

INTRODUÇÃO À ÁLGEBRA LINEAR



UNIVERSIDADE FEDERAL DE MINAS GERAIS

Reitor: Clélio Campolina Diniz

Vice-Reitora: Rocksane de Carvalho Norton

Pró-Reitoria de Graduação

Pró-Reitora: Antônia Vitória Soares Aranha

Pró-Reitor Adjunto: André Luiz dos Santos Cabral

Diretor do CAED: Fernando Fidalgo

Coordenador da UAB-UFMG: Wagner José Corradi Barbosa

Coordenador Adjunto UAB-UFMG: Hormindo Pereira de Souza Júnior

EDITORA UFMG

Diretor: Wander Melo Miranda

Vice-Diretor: Roberto Alexandre do Carmo Said

Conselho Editorial

Wander Melo Miranda (presidente)

Flávio de Lemos Carsalade

Heloisa Maria Murgel Starling

Márcio Gomes Soares

Maria das Graças Santa Bárbara

Maria Helena Damasceno e Silva Megale

Paulo Sérgio Lacerda Beirão

Roberto Alexandre do Carmo Said

HAMILTON PRADO BUENO

INTRODUÇÃO À ÁLGEBRA LINEAR

BELO HORIZONTE
EDITORIA UFMG
2009

© 2009, Hamilton Prado Bueno
© 2009, Editora UFMG
2011, 1^a reimpressão
Este livro ou parte dele não pode ser reproduzido por qualquer meio sem autorização escrita do Editor.

Bueno, Hamilton Prado.
B918i Introdução à álgebra linear / Hamilton Prado Bueno. – Belo Horizonte :
Editora UFMG, 2009.

101 p. : il. (Educação a Distância)

Inclui referências.
ISBN: 978-85-7041-796-1

1. Álgebra linear. I. Título. II. Série.

CDD: 512
CDU: 512.64

Elaborada pela DITTI – Setor de Tratamento da Informação
Biblioteca Universitária da UFMG

Este livro recebeu o apoio financeiro da Secretaria de Educação a Distância do MEC.

COORDENAÇÃO DE PRODUÇÃO DE TEXTOS DE MATEMÁTICA Dan Avritzer
ASSISTÊNCIA EDITORIAL Eliane Sousa e Euclídia Macedo
EDITORAÇÃO DE TEXTOS Maria do Carmo Leite Ribeiro
REVISÃO E NORMALIZAÇÃO Maria do Rosário Alves Pereira
REVISÃO DE PROVAS Maria do Rosário Alves Pereira e Renilde Silveira
PROJETO GRÁFICO Eduardo Ferreira
FORMATAÇÃO E CAPA Sérgio Luz
ILUSTRAÇÕES Cláudio Roberto
PRODUÇÃO GRÁFICA Warren Marilac
IMPRESSÃO Imprensa Universitária da UFMG

EDITORIA UFMG
Av. Antônio Carlos, 6.627 - Ala direita da Biblioteca Central - Térreo
Campus Pampulha - 31270-901 - Belo Horizonte - MG
Tel.: + 55 31 3409-4650 - Fax: + 55 31 3409-4768
www.editora.ufmg.br - editora@ufmg.br

PRÓ-REITORIA DE GRADUAÇÃO
Av. Antônio Carlos, 6.627 - Reitoria - 6º andar
Campus Pampulha - 31270-901 - Belo Horizonte - MG
Tel.: + 55 31 3409-4054 - Fax: + 55 31 3409-4060
www.ufmg.br - info@prograd.ufmg.br - educacaoadistancia@ufmg.br

Para

Franciele, Lilian e Paulo Henrique

Os Cursos de Graduação da UFMG, modalidade a distância, foram concebidos tendo em vista dois princípios fundamentais. O primeiro se refere à democratização do acesso à educação superior; o segundo consiste na formação de profissionais de alto nível, comprometidos com o desenvolvimento do país.

A coletânea da qual este volume faz parte visa dar suporte aos estudantes desses cursos. Cada volume está relacionado a um tema, eleito como estruturante na matriz curricular. Ele apresenta os conhecimentos mínimos que são considerados essenciais no estudo do tema. Isto não significa que o estudante deva se limitar somente ao estudo do volume. Ao contrário, ele é o ponto de partida na busca de um conhecimento mais amplo e aprofundado sobre o assunto. Nessa direção, cada volume apresenta uma bibliografia, com indicação de obras impressas e virtuais que deverão ser consultadas à medida que se fizer necessário.

Cada volume da coletânea está dividido em aulas, que consistem em unidades de estudo do tema tratado. Os objetivos, apresentados em cada início de aula, indicam as competências e habilidades que o estudante deve adquirir ao término de seu estudo. As aulas podem se constituir em apresentação, reflexões e indagações teóricas, em experimentos ou em orientações para atividades a serem realizadas pelos estudantes.

Para cada aula ou conjunto de aulas, foi elaborada uma lista de exercícios com o objetivo de levar o estudante a avaliar o seu progresso e a desenvolver estratégias de metacognição ao se conscientizar dos diversos aspectos envolvidos em seus processos cognitivos. Essa lista auxiliará o estudante a tornar-se mais autônomo, responsável, crítico, capaz de desenvolver sua independência intelectual. Caso ela mostre que as competências e habilidades indicadas nos objetivos não foram alcançadas, o aluno deverá estudar com mais afinco e atenção o tema proposto, reorientar seus estudos ou buscar ajuda dos tutores, professores especialistas e colegas.

Agradecemos a todas as instituições que colaboraram na produção desta coletânea. Em particular, agradecemos às pessoas (autores, coordenador da produção gráfica, coordenadores de redação, desenhistas, diagramadores, revisores) que dedicaram seu tempo, e esforço na preparação desta obra que, temos certeza, em muito contribuirá para a educação brasileira.

*Maria do Carmo Vila
Coordenadora do Centro de Apoio à Educação a Distância
UFMG*

Sumário

Apresentação	11
Ao Aluno	13
Aula 1 - Conceitos Fundamentais	15
1.1 Vetores na Física e na Matemática	15
1.2 Sistemas lineares e o método de Gauss-Jordan	21
1.3 Cálculo de determinantes	26
1.4 Exercícios	29
Aula 2 - O Espaço \mathbb{R}^n.....	31
2.1 Equação paramétrica do plano	31
2.2 Sistemas lineares em três variáveis	33
2.3 O espaço \mathbb{R}^n	35
2.4 Espaços vetoriais abstratos	37
2.5 Exercícios	40
Aula 3 - Subespaços do \mathbb{R}^n e Bases.....	41
3.1 Subespaços e combinações lineares	41
3.2 Bases	49
3.3 Dimensão	51
3.4 Exercícios	55
Aula 4 - Aplicações Lineares	57
4.1 Aplicações lineares e matrizes – Parte I	57
4.2 Espaço linha e espaço coluna	62
4.3 Multiplicação de matrizes	67
4.4 Exercícios	69
Aula 5 - O Teorema do Núcleo e da Imagem	71
5.1 Teorema (da dimensão) do núcleo e da imagem	71
5.2 Isomorfismo e inversas	72
5.3 Obtenção da inversa de uma matriz	75
5.4 Exercícios	78
Aula 6 - Mudanças de Base	79
6.1 Representação de um vetor em uma base	79
6.2 Aplicações lineares e matrizes – Parte II	81
6.3 Aplicação: diagonalização de uma matriz	86
6.4 Exercícios	89

Aula 7 - O Teorema de Cayley-Hamilton	91
7.1 Polinômios de aplicações lineares	91
7.2 Subespaços invariantes.....	92
7.3 O Teorema de Cayley-Hamilton	93
7.4 Aplicações	95
7.5 Exercícios.....	96
Referências	97
Índice remissivo.....	99
Sobre o autor.....	101

Apresentação

Este texto foi redigido para a disciplina Álgebra Linear I, do curso de licenciatura em Matemática da UFMG, na modalidade a distância.

Há alguns anos, escrevi um livro para um curso de Álgebra Linear. Ele destinava-se a alunos mais avançados, que já haviam estudado a parte básica dessa disciplina. Apesar de introduzir todos os conceitos de um primeiro curso de Álgebra Linear, essa parte do texto foi escrita com o propósito de ser sucinta e não detalhava resultados considerados conhecidos.

A oportunidade de escrever este texto tornou possível apresentar, com um enfoque que julgo adequado, tópicos de um primeiro curso de Álgebra Linear. Contudo, a estrutura do curso a distância fez com que meu texto se tornasse mais abstrato do que intencionava: para tornar possível o ensino de assuntos básicos da Álgebra Linear em apenas dois meses, a estrutura euclidiana – quer dizer, produtos internos e normas – foi relegada a um segundo curso, continuação deste. Assim, aspectos geométricos que julgo fundamentais foram postergados para esse futuro curso.

Para sanar essa deficiência, tentei apresentar uma visão geométrica dos assuntos abordados. Uma vez concluída a redação deste livro, julguei o resultado final como bastante aceitável, ainda mais levando em conta o exíguo período em que ele foi redigido. Com a oportunidade de ensiná-lo e então ouvir a opinião dos alunos sobre ele, melhorias podem ser implementadas e suas deficiências minoradas.

O desafio de escrever sobre tópicos básicos me fez procurar por demonstrações particularmente simples de resultados fundamentais. As demonstrações aqui apresentadas contaram com o auxílio do Prof. Helder Cândido Rodrigues. Não existem palavras de agradecimento capazes de dar uma ideia do quanto essa colaboração foi decisiva.

Muitos dos exercícios propostos neste texto tem sua origem nos livros de Álgebra Linear do Prof. Reginaldo Santos. Tomei a liberdade de utilizá-los, conhecedor do enorme desprendimento que o caracteriza.

Agradeço à Profa. Maria Cristina Costa Ferreira por sugestões e correções que muito aprimoraram este texto.

Hamilton Prado Bueno

Ao aluno

Os assuntos aqui apresentados dão continuidade ao livro *Geometria Analítica e Álgebra Linear: uma visão geométrica*, de Dan Avritzer. Assim, muitos resultados daquele texto são supostamente conhecidos. Por outro lado, meu enfoque sobre alguns assuntos difere do apresentado naquele texto. Isso motivou-me a apresentá-los novamente.

Há uma enorme diferença entre um texto de Geometria Analítica (mesmo com uma abordagem utilizando a Álgebra Linear, como no texto de Dan Avritzer) e um texto de Álgebra Linear. A apresentação torna-se muito mais abstrata, o que é uma imposição para se conseguir resultados mais gerais. Assim, ao estudar este texto, você se defrontará com resultados cujo significado terá dificuldades para compreender. Essa situação é normal e só será superada com muitas horas de estudo.

O curso apresenta muitos cálculos importantes. Nossa objetivo principal não é apenas que você seja capaz de realizá-los, mas que entenda o porquê de seu funcionamento. Quer dizer, você precisa compreender bem o assunto que está sendo exposto; caso contrário, chegará a contas que não terão qualquer significado.

A compreensão da estrutura da Álgebra Linear é o desafio que você terá que enfrentar neste curso. A demonstração dos resultados expostos (teorema, lemas, proposições) esclarece essa estrutura, não sendo apenas uma parte “desagradável e incompreensível” do assunto. Cada resultado deve ser lido e relido até que seu significado aflore. Se isso não for suficiente, tente seguir sua demonstração em um exemplo particular, que você mesmo pode elaborar. Além disso, você dispõe de contato quase direto com os monitores ou professores, contato esse que lhe ajudará na compreensão. Mas você terá que se dedicar bastante ao curso para superar suas dificuldades. Esse desafio é o mesmo enfrentado por alunos de cursos presenciais.

O fato de o curso ser a distância apresenta – o que à primeira vista parece paradoxal – a possibilidade de uma interação mais profunda entre você e seus professores ou tutores. Dúvidas podem ser sanadas rapidamente, graças à comunicação por meio da internet. Se desprezar esse canal de comunicação, o ritmo do curso fará com que você logo se sinta completamente desorientado. Assim, a sua participação no curso por meio desses canais de comunicação é decisiva para o seu aprendizado.

AULA 1

Conceitos Fundamentais

OBJETIVOS

No final desta aula, o aluno deverá ser capaz de:

1. distinguir o uso de vetores na Física e na Matemática;
2. resolver sistemas lineares pelo método de Gauss-Jordan;
3. calcular determinantes por meio do escalonamento de uma matriz.

1.1 VETORES NA FÍSICA E NA MATEMÁTICA

Nosso primeiro contato com vetores aconteceu antes de ingressarmos no ensino superior; provavelmente em uma aula de Física, vetores nos foram apresentados como “grandezas com módulo, direção e sentido” – tais como força, velocidade e aceleração. Aprendemos que vetores eram representados por setas: a direção sendo a de uma reta contendo aquela seta, o sentido indicado pela seta e o módulo como o tamanho da seta. Sendo mais preciso, cada reta do espaço define uma direção, convencionando-se que retas paralelas definem a mesma direção. Escolhido um segmento em uma reta, ao orientarmos esse segmento escolhemos um sentido para o vetor. Finalmente, o comprimento do segmento é o módulo desse vetor.

Assim, um vetor \vec{v} fica definido ao escolhermos dois pontos no espaço: o ponto inicial P_i e o ponto final P_f . Mas segmentos paralelos, com o mesmo sentido e o mesmo módulo, representam o mesmo vetor. No caso de termos $P_i = P_f$, temos o vetor $\vec{0}$. Note que ao vetor $\vec{0}$ não define uma direção!

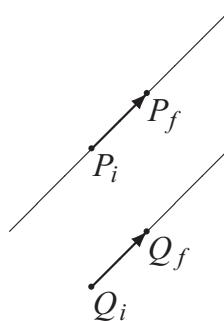


Figura 1.1: Os segmentos orientados $\overrightarrow{P_iP_f}$ e $\overrightarrow{Q_iQ_f}$ representam o mesmo vetor

Aprendemos ainda que vetores podem ser somados e multiplicados por escalares (o nome que passamos a utilizar para nos referir a números reais). A soma dos vetores \vec{u} e \vec{v} , definidos pelos pontos P_i , P_f e Q'_i , Q'_f é obtida ao transladarmos o vetor \vec{v} , de modo que seu ponto inicial seja o ponto P_f e seu ponto final o ponto Q_f ; o vetor $\vec{u} + \vec{v}$ é dado pelo segmento orientado unindo P_i e Q_f .

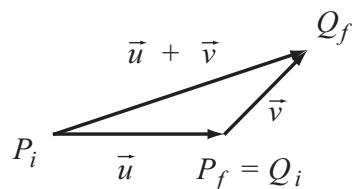


Figura 1.2: A adição de vetores

Exercício 1.1 Na figura 1.2, os vetores \vec{u} e \vec{v} pertencem ao “plano” do papel deste livro. Se os vetores pertencem ao espaço, ainda assim a figura está correta?

A multiplicação do vetor \vec{v} pelo escalar α é definida como o vetor $\alpha\vec{v}$ com a mesma direção do vetor \vec{v} , com o mesmo sentido, se $\alpha > 0$, e com sentido contrário, se $\alpha < 0$. O módulo do vetor $\alpha\vec{v}$ é definido por $\|\alpha\vec{v}\| = |\alpha| \|\vec{v}\|$, em que $|\alpha|$ é o valor absoluto do escalar α .

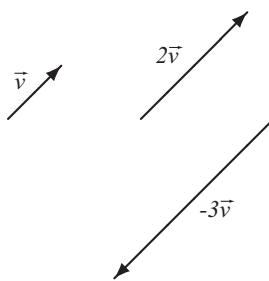


Figura 1.3: A multiplicação de um vetor por um escalar

Exercício 1.2 Como é definido $\alpha\vec{v}$, se $\alpha = 0$?

Se a ideia da soma de vetores é clara, a sua obtenção prática no caso de vetores definidos pelas coordenadas de seus pontos inicial e final não é tão simples: dados dois vetores $\vec{v} = \overrightarrow{P_i P_f}$ e $\vec{w} = \overrightarrow{Q_i Q_f}$ (isto é, dois segmentos orientados de retas), o vetor $\vec{v} + \vec{w}$ é obtido por meio de uma reta r , paralela à reta definida por \vec{v} , passando pelo ponto P_f . Nessa reta, obtemos dois pontos cuja distância ao ponto P_f é a mesma que a distância entre Q_i e Q_f . Ao escolhermos a solução R_f que define o mesmo sentido de $\overrightarrow{Q_i Q_f}$, o segmento orientado $\overrightarrow{P_f R_f}$ é o vetor $\vec{v} + \vec{w}$. Veja a Figura 1.4.

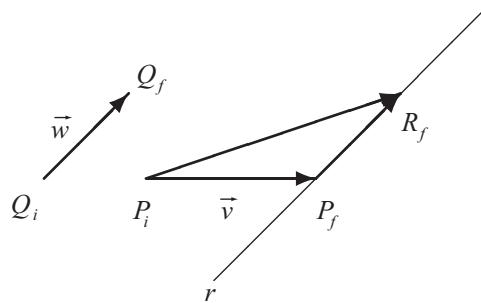


Figura 1.4: A construção geométrica da soma de vetores dados por suas coordenadas, a partir da definição

Observação 1.3 Não podemos descrever a reta r como $P_f + t\vec{w}$, pois (ainda) não sabemos como somar as coordenadas do ponto P_f com as coordenadas do vetor \vec{w} . Assim, ainda não sabemos como obter a solução $R_f = P_f + 1\vec{w}$! \triangleleft

Exercício 1.4 Considere os vetores $\vec{v} = \overrightarrow{P_i P_f}$ e $\vec{w} = \overrightarrow{Q_i Q_f}$ do plano, definidos pelos pontos $P_i = (0, 0)$, $P_f = (2, 3)$, $Q_i = (1, 5)$ e $Q_f = (7, 2)$. Utilizando o método descrito no parágrafo anterior, obtenha os pontos inicial R_i e final R_f do vetor $\vec{v} + \vec{w}$.

Exercício 1.5 Repita o exercício anterior no caso dos vetores espaciais $\vec{v} = \overrightarrow{P_i P_f}$ e $\vec{w} = \overrightarrow{Q_i Q_f}$, se $P_i = (0, 0, 3)$, $P_f = (2, 3, 7)$, $Q_i = (1, 5, 2)$ e $Q_f = (7, 2, 2)$.

Os cálculos feitos na solução dos exercícios anteriores nos mostram como é difícil operar com vetores desse modo.

A solução “matemática” para a resolução daqueles exercícios é bastante interessante e corresponde a uma utilização consciente do sistema de coordenadas cartesianas.

Em primeiro lugar, notamos que a cada vetor $\overrightarrow{Q_i Q_f}$ corresponde um único vetor $\overrightarrow{0P}$, cujo ponto inicial é a origem 0 e cujo ponto final é o ponto P . Se os vetores $\overrightarrow{Q_i Q_f}$ e $\overrightarrow{R_i R_f}$ são iguais, a eles corresponde o mesmo ponto P . Em outras palavras, cada vetor $\overrightarrow{P_i P_f}$ da Física é identificado com um único ponto P . Ou seja, um vetor da Física corresponde a um ponto P do sistema de coordenadas cartesianas.

Na Matemática, um vetor é um ponto do espaço (ou do plano).

Por esse motivo, matemáticos usualmente denotam vetores por u , v , w , ao invés de \vec{u} , \vec{v} , \vec{w} . Em algumas situações, é interessante distinguir entre vetores e pontos. Assim, quando queremos nos referir simplesmente ao ponto P (e não ao vetor definido por esse ponto), mantemos a notação de pontos: P , Q , R .

Por outro lado, quando nos referirmos a um vetor v no sentido da Física, manteremos a notação \vec{v} .

Para continuarmos, verificaremos duas propriedades básicas da adição de vetores: ela é comutativa e associativa. Ou seja, $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ e $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$. A comutatividade da adição é ilustrada pela Figura 1.5.

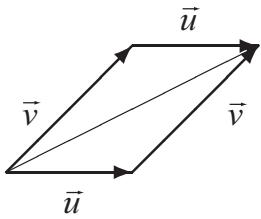


Figura 1.5: A adição de vetores é comutativa, pois vetores são somados de acordo com a “regra do paralelogramo”

Por outro lado, a associatividade é simples: suponhamos que $\vec{u} = \overrightarrow{P_i P_f}$, $\vec{v} = \overrightarrow{Q_i Q_f}$ e $\vec{w} = \overrightarrow{R_i R_f}$. Para simplificar a notação, vamos supor que $P_f = Q_i$ e $Q_f = R_i$. Então,

$$\begin{aligned} (\vec{u} + \vec{v}) + \vec{w} &= (\overrightarrow{P_i P_f} + \overrightarrow{Q_i Q_f}) + \overrightarrow{R_i R_f} = \overrightarrow{P_i Q_f} + \overrightarrow{R_i R_f} \\ &= \overrightarrow{P_i R_f} \\ &= \overrightarrow{P_i P_f} + \overrightarrow{Q_i R_f} = \overrightarrow{P_i P_f} + (\overrightarrow{Q_i Q_f} + \overrightarrow{R_i R_f}) = \vec{u} + (\vec{v} + \vec{w}). \end{aligned}$$

Confira esses cálculos na Figura 1.6.

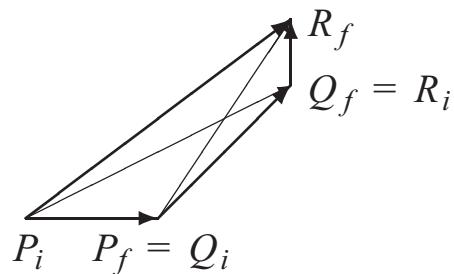


Figura 1.6: A adição de vetores é associativa

As propriedades comutativa e associativa da adição de vetores nos permitem reordenar termos de uma soma da maneira que mais nos convier e torna desnecessária a utilização de parênteses em uma adição de vetores. Isso é decisivo para o raciocínio que passaremos a descrever.

Consideremos os vetores $v = (x_0, y_0, z_0)$ e $w = (x_1, y_1, z_1)$. Denotaremos $v_1 = (x_0, 0, 0)$, $v_2 = (0, y_0, 0)$ e $v_3 = (0, 0, z_0)$. É geometricamente claro que $v = \vec{v}_1 + \vec{v}_2 + \vec{v}_3$ (veja a Figura 1.7). Usando a notação matemática, $v = v_1 + v_2 + v_3$.

Procedemos de maneira análoga com o vetor w .

Agora consideremos $v_1 = (x_0, 0, 0)$ e $w_1 = (x_1, 0, 0)$. Novamente é claro que $v_1 + w_1 = (x_0 + x_1, 0, 0)$. (Veja a Figura 1.8.) Assim, $v_1 + w_1 = (x_0 + x_1, 0, 0)$.

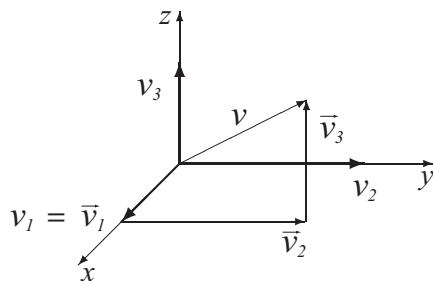
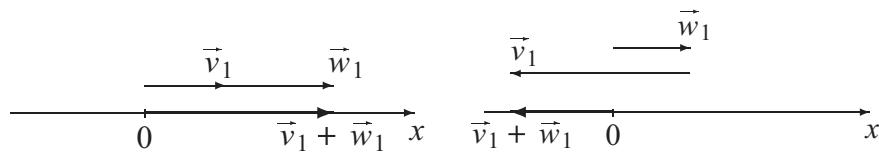
Figura 1.7: Vale a igualdade $v = v_1 + v_2 + v_3$ 

Figura 1.8: A soma de vetores em uma mesma direção é obtida ao somar suas coordenadas

Naturalmente, o mesmo procedimento também se aplica às somas $v_2 + w_2$ e $v_3 + w_3$. Ora, então temos

$$\begin{aligned} v + w &= (v_1 + v_2 + v_3) + (w_1 + w_2 + w_3) \\ &= (v_1 + w_1) + (v_2 + w_2) + (v_3 + w_3) \\ &= (x_0 + x_1, 0, 0) + (0, y_0 + y_1, 0) + (0, 0, z_0 + z_1) \\ &= (x_0 + x_1, y_0 + y_1, z_0 + z_1). \end{aligned}$$

Em outras palavras, o tratamento anterior nos mostra que podemos encontrar facilmente a soma de dois vetores, se conhecemos as coordenadas de ambos: basta somar as coordenadas correspondentes.

Exercício 1.6 Justifique: se $v = (x_0, y_0, z_0)$, então $\alpha v = (\alpha x_0, \alpha y_0, \alpha z_0)$.

Em particular, $-v = (-1)v = (-x_0, -y_0, -z_0)$, de modo que está definida a subtração de dois vetores: $v - w = v + (-w)$.

Uma vez resolvido o exercício anterior, falta apenas um passo para encontrarmos uma solução prática para os Exercícios 1.4 e 1.5.

Consideremos o vetor $\overrightarrow{P_i P_f}$ definido pelos pontos inicial $P_i = (x_0, y_0, z_0)$ e final $P_f = (x_1, y_1, z_1)$. Qual o ponto P que corresponde a esse vetor da Física? Em outras palavras, qual o vetor v da Matemática correspondente a esse vetor da Física?

Examinando a Figura 1.9, vemos que $\overrightarrow{0P_i} + \overrightarrow{P_i P_f} = \overrightarrow{0P_f}$. Assim, $\overrightarrow{P_i P_f} = \overrightarrow{0P_f} - \overrightarrow{0P_i}$. Como os vetores do lado direito da última igualdade têm seu ponto inicial na origem (correspondendo assim a vetores da Matemática), acabamos de verificar que $\overrightarrow{P_i P_f}$ corresponde ao vetor $v = (x_1 - x_0, y_1 - y_0, z_1 - z_0)$ da Matemática.

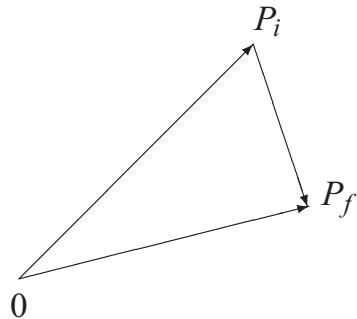
Exercício 1.7 Refaça os Exercícios 1.4 e 1.5.

Figura 1.9: O vetor $v = (x_1 - x_0, y_1 - y_0, z_1 - z_0)$ da matemática corresponde ao vetor $\overrightarrow{P_i P_f}$, se $P_i = (x_0, y_0, z_0)$ e $P_f = (x_1, y_1, z_1)$

Exercício 1.8 Verifique que todo procedimento descrito anteriormente permanece válido para vetores do plano. Em outras palavras, verifique que a adição dos vetores $u = (a, b)$ e $v = (c, d)$ é dada pelo vetor $u + v = (a + c, b + d)$ e que $\lambda u = (\lambda a, \lambda b)$. Se $\vec{u} = P_i \vec{P}_f$ for um vetor (da Física) determinado pelos pontos $P_i = (x_0, y_0)$ e $P_f = (x_1, y_1)$, verifique que ao vetor \vec{u} corresponde o vetor (da Matemática) $u = (x_1 - x_0, y_1 - y_0)$.

