Álgebra Linear Avançada

Notas de Aula

Daniel Miranda

Daniel Miranda

Versão 0.25
10 de Dezembro de 2023

UFABC

I have not yet any clear view as to the extent to which we are at liberty arbitrarily to create imaginaries and endow them with supernatural properties.

--- John Graves

10 de Dezembro de 2023

Conteúdo

Agradecimentos, 8

1 Grupos, Anéis e Corpos, 9

1.1 Operações binárias, 9

1.1.1 Definições e Equivalência, 9

1.1.2 Propriedades Algébricas de Operações Binárias, 11

1.2 Grupos, 14

1.2.1 Grupos de Permutações, 16

1.3 Corpos, 23

1.4 Anel dos Polinômios, 28

1.4.1 Anéis, 28

1.4.2 Anel de polinômios, 30

2 Espaços Vetoriais, 43

2.1 Definições e Exemplos, 44

2.2 Subespaços, 51

2.3 Bases e Dimensão, 58

2.4 Soma de Subespaços, 69

2.4.1 Soma Direta, 70

2.4.2 Decomposição em Somas Diretas, 72

2.5 Soma Direta Externa, 75

2.6 Coordenadas, 79

2.6.1 Espaços Linha e Coluna, 82

2.7 Bandeiras, 85

3 Transformações Lineares, 88

3.1 Definição e Exemplos, 88

3.1.1 Transformações de $\mathbb{K}^n$ a $\mathbb{K}^m$, 95

3.2 Isomorfismos, 100

3.3 Teorema do Núcleo-Imagem, 105

3.4 Representação Matricial dos Homomorfismos, 112

3.4.1 Mudança de Bases, 115

3.5 Somas Diretas e Projeções, 118

3.5.1 Soma Direta de Transformações Lineares, 118

3.5.2 Projeções, 120

3.6 Mudança de Corpo, 122

4 Espaços Quocientes, 128

4.1 Espaços Afins, 128

4.2 Espaço Quociente, 132

4.3 Teoremas de Isomorfismo, 136

4.3.1 Somas diretas e sequências exatas cindidas, 141

5 Espaços Duais, 144

5.1 Espaços Duais, 144

5.1.1 Bases duais, 146

5.1.2 Bidual, 149

5.2 Aniquiladores, 153

5.2.1 Aniquiladores e Somas Diretas, 154

5.3 Transposta de uma Transformação Linear, 156

5.4 Sistemas de Equações, 161

6 Espaços com Produto Interno, 164

6.1 Produto Interno, 165

6.2 Normas, 170

6.3 Bases Ortonormais, 176

6.4 Melhor Aproximação, 181

6.5 Projeção, 183

6.6 Adjuntas, 186

6.7 ⚡Espaços de Hilbert, 190

6.7.1 Espaços de Hilbert Separáveis e Séries de Fourier, 193

7 Determinantes , 202

7.1 Geometria dos Volumes, 202

7.2 Determinantes e Permutações, 208

7.3 Propriedades, 210

8 Estrutura dos Operadores Lineares, 220

8.1 Equivalência de matrizes, 221

8.2 Operadores Similares, 223

8.3 Autovalores e Autovetores, 226

8.3.1 Ação dos Polinômios em $\mathrm{Hom}(V, V)$, 233

8.4 Teoremas de Schur e Cayley-Hamilton, 234

8.4.1 Teorema de Cayley-Hamilton, 237

8.5 Teorema da Decomposição Primária, 239

8.6 Diagonalizabilidade, 243

9 Forma Normal de Jordan, 253

9.1 Operadores Nilpotentes, 253

9.2 Forma Normal de Jordan, 261

9.3 Cálculo da Forma de Jordan, 263

9.4 Forma de Jordan Real, 270

10 Forma Canônica Racional, 275

10.1 Decomposição Cíclica, 275

10.2 Forma Canônica Racional, 281

10.3 Forma Normal de Jordan Generalizada, 287

10.4 Calculando a Forma Canônica Racional, 290

A Teoria de Conjuntos, 296

A.1 Operações com famílias de conjuntos, 298

A.2 Relações, 302

A.2.1 Relação de Equivalência, 302

A.2.2 Relações de Ordem e o Lema de Zorn, 305

A.3 Cardinalidade, 308

A.3.1 Conjuntos Enumeráveis e não Enumeráveis, 309

B Matrizes e Sistemas Lineares , 311

B.1 Matrizes, 311

B.2 Sistemas de Equações Lineares , 315

B.3 Operações Elementares por Linhas , 317

Referências, 323

Índice Remissivo, 325

Índice Remissivo de Espaços Vetoriais, 329

Daniel Miranda
dmiranda@gmail.com
UFABC

Prefácio

Essas notas correspondem aos apontamentos para o curso de Álgebra Linear Avançada (MCTB002) ministradas remotamente durante a pandemia de 2020. Essas notas foram escritas em tempo real durante o decorrer do curso, ainda estão incompletas e podem apresentar erros ou seções incompletas.

Ficaríamos muito gratos se nos fossem enviadas sugestões de melhorias ou que nos fossem apontados erros porventura encontrados.

Estas notas são destinadas a estudantes de graduação e ou pós graduação que já cursaram pelo menos um curso de álgebra linear. Supõe-se que o leitor esteja familiarizado com a linguagem básica de espaços vetoriais, matrizes, sistemas lineares e determinantes.

Agradecimentos

Agradecemos a Mariana Rodrigues da Silveira e aos alunos Deborah Fabri e Guilherme Gigeck pelas correções.

Capítulo 1

Grupos, Anéis e Corpos

Neste capítulo introduziremos algumas estruturas algébricas que precisaremos no estudo de álgebra linear: grupos, anéis e corpos. Não necessitaremos de propriedades profundas desses objetos, que são estudados mais detalhadamente em outros cursos. Para o desenvolvimento subsequente será fundamental os grupos de permutação e em especial a definição de paridade para elementos desses grupos, os anéis de polinômios e o básico da teoria de corpos.

Nesse capítulo

▶ Operações binárias (p. 9)
▶ Grupos (p. 14)
▶ Corpos (p. 23)
▶ Anel dos Polinômios (p. 28)

1.1 Operações binárias

A adição de números inteiros pode ser vista como uma função associando a cada par ordenado $(a,\ b)$ de números inteiros o número inteiro $a+b$

$$a,\ b \xrightarrow{+} a+b.$$

A multiplicação de números inteiros tem a mesma descrição abstrata, enviando um par ordenado de números inteiros para um número inteiro. Essas operações compartilham características abstratas, incluindo associatividade, comutatividade e existência de elementos de identidade.

Esta seção discute algumas propriedades das operações binárias.

Operações e Estruturas Algébricas

Uma estrutura algébrica consiste em um conjunto não vazio A (denominado conjunto subjacente, ou domínio), uma coleção de operações em A (geralmente operações binárias) e um conjunto finito de identidades, conhecido como axiomas, que essas operações devem satisfazer. Algumas estruturas algébricas envolvem também em sua definição outra estrutura. Por exemplo na definição de espaços vetoriais está intrinsecamente envolvido um corpo.

No que segue apresentaremos as operações binárias e algumas identidades usuais: comutatividade, associatividade, existência de elemento neutro e de inverso, etc.

1.1.1 Definições e Equivalência

Definição 1.1

Seja A um conjunto não vazio. Uma **operação binária** em A é uma função

$$\mu\ :\ A \times A \to A.$$

De modo mais geral, dados A e B conjuntos não vazio também chamaremos de **operação binária** em A uma função

$$\mu\ :\ A \times B \to A.$$

Se a e b são elementos de A, geralmente escrevemos ab em vez de $\mu(a,\ b)$. Expressões como $a \cdot b$, $a + b$ ou $a * b$ são usadas para enfatizar o papel da operação, especialmente quando mais de uma operação binária estiver sendo considerada.

O par ordenado $(A, \cdot)$ significa um conjunto não vazio equipado com uma operação binária $\cdot$.

Exemplos 2.

1. *As operações de adição, multiplicação e subtração são operações binárias em $\mathbb{Z}$, o conjunto dos números inteiros.*
2. *A divisão não é uma operação binária em $\mathbb{Z}$, pois, por exemplo, $1 \div 2$ e $1 \div 0$ não representam números inteiros.*
3. *Adição, multiplicação e exponenciação são operações binárias no conjunto $\mathbb{Z}^+$ de números inteiros positivos.*
4. *Sejam X um conjunto arbitrário e X^X o conjunto de todas as funções $f : X \to X$. A composição da função, $\mu(g,\ f) = g \circ f$, é uma operação binária em X^X.*
5. *Seja X um conjunto arbitrário. O operador de interseção define uma operação binária em $A = \mathscr{P}(X)$, o conjunto das partes de X. Da mesma forma, o operador de união define uma operação binária em $\mathscr{P}(X)$.*

$\triangleleft$

Quando A é um conjunto finito de n elementos, uma operação binária pode ser representada por uma **tabela de Cayley**, uma lista tabular $n \times n$ de todos os produtos, com $\mu(a, b) = ab$ sendo colocado na coluna do a e na linha do b.

Exemplo 3. *$A = \{P, I\}$ um conjunto com dois elementos, que representam um inteiro genérico par P e inteiro genérico ímpar I. A tabela de Cayley*

$+$	P	I
P	P	I
I	I	P

expressa o fato que uma soma de dois inteiros pares ou dois inteiros ímpares é par, enquanto a soma de um número inteiro par e ímpar é ímpar.

$\triangleleft$

Exemplo 4. *Seja $A = \{a, b, c\}$ um conjunto com três elementos. As tabelas Cayley a seguir definem operações binárias em A:*

μ_1	a	b	c
a	a	c	b
b	b	a	c
c	c	b	a

μ_2	a	b	c
a	a	b	c
b	b	c	a
c	c	a	b

μ_3	a	b	c
a	a	a	a
b	a	b	c
c	a	c	b

Temos, por exemplo, $\mu_1(b,a) = b$ *(segunda linha, primeira coluna da primeira tabela), enquanto* $\mu_1(a,b) = c, \mu_2(a,b) = b$ *e* $\mu_3(a,b) = a$ *da primeira linha, segunda coluna das respectivas tabelas.*

$\triangleleft$

1.5 Definição Operação Induzida

Seja $(A,\ \cdot)$ um conjunto não vazio equipado com uma operação binária e seja $\phi : A \to B$ uma bijeção. A operação binária em B induzida por ϕ é definida por

$$\phi(a_1) * \phi(a_2) = \phi(a_1 \cdot a_2) \text{ para todos os } a_1 \text{e} a_2 \text{ em } A.$$

Esse relacionamento pode ser visualizado em um diagrama comutativo. Começando com um par $(a_1,\ a_2)$ em $A \times A$, existem duas maneiras de chegar a um elemento de B:

- multiplicando os elementos em A (lado esquerdo do diagrama) e mapeando o produto para B por ϕ (lado inferior do diagrama)
- ou mapeando os elementos individualmente para B (lado superior do diagrama) e multiplicando em B (lado direito).

O diagrama comuta porque essas duas composições de aplicações produzem o mesmo valor para todos os pares de entradas.

$$\begin{tikzcd}
\substack{(a_1,a_2)\\ A \times A} \arrow[r, "\phi\times\phi"] \arrow[d, "\cdot"'] & \substack{(b_1,b_2)\\ B \times B} \arrow[d, "*"] \\
A \arrow[r, "\phi"] & B
\end{tikzcd}$$

1.1.2 Propriedades Algébricas de Operações Binárias

Associatividade Por definição, uma operação binaria envolve dois elementos. Porém, em muitas situações, desejamos combinar três ou mais elementos. Isso gera uma ambiguidade potencial: quando escrevemos um produto abc, esse termo pode significar:

$$(ab)c = \mu(\mu(a,\ b), c) \text{ ou } a(bc) = \mu(a, \mu(b,c)).$$

Operação Induzida
A bijeção $\phi : A \to B$ carrega a operação binária de A para B. Essa construção singela é muito comum em diversos contextos e nas mais diversas áreas da matemática.

Diagramas comutativos
Diagramas são extremamente uteis para representar algumas propriedades algébricas. E em especial os diagramas comutativos que são representações pictóricas e diagramáticas do fato que diversas composições de funções são iguais.
Vale observar que existem diagramas que não comutam.

Associatividade
A propriedade associativa é a propriedade de algumas operações binárias que nos permite reorganizar os parênteses em uma expressão sem alterar o resultado.
Sem a propriedade associativa por exemplo, um produto de quatro elementos pode ser escrito, sem alterar a ordem dos fatores, de cinco maneiras possíveis:

$$((ab)c)d$$
$$(ab)(cd)$$
$$(a(bc))d$$
$$a((bc)d)$$
$$a(b(cd))$$

Se a operação for associativa, todas essas fórmulas produzirão o mesmo resultado e os parênteses podem ser considerados desnecessários e "o" produto pode ser escrito de forma inequívoca como:

$$abcd.$$

Em geral, essas expressões representam diferentes elementos de A. Para a operação binária μ_1 do Exemplo 4, temos

$$(ba)c = bc = c, \quad \text{e} \quad b(ac) = bb = a.$$

1.6 Definição

Uma operação binária μ em um conjunto A é dita **associativa** se $a(bc) = (ab)c$ para todos os a, b e c em A.

Exemplos 7.

1. *Adição e multiplicação são operações associativas no conjunto dos números inteiros, racionais, reais e complexos.*
2. *Subtração em $\mathbb{C}$ não é associativa: se $c \neq 0$, então*

   $$(a - b) - c \neq a - (b - c) = (a - b) + c.$$

   Da mesma forma, a divisão não é associativa no conjunto dos números complexos diferentes de zero.
3. *A potenciação em $\mathbb{R}$ não é associativa.*
4. *Seja $\mathcal{F}(X)$ as funções de X em X. Considere a operação de composição $f, g \mapsto f \circ g$. Essa operação é associativa.*

   Vamos mostrar que a composição de funções é associativa num contexto um pouco mais amplo. Sejam $f : X \to Y, g : Y \to Z, h : Z \to W$. Então: $((f \circ g) \circ h)(x) = h((f \circ g)(x)) = h(g(f(x)))$, e $(f \circ (g \circ h))(x) = (g \circ h)(f(x)) = h(g(f(x)))$.

$\triangleleft$

Identidade No conjunto $\mathbb{Z}$ dos números inteiros, $0 + a = a$ e $1a = a$. O conceito de elemento neutro ou identidade generaliza esse fato.

1.8 Definição

Seja (A, μ) um conjunto equipado com uma operação binária. Um elemento e em A é um **elemento neutro** ou **identidade** para μ se $ea = ae = a$ para todos os a em A.

Uma operação binária pode não possuir identidade. Porém, se possuir então essa será única: se e e e' são elementos neutros para μ, então $e = ee'$ (já que e' é uma identidade) e $ee' = e'$ (já que e é uma identidade), e assim $e = e'$.

Exemplo 9. *Não há identidade para subtração. Em outras palavras, não existe um número inteiro e tal que $a - e = e - a = a$ para todo número inteiro a.*

$\triangleleft$

Elementos inversos No conjunto dos números inteiros, o ato de adicionar a pode ser cancelado adicionando $-a$. O conceito de elemento inverso generaliza essa ideia.

1.10 Definição

Seja (A, μ) um conjunto equipado com uma operação binaria e assuma que existe uma identidade para μ. Se $a \in A$, um elemento b em A é dito um **inverso** de a (com relação a μ) se

$$ab = ba = e.$$

Não faz sentido perguntar sobre a existência de elementos inversos, a menos que μ tenha uma identidade. Além disso, mesmo que μ tenha uma identidade, um elemento específico a em A pode ou não ter um inverso.

1.11 Proposição

Se μ for associativa e possuir uma identidade e, cada elemento a terá no máximo um inverso.

Resumidamente, os inversos são únicos em relação a uma operação associativa. O inverso de a é normalmente indicado como a^{-1}. Porém, para uma operação binária comutativa, habitualmente designada $+$, o inverso de a é usualmente denotado por $-a$.

Bem, nosso objetivo ao definirmos e estudarmos operações algébricas é construirmos estruturas algébricas.

Uma **estrutura algébrica** consiste em um conjunto não vazio A, dito conjunto subjacente ou domínio, uma coleção de operações em A e um conjunto finito de identidades, conhecidas como axiomas, que essas operações devem satisfazer. A definição de algumas estruturas algébricas também envolvem outro conjunto auxiliar (como por exemplo os reais).

Um **homomorfismo** é um aplicação entre duas estruturas algébricas do mesmo tipo (como dois grupos, dois anéis ou dois espaços vetoriais) que preserva a estrutura. Ou seja. uma função $f : A \to B$ entre dois conjuntos A, B equipados com a mesma estrutura, de modo que, se $\cdot$ é uma operação da estrutura (que por simplificação, supomos ser uma operação binária), então

$$f(x \cdot y) = f(x) \cdot f(y)$$

para $x, y \in A$. Dizemos nesse caso que f preserva a operação ou que f é compatível com a operação.

Ressaltamos que um homomorfismo deve também preservar as constantes. Em particular, quando um elemento identidade é exigido num tipo de estrutura, o elemento identidade da primeira estrutura deve ser mapeado para o elemento identidade correspondente da segunda estrutura.

No que se segue apresentaremos alguns exemplos de estruturas algébricas: grupos, anéis, corpos e espaços vetoriais.

Exercícios

Ex. 1.1 --- Suponha que $\circ$ é uma operação binária associativa em A. Seja x um elemento fixo de A. Definimos outra relação binária $\circ^x$ em A por

$$a \circ^x b := a \circ (x \circ b)$$

Mostre que $\circ^x$ é associativa.

Ex. 1.2 --- Prove que Se μ for associativa e possuir uma identidade e, cada elemento a terá no máximo um inverso.

1.2 Grupos

1.1 Definição

Um **grupo** G é um conjunto não vazio com uma operação binária que satisfaz os axiomas a seguir:

G1 associatividade, ou seja, para todos $g, h, j \in G, g \cdot (h \cdot j) = (g\ h)j$;

G2 existe um único elemento $e \in G$ tal que para todo $g \in G$,

$$ge = eg = g,$$

e é denominado **elemento neutro** (ou **identidade** ou **zero** se a operação de G for comutativa;

G3 para cada $g \in G$, existe um único elemento $j \in G$ com a propriedade:

$$gj = jg = e,$$

Grupos de Transformações

Um dos principais exemplos de grupos que motiva o estudo de tal estrutura são os grupos de transformação: grupos que atuam em um determinado espaço X preservando sua estrutura inerente. No caso de grupos de permutação,X é um conjunto; para grupos de transformações lineares,Xé um espaço vetorial. Se em X tivermos a noção de distância temos os grupos de isometria de X.
O conceito de grupo de transformação está intimamente relacionado com o conceito de grupo de simetria: grupos de transformação frequentemente consistem em todas as transformações que preservam uma certa estrutura.
A observação fundamental que permite escolher os axiomas de grupo é que as transformações de um determinado objeto preservando alguma estrutura, por exemplo a distância, satisfazem: dadas duas transformações g e h de um objeto A, podemos realizar uma após a outra. A transformação composta resultante é chamada de produto $g \circ h$. Claramente, se g e h preservam a estrutura, então $g \circ h$ também. Bem como temos uma transformação identidade e associado a cada transformação g temos a transformação inversa g^{-1}.

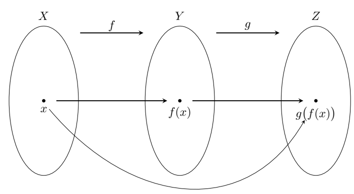


Figura 1.1 Composição de transformações que preserva uma estrutura preserva essa estrutura.

esse elemento é denominado inverso de g e geralmente é denotado por g^{-1} ou $-g$ se a operação de G for comutativa;

Se para todos os $g, h \in G, gh = hg$, então o grupo é denominado **comutativo** ou **abeliano**.

É importante observar que um grupo é na realidade um par $(G,\ (g, h) \to g \cdot h)$ consistindo em um conjunto não vazio G juntamente com uma função de $G \times G \to G$ e satisfazendo os axiomas G1-G3.

Normalmente, eliminamos o ponto que denota a operação e usamos simplesmente a concatenação. Algumas vezes (por exemplo num espaço vetorial), usamos a notação aditiva escrevendo $+$ e 0 para e e $-g$ para g^{-1}

As afirmações a seguir são facilmente deduzidas dos axiomas apresentados anteriormente e serão deixadas como exercício ao leitor.

1.2 Teorema

G seja um grupo, então temos

1. cancelamento - para todos $g, h, j \in G$, se $gh = gj$, ou $hg = jg$, então $h = j$
2. para todos $g, h \in G, (gh)^{-1} = h^{-1}g^{-1}$, e $(g^{-1})^{-1} = g$
3. A equação $gx = j$ possui uma única solução $x = g^{-1}j$

Exemplos 3. *Exemplos de grupos*

1. *os números inteiros $\mathbb{Z}$ com adição,*
2. *os números racionais positivos com multiplicação,*
3. *os números reais $\mathbb{R}$ com adição,*
4. *os números complexos $\mathbb{C}$ com adição,*
5. *os números complexos diferentes de zero com multiplicação,*
6. *os números inteiros que estão entre 0 e $n - 1$ inclusive, com adição módulo n. Esse grupo geralmente é indicado por $\mathbb{Z}/n\mathbb{Z}$, que*
7. *O conjunto $\{0\}$ com operação $0 + 0 = 0$ é um grupo abeliano.*
8. *o conjunto de permutações em um conjunto finito X, com composição como operação.*
9. *O conjunto das matrizes invertíveis com a operação de multiplicação de matrizes.*

◁

Nas seguintes definições, suponha que G seja um grupo. Se $H \subseteq G$ e H for um grupo com as operações de G, então H será denominado **subgrupo** de G. Por exemplo, os racionais com adição formam um subgrupo dos números reais com adição. Além disso, para todos os grupos G, o conjunto constituído apenas pelo elemento neutro forma um subgrupo e todo o grupo G é sempre um subgrupo de si mesmo. Se H for um subgrupo de G e atender à condição extra: $g^{-1}hg \in H$, sempre que $g \in G$ e $h \in H$, H será denominado de **subgrupo normal** de G. Se G for abeliano, todos os subgrupos são normais.

Exercícios

Ex. 1.3 --- Seja G seja um grupo. Mostre que

1. cancelamento - para todos $g, h, j \in G$, se $gh = gj$, ou $hg = jg$, então $h = j$
2. para todos $g, h \in G$, $(gh)^{-1} = h^{-1}g^{-1}$, e $(g^{-1})^{-1} = g$
3. A equação $gx = j$ possui uma única solução $x = g^{-1}j$

Ex. 1.4 --- Suponha que G seja um grupo e que $|G| = 2$. Mostre que $g * g = e$ para todo $g \in G$.

Ex. 1.5 --- Suponha que G seja um grupo e que $|G| = 3$. Mostre que $g * g * g = e$ para todo $g \in G$.

Ex. 1.6 --- Suponha que $x * y = e$ mostre que $y * x = e$.

1.2.1 Grupos de Permutações

1.4 Definição

O **grupo de permutações** S_n consiste em todas as bijeções σ: $[1,\ n] \to [1,\ n]$ onde $[1, n] = \{1, \ldots, n\}$, com a operação de composição

$$\sigma_1 \circ \sigma_2(k) = \sigma_1(\sigma_2(k)) \text{ para } 1 \leq k \leq n$$

como a operação do grupo. O elemento de identidade e é a aplicação identidade $e = \mathrm{I}_{[1,n]}$ de modo que $e(k) = k$, para todos os $k \in [1,\ n]$.

As permutações mais simples são os k-ciclos.

1.5 Definição

Uma lista ordenada $(i_1, \ldots, i_k)$ de k índices distintos em $[1, n] = \{1, \ldots, n\}$ determina um k-**ciclo** em S_n, a permutação que atua da seguinte maneira no conjunto $X = [1,\ n]$:

$$\sigma \text{ leva } \begin{cases} i_1 \to i_2 \to \ldots \to i_k \to i_1 \text{ para os elementos na lista} \\ j \to j \ \text{ para todos os } \ j \text{ que não estão na lista } i_1,\ \ldots,\ i_k\} \end{cases}$$

O 1-ciclo (k) é apenas a aplicação identidade I_X. O suporte de um k-ciclo é o conjunto de entradas $\mathrm{supp}(\sigma) = \{i_1,\ \ldots,\ i_k\}$. O suporte de um 1-ciclo (k) é o conjunto de um ponto $\{k\}$.

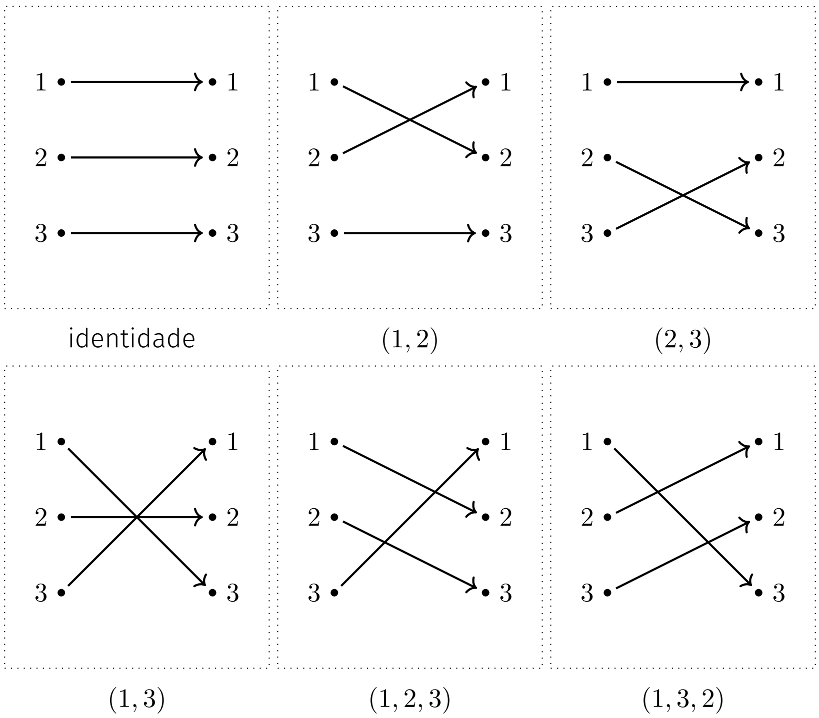


Figura 1.2 Grupo de permutação de três elementos.

A ordem das entradas no símbolo $\sigma = (i_1,\ \ldots,\ i_k)$ é importante, mas a notação do ciclo é ambígua: todos os k símbolos diferentes

$$(i_1,\ \ldots,\ i_k) = (i_2,\ \ldots,\ i_k,\ i_1) = (i_3,\ \ldots,\ i_k,\ i_1,\ i_2) = \ldots = (i_k,\ i_1,\ \ldots,\ i_{k-1})$$

obtido por ``mudanças cíclicas'' das entradas da lista em σ descrevem a mesma permutação em S_n.

Assim, $(1,\ 2,\ 3) = (2, 3, 1) = (3, 1, 2) \neq (1, 3, 2)$ porque $(1,\ 2,\ 3)$ envia $1 \to 2$ enquanto $(1,\ 3,\ 2)$ envia $1 \to 3$.

Uma maneira (complicada) de descrever os elementos gerais $\sigma \in S_n$ emprega uma matriz de dados para mostrar para onde cada $k \in [1,\ n]$ é mapeado:

$$\sigma = \left(\begin{matrix} 1 & 2 & 3 & n \\ j_1 & j_2 & j_3 & j_n \end{matrix} \right]$$

Uma notação mais eficiente é proporcionada pelo fato de que toda permutação σ pode ser decomposta de maneira única como um produto de ciclos com suportes disjuntos, o que significa que os fatores comutam.

O produto de dois ciclos $\sigma\tau = \sigma \circ \tau$ é uma composição de operadores; portanto, a ação de $\sigma\tau = \sigma \circ \tau$ em um elemento $k \in [1,\ n]$ é avaliado inserindo k no produto a partir da direita abaixo. Tomando $\sigma = (1, 2)$ e $\tau = (1, 2, 3)$ em S_5, temos

$$\sigma\tau \;\; k \to (1, 2)(1, 2, 3) \cdot k = (1, 2) \cdot ((1, 2, 3) \cdot k) = ((1, 2)(1, 2, 3) \cdot k)$$

Para determinar a composta, rastreamos o que acontece com cada k:

$$\begin{array}{cc} \multicolumn{2}{c}{(1,2)\ (1,2,3)} \\ 1 \to 2 \to 1 & 1 \to 1 \\ 2 \to 3 \to 3 & 2 \to 3 \\ 3 \to 1 \to 2 & 3 \to 2 \\ 4 \to 4 \to 4 & 4 \to 4 \\ 5 \to 5 \to 5 & 5 \to 5 \end{array}$$

Assim, o produto $(1,\ 2)(1, 2, 3)$ é igual a $(2,\ 3) = (1)(2, 3)(4)(5)$, quando incluímos 1-ciclos redundantes. Por outro lado, $(1,\ 2,\ 3)(1, 2) = (1, 3)$, o que mostra que os ciclos não precisam ser comutar se seus suportes se sobrepuserem. Como outro exemplo, temos

$$(1, 2, 3, 4)^2 \; = (1, 3)(2, 4)$$

que mostra que uma potência σ^k de um ciclo não precisa ser um ciclo, embora seja um produto de ciclos disjuntos. Agora apresentaremos o Teorema Fundamental da Decomposição em Ciclos.

1.6 Teorema Decomposição em Ciclos de Permutações

Todo $\sigma \in S_n$ que não é a identidade é um produto de ciclos disjuntos. Essa decomposição é única (a menos da ordem dos fatores da decomposição) se incluirmos os 1 ciclos necessários para contabilizar todos os índices $k \in [1,\ n]$.

Demonstração. Seja $\sigma \in S_n$ com σ não sendo a identidade.

Seja o primeiro menor elemento tal que $\sigma(a_1) \neq a_1$. Então, para algum $a_2, a_3, \ldots, a_k \in \{1, 2, \ldots, n\}$ temos que

$$\sigma(a_1) = a_2, \sigma(a_2) = a_3, \ldots, \sigma(a_{k-1}) = a_k.$$

Seja k tal que $\sigma(a_k) = a_i$ para algum $i \in \{1, 2, \ldots, k\}$. Se $\sigma(a_k) = a_2$, temos uma contradição, pois $\sigma(a_1) = \sigma(a_2)$ e logo σ não seria injetiva (e todas as permutações são bijetivas por definição).

Se $\sigma(a_k) = a_3$ ou $\sigma(a_k) = a_4$, ou $\sigma(a_k) = a_k$, então o mesmo tipo de contradição surge. Portanto, $\sigma(a_k) = a_1$ e, portanto:

$$a_1 \longrightarrow a_2 \longrightarrow \ldots \longrightarrow a_k \longrightarrow \quad \longrightarrow a_1$$

Agora seja $b_1 \in \{1, 2, \ldots, n\}$ o menor elemento tal que $\sigma(b_1) \neq b_1$ e tal que $b_1 \notin \{a_1, a_2, \ldots, a_k\}$. Repetimos o processo descrito acima para obter um ciclo $(b_1 b_2 \ldots b_j)$ onde $a_s \neq b_t$ para todo $s \in \{1, 2, \ldots, k\}$ e para todo $t \in \{1, 2, \ldots, j\}$.

Visto que $\{1, 2, \ldots, n\}$ é um conjunto finito, este processo deve eventualmente terminar decompondo σ como um produto finito de ciclos disjuntos.

■

Paridade de uma permutação. Em uma direção diferente, observamos que os 2 ciclos $(i,\ j)$ geram todo o grupo S_n no sentido de que todo $\sigma \in S_n$ pode ser escrito como um produto $\sigma = \tau_1 \cdot \ldots \cdot \tau_r$ de dois ciclos. No entanto, esses fatores não são necessariamente disjuntos e não precisam comutar, e essas decomposições estão longe de serem únicas, pois temos, por exemplo,

$$e = (1, 2)^2 = (1, 2)^4 = (1, 3)^2 \text{ etc.}$$

No entanto, um aspecto importante de tais fatorações é único, a saber, sua paridade

$$\operatorname{sgn}(\sigma) = (-1)^r$$

onde

$$r = \#\left[2 - \text{ ciclos na fatoração de } \sigma = (\tau_1,\ \ldots,\ \tau_r)\right].$$

Isso significa que os elementos $\sigma \in S_n$ se enquadram em duas classes disjuntas: **permutações pares** que podem ser escritas como um produto de um número par de 2 ciclos e **permutações ímpares.** Não é óbvio que as decomposições de em 2-ciclos de uma dada permutação tenham a mesma paridade. Provamos isso a seguir e mostramos como calcular $\operatorname{sgn}(\sigma)$ de maneira eficaz.

Primeiro observamos que sempre existe uma decomposição de uma permutação como produto de 2 ciclos. Pelo Teorema 1.2.1, basta mostrar que qualquer ciclo k pode ser decomposto. Para 1 ciclo, isso é óbvio, já que $(k) = e = (1, 2) \cdot (1, 2)$. Quando $k > 1$ é fácil verificar se

$$(1, 2, \ldots,\ k) = (1,\ k) \cdot \ldots \cdot (1, 3)(1, 2)$$

(com $k - 1$ fatores)

Depois de verificarmos que a paridade está bem definida, isso nos diz como reconhecer a paridade de qualquer ciclo k

$$\operatorname{sgn}(i_1,\ i_2, \ldots,\ i_k) = (-1)^{k-1} \text{ para todos os } k > 0$$

1.7 Teorema Paridade

Todas as decomposições $\sigma = \tau_1 \cdot \ldots \cdot \tau_r$ de uma permutação como um produto de 2 ciclos têm a mesma paridade $\operatorname{sgn}(\sigma) = (-1)^r$

Demonstração. O grupo S_n atua no espaço dos polinômios $\mathbb{K}[\boldsymbol{x}] = \mathbb{K}[x_1, \ldots, x_n]$ permutando as variáveis

$$(\sigma \cdot f)(x_1, \ldots,\ x_n) = f(x_{\sigma(1)}, \ldots,\ x_{\sigma(n)})$$

Por exemplo $(1, 2, 3) \cdot f(x_1,\ x_2,\ x_3,\ x_4,\ x_5) = f(x_2,\ x_3,\ x_1,\ x_4,\ x_5)$.

Sejam $x_1, \ldots, x_n$ n variáveis e considere

$$\Delta = \prod_{1 \le i < j \le n} (x_j - x_i).$$

Por exemplo, para $n = 4$, teríamos

$$\Delta = (x_2 - x_1)(x_3 - x_1)(x_4 - x_1)(x_3 - x_2)(x_4 - x_2)(x_4 - x_3).$$

Dada uma permutação $\sigma \in S_n$, defina uma função $f_\sigma \colon \{\Delta, -\Delta\} \to \{\Delta, -\Delta\}$ por

$$f_\sigma(\Delta) = \prod_{1 \le i < j \le n} (x_{\sigma(j)} - x_{\sigma(i)}),$$

e $f_\sigma(-\Delta) = -f_\sigma \Delta$.

Observe que, como σ é uma permutação, $f_\sigma(\Delta) = \Delta$ ou $f_\sigma(\Delta) = -\Delta$. Além disso, se σ, ρ são duas permutações, então $f_\sigma \circ f_\rho = f_{\sigma\rho}$, como é fácil de verificar.

Agora, vamos considerar o que uma transposição $\tau = (a, b)$ faz com Δ. Sem perda de generalidade, assumimos $a < b$.

Os fatores $(x_j - x_i)$ onde nem i nem j são iguais a a ou b permanecem inalterados.

Para os pares com exatamente um índice em $\{a, b\}$, temos duas classes: aquelas em que o outro índice está entre a e b, e aquelas em que o outro índice não está entre a e b.

- Se o outro índice estiver entre a e b, então $x_j - x_a$ é enviado para $-(x_b - x_j)$ e $x_b - x_j$ é enviado para $-(x_j - x_a)$; as duas mudanças de sinal se cancelam.
- Se o outro índice for maior que b, então $x_j - x_a$ e $x_j - x_b$ são trocados, sem mudanças de sinal.

- Se o outro índice for menor que a, então $x_a - x_i$ e $x_b - x_i$ são trocados, sem mudanças de sinal.

Finalmente, o fator $x_b - x_a$ é enviado para $-(x_b - x_a)$.

Em resumo, se τ é uma transposição, então $f_\tau(\Delta) = -\Delta$, $f_\tau(-\Delta) = \Delta$.

Agora pegue uma permutação arbitrária σ e expresse-a como um produto de transposições de duas maneiras diferentes:

$$\sigma = \tau_1 \cdots \tau_r = \rho_1 \cdots \rho_s.$$

Então

$$f_\sigma(\Delta) = f_{\tau_1 \cdots \tau_r}(\Delta) = f_{\tau_1} \circ \cdots \circ f_{\tau_r}(\Delta) = (-1)^r \Delta$$

e

$$f_\sigma(\Delta) = f_{\rho_1 \cdots \rho_s}(\Delta) = f_{\rho_1} \circ \cdots \circ f_{\rho_s}(\Delta) = (-1)^s \Delta.$$

Portanto, $(-1)^r\Delta = (-1)^s\Delta$, então r e s têm a mesma paridade.

