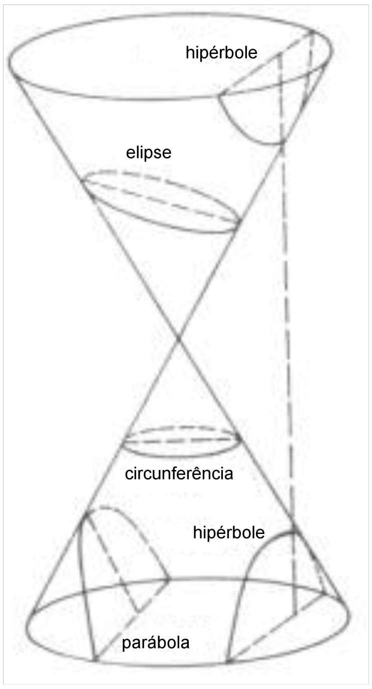


Figura 9.1: *Cone e suas secções.*

Pode-se verificar que, para um sistema apropriado de coordenadas xy nos planos secantes da figura 9.1, as equações dos três tipos de seções cônicas tomam as seguintes formas:

$$\begin{aligned} \text{Parábola:} \quad & y^2 = kx \quad \text{ou} \quad x^2 = ky \\ \text{Elipse:} \quad & \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \\ \text{Hipérbole:} \quad & \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad \text{ou} \quad \frac{y^2}{a^2} - \frac{x^2}{b^2} = 1 \end{aligned} \tag{9-4}$$

Seções cônicas com essas equações estão ilustradas nas figuras 9.1. Diz-se que uma seção cônica está em **posição canônica** em relação aos eixos coordenados se sua equação tomar uma das formas listadas nas Equações 9-4, para uma demonstração dessas equações vide [1], [3], [5], [7] e [8].

Exemplo 9.1 A equação $4x^2 + 9y^2 - 36 = 0$ pode ser escrita na forma

$$\frac{x^2}{9} + \frac{y^2}{4} = 1;$$

e em consequência seu gráfico é uma elipse com $a = 3$ e $b = 2$. Observe a singular de a e b: elipse tem interseções $(a,0)$ e $(-a,0)$ com eixo x e interseções $(0,b)$ e $(0,-b)$ com o eixo y. Também é claro que o gráfico é simétrico em relação ao eixo y, bem como em relação ao eixo x (substitua x por $-x$ ou y por $-y$).

A equação $9x^2 - 4y^2 + 36 = 0$ pode ser escrita na forma

$$\frac{y^2}{9} - \frac{x^2}{4} = 1;$$

assim, seu gráfico é uma hipérbole cujos ramos interceptam o eixo y nos pontos $(0,-3)$ e $(0,3)$. Seu gráfico também é simétrico em relação a cada eixo coordenado.

Finalmente, a equação $y^2 + 4x = 0$ pode ser escrita na forma

$$y^2 = -4x,$$

e, consequentemente, seu gráfico é uma parábola que se abre ao longo do eixo x negativo. A parábola tem a interseção única $(0,0)$ e é simétrica em relação ao eixo x (a substituição de y por $-y$ não altera a equação), apesar de não o ser em relação ao eixo y.

Além de parábola, elipse ou hipérbole, o gráfico de uma equação de segundo grau em x e y pode ser uma reta, um par de retas, um único ponto ou o conjunto vazio. Estes casos especiais, ilustrados no exemplo a seguir, são referidos como **cônicas degeneradas**.

Exemplo 9.2 O gráfico da equação $x^2 = 0$ é o eixo y. O gráfico da equação $y^2 - 1 = 0$ é o par de retas paralelas $y = 1$ e $y = -1$. O gráfico da equação $x^2 - y^2 = 0$ é o par de retas concorrentes $y = x$ e $y = -x$. O gráfico da equação $x^2 + y^2 = 0$ consiste apenas no ponto $(0,0)$. E o gráfico da equação $x^2 + y^2 + 1 = 0$ é o conjunto vazio.