Essa abordagem de vetores tem inúmeras vantagens. Mas também tem desvantagens: em alguns casos, o fato geométrico a ser descrito fica muito mais claro utilizando o conceito de vetor no sentido físico. Veja a Figura 1.10.

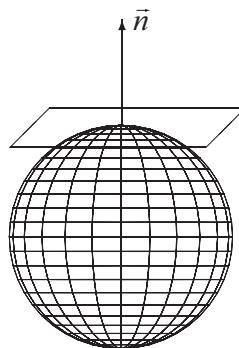


Figura 1.10: O vetor $\vec{n} = (0, 0, 1)$ é normal à esfera $x^2 + y^2 + z^2 = 1$ no ponto $(0, 0, 1)$

Exercício 1.9 Por que a figura anterior não corresponde ao sentido de vetor utilizado na Matemática?

Quando conveniente, ilustramos figuras utilizando o conceito físico de vetor. Essa situação ocorre com frequência no estudo da Geometria Analítica, abordada no livro *Geometria Analítica e Álgebra Linear: uma Visão Geométrica*, de Dan Avritzer. Você está convidado a rever as Aulas 1 a 4 do tomo II daquele livro.

1.2 SISTEMAS LINEARES E O MÉTODO DE GAUSS-JORDAN

Para $1 \leq i \leq m$ e $1 \leq j \leq n$, suponhamos conhecidos os valores a_{ij} e os valores b_j . Um *sistema linear* em m equações e n incógnitas procura a solução x_1, \dots, x_n que satisfaz

$$\begin{array}{lcl} a_{11}x_1 + \dots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n & = & b_m. \end{array}$$

Como sabemos, esse sistema pode ser escrito utilizando matrizes

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

ou,

$$Ax = b.$$

Se $b = 0$, o sistema é chamado *homogêneo*; se $b \neq 0$, o sistema é *não homogêneo*. Os sistemas $Ax = b$ e $Ax = 0$ relacionam-se de um modo especial, de modo que informações sobre as soluções de um fornecem dados importantes para a solução do outro. Por esse motivo, no estudo do sistema $Ax = b$, o sistema $Ax = 0$ é chamado *sistema homogêneo associado*.

Vamos estudar o sistema $Ax = b$. Para isso, mais sinteticamente ainda, representaremos esse sistema por uma única matriz, chamada *matriz aumentada do sistema*:

$$\mathbf{A} = (A | b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right).$$

É fácil verificar que as seguintes operações sobre as linhas da matriz **A** não alteram o conjunto de soluções do sistema $Ax = b$:

- (a) transpor as linhas i e j ;
- (b) multiplicar a linha i por um escalar não nulo;
- (c) substituir a linha j por sua soma com um múltiplo da linha i .¹

As operações (a), (b) e (c) são as *operações elementares* sobre as linhas da matriz **A**.

Consideremos então uma matriz satisfazendo as seguintes propriedades:

- se existir o primeiro elemento não nulo da linha i (chamado *pivô* da linha i) e se esse ocorrer na coluna j , então, se existir o pivô da linha $i + \ell$, esse ocorre numa coluna $k > j$, para todo $\ell \in \{1, \dots, m - i\}$;
- o pivô de cada linha é igual a 1, se ocorrer na matriz **A**.

¹ Com relação à operação (c), note que $x = (x_1, x_2, \dots, x_n)$ satisfaz

$$a_{i1}x_1 + \dots + a_{in}x_n = b_i$$

$$a_{j1}x_1 + \dots + a_{jn}x_n = b_j$$

se, e somente se, satisfizer

$$a_{i1}x_1 + \dots + a_{in}x_n = b_i$$

$$(a_{j1} + \alpha a_{i1})x_1 + \dots + (a_{jn} + \alpha a_{in})x_n = b_j + \alpha b_i$$

Dizemos então que essa matriz (ou o sistema) está na forma *escalonada* e uma sucessão de operações elementares utilizadas para levar uma matriz qualquer C até uma matriz na forma escalonada é um *escalonamento* da matriz C .

Dada uma matriz arbitrária $C = (c_{ij})$, a sucessiva aplicação de operações elementares (sobre suas linhas) pode levá-la até uma forma escalonada. De fato, se existir algum elemento não nulo na primeira coluna de C , ao aplicarmos as operações elementares (a) e (b), obtemos uma nova matriz $C' = (c'_{ij})$, com $c'_{11} = 1$. A aplicação da operação elementar (c) torna possível transformar em zero qualquer outro elemento não nulo da primeira coluna. O resultado então segue-se daí por indução sobre o número de linhas de C .

Contudo, dada uma matriz A , a sucessiva aplicação de operações elementares pode conduzir a diferentes formas escalonadas para essa matriz. É o que veremos no próximo exemplo.

Exemplo 1.10 Considere o sistema

$$\mathbf{A} = (A \mid b) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 2 & 2 & 2 & 2 \end{array} \right).$$

Subtraindo da segunda linha duas vezes a primeira e então dividindo por (-2) a (nova) segunda linha, obtemos

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \end{array} \right),$$

matriz que está na forma escalonada.

Por outro lado, trocando as duas linhas da matriz original \mathbf{A} , dividindo a (nova) primeira linha por 2 e então subtraindo a segunda linha, chegamos a

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{array} \right),$$

que também está na forma escalonada.

Assim, a uma mesma matriz podem corresponder diferentes formas escalonadas! Note, entretanto, que os pivôs são os mesmos nas duas formas obtidas. \triangleleft

Suponhamos agora que uma matriz C esteja na forma *escalonada*. Se cada pivô for o único elemento não nulo de sua coluna, dizemos que a matriz está em sua forma *escalonada reduzida por linhas*. Aplicando a operação elementar (c), podemos fazer com que uma matriz na forma escalonada atinja sua forma reduzida por linhas. De fato, consideremos o pivô da última linha não nula de C . A aplicação da operação elementar (c) torna possível zerar os elementos que estão acima do pivô, mantendo ainda a matriz na forma escalonada. A demonstração agora segue-se daí por indução, aplicando o mesmo procedimento ao pivô da penúltima linha não nula de C e assim sucessivamente.

A forma escalonada reduzida por linhas de uma matriz é única. Mostraremos esse resultado no Teorema 1.17.

Exemplo 1.11 Consideremos o sistema $Ax = b$, cuja matriz aumentada é dada por

$$\left(\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & b_1 \\ -1 & 0 & 0 & 0 & 0 & 0 & b_2 \\ 0 & 1 & 0 & 0 & 0 & 0 & b_3 \\ 1 & 1 & 1 & 1 & 0 & 0 & b_4 \\ 0 & 0 & 0 & 0 & 1 & -3 & b_5 \end{array} \right).$$

Queremos determinar para quais valores de b_1, \dots, b_5 o sistema tem solução. Se ele tiver solução, queremos determiná-la.

Levando a matriz aumentada do sistema à forma escalonada reduzida por linhas, obtemos

$$\left(\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & b_1 \\ 0 & 1 & 0 & 0 & 0 & 0 & b_3 \\ 0 & 0 & 1 & 1 & 0 & 0 & b_4 - b_1 - b_3 \\ 0 & 0 & 0 & 0 & 1 & -3 & b_5 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_1 + b_2 \end{array} \right).$$

A última linha nos mostra que esse sistema apenas possui solução se tivermos $b_1 + b_2 = 0$. Quer dizer, se tivermos $b_1 + b_2 \neq 0$, o sistema não tem solução.

Se o sistema tiver solução, podemos determiná-las.² Suponhamos, portanto, que $b_1 + b_2 = 0$. Escrevemos as variáveis correspondentes aos pivôs em termos das demais variáveis (chamadas *variáveis livres*):

$$\begin{aligned} x_1 &= b_1 \\ x_2 &= b_3 \\ x_3 &= (b_4 - b_1 - b_3) - x_4 \\ x_4 &= x_4 \\ x_5 &= b_5 + 3x_6 \\ x_6 &= x_6 \end{aligned}$$

Podemos escrever essa resposta de uma maneira que se mostrará bastante útil:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_3 \\ b_4 - b_1 - b_3 \\ 0 \\ b_5 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_6 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 3 \\ 1 \end{pmatrix}. \quad (1.1)$$

Quer dizer, para quaisquer valores de b_1, b_3, b_4 e b_5 , e para quaisquer valores escolhidos para as variáveis livres x_4 e x_6 , a solução do sistema $Ax = b$ (com $b_1 + b_2 = 0$) é dada pela expressão anterior. O sistema tem infinitas soluções, resultantes da escolha de valores para x_4 e x_6 . \triangleleft

Observação 1.12 O exemplo anterior deixa claro que a existência de soluções para um sistema $Ax = b$, sendo A uma matriz $m \times n$, não depende diretamente de m e n . Ele não possuirá solução se, na forma escalonada reduzida por linhas de $(A|b)$, não tivermos uma linha no formato $(0, c)$, com $c \neq 0$. (Esse é o formato da última linha no sistema anterior, se tivermos $b_1 + b_2 \neq 0$.) Se esse não for o caso, o sistema sempre possuirá solução: se existirem variáveis livres (quer dizer, o número de pivôs não for igual a n),

² Sendo mais explícito, estou dizendo que a denominação “sistema indeterminado”, utilizada no ensino médio, é inadequada.

então o sistema possuirá infinitas soluções, resultantes das infinitas escolhas de valores para as variáveis livres. Se não existirem variáveis livres, o sistema possuirá uma única solução. \triangleleft

Exercício 1.13 Escreva explicitamente o sistema considerado no Exemplo 1.11.

Exercício 1.14 Dê um exemplo de um sistema com uma equação e duas incógnitas que possua as mesmas soluções de um sistema com duas equações e duas incógnitas.

Exercício 1.15 Dê um exemplo de um sistema com duas equações e duas incógnitas que não possua solução.

É claro que, se estivermos tratando de um sistema homogêneo $Ax = 0$, não há necessidade de trabalhar com a matriz aumentada do sistema. É o que faremos ao tratar esse tipo de sistema.

Exercício 1.16 Justifique a afirmação anterior em termos de operações elementares.

Teorema 1.17 Qualquer que seja a matriz A , ela possui uma única forma escalonada reduzida por linhas.

Demonstração: Faremos indução no número de colunas da matriz A . Se A possuir uma única coluna, sua forma escalonada reduzida por linhas é

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

a primeira possibilidade ocorrendo se A possuir um elemento não nulo e a segunda, caso contrário.

Suponhamos, portanto, o resultado válido para qualquer matriz com $n - 1$ colunas e consideremos uma matriz A com n colunas. Suponhamos que R_1 e R_2 sejam formas escalonadas reduzidas por linhas da matriz A . Interpretando essas matrizes como matrizes aumentadas de um sistema, então $R_1 = (R|a)$ e $R_2 = (R|b)$, já que nossa hipótese de indução aplica-se às colunas de R . Agora consideramos duas opções para a coluna a : se ela possuir um pivô, então o sistema $Rx = a$ não possui solução, pois a linha de R correspondente à posição desse pivô seria identicamente nula, já que R_1 está na forma escalonada reduzida por linhas. Como as soluções de $Rx = a$ e $Rx = b$ são as mesmas, b também tem um pivô e, portanto, $a = b$, já que R_1 e R_2 estão na forma escalonada reduzida por linhas. Se em a não existir um pivô, então o sistema $Rx = a$ possui uma solução x_0 . Mas então $a = Rx_0 = b$, provando que $a = b$ e que $R_1 = R_2$. \square

Vamos agora explicitar a relação entre o sistema $Ax = b$ e seu sistema homogêneo associado, $Ax = 0$. Partimos de uma observação muito simples: o sistema homogêneo sempre tem solução! De fato, se tomarmos o valor de todas as incógnitas como sendo igual a zero, obtemos uma solução do sistema, chamada solução *trivial*.

A solução trivial será única, se não existirem variáveis livres. Isso implica, em particular, que a forma escalonada reduzida por linhas do sistema possui o mesmo número de pivôs e de incógnitas. Assim, desprezadas as possíveis linhas identicamente nulas da forma escalonada reduzida por linhas, o que resta é a matriz identidade.

Exercício 1.18 Justifique a afirmação feita no parágrafo anterior.

Se existirem variáveis livres (ou *uma* variável livre), o sistema $Ax = 0$ possuirá infinitas soluções, obtidas ao se atribuir diferentes valores a cada variável livre. É o que acontece no próximo resultado:

Teorema 1.19 Considere um sistema homogêneo $Ax = 0$. Se A for uma matriz $m \times n$, com $m < n$, então $Ax = 0$ possui infinitas soluções. Ou seja, qualquer sistema homogêneo com mais incógnitas do que equações possui infinitas soluções.

Demonstração: A forma escalonada reduzida por linhas de A possui um número r de pivôs que é, no máximo, igual ao número de equações. Assim, ela possui $n - r$ de variáveis livres e, portanto, infinitas soluções. \square

Definição 1.20 Sejam A uma matriz $m \times n$. Definimos o *núcleo* de A , denotado $\ker A$,³ como sendo o conjunto de soluções do sistema $Ax = 0$.

Teorema 1.21 Suponha que x_0 seja uma solução do sistema $Ax = b$, isto é, $Ax_0 = b$. Se x_1 também for uma solução do sistema $Ax = b$, então $x_1 = x_0 + z$, em que $z \in \ker A$. Em particular, se $Ax = 0$ só possuir a solução trivial, a solução de $Ax = b$ será única.

Demonstração: Suponhamos que $z \in \ker A$. Então $x_0 + z$ é solução do sistema $Ax = b$, pois $A(x_0 + z) = Ax_0 + Az = b + 0 = b$. Quer dizer, $x_0 + z$ é solução de $Ax = b$, para todo $z \in \ker A$.

Suponhamos agora que $Ax_1 = b$, ou seja, que x_1 seja também solução de $Ax = b$. Consideremos $x_1 - x_0$. Então $A(x_1 - x_0) = Ax_1 - Ax_0 = b - b = 0$, que dizer, $(x_1 - x_0) \in \ker A$. Denotando $z = x_1 - x_0$, temos $x_1 = x_0 + (x_1 - x_0) = x_0 + z$, o que completa a demonstração. \square

Voltando ao Exemplo 1.11, podemos agora interpretar a equação (1.1). O primeiro termo do lado direito (o termo dependente de b) é uma solução particular de $Ax = b$ (no caso $b_1 + b_2 = 0$). Os termos seguintes (correspondentes às variáveis livres x_4 e x_6) nos fornecem todas as soluções do sistema homogêneo associado. Observe que isso é imediato, pois corresponde à escolha $b_1 = \dots = b_6 = 0$.

³ A notação \ker vem do inglês: *kernel* quer dizer *núcleo*.

Exercício 1.22 Considere o sistema

$$\begin{array}{rcl} 3x & + & 2y & + & 3z & = & 8 \\ x & + & y & + & z & = & 3 \\ 2x & + & y & - & z & = & 2. \end{array}$$

Sabendo que $(1, 1, 1)$ é uma de suas soluções, ache todas as soluções do sistema.

1.3 CÁLCULO DE DETERMINANTES

Definição 1.23 Uma matriz quadrada é **triangular superior** se todas as suas entradas abaixo da diagonal principal são nulas.

Exercício 1.24 Mostre que o determinante de uma matriz triangular superior é o produto de suas entradas na diagonal principal. Defina matriz *triangular inferior* e mostre o resultado análogo.

Exercício 1.25 Justifique: se uma matriz quadrada estiver na forma escalonada, então ela é triangular superior.

O escalonamento de uma matriz (e o fato do determinante de uma matriz triangular superior ser o produto das entradas diagonais dessa matriz) nos fornece um método eficiente para o cálculo do determinante de uma matriz.

De fato, sabemos que a aplicação da operação fundamental (a) a uma matriz faz com que seu determinante seja multiplicado por -1 . A operação fundamental (b) faz com que o determinante seja multiplicado pelo valor c , enquanto a operação fundamental (c) não altera o valor do determinante. (As operações fundamentais (a), (b) e (c) foram descritas na seção 1.2.)

Vejamos um exemplo do cálculo do determinante de uma matriz:

Exemplo 1.26 Consideremos a matriz

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 1 & 3 & 3 \\ 4 & 3 & 2 & 1 \end{pmatrix}.$$

Multiplicando a primeira linha por -1 e somando à segunda e à terceira e, então, multiplicando a primeira linha por -4 e somando à quarta linha, não alteramos o valor do determinante:

$$\det A = \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 1 & 3 & 3 \\ 4 & 3 & 2 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & -1 & -2 & -3 \end{pmatrix}. \quad (1.2)$$

Continuando o escalonamento, obtemos (de acordo com as propriedades do determinante)

$$\det A = \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & -1 & -2 & -3 \end{pmatrix} = (2) \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & -2 \end{pmatrix}.$$

Então,

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

A última matriz é triangular superior, de modo que seu determinante é o produto de suas entradas na diagonal principal. Assim,

$$\det A = -2.$$

Observe, contudo, que também poderíamos obter o determinante da matriz A sem utilizar a operação fundamental (b). De (1.2) vem

$$\det A = \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & -1 & -2 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

o que implica que $\det A = -2$, pois a última matriz é triangular superior. \triangleleft

Exercício 1.27 Calcule o determinante da matriz

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

Algumas vezes, a divisão de uma matriz quadrada em *blocos* pode ser muito útil para o cálculo de seu determinante. Por exemplo, a matriz

$$Q = \begin{pmatrix} 0 & 1 & 2 & 1 & 2 & 3 & 3 & 3 & 1 & 0 \\ 1 & 1 & 2 & 2 & 2 & 2 & 3 & 3 & 1 & 1 \\ 1 & 2 & 2 & 1 & 3 & 1 & 3 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 & 3 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 & 2 & 2 & 7 & 1 & 1 \\ 0 & 0 & 0 & 0 & 3 & 4 & 3 & 4 & 1 & 4 \\ 0 & 0 & 0 & 0 & 3 & 1 & 4 & 1 & 5 & 6 \\ 0 & 0 & 0 & 0 & 6 & 3 & 4 & 1 & 2 & 1 \end{pmatrix}$$

pode ser escrita na forma

$$Q = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

em que

$$A = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \text{e} \quad D = \begin{pmatrix} 1 & 2 & 3 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 & 1 & 1 \\ 2 & 2 & 2 & 7 & 1 & 1 \\ 3 & 4 & 3 & 4 & 1 & 4 \\ 3 & 1 & 4 & 1 & 5 & 6 \\ 6 & 3 & 4 & 1 & 2 & 1 \end{pmatrix}.$$

Os blocos A e D ajudam no cálculo do determinante da matriz Q . De fato, vale o seguinte resultado:

Teorema 1.28 Seja Q uma matriz $n \times n$ com a forma

$$Q = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

em que A é uma matriz $m \times m$ e D uma matriz $(n-m) \times (n-m)$. Então,

$$\det Q = \det A \det D.$$

Demonstração: Utilizando as operações elementares (a) e (c), transformamos os blocos A e D em matrizes triangulares superiores. A operação elementar (c) não altera o determinante de Q (nem o de A ou D), mas a operação (a) inverte o sinal do determinante de Q . Mas, se ela for aplicada em uma linha da matriz A (ou da matriz Q), ela também altera o sinal do determinante de A (e respectivamente, de D). Assim, suponhamos que ao transformar A e D em matrizes triangulares superiores tenhamos feito j e k trocas de linhas, respectivamente. Teremos

$$(-1)^j (-1)^k \det Q = \begin{pmatrix} A' & B' \\ 0 & D' \end{pmatrix},$$

em que A' e D' são matrizes triangulares superiores. Como a matriz Q é triangular superior, seu determinante é o produto de suas entradas diagonais, de modo que

$$\begin{aligned} (-1)^j (-1)^k \det Q &= a'_{11} \cdots a'_{mm} d'_{11} \cdots d'_{(n-m)(n-m)} \\ &= \det A' \det D', \end{aligned}$$

em que a'_{11}, \dots, a'_{mm} são as entradas diagonais da matriz A' , enquanto $d'_{11}, \dots, d'_{(n-m)(n-m)}$ são as entradas diagonais de D' .

Mas $\det A = (-1)^j \det A'$ e $\det D = (-1)^k \det D'$, como vimos anteriormente. Substituindo na igualdade anterior, obtemos

$$\det Q = \det A \det D.$$

□

Exercício 1.29 Considere a matriz

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

(i) Utilizando operações elementares, leve essa matriz à forma

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

em que os blocos A e D são 2×2 .

(ii) Utilizando o item anterior, calcule seu determinante.

1.4 EXERCÍCIOS

1. Considere a matriz

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

Considere o sistema $Ax = b$, sendo b igual a

$$(a) \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}; \quad (b) \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}; \quad (c) \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}; \quad (d) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Encontre, em cada caso, todas as soluções do sistema $Ax = b$, se existirem. Em (a), imponha condições sobre b para que a solução exista.

2. Considere o sistema cuja matriz aumentada é

$$\left(\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array} \right).$$

Determine uma solução particular do sistema não homogêneo, bem como todas as soluções do sistema homogêneo associado. Escreva sua resposta como no Exemplo 1.11.

3. Utilizando o procedimento do Exemplo 1.26, calcule o determinante das matrizes

$$(a) \begin{pmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix}; \quad (b) \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ 5 & 9 & 1 & 6 \end{pmatrix}; \quad (c) \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 3 & 3 & 2 \end{pmatrix}.$$

AULA 2

O Espaço \mathbb{R}^n

OBJETIVOS

No final desta aula, o aluno deverá ser capaz de:

1. obter a representação paramétrica do plano;
2. reconhecer a importância de "sistemas de coordenadas" em casos simples;
3. efetuar operações básicas no \mathbb{R}^n .

2.1 EQUAÇÃO PARAMÉTRICA DO PLANO

Suponhamos que $n = (a, b, c)$ não seja o vetor nulo. Como sabemos, uma equação

$$ax + by + cz = d$$

representa um plano com vetor normal $n = (a, b, c)$. Se $d \neq 0$, então o plano não passa pela origem $(0, 0, 0)$ do sistema de coordenadas. Um ponto (x_0, y_0, z_0) do espaço pertence ao plano se, e somente se, $ax_0 + by_0 + cz_0 = d$.

Existe uma outra maneira de expressar a equação de um plano. Vejamos inicialmente em um exemplo:

Exemplo 2.1 Consideremos o plano π , de equação

$$x - 2y - 3z = 5.$$

Vamos proceder como no Exemplo 1.11. Para isso, levamos a forma matricial do sistema (com uma equação!) à forma escalonada reduzida por linhas e escrevemos as soluções do mesmo:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}. \quad (2.1)$$

Interpretando como no Teorema 1.21, o primeiro termo do lado direito da equação é uma solução particular do sistema: verifique que $(5, 0, 0)$ satisfaz a equação do plano. Os dois termos seguintes formam o conjunto de soluções do sistema homogêneo associado: $x - 2y - 3z = 0$.

Comecemos analisando esses dois termos, associados às variáveis livres λ e μ . Denotaremos $u = (2, 1, 0)$ e $v = (3, 0, 1)$. Esses vetores determinam 3 pontos: a origem, o ponto $P_1 = (2, 1, 0)$ e o ponto $P_2 = (3, 0, 1)$. Claramente esses pontos não estão alinhados, de modo que determinam um plano – justamente o plano de equação $x - 2y - 3z = 0$. Os vetores u e v determinam eixos coordenados, de modo que a origem e esses eixos formam um *sistema de coordenadas* nesse plano. (Na figura a seguir, esse plano é representado pelo plano do papel deste livro.) Consideremos um ponto $Q = (x, y, z)$ desse plano. Traçando retas paralelas aos eixos determinados por u e v e obtendo as intersecções dessas retas com os eixos coordenados, obtemos as coordenadas de Q nesse sistema de coordenadas.

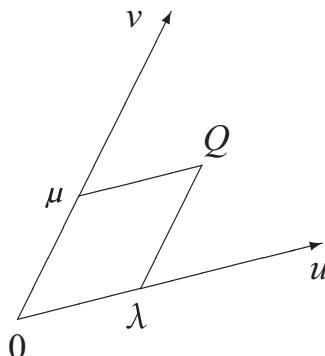


Figura 2.1: Os vetores u e v definem um sistema de coordenadas no plano por eles determinado

Acabamos de mostrar que todos os pontos daquele plano são descritos pelos parâmetros λ e μ . Um ponto $Q(x, y, z)$ só pertencerá ao plano $x + 2y + 3z = 0$ se for possível encontrar valores para λ e μ de modo que $\lambda u + \mu v$ nos dê as coordenadas de Q .

Quer dizer, o vetor $\vec{q} = (x, y, z)$ é combinação linear dos vetores \vec{u} e \vec{v} se, e somente se, o ponto $Q = (x, y, z)$ pertencer ao plano $x + 2y + 3z = 0$.

Note bem: apesar de todos os pontos do plano $x + 2y + 3z = 0$ serem pontos do espaço, conseguimos descrevê-los utilizando duas coordenadas: os valores de λ (que descreve a posição de um ponto do plano com respeito ao eixo gerado pelo vetor u) e μ (que descreve a posição de um ponto do plano com respeito ao eixo gerado pelo vetor v). A origem corresponde aos valores $(\lambda, \mu) = (0, 0)$. O valor $(\lambda, \mu) = (1, 0)$ corresponde ao ponto $P_1 = (2, 1, 0)$ do \mathbb{R}^3 . E assim por diante.

Agora consideremos todos os termos do lado direito da igualdade (2.1). Como vimos, os dois últimos criam um plano passando pela origem; o primeiro translada esse plano, de modo que ele passe pelo ponto $P_0 = (5, 0, 0)$. Veja a Figura 2.2.

A equação (2.1) é chamada equação paramétrica do plano π . □

Você deve se convencer que o exemplo anterior pode ser repetidos para qualquer plano $ax + by + cz = d$.

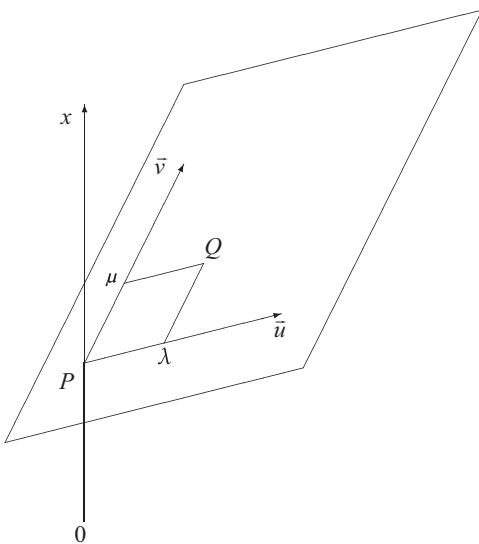


Figura 2.2: O ponto P translada o plano determinado pelos vetores u e v ; note a mudança de notação: aqui, \bar{u} e \bar{v}

Exercício 2.2 Considere o plano $x - y - z = 1$ e os vetores $u = (1, 1, 0)$ e $v = (1, 0, 1)$. Esboce o sistema de coordenadas no plano $x - y - z = 0$ determinado pelos vetores u e v . Ache valores para λ e μ de modo que $(2, 1, 1) = \lambda u + \mu v$. Esboce o plano $x - y - z = 1$.