■

1.8 Proposição

A aplicação de paridade $\operatorname{sgn} : S_n \to \{\pm 1\}$, definida por $\operatorname{sgn}(\sigma) = (-1)^r$ se σ puder ser decomposto como um produto de r dois ciclos, possui as seguintes propriedades algébricas

a) $\operatorname{sgn}(e) = +1$;

b) $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau)$;

c) $\operatorname{sgn}(\sigma^{-1}) = (\operatorname{sgn}(\sigma))^{-1} = \operatorname{sgn}(\sigma)($ pois $\operatorname{sgn} = \pm 1)$.

Demonstração. Obviamente $\operatorname{sgn}(e) = 1$, pois podemos escrever $e = (1, 2)^2$. Se $\sigma = c_1 \cdot \ldots \cdot c_r$ e $\tau = c'_1 \cdot \ldots \cdot c'_s$ em que c_i, c'_j são 2 ciclos, então $\sigma\tau = c_1 \cdot \ldots \cdot c_r c'_1 \cdot \ldots \cdot c'_s$ é um produto de ciclos $r + s$, provando (2.). A terceira propriedade segue porque

$$1 = \operatorname{sgn}(e) = \operatorname{sgn}(\sigma\sigma^{-1}) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\sigma^{-1})$$

pois os únicos valores de sgn são ± 1.

■

Exercícios

Ex. 1.7 --- Mostre que o produto de ciclos com suportes disjuntos comutam.

Ex. 1.8 --- Escreva

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 5 & 1 & 3 \end{bmatrix}$$

como um produto de ciclos disjuntos e que comutam.

Dica: Comece rastreando $1 \to 2 \to 4 \to \ldots$ até que um ciclo seja concluído; então alimente σ com o primeiro inteiro não incluído no ciclo anterior, etc.

Ex. 1.9 --- calcule a ação total dos seguintes produtos de ciclos

1. $(1,\ 2)(1,\ 3) \in S_3$;
2. $(1,\ 2)(1,\ 3) \in S_6$;
3. $(1,2)(1,2,3,4,5) \in S_5$;
4. $(1,2,3,4,5)(1,2) \in S_5$;
5. $(1,\ 2)^2 \in S_5$;
6. $(1,\ 2,\ 3)^2 \in S_5$.

Ex. 1.10 --- Escreva cada um dos produtos do exercício anterior como um produto de ciclos disjuntos.

Ex. 1.11 --- Determine as inversas σ^{-1} dos seguintes elementos em S_5

1. $(1,2)$;
2. Qualquer dois ciclo $(i_1,\ i_2)$ com $i_1 \neq i_2$;
3. $(1,2,3)$;
4. Qualquer k-ciclo $(i_1, \ldots\ i_k)$.

Ex. 1.12 --- Decomponha os seguintes produtos em S_n como produtos de ciclos disjuntos

1. $(1,\ 5)(1,\ 4)(1,\ 3)(1,\ 2)$;
2. $(1,\ 2)(1,\ 3)(1,\ 4)(1,\ 5)$;
3. $(1,\ k)(1,2,\ k-1)$.

Ex. 1.13 --- A ordem $o(\sigma)$ de uma permutação σ é o menor inteiro $m \geq 1$ tal que $\sigma^m = \sigma \cdot \ldots \cdot \sigma = e$.

1. Prove que todo k-ciclo tem ordem $o(\sigma) = k$.

2. Verifique que a r-ésima potência σ^r de um k-ciclo $\sigma = (i_1, \ldots, i_k)$ é um r-shift

1.3 Corpos

Um corpo é um conjunto munido de duas operações: adição e multiplicação e tal que essas operações se comportam de maneira similar as operações correspondentes nos números racionais, reais e complexos.

Corpos
Nesse texto estudaremos Álgebra Linear sobre corpos gerais e não sobre $\mathbb{R}$ ou $\mathbb{C}$ pois essa generalização não adiciona nenhuma complicação.

1.1 Definição

Um conjunto não vazio $\mathbb{K}$ munido de duas funções $(x, y) \to x + y$ e $(x, y) \to x \cdot y$ de $\mathbb{K} \times \mathbb{K}$ para $\mathbb{K}$ é dito um **corpo** se os nove axiomas a seguir forem atendidos:

K1 $x + y = y + x$ para todos os $x, y \in \mathbb{K}$.

K2 $x + (y + z) = (x + y) + z$ para todos os $x, y, z \in \mathbb{K}$.

K3 Existe um único elemento $0 \in \mathbb{K}$ tal que $x + 0 = x$ para todos os $x \in \mathbb{K}$.

K4 Para cada $x \in \mathbb{K}$, existe um único elemento $-x \in \mathbb{K}$, de modo que $x + (-x) = 0$.

K5 $xy = yx$ para todos os $x, y \in \mathbb{K}$.

K6 $x(yz) = (xy)z$ para todos os $x, y, z \in \mathbb{K}$

K7 Existe um único elemento $1 \neq 0$ em $\mathbb{K}$ tal que $1x = x$ para todos os $x \in \mathbb{K}$.

K8 Para cada $x \neq 0$ em $\mathbb{K}$, existe um único $y \in \mathbb{K}$, tal que $xy = 1$.

K9 $x(y + z) = xy + xz$ para todos os $x, y, z \in \mathbb{K}$.

Estritamente falando, um corpo é um tripla ordenada $(\mathbb{K}, (x, y) \to x + y, (x, y) \to xy)$ satisfazendo os axiomas K1-K9 acima. A operação binária $\mathbb{K} \times \mathbb{K} \to \mathbb{K}$ fdada por $(x, y) \to x + y$ é denominada **adição**, e a operação $(x, y) \to xy$ é denominada **multiplicação**. Ao se referir a algum corpo $(\mathbb{K}, x, y) \to x + y, (x, y) \to xy)$, as

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{1}$	$\bar{3}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{1}$	$\bar{4}$	$\bar{2}$
$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Tabela 1.1 Tabelas de Cayley da adição e multiplicação em $\mathbb{Z}_5$

referências à adição e multiplicação são eliminadas da notação e a letra $\mathbb{K}$ é usada para denotar o conjunto e as duas operações binárias que satisfazem os axiomas K1-K9. Embora esse procedimento seja um tanto ambíguo, ele não causa confusão em situações concretas. Em nosso primeiro exemplo abaixo, introduzimos uma notação que usaremos ao longo deste livro.

Exemplo 2. *Sejam $\mathbb{Q}$ o conjunto de números racionais, $\mathbb{R}$, o conjunto de números reais e $\mathbb{C}$, o conjunto de números complexos. Com a adição e multiplicação usual, $\mathbb{Q}, \mathbb{R}$ e $\mathbb{C}$ são todos corpos com $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.*

$\triangleleft$

Exemplo 3. *Seja $\mathbb{Q}(\sqrt{2}) = \{a+b\sqrt{2} \,|\, a, b \in \mathbb{Q}\}$. Dados $a+b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ e $c+d\sqrt{2}) \in \mathbb{Q}(\sqrt{2})$, definimos a soma e o produto, respectivamente, como:*

$$(a + b\sqrt{2}) + (c + d\sqrt{2}) \triangleq (a + c) + (b + d)\sqrt{2} \in \mathbb{Q}(\sqrt{2}) \tag{1.1}$$
$$(a + b\sqrt{2}) \cdot (c + d\sqrt{2}) \triangleq (ac + 2bd) + (ad + bc)\sqrt{2} \in \mathbb{Q}(\sqrt{2}). \tag{1.2}$$

Então $\mathbb{Q}(\sqrt{2})$ é um corpo.

$\triangleleft$

Os corpos dos Exemplos 2 e 3 são todos infinitos.

$\triangleq$

A expressão $p \triangleq q$ significa que p é definido como sendo q.

Exemplo 4. *Seja $\mathbb{Z}$ conjunto de números inteiros com a adição e a multiplicação usual. Seja p um primo positivo em $\mathbb{Z}$ e defina $\mathbb{Z}_p = \{\bar{0}, \bar{1}, \ldots, \overline{p-1}\}$. Então $\mathbb{Z}_p$ torna-se corpo (finito) se definimos a adição $\oplus$ e a multiplicação módulo p.*

$\triangleleft$

O leitor pode verificar facilmente que $(\mathbb{K}_p, \ \oplus, \ \cdot)$ satisfaz os axiomas K1-K9. Portanto, $\mathbb{K}_p$ é um corpo finito de cardinalidade p.

Exemplo 5. *o conjunto das matrizes 2×2 com entradas em números reais, $M(2, \mathbb{R})$, com adição e multiplicação de matrizes não é um corpo porque a multiplicação não é comutativa, não há divisão e existem divisores de 0.*

$\triangleleft$

1.6 Proposição

Seja $\mathbb{K}$ um corpo.

1. $\mathbb{K}$ possui pelo menos dois elementos 0 e 1 e $0 \neq 1$.
2. O número 0 (zero) é o único elemento neutro da soma.
3. O número 1 é o único elemento neutro da multiplicação.
4. Para todos os $a \in \mathbb{K}$, $a0 = 0a = 0$.
5. O oposto de um elemento é único.
6. O inverso de um elemento (não nulo) é único.
7. Para todos os $a, b \in \mathbb{K}$, se $ab = 0$, então $a = 0$ ou $b = 0$. Nesse caso dizemos que $\mathbb{K}$ não tem **divisores de zero**.
8. Dados quaisquer $a, b \in \mathbb{K}$, temos que $a(-b) = -ab$.

Demonstração. Provaremos 5. Dado $a \in \mathbb{K}$, sejam $a', a'' \in \mathbb{K}$ elementos tais que $a + a' = 0$ e $a + a'' = 0$. Então, usando oportunamente os axiomas acima, temos

$$a' = a' + 0 = a' + (a + a'') = (a' + a) + a'' = 0 + a'' = a''$$

Em outras palavras, provamos que existe um único elemento que cumpre o papel de oposto de a.

Provemos agora 7. Sejam dados $a, b \in \mathbb{K}$ quaisquer. Devemos mostrar que, se $ab = 0$, então ao menos um dos números a e b deve ser igual a 0. Se $a = 0$, não temos nada a provar. Suponhamos então que $a \neq 0$. Então, existe a^{-1} tal que $a.a^{-1} = 1$. Assim, de $ab = 0$, multiplicando ambos os membros por a^{-1}, obtemos

$$a^{-1}(ab) = a^{-1}.0$$

O lado direito, pela propriedade 3 (que supomos já ter sido provada), é igual a 0. Quanto ao lado direito, temos:

$$a^{-1}(ab) = (a^{-1}a)b = 1.b = b$$

Logo, voltando a juntar os lados direito e esquerdo, temos que $b = 0$. ■

Se $\mathbb{F}$ for um corpo e $\mathbb{F} \subset \mathbb{K}$ então $\mathbb{F}$ é dito **subcorpo** de e nesse caso $\mathbb{K}$ será dito extensão de $\mathbb{F}$. Por exemplo, $\mathbb{R}$ é um subcorpo de $\mathbb{C}$ e uma extensão de corpo dos racionais $\mathbb{Q}$.

Se $a \in \mathbb{K}$ e $a + a + \cdots + a = 0$ então a pode, ou não, ser igual a zero; por exemplo, se $\mathbb{K} = \mathbb{Q}$, então $a = 0$, mas se $\mathbb{K} = \mathbb{F}_5$, $5 = 1 + 1 + 1 + 1 + 1 = 0$ e claramente $1 \neq 0$.

Precisamos distinguir esses casos. Para esse fim faremos a seguinte definição

1.7 Definição

A característica de um corpo $\mathbb{K}$ é o menor n positivo tal que

$$\underbrace{1+\cdots+1}_{n \text{ termos}} = 0$$

se tal n existir; caso contrário diremos que o corpo tem característica zero.

Exemplos 8.

1. *Os corpos $\mathbb{Q}, \mathbb{R}$ e $\mathbb{C}$ têm a característica zero.*
2. *Os corpos $\mathbb{F}_p$ têm a característica p.*

$\triangleleft$

! Num corpo de característica p devemos tomar o cuidado em não dividir pelo elemento p ou algum de seus múltiplos.

Exercícios

Ex. 1.14 --- Seja $\mathbb{K}$ um conjunto formado exatamente por dois elementos, $\square$ e $\star$. Defina em $\mathbb{K}$ a operação binária $+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ como segue:

□ $\square + \square = \square$

□ $\square + \star = \star$

□ $\star + \square = \star$

□ $\star + \star = \square$

Defina também a operação: $\cdot \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ como:

□ $\square \cdot \square = \square$

□ $\square \cdot \star = \square$

□ $\star \cdot \square = \square$

□ $\star \cdot \star = \star$

Mostre que $(\mathbb{K}, +, \cdot)$ é um corpo.

Ex. 1.15 --- Seja $(\mathbb{K}, +, \cdot)$ um corpo. Mostre que:

1. Se 0_1 e 0_2 são elementos de $\mathbb{K}$ satisfazendo:

$$x + 0_1 = x \text{ e } x + 0_2 = x, \ \forall x \in \mathbb{K},$$

então $0_1 = 0_2$. Em outras palavras, existe um único elemento neutro da soma num corpo que denotamos por 0.

Ex. 1.16 --- Mostre que $\mathbb{Q}(\sqrt{2})$ é um corpo.

1. Dado $x \in \mathbb{K}$, se x_1 e x_2 são elementos de $\mathbb{K}$ satisfazendo:

$$x + x_1 = 0 \text{ e } x + x_2 = 0,$$

então $x_1 = x_2$. Em outras palavras, num corpo cada elemento x possui um único oposto, denotado por $-x$.

2. Se x_1 e x_2 são elementos de $\mathbb{K}$ que satisfazem:

$$x_1 \cdot x = x \text{ e } x_2 \cdot x = x, \ \forall x \in \mathbb{K},$$

então $x_1 = x_2$. Em outras palavras, um corpo possui uma única unidade, que é denotada por 1.

3. Dado $x \in \mathbb{K}$ com $x \neq 0$, se x_1 e x_2 são elementos de $\mathbb{K}$ satisfazendo:

$$x \cdot x_1 = 1 \text{ e } x \cdot x_2 = 1,$$

então $x_1 = x_2$. Em outras palavras, num corpo cada elemento não nulo x possui um único inverso multiplicativo, denotado por x^{-1}.

Ex. 1.17 --- Seja $(\mathbb{K}, +, \cdot)$ um corpo. Mostre que:

1. $0 \cdot x = 0$, para todo $x \in \mathbb{K}$.

2. Se x_1 e x_2 são elementos de $\mathbb{K}$ com $x_1 \cdot x_2 = 0$, então $x_1 = 0$ ou $x_2 = 0$.

3. $-x = (-1) \cdot x$, para todo $x \in \mathbb{K}$.

4. Se x, x_1 e x_2 são elementos de $\mathbb{K}$ com $x \neq 0$ e $x \cdot x_1 = x \cdot x_2$, então

$$x_1 = x_2.$$

Ex. 1.18 --- Sejam $\mathbb{K}$ um corpo e $\mathbb{F}$ um subcorpo de $\mathbb{K}$. Mostre que:

1. O elemento neutro da soma de $\mathbb{F}$ coincide com 0 elemento neutro da soma de $\mathbb{K}$;

2. A unidade multiplicativa de $\mathbb{F}$ coincide com a unidade multiplicativa de $\mathbb{K}$.

Ex. 1.19 --- Mostre que:

1. $\mathbb{Q}$ é um subcorpo de $\mathbb{R}$,

Ex. 1.20 --- Mostre que todo subcorpo de $\mathbb{C}$ contém $\mathbb{Q}$.

Ex. 1.21 --- Determine a característica do corpo do exercício 1 e dos corpos $\mathbb{C}, \mathbb{R}$ e $\mathbb{Q}$ e $\mathbb{Z}_p$.

Ex. 1.22 --- Mostre que se um corpo tem característica nula, então ele é infinito.

1.4 Anel dos Polinômios

1.4.1 Anéis

Anéis
A definição de um anel generaliza a aritmética dos números inteiros e dos polinômios sobre um corpo.

Nosso principal objetivo nessa seção é estudar propriedades dos anéis dos polinômios sobre um corpo. Estes anéis desempenharão papel fundamental na teoria dos operadores lineares. Começaremos pela definição de um anel.

1.1 Definição Anel

Um anel R é um conjunto com pelo menos dois elementos, 0 e 1, munido de duas operações: adição, denotada por $+$ e multiplicação, denotada por $\cdot$, que satisfaz os seguintes axiomas

R1 o conjunto R forma um grupo comutativo em relação à adição e com elemento neutro 0, denominado **zero**,

R2 a multiplicação é associativa

R3 existe um elemento **identidade**, denotado por 1, satisfazendo

(a) $1 \neq 0$ e

(b) $a \cdot 1 = 1 \cdot a = a$, para todos os $a \in R$.

R4 multiplicação se distribui sobre a adição, ou seja, se $a, b, c \in R$,

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{e} \quad (a + b) \cdot c = a \cdot c + b \cdot c$$

É usual omitir o símbolo $\cdot$ na multiplicação de dois elementos, escrevendo simplesmente ab para denotar o produto de a com b.

Um anel R é dito **commutativo** se $ab = ba$ para todos os $a, b \in R$

Um elemento $a \in R$, diferente de zero, é dito **divisor de zero** se existe $b \in R$, diferente de zero, tal que $ab = 0$.

Um **domínio de integridade** é um anel comutativo e sem divisores de zero.

Exemplos 2.

1. *Os **inteiros** com adição e multiplicação formam um anel comutativo denotado por* $\mathbb{Z}$.
2. *O conjunto das matrizes* 2×2 *com entradas reais,* $\mathcal{M}(2,\ \mathbb{R})$*, com adição e multiplicação de matrizes, é um anel.*

◁

Consideramos brevemente dois subsistemas num anel: **subanel** e **ideal**, que desempenham apenas um papel menor neste livro.

1.3 Definição Subanel e Ideal

- Dado um anel R, um subconjunto não vazio S de R é denominado **subanel** se for um anel com as operações de R.
- Um subconjunto não vazio I de R é denominado **ideal** de R se
  - $a, b \in I$, então $a + b \in I$
  - $a \in R$ e $c \in I$, então $ac \in I$ e $ca \in I$ e

Exemplos 4.

- *Se* R *é o anel de todas as matrizes reais de* 2×2 *e* S *é o subconjunto de matrizes triangulares superiores, então* S *é uma subanel;*
- *Se* R *é* $\mathbb{Z}$ *e* $I = 7\mathbb{Z}$, I *é o ideal de todos os números inteiros múltiplos de* 7.
- *O conjunto* $\{\mathbf{0}\}$ *é ideal para todos os anéis* R *e o elemento* 0 *pertence a todos os ideais* I *de* R.
- *Se* R *for um corpo, os únicos ideais de* R *serão* $\{\mathbf{0}\}$ *e* R *[se o* I *ideal contiver um elemento diferente de zero* a*, por exemplo, então* $1 = aa^{-1} \in I$ *e, portanto,* $c = 1c \in I$ *para todos os* $c \in R$*].*

◁

1.5 Definição Ideal Gerado

Se A é um conjunto qualquer de R, então definimos o **ideal gerado** por A como o menor ideal de R contendo A. Este ideal é a interseção de todos ideais que contém A e será denotado por $\langle A \rangle$.

Pode-se provar que $\langle A \rangle$ é composto de todas as somas finitas da forma $r_1 a_1 + + r_n a_n$ com $r_i \in R$ e $a_i \in A$.

1.6 Definição Domínio de Ideais Principais

Um **domínio de ideais principais** é um domínio de integridade no qual todo ideal é principal, ou seja, pode ser gerado por um único elemento.

1.4.2 Anel de polinômios

Começamos lembrando a definição de um polinômio em uma variável x e introduzimos alguma notação e terminologia que facilitarão a discussão.

1.7 Definição Polinômio

Seja $\mathbb{K}$ um corpo. Por um **polinômio** $p(x)$ com coeficientes em $\mathbb{K}$, queremos dizer uma expressão da forma $a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$, em que $a_i \in \mathbb{K}$ e x é um símbolo abstrato denominado variável. Os escalares a_i são os **coeficientes do polinômio** $p(x)$. O polinômio zero é o polinômio cujos coeficientes são zero. Denotamos esse polinômio por 0.

Suponha $f(x) \neq 0$. O maior número natural k, de modo que o coeficiente a_k não seja zero, é chamado de **grau** de $f(x)$, denotado $\operatorname{grau} f(x)$ e o termo $a_k x^k$ é chamado de termo principal. Se o coeficiente do termo principal for 1, diremos que o polinômio $f(x)$ é **mônico**.

Denotaremos por $\mathbb{K}[x]$ a coleção de todos os polinômios com entradas em $\mathbb{K}$ e por $\mathbb{K}[x]$ todos os polinômios de grau no máximo m.

O grau do polinômio zero é deixado indefinido ou então é definido como negativo (geralmente -1 ou ainda $-\infty$).

Definimos a soma de dois polinômios.

1.8 Definição Soma de Polinômios

Seja $f(x)$ e $g(x)$ sejam dois polinômios de grau k e l, respectivamente. Defina $m = \max\{k,\ l\}$ e desta forma $f(x)$ e $g(x)$ estão em $\mathbb{K}_{(m)}[x]$. Podemos então escrevê-los como $f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ e $g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$. Então a soma de $f(x)$ e $g(x)$ é

$$f(x)+g(x) = (a_m+b_m)x^m+(a_{m-1}+b_{m-1})x^{m-1}+\cdots+(a_1+b_1)x+(a_0+b_0).$$

Agora, definimos o produto de polinômios:

1.9 Definição Produto de Polinômios

Seja $f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ e $g(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$ seja polinômio com entradas em $\mathbb{K}$. Então o produto $f(x)g(x)$ é definido por

$$f(x)g(x) = \sum_{l=0}^{m+n} \left(\sum_{j+k=l} a_j b_k \right) x^l.$$

Ou seja, para obter o coeficiente de x^l no produto, você multiplica todos os termos $a_j x^j$ e $b_k x^k$, onde $j + k = l$ e soma esses termos.

Suponha que $f(x) \neq 0$ tenha o termo principal $a_m x^m$ e $g(x) \neq 0$ tenha o termo principal $b_n x^n$. Então $f(x)g(x)$ tem o termo principal $a_m b_n x^{m+n}$. Portanto, $f(x)g(x)$ é diferente de zero e possui um grau $m + n$.

O próximo teorema coleta as propriedades básicas da multiplicação.

1.10 Teorema

Seja $f, g, h \in \mathbb{K}[x]$. Então, valem as seguintes afirmações:

- a $(fg)h = f(gh)$. A multiplicação de polinômios é associativa.
- b $fg = gf$. A multiplicação de polinômios é comutativa.
- c O polinômio 1 é a identidade multiplicativa: $1 \cdot f = f \cdot 1 = f$.
- d $(f + g)h = fh + gh$. A multiplicação distribui-se sobre adição.
- e Se $f(x)g(x) = 0$, então $f(x) = 0$ ou $g(x) = 0$.
- f Se $f(x) \neq 0$ e $f(x)g(x) = f(x)h(x)$ então $g(x) = h(x)$.

Como consequência dos teorema anteriores, podemos concluir $\mathbb{K}[x]$ é um anel comutativo com identidade sobre $\mathbb{K}$.

1.11 Teorema Teorema da Divisão Euclidiana

Sejam $\mathbb{K}$ um corpo, $f(x)$ e $g(x)$ dois polinômios em $\mathbb{K}[x]$, com $g(x) \neq 0$. Então existem $q(x), r(x) \in \mathbb{K}[x]$ unicamente determinados, tais que $f(x) = g(x)q(x) + r(x)$ onde $r(x) = 0$ ou $\operatorname{grau} r(x) < \operatorname{grau} g(x)$.

Demonstração. Se $f(x) = 0$, então basta tomar $q(x) = r(x) = 0$, para obter o resultado desejado.

Assim, podemos supor que $\operatorname{grau} f(x) \geq \operatorname{grau} g(x)$. Consideremos então $f(x) = \sum_{i=0}^{n} a_i x_i$ e $g(x) = \sum_{i=0}^{m} b_i x_i$, com $a_n \neq 0 \neq b_m$ e $n \geq m$.

Vamos demonstrar por indução sobre $n = \operatorname{grau} f(x)$.

Se $n = \operatorname{grau} f(x) = 0$, então teremos $m = 0$, $f(x) = a_0$ e $g(x) = b_0$ serão polinômios invertíveis em $\mathbb{K}[x]$, de onde segue que

$$f(x) = a_0 = a_0(b_0)^{-1}b_0 + 0,$$

e o Teorema vale.

Suponhamos por hipótese de indução que o teorema vale para todo polinômio $l(x) \in \mathbb{K}[x]$, com $\operatorname{grau} l(x) \leq n-1$. Suponhamos $\operatorname{grau} f(x) = n$ e consideramos o polinômio $h(x) = f(x) - a_n b_m^{-1} x^{n-m} g(x) \in \mathbb{K}[x]$. Assim temos que $\operatorname{grau} h(x) \leq n-1$, e portanto pela hipótese de indução existem polinômios $q_1(x)$ e $r_1(x)$ em $\mathbb{K}[x]$ tais que $h(x) = g(x)q_1(x) + r_1(x)$, com $r_1(x) = 0$ ou $\operatorname{grau} r_1(x) < \operatorname{grau} g(x)$. Desta forma temos:

$$f(x) = h(x) + a_n b_m x^{n-m} g(x) + r_1(x)$$

e logo

$$f(x) = g(x)q_1(x) + r_1(x) + a_n b_m x^{n-m} g(x) = g(x)q_1(x) + a_n b_m x^{n-m}$$

O resultado agora segue tomando-se $q(x) = q_1(x) + a_n b_m$ e $r(x) = r_1(x)$.

Para ver que os polinômios $q(x)$ e $r(x)$ são unicamente determinados, suponhamos que

$$f(x) = g(x)q_1(x) + r_1(x) = g(x)q_2(x) + r_2(x),$$

$r_i(x) = 0$ ou $\operatorname{grau} r_i(x) < \operatorname{grau} g(x)$, para $i = 1, 2$. Da segunda igualdade segue que $g(x)(q_1(x) - q_2(x)) = r_2(x) - r_1(x)$. Se $q_1(x) \neq q_2(x)$, então o grau do polinômio da esquerda na última igualdade é maior ou igual ao grau de $g(x)$. Mas, o grau do polinômio da direita deve ser estritamente menor que o grau de $g(x)$, provocando uma contradição. Assim, devemos ter $q_1(x) = q_2(x)$ e, consequentemente, $r_1(x) = r_2(x)$. ■

Este teorema tem a seguinte consequência importante.

1.12 Corolário

Seja $\mathbb{K}$ um corpo. Então $\mathbb{K}[x]$ é um domínio de ideais principais.

Demonstração. Seja I um ideal de $\mathbb{K}[x]$. Se $I = 0$, então não temos nada a provar. Suponhamos $I \neq 0$. Seja $f(x) \in I$ um polinômio não nulo de menor grau possível. Provaremos que $I = \langle f(x) \rangle$. Para isso seja $h(x) \in I$, então existem polinômios $q(x)$ e $r(x)$ em $\mathbb{K}[x]$ tais que $h(x) = f(x)q(x) + r(x)$, onde $r(x) = 0$ ou $\operatorname{grau} r(x) < \operatorname{grau} f(x)$. Segue daí que $r(x)$ = $h(x)$-$f(x)q(x) \in I$, pois $h(x)$, $f(x) \in I$. Da minimalidade do grau de $f(x)$, segue que $r(x) = 0$, e portanto $h(x)$ = $f(x)q(x)$. Logo, $I \subseteq \langle f(x) \rangle$. A outra inclusão é obvia. Logo, $I = \langle f(x) \rangle$. ■

1.13 Definição Máximo Divisor Comum

Dados $f(x)$, $g(x) \in \mathbb{K}[x]$ definimos o **máximo divisor comum** entre $f(x)$ e $g(x)$ como um gerador do ideal $\langle f(x), g(x)\rangle$.

É imediato que dois mdc entre polinômios diferem por um produto por uma constante, isto é, se $d_1(x) = \mathrm{mdc}(f(x), g(x))$ e $d_2(x) = \mathrm{mdc}(f(x), g(x))$, então existe $u \in \mathbb{K}$ tal que $d_1(x) = ud_2(x)$. Portanto, para obtermos uma unicidade do mdc entre polinômios, podemos dizer que o mdc entre dois polinômios $f(x)$ e $g(x)$ é o gerador mônico do ideal $\langle f(x), g(x)\rangle$.

1.14 Lema Identidade de Bézout

Seja $d(x) = \mathrm{mdc}(f(x), g(x))$ em $\mathbb{K}[x]$. Então existem polinômios $r(x)$ e $s(x)$ tais que

$$(x) = f(x)r(x) + g(x)s(x)$$

Demonstração. Para calcular a identidade de Bezout para $gcd(f, g)$ observe que o conjunto I de polinômios da forma $af+bg$ é fechado sob adição e multiplicação, então é fechado sob resto, uma vez que esta é uma composição de tais operações: $f_i\ bmod g_i = f_i - q\, g_i$. Logo

Todo elemento do ideal $d(x) \in \langle f(x), g(x)\rangle$ é da forma $d(x) = f(x)r(x) + g(x)s(x)$, em particular o gerador. ∎

Vamos agora observar que d é o máximo divisor comum no sentido usual.

1.15 Proposição

Se $f(x)$, $g(x) \in \mathbb{K}[x]$ não são ambos zero, um polinômio d é um máximo divisor comum se, e somente se, divide $f(x)$ e $g(x)$, e possui o maior grau entre os polinômios que possuem esta propriedade.

Demonstração. Começamos observando que o polinômio de menor grau $d \in I$ é um gerador de I. O polinômio d divide todo $h \in I$. Caso contrário o resto da divisão de h por d estria em I e teria grau menor que d, contradizendo a hipótese que d tem o menor grau.

Portanto, se $f, g \in I$ então d é um divisor comum de $f, g,$. Vamos provar que d é o máximo divisor comum. Para isso seja outro divisor c. Como $c \mid f, g$ então $c \mid d = af + bg$, logo $\mathrm{grau}\, c \leq \mathrm{grau}\, d$. ∎

1.16 Proposição

Sejam $f(x), g(x), q(x), r(x) \in \mathbb{K}[x]$, onde $\mathbb{K}$ é um corpo. Se $f(x) = g(x)q(x) + r(x)$, então $\operatorname{mdc}(f(x), g(x)) = \operatorname{mdc}(g(x), r(x))$.

A prova é fácil e deixaremos como exercício para o leitor.

Observe agora que se $f(x)$ e $g(x)$ são dois polinômios em $\mathbb{K}[x]$, onde $\mathbb{K}$ é um corpo, então aplicando-se o teorema da divisão euclidiana repetidas vezes, obtemos: $f(x) = g(x)q_1(x) + r_1(x)$, com $r_1(x) = 0$ ou $\operatorname{grau} r_1(x) < \operatorname{grau} g(x)$ e $g(x) = r_1(x)q_2(x) + r_2(x)$ e assim sucessivamente até que algum destes restos $r_k(x)$ seja o polinômio nulo. É claro que em algum momento obteremos um tal resto, pois do contrário estaríamos construindo uma sequência monótona decrescente infinita de inteiros positivos.

Assim, supondo que $r_{k+1}(x) = 0$, devemos ter $\operatorname{mdc}(f(x), g(x)) = \operatorname{mdc}(g(x), r_1(x)) = \cdots = \operatorname{mdc}(r_{k-1}(x), r_k(x)) = r_k(x)$.

O algoritmo acima permite o cálculo do mdc entre dois polinômio, bem como permite que, isolando cada resto nas equações obtidas durante o processo, encontremos a combinação linear que expressa o máximo divisor comum em função dos dois polinômios iniciais.

Muitas vezes, queremos resolver equações polinomiais em um corpo, mas em muitos casos isso é impossível. Por exemplo, a equação polinomial $x^2 - 2 = 0$ não é solúvel em $\mathbb{Q}$, mas é solúvel em $\mathbb{R}$.

1.17 Definição Raiz

Seja $\mathbb{K}$ um corpo e $p(x)$ seja um polinômio com coeficientes em $\mathbb{K}$. Dizemos que $a \in \mathbb{K}$ é raiz de $p(x)$ se $p(a) = 0$.

O próximo resultado faz a conexão entre o problema de obter soluções de equações polinomiais e a teoria de fatoração polinômios:

1.18 Proposição

Sejam $\mathbb{K}$ um corpo, $f(x) \in \mathbb{K}[x]$e $\alpha \in \mathbb{K}$. Então α é uma raiz de $f(x)$ se, e somente se, $(x - \alpha)$ divide $f(x)$.

Demonstração. Do Teorema da divisão Euclidiana segue que existem polinômios $q(x)$ e $r(x)$ em $\mathbb{K}[x]$ tais que $f(x) = (x-\alpha)q(x)+r(x)$, onde $r(x) = 0$ ou $\operatorname{grau} r(x) < \operatorname{grau}(x - \alpha) = 1$. Portanto, podemos escrever $f(x) = (x - \alpha)q(x) + r_0$, onde $r_0 \in \mathbb{K}$.

Calculando $f(x)$ em α, temos $f(\alpha) = (\alpha-\alpha)q(\alpha)+r_0$, de onde segue que $r_0 = f(\alpha)$. Assim, se $x - \alpha$ divide $f(x)$, então $f(\alpha) = 0$, e reciprocamente. ■

1.19 Definição Multiplicidade da Raiz

Um elemento $\alpha \in \mathbb{K}$ é dito uma **raiz de multiplicidade** k de $p(x)$ se houver um polinômio $s(x)$ tal que $s(\alpha) \neq 0$ e $p(x) = (x - \alpha)^k s(x)$. Quando $k = 1$, α é dito **raiz simples** e quando $k \geq 2$ α é dito **raiz múltipla**.

Podemos definir uma derivada formal em $\mathbb{K}[x]$. Se $p(x) = \sum_{i=0}^{n} a_i x^i$, fazemos $= p'(x) = \sum_{i=1}^{n} i a_i x^{i-1}$. E usando a derivada formal podemos fornecer a seguinte caracterização da multiplicidade de uma raiz.

1.20 Proposição

Seja $p(x) \in \mathbb{K}[x]$. Então α é uma raiz de multiplicidade m de $p(x)$ se, e somente se, $p^{(m-1)}(\alpha) = 0$ mas $p^{(m)}(\alpha) \neq 0$, onde $p^{(k)}(x)$ denota a k-ésima derivada de $p(x)$.

Deixaremos a demonstração dessa proposição como exercício ao leitor

Como corolário da proposição 1.4.2 temos o seguinte resultado sobre o número de raízes de um polinômio.

1.21 Teorema

Sejam $\mathbb{K}$ um corpo e $f(x) \in \mathbb{K}[x]$ um polinômio de grau n. Então $f(x)$ possui no máximo n raízes em $\mathbb{K}$.

Uma classe particularmente importante de corpos são a dos corpos nos quais todo polinômio possui uma raiz.

1.22 Definição Algebricamente Fechado

Dizemos que um corpo$\mathbb{K}$ é algebricamente fechado se todo polinômio $f(x)$ em $\mathbb{K}[x]$ possui uma raiz em $\mathbb{K}$.

O exemplo canônico de um corpo algebricamente fechado é o corpo dos complexos.

1.23 Teorema Teorema Fundamental da Álgebra

Todo polinômio $f(x) \in \mathbb{C}[x]$ possui uma raiz em $\mathbb{C}$.

Como vimos acima, o problema de encontrarmos raízes de um polinômio $f(x)$ em um corpo $\mathbb{K}$ está associado ao fato de podermos fatorá-lo em produtos de outros polinômios em $\mathbb{K}[x]$, onde um dos quais, pelo menos, tem grau 1. Passaremos então a discutir este problema de fatoração a partir de agora. Nesse sentido, os resultados acima nos dão a seguinte consequência interessante.

1.24 Teorema

Sejam $\mathbb{K}$ um corpo algebricamente fechado e $f(x) \in \mathbb{K}[x]$ um polinômio de grau n. Então $f(x)$ se fatora em um produto de fatores lineares:

$$f(x) = c(x - \alpha_1)(x - \alpha_2)\dots(x - \alpha_n)$$

onde $\alpha i \in \mathbb{K}$ são as raízes de $f(x)$ em $\mathbb{K}$, e $c \in \mathbb{K}$ é o coeficiente líder do polinômio $f(x)$.

Demonstração. A demonstração é feita por indução sobre o grau de $f(x)$. ■

Sejam $f(x) \in \mathbb{K}[x]$ um polinômio não invertível. Dizemos que $f(x)$ é irredutível em $\mathbb{K}[x]$, se $f(x)$ só admite fatoração trivial, isto é, se $f(x) = g(x)h(x)$, então $h(x)$ é invertível em $\mathbb{K}[x]$ ou $g(x)$ é invertível em $\mathbb{K}[x]$.

Caso contrário, dizemos que $f(x)$ é redutível em $\mathbb{K}[x]$.

Também temos que $f(x)$ é um polinômio irredutível em $\mathbb{K}[x]$, se o fato de $f(x) = g(x)h(x)$, implicar em $\operatorname{grau} g(x) = 0$ ou $\operatorname{grau} h(x) = 0$, pois os únicos elementos invertíveis nestes anéis são exatamente os polinômios de grau zero.

Se $\mathbb{K}$ é um corpo algebricamente fechado, então segue do Teorema Fundamental da Álgebra que todo polinômio se fatora em produto de polinômios irredutíveis.