Observe que nenhuma das equações em 9-4 contém simultaneamente termo em x^2 e termo em x, nem há alguma que contenha simultaneamente termo em y^2 e termo em y. A presença de qualquer um desses pares em uma equação de segundo grau indica que o gráfico é uma seção cônica que sofreu translação a partir de sua posição canônica. Se não há presente um termo em xy, então se podem remover tais pares de termos da equação a fim de se identificar seu gráfico - através de um processo de *completamento de quadrados e de translação de coordenadas*.

Exemplo 9.3 Para se identificar o gráfico da equação de segundo grau

$$3x^2 - 2y^2 - 18x + 8y + 13 = 0,$$

reunimos os termos em x e os termos em y e completamos o quadrado em cada variável:

$$(3x^2 - 18x) - (2y^2 - 8y) + 13 = 0$$

$$\begin{aligned} 3(x^2-6x)-2(y^2-4y) &= -13 \\ 3(x^2-6x+9)-2(y^2-4y+4) &= -13+27-8 \end{aligned}$$

$$3(x-3)^2-2(y-2)^2=6.$$

Agora, fazemos a substituição

$$x' = x-3, \qquad y' = y-2,$$

que corresponde à escolha de um novo sistema de coordenadas transladado $x'y'$ cuja origem é o antigo ponto $x=3, y=2$. O resultado é a equação

$$3(x')2-2(y')^2=6,$$

ou seja,

$$\frac{(x')^2}{2}-\frac{(y')^2}{3}=1.$$

Essa última equação tem a forma da primeira equação em 9-4, com $a=\sqrt{2}$ e $b=\sqrt{3}$. Portanto, o gráfico da equação de segundo grau em questão é uma hipérbole transladada e está em posição canônica no sistema $x'y'$, conforme mostra a figura 9.2. Observamos que

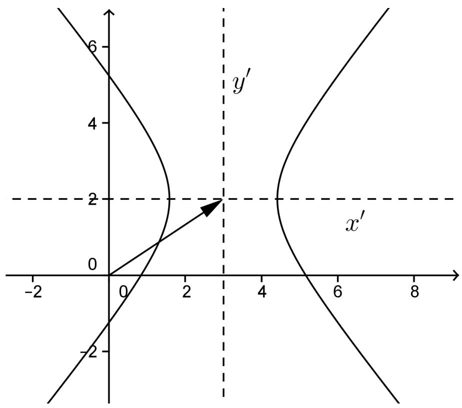


Figura 9.2: *Hipérbole Transladada.*

os eixos x' e y' são variedades afins, cujos subespaços vetoriais associados são os eixos x e y, e que a matriz A do operador auto-adjunto associado à forma quadrática é

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

e seus autovalores são $\lambda_1 = a = 3$ e $\lambda_2 = c = -2$.

Consideremos agora a presença de um termo xy em uma equação de segundo grau, que é chamada de **termo de produto cruzado das variáveis**.

A forma quadrática $q : \mathbb{R}^2 \to \mathbb{R}^2$ definida como na Equação 9-2 por $q(v) = \langle T(v), v \rangle$ (nesse caso temos $v = (x,y) \in \mathbb{R}^2$), só contém o termo de produto cruzado caso $b \neq 0$, nessas condições, pelo Teorema 8.5, temos que existem $\{u_1, u_2\} \subset \mathbb{R}^2$, uma base ortonormal de autovetores de T correspondente aos autovalores λ_1, λ_2. Nessas condições a matriz A do operador T na nova base $\{u_1, u_2\}$ é

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

e portanto

$$q(v') = \langle T(v'), v' \rangle = \lambda_1 (x')^2 + \lambda_2 (y')^2,$$

onde (x', y') são as coordenadas do vetor $v \in \mathbb{R}^2$ de acordo com a nova base $\{u_1, u_2\}$.