Exercício 2.3 Considere o plano π de equação $ax + by + cz = d$. (Isso quer dizer que $(a, b, c) \neq (0, 0, 0)!$) Procedendo como no Exemplo 2.1, encontre vetores u e v e as equações paramétricas do plano π . Verifique que os vetores u e v encontrados são linearmente independentes.

2.2 SISTEMAS LINEARES EM TRÊS VARIÁVEIS

Comecemos com um exemplo simples:

Exemplo 2.4 Consideremos o sistema

$$\begin{aligned} x + 2y + z &= 3 \\ x + y + 3z &= 1 \end{aligned} \tag{2.2}$$

Aplicando o método de Gauss-Jordan, chegamos à forma escalonada reduzida por linhas

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 1 & 1 & 3 & 1 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 5 & -1 \\ 0 & 1 & -2 & 1 \end{array} \right).$$

Portanto, a solução do sistema é

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -5 \\ 2 \\ 1 \end{pmatrix}. \tag{2.3}$$

A equação (2.3) nos dá a equação da reta, interseção dos planos dados.

Consideremos, por outro lado, o sistema homogêneo associado a (2.2). Sua solução é dada por

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} -5 \\ 2 \\ 1 \end{pmatrix}.$$

Essa solução estabelece um *sistema de coordenadas* na reta $r(z) = z(-5, 2, 1)$. Para o valor $z = 0$, estamos na origem. Para $z = 1$, estamos no ponto $(-5, 2, 1)$ e assim por diante. Assim, todos os pontos dessa reta são descritos por apenas um valor do parâmetro z . Quer dizer, se estamos interessados apenas em pontos dessa reta, qualquer ponto (x, y, z) dela pode ser representado utilizando uma única coordenada: o valor de z . \triangleleft

Em geral, ao procurarmos soluções de um sistema

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1 \\ a_{21}x + a_{22}y + a_{23}z &= b_2, \end{aligned}$$

estamos verificando se esses dois planos¹ são paralelos ou, caso contrário, determinando a reta formada pela interseção de ambos.

Expressando o sistema anterior em forma matricial, obtemos

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Um ponto (x_0, y_0, z_0) do espaço pertence à interseção dos planos (se essa existir) se ele satisfizer a equação matricial

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \quad (2.4)$$

A forma matricial (2.4) tem muitos significados e consequências. Nesta seção abordaremos um deles, ao notar que em (2.4) o ponto (x_0, y_0, z_0) aparece na forma de uma matriz coluna. Ou seja, se A for uma matriz $m \times 3$ (em que m é arbitrário), na equação matricial

$$Ax = b,$$

x representa um ponto do espaço:

$$(x_0, y_0, z_0) \longleftrightarrow \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}.$$

Ora, sistemas $Ax = b$ acontecem com matrizes $m \times n$ arbitrárias. A interpretação

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \longleftrightarrow (x_1, \dots, x_n),$$

¹ Estamos supondo que $(a_{11}, a_{12}, a_{13}) \neq (0, 0, 0)$ e $(a_{21}, a_{22}, a_{23}) \neq (0, 0, 0)$.

em que (x_1, \dots, x_n) designa um ponto genérico de um espaço abstrato (que denotaremos por \mathbb{R}^n) nos permitirá, como veremos no decorrer deste

curso, interpretar geometricamente o sistema $Ax = b$, de maneira semelhante à que fizemos nos Exemplos 2.1 e 2.4.

Essa última frase parece muito pretensiosa: se $n > 3$, como interpretar geometricamente a solução de um sistema $m \times n$, se não podemos vislumbrar o espaço \mathbb{R}^n ? Bom, esse é um dos objetivos deste curso, de modo que não podemos justificar nossa pretensão neste momento; mas também não podemos deixar de destacar a semelhança entre as equações (1.1), (2.1) e (2.3)!

2.3 O ESPAÇO \mathbb{R}^n

Definimos o conjunto \mathbb{R}^n por

$$\mathbb{R}^n = \{x = (x_1, x_2, \dots, x_n) : x_i \in \mathbb{R}\}.$$

O número de coordenadas x_1, \dots, x_n dependerá do problema considerado. Os casos $n = 2$ e $n = 3$ correspondem a pontos do plano e do espaço, respectivamente. Ao mostrarmos que algum resultado é válido para o \mathbb{R}^n , esse resultado será verdadeiro para qualquer valor de $n = \{1, 2, \dots\}$.

Se x e y são pontos do \mathbb{R}^n e λ , um número real, definimos

$$x + y = (x_1 + y_1, \dots, x_n + y_n) \quad (2.5)$$

$$\lambda x = (\lambda x_1, \dots, \lambda x_n). \quad (2.6)$$

(Designamos por (y_1, \dots, y_n) as coordenadas de $y \in \mathbb{R}^n$.)

Observe que essas definições estão de acordo com os resultados mostrados na seção 1.1.

Teorema 2.5 Se $x, y, z \in \mathbb{R}^n$ e $\lambda, \mu \in \mathbb{R}$, as seguintes propriedades são satisfeitas:

- (i) $x + y \in \mathbb{R}^n$ (fechamento);
- (ii) $(x + y) + z = x + (y + z)$ (associatividade);
- (iii) $x + y = y + x$ (comutatividade);
- (iv) existe $0 \in \mathbb{R}^n$ tal que $x + 0 = x$ (elemento neutro);
- (v) existe $(-x) \in \mathbb{R}^n$ tal que $x + (-x) = 0$ (inverso aditivo);
- (vi) $\lambda x \in \mathbb{R}^n$ (fechamento);
- (vii) $\mu(\lambda x) = (\mu\lambda)x$ (associatividade);
- (viii) $\lambda(x + y) = \lambda x + \lambda y$ (distributividade);
- (ix) $(\lambda + \mu)x = \lambda x + \mu x$ (distributividade);
- (x) $1x = x$ (regra da unidade).

Demonstração: Mostraremos apenas algumas dessas propriedades. Vejamos a propriedade comutativa:

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

Note que a adição no \mathbb{R}^n é feita adicionando cada coordenada; como em cada coordenada temos uma adição de números reais, que é comutativa, chegamos ao resultado desejado.

Vejamos a propriedade (viii):

$$\begin{aligned} \lambda(x + y) &= \lambda(x_1 + y_1, \dots, x_n + y_n) \\ &= (\lambda(x_1 + y_1), \dots, \lambda(x_n + y_n)) \\ &= (\lambda x_1 + \lambda y_1, \dots, \lambda x_n + \lambda y_n) \\ &= (\lambda x_1, \dots, \lambda x_n) + (\lambda y_1, \dots, \lambda y_n) \\ &= \lambda x + \lambda y. \end{aligned}$$

□

Observe que 0 tem, nessa definição, dois significados distintos: descreve tanto o vetor $0 = (0, 0, \dots, 0) \in \mathbb{R}^n$ como o escalar $0 \in \mathbb{R}$.

Exercício 2.6 Mostre as demais propriedades do Teorema 2.5.

Exercício 2.7 Considere a igualdade $0x = 0$ para todo $x \in \mathbb{R}^n$. Interprete e prove essa igualdade.

Definição 2.8 Por satisfazer as propriedades (i) – (x), dizemos que \mathbb{R}^n é um **espaço vetorial**. Os elementos de um espaço vetorial são chamados **vetores**. Elementos $\lambda \in \mathbb{R}$ são chamados **escalares**. Se $x, y \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, a soma $x + y$ é a **soma** dos vetores x e y , enquanto λx é a **multiplicação** do escalar λ pelo vetor x . Também chamamos um vetor $x \in \mathbb{R}^n$ de um **ponto** do \mathbb{R}^n .

Sinteticamente, dizemos que \mathbb{R}^n possui uma soma (de vetores) e uma multiplicação por escalar. Utilizaremos corriqueiramente a identificação

$$(x_1, \dots, x_n) \in \mathbb{R}^n \leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

que permite compreender um ponto do \mathbb{R}^n como uma matriz coluna e que faz corresponder às operações entre matrizes colunas as operações definidas no \mathbb{R}^n . Podemos até mesmo levar mais adiante essa interpretação, identificando pontos do \mathbb{R}^n com colunas de uma matriz!

Exemplo 2.9 Consideremos o sistema homogêneo representado matricialmente por

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

A cada vetor $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ corresponde a matriz coluna

$$(x_1, x_2, x_3, x_4) \longleftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

e cada coluna de A pode ser vista como um vetor do \mathbb{R}^4 .

Levando a matriz A à forma escalonada reduzida por linhas, obtemos

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

de modo que sua solução é

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_4 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Quer dizer, as soluções do sistema são os múltiplos do vetor $u = (-1, 0, 0, 1) \in \mathbb{R}^4$. Lembrando que a equação de uma reta passando pela origem no \mathbb{R}^3 é dada pelos múltiplos de um vetor fixo, não parece natural dizer que a solução desse sistema é uma *reta* (passando pela origem) no espaço \mathbb{R}^4 ?

Note: se estivermos interessados em vetores do \mathbb{R}^4 que pertencem a essa *reta*, então cada vetor fica caracterizado por uma única coordenada! \triangleleft

Exercício 2.10 Como feito nos Exemplos 2.1 e 2.4, indique o sistema de coordenadas estabelecido ao se obter a solução do Exemplo 2.9.

Exercício 2.11 Encontre as soluções do sistema

$$\begin{array}{rcl} x_1 + 2x_2 + 3x_3 + x_4 & = & 8 \\ x_1 + 3x_2 + 0x_3 + x_4 & = & 7 \\ x_1 + 0x_2 + 2x_3 + x_4 & = & 3 \end{array}$$

Interprete geometricamente as soluções desse sistema.

2.4 ESPAÇOS VETORIAIS ABSTRATOS

Na seção anterior definimos \mathbb{R}^n e mostramos que ele satisfaz as propriedades descritas no Teorema 2.5. Com base nessas propriedades, podemos definir abstratamente um *espaço vetorial*.

Definição 2.12 Um **espaço vetorial real** X é um conjunto cujos elementos (chamados **vetores**) podem ser somados e multiplicados por **escalares**, isto é, números reais.² Se $x, y, z \in X$ e $\lambda, \mu \in \mathbb{R}$, as seguintes propriedades devem ser satisfeitas pela adição e multiplicação por escalar:

² Também podemos admitir números complexos, isto é, escalares $\lambda \in \mathbb{C}$. Neste caso, temos um *espaço vetorial complexo*.

- (i) $x + y \in X$ (fechamento);
- (ii) $(x + y) + z = x + (y + z)$ (associatividade);
- (iii) $x + y = y + x$ (comutatividade);
- (iv) existe $0 \in X$ tal que $x + 0 = x$ (elemento neutro);
- (v) existe $(-x) \in X$ tal que $x + (-x) = 0$ (inverso aditivo);
- (vi) $\lambda x \in X$ (fechamento);
- (vii) $\mu(\lambda x) = (\mu\lambda)x$ (associatividade);
- (viii) $\lambda(x + y) = \lambda x + \lambda y$ (distributividade);
- (ix) $(\lambda + \mu)x = \lambda x + \mu x$ (distributividade);
- (x) $1x = x$ (regra da unidade).

Neste texto não daremos muita ênfase a espaços vetoriais abstratos. Mas alguns exemplos são importantes:

Exemplo 2.13 Seja $\mathcal{P} = \{a_0 + a_1t + \dots + a_nt^n\}$ o conjunto de polinômios em t com coeficientes reais e grau menor que ou igual a n . Com a adição usual de polinômios e a multiplicação de polinômio por um número real, \mathcal{P} é um espaço vetorial. \square

Exemplo 2.14 Seja $\mathbb{M}_{m \times n}$ o conjunto de todas as matrizes $m \times n$ com entradas reais. Com a adição de matrizes e multiplicação de uma matriz por um número $\lambda \in \mathbb{R}$, $\mathbb{M}_{m \times n}$ é um espaço vetorial. \square

Exercício 2.15 Prove as afirmações dos Exemplos 2.13 e 2.14.

Observe que o significado das operações de adição e multiplicação foi diferente em cada exemplo. No caso de um espaço vetorial abstrato, não sabemos como as operações de adição e multiplicação por escalar são realizadas; apenas sabemos que elas satisfazem as propriedades (i) – (x) da Definição 2.12. Muitas vezes, para salientarmos esse fato, denotamos essas operações como \oplus e \odot , respectivamente. Assim, a propriedade (iii) pode ser descrita por $x \oplus y = y \oplus x$, enquanto (vii), $\mu \odot (\lambda \odot x) = (\mu\lambda) \odot x$. Note que $\mu\lambda$ indica a operação usual de multiplicação de números reais.

O significado abstrato das operações de adição de vetores e multiplicação de vetor por escalar torna difícil a verificação de algumas propriedades que são óbvias no caso do \mathbb{R}^n . Temos, por exemplo:

Proposição 2.16 *Existe um único elemento neutro em um espaço vetorial. Quer dizer, se $x + 0 = x = x + 0'$ para todo $x \in X$, então $0 = 0'$.*

Demonstração: Como $0 \in X$, escolhendo $x = 0$ na igualdade $x + 0' = x$, concluímos que $0 + 0' = 0$. Por outro lado, escolhendo $x = 0' \in X$ na igualdade $x + 0 = x$, concluímos $0' + 0 = 0'$. Como a adição de vetores é comutativa, temos $0 = 0 + 0' = 0'$, mostrando que $0 = 0'$. \square

Observe que não podemos falar em *coordenadas* de um vetor do espaço X ! Quer dizer, não existe nenhuma expressão análoga a $0 = (0, \dots, 0) \in \mathbb{R}^n$.

Exercício 2.17 Mostre que, em um espaço vetorial X , vale a *lei do cancelamento*: se $x + a = x + b$, então $a = b$. Conclua, então, que o elemento inverso aditivo $(-x)$ de um elemento $x \in X$ é único.

Observação 2.18 Comparando as seções 2.3 e 2.4 com a seção 1.1, não podemos ignorar uma diferença: na seção 1.1 foi utilizado o conceito de *norma* de um vetor, conceito esse que não foi definido no espaço \mathbb{R}^n ou em espaços vetoriais abstratos. Também não foi definido o *produto escalar* de dois vetores. Isso aconteceu por uma razão muito simples: este curso trata apenas das propriedades *algebricas* de espaços vetoriais; propriedades topológicas, isto é, propriedades relacionadas com os conceitos de *distância* e *ortogonalidade* serão tratadas no curso de Álgebra Linear II. \triangleleft

2.5 EXERCÍCIOS

1. Em cada caso, encontre a equação paramétrica do plano

$$(a) \quad x + 2y + z = 3; \quad (b) \quad x - y + 2z = 5; \quad (c) \quad x + y + z = 1.$$

2. Encontre as soluções do sistema

$$(a) \quad \begin{cases} x + y + z = 7 \\ x - y + 2z = 2; \end{cases} \quad (b) \quad \begin{cases} x - 5y + 3z = 0 \\ 2x + y - z = 1. \end{cases}$$

3. Resolva o sistema homogêneo $Ax = 0$, a matriz A sendo dada por

$$(a) \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{pmatrix}; \quad (b) \quad \begin{pmatrix} 1 & 7 & 0 & 0 & -8 & -3 \\ 0 & 0 & 1 & 0 & 6 & 5 \\ 0 & 0 & 0 & 1 & 3 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Em cada caso, interprete geometricamente a solução obtida.

4. Seja \mathfrak{P} o conjunto de todos os polinômios com coeficientes reais (com todos os graus possíveis). Mostre que \mathfrak{P} é um espaço vetorial com a adição de polinômios e a multiplicação de um polinômio por um escalar definidos da maneira usual.

AULA 3

Subespaços do \mathbb{R}^n e Bases

OBJETIVOS

No final desta aula, o aluno deverá ser capaz de:

1. verificar se um subconjunto é um subespaço;
2. verificar se um conjunto é linearmente independente;
3. verificar que um conjunto é uma base de um subespaço;
4. operar com o conceito de dimensão de um subespaço.

3.1 SUBESPAÇOS E COMBINAÇÕES LINEARES

Definição 3.1 Seja S um subconjunto qualquer do \mathbb{R}^n . Dizemos que S é um **subespaço** do \mathbb{R}^n se, para quaisquer $x, y \in S$ e, para qualquer $\lambda \in \mathbb{R}$, temos

- (i) $x + y \in S$;
- (ii) $\lambda x \in S$.

Uma vez que $0x = 0$ para qualquer $x \in \mathbb{R}^n$, vemos que $0 \in \mathbb{R}^n$ é um elemento de qualquer subespaço do \mathbb{R}^n .

Exemplo 3.2 Considere o subconjunto $S = \{0\} \subset \mathbb{R}^n$. Então, S é um subespaço de \mathbb{R}^n , pois $0 + 0 = 0 \in S$ e $\lambda 0 = 0 \in S$, para todo $\lambda \in \mathbb{R}$. Considere também o subconjunto $S' = \mathbb{R}^n$. Claramente S' é um subespaço do \mathbb{R}^n . Os subespaços S e S' são chamados *subespaços triviais* do \mathbb{R}^n . \triangleleft

Exemplo 3.3 Seja $0 \neq v \in \mathbb{R}^n$ um vetor fixo e $Y = \{tv : t \in \mathbb{R}\}$. (No caso de $n = 2$ ou $n = 3$, sabemos que Y descreve uma reta passando pela origem.) Então Y é um subespaço de \mathbb{R}^n . De fato, se $y_1, y_2 \in Y$, então $y_1 = t_1v$ e $y_2 = t_2v$, para certos escalares t_1, t_2 . Logo, $y_1 + y_2 = t_1v + t_2v = (t_1 + t_2)v$ é um elemento de Y . Também, se $\lambda \in \mathbb{R}$, então $\lambda y_1 = \lambda(t_1v) = (\lambda t_1)v \in Y$. Isso mostra o afirmado. Imitando os casos $n = 2$ e $n = 3$, dizemos que Y descreve uma *reta* passando pela origem no espaço \mathbb{R}^n . Note que o vetor v gera um sistema de coordenadas em que os pontos da reta Y são descritos por um único parâmetro: o valor do número real t . Compare com o Exemplo 2.4. \triangleleft

Exercício 3.4 Sejam u, v vetores do \mathbb{R}^n e $Z = \{\alpha u + \beta v : \alpha, \beta \in \mathbb{R}\}$. Mostre que Z é um subespaço do \mathbb{R}^n . Agora suponha que u e v não sejam colineares, quer dizer, um não é múltiplo do outro. Justifique a denominação: Z é um *plano* passando pela origem no \mathbb{R}^n . Verifique que os vetores u, v geram um sistema de coordenadas em que os pontos do plano Z são descritos por dois parâmetros: α e β . (Compare com o Exemplo 2.1.)

Observação 3.5 No Exercício 3.4, você pode achar estranha a exigência dos vetores u, v não serem colineares. Ora, caso contrário, teríamos $u = \mu v$, por exemplo. Mas então

$$Z = \{\alpha\mu v + \beta v\} = \{(\alpha\mu + \beta)v\}$$

descreve o conjunto dos múltiplos escalares do vetor v . De acordo com o Exemplo 3.3, esse conjunto é uma *reta* no \mathbb{R}^n e todos os elementos desse conjunto são múltiplos de v . \square

Exercício 3.6 A afirmação “ \mathbb{R}^2 não é um subespaço do \mathbb{R}^3 ” é verdadeira. Discuta essa afirmação.

Exercício 3.7 Seja S um subconjunto do \mathbb{R}^n . Mostre que S é um subespaço se, e somente se, $\lambda x + y \in S$ para quaisquer $x, y \in S$ e $\lambda \in \mathbb{R}$.

Exercício 3.8 Seja X um subespaço do \mathbb{R}^n . Mostre que X satisfaz todas as propriedades da Definição 2.12 sendo, portanto, um espaço vetorial.

O exercício anterior garante assim que, com as operações de adição e multiplicação por escalar restritas aos elementos de um subespaço $S \subset \mathbb{R}^n$, esse conjunto é um espaço vetorial.

No próximo resultado identificamos qualquer solução x de $Ax = 0$ com um ponto do \mathbb{R}^n :

Proposição 3.9 Seja A uma matriz $m \times n$. Então, o núcleo de A ,

$$\ker A = \{x \in \mathbb{R}^n : Ax = 0\}$$

é um subespaço do \mathbb{R}^n .

Demonstração: Sejam $x_1, x_2 \in \ker A$. Então $Ax_1 = 0$ e $Ax_2 = 0$, de modo que $A(x_1 + x_2) = Ax_1 + Ax_2 = 0 + 0 = 0$. Similarmente, se $\lambda \in \mathbb{R}$, então $A(\lambda x_1) = \lambda Ax_1 = \lambda 0 = 0$. \square

Examine novamente os Exemplos 1.11, 2.1 e 2.4, passando sempre ao sistema homogêneo associado. Em todos eles encontramos subespaços do \mathbb{R}^n (em que o valor de n depende do exemplo).

Exercício 3.10 Se $b \neq 0$, então as soluções do sistema $Ax = b$ não formam um subespaço do \mathbb{R}^n .

Seja S um subespaço do \mathbb{R}^n . Nossa objetivo nesta Aula pode ser descrito como a introdução de um sistema (linear) de coordenadas em S , de maneira semelhante ao que foi feito no Exemplo 3.3 e no Exercício 3.4. Mas nosso caminho até alcançar esse objetivo ainda é longo...

Em primeiro lugar, precisamos definir o que é um *sistema de coordenadas* (linear) ou, como os matemáticos preferem denominar, uma *base*. Esse conceito depende, em certa extensão, do problema considerado: no caso de um plano, um sistema de coordenadas utiliza dois eixos coordenados; no espaço, usa três eixos. Dizer que uma base é um sistema de eixos coordenados parece bom, mas preferimos usar, provisoriamente, uma linguagem mais pictórica, por acreditarmos ser ela mais elucidativa. Assim, definimos provisoriamente uma base como um conjunto \mathcal{B} de vetores, satisfazendo:

- (i) os elementos de \mathcal{B} fazem parte do problema considerado;
- (ii) o conjunto \mathcal{B} contém todas as informações relevantes para o problema considerado;
- (iii) o conjunto \mathcal{B} não contém informações supérfluas.

No Exemplo 3.3, uma base seria constituída por um único vetor: o vetor v . Qualquer ponto da reta pode ser descrito utilizando apenas uma coordenada, de modo a obter-se o múltiplo adequado do vetor v . No Exercício 3.4, se fosse $u = \mu v$, então o vetor u seria supérfluo: todos os elementos do conjunto seriam múltiplos do vetor v e estaríamos na situação do Exemplo 3.3. Se u e v não forem colineares, uma base é constituída pelos vetores u e v . Qualquer ponto do plano é descrito por duas coordenadas (os valores de α e β) e, para se obter todos os pontos do plano, não podemos usar um número menor de coordenadas.

Existem várias questões que precisam ser tratadas para chegarmos ao conceito de base. É claro, precisamos dizer o que significa um conjunto ter todas as informações relevantes (para o problema considerado) e não ter informações supérfluas. Definidos esses conceitos, várias questões práticas passam a ser pertinentes: como obter um conjunto que possui todas as informações relevantes sobre o problema considerado? Como verificar se esse conjunto possui informações supérfluas? Como retirar do conjunto as informações supérfluas? Essas perguntas serão tratadas nesta Aula.

Começamos generalizando a construção feita no Exemplo 3.3 e no Exercício 3.4, agora utilizando qualquer número de vetores no \mathbb{R}^n :

Proposição 3.11 *Sejam v_1, v_2, \dots, v_k vetores quaisquer do \mathbb{R}^n . Então,*

$$\langle v_1, v_2, \dots, v_k \rangle = \{\alpha_1 v_1 + \dots + \alpha_k v_k : \alpha_1, \dots, \alpha_k \in \mathbb{R}\}$$

é um subespaço do \mathbb{R}^n , chamado subespaço gerado pelos vetores v_1, \dots, v_k .

Demonstração: Sejam $x, y \in \langle v_1, \dots, v_k \rangle$. Então, existem escalares $\alpha_1, \dots, \alpha_k$ e β_1, \dots, β_k tais que

$$x = \alpha_1 v_1 + \dots + \alpha_k v_k \quad \text{e} \quad y = \beta_1 v_1 + \dots + \beta_k v_k.$$

Consequentemente,

$$x + y = (\alpha_1 + \beta_1) v_1 + \dots + (\alpha_k + \beta_k) v_k$$

e

$$\lambda x = (\lambda \alpha_1) v_1 + \dots + (\lambda \alpha_k) v_k$$

são elementos de $\langle v_1, \dots, v_k \rangle$, provando o afirmado. \square

Exemplo 3.12 Sejam $v_1 = (1, 2, 1, 1)$, $v_2 = (2, 1, 1, 1)$, $v_3 = (1, 1, 1, 1)$, $v_4 = (0, 1, 0, 1)$ e $v_5 = (0, 0, 1, 0)$ vetores do espaço \mathbb{R}^4 . Vamos descrever o espaço $\langle v_1, \dots, v_5 \rangle$.

Temos, por definição,

$$\langle v_1, \dots, v_5 \rangle = \{\alpha v_1 + \beta v_2 + \gamma v_3 + \delta v_4 + \epsilon v_5\},$$

em que $\alpha, \beta, \gamma, \delta$ e ϵ são escalares. Assim, 0,5cm - 4cm

$$\begin{aligned}\langle v_1, \dots, v_5 \rangle &= \{(\alpha, 2\alpha, \alpha, \alpha) + (2\beta, \beta, \beta, \beta) + (\gamma, \gamma, \gamma, \gamma) + (0, \delta, 0, \delta) \\ &\quad + (0, 0, \epsilon, 0)\} \\ &= \{(\alpha + 2\beta + \gamma, 2\alpha + \beta + \gamma + \delta, \alpha + \beta + \gamma + \epsilon, \alpha + \beta \\ &\quad + \gamma + \delta)\}.\end{aligned}$$

Em particular, se $\alpha = 1$, $\beta = 0$, $\gamma = 0$, $\delta = 1$ e $\epsilon = 2$, temos que o vetor $(1, 3, 3, 2)$ pertence a $\langle v_1, \dots, v_5 \rangle$. \triangleleft

Não podemos denominar o espaço $\langle v_1, \dots, v_k \rangle$ de acordo com o número de elementos k utilizados na definição desse espaço. Por exemplo, se $k = 2$, $\langle v_1, v_2 \rangle$ pode não ser um plano, como vimos na Observação 3.5.

Exercício 3.13 Sejam $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$ e $v_3 = (1, 1, 0)$. Descreva o subespaço $\langle v_1, v_2, v_3 \rangle$.