1.25 Lema

Se $p, a, b \in \mathbb{K}[x]$ com p irredutível e tal que $p|ab$, então $p|a$ ou $p|b$.

Demonstração. Suponhamos que $p|ab$. Se $p \not| a$, então como p é irredutível temos que $\operatorname{mdc}(p, a) = 1$ e logo existem elementos $r, s \in \mathbb{K}[x]$ tais que $pr + as = 1$. Multiplicando agora esta última igualdade por b, obtemos $b = b(pr+as) = pbr+abs$ de onde segue que $p|b$. ■

1.26 Teorema Teorema de Fatoração Única

Dado $f(x) \in \mathbb{K}[x]$ e $\operatorname{grau} f(x) \geq 1$. Então existem polinômios irredutíveis mônicos $p_1(x), p_2(x), \ldots, p_t(x)$ unicamente determinados e $u \in \mathbb{K}$ tais que

$$f(x) = up_1(x)p_2(x)\ldots p_t(x)$$

com $\operatorname{grau} p_1(x) \leq \operatorname{grau} p_2(x) \leq \cdots \leq \operatorname{grau} p_t(x)$.

1.27 Proposição

Sejam $\mathbb{K}$ um corpo e $f(x)$ um polinômio em $\mathbb{K}[x]$ de grau igual a dois ou três. Então $f(x)$ é irredutível se, e somente se, $f(x)$ não possui raízes em $\mathbb{K}$.

Demonstração. Consideremos inicialmente $\operatorname{grau} f(x) = 2$. Suponhamos $f(x) = g(x)h(x)$, onde $g(x)$, $h(x) \in \mathbb{K}[x]$. Assim, temos $\operatorname{grau} g(x) + \operatorname{grau} h(x) = 2$, de onde decorre que $\operatorname{grau} g(x) = 0$ e $\operatorname{grau} h(x) = 2$, ou $\operatorname{grau} g(x) = 1$ e $\operatorname{grau} h(x) = 1$ ou $\operatorname{grau} g(x) = 2$ e $\operatorname{grau} h(x) = 0$. Se ocorrer o caso em que $\operatorname{grau} g(x) = \operatorname{grau} h(x) = 1$, então estes polinômios possuem raízes em$\mathbb{K}$ e estas são raízes de $f(x)$. As outras duas situações produzem fatorações triviais. Suponhamos agora que $\operatorname{grau} f(x) = 3$. Pelo mesmo tipo de argumento acima, $f(x) = g(x)h(x)$ é uma fatoração não trivial em $\mathbb{K}[x]$ se, e somente se, $\operatorname{grau} g(x) = 1$ ou $\operatorname{grau} h(x) = 1$, isto é, se, e somente se, $g(x)$ tem uma raiz em $\mathbb{K}$ ou $h(x)$ tem uma raiz em $\mathbb{K}$. Isto completa a prova da Proposição. ■

O critério acima não funciona em grau 4, como mostra o exemplo dado pelo polinômio $f(x) = x^4 + 3x^2 + 2$, que não possui raízes em $\mathbb{R}$, mas se fatora como $f(x) = (x^2 + 1)(x^2 + 2)$.

Passaremos a analisar separadamente a irredutibilidade em $\mathbb{R}[x]$, Para este caso, temos um teorema de classificação dos polinômios irredutíveis.

1.28 Teorema

Os únicos polinômios irredutíveis em $\mathbb{R}[x]$ são os lineares e os de grau dois que não possuem raízes em $\mathbb{R}$.

Pelo exposto acima, já sabemos que os polinômios lineares e os polinômios de grau dois que não possuem raízes em $\mathbb{R}$ são irredutíveis em $\mathbb{R}[x]$. O que temos que mostrar então é que estes são os únicos tais polinômios. Vamos fazer isto através de dois resultados auxiliares.

Primeiro observemos que se α é raiz de um polinômio de coeficientes reais então $\overline{\alpha}$ também é

1.29 Lema

Seja $f(x) = ax^2 + bx + c \in \mathbb{R}[x]$, com $a \neq 0$. Se $\alpha \in \mathbb{C}$ é uma raiz de $f(x)$, então $\overline{\alpha}$ também é uma raiz de $f(x)$.

Como consequência imediata deste lema, segue que as raízes complexas aparecem aos pares e, sendo assim, todo polinômio de grau ímpar com coeficientes reais tem pelo menos uma raiz real, de onde concluímos que são polinômios redutíveis em $\mathbb{R}[x]$. Falta então apenas analisar o caso dos polinômios de grau par e maior que dois. Para estes, temos o seguinte resultado.

1.30 Lema

Seja $f(x) \in \mathbb{R}[x]$ um polinômio com $\mathrm{grau} f(x)$ par e maior que 2, sem raízes em R. Então $f(x)$ possui pelo menos um fator irredutível de grau dois.

Demonstração. Seja $f(x)$ um polinômio como no enunciado deste Lema. Pelo Lema anterior, fatorando este polinômio em $\mathbb{C}[x]$, obtemos

$$f(x) = c(x - \alpha_1)(x - \alpha_1) \ldots (x - \alpha k)(x - \alpha k)$$

Observamos agora que o produto $(x - \alpha)(x - \alpha)$, onde $\alpha = a + bi \in \mathbb{C}$, produz um polinômio com coeficientes reais, a saber,

$$(x - \alpha)(x - \alpha) = (x - (a + bi))(x - (a - bi)) = x^2 - 2ax + (a^2 + b^2)$$

Logo, o resultado segue. ■

Supondo $f(x) = ax^2 + bx + c \in \mathbb{R}[x]$, com $a \neq 0$, chamamos $\Delta = b2 - 4ac$ o discriminante de $f(x)$. Assim, $f(x)$ não possui raízes em R se, e somente se, $\Delta < 0$. Resumindo tudo isto, podemos enunciar o seguinte

1.31 Teorema

Seja $f(x) \in \mathbb{R}[x]$. Então $f(x)$ é irredutível em $\mathbb{R}[x]$ se, e somente se, $\mathrm{grau}\, f(x) = 1$ ou, $\mathrm{grau}\, f(x) = 2$ e o discriminante de $f(x)$ é negativo.

Exercícios

Ex. 1.23 --- Seja $p(x) \in \mathbb{K}[x]$. Então α é uma raiz de multiplicidade m de $p(x)$ se, e somente se, $p^{(m-1)}(\alpha) = 0$ mas $p^{(m)}(\alpha) \neq 0$, onde $p^{(k)}(x)$ denota a k-ésima derivada de $p(x)$.

Ex. 1.24 --- Seja A um anel, mostre que para todo $a, b \in A$ vale:

1. $0 \cdot a = a \cdot 0 = 0$

2. $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$

3. $(-a) \cdot (-b) = a \cdot b$

   Se A tem identidade 1_A então:

4. $(-1_A) \cdot a = -a$

5. $(-1_A) \cdot (-1_A) = 1_A$

Ex. 1.25 --- Seja A um anel comutativo. Mostre que de fato o conjunto $(a) = \{x \cdot a \mid x \in A\}$ é um ideal de A para todo $a \in A$.

Ex. 1.26 --- Mostre que $(\mathbb{K}[x], +, \cdot)$ é um anel comutativo com identidade.

Ex. 1.27 --- Sejam $p, q \in \mathbb{K}[x]$ não nulos, mostre que:

1. Se $p \neq -q$ então $\operatorname{grau}(p+q) \leq \max\{\operatorname{grau}(p), \operatorname{grau}(q)\}$

2. $\operatorname{grau}(p \cdot q) = \operatorname{grau}(p) + \operatorname{grau}(q)$

Ex. 1.28 --- Seja $S \subseteq \mathbb{K}[x]$ um subconjunto não vazio, mostre que o conjunto

$$M(S) = \left\{ \sum_{p \in F} a_p \cdot p \in \mathbb{K}[x] \mid F \subseteq S \text{ é finito } a_p \in \mathbb{K}[x] \text{ para todo } p \in F \right\}$$

é um ideal de $\mathbb{K}[x]$ que contém S.

Ex. 1.29 --- Sejam $p_1, \ldots, p_n \in \mathbb{K}[x]$ polinômios coprimos dois a dois (i.e., $\operatorname{mdc}(p_i, p_j) = 1$ para todo $i \neq j$) tais que $p_i \mid q$ para todo i, onde $q \in \mathbb{K}[x]$. Mostre que $p_1 \cdot \ldots \cdot p_n \mid q$.

Ex. 1.30 --- Mostre que se $p \in \mathbb{K}[x]$ é irredutível e $p \nmid q$ então p e q são coprimos.

Ex. 1.31 --- Mostre que $p \in \mathbb{K}[x]$ é um polinômio primo se e somente se é irredutível.

Ex. 1.32 --- Mostre que os polinômios $x - \alpha$ e $x - \beta \in \mathbb{K}[x]$ são coprimos se, e somente se, $\alpha \neq \beta$.

Ex. 1.33 --- Sejam $p, q \in \mathbb{K}[x]$ polinômios não nulos e sejam $a_1, a_2, \ldots, a_m$ elementos distintos de $\mathbb{K}$. Se $p(a_i) = q(a_i)$ para cada $i = 1, 2, \ldots, m$ e m é maior do que os graus de p e de q, então $p = q$.

Ex. 1.34 --- Seja $p \in \mathbb{K}[x]$ um polinômio não constante. Mostre que

1. Se $\mathrm{grau}(p) = 1$ então p é irredutível.
2. Se $\mathrm{grau}(p) = 2$ ou 3 então p é irredutível se, e somente se, não tem raízes em $\mathbb{K}$.
3. É verdade que se $\mathrm{grau}(p) > 3$ então p é irredutível se e somente se não tem raízes em $\mathbb{K}$?

Ex. 1.35 --- Seja $T : \mathbb{K}^n \to \mathbb{K}^n$ um operador linear. Mostre que:

1. O conjunto de todos os polinômios p tais que $p(T) = 0$ formam um ideal não nulo $Z(T)$ de $\mathbb{K}[x]$.
2. Seja m_T o gerador mônico de $Z(T)$, mostre que se $p \in \mathbb{K}[x]$ é tal que $p(T) = 0$ então $m_T | p$.

Ex. 1.36 --- Seja $\mathbb{Q}$ o corpo dos racionais. Determine qual dos seguintes subconjuntos de $\mathbb{Q}[x]$ são ideais. Quando o conjunto for um ideal, ache seu gerador mônico.

1. O conjunto dos polinômios de grau impar.
2. O conjunto dos polinômios de grau maior igual que 5.
3. O conjunto dos polinômios tal que $p(0) = 0$.
4. O conjunto dos polinômios tal que $p(2) = p(4) = 0$.
5. A imagem do operador linear T definido por:

$$T\left(\sum_{i=0}^{n} c_i x^i\right) = \sum_{i=0}^{n} \frac{c_i}{i+1} x^{i+1}$$

Ex. 1.37 --- Ache o máximo divisor comum dos seguintes polinômios:

1. $2x^5 - x^3 - 3x^2 - 6x + 4$ e $x^4 + x^3 - x^2 - 2x - 2$

2. $3x^4 + 8x^2 - 3$ e $x^3 + 2x^2 + 3x + 6$

3. $x^4 - 2x^3 - 2x^2 - 2x - 3$ e $x^3 + 6x^2 + 7x + 1$

Ex. 1.38 --- O Teorema Fundamental da Álgebra diz que todo polinômio não constante $p \in \mathbb{C}[x]$ tem uma raiz em $\mathbb{C}$.

1. Use isso para mostrar que todo polinômio $p \in \mathbb{C}[x]$ se escreve como produto de fatores de grau 1..

2. Se p é um polinômio com coeficientes reais e α é uma raiz complexa de p, então o complexo conjugado $\overline{\alpha}$ é também uma raiz de p.

3. Seja α um número complexo com parte imaginária não nula. Mostre que $(x - \alpha)(x - \overline{\alpha})$ é um polinômio com coeficientes reais. Mostre que é irredutível sobre $\mathbb{R}$.

4. Mostre que todo polinômio com coeficientes reais se fatora em $\mathbb{R}[x]$ como produto de polinômio irredutíveis de graus 1 e 2.

5. Mostre que todo polinômio p de coeficientes reais que tem grau ímpar tem pelo menos uma raiz real.

Ex. 1.39 --- Mostre que se $p(x) = x^n - q \in \mathbb{Q}[x]$ e q é primo então $p(x)$ é irredutível sobre $\mathbb{Q}$. Isto mostra que existem polinômios irredutíveis de qualquer grau positivo em $\mathbb{Q}[x]$.

Ex. 1.40 --- [Fórmula de Taylor] A **derivada** do polinômio $p = a_0 + a_1x + \cdots + a_nx^n$ é o polinômio $p' = a_1 + 2a_2x \cdots + na_nx^{n-1}$. Definimos o operador linear $D : \mathbb{K}[x] \to \mathbb{K}[x]$ tal que $D(p) = p'$ e assim temos que as derivadas formais de ordem superior de p são $p'' = D^2(p)$, $p''' = D^3(p)$, etc. Seja $\mathbb{K}$ um corpo de característica 0, $a$um elemento de $\mathbb{K}$ e n um inteiro positivo. Se $p \in \mathbb{K}[x]$ com $\mathrm{grau}(p) \leq n$, mostre que

$$p = \sum_{k=0}^{n} \frac{D^k(p)}{k!}(a)(x - a)^k.$$

Dica: use o Teorema do Binômio, $(a + b)^m = \sum_{k=0}^{m} \begin{bmatrix} m \\ k \end{bmatrix} a^{m-k}b^k$.

Ex. 1.41 --- Seja c uma raiz do polinômio p, a multiplicidade de c como raiz de p é o maior inteiro positivo r tal que $(x - c)^r$ divide p. Seja $\mathbb{K}$ um corpo de característica 0 e $p \in \mathbb{K}[x]$ com $\mathrm{grau}(p) \leq n$. Mostre que o escalar c é uma raiz de p de multiplicidade r se, e somente se, $D^k(p)(c) = 0$ para todo $0 \leq k \leq r - 1$ e $D^r(p)(c) \neq 0$.

Ex. 1.42 --- Assumindo o teorema fundamental da álgebra prove que se p e q são polinômios sobre os complexos, então $\mathrm{mdc}(p, q) = 1$ se, e somente se, p e q não possuem raízes em comum.

Ex. 1.43 --- Seja p um polinômio mônico sobre os complexos. Prove que um polinômio tem todas as raízes distintas se, e somente se, p e p' (derivada) são coprimos.

Ex. 1.44 ---

1. Mostre que para qualquer $n \geq 2$, existe um polinômio $p \in \mathbb{K}[x]$ de $\mathrm{grau}(p) \leq n-1$ tal que x^n divide $1 + x - p^2$.

2. Deduza que para qualquer matriz nilpotente $N \in M_n(\mathbb{K})$, $I + N$ admite uma raiz quadrada, i.e.: existe $A \in M_n(\mathbb{K})$ tal que $I + N = A^2$.

3. Dê um exemplo para $n = 3$.

Ex. 1.45 --- [Resultado do tipo Bezout em $M_n(\mathbb{Z})$)] Sejam $A, B \in M_n(\mathbb{Z})$ matrizes invertíveis. Assuma que seus determinantes são inteiros coprimos (i.e., $\mathrm{mdc}(\det A, \det B) = 1$). Mostrar que existem matrizes $U, V \in M_n(\mathbb{Z})$ tais que $AU + BV = I_n$.

Capítulo 2

Espaços Vetoriais

When did mathematicians first generally accept the definition of a vetor space as a set of elements subject to suitably axiomatized operations of addition and multiplication by scalars, and not just as a set of n-tuples of scalars closed under these two operations? This abstract description did not come into early use, but has the evident advantages of conceptual simplicity and geometric invariance.

--- Mac Lane

Nesse capítulo

- Definições e Exemplos (p. 44)
- Subespaços (p. 51)
- Bases e Dimensão (p. 58)
- Soma de Subespaços (p. 69)
- Soma Direta Externa (p. 75)
- Coordenadas (p. 79)
- Bandeiras (p. 85)

A álgebra linear tem suas raízes históricas em diversos ramos da matemática. Na álgebra, suas raízes incluem equações lineares, formas bilineares e matrizes, bem como quatérnios e os hipercomplexos; em geometria, segmentos de reta direcionados (os vetores geométricos), geometria afim e projetiva; e na análise, equações diferenciais lineares e espaços de dimensão infinita de vários tipos. Finalmente, a álgebra linear tem suas raízes na física de Maxwell, Gibbs e Heaviside.

Dada a complexidade e variedade desses fenômenos matemáticos, não é surpreendente que a noção geral de espaço vetorial foi isolada e adotada relativamente tarde.

O tratamento moderno, mais abstrato, formulado pela primeira vez por Giuseppe Peano em 1888 na geometria. Não foi influente na época, nem quando Weyl o redescobriu em 1918. Por volta de 1920, foi redescoberto novamente por três analistas: Banach, Hahn e Wiener e uma algebrista, Noether. Então, a noção se desenvolveu rapidamente (MOORE, 1995).

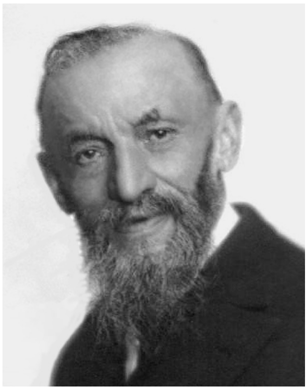


Figura 2.1 Giuseppe Peano

2.1 Definições e Exemplos

No resto do texto a menos que especificado o contrário assumiremos que $\mathbb{K}$ é um corpo arbitrário.

Espaços Vetoriais
Uma breve observação filosófica: nunca diremos exatamente o que são vetores ou espaços vetoriais, apenas como os vetores se comportam, de modo análogo ao que fizemos com os grupos e os corpos. Estes são exemplos de abstração, que aparecem em toda a matemática.
Essa estratégia é tremendamente poderosa: significa que podemos usar uma única teoria (álgebra linear) para lidar com muitos assuntos muito diferentes (vetores físicos, polinômios, matrizes, vetores populacionais na biologia, distribuições de probabilidade na probabilidade, funções na análise, etc.)

2.1 Definição Espaço Vetorial

Um espaço vetorial V sobre $\mathbb{K}$ é um conjunto não vazio, com duas funções, $(\boldsymbol{x}, \boldsymbol{y}) \to \boldsymbol{x} + \boldsymbol{y}$ de $V \times V$ para V (denominada adição) e $(\lambda_1, \boldsymbol{x}) \to \lambda_1 \boldsymbol{x}$ de $\mathbb{K} \times V$ a V (denominada multiplicação por escalar), que satisfazem os seguintes axiomas:

- **V1** $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$ para todos os $\boldsymbol{x}, \boldsymbol{y} \in V$.
- **V2** $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$ para todos os $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$.
- **V3** Existe um elemento $\mathbf{0} \in V$ tal que $\mathbf{0} + \boldsymbol{x} = \boldsymbol{x}$ para todo $\boldsymbol{x} \in V$.
- **V4** Para cada $\boldsymbol{x} \in V$, existe um $\boldsymbol{y} \in V$ tal que $\boldsymbol{x} + \boldsymbol{y} = \mathbf{0}$.
- **V5** $(\lambda_1 \lambda_2)\boldsymbol{x} = \lambda_1(\lambda_2 \boldsymbol{x})$ para todos os $\lambda_1, \lambda_2 \in \mathbb{K}$ e $\boldsymbol{x} \in V$.
- **V6** $\lambda_1(\boldsymbol{x} + \boldsymbol{y}) = \lambda_1 \boldsymbol{x} + \lambda_1 \boldsymbol{y}$ para todos os $\lambda_1 \in \mathbb{K}$ e $\boldsymbol{x}, \boldsymbol{y} \in V$.
- **V7** $(\lambda_1 + \lambda_2)\boldsymbol{x} = \lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{x}$ para todos os $\lambda_1, \lambda_2 \in \mathbb{K}$ e $\boldsymbol{x} \in V$.
- **V8** $1\boldsymbol{x} = \boldsymbol{x}$ para todos os $\boldsymbol{x} \in V$.

Como nos corpos, devemos comentar que um espaço vetorial sobre $\mathbb{K}$ é na realidade uma tripla $(V,\ (\boldsymbol{x}, \boldsymbol{y}) \to \boldsymbol{x}+\boldsymbol{y}, (\lambda, \boldsymbol{x}) \to \lambda \boldsymbol{x})$ consistindo em um conjunto não vazio V juntamente com duas funções de $V \times V \to V$ e $\mathbb{K} \times V \to V$ satisfazendo os axiomas V1-V8.

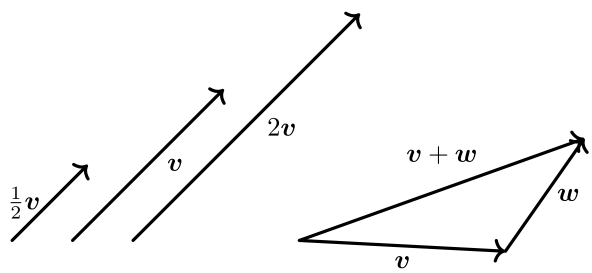


Figura 2.2 Representação Geométrica da multiplicação de um vetor por um escalar e soma de vetores. A intuição geométrica dessa representação será usada mesmo quando vetores forem polinômios, matrizes, etc.

Ressaltamos que existem, em geral, muitas maneiras de munir um determinado conjunto V com uma estrutura de um espaço vetorial sobre $\mathbb{K}$. No entanto, abandonaremos qualquer referência à adição e multiplicação escalar quando nenhuma confusão puder surgir e usaremos a notação V para indicar um determinado espaço vetorial sobre $\mathbb{K}$.

Se V for um espaço vetorial sobre $\mathbb{K}$, os elementos de V serão chamados **vetores** e os elementos de $\mathbb{K}$ **escalares**.

Apresentamos algumas consequências simples dessa definição.

2.2 Proposição

a) O elemento $\mathbf{0} \in V$ é único. Ou seja, existe um único $\mathbf{0}$ tal que $\mathbf{0}+\boldsymbol{x} = \boldsymbol{x}$ para todo $\boldsymbol{x} \in V$.

b) O elemento oposto é único. Ou seja, para todo $\boldsymbol{x} \in V$, existe um único $\boldsymbol{y} \in V$ tal que $\boldsymbol{x}+\boldsymbol{y} = \mathbf{0}$. Esse vetor será denotado por $-\boldsymbol{x}$.

c) $0\boldsymbol{x} = \lambda\mathbf{0} = \mathbf{0}$ para todos os $\lambda \in \mathbb{K}$ e $\boldsymbol{x} \in V$.

d) $(-1)\boldsymbol{x} = -\boldsymbol{x}$.

e) Se $\lambda\boldsymbol{x} = 0$, então $\lambda = 0$ ou $\boldsymbol{x} = 0$.

f) Se $\boldsymbol{x}+\boldsymbol{y} = \boldsymbol{z}+\boldsymbol{y}$ então $\boldsymbol{x} = \boldsymbol{z}$.

Demonstração. c) Como $0\boldsymbol{x} + 0\boldsymbol{x} = (0+0)\boldsymbol{x} = 0\boldsymbol{x}$, segue que $0\boldsymbol{x} = 0$. Analogamente, $\lambda\mathbf{0} + \lambda\mathbf{0} = \lambda(\mathbf{0}+\mathbf{0}) = \lambda\mathbf{0}$, ou seja, $\lambda\mathbf{0} = \mathbf{0}$.

d) De fato, $\boldsymbol{x}+(-1)\boldsymbol{x} = 1\boldsymbol{x}+(-1)\boldsymbol{x} = (1+(-1))\boldsymbol{x} = 0\boldsymbol{x} = 0$, de modo que o vetor $(-1)\boldsymbol{x}$ é o inverso de $\boldsymbol{x}$.

e) De fato, se $\lambda \neq 0$, então como $(\lambda\boldsymbol{x}) = \mathbf{0}$ temos que $\lambda^{-1}\mathbf{0} = \mathbf{0} = \lambda^{-1}(\lambda\boldsymbol{x}) = (\lambda^{-1}\lambda)\boldsymbol{x} = 1\boldsymbol{x} = \boldsymbol{x}$.

a), b) e f) Exercício.

■

Observação 3. *A expressão $\lambda_1\boldsymbol{x}_1 + \ldots + \lambda_n\boldsymbol{x}_n = \sum \lambda_i\boldsymbol{x}_i$ é definida unicamente para qualquer $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ e $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in V$: devido à associatividade da adição, não é necessário inserir parênteses indicando ordem para o cálculo de somas múltiplas. Analogamente, a expressão $\lambda_1\lambda_2\ldots\lambda_n\boldsymbol{x}$ está bem definida.*

Uma expressão da forma $\sum_{i=1}^{n} a_i\boldsymbol{x}_i$ é denominada de **combinação linear** dos vetores $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Os escalares a_i são chamados de **coeficientes** dessa combinação linear.

!

> **! Atenção**
> Uma combinação linear é necessariamente uma soma de um número **finito** de termos!

Exemplo 4 -Espaço zero-dimensional. *Nesse caso* $V = \{\mathbf{0}\}$. *Com a multiplicação por escalar definida como* $\lambda\mathbf{0} = \mathbf{0}$ *para todos os* $\lambda \in \mathbb{K}$ *e com a adição* $\mathbf{0} + \mathbf{0} = \mathbf{0}$.

$\triangleleft$

Ressaltamos que os espaços de dimensão zero sobre corpos distintos são espaços vetoriais distintos. O corpo deve especificado na definição do espaço vetorial.

Exemplo 5 -Corpo. $V = \mathbb{K}$ *com a adição e a multiplicação do corpo é um espaço vetorial.*

De maneira mais geral, dados um corpo $\mathbb{K}$ *e um subcorpo* $\mathbb{F}$ *deste, o corpo* $\mathbb{K}$ *pode ser visto como um espaço vetorial sobre* $\mathbb{F}$. *Por exemplo, o corpo de números complexos* $\mathbb{C}$ *é um espaço vetorial sobre o corpo dos números reais* $\mathbb{R}$, *e este é um espaço vetorial sobre o corpo dos números racionais* $\mathbb{Q}$.

$\triangleleft$

Exemplo 6 -Espaço de Coordenadas, $\mathbb{K}^n$. *Para cada* $n \in \mathbb{N}$, *temos o espaço vetorial* $\mathbb{K}^n = \{(x_1, \ldots, x_n) \,|\, x_i \in \mathbb{K}\}$ *consistindo em todas as ênuplas de elementos de* $\mathbb{K}$. *A adição de vetores e a multiplicação escalar são definidos componente a componente como*

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) \triangleq (x_1 + y_1, \ldots, x_n + y_n) \text{ e}$$
$$\lambda_1(x_1, \ldots, x_n) \triangleq (\lambda_1 x_1, \ldots, \lambda_1 x_n).$$

Em particular, quando $n = 1$, *temos que* $\mathbb{K}$ *é um espaço vetorial sobre* $\mathbb{K}$.

$\triangleleft$

Se A e B são dois conjuntos, vamos denotar o conjunto de funções de A para B por B^A. Assim, $B^A = \{f : A \to B \,|\, f \text{ é uma função }\}$.

No Exemplo 6, $\mathbb{K}^n$ pode ser visto como o conjunto de funções de $\{1, 2, \ldots, n\}$ a $\mathbb{K}$. Assim, $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{K}^n$ é identificado com a função $\delta_{\boldsymbol{x}} \in \mathbb{K}^{\{1,\ldots,n\}}$ definida como $\delta_{\boldsymbol{x}}(i) = x_i$ para $i = 1, \ldots, n$. Essas observações sugerem a seguinte generalização do Exemplo 6.

Exemplo 7 -Espaços de Funções. *Sejam* S *um conjunto arbitrário e* $\mathbb{K}(S) = \mathbb{K}^S$ *o conjunto de funções em* S *com valores no corpo* $\mathbb{K}$. *Se* $f : S \to \mathbb{K}$ *é uma função, então* $f(s)$ *indica o valor de* f *no elemento* $s \in S$.

A adição e multiplicação de funções por um escalar são definidas pontualmente:

- $(f + g)(s) \triangleq f(s) + g(s)$ *para todos os* $s \in S$,
- $(af)(s) \triangleq a(f(s))$ *para todos os* $a \in K, s \in S$.

Como já observamos se $S = \{1, \ldots, n\}$, *então* $\mathbb{K}(S)$ *pode ser identificado com* $\mathbb{K}^n$: *a função* f *está associada ao vetor formado por todos os valores de* f: $(f(1), \ \ldots \ , \ f(n))$. *As regras de adição e multiplicação são consistentes com relação a essa identificação.*

A todo elemento $s \in S$ podemos associar uma função delta δ_s centrada em $\{s\}$ definida como

$$\delta_s(s) = 1 \quad \text{e} \quad \delta_s(t) = 0 \text{ se } t \neq s.$$

Quando $S = (1, \ldots, n)$, usaremos a notação de delta Kronecker $\delta_i(k) = \delta_{ik}$.

Se o conjunto S for finito, uma função $f \in \mathbb{K}(S)$ poderá ser representada unicamente por uma combinação linear de funções delta:

$$f = \sum_{s \in S} f(s)\delta_s.$$

De fato, essa igualdade decorre do fato de que o lado esquerdo é igual ao lado direito em todos os pontos $s \in S$. Inversamente, se $f = \sum_{s\in S} a_s\delta_s$, então, tomando o valor no ponto s obtemos a $f(s) = a_s$

Se o conjunto S for infinito, esse resultado não será verdadeiro pois na definição de combinação linear consideramos apenas somas finitas. Somas de um número infinito de vetores em um espaço vetorial não são definidas, em geral! Algumas somas infinitas podem ser definidas em espaços vetoriais equipados com o conceito de topologia ou norma. Para esse fim veja a Seção 6.7.

$\triangleleft$

Exemplo 8 -Funções Vetoriais. *Seja V um espaço vetorial sobre $\mathbb{K}$ e A um conjunto arbitrário. Então, o conjunto V^A que consiste em todas as funções de A a V é um espaço vetorial sobre $\mathbb{K}$ quando definimos adição e multiplicação escalar pontualmente. Portanto, se $f, g \in V^A$, $f + g$ é a função de A a V definida como $(f + g)(a) \triangleq f(a) + g(a)$ para todos os $a \in A$. Para $\lambda \in \mathbb{K}$ e $f \in V^A$, λf é definido por $(\lambda f(a) \triangleq \lambda_1(f(a))$.*

$\triangleleft$

Se A for um conjunto finito de cardinalidade n no Exemplo 8, reduziremos nossa notação para o espaço vetorial V^A e simplesmente escreveremos V^n. Em particular, se $V = \mathbb{K}$, então $V^n = \mathbb{K}^n$ e recuperamos o exemplo em 6.

Exemplo 9 -Matrizes. *Vamos denotar o conjunto das matrizes $m \times n$ (a_{ij}) com coeficientes $a_{ij} \in \mathbb{K}$ por $\mathcal{M}_{m,n}(\mathbb{K})$. A adição usual de matrizes $(a_{ij}) + (b_{ij}) \triangleq (a_{ij} + b_{ij})$ e a multiplicação por escalar como $\lambda_1(a_{ij}) \triangleq (\lambda_1 a_{ij})$ fazem $\mathcal{M}_{m,n}(\mathbb{K})$ um espaço vetorial sobre $\mathbb{K}$.*

$\triangleleft$

Observe que nossa escolha de notação implica que $\mathbb{K}^n$ e $\mathcal{M}_{1,n}(\mathbb{K})$ são o mesmo espaço vetorial. Vetores dessa forma são também denominados **vetores linhas**. Embora agora tenhamos duas notações diferentes para o mesmo espaço vetorial, essa redundância é útil e não causará confusão na sequência.

Exemplo 10 -Vetores Colunas. *Seja $\mathbb{K}_n \triangleq \mathcal{M}_{n,1}(\mathbb{K})$ com as operações usuais de matrizes, então $\mathbb{K}_n$ é um espaço vetorial denominado espaço dos vetores colunas.*

Nesse caso um vetor $\boldsymbol{x} \in \mathbb{K}_n =$ *é dito* ***vetor coluna****, e é uma matriz consistindo de uma única coluna de* n *elementos.*

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

A matriz transposta do vetor coluna é um vetor linha e vice-versa.

$\triangleleft$

Exemplo 11 -Polinômios. *Seja* $\mathbb{K}[x]$ *o conjunto de todos os polinômios em uma variável* x *sobre* $\mathbb{K}$*. Assim, um elemento típico de* $\mathbb{K}[x]$ *é uma soma finita do formato* $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$*. Aqui* $n \in \mathbb{N} \cup \{\mathbf{0}\}$ *e* $a_0, \ldots, a_n \in \mathbb{K}$*. As noções usuais de adicionar dois polinômios e multiplicar um polinômio por uma constante, com a qual o leitor está familiarizado, estão bem definidas sobre qualquer corpo* $\mathbb{K}$*. Essas operações fornecem a* $\mathbb{K}[x]$ *a estrutura de um espaço vetorial sobre* $\mathbb{K}$*.*

$\triangleleft$

Muitos exemplos interessantes de espaços vetoriais vêm da análise. Aqui estão alguns exemplos típicos.

Exemplo 12 -Sequências. *Uma sequência com valores num corpo* $\mathbb{K}$ *é uma função* $s : \mathbb{N} \to \mathbb{K}$*. É usual para sequências escrevermos* $a_n \triangleq a(n)$*. O conjunto* $\mathrm{Seq}(\mathbb{K})$ *de todas as sequências infinitas com valores num corpo* $\mathbb{K}$ *é um espaço vetorial com as operações definidas termo a termo:*

$$(s_n) + (t_n) \triangleq (s_n + t_n) \quad e \tag{2.1}$$

$$a(s_n) \triangleq (as_n) \tag{2.2}$$

Também denotaremos o espaço $\mathrm{Seq}(\mathbb{K})$ *como* $\mathbb{K}^\infty$ *e veremos uma sequência como uma lista ordenada. Ou seja, veremos a sequência* (a_n) *como a lista ordenada* $(a_1, a_2, \cdots, a_i, \cdots)$ *e logo*

$$\mathbb{K}^\infty = \mathrm{Seq}(\mathbb{K}) = \{(a_1, a_2, \cdots, a_i, \cdots) \,|a_i \in \mathbb{K}, i \in \mathbb{N}\}.$$

com a soma e a multiplicação coordenada a coordenada.

$\triangleleft$

Exemplo 13 -Funções Contínuas e Diferenciáveis. *Seja* I *um intervalo fechado. Vamos denotar por* $C(I)$ *as funções reais contínuas em* I*. Denotaremos por* $C^k(I)$ *as funções de* $C(I)$ *que são* k*-vezes diferenciáveis no interior de* I*. Então* $C(I) \supseteq C^1(I) \supseteq C^2(I) \supseteq \cdots$*. Esses conjuntos são espaços vetoriais sobre* $\mathbb{R}$ *quando dotados da adição usual pontual* $(f+g)(x) = f(x) + g(x), x \in I$ *e multiplicação escalar* $(\lambda f)(x) = \lambda(f(x))$*.*

$\triangleleft$

Exemplo 14 -Funções integráveis. *Seja $B = [a_1, b_1] \times \ldots \times [a_n, b_n] \subseteq \mathbb{R}^n$ seja um retângulo fechado. Seja $\mathcal{R}(B)$ o conjunto de todas as funções a valores reais em B que são Riemann integráveis. Claramente $\mathcal{R}(B)$ é um espaço vetorial sobre $\mathbb{R}$ com a adição e multiplicação por escalar definidas como no Exemplo 13.*

$\triangleleft$

Observação 15. *Geralmente, é muito conveniente, mas não totalmente consistente, denotar o elemento zero e a adição em $\mathbb{K}$ e V pelos mesmos símbolos.*

Apresentaremos a seguir dois exemplos nos quais uma notação diferente é preferível.

Exemplo 16. *Seja $L = \{x \in \mathbb{R} \,|\, x > 0\}$. Consideramos L como um grupo abeliano com relação à multiplicação e introduzimos em L multiplicação por um escalar de $\mathbb{R}$ de acordo com a fórmula $(a,\ x) \to x^a$. É fácil verificar se todas as condições da Definição 2.1 são satisfeitas, embora na notação usual elas assumam uma forma diferente:*

- *o vetor zero em L é o elemento 1, ou seja, $\mathbf{0} = 1$.*
- *a identidade $1\boldsymbol{x} = \boldsymbol{x}$ adquire a forma $x^1 = x$*
- *a identidade $a(b\boldsymbol{x}) = (ab)\boldsymbol{x}$ adquire a forma $(x^b)^a = x^{ba}$*
- *a identidade $(a+b)\boldsymbol{x} = a\boldsymbol{x}+b\boldsymbol{x}$ neste contexto pode ser escrita como $x^{a+b} = x^a x^b$; etc.*

$\triangleleft$

Exemplo 17. *Seja V um espaço vetorial sobre o corpo dos números complexos $\mathbb{C}$. Defina um novo espaço vetorial $\overline{V}$ com o mesmo grupo aditivo V, mas com uma multiplicação por escalar diferente:*

$$(a, \boldsymbol{x}) \mapsto \overline{a}\boldsymbol{x},$$

onde $\overline{a}$ denota o conjugado complexo de a. Das fórmulas $\overline{a+b} = \overline{a} + \overline{b}$ e $\overline{ab} = \overline{a}\overline{b}$ segue sem maiores dificuldades que $\overline{V}$ é um espaço vetorial.