Reordenando os autovalores λ_1 e λ_2 ou simplesmente trocando o sinal ou de u_1 ou de u_2 caso seja mais conveniente, vide figura 9.3, podemos supor que u_2 é uma rotação 90° no sentido horário do vetor u_1, onde $u_1 = (\cos\theta, \operatorname{sen}\theta)$ e portanto $u_2 = (-\operatorname{sen}\theta, \cos\theta)$. Podemos então concluir que as coordenadas de $v = (x,y)$ na base canônica de $\mathbb{R}^2$, se relaciona com coordenadas de $v' = (x', y')$ na base ortonormal $\{u_1, u_2\}$ de $\mathbb{R}^2$ de acordo com a seguinte equação:

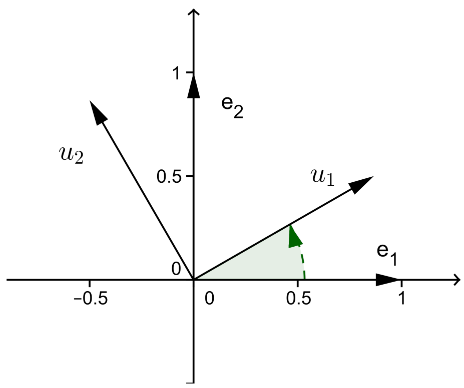


Figura 9.3: *Base ortonormal a partir de uma rotação anti-horária de um ângulo θ da base canônica.*

$$xe_1 + ye_2 = x'u_1 + y'u_2.$$

Aplicando o produto interno em ambos os lados pelo vetor e_1 e em seguida pela vetor e_2 obtemos

$$\begin{cases} x &= x'\langle u_1, e_1 \rangle + y' \langle u_2, e_1 \rangle \\ y &= x' \langle u_1, e_2 \rangle + y' \langle u_2, e_2 \rangle \end{cases} \iff \begin{cases} x &= x' \cos\theta - y' \operatorname{sen}\theta \\ y &= x' \operatorname{sen}\theta + y' \cos\theta \end{cases}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \qquad \text{ou} \qquad \begin{pmatrix} x \\ y \end{pmatrix} = B \begin{pmatrix} x' \\ y' \end{pmatrix}$$

onde $B = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$ é a matriz de passagem das coordenadas (x', y') para as coordenadas (x, y), para uma maior explanação sobre mudança de base vide [1], [3] e [8].

Definamos então um operador linear $R : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $R(x', y') = (x' \cos\theta - y' \operatorname{sen}\theta, x' \operatorname{sen}\theta + y' \cos\theta) = (x, y)$ que tem B como sua matriz associada. R é uma rotação, como já vimos no Capítulo 5, entretanto nesse caso, R retrata uma rotação dos eixos coordenados deixando o ponto em questão fixado.

Consideremos agora o caso de uma equação completa de segundo grau como

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0.$$

Considerando a base ortonormal $\{u_1, u_2\}$ de $\mathbb{R}^2$ definida acima, podemos reescrever a equação acima em função dos operadores lineares T e R do seguinte modo

$$\langle T(R(x', y')), R((x', y'))\rangle + \langle (d, e), R(x', y')\rangle + f = 0,$$

usando a adjunta R^* de R obtemos

$$\langle R^*(T(R(x', y'))), (x', y')\rangle + \langle R^*(d, e), (x', y')\rangle + f = 0,$$

que por sua vez é igual a

$$\lambda_1 (x')^2 + \lambda_2 (y')^2 + d'x' + e'y' + f = 0,$$

onde $d' = \langle (d, e), u_1\rangle$ e $e' = \langle (d, e), u_2\rangle$.

Observe que o termo constante na equação não se altera por essa mudança de coordenadas, ao passo que, os termos x' e y' podem ser eliminados por uma translação como foi feito no Exemplo 9.3. Temos também que a matriz C associada à transformação $R^* \circ T \circ R$ é o resultado da multiplicação das matrizes

$$\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ sen\theta & \cos\theta \end{pmatrix},$$

uma vez que as colunas de B são autovetores de A, obtemos então

$$\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \lambda_1 \cos\theta & -\lambda_2 \operatorname{sen}\theta \\ \lambda_1 \operatorname{sen}\theta & \lambda_2 \cos\theta \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Por definição de determinante, a equação acima nos mostra que $\det A = \det C = \lambda_1 \cdot \lambda_2$, uma vez que a matriz de R^* é a inversa da matriz de R. Com efeito

$$\det C = \det(B^t AB) = \det B^t \det A \det B.$$

Como $B^t = B^{-1}$, então

$$\begin{aligned}
\det C = \det(B^t AB) = \det(B^{-1}AB) &= \det B^{-1} \det A \det B \\
&= (\det B)^{-1} \det A \det B \\
&= \det A.
\end{aligned}$$