Definição 3.14 Um vetor v é **combinação linear** dos vetores v_1, \dots, v_k se existem escalares $\alpha_1, \dots, \alpha_k$ tais que

$$\alpha_1 v_1 + \dots + \alpha_k v_k = v. \quad (3.1)$$

Quer dizer, $v \in \langle v_1, \dots, v_k \rangle$ é o mesmo que v ser combinação linear dos vetores v_1, \dots, v_k .

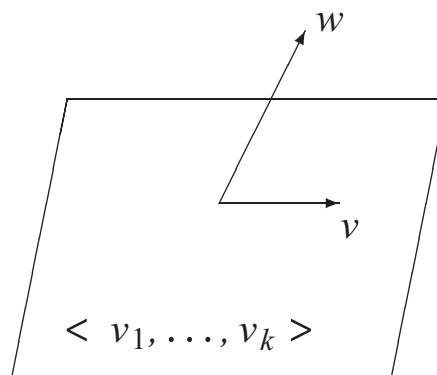


Figura 3.1: Se representarmos o subespaço $\langle v_1, \dots, v_k \rangle \subset \mathbb{R}^n$ como um plano, então $v \in \langle v_1, \dots, v_k \rangle$, enquanto $w \notin \langle v_1, \dots, v_k \rangle$

A equação vetorial (3.1) dá origem a um sistema não homogêneo, se $v \neq 0$. É o que veremos no próximo exemplo.

Exemplo 3.15 O vetor $(5, 2, 3, -1)$ é combinação linear dos vetores v_1, \dots, v_5 do Exemplo 3.12? Essa pergunta terá uma resposta afirmativa se existirem escalares $\alpha, \beta, \gamma, \delta$ e ϵ tais que

$$\alpha(1, 2, 1, 1) + \beta(2, 1, 1, 1) + \gamma(1, 1, 1, 1) + \delta(0, 1, 0, 1) + \epsilon(0, 0, 1, 0) = (5, 2, 3, -1)$$

ou seja, se

$$(\alpha + 2\beta + \gamma, 2\alpha + \beta + \gamma + \delta, \alpha + \beta + \gamma + \epsilon, \alpha + \beta + \gamma + \delta) = (5, 2, 3, -1).$$

A última igualdade dá origem ao sistema não homogêneo

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \epsilon \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 3 \\ -1 \end{pmatrix}.$$

Note que a matriz é formada tendo justamente os vetores v_1, \dots, v_5 como colunas. Escalonando a matriz aumentada do sistema, obtemos

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 0 & 0 & 5 \\ 2 & 1 & 1 & 1 & 0 & 2 \\ 1 & 1 & 1 & 0 & 1 & 3 \\ 1 & 1 & 1 & 1 & 0 & -1 \end{array} \right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 11 \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 0 & 2 & -10 \\ 0 & 0 & 0 & 1 & -1 & -4 \end{array} \right).$$

Assim, sua solução é dada por

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \epsilon \end{pmatrix} = \begin{pmatrix} 11 \\ 2 + \epsilon \\ -10 - 2\epsilon \\ -4 + \epsilon \\ \epsilon \end{pmatrix} = \begin{pmatrix} 11 \\ 2 \\ -10 \\ -4 \\ 0 \end{pmatrix} + \epsilon \begin{pmatrix} 0 \\ 1 \\ -2 \\ 1 \\ 1 \end{pmatrix}.$$

Como o sistema tem solução, vemos que o vetor $v = (5, 2, 3, -1)$ pertence ao espaço $\langle v_1, v_2, v_3, v_4, v_5 \rangle$. \triangleleft

Exercício 3.16

Sejam

$$v_1 = (2, 3, 1, 1), \quad v_2 = (-2, -2, 2, 1), \quad \text{e} \quad v_3 = (1, 0, -2, -1).$$

O vetor $v = (1, 1, 1, 1)$ pertence a $\langle v_1, v_2, v_3 \rangle$?

Agora estamos em condições de dizer quando um conjunto \mathcal{B} contém todas as informações relevantes sobre o subespaço S .

Definição 3.17 Um conjunto $\mathcal{B} = \{v_1, \dots, v_k\} \subset S$ é um **conjunto gerador** do subespaço $S \subset \mathbb{R}^n$ se todo elemento de S for combinação linear dos elementos de \mathcal{B} . Também dizemos que \mathcal{B} **gera** S .

Assim, se o problema que estivermos considerando tratar do subespaço S , estamos garantindo que \mathcal{B} tem todas as informações relevantes sobre S . Note que \mathcal{B} possui um número finito de elementos; se $S \neq \{0\}$, então S possui infinitos elementos.

Exercício 3.18 Mostre que $\{v_1, \dots, v_j\} \subset \mathbb{R}^n$ gera o subespaço $\langle v_1, \dots, v_j \rangle$

Como vimos no Exemplo 3.15, a equação (3.1) dá origem a um sistema não homogêneo, se $v \neq 0$. Como sabemos, o estudo de um sistema não homogêneo está intrinsecamente relacionado com o sistema homogêneo que lhe é associado. Esse é o objeto da próxima definição.

Definição 3.19 Dizemos que os vetores v_1, \dots, v_k do \mathbb{R}^n são **linearmente independentes** se

$$\alpha_1 v_1 + \dots + \alpha_k v_k = 0 \quad (3.2)$$

só tem a solução trivial $\alpha_1 = \dots = \alpha_k = 0$. Caso contrário, dizemos que os vetores v_1, \dots, v_k são **linearmente dependentes**.

Um conjunto $\mathcal{B} = \{v_1, \dots, v_k\}$ é linearmente independente, se os vetores v_1, \dots, v_k forem linearmente independentes; caso contrário, \mathcal{B} é linearmente dependente.

A equação (3.2) dá origem ao sistema homogêneo associado a (3.1). Assim, para verificar se o conjunto $\{v_1, \dots, v_k\}$ é linearmente independente, formamos a matriz $A = (v_1 \dots v_k)$, que tem os vetores v_1, \dots, v_k como colunas, e consideramos o sistema $Ax = 0$. Se esse sistema tiver apenas a solução trivial $x = 0$, então $\{v_1, \dots, v_k\}$ é linearmente independente; caso contrário, é linearmente dependente.

Consideraremos esse sistema em uma situação particularmente simples, que melhor nos fará entender a Definição 3.19:

Exemplo 3.20 Consideremos os vetores $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ e $v = (1, 1, 0)$. Queremos verificar se esses vetores são linearmente independentes ou linearmente dependentes. Para isso, consideraremos a igualdade vetorial

$$\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 v = 0. \quad (3.3)$$

Como já vimos, essa igualdade é o mesmo que o sistema homogêneo

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

(Repetimos: os vetores e_1 , e_2 e v constituem as *colunas* da matriz 3×3 , que denotaremos por A .)

A solução do sistema homogêneo é dada por

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \alpha_3 \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

Como o sistema homogêneo possui solução não trivial, concluímos que os vetores e_1 , e_2 e v são linearmente dependentes.

Substituindo a solução $(\alpha_1, \alpha_2, \alpha_3) = (-1, -1, 1)$ em (3.3), obtemos

$$-e_1 - e_2 + v = 0,$$

ou, o que é o mesmo,

$$v = e_1 + e_2. \quad (3.4)$$

Essa equação garante que v é combinação linear dos vetores e_1 e e_2 . Geometricamente, isso significa que v pertence ao plano π gerado pelos vetores e_1 e e_2 . Quer dizer, se tivermos os vetores e_1 e e_2 , então duas coordenadas bastarão para descrever todos os pontos do plano gerado por esses vetores, e o vetor v não é necessário. O vetor v é uma informação supérflua.

Compare esse exemplo com o Exercício 3.13. \triangleleft

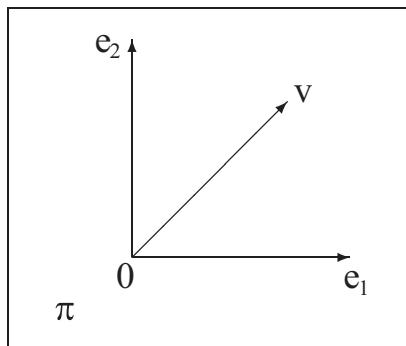


Figura 3.2: O vetor v é combinação linear dos vetores e_1 e e_2 , pois pertence ao plano gerado por estes vetores

Exercício 3.21

Verifique se os vetores

$$v_1 = (1, 1, 1), \quad v_2 = (1, 1, 0) \quad \text{e} \quad v_3 = (1, 0, 0)$$

são linearmente dependentes.

É claro que, no Exemplo 3.20, também podemos escrever e_1 (ou e_2) como combinação linear dos vetores restantes. Nesse caso, e_1 (ou, respectivamente, e_2) seria a informação supérflua.

Podemos formular abstratamente o que aconteceu no exemplo anterior. Veja que a equação (3.2) é tratada sem considerar as coordenadas dos vetores envolvidos!

Proposição 3.22 O conjunto $\{v_1, \dots, v_k\}$ é linearmente dependente se, e somente se, algum desses vetores é combinação linear dos vetores restantes.

Demonstração: Suponhamos que $\{v_1, \dots, v_k\}$ seja linearmente dependente. Então existem escalares $\alpha_1, \dots, \alpha_k$, nem todos nulos, tais que

$$\alpha_1 v_1 + \dots + \alpha_k v_k = 0.$$

Para simplificar a notação, vamos supor que $\alpha_1 \neq 0$. Nesse caso, temos

$$\alpha_1 v_1 = -\alpha_2 v_2 - \dots - \alpha_k v_k, \tag{3.5}$$

ou seja,

$$v_1 = \beta_2 v_2 + \dots + \beta_k v_k,$$

em que $\beta_i = \alpha_i / \alpha_1$ para $i = 2, \dots, k$. Assim, v_1 é combinação linear dos vetores $\{v_2, \dots, v_k\}$.

Reciprocamente, se (por exemplo) $v_1 = \alpha_2 v_2 + \dots + \alpha_k v_k$, então

$$1v_1 - \alpha_2 v_2 - \dots - \alpha_k v_k = 0,$$

e ao menos um dos escalares (qual seja, o coeficiente de v_1) é não nulo, mostrando que esse conjunto é linearmente dependente. \square

Note que, se tivermos $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$, podemos escrever qualquer vetor relacionado a um coeficiente $\alpha_i \neq 0$ como combinação linear dos vetores restantes: na demonstração anterior, esse fato é usado ao dividirmos a equação (3.5) pelo escalar $\alpha_1 \neq 0$.

Observação 3.23 Retirado um vetor “supérfluo” de um conjunto linearmente dependente, não podemos garantir que o conjunto restante seja formado apenas por vetores essenciais. Isto é, pode ser que o conjunto restante ainda seja linearmente dependente. Veremos, posteriormente, um método para retirar de uma vez todos os vetores supérfluos de um conjunto linearmente dependente. Veja a Observação 4.23. \triangleleft

Exercício 3.24 Verifique se o conjunto $\{(1, 1, 2, 1), (1, 1, 1, 1), (2, 1, 1, 1), (2, 1, 2, 1)\}$ é linearmente dependente ou linearmente independente. Se for linearmente dependente, escreva um dos vetores como combinação linear dos vetores restantes.

Exercício 3.25 Suponha que o vetor v pertença ao espaço $\langle v_1, \dots, v_k \rangle$. Mostre que o conjunto $\{v, v_1, \dots, v_k\}$ é linearmente dependente.

Exemplo 3.26 Sejam v_2, \dots, v_k vetores quaisquer do \mathbb{R}^n . Então o conjunto $\{0, v_2, \dots, v_k\}$ é linearmente dependente. (Aqui, 0 denota o vetor nulo.) Assim, qualquer conjunto que contenha o vetor nulo é linearmente dependente.

De fato, temos que

$$\alpha_1 0 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0$$

possui a solução não trivial $\alpha_1 = 1, \alpha_2 = \dots = \alpha_n = 0$. Você é capaz de exibir outras soluções? \triangleleft

Exercício 3.27 Suponha que o conjunto $\{v_1, \dots, v_k\} \subset \mathbb{R}^n$ seja linearmente dependente. Mostre que $\{v_1, \dots, v_k, v\} \subset \mathbb{R}^n$ é linearmente dependente, qualquer que seja o vetor $v \in \mathbb{R}^n$.

Observação 3.28 Pela Proposição 3.22, se $v \in \mathbb{R}^n$ não for combinação linear dos vetores v_1, \dots, v_k , então $v \notin \langle v_1, \dots, v_k \rangle$. Com a linguagem pictórica introduzida anteriormente, isso quer dizer que o vetor v traz uma informação que não está contida no subespaço $\langle v_1, \dots, v_k \rangle$.

Mas não podemos garantir que o conjunto $\{v, v_1, \dots, v_k\} \subset \mathbb{R}^n$ seja linearmente independente, pois não sabemos se seu subconjunto $\{v_1, \dots, v_k\}$ é linearmente independente, conforme o Exercício 3.27.

Se $\{v_1, \dots, v_k\}$ for linearmente independente e se v não pertencer a esse subespaço, então $\{v_1, \dots, v_k, v\}$ é linearmente independente. Esse resultado é uma consequência imediata da Proposição 3.22, mas vamos dar uma demonstração direta dele: suponhamos que

$$\alpha_1 v_1 + \dots + \alpha_k v_k + \beta v = 0. \quad (3.6)$$

Então $\beta = 0$ pois, caso contrário, v seria combinação linear dos vetores v_1, \dots, v_k . Mas então temos

$$\alpha_1 v_1 + \dots + \alpha_k v_k = 0$$

e, como esses vetores são linearmente independentes, $\alpha_i = 0$ para $i \in \{1, \dots, k\}$. Assim, todos os escalares em (3.6) são nulos, mostrando nossa afirmação. \triangleleft

Exercício 3.29 Seja $v \neq 0$ um vetor do \mathbb{R}^n . Mostre que o conjunto $\{v\}$ é linearmente independente.

3.2 BASES

Definição 3.30 Uma **base** $\mathcal{B} = \{v_1, \dots, v_k\}$ de um subespaço S do \mathbb{R}^n é um conjunto *ordenado* de vetores de S que gera esse subespaço e é linearmente independente.

Por conjunto *ordenado* queremos dizer que a ordem dos elementos é importante. Em outras palavras, os conjuntos ordenados $\{u, v, w\}$, $\{u, w, v\}$ e $\{w, u, v\}$ são distintos! Posteriormente explicaremos a razão de definirmos uma base como um conjunto ordenado. (Veja a Observação 6.3.)

Assim, para verificar que um conjunto \mathcal{B} é uma base do subespaço S , precisamos verificar três afirmações independentes:

- (i) os elementos de \mathcal{B} pertencem a S ;
- (ii) todo elemento de S é combinação linear dos elementos de \mathcal{B} ;
- (iii) o conjunto \mathcal{B} é linearmente independente.

Essas três exigências são a expressão, em termos matemáticos, das condições (i), (ii) e (iii) apresentadas na página 43.

Exemplo 3.31 Os vetores $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, ..., $e_n = (0, 0, \dots, 0, 1)$ formam uma base do espaço \mathbb{R}^n . De fato, se $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, temos

$$\begin{aligned} x = (x_1, \dots, x_n) &= x_1(1, 0, \dots, 0) + x_2(0, 1, 0, \dots, 0) \\ &\quad + \dots + x_n(0, \dots, 0, 1) \\ &= x_1 e_1 + x_2 e_2 + \dots + x_n e_n, \end{aligned}$$

mostrando que $\{e_1, \dots, e_n\}$ gera o \mathbb{R}^n . Além disso,

$$\begin{aligned} 0 &= \alpha_1 e_1 + \dots + \alpha_n e_n \\ &= (\alpha_1, 0, \dots, 0) + (0, \alpha_2, 0, \dots, 0) + \dots + (0, \dots, 0, \alpha_n) \\ &= (\alpha_1, \dots, \alpha_n), \end{aligned}$$

o que implica que $\alpha_1 = \dots = \alpha_n = 0$.

A base $\mathcal{B} = \{e_1, \dots, e_n\}$ é chamada *base canônica* do \mathbb{R}^n . \triangleleft

Exercício 3.32 Verifique se $\{v_1, v_2, v_3\} \subset \mathbb{R}^n$ é uma base do \mathbb{R}^3 , sendo

$$v_1 = (1, 3, 3), \quad v_2 = (2, 1, 1) \quad \text{e} \quad v_3 = (1, 1, 1).$$

A base canônica introduz no \mathbb{R}^n um sistema de coordenadas semelhante ao sistema com eixos x , y e z do \mathbb{R}^3 . Assim, a base canônica produz um sistema de coordenadas completamente natural. Para que estudar outras bases no \mathbb{R}^n ? A resposta é simples: a base canônica produz um sistema de coordenadas que pode não ser o mais adequado ao problema que estamos tratando. Com um sistema de eixos mais adequado, a obtenção da resposta para o nosso problema pode ser bem mais simples.

Além disso, muitas vezes estamos interessados em um subespaço particular do \mathbb{R}^n , subespaço esse que pode funcionar como um plano ou uma reta, por exemplo. Pode acontecer que nenhum (!) dos vetores da base canônica pertença a esse subespaço. Além disso, os pontos desse subespaço podem ser caracterizados por um número menor de coordenadas do que as n coordenadas utilizadas para caracterizar um ponto do \mathbb{R}^n : $x = (x_1, \dots, x_n)$. Essa situação ocorreu quando consideramos sistemas de coordenadas em retas e planos do \mathbb{R}^n , como no Exemplo 3.3 e Exercício 3.4.

Exemplo 3.33 O conjunto $\{v\}$ é uma base do subespaço

$$Y = \{tv : t \in \mathbb{R}\}$$

do Exemplo 3.3. De fato, esse conjunto é linearmente independente, de acordo com o Exercício 3.29. Além disso, qualquer elemento de Y é múltiplo de v , o que mostra que $\{v\}$ gera o subespaço Y . \triangleleft

Exemplo 3.34 Seja $\mathcal{B} = \{v_1, \dots, v_j\}$ um conjunto linearmente independente de vetores do \mathbb{R}^n . Então o subespaço $\langle v_1, \dots, v_j \rangle \subset \mathbb{R}^n$ tem \mathcal{B} como base. De fato, todo elemento de $\langle v_1, \dots, v_j \rangle$ é uma combinação linear de elementos de \mathcal{B} ; como esse conjunto é linearmente independente e está contido em $\langle v_1, \dots, v_j \rangle$, nossa afirmação está provada. \triangleleft

Exercício 3.35 Se u e v não forem colineares, mostre que $\{u, v\}$ é uma base do subespaço Z do Exercício 3.4.

Exercício 3.36 Mostre que o conjunto $\{(1, 1, 1, 1), (1, 1, 1, 0), (1, 1, 0, 0), (1, 0, 0, 0)\}$ é uma base do espaço \mathbb{R}^4 .

Proposição 3.37 Se $\mathcal{B} = \{v_1, \dots, v_k\}$ for a base de um subespaço $V \subset \mathbb{R}^n$, então cada vetor $v \in V$ é escrito de maneira única como combinação linear dos elementos de \mathcal{B} .

Demonstração: Suponhamos a existência de $v \in V$ tal que

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = v = \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_k v_k.$$

Queremos mostrar que $\alpha_i = \beta_i$ para $i \in \{1, \dots, k\}$. Ora, da igualdade anterior deduzimos que

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \dots + (\alpha_k - \beta_k)v_k = 0.$$

Como \mathcal{B} é linearmente independente, necessariamente temos $\alpha_i - \beta_i = 0$ para todo $i \in \{1, \dots, k\}$, o que prova nosso resultado. \square

3.3 DIMENSÃO

Agora passamos a considerar o conceito de dimensão de um subespaço; esse conceito pode ser utilizado para caracterizar todos os subespaços do \mathbb{R}^n .

Exemplo 3.38 Consideremos novamente o Exercício 3.4 no caso em que $u = (1, 0, 0)$ e $v = (0, 1, 0)$ são vetores do \mathbb{R}^3 . De acordo com o Exercício 3.35, $\{u, v\}$ é uma base do subespaço

$$Z = \{\alpha u + \beta v : \alpha, \beta \in \mathbb{R}\}.$$

Afirmamos que $\{u, u + v\}$ é outra base de Z . Seja $w = u + v = (1, 1, 0)$. Queremos mostrar que $\{u, w\}$ também é base de Z .

Temos que $w \in Z$, pois $w = 1u + 1v$. Assim, $\{u, w\} \subset Z$. Se $x \in Z$, então,

$$x = \alpha u + \beta v = \alpha(1, 0, 0) + \beta(0, 1, 0) = (\alpha, \beta, 0).$$

Então x é combinação linear de u e w . De fato,

$$(\alpha, \beta, 0) = \lambda_1 u + \lambda_2 w \Leftrightarrow (\alpha, \beta, 0) = (\lambda_1 + \lambda_2, \lambda_2, 0).$$

É claro então que $\lambda_2 = \beta$ e $\lambda_1 + \lambda_2 = \alpha$, o que implica $\lambda_1 = \alpha - \lambda_2 = \alpha - \beta$. Mostramos assim que qualquer elemento de Z é combinação linear dos elementos u e w .

Mas esses elementos também são linearmente independentes, pois

$$(0, 0, 0) = \lambda_1 u + \lambda_2 w \Leftrightarrow (0, 0, 0) = (\lambda_1 + \lambda_2, \lambda_2, 0).$$

Assim, $\lambda_1 = \lambda_2 = 0$ é a única solução de $0 = \lambda_1 u + \lambda_2 w$. \square

O Exemplo mostra que o mesmo subespaço do \mathbb{R}^n tem diferentes bases. Todas essas bases possuem algo em comum: o número de elementos. Utilizaremos, para mostrar esse fato, o seguinte resultado:

Teorema 3.39 Seja $\mathcal{B} = \{v_1, \dots, v_k\}$ uma base do subespaço $S \subset \mathbb{R}^n$. Então qualquer conjunto $\mathcal{C} \subset S$ com mais que k elementos é linearmente dependente.

Demonstração: Mostraremos esse resultado em uma situação particular, com uma notação mais simples. Para isso, suponhamos que $\mathcal{B} = \{v_1, v_2\}$ seja uma base de um subespaço $Y \subset \mathbb{R}^n$. Consideremos um conjunto $\mathcal{C} = \{u_1, u_2, u_3\} \subset Y$. Vamos mostrar que \mathcal{C} é linearmente dependente.

Para isso, consideremos a equação

$$\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 = 0. \quad (3.7)$$

Como \mathcal{B} é uma base de Y , cada elemento de \mathcal{C} é combinação linear dos elementos em \mathcal{B} . Assim,

$$\begin{aligned} u_1 &= \alpha_{11} v_1 + \alpha_{21} v_2 \\ u_2 &= \alpha_{12} v_1 + \alpha_{22} v_2 \\ u_3 &= \alpha_{13} v_1 + \alpha_{23} v_2 \end{aligned}$$

para determinados escalares α_{ij} , com $i \in \{1, 2\}$ e $j \in \{1, 2, 3\}$. Note bem: se conhecêssemos as coordenadas dos vetores envolvidos, poderíamos

determinar esses escalares. Mas, como os vetores v_1 e v_2 são arbitrários, só sabemos que esses escalares existem!

Substituindo essas expressões de u_1 , u_2 e u_3 em (3.7), obtemos

$$\lambda_1(\alpha_{11}v_1 + \alpha_{21}v_2) + \lambda_2(\alpha_{12}v_1 + \alpha_{22}v_2) + \lambda_3(\alpha_{13}v_1 + \alpha_{23}v_2) = 0,$$

ou seja,

$$(\lambda_1\alpha_{11} + \lambda_2\alpha_{12} + \lambda_3\alpha_{13})v_1 + (\lambda_1\alpha_{21} + \lambda_2\alpha_{22} + \lambda_3\alpha_{23})v_2 = 0.$$

Como \mathcal{B} é um conjunto linearmente independente, devemos ter

$$\begin{aligned}\lambda_1\alpha_{11} + \lambda_2\alpha_{12} + \lambda_3\alpha_{13} &= 0 \\ \lambda_1\alpha_{21} + \lambda_2\alpha_{22} + \lambda_3\alpha_{23} &= 0,\end{aligned}$$

isto é,

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Como o número de equações nesse sistema homogêneo é menor do que o número de incógnitas, o Teorema 1.19 garante que existe solução não trivial para o sistema, o que prova que \mathcal{C} é linearmente dependente. \square

Exercício 3.40 Diga se os conjuntos são linearmente dependentes ou linearmente independentes:

- (a) $S = \{(1,1), (2,1), (1,0)\} \subset \mathbb{R}^2$;
- (b) $R = \{(1,1,1,1), (0,0,0,0), (1,2,1,1)\} \subset \mathbb{R}^4$;
- (c) $P = \{(1,1,2), (1,0,1), (1,2,1), (0,0,1)\} \subset \mathbb{R}^3$.

Exercício 3.41 Se

$$\mathcal{B} = \{v_1, \dots, v_k\} \quad \text{e} \quad \mathcal{C} = \{u_1, \dots, u_k, u_{k+1}\},$$

demonstre o Teorema 3.39. Deduza daí então o caso $\mathcal{C} = \{u_1, \dots, u_k, u_{k+1}, \dots, u_{k+r}\}$, em que $r \geq 1$.

Corolário 3.42 Todas as bases de um subespaço $S \subset \mathbb{R}^n$ possuem o mesmo número de elementos.

Demonstração: Suponhamos que existam bases \mathcal{B} e \mathcal{B}' , com \mathcal{B} tendo menos elementos do que \mathcal{B}' . De acordo com o Teorema 3.39, o conjunto \mathcal{B}' seria linearmente dependente. Mas isso é impossível, pois \mathcal{B}' é uma base de S e, consequentemente, um conjunto linearmente independente. Chegamos a uma contradição que mostra que todas as bases devem ter o mesmo número de elementos. \square

Mas qualquer subespaço S do \mathbb{R}^n possui uma base? Note que já provamos que \mathbb{R}^n possui uma base. Assim, todos os elementos do \mathbb{R}^n (e, em particular, os elementos de S) podem ser escritos como combinação linear dos elementos da base do \mathbb{R}^n . Além disso, essa base é um conjunto linearmente independente. Isso não prova que S possui uma base?

Não! Há uma exigência que precisa ser cumprida: os elementos da base devem pertencer todos ao subespaço S .

Teorema 3.43 Seja $\mathcal{C} = \{u_1, \dots, u_j\} \subset S$ um conjunto linearmente independente. Então existem vetores u_{j+1}, \dots, u_{j+k} em S de modo que $\{u_1, \dots, u_j, u_{j+1}, \dots, u_{j+k}\}$ seja uma base de S .

Demonstração: Como S é um subespaço e $\mathcal{C} \subset S$, toda combinação linear de elementos de \mathcal{C} pertence a S . Assim,

$$\langle u_1, \dots, u_j \rangle \subset S.$$

Se $\langle u_1, \dots, u_j \rangle = S$, então \mathcal{C} é uma base de S , de acordo com o Exemplo 3.34. Caso contrário, existiria um vetor $u_{j+1} \in S$ tal que $u_{j+1} \notin \langle u_1, \dots, u_j \rangle$. De acordo com a Observação 3.28, o conjunto $\{u_1, \dots, u_{j+1}\}$ é linearmente independente.