$\triangleleft$

Exercícios

Ex. 2.1 --- Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Mostre que:

1. O vetor nulo $\mathbf{0}$ é único.
2. O vetor oposto $-\boldsymbol{v}$ a cada vetor $\boldsymbol{v} \in V$ é único.

3. $-1 \cdot \boldsymbol{v} = -\boldsymbol{v}$ para todo $\boldsymbol{v} \in V$, conclua que $-(-\boldsymbol{v}) = \boldsymbol{v}$.

4. Dados $\alpha \in \mathbb{K}$, $\boldsymbol{v} \in V$ temos $\alpha \cdot \boldsymbol{v} = 0 \Leftrightarrow \alpha = 0$ ou $v = 0$.

5. Dados $\alpha \in \mathbb{K}$, $\boldsymbol{v} \in V$ temos $\alpha \cdot \boldsymbol{v} = \boldsymbol{v} \Leftrightarrow \alpha = 1$ ou $v = 0$.

6. Mostre que se $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{z} + \boldsymbol{y}$ então $\boldsymbol{x} = \boldsymbol{z}$.

7. O axioma de comutatividade da soma pode ser deduzido dos outros axiomas da definição de espaço vetorial

Ex. 2.2 --- Determine se os seguintes conjuntos são espaços vetoriais sobre o corpo $\mathbb{K}$ especificado em cada caso.

1. $\mathbb{R}^3$ sobre $\mathbb{K} = \mathbb{R}$ com as operações

$$(x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$$
$$\alpha \cdot (x_1, x_2, x_3) = (\alpha \cdot x_1, x_2, x_3)$$

2. O conjunto $V = \{(a, b) \in \mathbb{R}^2 : a, b > 0\}$ sobre $\mathbb{K} = \mathbb{R}$ com as operações

$$(a, b) + (c, d) = (ac, bd)$$
$$\alpha \cdot (a, b) = (a^\alpha, b^\alpha),\ \alpha \in \mathbb{R}$$

3. O conjunto $\mathbb{Q}(\sqrt{2})$ dos elementos $a + b\sqrt{2}$ com $a, b \in \mathbb{Q}$ sobre $\mathbb{K} = \mathbb{Q}$ com as operações:

$$(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}$$
$$\alpha \cdot (a + b\sqrt{2}) = \alpha \cdot a + \alpha \cdot b\sqrt{2},\ \alpha \in \mathbb{Q}$$

Ex. 2.3 --- Mostre que $\mathbb{N}$ pode ser visto como um espaço vetorial sobre $\mathbb{Q}$. (Dica: use o fato que existe uma bijeção de $\mathbb{N}$ em $\mathbb{Q}$ e defina novas operações em $\mathbb{N}$)

Ex. 2.4 --- Seja V um espaço vetorial. Use os axiomas de espaço vetorial para provar que se $\boldsymbol{v} \in V$ e $n \in \mathbb{N}$ então

$$n\boldsymbol{v} = \underbrace{\boldsymbol{v} + \ldots + \boldsymbol{v}}_{n \text{ parcelas}}.$$

Ex. 2.5 --- Seja V um espaço vetorial. Sejam $\boldsymbol{u}, \boldsymbol{v} \in V$ não nulos. Prove que $\boldsymbol{v}$ é múltiplo de $\boldsymbol{u}$ se, e somente se,, $\boldsymbol{u}$ é múltiplo de $\boldsymbol{v}$.

Ex. 2.6 --- Prove detalhadamente que o espaço de funções $\mathbb{K}(S) = S^{\mathbb{K}}$ é um espaço vetorial com a adição e multiplicação de funções por um escalar são definidas pontualmente:

□$(f+g)(s) \triangleq f(s)+g(s)$ para todos os $s \in S$,

□$(af)(s) \triangleq a(f(s))$ para todos os $a \in K, s \in S$.

Ex. 2.7 --- Se S é um conjunto finito com n elementos, quantos elementos possui o espaço vetorial $S^{\mathbb{Z}_2}$?

Ex. 2.8 --- Defina a média $\boldsymbol{u} \star \boldsymbol{v}$ entre dois vetores $\boldsymbol{u}, \boldsymbol{v}$ no espaço vetorial V pondo $\boldsymbol{u} \star \boldsymbol{v} = \frac{1}{2}\boldsymbol{u} + \frac{1}{2}\boldsymbol{v}$. Prove que $(\boldsymbol{u} \star \boldsymbol{v}) \star \boldsymbol{w} = \boldsymbol{u} \star (\boldsymbol{v} \star \boldsymbol{w})$ se, e somente se,, $\boldsymbol{u} = \boldsymbol{w}$.

Ex. 2.9 --- Dados os espaços vetoriais V_1, V_2, considere o conjunto $V = V_1 \times V_2$ (produto cartesiano de V_1 por V_2), cujos elementos são os pares ordenados $v = (v_1, v_2)$, com $v_1 \in V_1$ e $v_2 \in V_2$. Defina operações que tornem V um espaço vetorial. Verifique a validade de cada um dos axiomas.

Ex. 2.10 --- Em $\mathbb{R}^2$ mantenhamos a definição de produto αv de um número por um vetor mas modifiquemos, de três maneiras diferentes, a definição de soma $u+v$ de vetores $u = (x, y)$ e $v = (x', y')$. Em cada tentativa, dizer quais axiomas de espaço vetorial continuam válidos e quais são violados:

1. $u + v = (x + y', x' + y)$,
2. $u + v = (xx', yy')$,
3. $u + v = (3x + 3x', 5x + 5x')$.

Ex. 2.11 --- Dado $\mathbb{K}$ um corpo finito com q elementos. Quantos elementos possui o espaço vetorial $\mathbb{K}^n$?

2.2 Subespaços

Na maioria das estruturas algébricas podemos definir subestruturas e os espaços vetoriais não são exceções. Suponha que V seja um espaço vetorial sobre $\mathbb{K}$.

2.1 Definição Subespaço Vetorial

Um subconjunto não vazio W de V é um **subespaço vetorial** de V se W

> for um espaço vetorial com as mesmas operações de adição de vetores e multiplicação escalar que V.

A seguinte proposição apesar de simples é extremamente útil.

> **2.2 Proposição**
>
> Um subconjunto W de V é um subespaço vetorial se W se for fechado em relação às operações de adição de vetores e multiplicação escalar de V.

Demonstração. Exercício. ■

Exemplo 3. *O conjunto dos polinômios de grau menor ou igual a n é subespaço de $\mathbb{K}[x]$. Esse espaço Vetorial será denotado por $\mathbb{K}_n[x]$.*

◁

Exemplos 4. *Os conjuntos $C([a,b])$, $C^k([a,b])$, e $\mathcal{R}([a,b])$ são todos os subespaços de $\mathbb{R}^{[a,b]}$.*

◁

Exemplo 5. *Vamos apresentar alguns exemplos de subespaços do espaço de sequências.*

O conjunto c_0 de todas as sequências de números complexos que convergem para 0 é um espaço vetorial, assim como o conjunto ℓ^∞ de todas as sequências complexas limitadas. Além disso, se p for um número inteiro positivo, o conjunto ℓ^p de todas as sequências complexas (s_n) para as quais

$$\sum_{n=1}^{\infty} |s_n|^p < \infty$$

é um espaço vetorial em operações definidas componente a componente. Para ver que o subespaço ℓ^p é fechado em relação a adição usaremos a desigualdade de Minkowski

$$\left(\sum_{n=1}^{\infty} |s_n + t_n|^p\right)^{1/p} \leq \left(\sum_{n=1}^{\infty} |s_n|^p\right)^{1/p} + \left(\sum_{n=1}^{\infty} |t_n|^p\right)^{1/p}$$

◁

Exemplo 6. *O conjunto*

$$(\mathbb{K}^\infty)_0 = \{(a_1, a_2, \cdots,) \,|a_i \in \mathbb{K} \text{ com apenas um número finito de termos não nulos } \}$$

é um subespaço de $\mathbb{K}^\infty$

◁

Exemplo 7 -Sistema Linear. *O espaço de solução de um sistema de equações lineares homogêneo. Seja A uma matriz $m \times n$ sobre $\mathbb{K}$. O conjunto de todas as matrizes $n \times 1$ X sobre $\mathbb{K}$, satisfazendo $AX = 0$ é um subespaço do espaço de todas as matrizes $n \times 1$ sobre $\mathbb{K}$.*

Para provar isso, devemos mostrar que $A(\lambda X + Y) = 0$ quando $AX = 0$, $AY = 0$ e λ é um escalar arbitrário em $\mathbb{K}$. Isso segue imediatamente do seguinte fato geral.

Se A é matriz $m \times n$ sobre $\mathbb{K}$ e B, C matrizes $n \times p$ sobre $\mathbb{K}$ então

$$A(\lambda B + C) = \lambda(AB) + AC$$

para todo escalar λ em $\mathbb{K}$.

$$\begin{aligned}
[A(\lambda B + C)]_{ij} &= \sum_k A_{ik}(\lambda B + C)_{kj} \\
&= \sum_k (\lambda A_{ik}B_{kj} + A_{ik}C_{kj}) \\
&= \lambda \sum_k A_{ik}B_{kj} + \sum_k A_{ik}C_{kj} \\
&= \lambda(AB)_{ij} + (AC)_{ij} \\
&= [\lambda(AB) + AC]_{ij}.
\end{aligned}$$

$\lhd$

Exemplo 8. *Considere o seguinte sistema de equações diferenciais lineares:*

$$x_1' = a_{11}x_1 + \cdots + a_{1n}x_n \tag{2.3}$$

$$\vdots \tag{2.4}$$

$$x_n' = a_{n1}x_1 + \cdots + a_{nn}x_n \tag{2.5}$$

com $x_1, \ldots, x_n \in C^1(I)$, onde I é um intervalo aberto em $\mathbb{R}$, x_i' indica a derivada de x_i e a_{ij} são escalares em $\mathbb{R}$. Defina $A = (a_{ij}) \in \mathrm{M}_{n\times n}(\mathbb{R})$. A matriz A é denominada de ***matriz do sistema****. Defina $\boldsymbol{x} = (x_1, \ldots, x_n)^t$. Com essa notação, nosso sistema de equações diferenciais se torna*

$$\boldsymbol{x}' = A\boldsymbol{x} \text{ com } \boldsymbol{x} \in C^1(I)^n.$$

O conjunto de soluções para o nosso sistema é $V = \{\boldsymbol{x} \in C^1(I)^n \,|\, \boldsymbol{x}' = A\boldsymbol{x}\}$. Claramente, V é um subespaço vetorial de $C^1(I)^n$.

$\lhd$

Se tivermos uma coleção $\mathcal{S} = \{W_i \,|\, i \in I\}$ de subespaços de V, então existem algumas maneiras óbvias de formar um novo subespaço a partir de $\mathcal{S}$.

Seja $\mathcal{S} = \{W_i \,|\, i \in I\}$ uma coleção de subespaços de V. A interseção, $\bigcap\limits_{i\in I} W_i$, dos subespaços em $\mathcal{S}$ é um subespaço vetorial de V.

Se $\mathcal{S}$ possuir a propriedade de que, para todo $i, j \in I$, existe um $k \in I$ tal que $W_i \cup W_j \subseteq W_k$, então claramente $\bigcup\limits_{i\in I} W_i$ é um subespaço de V.

Porém destacamos que em geral, a união de dois subespaços de V não é um subespaço vetorial de V. De fato, se W_1 e W_2 são subespaços de V, então $W_1 \cup W_2$ é um subespaço vetorial se, e somente se, $W_1 \subseteq W_2$ ou $W_2 \subseteq W_1$. Este fato é fácil de provar e será deixado como exercício.

2.9 Teorema

Seja V um espaço vetorial sobre um corpo infinito $\mathbb{K}$. Então V não pode ser a união de um número finito de subespaços próprios.

Demonstração. Suponha que $W_1, \ldots, W_n$ sejam subespaços próprios de V, tais que $V = W_1 \cup \cdots \cup W_n$. Mostraremos que tal fato é impossível. Lembramos ao leitor que um subespaço vetorial W de V é próprio se $W \neq V$. Assim, $V \setminus W \neq \emptyset$ para um subespaço vetorial W de V.

Podemos assumir, sem perda de generalidade, que $W_1 \subsetneq W_2 \cup \cdots \cup W_n$. Sejam $\boldsymbol{x} \in W_1 \setminus \bigcup_{i=2}^{n} W_i$ e $\boldsymbol{y} \in V \setminus W_1$. Como o corpo $\mathbb{K}$ é infinito e $\boldsymbol{y}$ não é o vetor nulo, $I = \{\boldsymbol{x} + \lambda_1 \boldsymbol{y} \mid \lambda_1 \in \mathbb{K}\}$ é um subconjunto infinito de V. Como existem apenas finitos subespaços W_i, existem $\{$ 1, $\ldots, n\}$ de forma que $I \cap W_j$ seja infinito. Suponha que $j \in \{2, \ldots, n\}$.

Então, existem dois escalares diferentes de zero $\lambda_1, \lambda_2 \in \mathbb{K}$, de modo que $\lambda_1 \neq \lambda_2$ e $\boldsymbol{x} + \lambda_1 \boldsymbol{y}, \boldsymbol{x} + \lambda_2 \boldsymbol{y} \in W_j$. Como W_j é um subespaço vetorial, $(\lambda_2 - \lambda_1)\boldsymbol{x} = \lambda_2(\boldsymbol{x} + \lambda_1 \boldsymbol{y}) - \lambda_1(\boldsymbol{x} + \lambda_2 \boldsymbol{y}) \in W_j$. Como $\lambda_2 - \lambda_1 \neq 0$, concluímos $\boldsymbol{x} \in W_j$. Mas esse fato é contrário à nossa escolha de $\boldsymbol{x} \notin W_2 \cup \cdots \cup W_n$. Assim, $j = 1$.

Agora, se $j = 1$, novamente existem dois escalares diferentes de zero $\lambda_1, \lambda_2 \in \mathbb{K}$ de modo que $\lambda_1 \neq \lambda_2$ e $\boldsymbol{x} + \lambda_1 \boldsymbol{y}, \boldsymbol{x} + \lambda_2 \boldsymbol{y} \in W_1$. Então $(\lambda_1 - \lambda_2)\boldsymbol{y} = (\boldsymbol{x} + \lambda_1 \boldsymbol{y}) - (\boldsymbol{x} + \lambda_2 \boldsymbol{y}) \in W_1$. Como $\lambda_1 - \lambda_2 \neq 0, \boldsymbol{y} \in W_1$.

Mas isso é impossível, pois o vetor $\boldsymbol{y}$ foi escolhido em $V \setminus W_1$. Concluímos assim que V não pode ser igual à união de $W_1, \ldots, W_n$.

■

Exemplo 10. *Se $\mathbb{K}$ é finito, então o teorema anterior é falso em geral. Por exemplo, seja $V = (\mathbb{F}_2)^2$. Então $V = W_1 \cup W_2 \cup W_3$, em que $W_1 = \{(0,0),(1,0)\}, W_2 = \{(0,0),(0,1)\}$ e $W_3 = \{(0,0),(1,1)\}$.*

◁

2.11 Definição Subespaço Gerado

Dado um subconjunto S de um espaço vetorial V o menor subespaço de V

contendo S

$$\langle S\rangle \triangleq \bigcap\{W \mid W \text{ é um subespaço de } V \text{ e } S \subseteq W\}$$

é denominado **subespaço gerado** por S.

Denotaremos por $\mathscr{P}(V)$ o conjunto de todos os subconjuntos de V e $\mathcal{S}(V)$ o conjunto de todos os subespaços de V. Então $\mathcal{S}(V) \subseteq \mathscr{P}(V)$ e temos uma função natural $\langle\cdot\rangle : \mathscr{P}(V) \to \mathcal{S}(V)$, que envia um subconjunto $S \in \mathscr{P}(V)$ para o subespaço vetorial $\langle S\rangle \in \mathcal{S}(V)$. Claramente, $\langle\cdot\rangle$ é uma aplicação sobrejetiva cuja restrição a $\mathcal{S}(V)$ é a identidade.

2.12 Definição Combinação Linear

Seja $B = \{\boldsymbol{v}_i \mid i \in I\}$ subconjunto de V. Um vetor $v \in V$ é uma combinação linear dos vetores em B se houver um conjunto escalares $\{c_i\}$, com apenas um número **finito** desses diferentes de 0 de modo que:

$$\boldsymbol{v} = \sum_{i\in I} c_i \boldsymbol{v}_i.$$

Observação 13. *Se escolhermos todos $c_i = 0$, obteremos*

$$\mathbf{0} = \sum c_i \boldsymbol{v}_i.$$

Esta é a combinação linear trivial dos vetores em B. Qualquer outra combinação linear é não trivial.

No caso, $B = \emptyset$, a única combinação linear que temos é a combinação linear vazia, cujo valor consideramos $0 \in V$ e que consideramos uma combinação linear trivial.

2.14 Teorema

A função $\langle\cdot\rangle : \mathscr{P}(V) \to \mathcal{S}(V)$ possui as seguintes propriedades

a Para $S \in \mathscr{P}(V), \langle S\rangle$ é o subespaço de V de todas as combinações lineares finitas de vetores de S. Assim,

$$\langle S\rangle = \left\{\sum_{i=1}^{n} x_i \boldsymbol{x}_i \mid x_i \in \mathbb{K}, \boldsymbol{x}_i \in S, n \geq 0\right\}$$

b Se $S_1 \subseteq S_2$, $\langle S_1\rangle \subseteq \langle S_2\rangle$.

c Se $\boldsymbol{x} \in \langle S\rangle$, existe um subconjunto finito $S' \subseteq S$ de modo que $\boldsymbol{x} \in \langle S'\rangle$.

d $S \subseteq \langle S\rangle$ para todos os $S \in \mathscr{P}(V)$.

$\boxed{e}$ Para cada $S \in \mathscr{P}(V), \langle\langle S\rangle\rangle = \langle S\rangle$.

$\boxed{f}$ Se $\boldsymbol{y} \in \langle S \cup \{\boldsymbol{x}\}\rangle$ e $\boldsymbol{y} \notin \langle S\rangle$, então $\boldsymbol{x} \in \langle S \cup \{\boldsymbol{y}\}\rangle$. Aqui $\boldsymbol{x}, \boldsymbol{y} \in V$ e $S \in \mathscr{P}(V)$.

Demonstração. $\boxed{a}$ Como $\langle S\rangle$ é subespaço então $\langle S\rangle$ é fechado em relação a combinações lineares e assim $\{\sum_{i=1}^{n} x_i\boldsymbol{x}_i \,|\, x_i \in \mathbb{K}, \boldsymbol{x}_i \in S, n \geq 0\} \subset \langle S\rangle$.

Como $\{\sum_{i=1}^{n} x_i\boldsymbol{x}_i \,|\, x_i \in \mathbb{K}, \boldsymbol{x}_i \in S, n \geq 0\}$ é um espaço vetorial contendo S, a igualdade segue.

As demonstrações das propriedades $\boxed{b}$ – $\boxed{e}$ serão deixadas como exercício. Vamos demonstrar $\boxed{f}$. Se $\boldsymbol{y} \in \langle S \cup \{\boldsymbol{x}\}\rangle - \langle S\rangle$, então $\boldsymbol{y}$ será uma combinação linear finita de vetores de $S \cup \{\boldsymbol{x}\}$. Além disso, $\boldsymbol{x}$ deve ocorrer com um coeficiente diferente de zero em tal combinação linear. Caso contrário, $\boldsymbol{y} \in \langle S\rangle$. Portanto, existem vetores $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in S$ e escalares diferentes de zero $x_1, \ldots, x_n, x_{n+1} \in \mathbb{K}$ de modo que $\boldsymbol{y} = x_1\boldsymbol{x}_1 + \cdots + x_n\boldsymbol{x}_n + x_{n+1}\boldsymbol{x}$. Como $x_{n+1} \neq 0$, podemos escrever $\boldsymbol{x}$ como uma combinação linear de $\boldsymbol{y}$ e $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Ou seja, $\boldsymbol{x} = x_{n+1}^{-1}\boldsymbol{y} - x_{n+1}^{-1}x_1\boldsymbol{x}_1 - x_{n+1}^{-1}x_n\boldsymbol{x}_n$. Assim, $\boldsymbol{x} \in \langle S \cup \{\boldsymbol{y}\}\rangle$.

■

Exercícios

Ex. 2.12 --- Prove que se se W se for fechado em relação as operações de adição de vetores e multiplicação escalar de V então W é um subespaço.

Ex. 2.13 --- Para os conjuntos seguintes, determine se os conjuntos dados são espaços vetoriais reais, se a adição e a multiplicação forem as usuais. Para aqueles que não forem diga quais axiomas de espaços vetoriais não são satisfeitos.

1. O conjunto dos polinômios de grau menor ou igual a n
2. O conjunto de todas as funções reais tais que $f(0) = f(1)$
3. O conjunto das funções tais que $f(0) = 1 + f(1)$
4. O conjunto das funções reais crescentes.
5. O conjunto das funções reais pares.
6. O conjunto das funções reais ímpares.
7. O conjunto das funções contínuas em $[0, 1]$ tais que $\displaystyle\int_0^1 f(x)dx = 0$

8. O conjunto das funções contínuas em $[0,1]$ tais que $\int_0^1 f(x)dx \geq 0$

9. O conjunto dos vetores (x,y,z) em $\mathbb{R}^3$ tais que $x=0$ ou $y=0$

10. O conjunto das matrizes 2×2 cujo traço é zero

11. O conjunto das matrizes 2×2 cujo determinante é zero

12. O conjunto das matrizes 2×2 que são simétricas,i.e., $A=A^t$

13. O conjunto dos vetores (x,y,z) que satisfaz a equação linear $a_1x+a_2y+a_3z=0$

Ex. 2.14 --- Considere o subconjunto S das funções de $\mathcal{F}(\mathbb{R},\mathbb{R})$ que são soluções da equação diferencial linear homogênea de ordem n com coeficientes constantes

$$y^{(n)}(t)+a_{n-1}y^{(n-1)}(t)+\cdots+a_1y'(t)+a_0y(t)=0$$

onde $a_0,a_1,\ldots,a_{n-1}\in\mathbb{R}$. Mostre que S é um subespaço vetorial de $\mathcal{F}(\mathbb{R},\mathbb{R})$.

Ex. 2.15 --- Seja $\mathcal{F}(X,\mathbb{K})$ o espaço das funções, onde $\mathbb{K}$ é um corpo. No caso particular em que $X=\mathbb{N}$ o espaço é denominado **espaço de sequências** e vamos denotá-lo por $\mathbb{K}^{\mathbb{N}}$. Um elemento $f\in\mathbb{K}^{\mathbb{N}}$ é uma função $f:\mathbb{N}\to\mathbb{K}$ dada por $n\mapsto x_n\in\mathbb{K}$ e será denotado por $f=(x_n)_{n\in\mathbb{N}}$. Quais dos seguintes conjuntos são subespaços de $\mathbb{K}^{\mathbb{N}}$?

1. O conjunto das sequências com apenas um número finito de coordenadas diferentes de zero.

2. Nenhuma coordenada igual a 1.

   Nos próximos itens $\mathbb{K}=\mathbb{R}$

3. O conjunto das séries de Cauchy, ou seja, as sequencias tais que dado $\epsilon>0$ existe $N>0$ tal que $|x_n-x_m|<\epsilon$ para $n,m>N$.

4. As sequências tais que $\sum_{n=1}^{\infty}|x_n|<\infty$.

5. As sequências limitadas, i.e. as sequencias $(x_n)_{n\in\mathbb{N}}$ para as quais existe $M>0$ tal que $|x_n|\leq M$ para todo $n\in\mathbb{N}$.

Ex. 2.16 --- Dado S um subespaço de V e $v\in V$. O conjunto $v+S=\{v+s : s\in S\}$ é chamado subespaço afim de V.

1. Quando um subespaço afim de V é subespaço de V?

2. Mostre que dois subespaços afim $x + S$ e $y + S$ ou são iguais ou são disjuntos.

Ex. 2.17 --- Prove que

1. Se $S_1 \subseteq S_2$, $\langle S_1 \rangle \subseteq \langle S_2 \rangle$.
2. Se $\boldsymbol{x} \in \langle S \rangle$, existe um subconjunto finito $S' \subseteq S$ de modo que $\boldsymbol{x} \in \langle S' \rangle$.
3. $S \subseteq \langle S \rangle$ para todos os $S \in \mathscr{P}(V)$.
4. Para cada $S \in \mathscr{P}(V), \langle \langle S \rangle \rangle = \langle S \rangle$.

Ex. 2.18 --- Seja $X \subseteq V$ um subconjunto de um espaço vetorial sobre $\mathbb{K}$ V. Mostre que $\langle X \rangle = \{\sum_{v \in F} \alpha_v \cdot v : F \subseteq X$ é um subconjunto finito e $\alpha_v \in \mathbb{K}$, $\forall v \in F\}$.

Ex. 2.19 --- Dê um contra-exemplo que a união de subespaços não é necessariamente um subespaço.

Ex. 2.20 --- Mostre que se W_1 e W_2 são subespaços de V, então $W_1 \cup W_2$ é um subespaço vetorial se, e somente se, $W_1 \subseteq W_2$ ou $W_2 \subseteq W_1$

2.3 Bases e Dimensão

2.1 Definição Dependência e Independência Linear

Suponha que V seja um espaço vetorial arbitrário sobre um corpo $\mathbb{K}$ e seja S um subconjunto de V.

- S é dito **linearmente dependente** sobre $\mathbb{K}$ se existir um subconjunto finito $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\} \subseteq S$ e escalares diferentes de zero $x_1, \dots, x_n \in \mathbb{K}$ de modo que $x_1\boldsymbol{x}_1 + +x_n\boldsymbol{x}_n = 0$.
- S é dito **linearmente independente** (sobre $\mathbb{K}$) se S não for linearmente dependente.

Portanto, se S for linearmente independente, sempre que $\displaystyle\sum_{i=1}^{n} x_i\boldsymbol{x}_i = 0$ com $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\} \subseteq S$ e $\{x_1, \dots, x_n\} \subseteq \mathbb{K}$ e, assim, $x_1 = \cdots x_n = 0$.

Nossa definição implica que o conjunto vazio $\emptyset$ é linearmente independente sobre $\mathbb{K}$, por vacuidade. Ao considerar questões de dependência, abandonaremos as palavras ``sobre $\mathbb{K}$'' sempre que $\mathbb{K}$ for claro a partir do contexto. No entanto, deve ser óbvio que, se mais de um corpo estiver envolvido, um determinado conjunto S pode ser linearmente dependente sobre um corpo e independente sobre outro. O exemplo a seguir torna isso claro.

Exemplo 2. *Suponha $V = \mathbb{R}$, o corpo dos números reais e sejam $\mathbb{K}_1 = \mathbb{Q}$ e $\mathbb{K}_2 = \mathbb{R}$. Então V é um espaço vetorial sobre $\mathbb{K}_1$ e $\mathbb{K}_2$. Seja $S = \{\boldsymbol{x}_1 = 1, \boldsymbol{x}_2 = \sqrt{2}\}$. É fácil ver que S é linearmente independente sobre $\mathbb{K}_1$. Mas, claramente, Sé linearmente dependente sobre $\mathbb{K}_2$ pois $(\sqrt{2})\boldsymbol{x}_1 + (-1)\boldsymbol{x}_2 = 0$.*

$\lhd$

2.3 Definição Base

Um subconjunto S de V é dito **base** de V se S for linearmente independente sobre $\mathbb{K}$ e $\langle S\rangle = V$.

! Se S for uma base de um espaço vetorial V', todo vetor diferente de zero $\boldsymbol{x} \in V$ poderá ser escrito unicamente no formato $\boldsymbol{x} = x_1\boldsymbol{x}_1 + \cdots + x_n\boldsymbol{x}_n$, em que $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subseteq S$ e $x_1, \ldots, x_n$ são escalares diferentes de zero em $\mathbb{K}$.

! **Atenção**
Mesmo que as bases possam ser infinitas, todo vetor do espaço vetorial deve ser escrito como **combinação linear finita** dos vetores da base!

Notação 4. *As bases serão denotadas por $\underline{B}, \underline{X}, \ldots$.*

Vejamos alguns exemplos de bases.

Exemplo 5. *O conjunto vazio $\emptyset$ é uma base para o subespaço nulo (0) de qualquer espaço vetorial V. Se considerarmos um corpo $\mathbb{K}$ como um espaço vetorial sobre si mesmo, qualquer elemento diferente de zero $\boldsymbol{x}$ de $\mathbb{K}$ é uma base de $\mathbb{K}$.*

$\lhd$

Exemplo 6. *Suponha $V = \mathbb{K}^n, n \in \mathbb{N}$. Para cada $i = 1, \ldots,$ n, seja $\boldsymbol{e}_i = (0, \ldots, 1, \ldots, 0)$. Portanto, $\boldsymbol{e}_i$ é a n-tupla cujas entradas são todas zero, exceto um 1 na i-ésima posição. Defina $\underline{B} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. Como $(x_1, \ldots, x_n) = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$, vemos $\underline{B}$ é uma base de $\mathbb{K}^n$. Vamos denominar $\underline{B}$ de base canônica (padrão) de $\mathbb{K}^n$.*

$\lhd$

Exemplo 7. *Suponha $V = (\mathbb{K}^\infty)_0 \triangleq (\mathbb{K}^{\mathbb{N}})_0$. Para cada $i \in \mathbb{N}$, seja $\underbrace{(0, \ldots, 1, \ldots,)}_{1 \text{ na } i\text{-ésima posição}}$. Ou seja, $\boldsymbol{e}_i$ é o vetor cujas entradas são todas zero, exceto um 1 na i-ésima posição.*

Defina $\underline{B} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n, \ldots\}$. Como todo vetor em $\boldsymbol{x} \in (\mathbb{K}^\infty)_0$ possui um número finito de coordenadas diferente de 0 $\boldsymbol{x} = (x_1, \ldots, x_n, 0, 0, \ldots) = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$, temos que $\underline{B}$ é uma base de $(\mathbb{K}^\infty)_0$. Vamos denominar $\underline{B}$ de base canônica de $(\mathbb{K}^\infty)_0$.

$\lhd$

Exemplo 8. *Seja $V = \mathcal{M}_{m,n}(\mathbb{K})$. Para qualquer $i = 1, \ldots, m$ e $j = 1, \ldots, n$, $\boldsymbol{e}_{ij}$ denota a matriz $m \times n$ cujas entradas são todas zero, exceto um 1 na posição (ij). Como $(a_{ij}) = \sum_{i,j} a_{ij}\boldsymbol{e}_{ij}$, temos que $\underline{B} = \{\boldsymbol{e}_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ é uma base para V. Os elementos $\boldsymbol{e}_{ij}$ em $\underline{B}$ são denominadas de unidades matriciais de $\mathcal{M}_{m,n}(\mathbb{K})$.*

◁

Exemplo 9. *Seja $V = \mathbb{K}[x]$. Seja $\underline{B}$ o conjunto dos monômios mônicos em x, assim, $\underline{B} = \{1, x, x^2, \ldots\}$ e é uma base de $\mathbb{K}[x]$.*

◁

Exemplo 10. *O conjunto dos números reais é um espaço vetorial sobre o corpo dos números racionais.*

Qual seria a base para esse espaço vetorial? Os elementos $\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7}, \ldots$, são linearmente independentes, mas certamente não geram $\mathbb{R}$, pois também precisamos de elementos como $e, e^2, e^3, \ldots$, que também formam um conjunto linearmente independente. De fato, como $\mathbb{Q}$ é enumerável, pode-se mostrar que o subespaço de $\mathbb{R}$ gerado por qualquer subconjunto enumerável de $\mathbb{R}$ deve ser enumerável. Como o próprio $\mathbb{R}$ é não-enumerável, nenhum conjunto enumerável pode ser uma base para $\mathbb{R}$ sobre $\mathbb{Q}$. Isso significa que qualquer base para $\mathbb{R}$ sobre $\mathbb{Q}$, deve ser não enumerável e assim difícil de descrever.

◁

Um dos fatos mais fundamentais é que todo espaço vetorial possui uma base:

2.11 Teorema Base

Todo espaço vetorial possui uma base.

De fato, temos um resultado ligeiramente mais forte: qualquer subconjunto linearmente independente S de V pode ser expandido para uma base. É esse fato que provaremos.

2.12 Teorema Extensão

Seja V um espaço vetorial sobre $\mathbb{K}$ e suponha que S seja um subconjunto linearmente independente de V. Então, existe uma base $\underline{B}$ de V tal que $\underline{B} \supseteq S$.

Demonstração. Denotaremos por $\mathcal{I}$ o conjunto de todos os subconjuntos independentes de V que contêm S. Assim,

$$\mathcal{I} = \{A \in \mathscr{P}(V) \mid A \supseteq S \text{ e A é linearmente independente sobre } \mathbb{K}\}.$$

Observamos que como $S \in \mathcal{I}$ então $\mathcal{I} \neq 0$. Ordenamos parcialmente $\mathcal{I}$ por inclusão. Assim, para $A_1, A_2 \in \mathcal{I}$,

$$A_1 \leq A_2 \text{ se, e somente se, } A_1 \subseteq A_2.$$

Dessa forma $(\mathcal{I}, \subseteq)$ é um conjunto parcialmente ordenado.

Suponha que $\mathcal{T} = \{A_i \,|\, i \in I\}$ seja uma coleção indexada de elementos de $\mathcal{S}$ que formam um subconjunto totalmente ordenado de $\mathcal{I}$. Mostraremos que $\mathcal{T}$ tem um limite superior. Para isso, defina $A = \bigcup_{i \in I} A_i$. Claramente, $A \in \mathscr{P}(V), S \subseteq A$ e $A_i \subseteq A$ para todos os $i \in I$. Se A não fosse linearmente independente, existiria um subconjunto finito $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subseteq A$ e escalares diferentes de zero $x_1, \ldots, x_n \in \mathbb{K}$ satisfazendo $x_1\boldsymbol{x}_1 + \cdots + x_n\boldsymbol{x}_n = 0$. Como $\mathcal{T}$ é totalmente ordenado, existe um índice $i_0 \in I$, de modo que $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subseteq A_{i_0}$. Então A_{i_0} seria linearmente dependente, o que é impossível, já que $A_{i_0} \in \mathcal{I}$.

Concluímos que A é linearmente independente e, consequentemente, $A \in \mathcal{I}$. Portanto, $\mathcal{T}$ tem um limite superior A em $\mathcal{T}$. Como $\mathcal{T}$ era arbitrário, agora podemos concluir que $(\mathcal{I}, \subseteq)$ é um conjunto indutivo. Aplicando o Lema de Zorn, temos que $\mathcal{I}$ tem um elemento maximal $\underline{\mathrm{B}}$. Como $\underline{\mathrm{B}} \in \mathcal{I}, \underline{\mathrm{B}} \supseteq S$ e $\underline{\mathrm{B}}$ é linearmente independente. Afirmamos que $\underline{\mathrm{B}}$ é de fato uma base de V. Para provar essa afirmação, precisamos apenas argumentar $\langle \underline{\mathrm{B}} \rangle = V$. Suponha que $\langle \underline{\mathrm{B}} \rangle \neq V$ então existiria um vetor $\boldsymbol{x} \in V \setminus \langle \underline{\mathrm{B}} \rangle$. Como $\boldsymbol{x} \notin \langle \underline{\mathrm{B}} \rangle$, o conjunto $\underline{\mathrm{B}} \cup \{\boldsymbol{x}\}$ é claramente linearmente independente. Consequentemente teríamos que $\underline{\mathrm{B}} \cup \{\boldsymbol{x}\} \in \mathcal{I}$ e que $\underline{\mathrm{B}} \cup \{\boldsymbol{x}\}$ é estritamente maior que $\underline{\mathrm{B}}$. Isso contrária o fato que $\underline{\mathrm{B}}$ é maximal em $\mathcal{I}$. Portanto, $\langle \underline{\mathrm{B}} \rangle = V$ e consequentemente $\underline{\mathrm{B}}$ é uma base de V contendo S. ■

do Teorema 2.3. Suponha $V \neq \{\mathbf{0}\}$ e seja $\boldsymbol{v} \neq \mathbf{0}$. Então $\{\boldsymbol{v}\}$ é um conjunto linearmente independente e pelo Teorema 2.3 pode ser completado a uma base de V. ■

Há uma grande desvantagem nessa demonstração de que todo espaço vetorial admite uma base: a menos que a dimensão seja finita ou pelo menos enumerável, ela não nos fornece nenhuma ideia de como encontrar uma base. De fato, esse é um problema sério com o conceito de base para espaços dimensionais infinitos em geral. Embora o Lema de Zorn nos permita demonstrar que existe uma base, na prática, esse fato pode ser inútil se não tivermos um procedimento para encontrar uma. Mas que elas existem...

Exemplo 13. *Uma base para o espaço vetorial $C^k(I)$ é não enumerável e não possui descrições simples. No entanto, como $\mathbb{R}[x] \subseteq C^k(I)$, o Teorema 2.3 garante que existe uma base de $C^k(I)$ que contém os monômios $1, x, x^2, \ldots$.*

◁

O Teorema 2.3 diz que qualquer subconjunto linearmente independente de V pode ser expandido para uma base de V. Há um resultado complementar, se algum subconjunto S de V gerar V, então S conterá uma base de V.