Portanto a cônica definida em $\mathbb{R}^2$ pela equação

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

pode, numa nova base ortonormal de $\mathbb{R}^2$, ser representada pela equação

$$\lambda_1(x')^2 + \lambda_2(y')^2 + d'x' + e'y' + f = 0,$$

onde $d' = \langle (d,e), u_1 \rangle$ e $e' = \langle (d,e), u_2 \rangle$, com u_1, u_2 autovetores associados a λ_1, λ_2 autovalores da matriz $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

Segundo os sinais desses autovalores, as seguintes possibilidades podem ocorrer:

1°) Se $\lambda_1\lambda_2 > 0$ esta equação representa uma elipse, ou suas degenerações (um ponto ou o vazio).
2°) Se $\lambda_1\lambda_2 < 0$ esta equação representa uma hipérbole ou sua degeneração (par de retas concorrentes).
3°) Se $\lambda_1\lambda_2 = 0$ esta equação representa uma parábola ou suas degenerações (par de retas paralelas, um reta ou o vazio).

Como o determinante de A é igual ao produto de seus autovalores $\lambda_1 \cdot \lambda_2$, então o sinal de $\lambda_1 \cdot \lambda_2$ é o mesmo que $-(b^2 - ac)$. Podemos assim, reescrever as possibilidades em função $-(b^2 - ac)$;

1°) uma elipse ou suas degenerações, se $b^2 - ac < 0$,
2°) uma parábola ou suas degenerações, se $b^2 - ac = 0$,
3°) uma hipérbole ou suas degenerações, se $b^2 - ac > 0$.

Para a análise mais detalhada das possibilidades de gráficos segundo os autovalores e os determinantes vide [1], [2], [5], [7] e [8].

Exemplo 9.4 Considere a seguinte equação do segundo grau

$$16x^2 + 24xy + 9y^2 + 15x - 20y = 0. \tag{9-5}$$

Temos que a matriz A, associada ao operador linear T da forma quadrática, da equação 9-5 é

$$A = \begin{pmatrix} 16 & 12 \\ 12 & 9 \end{pmatrix}$$

e o polinômio característico de T é

$$p(\lambda) = \lambda^2 - 25\lambda$$

cujas raízes, e portanto os autovalores de T, são $\lambda_1 = 25$ e $\lambda_2 = 0$. Os autovetores correspondentes a $\lambda_1 = 25$ são as soluções não triviais do sistema

$$\begin{pmatrix} -9 & 12 \\ 12 & -16 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Portanto, segue-se que $(4,3)$ é um autovetor associado a $\lambda_1 = 25$, de modo que $v_1 = \frac{1}{5}(4,3)$ é um outro vetor *unitário* associado a λ_1. Sendo v_1 um dos vetores que formam a nova base ortonormal $\mathcal{B}$ de $\mathbb{R}^2$, então podemos determinar o outro vetor v_2 por uma mera rotação de 90° do vetor v_1, assim obtemos $v_2 = \frac{1}{4}(-3,4)$ associado a λ_2, dessa maneira a equação 9-5 reescreve-se em termos das nova base $\mathcal{B} = \{\frac{1}{5}(4,3), \frac{1}{5}(-3,4)\}$ como

$$25(x')^2 - 25(y') = 0; \qquad \text{ou seja} \qquad y' = (x')^2.$$

Haja vista que, $d' = \langle (15,-20), \frac{1}{5}(4,3) \rangle = 0$ e $e' = \langle (15,-20), \frac{1}{5}(-3,4) \rangle = -25$. Dessa forma, vemos que o gráfico da equação 9-5 é uma parábola que sofreu rotação e é exibida pela figura a seguir. Os eixos x' e y' são obtidos a partir dos eixos x e y por uma rotação anti-horária de um ângulo

$$\theta = \arctan\left(\frac{y_{v_1}}{x_{v_1}}\right) = \arctan\left(\frac{3}{4}\right) \approx 36,87°.$$

Observe que mesmo analisando as possibilidades de seu gráfico utilizando os coeficientes não é suficiente para determinar seu gráfico, uma vez que $b^2 - ac = 144 - 144 = 0$ nos diz que pode ser tanto uma parábola quanto suas degenerações, entretanto, uma vez feito o cálculo, a informação da característica serve para uma rápida confirmação dos resultados.