Agora repetimos o raciocínio. Se $\langle u_1, \dots, u_j, u_{j+1} \rangle = S$, então o conjunto $\{u_1, \dots, u_{j+1}\}$ é uma base de S . Caso contrário, existiria $u_{j+2} \in S$ tal que $u_{j+2} \notin \langle u_1, \dots, u_j, u_{j+1} \rangle$ e, então, o conjunto $\{u_1, \dots, u_j, u_{j+1}, u_{j+2}\}$ seria linearmente independente.

Esse processo tem fim, pois qualquer conjunto com $n + 1$ elementos no \mathbb{R}^n é linearmente dependente, de acordo com o Teorema 3.39. Ou seja, temos

$$\langle u_1, \dots, u_j, \dots, u_{j+k} \rangle = S$$

para algum $k \in \{0, 1, \dots, n - j\}$. □

Corolário 3.44 Todo subespaço $S \neq \{0\}$ do \mathbb{R}^n possui uma base.

Demonstração: De fato, basta tomar um vetor $0 \neq v_1 \in S$ e aplicar o Teorema 3.43 ao conjunto $\mathcal{C} = \{v_1\}$. □

Note que o subespaço $\{0\} \subset \mathbb{R}^n$ não possui base. De fato, se \mathcal{B} fosse uma base do subespaço $\{0\}$, então 0 seria um elemento de \mathcal{B} . De acordo com o Exemplo 3.26, esse conjunto seria linearmente dependente!

Por outro lado, a demonstração do Corolário 3.44 garante que a existência de uma infinidade de bases para qualquer subespaço S do \mathbb{R}^n .

Definição 3.45 Seja S um subespaço do \mathbb{R}^n . Definimos a **dimensão** de S como o número de elementos de uma de suas bases. Escrevemos $\dim S = k$, se $\{v_1, \dots, v_k\}$ for uma base de S . Se $S = \{0\}$, dizemos que S tem dimensão 0.

Como todas as bases de um subespaço tem o mesmo número de elementos, está bem definida a dimensão do subespaço S .

Exemplo 3.46 Já vimos que o \mathbb{R}^n tem dois subespaços triviais: o $\{0\}$ e o próprio \mathbb{R}^n . Agora podemos caracterizar todos os subespaços do \mathbb{R}^n . Além desses, existem todas as retas passando pela origem (que são os subespaços de dimensão 1), todos os planos passando pela origem (que são os subespaços de dimensão 2), todos os subespaços de dimensão 3 e assim por diante, até chegarmos a todos os subespaços de dimensão $n - 1$. O próprio \mathbb{R}^n é seu único subespaço com dimensão n . (Note que, por falta de nomes adequados, mudamos a maneira de nos referir aos subespaços do \mathbb{R}^n : começamos com retas e planos e passamos a falar de subespaços de dimensão 3 etc.) △

De acordo com o Teorema 3.43, podemos obter uma base a partir de qualquer conjunto linearmente independente. Mas também podemos obter uma base a partir de qualquer conjunto que gera um subespaço:

Proposição 3.47 Suponha que o conjunto $\{v_1, \dots, v_k\} \subset S$ gere o subespaço $S \neq \{0\}$. Então um subconjunto de $\{v_1, \dots, v_k\}$ é uma base de S .

Demonstração: Se $\{v_1, \dots, v_k\}$ for linearmente dependente, um dos vetores desse conjunto é combinação linear dos vetores restantes. Retirando esse vetor, o conjunto restante continua gerando S . Continuamos retirando vetores que são combinação linear dos elementos restantes até obter um conjunto linearmente independente que continua gerando S . \square

A demonstração da Proposição 3.47 mostra como obter uma base a partir de um conjunto que gera um subespaço. Contudo, esse processo é muito trabalhoso: em cada etapa, precisamos verificar se o conjunto de vetores é linearmente independente; caso contrário, deve-se obter então um vetor como combinação linear dos vetores restantes, retirar esse vetor do conjunto e novamente verificar se o conjunto restante é linearmente independente, repetindo o processo, se necessário. Em cada etapa, retiramos apenas um vetor supérfluo. Gostaríamos, entretanto, de retirar de uma única vez todos os vetores supérfluos. Mostraremos como fazê-lo na Observação 4.23.

Se já sabemos qual a dimensão do subespaço $S \subset \mathbb{R}^n$, então a verificação de que um conjunto é uma base de S pode ser simplificada.

Teorema 3.48 Sejam $Y \subset \mathbb{R}^n$ um subespaço de dimensão k e

$$\mathcal{B} = \{v_1, \dots, v_k\} \subset Y.$$

Então, se uma das condições seguintes for satisfeita, \mathcal{B} é uma base de S :

- (i) o conjunto \mathcal{B} gera Y ;
- (ii) o conjunto \mathcal{B} é linearmente independente.

Demonstração: Suponhamos inicialmente que \mathcal{B} gere Y . Se esse conjunto fosse linearmente dependente, aplicando a Proposição 3.47 obteríamos então um subconjunto de \mathcal{B} , com menos que k elementos, que seria uma base de Y . Mas isso contradiz Y ter dimensão k .

Por outro lado, se \mathcal{B} fosse linearmente independente e não gerasse Y , existiria $v_{k+1} \in Y$, que não é combinação linear dos vetores de \mathcal{B} . De acordo com a Proposição 3.22 (ou a Observação 3.28), o conjunto $\{v_1, \dots, v_k, v_{k+1}\}$ seria linearmente independente. Mas isso contradiz o Teorema 3.39, pois existiria um conjunto linearmente independente com mais vetores do que a dimensão do espaço. \square

3.4 EXERCÍCIOS

- Determine os valores de a , b e c para que o vetor $v = (a, b, c)$ seja combinação linear dos vetores $u_1 = (1, 1, 1)$, $u_2 = (1, 2, 3)$ e $u_3 = (2, -1, 1)$.
- Sem utilizar determinantes, verifique se os vetores $(-1, 2, 3)$, $(-2, -3, 1)$ e $(3, 2, 1)$ são linearmente dependentes ou linearmente independentes.
- Seja

$$A = \begin{pmatrix} 1 & 3 \\ 4 & -3 \end{pmatrix}.$$

Determine $x \neq 0$ tal que $Ax = 3x$.

- Encontre um conjunto de geradores para o núcleo da matriz A , dada por

$$(a) \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 2 & 1 & 2 \\ 2 & 4 & 2 & 4 \\ 0 & 1 & 1 & 2 \end{pmatrix}; \quad (b) \begin{pmatrix} -1 & 2 & 2 & 0 \\ -1 & 2 & 1 & 0 \\ -1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- Encontre uma base para o núcleo de A , se

$$(a) \begin{pmatrix} 2 & 2 & 3 & 4 \\ 0 & 2 & 3 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad (b) \begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

- Considere os vetores $v_1 = (4, 2, -3)$, $v_2 = (2, 1, -2)$ e $v_3 = (-2, -1, 0)$ e defina $U = \langle v_1, v_2, v_3 \rangle$.
 - Mostre que $\{v_1, v_2, v_3\}$ é linearmente dependente;
 - calcule a dimensão de U ;
 - descreva geometricamente U ;
 - ache uma base para U .
- Sejam $v_1 = (1, 1, 1)$ e $v_2 = (1, 2, 1)$.
 - Determine $\langle v_1, v_2 \rangle$. Qual a dimensão desse subespaço do \mathbb{R}^3 ?
 - Determine um vetor $v_3 \in \mathbb{R}^3$ de modo que o conjunto $\{v_1, v_2, v_3\}$ seja uma base do \mathbb{R}^3 .
- Seja $V = \{(3a + 4b - 4c, 2a - 4b - 6c, -2a - 4b + 2c) : a, b, c \in \mathbb{R}\}$.
 - Mostre que V é um subespaço do \mathbb{R}^3 ;
 - ache uma base para V .
- Seja A uma matriz $m \times n$ e b um ponto do \mathbb{R}^m . Com a identificação usual de pontos com matrizes coluna, mostre que o conjunto

$$\text{im } A = \{b \in \mathbb{R}^m : \exists x \in \mathbb{R}^n \text{ com } Ax = b\}$$
 é um subespaço do \mathbb{R}^m .
- Seja $\mathcal{B} = \{v_1, \dots, v_k\}$ um subconjunto de vetores do subespaço $V \subset \mathbb{R}^n$. Suponha que cada vetor $v \in V$ se escreva, de maneira única, como combinação linear dos vetores de \mathcal{B} . Mostre que \mathcal{B} é uma base de V . (Compare com a Proposição 3.37.)

AULA 4

Aplicações Lineares

OBJETIVOS

No final desta aula, o aluno deverá ser capaz de:

1. reconhecer uma aplicação linear;
2. obter sua representação matricial;
3. obter bases para o núcleo e imagem de uma aplicação linear.

4.1 APlicações Lineares e Matrizes – PARTE I

Já vimos o tipo de conjunto que estudaremos neste curso: os (sub)espaços vetoriais. Agora veremos o tipo especial de funções entre esses conjuntos: as aplicações lineares.

Definição 4.1 Uma **aplicação linear** é uma função $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que

$$T(x + \lambda y) = Tx + \lambda Ty. \quad (4.1)$$

Essa definição merece vários comentários. Em primeiro lugar, notamos que estamos representando a imagem do ponto $x \in \mathbb{R}^n$ por Tx ao invés de $T(x)$. Essa será nossa prática: parênteses só serão utilizados para esclarecer o significado de uma expressão. Em segundo lugar, notamos que a soma de vetores e a multiplicação por escalar presentes em (4.1) possuem significados diferentes: no lado esquerdo da igualdade, a soma de vetores e a multiplicação por escalar ocorrem no espaço \mathbb{R}^n , enquanto no lado direito da igualdade acontecem no espaço \mathbb{R}^m . Finalmente, também é usual chamar uma aplicação linear de *transformação linear*.

Exemplo 4.2 Seja $I: \mathbb{R}^n \rightarrow \mathbb{R}^n$, dada por $Ix = x$. Então I é uma aplicação linear pois, para $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, temos $I(x + \alpha y) = x + \alpha y = Ix + \alpha Iy$. A aplicação linear I é chamada *aplicação identidade* (no espaço \mathbb{R}^n). \square

Exemplo 4.3 Seja $N: \mathbb{R}^n \rightarrow \mathbb{R}^m$ a aplicação definida por $Nx = 0$ para todo $x \in \mathbb{R}^n$. A aplicação N é chamada *aplicação nula*, qualquer que seja m , dimensão do espaço \mathbb{R}^m . Também se denota por 0 a aplicação nula: $0x = 0$ para todo $x \in \mathbb{R}^n$. (Note que esse é o terceiro emprego distinto do símbolo 0 .) \square

Exemplo 4.4 Na seção 2.2 vimos que a identificação de pontos (no caso, do \mathbb{R}^3) com matrizes coluna nos permitiu interpretar geometricamente um sistema linear em 3 incógnitas. Mas essa mesma identificação, no caso geral do \mathbb{R}^n , nos permite mais: ela nos permite interpretar uma matriz $m \times n$ com uma aplicação linear $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$. De fato, dado o ponto $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, escrevendo esse ponto como uma matriz coluna temos que

$$\begin{aligned} & \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix} = Ax \in \mathbb{R}^m. \end{aligned}$$

Assim, a matriz $A = (a_{ij})$ associa a cada ponto $x \in \mathbb{R}^n$, um ponto $Ax \in \mathbb{R}^m$. Para mostrar que essa função é linear (e, portanto, uma aplicação linear) basta relembrar propriedades básicas de matrizes:

$$A(x + \lambda y) = Ax + \lambda Ay.$$

□

Exercício 4.5 Considere a matriz

$$A = \begin{pmatrix} 2 & 1 & 7 & 5 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Quais são o domínio e o contra domínio da aplicação linear definida por A ? Calcule Ax , se $x = (1, 1, 1, 1)$.

Exemplo 4.6 Seja $f: \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = x^2$. A função f é uma aplicação linear? (Lembre-se: $\mathbb{R} = \mathbb{R}^1$ é um espaço vetorial!)

Uma vez que $f(x+y) = (x+y)^2 = x^2 + 2xy + y^2 = f(x) + f(y) + 2xy$, vemos que f não é linear. □

Exemplo 4.7 Consideremos $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ definida por

$$T(x_1, \dots, x_n) = (a_{11}x_1 + \dots + a_{1n}x_n, a_{21}x_1 + \dots + a_{2n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n),$$

em que os escalares a_{ij} , $i \in \{1, \dots, m\}$ e $j \in \{1, \dots, n\}$ são arbitrários. Um caso particular seria $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dada por

$$T(x_1, x_2, x_3) = (3x_1 + 2x_2, x_1 + x_2 + 2x_3).$$

Afirmamos que, no caso geral, T é uma aplicação linear. De fato, basta notar que, em cada uma das coordenadas de T , temos

$$a_{i1}(x_1 + \lambda y_1) + \dots + a_{in}(x_n + \lambda y_n) = (a_{i1}x_1 + \dots + a_{in}x_n) + \dots + \lambda(a_{i1}y_1 + \dots + a_{in}y_n).$$

Dessa igualdade, válida para todo $i \in \{1, \dots, m\}$, deduzimos que T é linear.

Se compararmos a definição de Tx com a expressão de Ax no Exemplo 4.6, não podemos deixar de perceber a semelhança: cada uma das coordenadas de Tx coincide com as coordenadas de Ax . Vamos explicar essa semelhança. Para isso, consideremos a base canônica $\mathcal{E} = \{e_1, \dots, e_n\}$ do \mathbb{R}^n . (Veja o Exemplo 3.31.)

Temos que $x = (x_1, \dots, x_n) = x_1e_1 + \dots + x_ne_n$. Assim, como cada x_i é um escalar e e_i um vetor, temos

$$Tx = T(x_1e_1 + \dots + x_ne_n) = x_1Te_1 + \dots + x_nTe_n.$$

De acordo com a definição de Tx , escrevendo a imagem de cada vetor Te_i como uma matriz coluna, obtemos

$$Te_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix}$$

e, portanto,

$$\begin{aligned} Tx &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \end{aligned} \quad (4.2)$$

A matriz $A = (a_{ij})$ é chamada *representação* de T com relação às bases canônicas do \mathbb{R}^n e \mathbb{R}^m . \triangleleft

Exercício 4.8 Considere $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definida por

$$Tx = T(x_1, x_2) = (3x_1 + x_2, x_1 + x_2, x_2).$$

Verifique que T é uma aplicação linear. Calcule a matriz A que representa T com relação às bases canônicas do \mathbb{R}^2 e \mathbb{R}^3 .

Exercício 4.9 Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear qualquer. Mostre que $T(0) = 0$.

Exemplo 4.10 Consideremos uma aplicação linear qualquer $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$. Vamos mostrar que, como no Exemplo 4.7, podemos identificar T com uma matriz. Para isso, notamos que, *escolhida* a base canônica do \mathbb{R}^n , temos, como antes,

$$Tx = x_1Te_1 + \dots + x_nTe_n. \quad (4.3)$$

Cada Te_i ($i \in \{1, \dots, n\}$) é um ponto do \mathbb{R}^m . Denotando

$$Te_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} \in \mathbb{R}^m$$

e substituindo em (4.3), chegamos, como no exemplo anterior, à igualdade

$$Tx = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = Ax.$$

Quer dizer, escolhendo as bases canônicas do \mathbb{R}^n e do \mathbb{R}^m , qualquer aplicação linear pode ser representada por uma matriz, chamada representação de T com relação a essas bases canônicas, de modo que $Tx = Ax$. (Posteriormente veremos porque a base canônica do \mathbb{R}^m foi apenas implicitamente utilizada neste exemplo.) \triangleleft

Salientamos o resultado que acabamos de mostrar:

Teorema 4.11 Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear. Então a matriz que representa T com relação às bases canônicas do \mathbb{R}^n e \mathbb{R}^m é dada por

$$(Te_1 \ Te_2 \ \dots \ Te_n),$$

em que Te_i é uma matriz coluna.

Cada aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ define uma matriz $m \times n$ e vice-versa. Assim, qual é a diferença entre aplicações lineares e matrizes? Essa diferença é sutil: uma aplicação linear não depende de escolha de bases no domínio e contra-domínio; se escolhermos essas bases, toda aplicação linear é representada por uma matriz. Dito de outra maneira, falamos de aplicação linear quando os espaços envolvidos (\mathbb{R}^n e \mathbb{R}^m) são considerados sem quaisquer bases. Escolhidas as bases nesses espaços, aí então obtemos matrizes. (Por enquanto, só escolhemos as bases canônicas no \mathbb{R}^n e \mathbb{R}^m .)

Exemplo 4.12 Seja $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ uma aplicação linear. Suponhamos que

$$Te_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad Te_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \text{e} \quad Te_3 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Então podemos determinar Tx , qualquer que seja o vetor $x \in \mathbb{R}^3$. De fato, se $x = (3, 2, 1) \in \mathbb{R}^3$, então

$$\begin{aligned} Tx &= 3Te_1 + 2Te_2 + 1Te_3 \\ &= 3\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2\begin{pmatrix} 3 \\ 1 \end{pmatrix} + 1\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 10 \\ 7 \end{pmatrix}. \end{aligned} \quad \triangleleft$$

O Exemplo 4.12 nos mostra que aplicações lineares $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções muito rígidas: basta conhecermos seu valor em todos os pontos da base canônica $\{e_1, \dots, e_n\}$ do \mathbb{R}^n para conhecermos seu valor em qualquer ponto do \mathbb{R}^n ! (No caso de $T: \mathbb{R} \rightarrow \mathbb{R}$, isso significa que basta conhecer a imagem de $Te_1 = T(1)$ para conhecermos seu valor em qualquer ponto de \mathbb{R} . Compare essa situação com a de uma função qualquer $f: \mathbb{R} \rightarrow \mathbb{R}$.) Mas a base canônica não tem nada de extraordinário: o mesmo resultado vale para qualquer base do \mathbb{R}^n .

Teorema 4.13 Seja $\{v_1, \dots, v_n\}$ uma base do \mathbb{R}^n . Fixados vetores arbitrários $w_1, \dots, w_n \in \mathbb{R}^m$, existe uma única aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que $Tv_i = w_i$.

Demonstração: Dado $v \in \mathbb{R}^n$, existem únicos escalares $\alpha_1, \dots, \alpha_n$ tais que $v = \alpha_1 v_1 + \dots + \alpha_n v_n$. Definimos então

$$Tv = \alpha_1 w_1 + \dots + \alpha_n w_n.$$

Temos que T é linear. De fato, se $v = \alpha_1 v_1 + \dots + \alpha_n v_n$ e $u = \beta_1 v_1 + \dots + \beta_n v_n$, então $v + \gamma u = (\alpha_1 + \gamma\beta_1)v_1 + \dots + (\alpha_n + \gamma\beta_n)v_n$, de modo que

$$\begin{aligned} T(v + \gamma u) &= (\alpha_1 + \gamma\beta_1)w_1 + \dots + (\alpha_n + \gamma\beta_n)w_n \\ &= \alpha_1 w_1 + \dots + \alpha_n w_n + \gamma(\beta_1 w_1 + \dots + \beta_n w_n) \\ &= Tv + \gamma Tu. \end{aligned}$$

Suponhamos agora que $S: \mathbb{R}^n \rightarrow \mathbb{R}^m$ seja uma aplicação linear com $Sv_i = w_i$. Vamos mostrar que $Sv = Tv$ para todo $v \in \mathbb{R}^n$, o que garante a unicidade de T . De fato, se $v = \alpha_1 v_1 + \dots + \alpha_n v_n$, então temos,

$$\begin{aligned} Sv &= S(\alpha_1 v_1 + \dots + \alpha_n v_n) \\ &= \alpha_1 S v_1 + \dots + \alpha_n S v_n = \alpha_1 w_1 + \dots + \alpha_n w_n = Tv, \end{aligned}$$

mostrando o afirmado. \square

Exercício 4.14 Sejam $\mathcal{B} = \{v_1 = (1, 1, 0), v_2 = (1, 2, 0), v_3 = (1, 1, 1)\}$. Ache a aplicação linear $T: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ tal que $Tv_1 = (1, 0, 0, 0)$, $Tv_2 = (0, 0, 0, 1)$ e $Tv_3 = (1, 1, 1, 1)$.

A rigidez de uma aplicação linear produz resultados surpreendentes. Por exemplo,

Proposição 4.15 Sejam $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear e $\{v_1, \dots, v_n\}$ um conjunto qualquer. Se $\{Tv_1, \dots, Tv_n\}$ for linearmente independente, então $\{v_1, \dots, v_n\}$ é linearmente independente e, portanto, uma base do \mathbb{R}^n .

Demonstração: Suponhamos que

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0. \quad (4.4)$$

Queremos provar que $\alpha_i = 0$ para todo $i \in \{1, \dots, n\}$. Aplicando T em ambos os lados da igualdade (4.4), obtemos

$$0 = T(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1(Tv_1) + \dots + \alpha_n(Tv_n).$$

Como os vetores Tv_1, \dots, Tv_n são linearmente independentes, temos $\alpha_1 = \dots = \alpha_n = 0$, como queríamos provar. O fato de $\{v_1, \dots, v_n\}$ ser uma base do \mathbb{R}^n é consequência do Teorema 3.48. \square

4.2 ESPAÇO LINHA E ESPAÇO COLUNA

Definimos agora os dois subespaços fundamentais associados a uma aplicação linear:

Definição 4.16 Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear. Definimos a **imagem** de T , denotada por $\text{im } T$, por

$$\text{im } T = \{y \in \mathbb{R}^m : y = Tx\}.$$

Definimos o **núcleo** de T , denotado por $\ker T$, por

$$\ker T = \{x \in \mathbb{R}^n : Tx = 0\}.$$

O núcleo e a imagem de T são subespaços vetoriais do domínio de T (isto é, o \mathbb{R}^n) e do contradomínio de T (isto é, o \mathbb{R}^m), respectivamente. De fato, se $x_1, x_2 \in \ker T$ e $\lambda \in \mathbb{R}$, então $T(x_1 + \lambda x_2) = T(x_1) + \lambda T(x_2) = 0 + \lambda 0 = 0$, provando que $x_1 + \lambda x_2 \in \ker T$. Se $y_1, y_2 \in \text{im } T$, então existem $x_1, x_2 \in X$ tais que $y_1 = T(x_1)$ e $y_2 = T(x_2)$. Logo, se $\lambda \in \mathbb{R}$, $y_1 + \lambda y_2 = T(x_1) + \lambda T(x_2) = T(x_1 + \lambda x_2)$, o que mostra que $y_1 + \lambda y_2 \in \text{im } T$.

O subespaço $\ker T$ já havia sido definido para uma matriz A . Existe alguma diferença entre as duas definições dadas? Como dissemos, ao definirmos $\ker T$, o espaço \mathbb{R}^n está sendo considerado sem qualquer base (ou seja, sem qualquer sistema de eixos coordenados); se A for a matriz que representa T (com relação às bases canônicas do \mathbb{R}^n e \mathbb{R}^m), então os elementos de $\ker T$ estão sendo descritos em relação à base canônica do \mathbb{R}^n . Os elementos de cada conjunto são os mesmos, mas não existem coordenadas envolvidas na definição de $\ker T$!

Muitos problemas envolvem a obtenção de bases para esses dois subespaços. Em geral, partimos da base canônica para então obtermos uma base adequada para cada um deles. Por esse motivo, muito do que faremos utiliza a linguagem de matrizes.

Por isso, nesta seção estudaremos mais detalhadamente a matriz $A = (a_{ij})$. Como sabemos, ela pode ser vista por meio de suas linhas ou colunas:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = (c_1 \dots c_n) = \begin{pmatrix} \ell_1 \\ \vdots \\ \ell_m \end{pmatrix}. \quad (4.5)$$

Os vetores coluna c_1, \dots, c_n são naturalmente identificados com vetores do \mathbb{R}^m . Se $\mathcal{C} = \{c_1, \dots, c_n\}$, chamamos de *espaço coluna* o espaço gerado por \mathcal{C} , isto é, $\langle \mathcal{C} \rangle \subset \mathbb{R}^m$.

Por outro lado, podemos interpretar as linhas de A como elementos do próprio espaço \mathbb{R}^n . Se denotarmos $\mathcal{L} = \{\ell_1, \dots, \ell_m\} \subset \mathbb{R}^n$, chamamos de *espaço linha* o espaço gerado por \mathcal{L} , isto é, $\langle \mathcal{L} \rangle \subset \mathbb{R}^n$.

Observação 4.17 Um vetor é identificado com uma matriz coluna:

$$(x_1, \dots, x_n) \longleftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

A identificação de um vetor com uma matriz linha

$$(x_1, \dots, x_n) \longleftrightarrow (x_1 \ x_2 \ \dots \ x_n),$$

como feita na definição do espaço linha, é excepcional e deve ser evitada, pelo menos neste primeiro curso. \triangleleft

Exemplo 4.18 Considere a matriz

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

O espaço coluna de A é o subespaço do \mathbb{R}^2

$$\langle \mathcal{C} \rangle = \left\{ \alpha \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ 2 \end{pmatrix} : \alpha, \beta, \gamma \in \mathbb{R} \right\}.$$

Como as duas primeiras colunas de A são dadas por vetores linearmente independentes, é claro que $\langle \mathcal{C} \rangle = \mathbb{R}^2$.

Por outro lado, o espaço linha de A é o subespaço do \mathbb{R}^3

$$\langle \mathcal{L} \rangle = \{ \alpha(2, 1, 1) + \beta(1, 1, 2) : \alpha, \beta \in \mathbb{R} \}.$$

Como as duas primeiras linhas de A são dadas por vetores linearmente independentes, $\langle \mathcal{L} \rangle$ é um plano do \mathbb{R}^3 . Note que as dimensões de $\langle \mathcal{L} \rangle$ e $\langle \mathcal{C} \rangle$ são iguais. \triangleleft

Mas, qual o significado dos subespaços $\langle \mathcal{C} \rangle$ e $\langle \mathcal{L} \rangle$? Começamos interpretando o espaço coluna de uma matriz.

Lema 4.19 Considere o sistema linear não homogêneo $Ax = b$, em que $A = (a_{ij})$ é uma matriz $m \times n$. Então são equivalentes:

- (i) existe solução x para $Ax = b$;
- (ii) o vetor b é combinação linear das colunas de A .

Demonstração: Basta notar que o sistema $Ax = b$ é equivalente à equação

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad (4.6)$$

de acordo com a equação (4.2).

Quer dizer, se $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ for uma solução de $Ax = b$, então b é combinação linear das colunas de A por meio dos escalares x_1, \dots, x_n . Reciprocamente, se existirem escalares x_1, \dots, x_n de modo que b seja combinação linear das colunas de A , então o vetor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ é solução de $Ax = b$. \square

Em outras palavras, acabamos de mostrar que $\langle \mathcal{C} \rangle$ é o subespaço $\text{im } A$.