2.14 Teorema Redução

Seja V um espaço vetorial sobre $\mathbb{K}$ e suponha que $V = \langle S \rangle$. Então S contém uma base de V.

Demonstração. Se $S = \emptyset$ ou $\{\mathbf{0}\}$, então $V = (0)$. Nesse caso, $\emptyset$ é uma base de V contida em S. Portanto, podemos supor que S contém um vetor diferente de zero $\boldsymbol{x}$.

Seja $\mathcal{I} = \{A \subseteq S \,|\, A \text{ linearmente independente sobre } \mathbb{K}\}$. Claramente, $\{\boldsymbol{x}\} \in \mathcal{I}$. Ordenaremos parcialmente $\mathcal{I}$ por inclusão.

Se $\mathcal{T} = \{A_i \,|\, i \in I\}$ é um subconjunto totalmente ordenado de $\mathcal{I}$, então $\bigcup_{i \in I} A_i$ é um limite superior para $\mathcal{T}$ em $\mathcal{I}$.

Portanto, $(\mathcal{I}, \subseteq)$ é indutivo. Aplicando o Lema de Zorn temos que $\mathcal{I}$ possui um elemento maximal $\underline{\mathrm{B}}$.

Afirmamos que $\underline{\mathrm{B}}$ é uma base para V. Como $\underline{\mathrm{B}} \in \mathcal{I}, \underline{\mathrm{B}} \subseteq S$ e $\underline{\mathrm{B}}$ é linearmente independente sobre $\mathbb{K}$. Se $\langle \underline{\mathrm{B}} \rangle = V$, então $\underline{\mathrm{B}}$ é uma base de V, e o resultado segue. Suponha que $\langle \underline{\mathrm{B}} \rangle \neq V$. Então $S \not\subseteq \langle \underline{\mathrm{B}} \rangle$ ou caso contrário $V = \langle S \rangle \subseteq \langle \langle \underline{\mathrm{B}} \rangle \rangle = \langle \underline{\mathrm{B}} \rangle$. Portanto, existe um vetor $\boldsymbol{y} \in S \setminus \langle \underline{\mathrm{B}} \rangle$. Claramente, $\underline{\mathrm{B}} \cup \{\boldsymbol{y}\}$ é linearmente independente sobre $\mathbb{K}$. Assim, $\underline{\mathrm{B}} \cup \{\boldsymbol{y}\} \in \mathcal{I}$. Mas $\boldsymbol{y} \notin \langle \underline{\mathrm{B}} \rangle$ implica $\boldsymbol{y} \notin \underline{\mathrm{B}}$. Portanto, $\underline{\mathrm{B}} \cup \{\boldsymbol{y}\}$ é estritamente maior que $\underline{\mathrm{B}}$ em $\mathcal{I}$. Como $\underline{\mathrm{B}}$ é maximal, isso é uma contradição. Portanto, $\langle \underline{\mathrm{B}} \rangle = V$ e nossa prova está concluída. ■

Um espaço vetorial V possui muitas bases diferentes. Por exemplo, $\underline{\mathrm{X}} = \{(0, \dots, \lambda, \dots, 0) + \boldsymbol{e}_i \,|\, i = 1, \dots, n\}$ é claramente uma base para $\mathbb{K}^n$ para qualquer $\lambda \neq -1$ em $\mathbb{K}$. O que todas as bases de V têm em comum é sua cardinalidade. Provamos esse fato em nosso próximo teorema.

2.15 Lema Troca de Steinitz

Seja $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ uma coleção de vetores linearmente independentes em um espaço vetorial V. Seja $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ seja uma coleção de vetores que geram V. Então, $m \leq n$ e, além disso, possivelmente após reordenar os vetores $\boldsymbol{w}_i$, o conjunto $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m, \boldsymbol{w}_{m+1}, \dots, \boldsymbol{w}_n\}$ gera V.

Demonstração. Suponha que $V = \boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ e $W = \boldsymbol{w}_1, \dots, \boldsymbol{w}_n$. Desejamos mostrar que $m \leq n$, e que após rearranjar os $\boldsymbol{w}_j$ se necessário, o conjunto $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m, \boldsymbol{w}_{m+1}, \dots, \boldsymbol{w}_n$ gera V. Vamos proceder por indução em m.

Para o caso base, suponha que m seja zero. Nesse caso, a afirmação é verdadeira porque não há vetores $\boldsymbol{v}_i$, e o conjunto $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ gera V por hipótese.

Para o passo indutivo, assuma que a proposição é verdadeira para $m-1$. Pela hipótese de indução, podemos reorganizar os $\boldsymbol{w}_i$ de modo que $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m-1}, \boldsymbol{w}_m, \ldots, \boldsymbol{w}_n$ gere V. Como $\boldsymbol{v}_m \in V$, existem coeficientes $a_1, \ldots, a_n$ tais que

$$\boldsymbol{v}_m = \sum_{j=1}^{m-1} a_j \boldsymbol{v}_j + \sum_{j=m}^{n} a_j \boldsymbol{w}_j.$$

Pelo menos um dos $a_m, \ldots, a_n$ deve ser diferente de zero, pois caso contrário, essa igualdade contradiria a independência linear de $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$; portanto, $m \leq n$. Reordenando os $a_m\boldsymbol{w}_m, \ldots, a_n\boldsymbol{w}_n$ se necessário, podemos assumir que a_m é não nulo. Portanto, temos

$$\boldsymbol{w}_m = \frac{1}{a_m}\left(\boldsymbol{v}_m - \sum_{j=1}^{m-1} a_j \boldsymbol{v}_j - \sum_{j=m+1}^{n} a_j \boldsymbol{w}_j\right).$$

Em outras palavras, $\boldsymbol{w}_m$ está no espaço gerado por $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{w}_{m+1}, \ldots, \boldsymbol{w}_n$. Visto que esse espaço contém cada um dos vetores $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m-1}, \boldsymbol{w}_m, \boldsymbol{w}_{m+1}, \ldots, \boldsymbol{w}_n$, pela hipótese de indução ele contém V. ■

Existem duas afirmações aqui. A primeira é que $m \leq n$. Em palavras, isso diz que qualquer conjunto linearmente independente é no máximo tão grande quanto qualquer conjunto que gera.

A segunda parte é que, se pegarmos um monte de vetores linearmente independentes e qualquer conjunto de vetores que geram, podemos emprestar parte do conjunto que gera para estender o conjunto linearmente independente para um conjunto que gere todo o espaço vetorial.

A conclusão realmente profunda do Lema da Troca de Steinitz é que faz sentido falar sobre a dimensão de um espaço vetorial.

2.16 Teorema da Invariância da Dimensão

Seja V um espaço vetorial sobre $\mathbb{K}$ e suponha que $\underline{B}_1$ e $\underline{B}_2$ sejam duas bases de V. Então $|\underline{B}_1| = |\underline{B}_2|$.

Demonstração. Dividimos essa prova em dois casos.

Caso 1: suponha que V tenha uma base $\underline{B}_1$ finita.

Nesse caso, provaremos que $|\underline{B}_2| = |\underline{B}_1|$. Suponha que $\underline{B}_1 = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$. É claramente suficiente mostrar que $|\underline{B}_2| = n$. Supomos que $|\underline{B}_2| \neq n$ e derivamos uma contradição. Existem duas possibilidades a serem consideradas aqui: $|\underline{B}_2| = m < n$ ou $|\underline{B}_2| > n$. Vamos primeiro supor que $\underline{B}_2 = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$ com $m < n$.

Então pelo Lema 2.3 $m \geq n$ e a contradição segue.

Agora suponha que $|\underline{B}_2| > n$ ($|\underline{B}_2|$ pode ser infinito aqui).

Como $\underline{B}_2$ é linearmente independente todo conjunto finito de $\underline{B}_2$ é linearmente independente. Como $\underline{B}_1$ gera, todo conjunto finito de $\underline{B}_2$ tem no máximo $|\underline{B}_1|$ elementos e logo $|\underline{B}_2| < |\underline{B}_1|$. Consequentemente $|\underline{B}_2| < n$ e a contradição segue.

Caso 2: suponha que nenhuma base de V seja finita.

Nesse caso, $\underline{B}_2$ e $\underline{B}_1$ são conjuntos infinitos. Seja $\boldsymbol{x} \in \underline{B}_2$. Como $\underline{B}_1$ é uma base de V, existe um único subconjunto finito $I_x \subseteq \underline{B}_1$, tal que $\boldsymbol{x} \in \langle I_x \rangle$ e $\boldsymbol{x} \notin \langle I' \rangle$ para qualquer subconjunto próprio I' de I_x. Portanto, temos uma função bem definida $\varphi : \underline{B}_2 \to \mathscr{P}(\underline{B}_1)$ fornecida por $\varphi(\boldsymbol{x}) = I_x$.

Como $\underline{B}_2$ é infinito, podemos aplicar o Lema A.3.1 e concluir que $|\underline{B}_2| \geq |\bigcup_{\boldsymbol{x} \in \underline{B}_1} I_{\boldsymbol{x}}|$.

Como $\boldsymbol{x} \in \langle I_{\boldsymbol{x}} \rangle$ para todos os $\boldsymbol{x} \in \underline{B}_2$, $V = \langle \bigcup_{\boldsymbol{x} \in \underline{B}_2} I_{\boldsymbol{x}} \rangle$. Portanto, $\bigcup_{\boldsymbol{x} \in \underline{B}_2} I_x$ é um subconjunto de $\underline{B}_1$ que gera o espaço V. Logo concluímos que $\bigcup_{\boldsymbol{x} \in \underline{B}_2} I_x = \underline{B}_1$. Em particular, $|\underline{B}_2| \geq |\underline{B}_1|$. Invertendo os papeis de $\underline{B}_2$ e $\underline{B}_2$ temos que $|\underline{B}_1| \geq |\underline{B}_2|$. Isso completa a prova do teorema. ■

2.17 Definição Dimensão

A cardinalidade comum de qualquer base de V é denominada **dimensão** de V. Escreveremos $\dim V$ para a dimensão de V. Se queremos enfatizar em que corpo estamos, então usaremos a notação $\dim_{\mathbb{K}} V$ para a dimensão do espaço vetorial sobre K. Assim, $\dim V = |\underline{B}|$, onde $\underline{B}$ é uma base qualquer de V quando o corpo base $\mathbb{K}$ é subentendido.

Vamos verificar as dimensões de alguns dos nossos exemplos anteriores. No Exemplo 2, $\dim_{\mathbb{K}_2} V = 1$ e $\dim_{\mathbb{K}_1}(V) = |\mathbb{R}|$, a cardinalidade de $\mathbb{R}$. No Exemplo 5, $\dim_{\mathbb{K}}(0) = 0$. No Exemplo 6, $\dim \mathbb{K}^n = n$. No exemplo 8, $\dim \mathcal{M}_{m,n}(\mathbb{K}) = m \cdot n$. No Exemplo 9, $\dim V = |\mathbb{N}|$, a cardinalidade de $\mathbb{N}$.

Se a dimensão de um espaço vetorial V for infinita, como nos Exemplos 2 e 9, em geral não faremos nenhuma tentativa de distinguir qual a cardinalidade de $\dim V$. Em vez disso, escreveremos simplesmente $\dim V = \infty$. Se V tiver uma base finita $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$, diremos que V é um espaço vetorial de dimensão finita e escreveremos $\dim V < \infty$, ou, mais precisamente, $\dim V = n < \infty$. Por exemplo, $\dim_{\mathbb{R}} C^k(I) = \infty$, enquanto $\dim_{\mathbb{R}} \mathbb{R}^n = n < \infty$.

2.18 Teorema

Seja V um espaço vetorial sobre $\mathbb{K}$.

a Se W for um subespaço de V, então $\dim W \leq \dim V$.

$\boxed{b}$ Se V for de dimensão finita e W é um subespaço de V tal que $\dim W = \dim V$, $W = V$.

Demonstração. O Teorema 2.3 diz que qualquer base de um subespaço W de V pode ser completada para uma base de V. Isso prova imediatamente os itens $\boxed{a}$ e $\boxed{b}$. ∎

Se V não é de dimensão finita, então $\boxed{b}$ é falso em geral. Um exemplo simples ilustra esse ponto.

Exemplo 19. *Seja* $V = \mathbb{K}[x]$ *e seja* W *o subespaço de* V *que consiste em todos os polinômios pares. Assim,* $W = \left\{\sum a_i x^{2i} \,|\, a_i \in \mathbb{K}\right\}$. *Uma base de* W *é claramente os monômios de potências pares de* x. *Assim,* $\dim V = \dim W$, *mas* $W \neq V$.

◁

Exercícios

Ex. 2.21 --- Suponha que S é finito. Exiba uma base para o espaço de funções $\mathbb{K}(S) = S^{\mathbb{K}}$

Ex. 2.22 --- Dado $L \subset V$, prove que se $\dim(L) = \dim(V) < \infty$. então $L = V$.

Ex. 2.23 ---

1. Prove que os únicos subespaços de $\mathbb{R}$, são o próprio $\mathbb{R}$ e o subespaço nulo
2. Prove que todos os subespaços de $\mathbb{R}^2$ são o próprio $\mathbb{R}^2$, o subespaço nulo ou o subespaço consistindo de um múltiplo de um vetor fixo em $\mathbb{R}^2$.
3. Quais são todos os subespaços de $\mathbb{R}^3$?

Ex. 2.24 --- Dado L um espaço n-dimensional sobre um corpo finito com q-elementos.

1. Calcule o número de subespaços k-dimensionais em L. Para $1 \leqslant k \leqslant n$
2. Calcule o número de pares de subespaços L_1 e L_2 com $\dim(L_1)$, $\dim(L_2)$ e $\dim(L_1 \cap L_2)$ fixos. Verifique que quando $q \to \infty$ o número de pares relativos em posição geral sobre o número total de pares com $\dim(L_1)$, $\dim(L_2)$ dados se aproxima a 1.

Ex. 2.25 --- Seja $X \subseteq V$ um subconjunto de um espaço vetorial sobre $\mathbb{K}$ V. Mostre que $\langle X \rangle = \{\sum_{v \in F} \alpha_v \cdot v : F \subseteq X$ é um subconjunto finito e $\alpha_v \in \mathbb{K}, \forall v \in F\}$.

Ex. 2.26 --- Seja $\underline{B}$ um subconjunto de um espaço vetorial V. Mostre que $\underline{B}$ é linearmente dependente se, e somente se, existir $v \in \underline{B}$ que pode ser escrito como combinação linear dos elementos de $\underline{B} \setminus \{v\}$.

Ex. 2.27 --- Se $\underline{E} = \{\boldsymbol{e}_i\}_{i\in I}$ é uma base para V e $\underline{F} = \{\boldsymbol{f}_j\}_{j\in J}$ é uma base para W. Mostre que $\{(\boldsymbol{e}_i, 0)\}_{i\in I} \cup \{(0, \boldsymbol{f}_j)\}_{j\in J}$ é base para $V \boxplus W$.

Ex. 2.28 --- Seja V um espaço vetorial sobre $\mathbb{K}$ de dimensão não necessariamente finita e seja $\underline{B}$ um conjunto linearmente independente em V. Mostre que se existir um elemento $v \in V$ que não seja combinação linear de elementos de $\underline{B}$, então o conjunto $\underline{B} \cup \{v\}$ é linearmente independente.

Ex. 2.29 --- Seja V um espaço vetorial sobre $\mathbb{K}$. Mostre que:

1. Dados vetores v_1, v_2 não nulos. Então os vetores v_1, v_2 são linearmente independentes se, e somente se, $\langle v_1\rangle \cap \langle v_2\rangle = \{0\}$.

2. Dados vetores v_1, v_2, v_3 não nulos. Prove que $\langle v_1\rangle \cap \langle v_2\rangle \cap \langle v_3\rangle = \{0\}$ não implica que os vetores v_1, v_2, v_3 sejam linearmente independentes.

3. Dado $S = \{s_1, \dots, s_n\} \subset V$. Prove que S é linearmente independente se, e somente se, $\langle S \backslash s_i\rangle \neq \langle S\rangle$ para todo $s_i \in S$.

4. Prove que se $A, B \subset V$. Então $\langle A\rangle + \langle B\rangle = \langle A \cup B\rangle$.

Ex. 2.30 --- Mostre que se os coeficientes $a_1, \dots, a_n$ não são todos iguais a zero, o hiperplano

$$H = \{(x_1, \dots, x_n) \in \mathbb{R}^n \,|\, a_1x_1 + \cdots + a_nx_n = 0\}$$

é um subespaço vetorial de dimensão $n-1$ em $\mathbb{R}^n$.

Ex. 2.31 --- Prove que em qualquer conjunto de vetores S existe um subconjunto S' linearmente independente tal que $\langle S\rangle = \langle S'\rangle$.

Ex. 2.32 --- Se $\mathbb{K} = \mathbb{Z}_2$, o subconjunto $\{(\bar{1},\bar{1},\bar{0}), (\bar{1},\bar{0},\bar{1}), (\bar{0},\bar{1},\bar{1})\}$ de $\mathbb{K}^3$ é linearmente dependente? e se $\mathbb{K} = \mathbb{Z}_{13}$?

Ex. 2.33 --- Seja $V = \mathcal{F}(\mathbb{R}, \mathbb{C})$ o $\mathbb{C}$-espaço vetorial de todas as funções de $\mathbb{R}$ em $\mathbb{C}$. Prove que $\{f_1, f_2, f_3\}$ é linearmente independente onde $f_1(x) = 1$,$f_2(x) = e^{ix} = \cos(x) + i\,\mathrm{sen}(x)$ e $f_3(x) = e^{-ix}$ para todo $x \in \mathbb{R}$.

Ex. 2.34 --- Considere o espaço das funções $\mathcal{F}(\mathbb{R}, \mathbb{R})$. Mostre que os seguintes subconjuntos sejam linearmente independentes.

1. $(f_n)_{n\in\mathbb{N}}$ onde $f_n : x \mapsto e^{nx}$

2. $(f_a)_{a\in\mathbb{R}}$ onde $f_a : x \mapsto |x-a|$

Ex. 2.35 --- Seja V um espaço vetorial sobre $\mathbb{R}$ e considere no conjunto $V_{\mathbb{C}} = \{(u,v) : u,v \in V\}$ as seguintes operações de adição e multiplicação por um número complexo:

$$(u_1,v_1)+(u_2,v_2) = (u_1+u_2, v_1+v_2)$$
$$(\alpha+i\beta)\cdot(u,v) = (\alpha u-\beta v, \beta u+\alpha v)$$

1. Mostre que $V_{\mathbb{C}}$ é um espaço vetorial sobre $\mathbb{C}$.

2. Seja $\{v_1, v_2, \ldots, v_n\} \subseteq V$ um subconjunto linearmente independente. Mostre que $\{(v_1,0),(v_2,0),\ldots,(v_n,0)\}$ e $\{(0,v_1),(0,v_2),\ldots,(0,v_n)\}$ são subconjuntos linearmente independente.em $V_{\mathbb{C}}$.

Ex. 2.36 --- Para um $\mathbb{C}$-espaço vetorial V, denotaremos por $V_{\mathbb{R}}$ o conjunto V visto como $\mathbb{R}$-espaço vetorial. Mostre que se $\{v_1, v_2, \ldots, v_n\}$ for um subconjunto linearmente independente.em V então $\{v_1, v_2, \ldots, v_n\}$ e $\{v_1, v_2, \ldots, v_n\} \cup \{iv_1, iv_2, \ldots, iv_n\}$ são subconjuntos linearmente independente.em $V_{\mathbb{R}}$.

Ex. 2.37 --- Seja V um espaço vetorial, mostre que são equivalentes:

1. B é uma base de V

2. B é um conjunto linearmente independente maximal (Se A é um subconjunto linearmente independente de V, dizemos que é um conjunto linearmente independente máximo se a adição de qualquer vetor a A resultar em um conjunto que não é linearmente independente.)

3. B é um conjunto gerador minimal (Se A é um conjunto que gera V dizemos que é um conjunto gerador minimal se a remoção de qualquer vetor em tudo de A resultará em um conjunto que não gera V).

Ex. 2.38 --- Seja $V = \langle X \rangle$ mostre que existe uma base B de V tal que B $\subseteq X$. Dica: Defina um conjunto parcialmente ordenado e use o lema de Zorn.

Ex. 2.39 --- Ache uma base de $\mathcal{M}_{n,n}(\mathbb{C})$ como espaço vetorial sobre $\mathbb{R}$. Qual é a $\dim_{\mathbb{R}}(\mathcal{M}_{n,n}(\mathbb{C}))$?

Ex. 2.40 --- Seja S o $\mathbb{R}$-espaço vetorial do Exercício 2.14. Mostre que $\dim_{\mathbb{R}} S = n$.

Dica: Use o Teorema de Existência e Unicidade de soluções: Considere a equação

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1 y'(t) + a_0 y(t) = 0 \tag{2.6}$$

onde $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$. Dados $A_0, A_1, \ldots, A_{n-1} \in \mathbb{R}$, existe uma única solução $y : \mathbb{R} \to \mathbb{R}$ da equação (2.6) verificando $y(0) = A_0$, $y'(0) = A_1, \ldots, y^{(n-1)}(0) = A_{n-1}$ (condições iniciais da equação (2.6)).

Construa n soluções que formarão uma base do espaço das soluções de (2.6), da seguinte forma: considerando as condições iniciais $A_0 = 1$ e $A_1 = \cdots = A_{n-1} = 0$ o Teorema de Existência e Unicidade garante que existe uma única solução $y_1 : \mathbb{R} \to \mathbb{R}$ de (2.6) que verifica as condições $y_1(0) = 1$ e $y_1'(0) = \cdots = y_1^{(n-1)}(0) = 0$. Repita o procedimento considerando as condições iniciais $A_i = 1$ e $A_j = 0$ para todo $j \neq i$ com i variando de 1 a $n-1$.

Ex. 2.41 --- Determine se os espaços abaixo têm dimensão finita. Se sim determine a dimensão e uma base para o espaço:

1. O conjunto de todas as sequências reais.
2. O conjunto das sequências reais que satisfazem $a_k = a_{k-1} + a_{k-2}$ para $k \geqslant 3$.
3. $\mathbb{C}^n$ visto como um espaço vetorial sobre $\mathbb{C}$ e visto como um espaço vetorial sobre $\mathbb{R}$.
4. O conjunto das sequências com apenas um número finito de termos não nulos.
5. O espaço das funções em $\mathcal{F}(X, \mathbb{R})$, $|X| < \infty$ que se anulam em todos os pontos de um subconjunto $X_0 \subset X$.
6. O espaço das funções $C([0, 1], \mathbb{R})$.

Ex. 2.42 --- Dado um corpo $\mathbb{F}$. Um subcorpo $\mathbb{K}$ é um subconjunto de $\mathbb{F}$ que é corpo quando restringimos as operações de $\mathbb{F}$ a $\mathbb{K}$.

1. Mostre que $\mathbb{F}$ é espaço vetorial sobre $\mathbb{K}$.
2. Suponha que V é um subespaço m-dimensional sobre $\mathbb{F}$. Suponha que $\mathbb{F}$ é um espaço n-dimensional sobre $\mathbb{K}$. Qual a dimensão de V sobre $\mathbb{K}$?

Ex. 2.43 --- Seja W um subespaço de V com $\dim(V) < \infty$. Prove que:

1. $\dim(W) \leq \dim(V)$

2. $\dim(W) = \dim(V)$ se, e somente se, $W = V$.

Ex. 2.44 --- Mostre que $\mathbb{R}$ é um espaço vetorial de dimensão infinita sobre $\mathbb{Q}$.

2.4 Soma de Subespaços

O conjunto $\mathcal{S}(V)$ de todos os subespaços de um espaço vetorial V possui uma estrutura rica que apresentamos a seguir. Essa estrutura motiva e justifica algumas das definições que apresentamos. Em $\mathcal{S}(V)$ temos uma relação de ordem dada pela inclusão de conjuntos. Com essa ordem $\mathcal{S}(V)$ é parcialmente ordenado pela inclusão de conjuntos. Nessa ordem, o subespaço zero $\{\mathbf{0}\}$ é o menor elemento em $\mathcal{S}(V)$ e o espaço V é o maior elemento.

Se $S, T \in \mathcal{S}(V)$, então é fácil de demonstrar que $S \cap T$ é o maior subespaço de V contido simultaneamente em S e T. Em termos de inclusão de conjuntos, $S \cap T$ é o maior limite inferior de S e T. Da mesma forma, se $\{S_i \mid i \in I\}$ é uma coleção de subespaços de V, então sua interseção é o maior limite inferior dos subespaços

$$\inf\{S_i \mid i \in I\} = \bigcap_{i \in I} S_i$$

Para determinar o menor subespaço de V contendo os subespaços S e T, fazemos a seguinte definição.

2.1 Definição Soma

Se S e T são subespaços de V. A **soma** $S + T$ é definida por

$$S + T \triangleq \{\boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{x} \in S,\ \boldsymbol{y} \in T\}$$

Em geral, a soma de qualquer coleção $\{S_i \mid i \in I\}$ de subespaços é o conjunto de todas as somas **finitas** de vetores da união $\bigcup_{i \in I} S_i$

$$\sum_{i \in I} S_i \triangleq \{\boldsymbol{x}_1 + \cdots + \boldsymbol{x}_n \mid \boldsymbol{x}_j \in \bigcup_{i \in I} S_i\}$$

2.2 Proposição

Dado uma coleção $\{S_i \mid i \in I\}$ de subespaços. Então são equivalentes:

1. W é o menos subespaço contendo S_i para $i \in I$

2. $W = \langle \bigcup\{S_i \mid i \in I\}\rangle$
3. $W = \sum_{i\in I} S_i$

A demonstração será deixada como exercício.

2.4.1 Soma Direta

Como veremos, existem muitas maneiras de construir novos espaços vetoriais a partir dos antigos.

Somas Diretas Internas

2.3 Definição Soma Direta

Seja V é um espaço vetorial. Dizemos que V é a soma direta (interna) de uma família $\mathcal{S} = \{S_i \mid i \in I\}$ de subespaços de V se todo vetor $v \in V$ puder ser escrito, de uma maneira única (exceto pela ordem), como uma soma finita de vetores dos subespaços em $\int$, isto é, se para todo $\boldsymbol{v} \in V$,

$$\boldsymbol{v} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_n \text{ com } \boldsymbol{v}_i \in S_i \text{ e}$$
$$\boldsymbol{v} = \boldsymbol{v}'_1 + \cdots + \boldsymbol{v}'_m, \text{ com } \boldsymbol{v}'_i \in S_i$$

Então $m = n$ e (após trocar os índices se necessário) $\boldsymbol{v}_i = \boldsymbol{v}'_i$ para $i = 1, \ldots, n$.

Se V é a soma direta de $\mathcal{S}$, escrevemos

$$V = \bigoplus_{i\in I} S_i.$$

Se $\mathcal{S} = \{S_1, \ldots, S_n\}$ é uma família finita, podemos denotar a soma direta como

$$V = S_1 \oplus \cdots \oplus S_n$$

Observe que uma soma é direta se, e somente se, sempre que $\boldsymbol{u}_{i_1} + \cdots + \boldsymbol{u}_{i_n} = 0$ em que $\boldsymbol{u}_{i_j} \in S_{i_j}$ e $i_j \neq i_k$ então $\boldsymbol{u}_{i_j} = 0$ para todos os j, isto é, se, e somente se, 0 tiver uma representação única como uma soma de vetores de subespaços distintos. Essa afirmação motiva a seguinte definição e a próxima proposição.

2.4 Definição Espaços Independentes

Seja V um espaço vetorial e $\{S_1, \ldots, S_k\}$ seja um conjunto de subespaços de V. Esse conjunto de espaços é **independente** se $s_1 + \cdots + s_k = 0$ com $s_i \in S_i$ implica $s_i = 0$ para todo i.

Assim dados um espaço vetorial V e um conjunto $\{S_1, \ldots, S_k\}$ de subespaços de V. Então V é a soma direta $V = S_1 \oplus \cdots \oplus S_k$ se, e somente se,

a $V = S_1 + \cdots + S_k$ e

b $\{S_1, \ldots, S_k\}$ são independentes.

A seguinte caracterização de somas diretas é bastante útil.

2.5 Teorema

Um espaço vetorial V é a soma direta de uma família $\mathcal{S} = \{S_i \,|\, i \in I\}$ de subespaços se, e somente se,

a V é a soma dos S_i, $V = \sum_{i\in I} S_i$

b Para cada $i \in I$, $S_i \cap (\sum_{j\neq i} S_j) = \{\mathbf{0}\}$

Demonstração. Suponha primeiro que V seja a soma direta de $\mathcal{S}$. Então *a* é verdadeiro e se

$$\boldsymbol{v} \in S_i \cap (\sum_{j\neq i} S_j)$$

então $\boldsymbol{v} = \boldsymbol{s}_i$ para algum $\boldsymbol{s}_i \in S_i$ e

$$\boldsymbol{v} = \boldsymbol{s}_{j_1} + \cdots + \boldsymbol{s}_{j_n}$$

onde $\boldsymbol{s}_{j_k} \in S_{j_k}$ e $j_k \neq i$ para todos os $k = 1, \ldots, n$. Portanto, pela unicidade das representações de soma direta, $\boldsymbol{s}_i = 0$ e, portanto, $\boldsymbol{v} = 0$. Logo demonstramos *b*.

Para a recíproca, suponha que *a* e *b* sejam válidos. Precisamos apenas verificar a condição de unicidade

$$\boldsymbol{v} = \boldsymbol{s}_{j_1} + \cdots + \boldsymbol{s}_{j_n} \quad \text{e} \tag{2.7}$$

$$\boldsymbol{v} = \boldsymbol{t}_{k_1} + \cdots + \boldsymbol{t}_{k_m} \tag{2.8}$$

onde $\boldsymbol{s}_{j_i} \in S_{j_i}$ e $\boldsymbol{t}_{k_i} \in S_{k_i}$ e incluindo termos adicionais iguais a 0, podemos assumir que o índice define $\{j_1, \ldots, j_n\}$ e $\{k_1, \ldots, k_m\}$ são o mesmo conjunto $\{i_1, \ldots, i_p\}$, isto é

$$\boldsymbol{v} = \boldsymbol{s}_{i_1} + \cdots + \boldsymbol{s}_{i_p} \quad \text{e} \tag{2.9}$$

$$\boldsymbol{v} = \boldsymbol{t}_{i_1} + \cdots + \boldsymbol{t}_{i_p} \tag{2.10}$$

Mas $(\boldsymbol{s}_{i_1} - \boldsymbol{t}_{i_1}) + \cdots + (\boldsymbol{s}_{i_p} - \boldsymbol{t}_{i_p}) = 0$. Portanto, cada termo $\boldsymbol{s}_{i_u} - \boldsymbol{t}_{i_u} \in S_{i_u}$ é uma soma de vetores de subespaços diferentes de S_{i_u}, que só pode acontecer se $\boldsymbol{s}_{i_u} - \boldsymbol{t}_{i_u} = 0$. Portanto, $\boldsymbol{s}_{i_u} = \boldsymbol{t}_{i_u}$ para todos os i_u e assim demonstramos que V é a soma direta de $\mathcal{S}$. ∎

Se tivermos apenas dois subespaços $\{W_1, W_2\}$, essa condição simplesmente indicará $W_1 \cap W_2 = \{\mathbf{0}\}$. Se tivermos mais de dois subespaços, essa condição é mais forte que a condição $W_i \cap W_j = \{\mathbf{0}\}$ para $i \neq j$.

Exemplo 6. *Seja $\mathcal{M}_{n,n}(\mathbb{K})$ o espaço vetorial das matrizes $n \times n$ com entradas em $\mathbb{K}$. Seja $\mathfrak{sl}_n(\mathbb{K}) \subset \mathcal{M}_{n,n}(\mathbb{K})$ é o subespaço de matrizes cujos traço é zero. Seja $D \subset \mathcal{M}_{n,n}(\mathbb{K})$ o subespaço gerado pela matriz identidade I_n.*

É fácil verificar se $\mathfrak{sl}_n(\mathbb{K}) \cap D = \{\mathbf{0}\}$. Para ver que $\mathfrak{sl}_n(\mathbb{K}) + D = \mathcal{M}_{n,n}(\mathbb{K})$, considere qualquer matriz $A = [a_{ij}]$. Defina $\alpha = \operatorname{tr}(A)$ e $B = A - \alpha\,\mathrm{I}_n$. Então, claramente, $A = B + \alpha\,\mathrm{I}_n$, onde $\operatorname{tr}(B) = 0$ e $\alpha\,\mathrm{I}_n \in D$. Portanto, $\mathcal{M}_{n,n}(\mathbb{K}) = \mathfrak{sl}_n(\mathbb{K}) \oplus D$.

◁

Exemplo 7. *Toda matriz $A \in \mathcal{M}_{n,n}(\mathbb{K})$ pode ser decomposta como*

$$A = \frac{1}{2}(A + A^{\mathrm{t}}) + \frac{1}{2}(A - A^{\mathrm{t}}) = B + C \tag{2.11}$$

onde A^{t} é a transposta de A. É fácil verificar que B é uma matriz simétrica e C é assimétrica e assim temos uma decomposição de A como a soma de uma matriz simétrica e uma matriz simétrica.

Como os conjuntos $\mathrm{Sim}_{n,n}$ e $\mathrm{ASim}_{n,n}$ de todas as matrizes simétricas e assimétricas em $\mathcal{M}_{n,n}(\mathbb{K})$ são subespaços de $\mathcal{M}_{n,n}(\mathbb{K})$, temos que

$$\mathcal{M}_{n,n}(\mathbb{K}) = \mathrm{Sim}_{n,n} + \mathrm{ASim}_{n,n}$$

Além disso, se $S + T = S' + T'$, onde S e S' são simétricas e T e T' são antissimétricas, então a matriz

$$U = S - S' = T' - T$$

é simétrica e antissimétrica. Portanto, se $\operatorname{char}(\mathbb{K}) \neq 2$, devemos ter $U = 0$ e, portanto, $S = S'$ e $T = T'$. Assim se $\operatorname{char}(\mathbb{K}) \neq 2$ a soma é direta:

$$\mathcal{M}_{n,n}(\mathbb{K}) = \mathrm{Sim}_{n,n} \oplus \mathrm{ASim}_{n,n}$$

◁

2.4.2 Decomposição em Somas Diretas

2.8 Proposição

Seja V um espaço vetorial e seja $\{W_1, \dots, W_k\}$ um conjunto de subespaços de V. Seja, $\underline{\mathrm{B}}_i$ seja uma base de W_i, para todo i, e seja, $\underline{\mathrm{B}} = \cup \underline{\mathrm{B}}_i$. Então

- *a* $\underline{\mathrm{B}}$ gera V se, e somente se, $V = W_1 + \cdots + W_k$.
- *b* $\underline{\mathrm{B}}$ é linearmente independente se, e somente se, $\{W_1, \dots, W_k\}$ for independente.

$\boxed{c}$ $\underline{B}$ é uma base para V se, e somente se, $V = W_1 \oplus \cdots \oplus W_k$.

A demonstração da proposição será deixada como exercício

2.9 Corolário

Sejam V um espaço vetorial de dimensão finita e $\{W_1,\ldots, W_k\}$ um conjunto de subespaços com $V = W_1 \oplus \cdots \oplus W_k$. Então $\dim(V) = \dim(W_1) + \cdots + \dim(W_k)$.

A demonstração da proposição será deixada como exercício

2.10 Definição Complemento

Seja V um espaço vetorial e W_1 seja um subespaço de V. Então W_2 é um **complemento** de W_1 se $V = W_1 \oplus W_2$.

2.11 Teorema

Seja V um espaço vetorial e W seja um subespaço de V. Então W_1 possui um complemento W'. Além disso temos $\dim V = \dim W + \dim W'$

Demonstração. Suponha que W seja um subespaço de V. Seja $\underline{B}$ uma base de W. Pelo Teorema 2.3, existe uma base $\underline{B}'$ de V tal que $\underline{B} \subseteq \underline{B}'$. Seja $W' = \langle \underline{B}' - \underline{B} \rangle$. Como $\underline{B}' = \underline{B} \cup (\underline{B}' - \underline{B})$, $V = \langle \underline{B}' \rangle = \langle \underline{B} \rangle + \langle \underline{B}' - \underline{B} \rangle = W + W'$. Como $\underline{B}'$ é linearmente independente e $\underline{B} \cap (\underline{B}' - \underline{B}) = 0$, $\langle \underline{B} \rangle \cap \langle \underline{B}' - \underline{B} \rangle = (0)$. Assim, $W \cap W' = (0)$, e a prova de $\boxed{c}$ está completa.