Exemplo 9.5 Com relação ao gráfico da seguinte equação do segundo grau

$$34x^2 - 24xy + 41y^2 - 40x - 30y - 25 = 0. \tag{9-6}$$

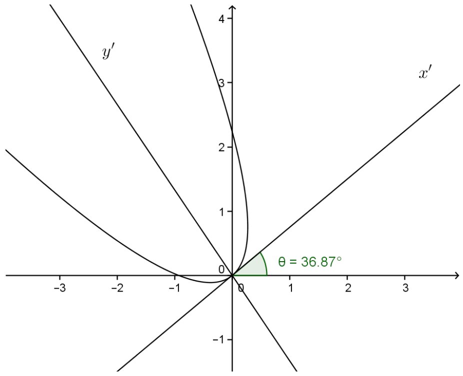


Figura 9.4: *Parábola após rotação.*

Temos que a matriz A, associada ao operador linear T da forma quadrática, da equação 9-6 é

$$A = \begin{pmatrix} 34 & -12 \\ -12 & 41 \end{pmatrix}$$

e o polinômio característico de T é

$$p(\lambda) = \lambda^2 - 75\lambda + 1250$$

cujas raízes, e portanto os autovalores de T, são $\lambda_1 = 25$ e $\lambda_2 = 50$.

Os autovetores correspondentes a $\lambda_1 = 25$ são as soluções não triviais do sistema

$$\begin{pmatrix} 9 & -12 \\ -12 & 16 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Portanto, segue-se que $(4,3)$ é um autovetor associado a $\lambda_1 = 25$, de modo que $v_1 = \frac{1}{5}(4,3)$ é um outro vetor *unitário* associado a λ_1. Sendo v_1 um dos vetores que formam a nova base ortonormal $\mathcal{B}$ de $\mathbb{R}^2$, então podemos determinar o outro vetor v_2 por uma mera rotação de $90°$ do vetor v_1, assim obtemos $v_2 = \frac{1}{4}(-3,4)$ associado a λ_2, dessa maneira a equação 9-6 reescreve-se em termos da nova base $\mathcal{B} = \{\frac{1}{5}(4,3), \frac{1}{5}(-3,4)\}$ como

$$25(x')^2 + 50(y')^2 - 50x' - 25 = 0; \qquad \text{ou seja} \qquad y' = (x')^2$$

Haja vista que, $d' = \langle(-40,-30), \frac{1}{5}(4,3)\rangle = -50$ e $e' = \langle(-40,-30), \frac{1}{5}(-3,4)\rangle = 0$. Para identificar o gráfico dessa equação transformada, completamos o quadrado como feito no Exemplo 9.3:

$$\begin{array}{rcl} 25[(x')^2 - 2x'] + 50(y')^2 - 25 & = & 0 \\ 25[(x')^2 - 2x' + 1] + 50(y')^2 - 25 + 25 & = & 0 \\ 25(x'-1)^2 + 50(y')^2 = 50 & & \end{array}$$

$$\frac{(x'-1)^2}{2} + \frac{(y')^2}{1} = 1.$$

Dessa forma, vemos que o gráfico da equação 9-6 é uma elipse que sofreu rotação seguida de uma translação por um vetor $v = (1,0)_{\mathcal{B}}$. Observe que as coordenadas de v estão na base $\mathcal{B}$, então teríamos a princípio que determinar v na base canônica e para isso deveríamos usar a matriz de passagem das coordenadas (x', y') para as coordenadas (x, y), entretanto basta observar que v retrata simplesmente o vetor v_1 da base $\mathcal{B}$, então a translação se dá de fato pelo vetor $v = v_1 = \frac{1}{5}(4,3)$ e é exibida pela figura a seguir 9.5. Os eixos x' e y' são obtidos a partir dos eixos x e y por uma rotação anti-horária de um ângulo

$$\theta = \arctan\left(\frac{y_{v_1}}{x_{v_1}}\right) = \arctan\left(\frac{3}{4}\right) \approx 36,87^\circ.$$