Definição 4.20 Se $A = (a_{ij})$ for uma matriz $m \times n$, definimos a **transposta** de A como a matriz $n \times m$ $A^t = (a_{ij}^t)$, com $a_{ij}^t = a_{ji}$.

Logo, se A for a matriz dada por (4.5), então

$$A^t = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}.$$

Assim, as colunas da matriz A^t são justamente as linhas da matriz A . Como consequência imediata do Lema 4.19, temos que

$$\langle \mathcal{L} \rangle = \text{im } A^t. \quad (4.7)$$

Utilizando essas interpretações para os espaços linha e coluna de uma matriz, estamos em condições de resolver um problema fundamental: dada uma matriz A , encontrar bases para os espaços $\ker A$ e $\text{im } A$. Para isso, enunciamos o seguinte

Corolário 4.21 Seja A uma matriz $m \times n$ e R sua forma escalonada reduzida por linhas. Então as colunas de A correspondentes aos pivôs de R formam uma base do espaço coluna de A .

Demonstração: Considere a matriz aumentada $(A|b)$, com $b \in \text{im } A$ e sua forma escalonada reduzida por linhas $(R|b')$. É fácil ver que as colunas de R correspondentes aos pivôs formam uma base do espaço coluna de R . Quer dizer, existe uma solução x de $(R|b')$ com coordenadas não nulas apenas nas posições correspondentes aos pivôs. Como as soluções de $Ax = b$ e $Rx = b'$ são as mesmas, a equação (4.6) garante que combinações lineares de colunas de R produzem combinações lineares das colunas correspondentes de A . Em particular, as colunas de A correspondentes aos pivôs de R são linearmente independentes e geram o espaço coluna de A . \square

Estude o próximo exemplo e então releia a demonstração do Corolário 4.21.

Exemplo 4.22 Vamos obter bases para o núcleo e a imagem da matriz

$$A = \begin{pmatrix} 3 & 1 & 2 & 4 & 1 \\ 1 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 1 & 1 \end{pmatrix}.$$

Para obter o núcleo, resolvemos $Ax = 0$ levando a matriz A a sua forma escalonada reduzida por linhas:

$$\begin{pmatrix} 3 & 1 & 2 & 4 & 1 \\ 1 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1/2 & 0 & -5 \\ 0 & 1 & 1/2 & 0 & 4 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}.$$

Escrevendo as variáveis correspondentes aos pivôs em termos da variáveis livres, obtemos a solução geral do sistema homogêneo associado.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = x_3 \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 5 \\ -4 \\ 0 \\ -3 \\ 1 \end{pmatrix} \quad (4.8)$$

Assim, podemos concluir que o núcleo de A é um subespaço de dimensão 2, gerado pelos vetores

$$u_1 = \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^5 \quad \text{e} \quad u_2 = \begin{pmatrix} 5 \\ -4 \\ 0 \\ -3 \\ 1 \end{pmatrix} \in \mathbb{R}^5.$$

De acordo com o Corolário 4.21, temos que os vetores

$$v_1 = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad \text{e} \quad v_3 = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}$$

formam uma base da imagem de A , pois estes vetores coluna correspondem às posições dos pivôs na forma escalonada reduzida por linhas de A .

Vamos aproveitar este exemplo e ilustrar a demonstração do Corolário 4.21. Se quisermos escrever a última coluna da matriz A como combinação linear das colunas de A correspondentes aos pivôs, basta tomar o vetor correspondente à variável livre x_5 . Esse vetor tem a quinta coordenada não nula.

$$\begin{pmatrix} 5 \\ -4 \\ 0 \\ -3 \\ 1 \end{pmatrix}.$$

Essa é uma solução de $Ax = 0$. Assim,

$$5 \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} - 4 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + 0 \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} - 3 \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Ou seja,

$$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = -5 \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + 4 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}.$$

Veja que esses mesmos escalares escrevem a última coluna da matriz R como combinação linear das colunas correspondentes aos pivôs:

$$\begin{pmatrix} -5 \\ 4 \\ 3 \end{pmatrix} = -5 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad \square$$

Observação 4.23 Resulta do Corolário 4.21 um método simples para a obtenção de uma base do subespaço $\langle v_1, \dots, v_k \rangle \subset \mathbb{R}^n$: basta formar a matriz $A = (v_1 \ v_2 \ \dots \ v_k)$ e então selecionar as colunas correspondentes aos pivôs da sua forma escalonada reduzida por linhas. Compare com a Observação 3.23.

Um outro método para obter-se uma base para a imagem da matriz A consiste em escrever suas colunas como linhas, ou seja, considerar a matriz transposta. Uma vez que o processo de escalonamento de uma matriz (no caso, de A^t) produz apenas combinações lineares das linhas envolvidas, faz com que o espaço gerado por essas linhas seja conservado no processo de escalonamento. Assim, as linhas não nulas da forma escalonada de A^t

são linearmente independentes e, escritas novamente como colunas, produzem uma base para $\text{im } A$.

Note, contudo, que ao utilizar esse método, temos que escalarizar duas matrizes distintas: a matriz A , para obter uma base para $\ker A$, e a matriz A^t , para produzir uma base para $\text{im } A$. Quer dizer, a aplicação do Corolário 4.21 é um método muito mais eficaz! \square

Vamos agora relacionar as dimensões dos subespaços $\langle \mathcal{C} \rangle$ e $\langle \mathcal{L} \rangle$ de uma matriz A . Mostraremos que esses espaços têm a mesma dimensão; isso é um fato notável, pois eles são subespaços de espaços vetoriais diferentes!

Teorema 4.24 *Dada uma matriz $m \times n$, seu espaço linha tem a mesma dimensão de seu espaço coluna.*

Demonstração: Suponhamos que os vetores

$$b_1 = (b_{11}, b_{12}, \dots, b_{1n}), b_2 = (b_{21}, b_{22}, \dots, b_{2n}), \dots, b_r = (b_{r1}, b_{r2}, \dots, b_{rn})$$

formem uma base do espaço linha da matriz A . Então cada linha ℓ_i de A é combinação linear desses elementos:

$$\begin{aligned}\ell_1 &= \lambda_{11}b_1 + \dots + \lambda_{1r}b_r \\ \ell_2 &= \lambda_{21}b_1 + \dots + \lambda_{2r}b_r \\ &\vdots = \vdots \\ \ell_m &= \lambda_{m1}b_1 + \dots + \lambda_{mr}b_r\end{aligned}$$

Igualando a componente j de cada uma dessas equações, obtemos

$$\begin{aligned}a_{1j} &= \lambda_{11}b_{1j} + \lambda_{12}b_{2j} + \dots + \lambda_{1r}b_{rj} \\ a_{2j} &= \lambda_{21}b_{1j} + \lambda_{22}b_{2j} + \dots + \lambda_{2r}b_{rj} \\ &\vdots = \vdots \\ a_{mj} &= \lambda_{m1}b_{1j} + \lambda_{m2}b_{2j} + \dots + \lambda_{mr}b_{rj}.\end{aligned}$$

Assim,

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = b_{1j} \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \\ \vdots \\ \lambda_{m1} \end{pmatrix} + b_{2j} \begin{pmatrix} \lambda_{12} \\ \lambda_{22} \\ \vdots \\ \lambda_{m2} \end{pmatrix} + \dots + b_{rj} \begin{pmatrix} \lambda_{1r} \\ \lambda_{2r} \\ \vdots \\ \lambda_{mr} \end{pmatrix},$$

mostrando que as colunas de A são combinações lineares dos r vetores

$$\begin{pmatrix} \lambda_{11} \\ \lambda_{21} \\ \vdots \\ \lambda_{m1} \end{pmatrix}, \dots, \begin{pmatrix} \lambda_{1r} \\ \lambda_{2r} \\ \vdots \\ \lambda_{mr} \end{pmatrix}.$$

Isso quer dizer que o espaço coluna tem dimensão, no máximo, igual a r , ou seja,

$$\dim \langle \mathcal{C} \rangle \leq \dim \langle \mathcal{L} \rangle.$$

Procedendo da mesma maneira com relação a uma base do espaço coluna, mostramos que

$$\dim \langle \mathcal{L} \rangle \leq \dim \langle \mathcal{C} \rangle.$$

Assim, essas duas dimensões são iguais.¹ \square

¹ Podemos notar que mostramos $\langle \mathcal{C} \rangle \leq \dim \langle \mathcal{L} \rangle$ para qualquer matriz. Aplicando esse fato à matriz A^t , obtemos o resultado.

Definição 4.25 Definimos o **posto** da matriz A , denotado por $\text{posto } A$, como sendo

$$\dim \langle \mathcal{C} \rangle = \dim \langle \mathcal{L} \rangle.$$

Se A for uma representação matricial da aplicação linear T , definimos

$$\text{posto } T = \text{posto } A.$$

O seguinte resultado decorre imediatamente do Teorema 4.24:

Corolário 4.26 Seja A uma matriz $m \times n$. Então,

$$\dim(\text{im } A) = \dim(\text{im } A^t).$$

4.3 MULTIPLICAÇÃO DE MATRIZES

Sem dúvida, uma das definições mais estranhas da Matemática parece ser a da multiplicação de matrizes: dadas as matrizes $B = (b_{ik})$, $p \times m$, e $A = (a_{kj})$, $m \times n$, e então o produto BA é a matriz $d = (d_{ij})$, definida por

$$d_{ij} = \sum_{k=1}^m b_{ik} a_{kj} = (b_{i1} \ b_{i2} \ \dots \ b_{im}) \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix},$$

correspondente à multiplicação da linha ℓ_i da matriz B pela coluna c_j da matriz A .

Historicamente, a definição da multiplicação de matrizes foi imposição de um fato muito mais elementar: a composição de funções (no caso, aplicações lineares). Começamos mostrando que a composição de aplicações lineares também define uma aplicação linear:

Lema 4.27 Sejam $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $S: \mathbb{R}^m \rightarrow \mathbb{R}^p$ duas aplicações lineares. Então a composta $S \circ T: \mathbb{R}^n \rightarrow \mathbb{R}^p$ é uma aplicação linear:

$$\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m \xrightarrow{S} \mathbb{R}^p.$$

Demonstração: De fato, se $x + \lambda y \in \mathbb{R}^n$, então

$$\begin{aligned} (S \circ T)(x + \lambda y) &= S(T(x + \lambda y)) \\ &= S(Tx + \lambda Ty) = S(Tx) + \lambda S(Ty) \\ &= (S \circ T)x + \lambda(S \circ T)y, \end{aligned}$$

provando o afirmado. \square

Suponhamos que $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $S: \mathbb{R}^m \rightarrow \mathbb{R}^p$ sejam representadas (com relação às bases canônicas dos espaços envolvidos) pelas matrizes A e B , respectivamente. Assim,

$$Tx = Ax \quad \forall x \in \mathbb{R}^n \quad \text{e} \quad Sy = By \quad \forall y \in \mathbb{R}^m. \quad (4.9)$$

Uma vez que $Tx \in \mathbb{R}^m$, temos então que

$$(S \circ T)x = S(Tx) = B(Ax) = (BA)x,$$

vemos que a composta $S \circ T$ é representada pela matriz BA . Esse fato justifica uma notação frequentemente usada: ao invés de denotarmos $S \circ T$, escrevemos simplesmente ST .

Mas, como obter a definição da multiplicação de matrizes BA a partir da composição ST de aplicações lineares? É o que passaremos a mostrar.

Partimos da linearidade da aplicação S :

$$Sx = S(x_1e_1 + \dots + x_m e_m) = x_1Se_1 + \dots + x_m Se_m. \quad (4.10)$$

Denotando

$$Se_j = \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{pmatrix},$$

concluímos que

$$Sx = \begin{pmatrix} b_{11}x_1 + \dots + b_{1m}x_m \\ b_{21}x_1 + \dots + b_{2m}x_m \\ \vdots \\ b_{p1}x_1 + \dots + b_{pm}x_m \end{pmatrix}, \quad (4.11)$$

expressão resultante de (4.10) e da definição da adição de vetores no \mathbb{R}^m .

Suponhamos que $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $S: \mathbb{R}^m \rightarrow \mathbb{R}^p$ sejam representadas (com relação às bases canônicas dos espaços envolvidos) pelas matrizes A e B , respectivamente. Quer dizer,

$$A = (Te_1 \ Te_2 \ \dots \ Te_n) = (c_1 \ c_2 \ \dots \ c_n)$$

e

$$B = (Se_1 \ Se_2 \ \dots \ Se_m) = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pm} \end{pmatrix},$$

em que a matriz A está sendo denotada por suas colunas.

A aplicação linear $ST: \mathbb{R}^n \rightarrow \mathbb{R}^p$ é representada, com relação às bases canônicas do \mathbb{R}^n e \mathbb{R}^p , por uma matriz D (denotada por suas colunas) definida por

$$D = ((ST)e_1 \ (ST)e_2 \ \dots \ (ST)e_n).$$

Para obter D , notamos que $(ST)e_j = S(Te_j) = Sc_j$, o que implica

$$D = (Sc_1 \ Sc_2 \ \dots \ Sc_n).$$

Uma vez que $(ST)x = (BA)x$, obtemos a definição do produto BA , de acordo com a equação (4.11): a j -ésima coluna de BA é dada por

$$\begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pm} \end{pmatrix} \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = \begin{pmatrix} b_{11}a_{1j} + b_{12}a_{2j} + \dots + b_{1m}a_{mj} \\ b_{21}a_{1j} + b_{22}a_{2j} + \dots + b_{2m}a_{mj} \\ \vdots \\ b_{p1}a_{1j} + b_{p2}a_{2j} + \dots + b_{pm}a_{mj} \end{pmatrix}.$$

4.4 EXERCÍCIOS

1. Considere as aplicações do \mathbb{R}^2

$$(a) \quad \begin{aligned} \Pi_1: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ \Pi_1(x_1, x_2) &= (x_1, 0) \end{aligned}$$

$$(b) \quad \begin{aligned} \Pi_2: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ \Pi_2(x_1, x_2) &= (0, x_2) \end{aligned}$$

$$(c) \quad \begin{aligned} R_1: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ R_1(x_1, x_2) &= (x_1, -x_2) \end{aligned}$$

$$(d) \quad \begin{aligned} R_2: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ R_2(x_1, x_2) &= (-x_1, x_2) \end{aligned}$$

Mostre que cada uma dessas aplicações é linear e encontre sua representação matricial (com relação à base canônica do \mathbb{R}^2). Interprete-as geometricamente.

2. Verifique se a aplicação linear definida pela matriz A é injetora, se

$$(a) \quad A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix}; \quad (b) \quad A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 1 \end{pmatrix}; \quad (c) \quad A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. Escreva a equação

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} + x_3 \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}$$

na forma de um sistema não homogêneo $Ax = b$. Resolva-o, se esse sistema possuir solução.

4. Encontre bases para o espaço linha e para o espaço coluna da matriz A , dada por

$$(a) \quad \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}; \quad (b) \quad \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad (c) \quad \begin{pmatrix} 1 & -2 & 1 \\ 2 & 1 & -2 \\ 0 & -1 & 1 \end{pmatrix}.$$

5. Encontre bases para o núcleo e imagem da matriz

$$\begin{pmatrix} 1 & 2 & 2 & 3 & 1 & 4 \\ 2 & 4 & 5 & 5 & 4 & 9 \\ 3 & 6 & 7 & 8 & 5 & 9 \end{pmatrix}.$$

Determine o posto dessa matriz.

6. Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear e $\{v_1, \dots, v_n\}$ uma base do \mathbb{R}^n . Mostre que $\text{im } T$ é gerada pelos vetores Tv_1, Tv_2, \dots, Tv_n .
7. Seja $A = (c_1 \ c_2 \ \dots \ c_n)$ uma matriz quadrada, dada por suas colunas. Mostre que $\{c_1, \dots, c_n\}$ é linearmente dependente se, e somente se, $\det A = 0$.
8. Sejam A, B matrizes $n \times n$. Mostre que $AB = 0$ se, e somente se, o espaço coluna de B estiver contido no núcleo de A .

AULA 5

O Teorema do Núcleo e da Imagem

OBJETIVOS

No final desta aula, o aluno deverá ser capaz de:

1. aplicar o teorema (da dimensão) do núcleo e da imagem;
2. obter a inversa de uma matriz por meio da forma escalonada reduzida por linhas.

5.1 TEOREMA (DA DIMENSÃO) DO NÚCLEO E DA IMAGEM

Nesta seção mostraremos um dos resultados mais importantes da Álgebra Linear, com implicações importantes no estudo de sistemas lineares.

Teorema 5.1 (do Núcleo e da Imagem)

Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear. Então,

$$n = \dim \ker T + \dim \text{im } T.$$

Demonstração: Se $\ker T \neq \{0\}$, seja $\{x_1, \dots, x_j\}$ uma base de $\ker T$. Esse é um conjunto linearmente independente no espaço \mathbb{R}^n , de modo que podemos aplicar o Teorema 3.43 e obter uma base

$$\mathcal{B} = \{x_1, \dots, x_j, w_{j+1}, \dots, w_n\}$$

do \mathbb{R}^n . (Se $\ker T = \{0\}$, simplesmente tomamos uma base do \mathbb{R}^n .)

Afirmamos que $\{Tw_{j+1}, \dots, Tw_n\}$ é uma base de $\text{im } T \subset Y$. De fato, suponhamos que

$$\alpha_{j+1}Tw_{j+1} + \dots + \alpha_nTw_n = 0.$$

Daí decorre que $T(\alpha_{j+1}w_{j+1} + \dots + \alpha_nw_n) = 0$; definindo

$$w = \alpha_{j+1}w_{j+1} + \dots + \alpha_nw_n,$$

concluímos que $w \in \ker T$. Como $\{x_1, \dots, x_j\}$ é uma base de $\ker T$, temos que

$$w = \beta_1x_1 + \dots + \beta_jx_j.$$

Ou seja,

$$\beta_1x_1 + \dots + \beta_jx_j - \alpha_{j+1}w_j - \dots - \alpha_nw_n = 0.$$

Como \mathcal{B} é uma base do \mathbb{R}^n , concluímos que $\beta_1 = \dots = \beta_j = -\alpha_{j+1} = \dots = -\alpha_n = 0$. Isso mostra que os vetores Tw_{j+1}, \dots, Tw_n são linearmente independentes.

Seja agora $y \in \text{im } T$. Então existe $x \in \mathbb{R}^n$ tal que $Tx = y$. Como \mathcal{B} é base de \mathbb{R}^n , $x = \alpha_1x_1 + \dots + \alpha_jx_j + \alpha_{j+1}w_{j+1} + \dots + \alpha_nw_n$ e, portanto,

$$\begin{aligned} y = Tx &= \alpha_1Tx_1 + \dots + \alpha_jTx_j + \alpha_{j+1}Tw_{j+1} + \dots + \alpha_nTw_n \\ &= \alpha_{j+1}Tw_{j+1} + \dots + \alpha_nTw_n, \end{aligned}$$

pois $Tx_i = 0$, $i \in \{1, \dots, j\}$. Isso mostra que $\{Tw_{j+1}, \dots, Tw_n\}$ gera $\text{im } T$ e conclui a prova. \square

5.2 ISOMORFISMO E INVERSAS

Definição 5.2 Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação *linear*. Dizemos que T é um **isomorfismo** se T for uma bijeção.¹

Em geral, não é fácil verificar que uma função $f: X \rightarrow Y$ é injetora ou sobrejetora. No caso de aplicações lineares, essa verificação é muito mais simples, como veremos.

Vamos mostrar que, se $n \neq m$, então $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ não é um isomorfismo.

Lema 5.3 Se $m > n$, então $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ não é sobrejetora.

Demonstração: Pelo Teorema do Núcleo e da Imagem, temos que

$$\dim \text{im } T = n - \dim \ker T \leq n.$$

Como $n < m$, não podemos ter $\dim \text{im } T = m$. \square

Exemplo 5.4 Considere a aplicação linear $S: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dada por

$$S(x_1, x_2) = (x_1, x_2, x_1 + x_2).$$

É fácil verificar que S é injetora. Mas S não é sobrejetora, pois não existe $x = (x_1, x_2) \in \mathbb{R}^2$ tal que $Sx = (1, 1, 0)$. \triangleleft

Lema 5.5 Uma aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ é injetora se, e somente se, $\ker T = \{0\}$.

Demonstração: Se existir $x \neq 0$ tal que $Tx = 0$, então, T não é injetora, pois também temos $T(0) = 0$.

Suponhamos agora $\ker T = \{0\}$ e que existam x_1, x_2 tais que $Tx_1 = Tx_2$. Daí decorre que $Tx_1 - Tx_2 = 0$, ou seja, $T(x_1 - x_2) = 0$. A última igualdade garante que $x_1 - x_2 \in \ker T$, isto é, $x_1 - x_2 = 0$ e, portanto, $x_1 = x_2$. Isso mostra que T é injetora. \square

Exercício 5.6 Seja $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $T(x, y) = (2x + y, 3x + 2y)$. Mostre que T é injetora. Conclua que T é um isomorfismo. Obtenha a expressão de T^{-1} .

Corolário 5.7 Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear. Se $m < n$, então, $\ker T \neq \{0\}$.

Demonstração: Pelo Teorema do Núcleo e da Imagem, temos

$$n = \dim \ker T + \dim \text{im } T \leq \dim \ker T + m,$$

pois $\dim \text{im } T \leq m$. Como $m < n$, temos $\dim \ker T \geq 1$. \square

Observe que o Corolário 5.7 nada mais é do que uma paráfrase do Teorema 1.19. De fato, a aplicação T pode ser representada por uma matriz A , $m \times n$. Pelo Teorema 1.19, $Ax = 0$ possui infinitas soluções. Ou seja, $\dim \ker A \geq 1$ ou, o que é o mesmo, $\dim \ker T \geq 1$.

Resulta imediatamente dos Lemas 5.3 e 5.5 que

Teorema 5.8 Uma aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ só pode ser um isomorfismo se tivermos $n = m$.

Exemplo 5.9 A aplicação linear $S: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $S(x_1, x_2, x_3) = (x_1, x_2, 0)$ não é uma bijeção, pois $S(0, 0, 1) = (0, 0, 0)$. \triangleleft

A verificação de que uma aplicação $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ é um isomorfismo é bastante simples:

Teorema 5.10 Uma aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ é injetora se, e somente se, for sobrejetora.

Demonstração: De fato, se tivermos $\ker T = \{0\}$, então

$$\dim \text{im } T = n - \dim \ker T = n.$$

Como a dimensão da imagem de T é igual a dimensão do contradomínio de T , temos que T é sobrejetora. Da mesma forma, se T for sobrejetora, então $\dim \text{im } T = n$ e daí deduzimos, como antes, $\dim \ker T = 0$. \square

A formulação do Teorema 5.10 em termos de sistemas lineares é a seguinte:

Corolário 5.11 Seja A uma matriz $n \times n$. Então o sistema não homogêneo $Ax = b$ tem solução única para todo $b \in Y$ se, e somente se, o sistema homogêneo $Ax = 0$ tiver solução única.

(Note que o Corolário 5.11 nada mais é do que o Teorema 1.21.)

De acordo com o Corolário 4.26, temos $\dim \text{im } A = \dim \text{im } A^t$ para toda matriz A . Em geral, temos $\dim \ker A \neq \dim \ker A^t$, a igualdade sendo válida apenas para matrizes quadradas:

Corolário 5.12 Seja A uma matriz $n \times n$. Então,

$$\dim(\ker A) = \dim(\ker A^t).$$

Demonstração: De fato, denotando $r := \dim(\text{im } A) = \dim(\text{im } A^t)$, a aplicação do Teorema do Núcleo e da Imagem garante que:

$$\dim(\ker A) = n - r \quad \text{e} \quad \dim(\ker A^t) = n - r.$$

Daí decorre o afirmado. \square

Suponhamos que $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ seja um isomorfismo. (Isso quer dizer, em particular, que T é linear.) Sabemos que $T^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ existe. Também temos que T^{-1} é uma aplicação linear?

Proposição 5.13 Suponhamos que $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ seja um isomorfismo. Então $T^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma aplicação linear.

Demonstração: Como T é linear, vale $T(x_1 + \alpha x_2) = Tx_1 + \alpha Tx_2$. Aplicando T^{-1} em ambos os lados dessa igualdade, obtemos

$$x_1 + \alpha x_2 = T^{-1}(Tx_1 + \alpha Tx_2).$$

Denotando $z_1 = Tx_1$ e $z_2 = Tx_2$, temos $T^{-1}z_1 = x_1$, $T^{-1}z_2 = x_2$ e

$$T^{-1}(z_1 + \alpha z_2) = x_1 + \alpha x_2 = T^{-1}z_1 + \alpha T^{-1}z_2,$$

provando a linearidade de T^{-1} . \square

Observação 5.14 Um isomorfismo $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ (veja o Teorema 5.8) corresponde a um “tradutor perfeito” entre o domínio de T (que é todo o \mathbb{R}^n) e a imagem de T (que também é todo o \mathbb{R}^n). O resultado de cada operação realizada no domínio de T corresponde, por meio de T , à mesma operação realizada na imagem de T , por meio da lei $T(x_1 + \alpha x_2) = Tx_1 + \alpha Tx_2$. Além disso, o resultado de cada operação realizada em $\text{im } T$ corresponde à mesma operação realizada no domínio de T , por meio da lei: $T^{-1}(y_1 + \alpha y_2) = T^{-1}y_1 + \alpha T^{-1}y_2$.

Comparemos com uma bijeção simples: $f: \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = x^3$. Para verificar que f é injetora temos que provar que $f(x_1) = f(x_2)$ implica $x_1 = x_2$. Ou seja, de $x_1^3 = x_2^3$ devemos obter $x_1 = x_2$. Ora,

$$x_1^3 = x_2^3 \Leftrightarrow \sqrt[3]{x_1^3} = \sqrt[3]{x_2^3} \Leftrightarrow x_1 = x_2.$$

Seja agora $y \in \mathbb{R}$. Queremos provar que $y \in \text{im } f$, ou seja, que existe $x \in \mathbb{R}$ tal que $f(x) = y$. Essa igualdade é o mesmo que $x^3 = y$, de onde vem $x = \sqrt[3]{y}$. Assim, f é uma bijeção. Mas $8 = f(2) = f(1+1) \neq f(1) + f(1) = 1^3 + 1^3 = 1 + 1 = 2$.

Para verificar que uma aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma bijeção, não precisamos verificar que T é injetora e sobrejetora: basta verificar que T é injetora, de acordo com Teorema 5.10! \triangleleft

Exercício 5.15 Na demonstração do Teorema do Núcleo e da Imagem 5.1, mostre que T estabelece um isomorfismo entre $\langle w_{k+1}, \dots, w_n \rangle$ e $\text{im } T$.