■

Observação 12. *O subespaço W' de V construído no do Teorema 2.3 é chamado um **complemento** de W em V. Ressaltamos que em geral o complemento não é único(embora sejam isomorfos). Exceto quando $W_1 = \{\mathbf{0}\}$ (onde $W_2 = V$) ou $W_1 = V$ (onde $W_1 = \{\mathbf{0}\}$), o subespaço W_2 nunca é único. Sempre podemos escolher uma maneira diferente de estender, $\underline{B}_1$ para uma base de V, a fim de obter um W_2 diferente.*

2.13 Teorema

Se V for de dimensão finita e W_1 e W_2 são subespaços de V, então

$$\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim W_1 + \dim W_2.$$

Demonstração. Seja $\underline{B}_0 = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ uma base de $W_1 \cap W_2$. No caso particular em que $W_1 \cap W_2 = (0)$, faremos $\underline{B}_0 = \emptyset$. Pelo Teorema da Extensão 2.3, podemos completar $\underline{B}_0$ para uma base $\underline{B}_1 = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$ de W_1. De modo análogo, podemos completar $\underline{B}_0$ para uma base $\underline{B}_2 = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_p\}$ de W_2. Assim, $\dim W_1 \cap W_2 = n, \dim W_1 = n + m,$ e $\dim W_2 = n + p.$

Afirmamos que $\underline{B} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_m, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_p\}$ é uma base de $W_1 + W_2$. Claramente $\langle \underline{B} \rangle = W_1 + W_2$.

Precisamos apenas argumentar que $\underline{B}$ é linearmente independente. Suponha que

$$\sum_{i=1}^{n} x_i \boldsymbol{x}_i + \sum_{i=1}^{m} y_i \boldsymbol{y}_i + \sum_{i=1}^{p} z_i \boldsymbol{z}_i \text{ para alguns } x_i, y_i, z_i \in \mathbb{K}.$$

Podemos reescrever a equação anterior como

$$\sum_{i=1}^{p} z_i \boldsymbol{z}_i = -\sum_{i=1}^{n} x_i \boldsymbol{x}_i - \sum_{i=1}^{m} y_i \boldsymbol{y}_i = 0 \text{ para alguns } x_i, y_i, z_i \in \mathbb{K}.$$

Comparando ambos os lados da equação temos que $\sum_{i=1}^{p} z_i \boldsymbol{z}_i \in W_1 \cap W_2 = \langle \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \rangle$. Logo podemos escrever $\sum_{i=1}^{p} z_i \boldsymbol{z}_i$ como:

$$\sum_{i=1}^{p} z_i \boldsymbol{z}_i = \sum_{i=1}^{n} w_i \boldsymbol{x}_i$$

Mas como $\underline{B}_2$ é base os coeficientes coeficientes z_i são zero.

Assim, $\sum_{i=1}^{n} x_i \boldsymbol{x}_i = \sum_{i=1}^{m} y_i \boldsymbol{y}_i$. Como $\underline{B}_1$ é uma base de W_1, concluímos que $x_1 = \cdots = x_n = y_1 = \cdots = y_m = 0.$

Logo $\underline{B}$ é linearmente independente. Assim, $\dim(W_1 + W_2) = |\underline{B}| = n + m + p$ e a prova de $\boxed{d}$ segue. ■

Alguns comentários sobre o Teorema 2.4.2. Ele é verdadeiro, independentemente de V ser finito dimensional ou não. A prova é a mesma que a fornecida acima quando $\dim(W_1 + W_2) < \infty$. Se $\dim(W_1 + W_2) = \infty$, então W_1 ou W_2 é um subespaço de dimensão infinita com a mesma dimensão que $W_1 + W_2$. Assim, o resultado ainda é verdadeiro, mas bastante desinteressante.

Finalmente, o Teorema 2.4.2 pode ser generalizada para uma família finita de subespaços $W_1, \ldots, W_k$ de V.

2.14 Corolário

Seja V um espaço vetorial de dimensão finita da dimensão n. Suponha que $W_1, \dots, W_k$ sejam subespaços de V. Para cada $i = 1, \dots, k$, defina $c_i = n - \dim W_i$. Então

- $\dim(W_1 \cap \cdots \cap W_k) = n - \sum_{i=1}^{k} c_i + \sum_{j=1}^{k-1} n - \dim((W_1 \cap \cdots \cap W_j) + W_{j+1}$
- $\dim(W_1 \cap \cdots \cap W_k) \geq n - \sum_{i=1}^{k} c_i$.
- $\dim(W_1 \cap \cdots \cap W_k) = n - \sum_{i=1}^{k} c_i$ se, e somente se, for para todos $i = 1, \dots, k$,

$$W_i + (\bigcap_{j \neq i} W_j) = V$$

Demonstração. A parte $\boxed{a}$ segue do Teorema 2.4.2 $\boxed{d}$ por indução. As partes $\boxed{b}$ e $\boxed{c}$ são consequências fáceis de $\boxed{a}$. Deixamos os detalhes técnicos para um exercício no final desta seção. ■

2.5 Soma Direta Externa

Se V e W forem espaços vetoriais sobre um corpo $\mathbb{K}$, o produto cartesiano

$$V \times W = \{(\boldsymbol{v},\ \boldsymbol{w}) : \boldsymbol{v} \in V,\ \boldsymbol{w} \in W\}$$

pode ser munido de uma estrutura de espaço vetorial definindo as operações

$$(\boldsymbol{v}_1,\ \boldsymbol{w}_1) + (\boldsymbol{v}_2,\ \boldsymbol{w}_2) \triangleq (\boldsymbol{v}_1 + \boldsymbol{v}_2,\ \boldsymbol{w}_1 + \boldsymbol{w}_2) \tag{2.12}$$

$$\lambda(\boldsymbol{v},\ \boldsymbol{w}) \triangleq (\lambda \boldsymbol{v},\ \lambda \boldsymbol{w})\ . \tag{2.13}$$

O produto cartesiano $V \times W$, munido das operações de espaço vetorial acima, é denominado de soma direta externa de V e W.

A construção anterior pode ser generalizada facilmente.

2.1 Definição Soma Direta Externa

Sejam $V_1, \dots, V_n$ espaços vetoriais sobre um corpo $\mathbb{K}$. A soma direta externa de 1, denotado por $V = V_1 \boxplus \cdots \boxplus V_n$ é o espaço vetorial $V = V_1 \times \cdots \times V_n$ cujos elementos são as ênuplas ordenadas, i.e.,

$$V_1 \boxplus \cdots \boxplus V_n = \{(\boldsymbol{x}_1,\ \dots,\ \boldsymbol{x}_n) \,|\, \boldsymbol{x}_i \in V_i,\ i = 1,\ \dots,\ n\}$$

com as operações componentes a componentes

$$(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) + (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \triangleq (\boldsymbol{y}_1 + \boldsymbol{x}_1, \ldots, \boldsymbol{y}_n + \boldsymbol{x}_n) \tag{2.14}$$

$$\lambda(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \triangleq (\lambda\boldsymbol{x}_1, \ldots, \lambda\boldsymbol{x}_n) \tag{2.15}$$

Exemplo 2. *O espaço vetorial $\mathbb{K}^n$ é a soma direta externa de n cópias de $\mathbb{K}$, isto é,*

$$\mathbb{K}^n = \mathbb{K} \boxplus \cdots \boxplus \mathbb{K}$$

◁

Essa construção pode ser generalizada novamente para qualquer coleção de espaços vetoriais, generalizando a ideia de que um ênupla ordenada $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ é apenas uma função $f : \{1, \ldots, n\} \to \cup V_i$ do conjunto de índices $\{1, \ldots, n\}$ para a união dos espaços com a propriedade que $f(i) \in V_i$.

2.3 Definição Produto Direto

Seja $\mathcal{F} = \{V_i \mid i \in I\}$ família de espaços vetoriais acima de $\mathbb{K}$. O **produto direto** de $\mathcal{F}$ é o espaço vetorial

$$\prod_{i\in I} V_i \triangleq \{f : K \to \bigcup_{i\in I} V_i \mid f(i) \in V_i\}$$

pensado como um subespaço do espaço vetorial de todas as funções de $I \to \cup V_i$.

Será mais útil restringir o conjunto de funções àquelas com suporte finito.

2.4 Definição Suporte

Seja $\mathcal{F} = \{V_i \mid i \in I\}$ uma família de espaços vetoriais sobre $\mathbb{K}$. O suporte de uma função $f : K \to \cup V_i$ é o conjunto

$$\mathrm{suporte}(f) \triangleq \{i \in I \mid f(i) \neq 0\}$$

Assim, uma função f tem suporte finito se $f(i) = 0$ para todos, exceto um número finito de $i \in I$.

2.5 Definição Soma Direta Externa

A **soma direta externa** da família $\mathcal{F}$ é o espaço vetorial

$$\boxed{+}_{i\in I} V_i \triangleq \{f : I \to \bigcup_{i\in I} V_i \mid f(i) \in V_i f \text{ possui suporte finito } \}$$
$$= \{f : I \to \bigcup_{i\in I} V_i \mid f(i) \in V_i, f(i) \text{ é o exceto por um número finito de termos.}\}$$

visto como um subespaço do espaço vetorial de todas as funções de $I \to \cup V_i$.

Um caso especial importante ocorre quando $V_i = V$ para todos os $i \in I$. Se permitirmos que V^I denote o conjunto de todas as funções de I a V e $(V^I)_0$ denote o conjunto de todas as funções em V^I que possuem suporte finito então

$$\prod_{i\in I} V = V^I \qquad \text{e} \qquad \boxed{+}_{i\in I} V = (V^I)_0$$

Observe que o produto direto e a soma direta externa são os mesmos para uma família de espaços vetoriais finita.

Exemplos 6. *O espaço vetorial $\mathbb{K}^\infty$ é produto direto de $\mathbb{K}$ e $(\mathbb{K}^\infty)_0$ é soma direta externa:*

$$\mathbb{K}^\infty = \prod_{i\in\mathbb{N}} \mathbb{K} \quad e \quad (\mathbb{K}^\infty)_0 = \boxed{+}_{i\in\mathbb{N}} \mathbb{K}$$

◁

Exercícios

Ex. 2.45 --- Mostre que se $U \subset S$ então

$$S \cap (T + U) = (S \cap T) + U$$

Essa propriedade é denominada Lei Modular para o reticulado $\mathcal{S}(V)$.

Ex. 2.46 --- Para quais espaços vetoriais a lei distributiva de subespaços

$$S \cap (T + U) = (S \cap T) + (S \cap U)$$

é verdadeira? Essa propriedade é denominada Lei Modular para o reticulado $\mathcal{S}(V)$.

Ex. 2.47 --- Dado S um subespaço de V e $v \in V$. O conjunto $v+S = \{v+s : s \in S\}$ é chamado subespaço afim de V.

1. Quando um subespaço afim de V é subespaço de V?

2. Mostre que dois subespaços afim $x + S$ e $y + S$ ou são iguais ou são disjuntos.

Ex. 2.48 --- Seja $X \subseteq V$ um subconjunto de um espaço vetorial sobre $\mathbb{K}$ V. Mostre que $\langle X \rangle = \{\sum_{v \in F} \alpha_v \cdot v : F \subseteq X$ é um subconjunto finito e $\alpha_v \in \mathbb{K}$, $\forall v \in F\}$.

Ex. 2.49 --- Prove que

1. Se $S_1 \subseteq S_2$, $\langle S_1 \rangle \subseteq \langle S_2 \rangle$.
2. Se $\boldsymbol{x} \in \langle S \rangle$, existe um subconjunto finito $S' \subseteq S$ de modo que $\boldsymbol{x} \in \langle S' \rangle$.
3. $S \subseteq \langle S \rangle$ para todos os $S \in \mathscr{P}(V)$.
4. Para cada $S \in \mathscr{P}(V), \langle \langle S \rangle \rangle = \langle S \rangle$.
5. Se $\boldsymbol{y} \in \langle S \cup \{\boldsymbol{x}\} \rangle$ e $\boldsymbol{y} \notin \langle S \rangle$, então $\boldsymbol{x} \in \langle S \cup \{\boldsymbol{y}\} \rangle$. Aqui $\boldsymbol{x}, \boldsymbol{y} \in V$ e $S \in \mathscr{P}(V)$.

Ex. 2.50 --- Dê um contra-exemplo que a união de subespaços não é necessariamente um subespaço.

Ex. 2.51 --- Mostre que se W_1 e W_2 são subespaços de V, então $W_1 \cup W_2$ é um subespaço vetorial se, e somente se, $W_1 \subseteq W_2$ ou $W_2 \subseteq W_1$

Exercícios

Ex. 2.52 --- Seja V um espaço vetorial e seja $\{W_1, \ldots, W_k\}$ um conjunto de subespaços de V. Seja, $\underline{\mathrm{B}}_i$ seja uma base de W_i, para todo i, e deixe, $\underline{\mathrm{B}} = \cup \underline{\mathrm{B}}_i$. Mostre que

1. $\underline{\mathrm{B}}$ gera V se, e somente se, $V = W_1 + \cdots + W_k$.
2. $\underline{\mathrm{B}}$ é linearmente independente se, e somente se, $\{W_1, \ldots, W_k\}$ for independente.
3. $\underline{\mathrm{B}}$ é uma base para V se, e somente se, $V = W_1 \oplus \cdots \oplus W_k$.

Ex. 2.53 --- Seja V um espaço vetorial da dimensão n e seja $\{W_1, \ldots, W_k\}$ um conjunto de subespaços. Sejam $n_i = \dim (W_i)$.

1. Se $n_1 + \cdots + n_k > n$, então $\{W_1, \ldots, W_k\}$ não é independente.
2. Se $n_1 + \cdots + n_k < n$, então $V \neq W_1 + \cdots + W_k$.
3. Se $n_1 + \cdots + n_k = n$, o seguinte é equivalente:

   a) $V = W_1 \oplus \cdots \oplus W_k$.

b) $V = W_1 + \cdots + W_k$

c) $\{W_1, \ldots, W_k\}$ é independente.

Ex. 2.54 --- Sejam V um espaço vetorial de dimensão finita $\{W_1,. \ . \ . \ , W_k\}$ um conjunto de subespaços com $V = W_1 \oplus \cdots \oplus W_k$. Então $\dim(V) = \dim(W_1) + \cdots + \dim(W_k)$.

1. Ache o subespaço complementar a $L_3 = \{p(x) \in P_n : p(1) = 0\}$ em P_n o espaço vetorial dos polinômios de grau menor ou igual que n

Ex. 2.55 --- Seja V um espaço vetorial de dimensão finita da dimensão n. Suponha que $W_1, \ldots, W_k$ sejam subespaços de V. Para cada $i = 1, \ldots, k$, defina $c_i = n - \dim W_i$. Então

1. $\dim(W_1 \cap \cdots \cap W_k) = n - \sum_{i=1}^{k} c_i + \sum_{j=1}^{k-1} n - \dim((W_1 \cap \cdots \cap W_j) + W_{j+1}$

2. $\dim(W_1 \cap \cdots \cap W_k) \geq n - \sum_{i=1}^{k} c_i$.

3. $\dim(W_1 \cap \cdots \cap W_k) = n - \sum_{i=1}^{k} c_i$ se, e somente se, for para todos $i = 1, \ldots, k$,
$$W_i + (\bigcap_{j \neq i} W_j) = V$$

2.6 Coordenadas

Suponha que V seja um espaço vetorial de dimensão finita n sobre $\mathbb{K}$.

2.1 Definição Base Ordenada

Uma **base ordenada** para V é uma n-tupla ordenada $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ de vetores para os quais o conjunto $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ é uma base para V.

Se $\underline{\mathrm{x}} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ for uma base ordenada para V, então para todo $\boldsymbol{y} \in V$ haverá uma única n-tupla $(y_1, \ldots, y_n)$ de escalares para os quais

$$\boldsymbol{y} = y_1\boldsymbol{x}_1 + \cdots + y_n\boldsymbol{x}_b$$

Assim, se $\underline{\mathrm{x}} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ é uma base ordenada de V, temos uma função natural

$[\,\cdot\,]_{\underline{\mathrm{x}}} : V \to \mathbb{K}_n$, denominado aplicação de coordenadas, definida da seguinte forma.

2.2 Definição Coordenadas de um vetor

Se $\underline{\mathrm{x}} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ é uma base de V, então definimos as **coordenadas** do vetor $\boldsymbol{y} = \sum_{i=1}^{n} y_i \boldsymbol{x}_i$ na base $\underline{\mathrm{x}}$ como

$$[\boldsymbol{y}]_{\underline{\mathrm{x}}} \triangleq (y_1, \ldots, y_n)^t \in \mathbb{K}_n$$

Como $\underline{\mathrm{x}}$ é uma base de V, a representação de um determinado vetor $\boldsymbol{y}$ como uma combinação linear de $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ é única e assim temos que a função está bem definida.

A função $[\,\cdot\,]_{\underline{\mathrm{x}}} : V \to \mathbb{K}_n$ é uma bijetiva e preserva a adição de vetores e a multiplicação escalar. Consequentemente, $[\lambda_1 \boldsymbol{y} + \lambda_2 \boldsymbol{z}]_{\underline{\mathrm{x}}} = \lambda_1 [\boldsymbol{y}]_{\underline{\mathrm{x}}} + \lambda_2 [\boldsymbol{z}]_{\underline{\mathrm{x}}}$ para todos os $\lambda_1, \lambda_2 \in \mathbb{K}$ e $\boldsymbol{y}, \boldsymbol{z} \in V$. Como veremos em breve $[\,\cdot\,]_{\underline{\mathrm{x}}}$ é um exemplo de transformação linear.

2.3 Definição Componentes

O vetor coluna $[\boldsymbol{y}]_{\underline{\mathrm{x}}}$ costuma ser denominado **componentes** de $\boldsymbol{y}$ na base $\underline{\mathrm{x}}$.

Exemplo 4. *Seja $V = \mathcal{M}_{2,2}(\mathbb{K})$ e considere os vetores*

$$\boldsymbol{a}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{a}_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \boldsymbol{a}_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{a}_4 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

É fácil ver que $\underline{\mathrm{A}} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4)$ é uma base ordenada para $\mathcal{M}_{2,2}(\mathbb{K})$. Em particular qualquer matriz $A = (a_{ij})$ pode ser escrita como

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \frac{a_{12} + a_{21}}{2} \boldsymbol{a}_1 + \frac{a_{12} - a_{21}}{2} \boldsymbol{a}_2 + \frac{a_{11} + a_{22}}{2} \boldsymbol{a}_3 + \frac{a_{11} - a_{22}}{2} \boldsymbol{a}_4$$

E logo

$$\left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right]_{\underline{\mathrm{A}}} = \begin{bmatrix} \frac{a_{12}+a_{21}}{2} \\ \frac{a_{12}-a_{21}}{2} \\ \frac{a_{11}+a_{22}}{2} \\ \frac{a_{11}-a_{22}}{2} \end{bmatrix}$$

◁

Suponha que $\underline{\mathrm{x}} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ e $\underline{\mathrm{y}} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ são duas bases de V. Há um relação simples entre as componentes de um determinado vetor $\boldsymbol{z}$ nas $\underline{\mathrm{x}}$ e $\underline{\mathrm{y}}$. Para entender essa relação definimos

2.5 Definição Matriz Mudança de Base

A **matriz mudança de base** de $\underline{Y}$ para $\underline{X}$, denotada por $\mathrm{M}_{\underline{Y}\to\underline{X}}$, é a matriz $n \times n$ cujas colunas são definidas pela seguinte equação:

$$\mathrm{M}_{\underline{Y}\to\underline{X}} \triangleq \left([\boldsymbol{y}_1]_{\underline{X}} \;\middle|\; \cdots \;\middle|\; [\boldsymbol{y}_n]_{\underline{X}} \right)$$

Na equação 2.6, a i-ésima coluna de $\mathrm{M}_{\underline{Y}\to\underline{X}}$ é a matriz $n \times 1$ $[\boldsymbol{y}_i]_{\underline{X}}$.

A multiplicação por $\mathrm{M}_{\underline{Y}\to\underline{X}}$ induz uma aplicação de $\mathbb{K}_n$ a $\mathbb{K}_n$ que conecta as componentes nas bases $\underline{X}$ e $\underline{Y}$.

2.6 Teorema Mudança de Base

$$\mathrm{M}_{\underline{Y}\to\underline{X}}[\boldsymbol{z}]_{\underline{Y}} = [\boldsymbol{z}]_{\underline{X}} \text{ para todos os } \boldsymbol{z} \in V.$$

Demonstração. Vamos denotar a i-ésima coluna de qualquer matriz M por $\mathrm{Col}_i(M)$. Então, para todo $i = 1, \ldots, n$, temos

$$\mathrm{M}_{\underline{Y}\to\underline{X}}[\boldsymbol{y}_i]_{\underline{Y}} = \mathrm{M}_{\underline{Y}\to\underline{X}}(0, \ldots,\ 0, 1, 0, \ldots,\ 0)^t = \mathrm{Col}_i\left(\mathrm{M}_{\underline{Y}\to\underline{X}}\right) = [\boldsymbol{y}_i]_{\underline{X}}.$$

Portanto, o teorema é verdadeiro para $\boldsymbol{y} \in \underline{Y}$.

Agora, já observamos que $[\,\cdot\,]_{\underline{Y}}$ e $[\,\cdot\,]_{\underline{X}}$ preserva a adição de vetor e multiplicação por escalar. O mesmo acontece com a multiplicação por $\mathrm{M}_{\underline{Y}\to\underline{X}}$ como uma aplicação em $\mathbb{K}_n$. Como qualquer $\boldsymbol{y} \in V$ é uma combinação linear dos vetores em $\underline{Y}$, concluímos que $\mathrm{M}_{\underline{Y}\to\underline{X}}[\boldsymbol{z}]_{\underline{Y}} = [\boldsymbol{z}]_{\underline{X}}$ para todo $\boldsymbol{y} \in V$. ∎

Podemos representar as afirmações do Teorema 2.6 em termos do seguinte diagrama comutativo:

$$\begin{tikzcd}
 & V \arrow[dl, "{[\,\cdot\,]_{\underline{Y}}}"'] \arrow[dr, "{[\,\cdot\,]_{\underline{X}}}"] & \\
\mathbb{K}_n \arrow[rr, "\mathrm{M}_{\underline{Y}\to\underline{X}}"'] & & \mathbb{K}_n
\end{tikzcd} \tag{2.16}$$

Por um diagrama, queremos dizer uma coleção de espaços vetoriais e aplicações entre esses espaços. As aplicações são representadas como setas. Dizemos que um diagrama é comutativo se duas sequências de aplicações (ou seja, composições de funções no diagrama) que se originam no mesmo espaço e terminam no mesmo espaço forem iguais. Portanto, o diagrama 2.16 é comutativo se, e somente se, os dois caminhos de V para $\mathbb{K}_n$, são iguais. É exatamente isso que o Teorema 2.6 diz.

Exemplo 7. *Seja* $V = \mathbb{K}_3[x]$ *o espaço de todos os polinômios de grau menor ou igual a* 3. *Sejam* $\underline{B} = (p_1, p_2, p_3, p_4) = (1, x, x^2, x^3)$ *e* $\underline{C} = (q_1, q_2, q_3, q_4) = (1, (x-1), (x-1)^2, (x-1)^3)$ *duas bases ordenadas de* D. *Então*

$$\begin{aligned}
q_1 &= p_1\\
q_2 &= -1p_1 + p_2\\
q_3 &= p_1 - 2p_2 + p_3\\
q_4 &= -1p_1 + 3p_2 - 3p_3 + p_4.
\end{aligned}$$

Então a matriz mudança de base é dada por

$$\mathrm{M}_{\underline{C}\to\underline{B}} = \begin{bmatrix} 1 & -1 & 1 & -1\\ 0 & 1 & -2 & 3\\ 0 & 0 & 1 & -3\\ 0 & 0 & 0 & 1 \end{bmatrix}$$

e sua inversa é

$$\mathrm{M}_{\underline{B}\to\underline{C}} = \begin{bmatrix} 1 & 1 & 1 & 1\\ 0 & 1 & 2 & 3\\ 0 & 0 & 1 & 3\\ 0 & 0 & 0 & 1 \end{bmatrix}$$

◁

2.6.1 Espaços Linha e Coluna

Seja $A \in \mathcal{M}_{n,m}(\mathbb{K})$. As linhas de A geram um subespaço de $\mathbb{K}^n$ conhecido como espaço de linhas de A e denotado $\mathrm{EspLin}(A)$ enquanto as colunas de A geram um subespaço de $\mathbb{K}^m$ conhecido como espaço de colunas de A e denotado por $\mathrm{EspCol}(A)$.

As dimensões desses espaços são denominadas posto por linha e posto por coluna, respectivamente. Denotamos o posto por linha por $\mathrm{posto}_l(A)$ e o posto por coluna por $\mathrm{posto}_c(A)$.

Demonstraremos que posto por linha de uma matriz é igual ao posto por coluna, apesar de se $m \neq n$, o espaço da linha e o espaço da coluna não estarem no mesmo espaço vetorial.

Nossa demonstração desse fato depende da seguinte proposição sobre matrizes.

2.8 Proposição

Seja $A \in \mathcal{M}_{m,n}(\mathbb{K})$. As operações elementares por coluna não afetam posto por linha de A. Da mesma forma, as operações elementares por linha não afetam posto por coluna de A.

Demonstração. Vamos demonstrar a primeira afirmação.

Observe que uma relação de dependência linear entre algumas linhas de uma matriz A é equivalente à existência de um vetor não nulo $\boldsymbol{v}$ com $\boldsymbol{v}A = 0$. Se E é uma matriz elementar e $B = AE$, então $\boldsymbol{v}B = \boldsymbol{v}AE = \mathbf{0}$; inversamente, se $\boldsymbol{v}B = 0$, então $\boldsymbol{v}A = \boldsymbol{v}BE^{-1} = \mathbf{0}$. Portanto, se um conjunto de linhas de A for linearmente dependente, o conjunto correspondente de linhas de B também será linearmente dependente e vice-versa.

Sejam $\boldsymbol{e}_i$ são os vetores de base padrão em $\mathbb{K}^m$. Como a i-ésima linha de A é dada por $\boldsymbol{e}_i A$ temos que espaço de linha de A é

$$\mathrm{EspLin}(A) = \langle \boldsymbol{e}_1 A,\ \ldots,\ \boldsymbol{e}_n A\rangle.$$

Executar uma operação de coluna elementar em A é equivalente a multiplicar A à direita por uma matriz elementar E. Portanto, o espaço da linha de AE é

$$\mathrm{EspLin}(AE) = \langle \boldsymbol{e}_1 AE,\ \ldots,\ \boldsymbol{e}_n AE\rangle$$

e como E é invertível,

$$\mathrm{posto}_l(A) = \dim\left(\mathrm{EspLin}(A)\right) = \dim\left(\mathrm{EspLin}(AE)\right) = \mathrm{posto}_l(AE)$$

como desejado.

A demonstração da segunda afirmação é análoga a demonstração da primeira afirmação.

■

2.9 Teorema

Se $A \in \mathcal{M}_{m,n}$, $\mathrm{posto}_l(A) = \mathrm{posto}_c(A)$. Esse número é denominado **posto** de A e é denotado por $\mathrm{posto}(A)$.

Demonstração. Pelo lema anterior, podemos reduzir A para a forma escalonada reduzida por coluna sem alterar o posto por linha e o por coluna. Em seguida, podemos reduzir ainda mais A para forma escalonada reduzida por linha sem afetar nenhum dos postos. Logo a matriz resultante M tem os mesmos postos por linha e coluna que A. Mas M é uma matriz com 1's seguidos por 0's na diagonal principal e 0's em outros lugares. Consequentemente,

$$\mathrm{posto}_l(A) = \mathrm{posto}_l(M) = \mathrm{posto}_c(M) = \mathrm{posto}_c(A)$$

como desejado. ■

Exercícios

Ex. 2.56 --- Seja $P_n(x)$ o conjunto de todos os polinômios com coeficientes num corpo F de grau menor ou igual a n. Mostre que:

1. $1, x, \ldots x^n$ é uma base para L. As coordenadas do polinômio nessa base são os seus coeficientes.

2. $1, x-a, (x-a)^2, \ldots .(x-a)^n$ uma base de L. Se $\mathrm{char}(K) = p > n$ então as coordenadas do polinômio nessa base são $\{f(a), f'(a), \frac{f''(a)}{2}, \ldots \frac{f^n(a)}{n!}\}$

Ex. 2.57 ---

1. Verifique que $\underline{B} = (1+x, 1+x^2, 1+2x-2x^2)$ é uma base para $\mathbb{K}_2(x)$.

2. Calcule as coordenadas dos vetores $[1]_{\underline{B}}, [x]_{\underline{B}}, [x^2]_{\underline{B}}$.

Ex. 2.58 --- Sejam $\boldsymbol{f}_1(x) = -\dfrac{1}{6}(x-1)(x-2)(x-3)$, $\boldsymbol{f}_2(x) = \dfrac{1}{2}x(x-2)(x-3)$, $\boldsymbol{f}_3(x) = -\dfrac{1}{2}x(x-1)(x-3)$, $\boldsymbol{f}_4(x) = \dfrac{1}{6}x(x-1)(x-2)$.

1. Prove que $\underline{B} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3, \boldsymbol{f}_4)$ é uma base para $\mathbb{R}_{(3)}[x]$.

2. Se $g(X) \in \mathbb{R}_{(3)}[x]$, prove que $[g]_{\underline{B}} = \begin{bmatrix} g(0) \\ g(1) \\ g(2) \\ g(3) \end{bmatrix}$

Ex. 2.59 --- Seja $\underline{B} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3, \boldsymbol{f}_4)$ seja a base de $\mathbb{R}_{(3)}[x]$ do Exercício anterior. Calcule os vetores de coordenadas da base canônica , $(1, x, x^2, x^3)$ em relação a $\underline{B}$.

Ex. 2.60 --- Seja $\underline{B}$ uma base para o espaço vetorial finito dimensional V sobre o corpo $\mathbb{K}$ e sejam $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ seja uma sequência de vetores em V. Prove que $\langle(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)\rangle = V$ se, e somente se, $(\langle[\boldsymbol{u}_1]_{\underline{B}}, \ldots, \boldsymbol{u}_k]_{\underline{B}}\rangle = \mathbb{K}^n$.

Ex. 2.61 --- Seja B uma base para o espaço vetorial dimensional n V sobre o corpo $\mathbb{K}$ e sejam $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ seja uma sequência de vetores em V. Prove que $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ é uma base para V se, e somente se, $([\boldsymbol{u}_1]_B, \ldots$, for um base para $\mathbb{K}^n$.

2.7 Bandeiras

2.1 Definição Bandeira

Uma sequência de subespaços encaixantes ascendente $V_0 \subset V_1 \subset \ldots \subset V_n$ do espaço V é denominada de **bandeira** ou **cadeia ascendente**.

O número n é denominado **comprimento da bandeira** $V_0 \subset V_1 \subset \ldots \subset V_n$.

2.2 Definição Bandeira Maximal

Uma bandeira $V_0 \subset V_1 \subset \ldots \subset V_n \subset \ldots$ é dita **maximal** se $V_0 = \{0\}, \cup V_i = V$ e se nenhum subespaço puder ser inserido entre V_i, V_{i+1} (para qualquer i) isto é se $V_i \subset W \subset V_{i+1}$, então $W = V_i$ ou $W = V_{i+1}$.

Exemplo 3. *Em $P_n(x)$ temos a bandeira maximal:*

$$\{\mathbf{0}\} \subset P_0(x) \subset P_1(x) \subset \cdots \subset P_n(x)$$

◁

Um bandeira de comprimento n pode ser construído a partir de uma base ordenada $\underline{E} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ do espaço V definindo $V_0 = \{\mathbf{0}\}$ e $V_i = \langle\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i\}\rangle$ para $i = 1, \ldots, n$). Essa bandeira é denominada **bandeira canônica associada a base** $\underline{E}$.

Essa bandeira é máximal e num espaço vetorial de dimensão finita todas as bandeiras maximais podem ser construídas assim.

2.4 Teorema

A dimensão do espaço vetorial V é igual ao comprimento de qualquer bandeira máximal de V.

Demonstração. Vamos provar apenas o caso em que $\dim V$ é finita.

Seja $V_0 \subset V_1 \subset V_2 \subset \ldots$ uma bandeira máximal em V. Para $i \in I$, selecionamos um vetor $\boldsymbol{e}_i \in V_i \backslash V_{i-1}$. Mostraremos que o conjunto $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i\}$ é uma base do espaço V_i.

Primeiro, o subespaço gerado por $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{i-1}\}$ está contido em V_{i-1} e $\boldsymbol{e}_i$ não está em V_{i-1}, e segue por indução em i que o conjunto $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i\}$ é linearmente independente, para todos os i.

Agora mostraremos por indução que $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i\}$ gera V_i. Suponha por hipótese indutiva que isso seja verdade para $i-1$ e seja $V' = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_i \rangle$. Então $V_{i-1} \subset V'$ de acordo com a hipótese de indução e $V_{i-1} \neq V'$ pois $\boldsymbol{e}_i \notin V_{i-1}$. A condição de maximalidade da bandeira implica então que $V' = V_i$.

Assim, se $V_0 \subset V_1 \subset \ldots \subset V_n = V$ for uma bandeira maximal finita em V, então, os vetores $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}, \boldsymbol{e}_i \in V_i \backslash V_{i-1}$, formam uma base de V e logo $n = \dim V$.

■

Uma bandeira em um espaço de uma dimensão finita, V, pode ser estendida a bandeira máximal e, portanto, seu comprimento é sempre menor ou igual a $\dim V$. De fato, se continuamos a inserir subespaços intermediários no bandeira inicial este processo não pode continuar indefinidamente, porque o conjunto de vetores $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i\rangle$, $\boldsymbol{e}_i \in V_i \backslash V_{i-1}$ é um conjunto linearmente independente. Portanto, o comprimento do bandeira não pode exceder $\dim V$.

Dizemos que um espaço vetorial satisfaz a condição da cadeia ascendente (acc) se para qualquer sequência ascendente de subespaços

$$W_1 \subset W_2 \subset W_3 \subset \cdots,$$

eventualmente se estabiliza, isto é, existe um número inteiro positivo n tal que

$$W_n = W_{n+1} = W_{n+2} = \cdots.$$

Da mesma forma, diz-se que P satisfaz a condição da cadeia descendente (dcc) se todas as sequências descendentes de subespaços

$$W_1 \supset W_2 \supset W_3 \supset \cdots$$

eventualmente se estabiliza.

2.5 Teorema

Seja V um espaço vetorial sobre $\mathbb{K}$. Então, são equivalentes:

a) V é de dimensão finita;

b) V tem uma bandeira maximal;

c) V satisfaz a condição da cadeia ascendente (acc);

d) V satisfaz a condição da cadeia descendente (dcc).

Exercícios

Ex. 2.62 --- Mostre que se V é um espaço vetorial sobre $\mathbb{K}$. Então, são equivalentes:

1. V é de dimensão finita;

2. V tem uma bandeira maximal;

3. V satisfaz a condição da cadeia ascendente (acc);

4. V satisfaz a condição da cadeia descendente (dcc).

Ex. 2.63 --- Seja $0 = V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_n = W_1$ uma bandeira maximal para W_1 e $0 = L_0 \subsetneq L_1 \subsetneq \ldots \subsetneq L_m = W_2$ uma bandeira maximal para W_2. Mostre que

$$0 = V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_n \subsetneq V_n \oplus L_1 \subsetneq V_n \oplus L_2 \subsetneq \ldots \subsetneq V_n \oplus L_m = W_1 \oplus W_2 = V$$

é bandeira maximal para $V = W_1 \oplus W_2$. Conclua (mais uma vez) que dimensão da soma direta de espaços vetoriais de dimensão finita tem dimensão finita igual a soma das dimensões.

Capítulo 3

Transformações Lineares

Nesse capítulo

- Definição e Exemplos (p. 88)
- Isomorfismos (p. 100)
- Teorema do Núcleo-Imagem (p. 105)
- Representação Matricial dos Homomorfismos (p. 112)
- Somas Diretas e Projeções (p. 118)
- Mudança de Corpo (p. 122)

Neste capítulo vamos nos dedicar a um tipo especial de funções entre espaços vetoriais: as lineares.

As transformações lineares são as funções que preservam a estrutura vetorial - ou seja, as operações e e os axiomas de adição e multiplicação por um escalar. E o estudo das propriedades dessas transformações lineares é um dos objetos centrais da álgebra linear e nesse texto culminará com as Formas Normal de Jordan e na Forma Racional.

As transformações lineares são vitais em praticamente todas as áreas da ciência e outras áreas da matemática. Praticamente todas as áreas da ciência moderna contêm modelos onde as equações são aproximadas por equações e transformações lineares (aproximando a função pela sua derivada/"argumentos de expansão de Taylor").

3.1 Definição e Exemplos

3.1 Definição Transformação Linear

Seja V e W espaços vetoriais sobre um corpo $\mathbb{K}$. Uma função $T : V \to W$ é denominada de **transformação linear** ou **homomorfismo** se

$$T(\lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y}) = \lambda_1 T(\boldsymbol{x}) + \lambda_2 T(\boldsymbol{y}) \text{ para todos os } \lambda_1, \lambda_2 \in \mathbb{K} \text{ e } \boldsymbol{x}, \boldsymbol{y} \in V.$$

$$\begin{tikzcd}
\boldsymbol{x}, \boldsymbol{y} \arrow[r, "T"] \arrow[d, "+"'] & T(\boldsymbol{x}),\ T(\boldsymbol{y}) \arrow[d, "+"] \\
\boldsymbol{x} + \boldsymbol{y} \arrow[r, "T"'] & T(\boldsymbol{x}) + T(\boldsymbol{y})
\end{tikzcd}
\qquad
\begin{tikzcd}
\boldsymbol{x} \arrow[r, "T"] \arrow[d, "\cdot\lambda"'] & T(\boldsymbol{x}) \arrow[d, "\cdot\lambda"] \\
\lambda \boldsymbol{x} \arrow[r, "T"'] & \lambda T(\boldsymbol{x})
\end{tikzcd}$$

Uma transformação linear de V para V é dita **operador linear** em V. E uma transformação linear de V para $\mathbb{K}$ é denominada de **funcional linear** em V.