A característica $b^2 - ac = 144 - 1394 = -1250 < 0$ serve para confirmar nossas expec-

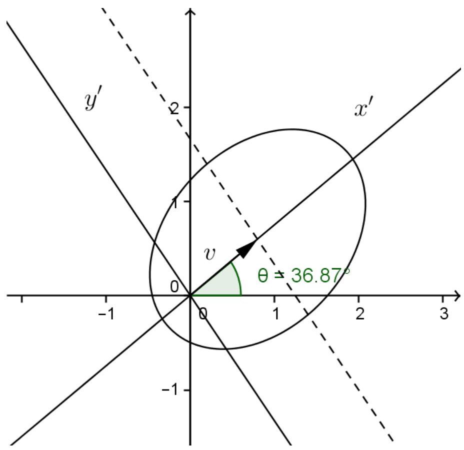


Figura 9.5: *Elipse após rotação seguida de uma translação.*

tativa de que era uma elipse.

CAPÍTULO 10

Conclusão

Nos Capítulos anteriores desenvolvemos conceitos fundamentais, nem todos básicos, sobre Álgebra Linear, cujo o foco foi os operadores lineares de $\mathbb{R}^2$. Através de vários exemplos, exploramos um vasto conjunto de transformações geométricas no plano. Desta maneira, explicitamos, a nível de ensino médio, uma outra abordagem das matrizes sem ficar restrito à resolução de sistemas lineares.

Através do Teorema Espectral para operadores auto-adjuntos do espaços de dimensão 2, este trabalho apresenta, nos Capítulos 8 e 9, uma justificativa algébrica para os cálculos que aparecem no estudo das cônicas. Poder-se-ia dizer que tal justificativa não cabe à nível de ensino médio, uma vez que, no espaço $\mathbb{R}^2$ a noção geométrica intuitiva dos alunos é suficiente para que aceitem os processo aritmético que aparecem nos estudos das cônicas. Entretanto, a compreensão do Teorema Espectral, caso particular apresentado nesse trabalho, pelos professores, que ensinam as cônicas, se justifica, uma vez que a Matemática é uma ciência dedutiva e, não, descritiva.

Referências Bibliográficas

[1] BOLDRINI, J. L.; COSTA, S. I. R.; RIBEIRO, V. L. F. F.; WETZLER, H. G. **Álgebra Linear**. Harbra, São Paulo, 1980.

[2] HEFEZ, A.; DE SOUZA FERNANDEZ, C. **Introducao a Algebra Linear**. SBM, Rio de Janeiro, 2012.

[3] HEGENBERG, L. **Matrizes, Vetores e Geometria Analítica**. Almeida Neves, Rio de Janeiro, 1971.

[4] LIMA, E. L. **Álgebra Linear**. Impa, Rio de Janeiro, 2008.

[5] LIPSCHUTZ, S. **Álgebra Linear**. MacGraw-Hill, Rio de Janeiro, 1980.

[6] MORGADO, A. C.; JÚDICE, E. D.; WAGNER, E.; LIMA, E. L.; DE CARVALHO, J. B. P.; CARNEIRO, J. P. Q.; GOMES, M. L. M.; CARVALHO, P. C. P. **Exame de textos: Análise de livros de Matemática para o ensino médio**. SBM, Rio de Janeiro, 2001.

[7] NACHBIN, L. **Introducão a Álgebra**. MacGraw-Hill, Rio de Janeiro, 1971.

[8] PENNEY, D. E.; C.H. EDWARDS, J. **Introducão à Álgebra Linear**. LTC-Livros T'ecnicos e Cient'ificos Editora S.A, Rio de Janeiro, 1998.