5.3 OBTENÇÃO DA INVERSA DE UMA MATRIZ

Dado um isomorfismo $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, como obter a aplicação inversa $T^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}^n$? Em geral, é mais fácil obter uma matriz A que representa T e então a inversa da matriz A . Relembreamos:

Definição 5.16 Uma matriz B é inversa da matriz quadrada A se

$$AB = BA = I,$$

em que I é a matriz identidade. Se A possuir inversa, dizemos que A é invertível.

É fácil ver que uma matriz invertível A possui uma única inversa, que denotaremos por A^{-1} . De fato, suponhamos que $AB = BA = I$ e $AC = CA = I$. Queremos mostrar que $B = C$. Temos

$$B = B(AC) = (BA)C = IC = C,$$

provando o afirmado.

Proposição 5.17 Seja A uma matriz quadrada. Suponhamos que exista uma matriz quadrada B tal que $BA = I$. Então

$$AB = I,$$

de modo que $B = A^{-1}$ é a inversa de A .

Demonstração: Da igualdade $BA = I$ deduzimos que $\ker A = \{0\}$. De fato, se existisse $x \neq 0$ tal que $Ax = 0$, então $BAx = B(0) = 0$, o que contradiz $BAx = Ix = x \neq 0$. Como A é uma matriz quadrada, isso quer dizer que A é sobrejetora, de acordo com o Teorema 5.10. Portanto, A possui inversa A^{-1} . Como $BA = I$, multiplicando essa igualdade à direita por A^{-1} , vem $BAA^{-1} = A^{-1}$, ou seja, provamos que $B = A^{-1}$. \square

O cálculo da inversa de uma matriz por meio da adjunta (clássica) da matriz quadrada A não é um método eficiente. Vamos mostrar uma maneira mais simples de obter a inversa de uma matriz quadrada.

Decorre do Lema 5.5 que uma matriz quadrada A é invertível se, e somente se, $\ker A = \{0\}$. Em termos da forma escalonada reduzida por linhas da matriz A , isto quer dizer que não existem variáveis livres, pois a existência destas garante a existência de infinitos elementos x tais que $Ax = 0$. (Veja, a esse respeito, a Observação 1.12.) Uma vez que a forma escalonada R da matriz $n \times n$ A possui n pivôs, vemos que $R = I$, em que I é a matriz identidade.

Definição 5.18 Uma matriz E é **elementar** se puder ser obtida da matriz identidade $m \times m$ por meio da aplicação de uma operação elementar.

O próximo resultado mostra que a aplicação de uma operação elementar sobre as linhas da matriz A é equivalente à multiplicação dessa matriz por uma matriz elementar.

Proposição 5.19 Seja e uma operação elementar sobre (as linhas de) a matriz A , $m \times n$ e E a matriz elementar igual a $e(I)$, sendo I a matriz identidade $m \times m$. Então $e(A) = EA$.

Demonstração: A demonstração deve ser feita para todos os tipos de operação elementar. Consideraremos apenas a aplicação de uma operação elementar (c): a linha j será substituída pela soma da linha j com λ vezes a linha i . A matriz E , nesse caso, é dada por

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \lambda & \dots & 1 & \dots & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \dots & & & & 1 \end{pmatrix} \leftarrow \text{linha } j$$

↑
coluna j

Então,

$$\begin{aligned} EA &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \lambda & \dots & 1 & \dots & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \dots & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{j1} + \lambda a_{i1} & a_{j2} + \lambda a_{i2} & \dots & a_{jn} + \lambda a_{in} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \end{aligned}$$

que é justamente $e(A)$. \square

Ora, uma sequência de operações elementares produz a forma escalonada reduzida por linhas da matriz invertível A , isto é, produz a matriz identidade. Quer dizer,

$$e_k e_{k-1} \dots e_1(A) = I.$$

Exercício 5.20 Justifique a afirmação anterior. Isto é, mostre que a forma escalonada reduzida por linhas de uma matriz $n \times n$ invertível é a matriz identidade $n \times n$.

Em termos das matrizes elementares, isso quer dizer que

$$E_k E_{k-1} \dots E_1 A = I.$$

Definindo a matriz $B = E_k E_{k-1} \dots E_1$, acabamos de mostrar que $BA = I$. Mas isso significa que $B = A^{-1}$, de acordo com a Proposição 5.17. Para obter B , basta então aplicar a sequência de operações elementares $e_k \dots e_1$ à matriz I , pois

$$e_k \dots e_1(A) = E_k \dots E_1 I = E_k \dots E_1 = B.$$

Exemplo 5.21 Para obter a inversa da matriz

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 4 & 0 \\ 2 & 3 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

consideramos $(A|I)$, em que I é a matriz identidade 4×4 e levamos a matriz A à sua forma escalonada reduzida por linhas:

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 4 & 0 & 0 & 1 & 0 & 0 \\ 2 & 3 & 5 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & 2 & -1 & 0 & 0 \\ 0 & 1 & 3 & 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \rightarrow$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 3 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 5 & 0 & -4 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 3 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -\frac{4}{5} & \frac{1}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \rightarrow$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \frac{7}{5} & \frac{2}{5} & -\frac{3}{5} & 0 \\ 0 & 1 & 0 & 0 & \frac{2}{5} & -\frac{3}{5} & \frac{2}{5} & 0 \\ 0 & 0 & 1 & 0 & -\frac{4}{5} & \frac{1}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right)$$

A matriz do lado direito é a inversa da matriz A . \square

5.4 EXERCÍCIOS

1. Dê exemplos de aplicações lineares $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ tais que
 - $\ker T = \{x \in \mathbb{R}^3 : x_3 = -x_1\}$;
 - $\text{im } T = \{x \in \mathbb{R}^3 : x_1 = x_2\}$.
2. Seja $\ell: \mathbb{R}^3 \rightarrow \mathbb{R}$ uma aplicação linear.
 - Mostre que existem escalares a, b e c tais que $\ell(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$;
 - descreva geometricamente todas as possibilidades para $\ker \ell$.
3. Verdadeiro ou falso? Se for verdadeiro, justifique. Se for falso, dê um contraexemplo.
Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear.
 - Uma base de $\text{im } T$ é obtida ao se completar uma base de $\ker T$, de modo a ter uma base do \mathbb{R}^n e então calcular a imagem desses vetores adicionais por T ;
 - uma base de $\text{im } T$ é Te_1, \dots, Te_n ;
 - toda função $f: \mathbb{R} \rightarrow \mathbb{R}$ é bijetora.
4. Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear. Mostre que T é injetora se, e somente se, a imagem de todo conjunto linearmente independente (no \mathbb{R}^n) for um conjunto linearmente independente (no \mathbb{R}^m).
5. Encontre, se possível, as inversas das seguintes matrizes:

$$(a) \begin{pmatrix} 1 & 2 & 2 \\ 1 & 3 & 1 \\ 1 & 3 & 2 \end{pmatrix}; \quad (b) \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 3 & 3 & 2 \end{pmatrix}; \quad (c) \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ 5 & 9 & 1 & 6 \end{pmatrix}.$$

6. Encontre todos os valores de a para os quais a matriz

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & a \end{pmatrix}$$

possui inversa.

AULA 6

Mudanças de Base

OBJETIVOS

No final desta aula, o aluno deverá ser capaz de:

1. obter as coordenadas de um vetor em diferentes bases;
2. obter a representação matricial de uma aplicação linear em diferentes bases;
3. saber como essas representações matriciais se relacionam.

6.1 REPRESENTAÇÃO DE UM VETOR EM UMA BASE

Como dissemos, uma base no \mathbb{R}^n equivale a um sistema referencial, isto é, a um sistema de coordenadas. Mudando o sistema de coordenadas, mudam as coordenadas do ponto, mas o ponto permanece o mesmo.

Exemplo 6.1 Sejam $v_1 = (1, 1)$ e $v_2 = (-1, 1)$. É claro que $\mathcal{B} = \{v_1, v_2\}$ é uma base do \mathbb{R}^2 . Consideremos $x = (2, 2) \in \mathbb{R}^2$. Na base \mathcal{B} , temos $x = 2v_1 + 0v_2$. Quer dizer, as coordenadas de x na base \mathcal{B} são 2 e 0. \triangleleft

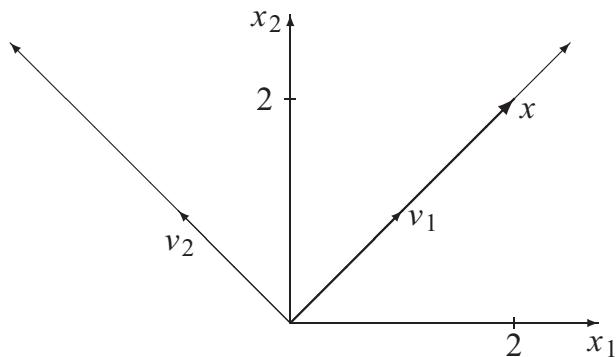


Figura 6.1: As coordenadas de $x = (2, 2)$ na base \mathcal{B} são 2 e 0, pois $x = 2v_1 + 0v_2$

O exemplo anterior nos coloca algumas questões. Em primeiro lugar, como representar as coordenadas do ponto x na base \mathcal{B} ? E, considerada uma base \mathcal{B} do \mathbb{R}^n , como obter as coordenadas do ponto x na base \mathcal{B} ?

Definição 6.2 Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base do \mathbb{R}^n . Se $x \in \mathbb{R}^n$, então existem (únicos) escalares $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tais que

$$x = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

O vetor

$$[x]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 e_1 + \dots + \lambda_n e_n \in \mathbb{R}^n$$

é chamado **representação** de x na base \mathcal{B} e $\lambda_1, \dots, \lambda_n$ as **coordenadas** de x na base \mathcal{B} .

Se \mathcal{E} for a base canônica do \mathbb{R}^n , escrevemos simplesmente x ao invés de $[x]_{\mathcal{E}}$.

Observação 6.3 A Definição 6.2 merece alguns comentários. Em primeiro lugar, ao escrever o vetor x como combinação linear dos vetores da base \mathcal{B} , estamos justamente encontrando as coordenadas do vetor x no sistema de coordenadas formado por \mathcal{B} .

Um segundo fato precisa ser ressaltado: uma base do \mathbb{R}^n é um conjunto ordenado. Apenas essa ordenação permite dar sentido à representação de um vetor em uma base. \triangleleft

Exemplo 6.4 Seja $x = (3, 2) \in \mathbb{R}^2$. Considere a base

$$\mathcal{B} = \{v_1 = (1, 1), v_2 = (-1, 1)\}$$

do \mathbb{R}^2 . Para encontrar $[x]_{\mathcal{B}}$, resolvemos o sistema

$$x = \lambda_1 v_1 + \lambda_2 v_2,$$

isto é,

$$(3, 2) = \lambda_1(1, 1) + \lambda_2(-1, 1).$$

Como o sistema tem apenas duas incógnitas, não é necessário utilizar o método de Gauss-Jordan. Contudo, vamos aplicar esse método, pois ele é bastante útil em sistemas maiores. Levamos a matriz aumentada do sistema anterior à forma escalonada reduzida por linhas:¹

$$\left(\begin{array}{cc|c} 1 & -1 & 3 \\ 1 & 1 & 2 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} 1 & -1 & 3 \\ 0 & 2 & -1 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} 1 & 0 & \frac{5}{2} \\ 0 & 1 & -\frac{1}{2} \end{array} \right).$$

Assim,

$$[x]_{\mathcal{B}} = \begin{pmatrix} \frac{5}{2} \\ -\frac{1}{2} \end{pmatrix}.$$

Proposição 6.5 Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base do \mathbb{R}^n . Se $x = \lambda_1 v_1 + \dots + \lambda_n v_n$, a aplicação $B: \mathbb{R}^n \rightarrow \mathbb{R}^n$ dada por

$$Bx = [x]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 e_1 + \dots + \lambda_n e_n \in \mathbb{R}^n$$

¹ Repetindo: os vetores dados aparecem como *colunas* da matriz!

é um isomorfismo. Para ressaltar que a imagem de B está sendo considerada com a base canônica, escreveremos $B: \mathbb{R}^n \rightarrow (\mathbb{R}^n, \mathcal{E})$.

Demonstração: Se $x = \lambda_1 v_1 + \dots + \lambda_n v_n$ e $y = \gamma_1 v_1 + \dots + \gamma_n v_n$, então,

$$B(x + \alpha y) = B((\lambda_1 + \alpha \gamma_1)v_1 + \dots + (\lambda_n + \alpha \gamma_n)v_n)$$

$$= \begin{pmatrix} \lambda_1 + \alpha \gamma_1 \\ \lambda_2 + \alpha \gamma_2 \\ \vdots \\ \lambda_n + \alpha \gamma_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} + \alpha \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix} = Bx + \alpha By,$$

mostrando a linearidade de B .

Se $Bx = 0$, então $x = 0v_1 + \dots + 0v_n = 0$, mostrando que $\ker B = \{0\}$. O resultado decorre então do Teorema 5.10. \square

Observação 6.6 Uma vez que $v_i = 0v_1 + \dots + 0v_{i-1} + 1v_i + 0v_{i+1} + \dots + 0v_n$, vemos que

$$Bv_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = e_i,$$

o i -ésimo vetor da base canônica do \mathbb{R}^n . \square

6.2 APLICAÇÕES LINEARES E MATRIZES – PARTE II

Na primeira seção da Aula 4 mostramos como associar a cada aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma matriz $A = (a_{ij})$, que representa T com relação às bases canônicas do \mathbb{R}^n e \mathbb{R}^m . Mostraremos agora que a mesma associação entre aplicações lineares e matrizes é válida para quaisquer escolhas de bases \mathcal{B} do espaço \mathbb{R}^n e \mathcal{C} do espaço \mathbb{R}^m .

Escolhendo uma base arbitrária $\mathcal{B} = \{v_1, \dots, v_n\}$ do espaço \mathbb{R}^n e escrevendo $x = \lambda_1 v_1 + \dots + \lambda_n v_n$, já vimos que a aplicação

$$B: \mathbb{R}^n \rightarrow (\mathbb{R}^n, \mathcal{E})$$

definida por

$$Bx = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$$

é um isomorfismo.

Da mesma forma, ao se escolher uma base $\mathcal{C} = \{w_1, \dots, w_m\}$ no espaço \mathbb{R}^m , obtém-se um isomorfismo C entre \mathbb{R}^m e \mathbb{R}^m . (Note que as imagens das aplicações B e C chegam, respectivamente, no \mathbb{R}^n e \mathbb{R}^m , esses espaços considerados com suas bases canônicas.)

Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear. Vamos considerar uma base \mathcal{B} no espaço \mathbb{R}^n e uma base \mathcal{C} no espaço \mathbb{R}^m . Nosso objetivo é mostrar como associar uma matriz à aplicação T , levando em conta as bases escolhidas \mathcal{B} e \mathcal{C} .

Temos o seguinte diagrama (as setas verticais *sempre* indicarão isomorfismos):

$$\begin{array}{ccc} & T & \\ \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ B \downarrow & & \downarrow C \\ (\mathbb{R}^n, \mathcal{E}) & \longrightarrow & (\mathbb{R}^m, \mathcal{E}) \\ A & & \end{array} \quad (6.1)$$

em que o mesmo símbolo \mathcal{E} está representando as bases canônicas dos espaços \mathbb{R}^n e \mathbb{R}^m .

A aplicação linear A é definida como composta de aplicações lineares (estamos usando a notação de composta para enfatizar)

$$A = C \circ T \circ B^{-1}$$

e é representada por uma matriz, de acordo com o que vimos na primeira seção da Aula 4. É usual chamar essa matriz de *representação da aplicação linear* T com respeito às bases \mathcal{B} e \mathcal{C} (dos espaços \mathbb{R}^n e \mathbb{R}^m , respectivamente) e denotá-la por $T_{\mathcal{B}}^{\mathcal{C}}$. Veja que $A = T_{\mathcal{B}}^{\mathcal{C}}$ é uma aplicação linear de $(\mathbb{R}^n, \mathcal{E})$ no espaço $(\mathbb{R}^m, \mathcal{E})$.

Exemplo 6.7 Consideremos os espaços \mathbb{R}^n e \mathbb{R}^m com as bases $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_m\}$, respectivamente. Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação linear. Vejamos como obter $T_{\mathcal{B}}^{\mathcal{C}}$. Para isso, usamos o diagrama (6.1):

$$\begin{array}{ccc} & T & \\ \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ B \downarrow & & \downarrow C \\ (\mathbb{R}^n, \mathcal{E}) & \longrightarrow & (\mathbb{R}^m, \mathcal{E}) \\ T_{\mathcal{B}}^{\mathcal{C}} & & \end{array} .$$

Como vimos na primeira seção da Aula 4, a i -ésima coluna da matriz procurada é obtida ao se calcular $T_{\mathcal{B}}^{\mathcal{C}} e_i = (CTB^{-1})e_i$. Mas, $Bv_i = e_i$, de modo que $(CTB^{-1})e_i = (CT)B^{-1}e_i = (CT)v_i$. Como C é a aplicação que associa a $Tv_i \in \mathbb{R}^m$ às suas coordenadas na base \mathcal{C} , temos que a i -ésima coluna da matriz procurada é $[Tv_i]_{\mathcal{C}}$. \triangleleft

No caso de aplicações lineares $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, a mesma base \mathcal{B} pode ser escolhida tanto no domínio quanto no contradomínio. Nesse caso, denotamos $T_{\mathcal{B}}$ ao invés de $T_{\mathcal{B}}^{\mathcal{B}}$.

Exemplo 6.8 Considere a aplicação linear $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por

$$T(x, y) = (4x - 2y, 2x + y).$$

Seja \mathcal{B} a base do \mathbb{R}^2 formada pelos vetores $v_1 = (1, 1)$ e $v_2 = (-1, 0)$. Vamos achar a matriz que representa T com relação à base \mathcal{B} . (Quer dizer, estamos utilizando a mesma base no domínio e no contradomínio e procuramos a matriz $T_{\mathcal{B}}$.) Para isso, calculamos

$$T(v_1) = (2, 3) = 3(1, 1) + (-1, 0) = 3v_1 + v_2.$$

Note que escrevemos a imagem de $T(v_1)$ na base \mathcal{B} , utilizada também no contradomínio. De acordo com a notação introduzida na Definição 6.2, temos,

$$[T(v_1)]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Da mesma forma, $T(v_2) = (-4, -2) = -2(1, 1) + 2(-1, 0) = -2v_1 + 2v_2$ e, portanto,

$$[T(v_2)]_{\mathcal{B}} = \begin{pmatrix} -2 \\ 2 \end{pmatrix}.$$

Assim,

$$T_{\mathcal{B}} = ([Tv_1]_{\mathcal{B}} \ [Tv_2]_{\mathcal{B}}) = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}.$$

As colunas de $T_{\mathcal{B}}$ são as imagens dos vetores da base \mathcal{B} , escritas na própria base \mathcal{B} utilizada, nesse caso, também no contradomínio.

Se quisermos calcular a imagem do vetor $(1, 2) = 1e_1 + 2e_2 \in \mathbb{R}^2$ utilizando a matriz $T_{\mathcal{B}}$, primeiro expressamos esse vetor na base \mathcal{B} :

$$(1, 2) = 2(1, 1) + 1(-1, 0) = 2v_1 + v_2.$$

Calculando

$$T_{\mathcal{B}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix},$$

obtemos a “resposta” na base \mathcal{B} . Se quisermos a resposta na base canônica, precisamos escrever o resultado obtido nesta base:

$$4v_1 + 4v_2 = 4(1, 1) + 4(-1, 0) = (0, 4) = 0e_1 + 4e_2,$$

que é o mesmo resultado que obtemos ao calcular $T(1, 2)$ utilizando a expressão $T(x, y) = (4x - 2y, 2x + y)$. \triangleleft

Escolhidas as bases \mathcal{B} do \mathbb{R}^n e \mathcal{C} do \mathbb{R}^m , associamos, assim, a cada aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ a matriz $T_{\mathcal{B}}^{\mathcal{C}}$, cuja expressão depende das bases \mathcal{B} e \mathcal{C} . Uma vez que cada escolha de base \mathcal{B} no \mathbb{R}^n produz um isomorfismo diferente entre $(\mathbb{R}^n, \mathcal{B})$ e $(\mathbb{R}^n, \mathcal{E})$ e o mesmo acontece com $(\mathbb{R}^m, \mathcal{C})$ e $(\mathbb{R}^m, \mathcal{E})$, vemos que existem muitas maneiras distintas de representar uma transformação linear por meio de uma matriz. Como se relacionam essas diferentes matrizes que representam a aplicação linear T ?

Para responder a essa pergunta, começamos estudando como se relacionam as representações de x em bases $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\bar{\mathcal{B}} = \{v'_1, \dots, v'_n\}$ do espaço \mathbb{R}^n . O mesmo procedimento anterior pode ser utilizado (I denota a aplicação identidade):

$$\begin{array}{ccc} & I & \\ \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ B & \downarrow & \downarrow \bar{B} \\ (\mathbb{R}^n, \mathcal{E}) & \longrightarrow & (\mathbb{R}^n, \mathcal{E}) \\ P_{\mathcal{B}}^{\bar{\mathcal{B}}} & & \end{array}.$$

(Para sermos coerentes com a notação anterior, deveríamos escrever $I_{\mathcal{B}}^{\bar{\mathcal{B}}}$ ao invés de $P_{\mathcal{B}}^{\bar{\mathcal{B}}}$. Entretanto, é usual denotar esse tipo de matriz pela letra P .)

De acordo com o Exemplo 6.7, a i -ésima coluna de $P_{\mathcal{B}}^{\bar{\mathcal{B}}}$ é obtida calculando-se a expressão de $\bar{B}IB^{-1}(e_i) = \bar{B}I(v_i) = \bar{B}(v_i) = [v_i]_{\bar{\mathcal{B}}}$. A matriz $P_{\mathcal{B}}^{\bar{\mathcal{B}}}$ é chamada *matriz mudança*² da base \mathcal{B} para a base $\bar{\mathcal{B}}$. Dadas as coordenadas de x na base \mathcal{B} , isto é, $[x]_{\mathcal{B}}$, as coordenadas de x na base $\bar{\mathcal{B}}$ são dadas por

$$P_{\mathcal{B}}^{\bar{\mathcal{B}}} [x]_{\mathcal{B}} = [x]_{\bar{\mathcal{B}}}. \quad (6.2)$$

Claramente a matriz $P_{\mathcal{B}}^{\bar{\mathcal{B}}}$ possui inversa $P_{\mathcal{B}}^{\mathcal{B}}$.

² Alguns autores preferem chamar essa matriz de “matriz de passagem” da base \mathcal{B} para a base $\bar{\mathcal{B}}$. Assim, a terminologia utilizada por eles fica invertida com relação à nossa.

Exercício 6.9 Utilizando o diagrama anterior, justifique a última afirmação.

Exemplo 6.10 Voltemos ao Exemplo 6.8. Vamos obter a matriz $P_{\mathcal{E}}^{\mathcal{B}}$, matriz mudança da base \mathcal{E} para a base \mathcal{B} . Ela é calculada pelo mesmo método: escrevemos a imagem dos vetores e_1, e_2 pela aplicação identidade na base \mathcal{B} . Temos

$$(1, 0) = 0(1, 1) - 1(-1, 0) = 0v_1 - v_2$$

e

$$(0, 1) = 1(1, 1) + 1(-1, 0) = 1v_1 + 1v_2.$$

A matriz $P_{\mathcal{E}}^{\mathcal{B}}$ é, então,

$$P_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}.$$

□

Exercício 6.11 Continuando o Exemplo 6.8, calcule a matriz $P_{\mathcal{B}}^{\mathcal{E}}$. Verifique que $(P_{\mathcal{E}}^{\mathcal{B}})^{-1} = P_{\mathcal{B}}^{\mathcal{E}}$.

É usual condensar o diagrama (6.1) e escrever

$$(\mathbb{R}^n, \mathcal{B}) \xrightarrow{T_{\mathcal{B}}^{\mathcal{C}}} (\mathbb{R}^m, \mathcal{C}),$$

salientando as bases utilizadas para se produzir a matriz $T_{\mathcal{B}}^{\mathcal{C}}$. Note, contudo, que $(\mathbb{R}^n, \mathcal{B})$ é apenas uma notação para o espaço \mathbb{R}^n com a base canônica!

Passemos então à pergunta feita anteriormente, utilizando essa nova notação: como se relacionam as matrizes $T_{\mathcal{B}}^{\mathcal{C}}$ e $T_{\mathcal{B}}^{\tilde{\mathcal{C}}}$? Temos o diagrama

$$\begin{array}{ccc} T_{\mathcal{B}}^{\mathcal{C}} & & \\ (\mathbb{R}^n, \mathcal{B}) & \longrightarrow & (\mathbb{R}^m, \mathcal{C}) \\ \downarrow & & \downarrow \\ P_{\mathcal{B}}^{\mathcal{C}} & & Q_{\mathcal{C}}^{\mathcal{C}} \\ (\mathbb{R}^n, \tilde{\mathcal{B}}) & \longrightarrow & (\mathbb{R}^m, \tilde{\mathcal{C}}) \\ T_{\mathcal{B}}^{\tilde{\mathcal{C}}} & & \end{array}.$$

Esse diagrama, cujas componentes são matrizes, nos mostra que

$$T_{\mathcal{B}}^{\mathcal{C}} = [Q_{\mathcal{C}}^{\mathcal{C}}]^{-1} T_{\mathcal{B}}^{\tilde{\mathcal{C}}} P_{\mathcal{B}}^{\mathcal{C}} = Q_{\mathcal{C}}^{\mathcal{C}} T_{\mathcal{B}}^{\tilde{\mathcal{C}}} P_{\mathcal{B}}^{\mathcal{C}}.$$

O caso em que $n = m$ permite que se tome a mesma base no domínio e contradomínio. Nesse caso, a relação entre $T_{\mathcal{B}}$ e $T_{\tilde{\mathcal{B}}}$ é dada por

$$T_{\tilde{\mathcal{B}}} = [P_{\tilde{\mathcal{B}}}^{\mathcal{B}}]^{-1} T_{\mathcal{B}} P_{\tilde{\mathcal{B}}}^{\mathcal{B}} = P_{\tilde{\mathcal{B}}}^{\mathcal{B}} T_{\mathcal{B}} P_{\tilde{\mathcal{B}}}^{\mathcal{B}},$$

para qualquer outra base $\tilde{\mathcal{B}}$ do \mathbb{R}^n .