Começaremos apresentando alguns exemplos.

Exemplos 2.

- *A aplicação* $0 : V \to W$ *que envia todos os vetores para* $\mathbf{0} \in W$ *é uma aplicação linear. Vamos denominar essa aplicação de* ***transformação nula****.*
- *A transformação linear* ***identidade*** $\mathrm{I}_V : V \to V$ *é a transformação linear definida como* $\mathrm{I}_V(\boldsymbol{v}) = \boldsymbol{v}$ *para todo* $\boldsymbol{v} \in V$*. Quando não for gerar confusão escreveremos simplesmente* I*.*

$\triangleleft$

Em geral, uma aplicação que preserva uma estrutura algébrica é dita homomorfismo, e essa é a razão pela qul as transformações lineares também serão ditas homomorfismos (lineares). Por essa mesma razão temos as seguintes nomenclaturas:

3.3 Definição

Usaremos as seguintes designações:

- endomorfismo para operador linear.
- monomorfismo para transformação linear injetiva.
- epimorfismo para transformação linear sobrejetiva
- isomorfismo para transformação linear bijetiva.
- automorfismo para operador linear bijetivo.

Exemplo 4. *Se* V *é um espaço de dimensão finita com base ordenada* $\underline{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$*, então* $[\,\cdot\,]_{\underline{x}}$*:* $V \to \mathbb{K}_n$ *é uma transformação linear bijetiva, ou seja, um isomorfismo.*

$\triangleleft$

Exemplo 5. *A transposição,* $A \to A^t$ *é uma aplicação linear de* $\mathcal{M}_{m,n}(\mathbb{K}) \to \mathcal{M}_{n,m}(\mathbb{K})$*.*

$\triangleleft$

Exemplo 6. *Suponha que* $V = \mathcal{M}_{m,n}(\mathbb{K})$ *e* $A \in \mathcal{M}_{m,m}(\mathbb{K})$*. Então a multiplicação por* A *(necessariamente à esquerda) induz uma transformação linear* $\mathrm{L}_A : V \to V$ *dada por* $\mathrm{L}_A(B) = AB$ *para* $B \in V$*.*

$\triangleleft$

Os exemplos 4 e 6 mostram que o diagrama comutativo em 2.16 consiste em transformações lineares.

Exemplo 7. *Seja* $V = C^\infty(\mathbb{R})$*, o espaço vetorial das funções a valores reais infinitamente diferenciáveis.*

1. *Para um número real a,seja $\mathrm{E}_a : V \to \mathbb{R}$ a avaliação em a, ou seja, $\mathrm{E}_a(f(x)) = f(a)$. Então E_a é uma transformação linear. Também temos a transformação linear $\overline{\mathrm{E}}_a$: $V \to V$, em que $\overline{\mathrm{E}}_a(f(x))$ é a função constante cujo valor é $f(a)$.*
2. *Seja D: $V \to V$ a diferenciação, isto é, $\mathrm{D}(f(x)) = f'(x)$. Então D é uma transformação linear.*
3. *Para um número real a, seja I_a: $V \to V$ a integração definida começando em $t = a$, ou seja,*
$$\mathrm{I}_a(f)(x) = \int_a^x f(t)dt.$$
Então I_a é uma transformação linear.
4. *Também temos a transformação linear $I_a^b = \mathrm{E}_b \circ \mathrm{I}_a$. Dessa forma*
$$I_a^b(f(x)) = (\mathrm{E}_b \circ \mathrm{I}_a)(f(x)) = \int_a^b f(x)dx.$$

◁

3.8 Teorema Fundamental do Cálculo

1. $\mathrm{D} \circ \mathrm{I}_a = \mathrm{I}$.
2. $\mathrm{I}_a \circ \mathrm{D} = \mathrm{I} - \overline{\mathrm{E}}_a$.

Exemplo 9. *Suponha $V = \mathbb{K}[x]$. Podemos definir uma derivada formal em V da seguinte maneira: Se $p(x) = \sum_{i=0}^n a_i x^i$, fazemos $D[p(x)] = p'(x) = \sum_{i=1}^n i a_i x^{i-1}$. O leitor pode verificar facilmente que essa transformação, que é denominada derivação formal em $\mathbb{K}[x]$, é uma transformação linear.*

◁

Exemplo 10. *Suponha $V = \mathcal{R}(B)$ como no Exemplo 14. Então $T(f) = \int_B f dA$. É uma transformação linear de V para $\mathbb{R}$.*

◁

Exemplo 11. *Se X é uma variável aleatória com função de densidade de probabilidade de $f(x)$, o valor esperado é definido como a integral:*
$$\mathbb{E}[X] = \int_{\mathbb{R}} x f(x)\, dx,$$
é um funcional linear.

◁

Exemplo 12. *Seja $V = \mathbb{K}^\infty$ o espaço vetorial definido no Exemplo 6. Definimos o shift para a esquerda L: $V \to V$ e o shift para a direita $\mathbb{R}$: $V \to V$ como*
$$\mathrm{L}([a_1,\ a_2,\ a_3,\ \ldots]) = [a_2,\ a_3,\ a_4,\ \ldots],$$

$$\mathrm{R}([a_1,\ a_2,\ a_3,\ \ldots]) = [0,\ a_1,\ a_2,\ \ldots].$$

Observamos que L *e* $\mathbb{R}$ *são transformações lineares.*

Podemos restringir L *e* $\mathbb{R}$ *para o subespaço* $W = (\mathbb{K}^\infty)_0$ *de* V *definido no Exemplo* 7. *E assim obter que* $\mathrm{L}\colon\ W \to W$ *e* $\mathbb{R}\colon\ W \to W$ *também são transformações lineares.*

Nesse mesmo exemplo vimos que $\underline{B} = \{\boldsymbol{e}_i, i \in \mathbb{N}\}$ *onde* $\boldsymbol{e}_i = \underbrace{(0, \ldots,\ 1, \ldots,)}_{\text{1 na i-ésima posição}}$ *é uma base de* $(\mathbb{K}^\infty)_0$.

As transformações L *e* $\mathbb{R}$ *agem do seguinte modo nos vetores da base:*

$$\mathrm{R}(\boldsymbol{e}_i) = \boldsymbol{e}_{i+1} \tag{3.1}$$

$$\mathrm{L}(\boldsymbol{e}_i) = \begin{cases} \boldsymbol{e}_{i-1} & \text{se } i > 1 \\ \mathbf{0} & \text{se } i = 1 \end{cases} \tag{3.2}$$

$$\boldsymbol{e}_1 \xrightarrow{\mathrm{R}} \boldsymbol{e}_2 \xrightarrow{\mathrm{R}} \boldsymbol{e}_3 \xrightarrow{\mathrm{R}} \ldots \xrightarrow{\mathrm{R}} \boldsymbol{e}_n \xrightarrow{\mathrm{R}} \boldsymbol{e}_{n+1} \xrightarrow{\mathrm{R}} \ldots$$

$$\mathbf{0} \xleftarrow[\mathrm{L}]{} \boldsymbol{e}_1 \xleftarrow[\mathrm{L}]{} \boldsymbol{e}_2 \xleftarrow[\mathrm{L}]{} \boldsymbol{e}_3 \xleftarrow[\mathrm{L}]{} \ldots \xleftarrow[\mathrm{L}]{} \boldsymbol{e}_n \xleftarrow[\mathrm{L}]{} \boldsymbol{e}_{n+1} \xleftarrow[\mathrm{L}]{} \ldots$$

◁

3.13 Proposição

Sejam $T, S : V \to W$ e $L : W \to Z$ transformações lineares e $\lambda_1, \lambda_2 \in \mathbb{K}$. Então

- **a** $T(\mathbf{0}) = \mathbf{0}$.
- **b** A a composta $L \circ T : V \to Z$ é uma transformação linear.
- **c** $\lambda_1 T + \lambda_2 S$ é uma transformação linear de V em W.
- **d** Se T for invertível, então $T^{-1} : W \to V$ é uma transformação linear.

Demonstração. **c** Se $a, b \in \mathbb{K}$ e $\boldsymbol{x}, \boldsymbol{y} \in V$, então

$$\begin{aligned} (\lambda_1 T + \lambda_2 S)(a\boldsymbol{x} + b\boldsymbol{y}) &= \lambda_1 T(a\boldsymbol{x} + b\boldsymbol{y}) + \lambda_2 S(a\boldsymbol{x} + b\boldsymbol{y}) && (3.3)\\ &= \lambda_1 aT(\boldsymbol{x})\lambda_1 bT(\boldsymbol{y}) + \lambda_2 aS(\boldsymbol{x}) + \lambda_2 bS(\boldsymbol{y}) && (3.4)\\ &= a(\lambda_1 T(\boldsymbol{x}) + \lambda_2 S(\boldsymbol{x})) + b(\lambda_1 T(\boldsymbol{y}) + \lambda_2 S(\boldsymbol{y})) && (3.5)\\ &= a(\lambda_1 T + \lambda_2 S)(\boldsymbol{x}) + b(\lambda_1 T + \lambda_2 S)\boldsymbol{y}). && (3.6) \end{aligned}$$

Portanto, $\lambda_1 T + \lambda_2 S$ é linear.

d Seja $T : V \to W$ uma transformação linear bijetiva. Então $T^{-1} : W \to V$ é uma função bem definida e, já que quaisquer dois vetores $\boldsymbol{y}_1$ e $\boldsymbol{y}_2$ em W têm o formato $\boldsymbol{y}_1 = T\boldsymbol{x}_1$ e $\boldsymbol{y}_2 = T\boldsymbol{x}_2$, temos

$$
\begin{aligned}
T^{-1}(\lambda_1\boldsymbol{y}_1 + \lambda_2\boldsymbol{y}_2) &= T^{-1}(\lambda_1 T\boldsymbol{x}_1 + \lambda_2 T\boldsymbol{x}_2) && (3.7)\\
&= T^{-1}(T(\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2)) && (3.8)\\
&= \lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 && (3.9)\\
&= \lambda_1 T^{-1}(\boldsymbol{y}_1) + \lambda_2 T^{-1}(\boldsymbol{y}_2) && (3.10)
\end{aligned}
$$

o que mostra que T^{-1} é linear. ■

Neste ponto, vamos introduzir um nome para a coleção de todas as transformações lineares de V para W.

3.14 Definição Espaço das Transformações Lineares

Seja V e W espaços vetoriais sobre $\mathbb{K}$. O espaço vetorial de todas as transformações lineares de V para W será denotado por $\mathrm{Hom}_{\mathbb{K}}(V, W)$ ou $\mathcal{L}_{\mathbb{K}}(V, W)$. E o conjunto de todos os operadores lineares em V será denotado por $\mathrm{Hom}_{\mathbb{K}}(V, V)$.

Quando o corpo base $\mathbb{K}$ estiver claro a partir do contexto, escreveremos simplesmente $\mathrm{Hom}(V, W)$ em vez de $\mathrm{Hom}_{\mathbb{K}}(V, W)$. Portanto, $\mathrm{Hom}(V, W)$ é o subespaço do espaço vetorial W^V (Exemplo 8) consistindo em todas as transformações lineares de V para W.

3.15 Teorema

$\mathrm{Hom}(V, W)$ é um subespaço vetorial de W^V.

Como qualquer $T \in \mathrm{Hom}(V, W)$ tem a propriedade que $T(\mathbf{0}) = \mathbf{0}$, temos que $\mathrm{Hom}(V, W)$ é um subespaço próprio de W^V sempre que $W \neq \{\mathbf{0}\}$.

Existem dois subespaços vetoriais muito importantes associados a uma transformação linear T de V para W.

3.16 Definição Núcleo e Imagem

Seja $T \in \mathrm{Hom}(V,\ W)$.

1 O subespaço

$$\ker T = \{v \in V | T\boldsymbol{v} = 0\}$$

é denominado **núcleo** de T

2 O subespaço

$$\operatorname{im} T = \{T\boldsymbol{v} | \boldsymbol{v} \in V\}$$

é denominado **imagem** de T.

3 A dimensão do $\ker T$ é a **nulidade** de T e é denotada por $\operatorname{nul} T$.

4 A dimensão de $\operatorname{im} T$ é o **posto** de T e é indicada por $\operatorname{posto} T$.

É um exercício de rotina mostrar que $\ker T$ é um subespaço de V e que $\operatorname{im} T$ é um subespaço de W. Além disso, temos o seguinte.

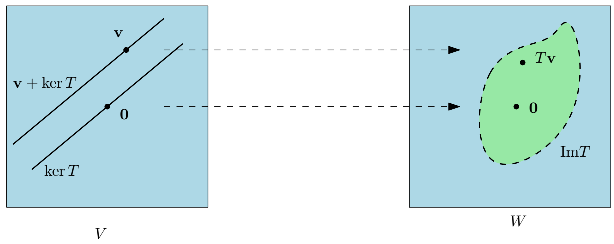


Figura 3.1 Uma aplicação linear $T : V \to W$ envia todos os pontos de uma classe $\boldsymbol{v} + \ker T$ do núcleo para um único ponto $T\boldsymbol{v}$ em W. Diferentes classes são levadas para diferentes pontos, e todas as imagens formam $\operatorname{im} T$. a classe do zero $\mathbf{0} + \ker T$ colapsa na origem $0 \in W'$

3.17 Teorema

Seja $T \in \operatorname{Hom}(V, W)$. Então

a T é injetiva se, e somente se, $\ker T = \{\mathbf{0}\}$

b T é sobrejetiva se, e somente se, $\operatorname{im} T = W$

Demonstração. Para ver a validade da primeira afirmação, observe que $T\boldsymbol{u} = T\boldsymbol{v} \Leftrightarrow T(\boldsymbol{u}-\boldsymbol{v}) = 0 \Leftrightarrow \boldsymbol{u}-\boldsymbol{v} \in \ker T$ Portanto, se $\ker T = \{\mathbf{0}\}$, $T\boldsymbol{u} = T\boldsymbol{v} \Leftrightarrow \boldsymbol{u} = \boldsymbol{v}$, o que mostra que T é injetiva. Por outro lado, se T for injetiva e $\boldsymbol{u} \in \ker T$, $T\boldsymbol{u} = T\mathbf{0}$ e, portanto, $\boldsymbol{u} = \mathbf{0}$. Isso mostra que $\ker T = \{\mathbf{0}\}$.

A segunda afirmação é direta da definição de sobrejetividade. ■

O seguinte teorema é extremamente útil

3.18 Teorema da Extensão por Linearidade

Sejam V e W espaços vetoriais sobre $\mathbb{K}$ e suponha $\underline{x} = (\boldsymbol{x}_i \,|\, i \in I)$ é uma base de V. Se $(\boldsymbol{y}_i \,|\, i \in I)$ é qualquer subconjunto de W, existe uma única

$T \in \mathrm{Hom}(V, W)$ tal que $T(\boldsymbol{x}_i) = \boldsymbol{y}_i$ para todos os $i \in I$.

Demonstração. Vamos estender T a V por linearidade, ou seja,

$$\begin{aligned} T(x_1\boldsymbol{x}_1 + \cdots + x_n\boldsymbol{x}_n) &= x_1T(\boldsymbol{x}_1) + \cdots + x_nT(\boldsymbol{x}_n) \\ &= x_1\boldsymbol{y}_1 + \cdots + x_n\boldsymbol{y}_n \end{aligned}$$

Esse processo fornece uma única transformação linear. Vejamos detalhadamente como:

Seja $\boldsymbol{x} \in V$. Então

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i\boldsymbol{x}_i,$$

onde $x_1, x_2, \ldots, x_n$ são escalares unicamente determinados. Definimos $T : V \to W$ por $T(\boldsymbol{x}) = \sum_{i=1}^{n} x_i\boldsymbol{y}_i$ motivado pelos argumentos anteriores.

☐ T é linear: Suponha que $\boldsymbol{u}, \boldsymbol{v} \in V$ e $\lambda \in \mathbb{K}$. Então escrevemos

$\boldsymbol{u} = \sum_{i=1}^{n} u_i\boldsymbol{x}_i$ e $\boldsymbol{v} = \sum_{i=1}^{n} v_i\boldsymbol{x}_i$

Logo

$$\lambda\boldsymbol{u} + \boldsymbol{v} = \sum_{i=1}^{n} (\lambda u_i + v_i)\boldsymbol{x}_i.$$

Então

$$T(\lambda\boldsymbol{u} + \boldsymbol{v}) = \sum_{i=1}^{n} (\lambda u_i + v_i)\boldsymbol{y}_i = \lambda \sum_{i=1}^{n} u_i\boldsymbol{y}_i + \sum_{i=1}^{n} v_i\boldsymbol{y}_i = \lambda T(\boldsymbol{u}) + T(\boldsymbol{v})\ .$$

☐ Claramente

$$T(\boldsymbol{x}_i) = \boldsymbol{y}_i \text{ para } i = 1, 2, \ldots, n.$$

☐ T é única: Suponha que $S : V \to W$ é linear e $S(\boldsymbol{x}_i) = \boldsymbol{y}_i$ para $i = 1, 2, \ldots, n$. Então para $\boldsymbol{x} \in V$ com

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i\boldsymbol{x}_i,$$

temos

$$S(\boldsymbol{x}) = \sum_{i=1}^{n} x_iS(\boldsymbol{x}_i) = \sum_{i=1}^{n} x_i\boldsymbol{y}_i = T(\boldsymbol{x}).$$

Logo $S = T$.

■

O teorema anterior admite a seguinte generalização para conjuntos linearmente independentes.

3.19 Teorema da Extensão para Conjuntos Linearmente Independentes

Sejam V e W espaços vetoriais sobre $\mathbb{K}$ e suponha $\{\boldsymbol{x}_i \,|\, i \in I\}$ é um conjunto linearmente independente de V. Se $\{\boldsymbol{y}_i \,|\, i \in I\}$ é qualquer subconjunto de

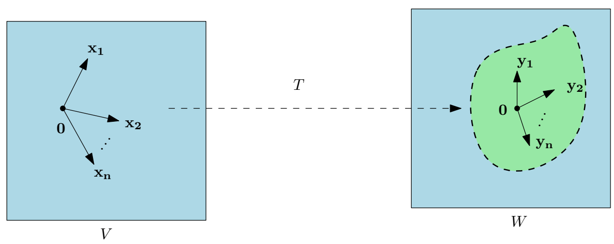


Figura 3.2 Uma aplicação linear $T : V \to W$ fica completamente determinada pelos valores que assume numa base. Essa é uma ferramenta útil para construirmos transformações lineares.

> W, existe $T \in \operatorname{Hom}(V, W)$ tal que $T(\boldsymbol{x}_i) = \boldsymbol{y}_i$ para todos os $i \in I$.

Essa transformação não é necessariamente única.

> **3.20 Teorema**
>
> Sejam V e W espaços vetoriais sobre $\mathbb{K}$, e suponha que V tenha dimensão finita $\dim V = n$. Então todas as bases ordenadas $\underline{\mathrm{x}}$ de V determinam um isomorfismo $\tau(\underline{\mathrm{x}}) : \operatorname{Hom}(V, W) \to W^n$.

Demonstração. Seja uma base ordenada $\underline{\mathrm{x}} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ de V. Defina $\tau(\underline{\mathrm{x}}) : \operatorname{Hom}(V, W) \to W^n$ por

$$\tau(\underline{\mathrm{x}})(T) = (T(\boldsymbol{x}_1), \ldots, T(\boldsymbol{x}_n)).$$

O fato de que $\tau(\underline{\mathrm{x}})$ é uma transformação linear é óbvio. Podemos definir a aplicação inversa ϕ: $W^n \to \operatorname{Hom}(V, W)$ por

$$\phi((\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)) = T,$$

Sendo T a única transformação linear que satisfaz $T(\boldsymbol{x}_i) = \boldsymbol{y}_i$. Portanto, $\tau(\underline{\mathrm{x}})$ é um isomorfismo. ■

3.1.1 Transformações de $\mathbb{K}^n$ a $\mathbb{K}^m$

Para qualquer matriz $A \in \mathcal{M}_{m,n}(\mathbb{K})$, a aplicação de multiplicação

$$T_A(\boldsymbol{v}) = A\boldsymbol{v}$$

é uma transformação linear de $\mathbb{K}^n$ a $\mathbb{K}^m$.

De fato, qualquer transformação linear $T \in \operatorname{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ tem essa forma, ou seja, T é a multiplicação por uma matriz. Para vermos isso fixemos as bases canônicas $\underline{E} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ e $\underline{F} = \{\boldsymbol{f}_1, \dots, \boldsymbol{f}_m\}$ de $\mathbb{K}^n$ e $\mathbb{K}^m$ respectivamente e assim definirmos a matriz

$$[T]_{\underline{E},\underline{F}} \triangleq (T\boldsymbol{e}_1|\cdots|T\boldsymbol{e}_n)$$

e assim temos

$$[T]_{\underline{E},\underline{F}}[T]_{\underline{E},\underline{F}}\boldsymbol{f}_i = (T\boldsymbol{e}_1|\cdots|T\boldsymbol{e}_n)\boldsymbol{f}_i = \mathrm{Col}_i(T\boldsymbol{e}_1|\cdots|T\boldsymbol{e}_n) = T\boldsymbol{f}_i$$

e consequentemente, por coincidirem numa base, temos que $T = T_A$, onde

$$A = [T]_{\underline{E},\underline{F}} = \left[T\boldsymbol{e}_1 \;\middle|\; \cdots \;\middle|\; T\boldsymbol{e}_n\right]$$

3.21 Teorema

a Se A for uma matriz $m \times n$ sobre $\mathbb{K}$, então $T_A \in \mathrm{Hom}(\mathbb{K}^n,\ \mathbb{K}^m)$.

b Se $T \in \mathrm{Hom}(\mathbb{K}^n,\ \mathbb{K}^m)$, então $T = T_A$, em que

$$A = \left[T\boldsymbol{e}_1 \;\middle|\; \cdots \;\middle|\; T\boldsymbol{e}_n\right]$$

A matriz A é denominada matriz de T.

Exemplos 22. *Vamos considerar algumas transformações de $\mathbb{R}^2$ em $\mathbb{R}^2$ e descrever seu comportamento geométrico.*

1 *A primeira é a homotetia $H_\lambda[v] = \lambda v$ que multiplica todos os vetores em $\mathbb{R}^2$ por λ. Essa transformação admite a mesma representação matricial em todas as bases:*

$$\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

2 *A segunda é a homotetia por fatores diferentes. Nesse caso seja $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ a base canônica de $\mathbb{R}^2$. Seja $H_{\lambda,\mu}$ a única transformação linear tal que $H_{\lambda,\mu}\boldsymbol{e}_1 = \lambda\boldsymbol{e}_1$e $H_{\lambda,\mu}\boldsymbol{e}_2 = \mu\boldsymbol{e}_2$. Essa transformação admite a seguinte representação matricial na base canônica*

$$\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$$

3 *Uma rotação em $\mathbb{R}^2$ pelo ângulo θ denotada R_θ é a transformação que na base canônica é representada pela matriz*

$$R_\theta = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

4 *Reflexão no reta xy. Seja a transformação que deixa o eixo x invariante e reflete o eixo y em torno de x. Seja $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ a base canônica de $\mathbb{R}^2$ então essa transformação é caracterizada por $T\boldsymbol{e}_1 = \boldsymbol{e}_1$ e $T\boldsymbol{e}_2 = -\boldsymbol{e}_2$ e admite a seguinte representação matricial na base canônica.*

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

$\boxed{5}$ *Uma transformação de cisalhamento é uma transformação do plano com a propriedade de que existe um vetor* $\boldsymbol{e}_1$, *tal que* $T(\boldsymbol{e}_1) = \boldsymbol{e}_1$ *e* $T(\boldsymbol{e}_2) - \boldsymbol{e}_2$ *é um múltiplo de* $\boldsymbol{e}_1$ *para todo*$\boldsymbol{e}_2$. *Ou seja, é a transformação linear que leva* $(\boldsymbol{e}_1, \boldsymbol{e}_2) \mapsto (\boldsymbol{e}_1, \boldsymbol{e}_2 + m\boldsymbol{e}_1)$. *Nessa base*

$$C_{\boldsymbol{e}_1,\boldsymbol{e}_2.m} = \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}$$

$\triangleleft$

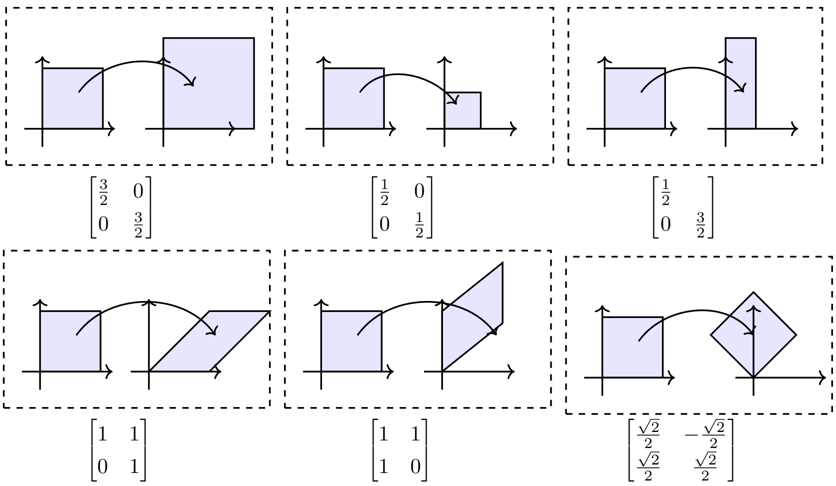


Figura 3.3 Transformações dos Exemplos 22

O resultado do Teorema 3.1.1 pode ser generalizado para $(\mathbb{K}^\infty)_0$. Nesse caso se matrizes infinitas são usadas para descrever transformações lineares, somente as matrizes cujas colunas têm apenas um número finito de entradas diferentes de zero podem ser usadas. Se uma matriz A descreve uma transformação linear $T : (\mathbb{K}^\infty)_0 \to (\mathbb{K}^\infty)_0$ então as colunas de A descrevem as imagens por T dos vetores da base, e isso só faz sentido se essas colunas tiverem um número finito de entradas diferentes de zero. Entretanto, não há restrição nas linhas de A: no produto $A\boldsymbol{v}$, existem apenas finitos coeficientes diferentes de $\boldsymbol{v}$ envolvidos; portanto, cada uma de suas entradas, mesmo que seja dada como uma quantidade infinita de entradas, envolve apenas um número finito de termos diferentes de zero e, portanto, está bem definido. Os produtos de duas matrizes desse tipo estão bem definidos e correspondem à composição de transformações lineares.

Exemplo 23. *O shift para a esquerda admite a representação matricial*

$$\mathrm{L} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 0 & \ddots & \\ \vdots & \vdots & & & & \ddots \end{bmatrix}$$

$\triangleleft$

Exercícios

Ex. 3.1 --- Seja $T : U \to V$ uma transformação linear. Mostre que

1. $T(0_U) = 0_V$, onde 0_U e 0_V denotam os vetores nulos de U e V, respectivamente.

2. $T(-u) = -T(u)$, para todo $u \in U$.

3. $T(\sum_{i=1}^{m} \alpha_i u_i) = \sum_{i=1}^{m} \alpha_i T(u_i)$, onde $\alpha_i \in \mathbb{K}$ e $u_i \in U$ para $i = 1, \ldots, m$.

Ex. 3.2 ---

1. Prove que a derivada formal do exemplo 9 é uma transformação linear.

   Considere $D : \mathbb{K}_n[x] \to \mathbb{K}_n[x]$ tal que $D[p(x)] \triangleq p'(x)$.

2. Prove que $D^n = 0_{\mathbb{K}_n[x]}$. Uma transformação T tal que $T^k = 0$ para algum k é dita nilpotente.

3. Calcule $\ker D^k$ para $k = 1, \ldots n$.

Ex. 3.3 --- Considere $\mathrm{L} : (\mathbb{K}^\infty)_0 \to (\mathbb{K}^\infty)_0$ definida como

$$\mathrm{L}([a_1,\ a_2,\ a_3,\ \ldots]) = [a_2,\ a_3,\ a_4,\ \ldots],$$

1. Mostre que L é uma transformação linear.

2. Calcule $\ker \mathrm{L}^k$ para $k \in \mathbb{N}$.

3. Mostre que para todo vetor $\boldsymbol{v} \in (\mathbb{K}^\infty)_0$ existe $k \in \mathbb{N}$ tal que $\mathrm{L}^k(v) = \mathbf{0}$, mas L não é nilpotente.

4. Calcule o núcleo de $\mathrm{L}^n = \underbrace{L \circ \ldots \circ L}_{n \text{ vezes}}$.

Ex. 3.4 --- Dado um espaço vetorial $V = L_1 \oplus L_2$. Prove que os seguintes operadores são lineares:

1. Dado $v \in V$ e seja a decomposição de v, $v = v_1 + v_2$, seja $\sigma : V \to V$ tal que $\sigma(v_1 + v_2) = v_1$. Esse operador é chamado projeção no espaço L_1.

2. Dado $v \in V$ e seja a decomposição de v, $v = v_1 + v_2$, seja $\tau : V \to V$ tal que $\tau(v_1 + v_2) = v_1 - v_2$. Esse operador é chamado reflexão no espaço L_1 paralelo a L_2.

Ex. 3.5 --- Sejam V e W espaços vetoriais sobre $\mathbb{K}$ e suponha $\{\boldsymbol{x}_i \,|\, i \in \Delta\}$ é um conjunto linearmente independente de V e que $\{\boldsymbol{y}_i \,|\, i \in \Delta\}$ é qualquer subconjunto de W.

1. Mostre que existe $T \in \mathrm{Hom}(V, W)$ tal que $T(\boldsymbol{x}_i) = \boldsymbol{y}_i$ para todos os $i \in \Delta$.
2. Forneça exemplos que mostrem que a T não é necessariamente única.

Ex. 3.6 --- Mostre que:

1. Se $T, G : U \to V$ são transformações lineais então $T + G : U \to V$ é uma transformação linear.
2. Se $T : U \to V$ é uma transformação linear então $\alpha T : U \to V$ é uma transformação linear para todo $\alpha \in \mathbb{K}$.
3. Se $T : U \to V$ e $G : V \to W$ são transformações lineais então $G \circ T : U \to W$ é uma transformação linear.

Ex. 3.7 --- Seja $V = (0, \infty)$, dados $x, y \in V$ e $\alpha \in \mathbb{R}$ defina as operações:

$$x + y := x \cdot y$$
$$\alpha \cdot x := x^\alpha$$

Verifique que V com essas operações é um $\mathbb{R}$-espaço vetorial e que $T : V \to \mathbb{R}$ dada por $T(x) = \ln(x)$ é uma transformação linear.

Ex. 3.8 --- Seja $T : U \to V$ uma transformação linear e $U' \subseteq U, V' \subseteq V$ subespaços vetoriais. Mostre que $T(U')$ é um subespaço de V e que $T^{-1}(V')$ (pré-imagem de V' por T) é um subespaço de U. Conclua que $\mathrm{Im}(T)$ é um subespaço de V e $\ker(T)$ é um subespaço de U.

Ex. 3.9 --- Seja V e W espaços vetoriais sobre $\mathbb{K}$. Suponha que $T : V \to W$ seja uma função. Mostre que $T \in \mathrm{Hom}(V, W)$ se e somente se o gráfico $G_T = \{(\boldsymbol{x}, T(\boldsymbol{x})) \in V \times W\}$de T for um subespaço de $V \times W$.

Ex. 3.10 --- Seja $\mathbb{K}_n[x]$ o espaço dos polinômios de grau menor que n e seja A_h o operador diferença

$$A_h(p(x)) := \frac{p(x+h) - p(x)}{h}$$

sendo h um número fixo não nulo. Ache o núcleo e a imagem desse operador.

Ex. 3.11 --- Sejam U um espaço vetorial sobre $\mathbb{K}$ e $T : U \to U$ uma transformação linear tal que $T \circ T = T$. Seja $W = \{x \in U : T(x) = x\}$ e $V = \{x \in U : T(x) = 0\}$. Prove que:

1. $U = W \oplus V$
2. $T(U) = W$
3. $T(V) = \{0\}$.

Ex. 3.12 --- Existe um mapa $f : \mathbb{R} \to \mathbb{R}$ satisfazendo $f(x+y) = f(x)+f(y)$, $f(x) = 0$ se x é racional $f(\sqrt{2}) = 1$.

Ex. 3.13 --- Calcule todos os funcionais lineares de $\mathbb{Z}_2^3$. Qual a dimensão do espaço dos funcionais lineares sobre $\mathbb{Z}_2^3$?

Ex. 3.14 --- Seja $T \in \mathrm{Hom}(V, V)$, e seja $L \subset V$ o subespaço de V tal que $L = \{v : f(T(v) = 0, \forall f \in V^*\}$. Prove que $L = \ker(T)$.

Ex. 3.15 --- Seja T a função de $\mathbb{R}^3$ em $\mathbb{R}^3$ definida por:

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2, -x_1 - 2x_2 + 2x_3)$$

1. Verifique que T é uma transformação linear
2. Determine a imagem de T
3. Determine o posto de T

Ex. 3.16 --- Dado $M_{n\times n}$(K) o espaço vetorial das matrizes $n \times n$ sobre K e seja B uma matriz fixa em $M_{n\times n}(K)$. Se $T(A) = \mathrm{AB} - \mathrm{BA}$, prove que T(A) é uma transformação linear de $M_{n\times n}$(K) em $M_{n\times n}$(K). Determine a imagem e o posto de T.

Ex. 3.17 --- Se $U \xrightarrow{T} V \xrightarrow{S} W$ são espaços vetoriais de dimensão finita e aplicações lineares, então

$$\dim\ker(S \circ T) \leq \dim\ker S + \dim\ker T.$$

3.2 Isomorfismos

3.1 Definição Isomorfismo

Uma transformação linear bijetiva $T : V \to W$ é dita **isomorfismo** de V a W. Quando existe um isomorfismo de V a W, dizemos que V e W são isomorfos e escrevemos $V \cong W$.

Dois objetos espaços vetoriais V e W são isomorfos se "forem essencialmente os mesmos", pelo menos do ponto de vista vetorial, o que significa que uma vez podem identificá-los um com o outro de uma maneira razoável.

Exemplo 2. 1 *Seja V um espaço vetorial então a aplicação identidade $\mathrm{I}_V : V \to V$ definida como $\mathrm{I}_V(c) = v$ é um isomorfismo. E assim*

$$V \cong V$$

2 *Se $T : V \to W$ é um isomorfismo então $T^{-1} : W \to V$ é um isomorfismo. Ou seja,*

$$V \cong W \Leftrightarrow W \cong V$$

3 *Se $T : V \to W$ e $S : W \to Z$ são isomorfismos então $S \circ T : V \to Z$ é um isomorfismo. Ou seja,*

$$V \cong W \text{ e } W \cong Z \Rightarrow V \cong Z$$

Logo o isomorfismo é uma relação de equivalência no conjunto de todos os espaços vetoriais sobre um corpo.

◁

Exemplo 3. *$\mathbb{K}_n \cong \mathcal{M}_{1,n}(\mathbb{K})$ através da transposição $A \to A^t$. Já mencionamos que $\mathbb{K}^n = \mathcal{M}_{1,n}(\mathbb{K})$. Assim, os espaços vetoriais $\mathbb{K}_n$, $\mathcal{M}_{1,n}(\mathbb{K})$ e $\mathbb{K}^n$ são isomorfos entre si.*

$$\mathbb{K}_n \cong \mathcal{M}_{1,n}(\mathbb{K}) \cong \mathbb{K}^n$$

◁

Notação 4. *A partir desse ponto, os vetores em $\mathbb{K}^n$ serão representados por vetores linhas ou colunas!*

Ou seja, $\boldsymbol{x} \in \mathbb{K}^n$ poderá ser escrito como

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

ou

$$\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}.$$

Exemplo 5. *No Exemplo 4 mostramos que dado um espaço vetorial V sobre $\mathbb{K}$ de dimensão n então V é isomorfo a $\mathbb{K}^n$,*

$$V \cong \mathbb{K}^n.$$

E consequentemente todos os espaços vetoriais de dimensão n são isomorfos entre si, i.e., a menos de isomorfismo existe um único espaço vetorial de dimensão n sobre $\mathbb{K}$.

◁

Em nosso próximo teorema, precisaremos de uma descrição isomorfa do espaço vetorial V^n introduzida no Exemplo 8.

Exemplo 6. *Neste exemplo, construímos um espaço vetorial isomorfo a V^n. Seja V um espaço vetorial sobre $\mathbb{K}$ e seja $n \in \mathbb{N}$. Considere a soma direta*

$$\underbrace{V \boxplus \cdots \boxplus V}_{n \text{ vezes}}$$

Suponha que A seja qualquer conjunto finito com cardinalidade n. Sem perda de generalidade, podemos assumir $A = \{1, \ldots, n\}$. Então $V^A = V^n$.