Observação 6.12 Dada uma aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$, a escolha de bases \mathcal{B} e \mathcal{C} no \mathbb{R}^n pode fazer com que a representação matricial de T assuma formas bem gerais. Por exemplo, se T for invertível, $T_{\mathcal{B}}^{\mathcal{C}}$ pode ser a matriz identidade! Assim, a representação de T em bases completamente arbitrárias quase não nos passa informação relevante sobre a aplicação T . Voltaremos a esse assunto no próximo curso de Álgebra Linear. □

Exemplo 6.13 Voltemos ao Exemplo 6.8. Temos a aplicação linear $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por

$$T(x, y) = (4x - 2y, 2x + y).$$

A representação de T na base canônica é a matriz cujas colunas são

$$T(1, 0) = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \quad \text{e} \quad T(0, 1) = \begin{pmatrix} -2 \\ 1 \end{pmatrix},$$

ou seja,

$$T_{\mathcal{E}} = \begin{pmatrix} 4 & -2 \\ 2 & 1 \end{pmatrix}.$$

Por outro lado, no Exemplo 6.8 calculamos a matriz $T_{\mathcal{B}}$:

$$T_{\mathcal{B}} = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}.$$

Vamos entender a relação entre as matrizes $T_{\mathcal{E}}$ e $T_{\mathcal{B}}$. Para isso, analisamos o seguinte diagrama

$$\begin{array}{ccc} & T_{\mathcal{E}} & \\ (\mathbb{R}^2, \mathcal{E}) & \xrightarrow{\quad} & (\mathbb{R}^2, \mathcal{E}) \\ P_{\mathcal{E}}^{\mathcal{B}} \downarrow & & \downarrow P_{\mathcal{E}}^{\mathcal{B}} \\ (\mathbb{R}^2, \mathcal{B}) & \xrightarrow{\quad} & (\mathbb{R}^2, \mathcal{B}) \\ & T_{\mathcal{B}} & \end{array}.$$

A matriz $P_{\mathcal{E}}^{\mathcal{B}}$ é a matriz mudança da base \mathcal{E} para a base \mathcal{B} calculada no Exemplo 6.10:

$$P_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}.$$

O diagrama anterior garante que

$$T_{\mathcal{E}} = [P_{\mathcal{E}}^{\mathcal{B}}]^{-1} T_{\mathcal{B}} P_{\mathcal{E}}^{\mathcal{B}},$$

ou seja,

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

Se calcularmos a inversa da matriz $P_{\mathcal{E}}^{\mathcal{B}}$, verificaremos esse fato. Entretanto, é fácil obter $P_{\mathcal{B}}^{\mathcal{E}}$. Essa matriz tem como colunas a expressão dos vetores v_1 e v_2 na base canônica. Assim, é claro que

$$P_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Verifique que $P_{\mathcal{B}}^{\mathcal{E}} = [P_{\mathcal{E}}^{\mathcal{B}}]^{-1}$ e que $T_{\mathcal{E}} = P_{\mathcal{B}}^{\mathcal{E}} T_{\mathcal{B}} P_{\mathcal{E}}^{\mathcal{B}}$.

Note que a última igualdade é coerente: dado um vetor $x = [x]_{\mathcal{E}}$ (na base canônica), podemos calcular diretamente $T_{\mathcal{E}}x = T_{\mathcal{E}}[x]_{\mathcal{E}} = [Tx]_{\mathcal{E}} = Tx$, de acordo com a nossa convenção de denotar $[y]_{\mathcal{E}}$ simplesmente por y . Por outro lado, calculando pelo lado direito da igualdade, temos

$$P_{\mathcal{B}}^{\mathcal{E}} T_{\mathcal{B}} P_{\mathcal{E}}^{\mathcal{B}} x = P_{\mathcal{B}}^{\mathcal{E}} T_{\mathcal{B}} P_{\mathcal{B}}^{\mathcal{B}} [x]_{\mathcal{E}} = P_{\mathcal{B}}^{\mathcal{E}} T_{\mathcal{B}} [x]_{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{E}} [Tx]_{\mathcal{B}} = [Tx]_{\mathcal{E}} = Tx,$$

o que produz o mesmo resultado. \triangleleft

6.3 APLICAÇÃO: DIAGONALIZAÇÃO DE UMA MATRIZ

Definição 6.14 Seja A uma matriz quadrada. O **polinômio característico da matriz A** é o polinômio

$$p(z) = \det(zI - A).$$

Exemplo 6.15 Vamos calcular o polinômio característico da matriz

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

Para isso, devemos calcular o determinante da matriz

$$zI - A = z \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} z-1 & 1 \\ -2 & z-4 \end{pmatrix}.$$

Assim,

$$p(z) = \det \begin{pmatrix} z-1 & 1 \\ -2 & z-4 \end{pmatrix} = (z-1)(z-4) + 2 = z^2 - 5z + 6. \quad \square$$

Exercício 6.16 Calcule o polinômio característico $p(z)$ da matriz

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

Vamos agora mostrar que podemos definir o polinômio característico de uma aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$. Para isso, escolhemos uma base \mathcal{B} e calculamos a matriz $A = T_{\mathcal{B}}$. Sabemos calcular o polinômio característico da matriz A . Mas esse polinômio característico não depende da base \mathcal{B} escolhida? Vamos mostrar que não. Para isso, se \mathcal{C} for uma outra base do \mathbb{R}^n e se a matriz $B = T_{\mathcal{C}}$ for a representação de T na base \mathcal{C} , já vimos que $A = P^{-1}BP$, sendo $P = P_{\mathcal{B}}^{\mathcal{C}}$ a matriz de mudança da base \mathcal{B} para a base \mathcal{C} . Temos que

$$\begin{aligned} \det(zI - A) &= \det(P^{-1}(zI - B)P) \\ &= \det P^{-1} \det(zI - B) \det P = \det(zI - B) \det(P^{-1}P) \\ &= \det(zI - B), \end{aligned}$$

mostrando que qualquer representação de T numa base do \mathbb{R}^n possui o mesmo polinômio característico. Em consequência, podemos definir:

Definição 6.17 O **polinômio característico da aplicação linear**

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^n$$

é o polinômio característico de qualquer uma de suas representações matriciais $T_{\mathcal{B}}$, sendo \mathcal{B} uma base do \mathbb{R}^n .

Uma vez definido o polinômio característico de uma aplicação linear T , estamos em condições de definir autovalores e autovetores de T .

Definição 6.18 As raízes $\lambda \in \mathbb{K}$ do polinômio característico são os *autovalores* de T . Se λ for um autovalor de T , os elementos $0 \neq v \in \ker(\lambda I - T)$ são os *autovetores* de T associados ao autovalor λ .

Se $0 \neq v$ for um autovetor de T associado ao autovalor λ , então $(\lambda I - T)v = 0$, ou seja, $Tv = \lambda v$. Lembramos que, dada uma matriz quadrada A , existe $v \neq 0$ tal que $Av = 0$ se, e somente se, $\det A = 0$. Assim, os autovalores de T nos fornecem justamente os valores λ tais que $\lambda I - T$ não é invertível. Mais precisamente, se B for uma representação matricial de T e λ um autovalor de T , então a matriz $A = \lambda I - B$ não é invertível.

Dada uma aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, a obtenção de uma base \mathcal{B} do \mathbb{R}^n formada por autovetores de T resolve um problema importante: nessa base, a representação de T , isto é, $T_{\mathcal{B}}$, é particularmente simples.

De fato, se $\mathcal{B} = \{v_1, \dots, v_n\}$ for uma base de autovetores de $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ associados, respectivamente, aos autovalores $\lambda_1, \dots, \lambda_n$, então,

$$\begin{aligned} T_{\mathcal{B}} &= ([Tv_1]_{\mathcal{B}} [Tv_2]_{\mathcal{B}} \dots [Tv_n]_{\mathcal{B}}) \\ &= ([\lambda_1 v_1]_{\mathcal{B}} [\lambda_2 v_2]_{\mathcal{B}} \dots [\lambda_n v_n]_{\mathcal{B}}) \\ &= \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}. \end{aligned}$$

Exemplo 6.19 Seja $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por

$$T(x_1, x_2) = (x_1 + 4x_2, 2x_1 + 3x_2).$$

Encontramos o polinômio característico de T ao escolher uma de suas representações matriciais. Escolhendo a base canônica \mathcal{E} do \mathbb{R}^2 , encontramos

$$B = T_{\mathcal{E}} = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}.$$

O polinômio característico de T é, portanto,

$$\det(zI - B) = \det \begin{pmatrix} z-1 & -4 \\ -2 & z-3 \end{pmatrix} = z^2 - 4z - 5.$$

Os autovalores de T são as raízes de $p(z) = z^2 - 4z - 5$, que são dadas por $\lambda = -1$ e $\mu = 5$. Para encontrarmos os autovetores associados a $\lambda = -1$, basta resolvemos o sistema $(-I - B)v = 0$, obtido ao substituir $z = \lambda = -1$ em $(zI - B)$. Logo, temos

$$\begin{pmatrix} -2 & -4 \\ -2 & -4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix},$$

que nos fornece a solução

$$v = \alpha \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Tomamos $\alpha = 1$ e o vetor

$$v_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix},$$

pois qualquer outro autovetor associado a $\lambda = -1$ é múltiplo de v_1 .

Do mesmo modo, encontramos os autovetores associados a $\mu = 5$ resolvendo o sistema $(5I - B)v = 0$, obtido ao substituir $z = \mu = 5$ em $zI - B$, ou seja,

$$\begin{pmatrix} 4 & -4 \\ -2 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix},$$

cuja solução é

$$v = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Tomamos $\alpha = 1$ e o vetor

$$v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

pois qualquer outro autovetor associado a $\mu = 5$ é múltiplo de v_2 .

Encontramos assim a base $\mathcal{B} = \{v_1, v_2\}$, formada por autovetores da aplicação T . A representação de T na base \mathcal{B} é:

$$T_{\mathcal{B}} = \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix}$$

(justifique!) e nos mostre como a escolha de uma base adequada pode simplificar a representação matricial da aplicação linear T . \triangleleft

Exemplo 6.20 Nem todas as matrizes (ou aplicações lineares) possuem uma base formada por autovalores. De fato, consideremos a matriz

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

O polinômio característico de A é $p(z) = z^2$, de modo que o único autovetor de A é $\lambda = 0$. Associado a esse autovetor existe apenas um autovetor linearmente independente: o sistema $Av = 0$ possui a solução

$$v = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

que nos fornece, ao escolhermos $\alpha = 1$, o autovetor

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Qualquer outro autovetor de A é múltiplo de v_1 , de modo que A não possui uma base formada por autovetores. \triangleleft

Observação 6.21 A existência de uma base \mathcal{B} do \mathbb{R}^n na qual uma aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ tem representação matricial $T_{\mathcal{B}}$ diagonal é um dos principais objetivos do curso de Álgebra Linear II. \triangleleft

Exercício 6.22 Considere a matriz

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 18 \end{pmatrix}.$$

Mostre que existe uma base $\mathcal{B} = \{v_1, v_2, v_3\}$ do \mathbb{R}^3 formada por autovetores da matriz A . Encontre a representação de A na base \mathcal{B} e a matriz $P_{\mathcal{B}}^{\mathcal{E}}$.

6.4 EXERCÍCIOS

1. Considere a aplicação linear $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida por meio da matriz

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

Seja $\mathcal{B} = \{v_1 = (1, 0, 0), v_2 = (1, 2, 0), v_3 = (0, -2, 1)\}$.

- (a) Mostre que \mathcal{B} é uma base do \mathbb{R}^3 ;
- (b) calcule $B = T_{\mathcal{B}}$;
- (c) calcule as matrizes mudança de base $P_{\mathcal{B}}^{\mathcal{E}}$ e $P_{\mathcal{E}}^{\mathcal{B}}$;
- (d) escreva B em termos de A utilizando essas matrizes mudança de base.

2. Seja $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ a aplicação linear definida por

$$T(x_1, x_2, x_3) = (x_1 + x_2 + 2x_3, x_2 + 2x_3, x_3).$$

- (a) Ache $A = T_{\mathcal{E}}$, em que \mathcal{E} é a base canônica do \mathbb{R}^3 ;
- (b) calcule A^{-1} ;
- (c) obtenha então a expressão de T^{-1} .

3. Sejam $T(x, y, z) = (x + y + z, y + z, x)$ e $\mathcal{B} = \{(1, 0, 1), (0, 1, 0), (-1, 0, 1)\}$. Então:

- (a) ache a matriz $T_{\mathcal{B}}$;
- (b) usando essa matriz, especifique uma base para $\ker T$ e $\text{im } T$;
- (c) calcule $T(1, 1, 1)$ utilizando a representação matricial calculada em (a).

4. Considere a matriz

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

- (a) Encontre o polinômio característico de A ;
- (b) obtenha os autovalores de A e uma base \mathcal{B} do \mathbb{R}^3 formada por autovetores de A ;
- (c) ache a matriz $T_{\mathcal{B}}$;
- (d) calcule $T_{\mathcal{E}}$, a matriz $P_{\mathcal{B}}^{\mathcal{E}}$ e sua inversa $P_{\mathcal{E}}^{\mathcal{B}}$;
- (e) verifique que $P_{\mathcal{E}}^{\mathcal{B}} = (P_{\mathcal{B}}^{\mathcal{E}})^t$;
- (f) utilizando as matrizes do item anterior, ache a relação entre $T_{\mathcal{E}}$ e $T_{\mathcal{B}}$.

5. Seja

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{pmatrix}.$$

Obtenha os autovalores de B e uma base \mathcal{B} do \mathbb{R}^3 formada por autovetores de B . Encontre a representação de B na base \mathcal{B} .

AULA 7

O Teorema de Cayley-Hamilton

OBJETIVOS

No final desta aula, o aluno deverá ser capaz de:

1. calcular polinômios de matrizes e de aplicações lineares;
2. calcular o polinômio característico de uma aplicação linear;
3. aplicar o Teorema de Cayley-Hamilton em situações simples.

7.1 POLINÔMIOS DE APLICAÇÕES LINEARES

Seja A uma matriz quadrada. Então faz sentido calcular A^2 e, em geral, a potência A^k , qualquer que seja o número natural $k \geq 1$.

Exemplo 7.1 Seja A a matriz quadrada

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Então

$$A^2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}, \quad \text{e} \quad A^3 = \begin{pmatrix} 13 & 14 \\ 14 & 13 \end{pmatrix},$$

de modo que

$$3A^3 + 1A^2 + 5I = \begin{pmatrix} 49 & 46 \\ 46 & 49 \end{pmatrix}.$$

Observe que $3A^3 + 1A^2 + 5I$ resulta da substituição de $z = A$ no polinômio $3z^3 + z^2 + 5$, com a convenção de que a constante 5 é substituída por $5I$, em que I é a matriz identidade 2×2 . \square

Um pouco mais geralmente, seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma aplicação linear. Podemos definir, para todo $k \in \mathbb{N}$, $T^k = T \circ T^{k-1} = T(T^{k-1})$, se $k \geq 1$ e $T^0 = I$, em que $I: \mathbb{R}^n \rightarrow \mathbb{R}^n$ é a aplicação identidade. Denotaremos $q \in \mathbb{R}[z]$ para indicar que o polinômio

$$q(z) = a_k z^k + a_{k-1} z^{k-1} + \dots + a_1 z + a_0 z^0$$

tem coeficientes no corpo \mathbb{R} . Podemos então calcular

$$q(T) := a_k T^k + a_{k-1} T^{k-1} + \dots + a_1 T + a_0 I.$$

(Aqui, I a aplicação identidade $I: \mathbb{R}^n \rightarrow \mathbb{R}^n$). Note que $q(T)$ é uma aplicação linear do \mathbb{R}^n em \mathbb{R}^n , que é representada por uma matriz $n \times n$ ao se escolher uma base do \mathbb{R}^n .

Exercício 7.2 Seja

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Calcule $q(A)$, se $q(z) = z^5 + 3z^2 + z + 1$.

7.2 SUBESPAÇOS INVARIANTES

Definição 7.3 Um subespaço $W \subset \mathbb{R}^n$ é **invariante** por $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ se $T(W) \subset W$, isto é, $Tw \in W$ para todo $w \in W$.

Exemplo 7.4 Considere a aplicação $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida por

$$T(x_1, x_2, x_3) = (x_1 - x_2, x_1 + x_2, 3x_3)$$

e o subespaço $W = \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\} \subset \mathbb{R}^3$. Afirmamos que W é invariante por T . De fato, se $w \in W$, então $w = (x_1, x_2, 0)$ para $x_1, x_2 \in \mathbb{R}$. Logo,

$$Tw = T(x_1, x_2, 0) = (x_1 - x_2, x_1 + x_2, 0) \in W,$$

mostrando o afirmado. \triangleleft

Para verificar que um subespaço $W \subset \mathbb{R}^n$ é invariante pela aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ basta considerar uma base $\{w_1, \dots, w_m\}$ do espaço W .

Proposição 7.5 Seja $\mathcal{B} = \{w_1, \dots, w_m\}$ uma base do subespaço $W \subset \mathbb{R}^n$. Então W é invariante pela aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ se, e somente se, $Tw_i \in W$ para todo $i \in \{1, \dots, m\}$.

Demonstração: É claro que se W for invariante por T , então $Tw_i \in W$ para todo $i \in \{1, \dots, m\}$. Reciprocamente, se $w \in W$, então existem (únicos) escalares $\alpha_1, \dots, \alpha_n$ tais que $w = \alpha_1 w_1 + \dots + \alpha_n w_n$. Como T é linear,

$$Tw = T(\alpha_1 w_1 + \dots + \alpha_n w_n) = \alpha_1 (Tw_1) + \dots + \alpha_n (Tw_n).$$

Como $Tw_i \in W$ para $i \in \{1, \dots, m\}$, concluímos que $Tw \in W$, pois W é um subespaço. \square

Seja $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma aplicação linear e $W \subset \mathbb{R}^n$ um subespaço. Suponhamos que $T(W) \subset W$, isto é, $Tw \in W$ para todo $w \in W$. Consideremos uma base $\{w_1, \dots, w_m\}$ de W . De acordo com o Teorema 3.43, existem vetores $\{x_{m+1}, \dots, x_n\}$ de modo que

$$\mathcal{B} = \{w_1, \dots, w_m, x_{m+1}, \dots, x_n\}$$

seja uma base do \mathbb{R}^n . Vamos representar T nessa base, que dizer, vamos obter $T_{\mathcal{B}} = T_{\mathcal{B}}^{\mathcal{B}}$. Temos que

$$T_{\mathcal{B}} = ([Tw_1]_{\mathcal{B}} \dots [Tw_m]_{\mathcal{B}} [Tx_{m+1}]_{\mathcal{B}} \dots [Tx_n]_{\mathcal{B}}).$$

Como $Tw_j \in W$, temos que

$$[Tw_j]_{\mathcal{B}} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Por outro lado, nada sabemos sobre Tx_j para $j \in \{m+1, \dots, n\}$. Assim,

$$[Tx_j]_{\mathcal{B}} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \\ b_{(m+1)j} \\ \vdots \\ b_{nj} \end{pmatrix}.$$

Em outras palavras, a matriz $T_{\mathcal{B}}$ tem a forma

$$T_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & a_{1(m+1)} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2m} & a_{2(m+1)} & \dots & a_{2n} \\ \vdots & \dots & \ddots & \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} & a_{m(m+1)} & \dots & a_{mn} \\ 0 & 0 & \dots & 0 & a_{(m+1)(m+1)} & \dots & a_{(m+1)n} \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & 0 & b_{n(m+1)} & \dots & b_{nn} \end{pmatrix},$$

que pode ser escrita na forma

$$T_{\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}.$$

7.3 O TEOREMA DE CAYLEY-HAMILTON

Teorema 7.6 (Cayley-Hamilton)

Se $p \in \mathbb{R}[z]$ for o polinômio característico de $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, então $p(T) = 0$.

Demonstração: Seja $0 \neq v \in \mathbb{R}^n$ arbitrário. Queremos mostrar que $p(T)v = 0$. Seja m o maior natural tal que o conjunto

$$S = \{v, Tv, \dots, T^{m-1}v\}$$

é linearmente independente. Então,

$$T^m v = \alpha_0 v + \dots + \alpha_{m-1} T^{m-1} v. \quad (7.1)$$

Seja $W = \langle S \rangle$ o subespaço gerado por S . Então os elementos de S formam uma base de W . O subespaço W é invariante por T , de acordo com a Proposição 7.5 e a igualdade (7.1).

Completamos então a base S de modo a obter uma base \mathcal{B} do \mathbb{R}^n . Como vimos na seção 7.2, a representação de $T_{\mathcal{B}}$ é dada por

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

em que a matriz A é definida por

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha_0 \\ 1 & 0 & \cdots & 0 & \alpha_1 \\ 0 & 1 & \ddots & 0 & \alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \alpha_{m-1} \end{pmatrix}.$$

O polinômio característico de $T_{\mathcal{B}}$ é

$$p(z) = \det \begin{pmatrix} zI - A & B \\ 0 & zI - D \end{pmatrix} = \det(zI - A) \det(zI - D) = q_1(z)q_2(z),$$

como consequência do Teorema 1.28. (Em cada expressão, os tamanhos das matrizes I são diferentes.)

Vamos calcular $q_1(z) = \det(zI - A)$. Temos:

$$\begin{aligned} \det(zI - A) &= \det \begin{pmatrix} z & 0 & \cdots & 0 & -\alpha_0 \\ -1 & z & \cdots & 0 & -\alpha_1 \\ 0 & -1 & \ddots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & z - \alpha_{m-1} \end{pmatrix} \\ &= z \det \begin{pmatrix} z & 0 & \cdots & -\alpha_1 \\ -1 & z & \cdots & -\alpha_2 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \cdots & -1 & z - \alpha_{m-1} \end{pmatrix} + \\ &\quad (-\alpha_0)(-1)^{m+1} \det \begin{pmatrix} -1 & z & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix}. \end{aligned}$$

Como o determinante da última matriz é $(-1)^{m-1}$, o último termo é justamente $-\alpha_0$. Assim,

$$\det(zI - A) = z \det \begin{pmatrix} z & 0 & \cdots & -\alpha_1 \\ -1 & z & \cdots & -\alpha_2 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \cdots & -1 & z - \alpha_{m-1} \end{pmatrix} - \alpha_0.$$

Procedendo do mesmo modo, obtemos

$$q_1(z) = \det(zI - A) = z^m - \alpha_{m-1}z^{m-1} - \dots - \alpha_0.$$

Substituindo z por T e então calculando $q_1(T)v$, a equação (7.1) nos mostra que $q_1(T)v=0$. Assim, $p(T)v = q_2(T)q_1(T)v = 0$. Como v foi escolhido arbitrariamente, concluímos que $p(T) = 0$. \square

7.4 APLICAÇÕES

Exemplo 7.7 Consideremos a matriz

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

O polinômio característico de A é

$$p(z) = (z - 2)(z + 1)(z - 3) = z^3 - 4z^2 + z + 6.$$

Se quisermos calcular, por exemplo, $q(A)$, sendo $q(z) = z^8 - 4z^7 + z^6 + 13z^5 - 28z^4 + 7z^3 + 42z^2 + z + 1$, podemos efetuar a divisão euclidiana de $q(z)$ por $p(z)$ e obter

$$q(z) = (z^5 + 7z^2)p(z) + z + 1.$$

O Teorema de Cayley-Hamilton garante que $p(A) = 0$; assim, $q(A) = (A^5 + 7A^2)p(A) + A + I = A + I$. Logo,

$$q(A) = A + I = \begin{pmatrix} 3 & 0 & 0 \\ 3 & 0 & 0 \\ 0 & 4 & 4 \end{pmatrix}.$$

□

7.5 EXERCÍCIOS

1. Considere a matriz

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

- (a) Mostre que, se $m(z) = (z - 1)(z - 4)$, então $m(A) = 0$. (O polinômio m não é o polinômio característico de A .)
- (b) procedendo como no Exemplo 7.7, calcule A^{1000} . Para isso, efetue a divisão euclidiana de z^{1000} por $m(z)$ e obtenha o resto $r(z) = \frac{c-1}{3}z + \frac{4-c}{3}$, em que $c = 4^{1000}$. Conclua que

$$A^{1000} = \begin{pmatrix} \frac{c+2}{3} & \frac{c-1}{3} & \frac{c-1}{3} \\ \frac{c^2-1}{3} & \frac{c+2}{3} & \frac{c-1}{3} \\ \frac{c-1}{3} & \frac{c-1}{3} & \frac{c+2}{3} \end{pmatrix}.$$

2. Se $k \in \{1, 2, \dots\}$, calcule A^k , se

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

Referências

- [1] AVRITZER, D. *Geometria Analítica e Álgebra Linear: uma visão geométrica*. Tomos I e II. Editora UFMG: Belo Horizonte, 2009.
- [2] BUENO, H. *Álgebra Linear – um segundo curso*. Rio de Janeiro: Sociedade Brasileira de Matemática, 2006.
- [3] SANTOS, R. J. *Um curso de Geometria Analítica e Álgebra Linear*. Belo Horizonte: Imprensa Universitária da UFMG, 2007.
- [4] SANTOS, R. J. *Introdução à Álgebra Linear*. Belo Horizonte: Imprensa Universitária da UFMG, 2008.
- [5] SANTOS, R. J. *Álgebra Linear e aplicações*. Belo Horizonte: Imprensa Universitária da UFMG, 2006.

Índice remissivo

- base
 - canônica do \mathbb{R}^n , 35
 - de um subespaço, 35
- cálculo de determinantes
 - de matrizes em bloco, 13
 - por meio do escalonamento, 12
- combinação linear, 30
- conjunto
 - gerador, 31
 - ordenado, 35
- dimensão
 - de um subespaço, 39
- espaço
 - vetorial
 - \mathbb{R}^n , 21
 - complexo, 23
 - real, 23
- Gauss-Jordan, 7
- método de Gauss-Jordan, 7
- matriz
 - aumentada de um sistema, 7
 - escalonamento de uma, 8
 - forma escalonada, 8
 - reduzida por linhas, 8
- núcleo
 - de uma matriz, 11
- operações elementares
 - sobre as linhas de uma matriz, 7
- pivô, 7
- regra
 - da unidade, 21, 24
- sistema linear, 7
 - escalonamento, 8
 - forma escalonada, 8
 - reduzida por linhas, 8
- homogêneo, 7
 - matriz aumentada de um, 7
 - não homogêneo, 7
 - homogêneo associado, 7
- operações elementares, 7
- pivô, 7
- variável livre, 9
- subespaço, 27
 - gerado por um conjunto de vetores, 29
- teorema
 - de unicidade da forma escalonada reduzida por linhas, 10
- variável livre, 9
- vetor, 23
 - multiplicação por escalar, 22
- vetores
 - linearmente dependentes, 32
 - linearmente independentes, 32
 - soma de, 22

Sobre o autor

HAMILTON PRADO BUENO doutorou-se em Matemática na PUC-Rio. É professor do Departamento de Matemática da UFMG desde 1979. Atua em pesquisa sobre equações diferenciais, área na qual tem publicado regularmente. Dedica-se também à produção de livros didáticos de Matemática.



Para obter mais
informações sobre
outros títulos da
EDITORA UFMG,
visite o site

www.editora.ufmg.br

A presente edição foi composta pela Editora UFMG, em caracteres Chaparral Pro e Optima Std, e impressa pela Imprensa Universitária da UFMG, em sistema offset 90g (miolo) e cartão supremo 250g (capa), em 2011.