Existe um isomorfismo natural $T : V \boxplus \cdots \boxplus V \to V^n$ dado por $T((\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)) = f \in V^n$, onde $f(i) = \boldsymbol{x}_i$ para todos os $i = 1, \ldots, n$. O fato de que T é um isomorfismo é um exercício fácil, que deixamos ao leitor. Assim

$$V^n \cong \underbrace{V \boxplus \cdots \boxplus V}_{n \text{ vezes}}$$

Observe que para definirmos o isomorfismo anterior não tivemos que realizar nenhuma escolha de bases!!!

◁

Notação 7 -Isomorfismo natural. *Apesar da palavra "**natural**" possuir um significado preciso na Teoria das Categorias, em nosso uso queremos dizer apenas que o isomorfismo não depende da escolha de bases. Se um isomorfismo depender de uma base diremos que o isomorfismo é **acidental**.*

*Usaremos a palavra "**canônico**" para referir a uma escolha ou representação bem usual. Como a base canônica de $\mathbb{R}^n$. Destacamos que o uso do termo canônico não é pacificado na literatura sendo algumas vezes usado nesse sentido e outras vezes usado como sinônimo de natural.*

A partir desse ponto, identificaremos os espaços vetoriais $V \boxplus \cdots \boxplus V$ (n vezes), V^n e V^A com $|A| = n$ e escreva apenas V^n para representar qualquer um desses espaços.

Exemplo 8. *Dado $(\mathbb{K}^\infty)_0 = \{$sequências sobre $\mathbb{K}$ com com um número finito de termos não nulos$\}$, então*

$$\mathbb{K}[x] \cong (\mathbb{K}^\infty)_0$$

Um isomorfismo é dado por:

$$T[a_0 + a_1 x + \dots a_n x^n] = [a_0, a_1, \dots a_n, 0, \dots]$$

$\lhd$

Os espaços vetoriais isomorfos compartilham muitas propriedades, como mostra o próximo teorema. Se $T \in \mathrm{Hom}(V,\ W)$ e $S \subseteq V$, escreveremos $TS = \{Ts | s \in S\}$

3.9 Teorema

Sejam $T \in \mathrm{Hom}(V,\ W)$ um isomorfismo e $S \subseteq V$. Então

- *a* S gera V se, e somente se, $T(S)$ gera W.
- *b* S é linearmente independente em V se, e somente se, $T(S)$ for linearmente independente em W.

Um isomorfismo pode ser caracterizado como uma transformação linear $T : V \to W$ que leva uma base para V para uma base para W.

3.10 Teorema

Sejam V e W espaços vetoriais sobre $\mathbb{K}$. Então

- *a* uma transformação linear $T \in \mathrm{Hom}(V,\ W)$ é um isomorfismo se, e somente se, existir uma base $\underline{B}$ para V de modo que $T(\underline{B})$ é uma base para W. Nesse caso, T leva qualquer base de V para uma base de W.
- *b* $V \cong W$ se, e somente se, $\dim V = \dim W$.

O teorema a seguir diz que, a menos de isomorfismo, existe apenas um espaço vetorial de cada dimensão sobre um determinado corpo.

3.11 Teorema de Classificação dos Espaços Vetoriais

- *a* Se n é um número natural, então qualquer espaço vetorial n-dimensional sobre $\mathbb{K}$ é isomorfo a $\mathbb{K}^n$.
- *b* Se κ é qualquer número cardinal e se B for um conjunto de cardinalidade κ, qualquer espaço vetorial dimensional κ sobre $\mathbb{K}$ é isomorfo ao espaço vetorial $(\mathbb{K}^B)_0$ de todas as funções de B com valores em $\mathbb{K}$

Problema de Classificação

Um teorema de classificação responde ao seguinte problema de classificação "Quais são os objetos de um determinado tipo, a menos de equivalência?". Ele fornece uma enumeração não redundante: cada objeto é equivalente a exatamente uma classe.

e com suporte finito.

Demonstração. No Exemplo 5, vimos que qualquer espaço vetorial n-dimensional é isomorfo a $\mathbb{K}^n$.

Assim, suponha que B seja um conjunto de cardinalidade κ e permita que $(\mathbb{K}^B)_0$ seja o espaço vetorial de todas as funções de B a $\mathbb{K}$ com suporte finito. Deixamos ao leitor mostrar que as funções $\delta_b \in (\mathbb{K}^B)_0$ definidas para todos os $b \in B$ por

$$\delta_b(x) = \begin{cases} 1 & \text{se } x = b \\ 0 & \text{se } x \neq b \end{cases}$$

formam uma base para $(\mathbb{K}^B)_0$, denominada base canônica. Portanto, $\dim((\mathbb{K}^B)_0) = |B|$.

Segue-se que, para qualquer número cardinal κ, existe um espaço vetorial da dimensão κ. Além disso, qualquer espaço vetorial da dimensão κ é isomorfo a $(\mathbb{K}^B)_0$. ■

Exercícios

Ex. 3.18 --- Mostre que

$$V^n \cong \underbrace{V \boxplus \cdots \boxplus V}_{n \text{ vezes}}$$

Ex. 3.19 --- Seja $V = A + B$ e seja a soma direta externa $E = A \boxplus B$. Defina a aplicação $\tau : A \boxplus B \to A + B$ como $\tau : (a, b) = a + b$.

1. Prove que tal é linear.
2. Qual é o núcleo de τ?
3. Quando τ é um isomorfismo.

Ex. 3.20 --- Mostre que

$$\mathbb{K}[x] \cong (\mathbb{K}^\infty)_0$$

Ex. 3.21 --- Sejam V um espaço vetorial sobre $\mathbb{K}$ com dimensão finita e $T : V \to V$ uma transformação linear. Suponha que exista $G : V \to V$ tal que $T \circ G = \mathrm{I}_V$. Prove que T é um isomorfismo e $G = T^{-1}$. Dê um exemplo que mostre que isso é falso quando a dimensão de V não for finita.

Ex. 3.22 --- Encontre um espaço vetorial V e decomposições $V = A \oplus B = C \oplus D$ com $A \cong C$, mas $B \ncong D$. Portanto, $A \cong C$ não implica que $A^c \cong C^c$.

Ex. 3.23 --- Sejam V, A, B, C espaços vetoriais, sejam $F\colon\ A \to C$ e $G\colon\ B \to C$ sejam aplicações lineares e defina:

$$A \times c^B := \{(a,\ b) \in A \times B | F(a) = G(b)\}.$$

Mostre que:

1. O espaço $\mathrm{Hom}(V,\ A \oplus B)$ é isomorfo ao espaço $\mathrm{Hom}(V,\ A) \oplus \mathrm{Hom}(V,\ B)$

2. O espaço $\mathrm{Hom}(V, vezesa\ c^{B)}$ é isomorfo ao espaço:

   $\mathrm{Hom}(V,\ A) \times \mathrm{Hom}(V, C)\,\mathrm{Hom}(V,\ B) = \{(S,\ T) \in \mathrm{Hom}(V,\ A) \times \mathrm{Hom}(V,\ B) | FS = GT\}$

3.3 Teorema do Núcleo-Imagem

3.1 Teorema Teorema do Núcleo-Imagem

Sejam V e W espaços vetoriais de dimensão finita sobre $\mathbb{K}$ e $T : V \to W$ uma transformação linear, então:

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T)$$

Demonstração. Seja $\underline{\mathrm{B}}_1 = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$ uma base para o núcleo de T, ou seja, $\dim(\ker T) = r$. O núcleo de T é um subespaço de V. Pelo Teorema da Extensão, a base $\underline{\mathrm{B}}_1$ pode ser completada para uma base de V, $\underline{\mathrm{B}}_2 = \{\boldsymbol{u}_1, \ldots,\ \boldsymbol{u}_r,\ \boldsymbol{v}_1, \ldots,\ \boldsymbol{v}_s\}$. Vamos mostrar que $\underline{\mathrm{B}} = \{T(\boldsymbol{v}_1), \ldots,\ T(\boldsymbol{v}_s)\}$ é uma base para a imagem de T.

Dado $\boldsymbol{v} \in \operatorname{im} T$, então existe $\boldsymbol{u} \in V$ tal que $T(\boldsymbol{u}) = \boldsymbol{v}$. Mas, o vetor $\boldsymbol{u}$ pode ser escrito como combinação linear dos elementos da base B_2 de V, ou seja, podemos escrever $\boldsymbol{u}$ como

$$\boldsymbol{u} = a_1\boldsymbol{u}_1 + \ldots + a_r\boldsymbol{u}_r + b_1\boldsymbol{v}_1 + \ldots + b_s\boldsymbol{v}_s$$

Logo

$$\boldsymbol{v} = T(\boldsymbol{u}) = T(a_1\boldsymbol{u}_1 + \ldots + a_r\boldsymbol{u}_r + b_1\boldsymbol{v}_1 + \ldots + b_s\boldsymbol{v}_s) = \tag{3.11}$$

$$= a_1T(\boldsymbol{u}_1) + \ldots + a_rT(\boldsymbol{u}_r) + b_1T(\boldsymbol{v}_1) + \ldots + b_sT(\boldsymbol{v}_s) = b_1T(\boldsymbol{v}_1) + \ldots + b_sT(\boldsymbol{v}) \tag{3.12}$$

pois, como $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ pertencem ao núcleo de T, segue que $T(\boldsymbol{u}_1), \ldots, T(\boldsymbol{u}_r) = \mathbf{0}$.

Assim, dado um $\boldsymbol{v} \in \operatorname{im} T$, mostramos que ele pode ser escrito como combinação linear dos elementos do conjunto B, logo, $\operatorname{im} T = [\{T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_s)\}]$;

Falta mostrar que $\underline{\mathrm{B}}$ é linearmente independente. Considere a combinação linear nula:

$$b_1T(\boldsymbol{v}_1) + \ \ldots + b_sT(\boldsymbol{v}_s) = \mathbf{0} \Rightarrow T(b_1\boldsymbol{v}_1 + \ \ldots + b_s\boldsymbol{v}_s) = \mathbf{0}$$

Dessa forma, $b_1\boldsymbol{v}_1+\ldots+b_s\boldsymbol{v}_s \in \ker T$. Logo, esse elemento pode ser escrito como combinação linear dos elementos da base do núcleo, $\underline{\mathrm{B}}_1$:

$$b_1\boldsymbol{v}_1+\ldots+b_s\boldsymbol{v}_s = a_1\boldsymbol{u}_1+\ldots+a_r\boldsymbol{u}_r \Rightarrow a_1\boldsymbol{u}_1+\ldots+a_r\boldsymbol{u}_r+(-b_1)\boldsymbol{v}_1+\ldots+(-b_s)\boldsymbol{v}_s = \boldsymbol{0}$$

Como $\underline{\mathrm{B}}_2 = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r,\ \boldsymbol{v}_1, \ldots, \boldsymbol{v}_s\}$ é base de V, então é linearmente independente, logo todos os escalares da última igualdade são nulos. E logo $b_1 == b_s = \boldsymbol{0}$. Assim, provamos que $\underline{\mathrm{B}}$ é uma base para $\operatorname{im} T$, desta forma $dim(\operatorname{im} T) = s$.

Como $\dim V = r + s$, então:

$$\dim V = r + s = \dim(\ker T) + \dim(\operatorname{im} T)$$

■

Segunda Demonstração Vamos apresentar uma formulação mais abstrata do Teorema do Núcleo Imagem, bem como outra demonstração desse fato.

Seja $T \in \operatorname{Hom}(V,\ W)$. Como qualquer subespaço de V possui um complemento, podemos escrever $V = \ker T \oplus \ker T^{\mathsf{c}}$ onde $\ker T^{\mathsf{c}}$ é um complemento de $\ker T$ em V. Segue que

$$\dim(V) = \dim(\ker T) + \dim(\ker T^{\mathsf{c}}).$$

Dessa forma a restrição de T a $\ker T^{\mathsf{c}}$, $T^{\mathsf{c}} : \ker T^{\mathsf{c}} \to W$ é injetiva, já que $\ker(T^{\mathsf{c}}) = \ker T \cap \ker T^{\mathsf{c}} = \{\boldsymbol{0}\}$ Além disso, $\operatorname{im}(T^{\mathsf{c}}) \subseteq \operatorname{im} T$. Para inclusão inversa, se $T\boldsymbol{v} \in \operatorname{im} T$, então como $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ para $\boldsymbol{u} \in \ker T$ e $\boldsymbol{w} \in \ker T^{\mathsf{c}}$, temos

$$T\boldsymbol{v} = T\boldsymbol{u} + T\boldsymbol{w} = T\boldsymbol{w} = T^{\mathsf{c}}\boldsymbol{w} \in \operatorname{im}(T^{\mathsf{c}})$$

Assim, $\operatorname{im}(T^{\mathsf{c}}) = \operatorname{im} T$. Segue que

$$\ker T^{\mathsf{c}} \cong \operatorname{im} T$$

Provamos o seguinte teorema.

3.2 Teorema Teorema do Núcleo-Imagem

Seja $T \in \operatorname{Hom}(V,\ W)$.

a Qualquer complemento de $\ker T$ é isomorfo a $\operatorname{im} T$

b $\dim(\ker T) + \dim(\operatorname{im} T) = \dim(V)$

Se $A \in \mathcal{M}_{m,n}$, como a imagem de T_A é o espaço da coluna de A, temos

$$\dim(\ker(T_A)) + \operatorname{posto}(A) = \dim(\mathbb{K}^n)$$

Isso fornece o seguinte resultado útil.

3.3 Teorema

Seja $A \in \mathcal{M}_{n,m}(\mathbb{K})$:

1. $T_A : \mathbb{K}^n \to \mathbb{K}^m$ é injetiva se, e somente se, $\operatorname{posto}(A) = n$.
2. $T_A : \mathbb{K}^n \to \mathbb{K}^m$ é sobrejetiva se, e somente se, $\operatorname{posto}(A) = m$.

3.4 Teorema

Sejam V e W espaços vetoriais sobre $\mathbb{K}$ e suponha que $T \in \operatorname{Hom}(V, W)$. Então

a. Se T for sobrejetiva então $\dim V \geq \dim W$.

b. Se $\dim V = \dim W < \infty$, T é um isomorfismo se, e somente se T for injetiva ou T for sobrejetiva.

Demonstração. a Segue imediatamente do Teorema 2.3.

b Se T é um isomorfismo então T é injetiva e sobrejetiva. Se T é injetiva $\ker T = \mathbf{0}$ e pelo Teorema do Núcleo-Imagem $\dim \operatorname{im} T = n$ e logo $\operatorname{im} T = W$ e é isomorfismo. Se T é sobrejetiva $\operatorname{im} T = W$ e pelo Teorema do Núcleo-Imagem $\dim \ker T = 0$ e logo T é isomorfismo.

■

Observe o resultado do b não é verdadeiro se os espaços vetoriais não tiverem dimensões finitas.

3.5 Proposição

Se $U \xrightarrow{T} V \xrightarrow{S} W$ são espaços vetoriais de dimensão finita e aplicações lineares, então

$$\dim \ker (S \circ T) \leq \dim \ker S + \dim \ker T.$$

Demonstração. A demonstração será deixada como exercício. ■

Terminamos esta seção com uma generalização do Teorema 3.3 b. Precisamos da seguinte definição.

3.6 Definição Complexo de Cadeias

Um **complexo de cadeias** $C = \{(V_i, d_i) \,|\, i \in \mathbb{Z}\}$ de espaços vetoriais sobre $\mathbb{K}$, é uma sequência infinita $\{V_i\}$ de espaços vetoriais V_i, um para todo número inteiro $i \in \mathbb{Z}$, juntamente com uma sequência $\{d_i\}$ de transformações lineares, $d_i \in \mathrm{Hom}(V_i, V_{i-1})$ para $i \in \mathbb{Z}$, de modo que $d_{i+1}d_i = 0$ para todos os $i \in \mathbb{Z}$.

A condição $d_{i+1}d_i = 0$ nos diz que $\mathrm{im}\, d_{i+1} \subset \ker d_i$

Geralmente, desenhamos um complexo de cadeias como uma sequência infinita de espaços e aplicações da seguinte maneira:

$$C : \cdots \longrightarrow V_{i+1} \xrightarrow{d_{i+1}} V_i \xrightarrow{d_1} V_{i-1} \longrightarrow \cdots \tag{3.13}$$

Se um complexo de cadeias C tiver apenas um número finito de termos diferentes de zero, poderemos simplificar a notação e escrever C como

$$C : 0 \longrightarrow V_n \xrightarrow{d_n} V_{n-1} \xrightarrow{d_{n-1}} \ldots \longrightarrow V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.14}$$

Entende-se aqui que todos os outros espaços vetoriais e aplicações que não aparecem explicitamente na Equação 3.14 são zeros.

3.7 Definição Exato

Um complexo de cadeias

$$C : \cdots \longrightarrow V_{i+1} \xrightarrow{d_{1+1}} V^i \xrightarrow{d_i} V_{i-1} \longrightarrow \cdots \tag{3.15}$$

é dito **exato** se $\mathrm{im}\, d_{i+1} = \ker d_i$ para todo $i \in \mathbb{Z}$.

3.8 Proposição

Seja $T : X \to Y$ um operador linear. Então:

1. A sequência $X \to^T Y \to 0$ é exata se, e somente se, T é sobrejetivo.
2. A sequência $0 \to X \to^T Y$ é exata se, e somente se, T é injetivo.
3. A sequência $0 \to X \to^T Y \to 0$ é exata se, e somente se, T for bijetivo.

Exemplo 9. *Sejam V e W espaços vetoriais sobre $\mathbb{K}$, e seja $T \in \mathrm{Hom}(V, W)$. Então*

$$C : 0 \longrightarrow \ker T \xrightarrow{i} V \xrightarrow{T} \mathrm{im}\, T \longrightarrow 0 \tag{3.16}$$

é um complexo de cadeia exato. Na equação anterior i indica a inclusão de $\ker T$ em V.

◁

Podemos generalizar o exemplo ligeiramente a seguir:

3.10 Definição Sequência Exata Curta

Uma **sequência exata curta**, é um complexo de cadeia exato C da seguinte forma:

$$C : 0 \longrightarrow V_2 \xrightarrow{d_2} V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.17}$$

Portanto, no Exemplo 9 temos uma sequência exata curta com $V_2 = \ker T, \mathrm{d}_2 = i, V_1 = V$ etc. Claramente, o complexo de cadeia C da Equação 3.18 é uma sequência exata curta se, e somente se, d_2 é injetiva, d_1 é sobrejetiva e $\operatorname{im} d_2 = \ker d_1$.

- O Teorema 3.3 $\boxed{c}$ implica que, se C for uma sequência exata curta,

$$\dim V_2 - \dim V_1 + \dim V_0 = 0.$$

Essa é outra forma do Teorema do Núcleo Imagem:

3.11 Teorema Teorema do Núcleo-Imagem

Dada uma sequência exata curta

$$C : 0 \longrightarrow V_2 \xrightarrow{d_2} V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.18}$$

então

$$\dim V_2 - \dim V_1 + \dim V_0 = 0.$$

Agora podemos provar a seguinte generalização desse resultado:

3.12 Teorema

Suponha que

$$C : 0 \longrightarrow V_n \xrightarrow{d_n} V_{n-1} \xrightarrow{d_{n-1}} \dots \longrightarrow V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.19}$$

seja um complexo de cadeia exato. Então $\sum_{i=0}^{n} (-1)^1 \dim V_i = 0.$

Demonstração. O complexo de cadeias C se decompõe nas seguintes sequências exatas curtas

$$C_1 : 0 \longrightarrow \ker d_1 \longrightarrow V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \tag{3.20}$$

$$C_2 : 0 \longrightarrow \ker d_2 \longrightarrow V_2 \xrightarrow{d_2} \ker d_1 \longrightarrow 0$$

$$\vdots$$

$$C_n : 0 \longrightarrow \ker d_n \longrightarrow V_n \xrightarrow{d_n} \ker d_{n-1} \longrightarrow 0$$

Aplicando o Teorema 3.3 $\boxed{c}$ a cada C_i e adicionarmos os resultados, obteremos $\sum_{i=0}^{n}(-1)^i \dim V_i = 0$. ■

Exercícios

Ex. 3.24 --- Seja $T : \mathbb{K}_n[x] \to \mathbb{K}_n[x]$ o operador linear definido por $Tp(x) = 3p(x) - 6p'(x) + p''(x)$.

1. Mostre que $\ker T = \{\mathbf{0}\}$.
2. Conclua que para todo polinômio $q(x)$ existe um polinômio $p(x)$ tal que $3p(x) - 6p'(x) + p''(x) = q(x)$.

Ex. 3.25 --- Sejam V um espaço vetorial de dimensão finita e T :V→ V uma aplicação linear. Mostre que

$$V = \ker T \oplus \operatorname{im} T$$

se, e somente se, $\ker T = \ker T^2$.

Ex. 3.26 --- Sejam V e W espaços vetoriais de dimensão finita e T :V→ W uma transformação linear. Prove que, se T é sobrejetiva, então $\dim(V) \geq \dim(W)$.

Ex. 3.27 --- Sejam V e W espaços vetoriais de dimensão finita e T :V→ W uma transformação linear. Prove que, se T é injetiva, $\dim(V) \leq \dim(W)$.

Ex. 3.28 --- Sejam V e W espaços vetoriais de dimensão finita e T :V→ W uma transformação linear sobrejetiva. Prove que existe uma transformação linear S :W→V tal que $T \circ S = \mathrm{I}_W$.

Ex. 3.29 --- Sejam V e W espaços vetoriais de dimensão finita e T :V→ W uma transformação linear injetiva. Prove que existe uma transformação linear S :W→Vtal que $S \circ T = \mathrm{I}_V$.

Ex. 3.30 --- Prove que toda transformação linear A pode ser escrita como o produto $A = TS$, onde S é uma transformação linear sobrejetiva e T é uma transformação linear injetiva. Vale também uma decomposição do tipo $A = ST$, com S sobrejetiva e T injetiva?

Ex. 3.31 --- Seja V um espaço vetorial n-dimensional sobre um corpo $\mathbb{K}$ e T um operador em V. Prove que $\ker(T^n) = \ker(T^{n+1})$ e $\operatorname{im}(T^n) = \operatorname{im}(T^{n+1})$.

Ex. 3.32 --- Seja V um espaço vetorial n-dimensional sobre um corpo $\mathbb{K}$ e T um operador em V. Prove que $V = \operatorname{im}(T^n) \oplus \ker(T^n)$

Ex. 3.33 --- Uma *sequência exata* de $\mathbb{K}$-espaços vetoriais V_i de dimensão finita

$$V_0 \stackrel{T_0}{\to} V_1 \stackrel{T_1}{\to} V_2 \stackrel{T_2}{\to} V_3 \stackrel{T_3}{\to} \cdots \stackrel{T_n}{\to} V_n$$

é uma sequência de transformações lineares T_i tais que $\ker T_{i+1} = \operatorname{Im} T_i$ para $i = 0, \dots n$.

1. Mostre que a sequência $0 \stackrel{T_0}{\to} V_1 \stackrel{T_1}{\to} V_2$ é exata se, e somente se, T_1 é injetiva.

2. Mostre que a sequência $V_0 \stackrel{T_0}{\to} V_1 \stackrel{T_1}{\to} 0$ é exata se, e somente se, T_0 é sobrejetiva.

3. Considere a sequência exata

$$0 \stackrel{T_0}{\to} V_1 \stackrel{T_1}{\to} V_2 \stackrel{T_2}{\to} V_3 \stackrel{T_3}{\to} V_4 \stackrel{T_4}{\to} 0$$

   na qual $\dim_{\mathbb{K}} V_i = d_i$, mostre que $d_1 + d_3 = d_2 + d_4$.

Ex. 3.34 --- Um *complexo* é definido com sendo um par (C, d) onde $C = (C_i)_{i \in \mathbb{Z}}$ é uma família de espaços vetoriais, junto com uma família de $d = (d_i)_{i \in \mathbb{Z}}$ de aplicações lineais $d_i : C_i \to C_{i+1}$ satisfazendo a seguinte condição $d_{i+1} \circ d_i = 0$. Um complexo é usualmente representado como segue:

$$C : \cdots \xrightarrow{d_{i-2}} C_{i-1} \xrightarrow{d_{i-1}} C_i \xrightarrow{d_i} C_{i+1} \xrightarrow{d_{i+1}} \cdots$$

1. O *i-ésimo espaço de cohomologia* de um complexo (C, d) é definido como o quociente $H^i(C, d) := \ker(d_i)/\operatorname{Im}(d_{i-1})$. Mostre que este espaço está bem definido, i.e. mostre que para todo $i \in \mathbb{Z}$ temos $\operatorname{Im}(d_{i-1}) \subseteq \ker(d_i)$.

2. Suponha que o complexo (C, d) tem dimensão finita (i.e. $\dim(C_i) < \infty$ para todo $i \in \mathbb{N}$) e é limitado, ou seja, existem $n, m \in \mathbb{Z}$ tais que $C_i = \{0\}$

se $i \notin \{m, \dots, n\}$. Neste caso podemos definir a *característica de Euler* de (C, d) como sendo

$$\chi(C,d) := \sum_{k=m}^{n} (-1)^k \dim(C_k).$$

Mostre que $\chi(C,d) = \sum_{k=m}^{n} (-1)^k \dim(H^k(C,d))$.

3.4 Representação Matricial dos Homomorfismos

3.1 Teorema

Suponha que V e W sejam espaços vetoriais tais que $\dim V = n$ e $\dim W = m$. Se $\underline{X} = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$ for uma base de V e $\underline{Y} = \{\boldsymbol{y}_1, \dots, \boldsymbol{y}_m\}$ for uma base de W então o par $(\underline{X}, \underline{Y})$ determina um isomorfismo $[\,\cdot\,]_{\underline{X},\underline{Y}} : \mathrm{Hom}(V,W) \to \mathcal{M}_{m,n}(\mathbb{K})$ definida como

$$T \mapsto [T]_{\underline{X},\underline{Y}}$$

onde $[T]_{\underline{X},\underline{Y}} = \left([T(\boldsymbol{x}_1)]_{\underline{Y}} \,\middle|\, \cdots \,\middle|\, [T(\boldsymbol{x}_n)]_{\underline{Y}}\right)$ é a matriz cuja i-ésima coluna é o vetor $[T(\boldsymbol{x}_i)]_{\underline{Y}}$ Além disso, o diagrama comuta

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "(\,\cdot\,)_{\underline{X}}"'] & W \arrow[d, "(\,\cdot\,)_{\underline{Y}}"] \\ \mathbb{K}_n \arrow[r, "{[T]_{\underline{X},\underline{Y}}}"] & \mathbb{K}_n \end{tikzcd} \tag{3.21}$$

Demonstração. **Linear** Se $T_1, T_2 \in \mathrm{Hom}(V,W)$ e $\lambda_1, \lambda_2 \in \mathbb{K}$ então

$$[\lambda_1 T_1 + \lambda_2 T_2]_{\underline{X},\underline{Y}} = \left([(\lambda_1 T_1 + \lambda_2 T_2)(\boldsymbol{x}_1)]_{\underline{Y}} \,\middle|\, \cdots \,\middle|\, [(\lambda_1 T_1 + \lambda_2 T_2)(\boldsymbol{x}_n)]_{\underline{Y}}\right) \tag{3.22}$$

$$= \left(\lambda_1 [T_1(\boldsymbol{x}_1)]_{\underline{Y}} + \lambda_2 [T_2(\boldsymbol{x}_1)]_{\underline{Y}} \,\middle|\, \cdots \,\middle|\, \lambda_1 [T_1(\boldsymbol{x}_n)]_{\underline{Y}} + \lambda_2 [T_2(\boldsymbol{x}_n)]_{\underline{Y}}\right) \tag{3.23}$$

$$= \lambda_1 \left([T_1(\boldsymbol{x}_1)]_{\underline{Y}} \,\middle|\, \cdots \,\middle|\, [T_1(\boldsymbol{x}_n)]_{\underline{Y}}) + \lambda_2 ([T_2(\boldsymbol{x}_1)]_{\underline{\boldsymbol{y}}} \,\middle|\, \cdots \,\middle|\, [T_2(\boldsymbol{x}_n)]_{\underline{Y}} \right) \tag{3.24}$$

$$= \lambda_1 \, [T_1]_{\underline{X},\underline{Y}} + \lambda_2 \, [T_2]_{\underline{X},\underline{Y}} \tag{3.25}$$

Portanto, $[\,\cdot\,]_{\underline{X},\underline{Y}}$ é de fato uma transformação linear de $\mathrm{Hom}(V,W)$ para $\mathcal{M}_{m,n}(\mathbb{K})$.

Isomorfismo

Suponha que $T \in \ker\, [\,\cdot\,]_{\underline{X},\underline{Y}}$ Então $[T]_{\underline{X},\underline{Y}} = 0$. Em particular, $[T\boldsymbol{x}_i]_{\underline{Y}} = 0$ para todos os $i = 1, \dots, n$. O Teorema 5.2.1 implica $T(\boldsymbol{x}_i) = \boldsymbol{0}$. Assim, $T = 0$, e concluímos que $[\,\cdot\,]_{\underline{X},\underline{Y}}$ é uma transformação linear injetiva.

A transformação linear $[\,\cdot\,]_{\underline{X},\underline{Y}}$ também é sobrejetiva. Para ver isso, seja $A = (x_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$. Seja $\boldsymbol{z}_i = \sum_{j=1}^{m} x_{ji}\boldsymbol{y}_j$ para $i = 1, \ldots, n$. Então $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\} \subseteq W$ e $[\boldsymbol{z}_i]_{\boldsymbol{y}} = (x_{li}, \ldots, x_{mi})^{\mathrm{t}} = \mathrm{Col}_i(A)$ para todos os $i = 1, \ldots, n$. A partir do Teorema 3.1, temos que existe uma única $T \in \mathrm{Hom}(V, W)$ tal que $T(\boldsymbol{x}_i) = \boldsymbol{z}_i$ por $i = 1, \ldots, n$. Assim, $[T]_{\underline{X},\underline{Y}} = A$ e $[\,\cdot\,]_{\underline{X},\underline{Y}}$ é sobrejetiva.

Comuta

Precisamos apenas argumentar que o diagrama é comutativo. Para qualquer $\boldsymbol{x}_i \in \underline{X}$, temos

$$\begin{aligned}([\,\cdot\,]_{\underline{Y}}T)(\boldsymbol{x}_i) &= [T(\boldsymbol{x}_i)]_{\underline{Y}} \\ &= \mathrm{Col}_i([T]_{\underline{X},\underline{Y}}) \\ &= [T]_{\underline{X},\underline{Y}}\,(0, \ldots, 1, \ldots, 0)^t \\ &= [T]_{\underline{X},\underline{Y}}\,[\boldsymbol{x}_i]_{\underline{X}}.\end{aligned}$$

■

3.2 Definição Matriz da Transformação Linear

A matriz $[T]_{\underline{X},\underline{Y}}$ é denominada a **matriz da transformação linear** T em relação às bases $\underline{X}$ e $\underline{Y}$.

Como as setas verticais em 3.21 e $[\,\cdot\,]_{\underline{X},\underline{Y}}$ são isomorfismos, V, $W, \mathrm{Hom}(V, W)$ e T geralmente são identificados com $\mathbb{K}_n, \mathbb{K}_m, \mathcal{M}_{m,n}(\mathbb{K})$ e $A = [T]_{\underline{X},\underline{Y}}$. Assim, a distinção entre uma transformação linear e uma matriz é frequentemente obscurecida na literatura.

Exemplo 3. *Seja $V = \mathbb{K}_3[x]$ o espaço de todos os polinômios de grau menor ou igual a 3. O operador de diferenciação formal D do Exemplo 9 leva V em V, pois a derivação diminui os graus dos polinômios. Seja $\underline{B} = (p_1, p_2, p_3, p_4) = (1, x, x^2, x^3)$ a base ordenada canônica de V. Então*

$$\begin{aligned} Dp_1 &= 0, & Dp_1 &= 0p_1 + 0p_2 + 0p_3 + 0p_4 \\ Dp_2 &= 1, & Dp_2 &= 1p_1 + 0p_2 + 0p_l + 0p_4 \\ Dp_3 &= 2x, & Dp_3 &= 0p_1 + 2p_2 + 0p_3 + 0p_4 \\ Dp_4 &= 3x^2, & Dp_4 &= 0p_1 + 0p_2 + 3p_3 + 0p_4 \end{aligned}$$

e logo a matriz de D na base ordenada $\underline{B}$ é

$$[D]_{\underline{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

◁

3.4 Corolário

Seja $\underline{v}$, $\underline{w}$ sejam bases para V e W, respectivamente. Seja $T, S \in \mathrm{Hom}(V, W)$. Então

1. $[cT]_{\underline{v},\,\underline{w}} = c[T]_{\underline{v},\,\underline{w}}$
2. $[T+S]_{\underline{v},\,\underline{w}} = [T]_{\underline{v},\,\underline{w}} + [S]_{\underline{v},\,\underline{w}}$.

Sejam A uma matriz $m \times n$ com colunas $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ e e $c = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ um vetor em $\mathbb{K}_n$. Então, o produto de A por $\boldsymbol{x}$ é

$$A\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n.$$

Uma consequência imediata dessa representação é

3.5 Teorema

Seja V um espaço vetorial dimensional n com base $\underline{v}$, W um espaço vetorial dimensional com base $\underline{w}$ e $T : V \to W$ uma transformação linear. Então, para um vetor arbitrário $v \in V$

$$[T(\boldsymbol{v})]_{\underline{vw}} = [T]_{\underline{v},\,\underline{w}}\,[v]_{\underline{v}}.$$

Seja A uma matriz $m \times n$ com colunas $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$

$$A = \left(\boldsymbol{a}_1 \,\middle|\, \boldsymbol{a}_2 \,\middle|\, \ldots \,\middle|\, \boldsymbol{a}_n\right)$$

e B a matriz $p \times m$. Então, o produto de B e A é definido como a matriz $p \times n$ cuja j-ésima coluna é $B a_j$. Portanto,

$$BA = \left(B\boldsymbol{a}_1 \,\middle|\, B\boldsymbol{a}_2 \,\middle|\, \ldots \,\middle|\, B\boldsymbol{a}_n\right).$$

Como consequência dessa definição, temos:

3.6 Teorema

Sejam V, W, X espaços vetoriais finito dimensionais com bases $\underline{v}$, $\underline{w}$ e$\underline{x}$ respectivamente e $T: V \to W$ e $S: W \to X$ transformações lineares. Então

$$[S \circ T]_{\underline{\mathsf{V}},\ \underline{\mathsf{X}}} = [S]_{\underline{\mathsf{W}},\ \underline{\mathsf{X}}}[T]_{\underline{\mathsf{V}}\ \underline{\mathsf{W}}}$$

Demonstração. Para esse fim, calculamos o vetor de coordenadas de $(S \circ T)(\boldsymbol{v}_j)$ em relação à base $\underline{\mathsf{X}}$. Vamos definir $[T]_{\underline{\mathsf{V}},\ \underline{\mathsf{W}}} = A$ e $[S]_{\underline{\mathsf{W}},\ \underline{\mathsf{X}}} = B$. Pela definição de composição

$$(S \circ T)(\boldsymbol{v}_j) = S(T(\boldsymbol{v}_j)) \ .$$

Tomando vetores de coordenadas, obtemos

$$[(S \circ T)(\boldsymbol{v}_j)]_{\underline{\mathsf{X}}} = [S(T(\boldsymbol{v}_j))]_{\underline{\mathsf{X}}}.$$

Pelo Teorema 3.4, segue-se que

$$[S(T(\boldsymbol{v}_j))]_{\underline{\mathsf{X}}} = B[T(\boldsymbol{v}_j)]_{\underline{\mathsf{W}}}.$$

Pela definição de $[T]_{\underline{\mathsf{V}},\ \underline{\mathsf{W}}}$, segue-se que

$$[T(\boldsymbol{v}_j)]_{\underline{\mathsf{W}}} = a_j,$$

e, portanto, a j-ésima coluna de $[S \circ T](\underline{\mathsf{V}},\ \underline{\mathsf{X}})$ é Ba_j. ■

O teorema anterior é a motivação para o produto de matrizes.

3.4.1 Mudança de Bases

A representação matricial $[T]_{\underline{\mathsf{X}},\underline{\mathsf{Y}}}$ de T depende, é claro, das bases específicas $\underline{\mathsf{X}}$ e $\underline{\mathsf{Y}}$ escolhidas. É fácil acompanhar como $[T]_{\underline{\mathsf{X}},\underline{\mathsf{Y}}}$ muda com $\underline{\mathsf{X}}$ e $\underline{\mathsf{Y}}$

3.7 Teorema Mudança de Base

Seja V e W os espaços vetoriais de dimensão finita sobre $\mathbb{K}$ das dimensões n e m, respectivamente. Suponha que $\underline{\mathsf{X}}$ e $\underline{\mathsf{X}'}$ sejam duas bases de V e $\underline{\mathsf{Y}}$ e $\underline{\mathsf{Y}'}$ duas bases de W. Então, para todo $T \in \mathrm{Hom}(V, W)$, temos

$$[T]_{\underline{\mathsf{X}'},\underline{\mathsf{Y}'}} = \mathrm{M}_{\underline{\mathsf{Y}}\to\underline{\mathsf{Y}'}}\,[T]_{\underline{\mathsf{X}},\underline{\mathsf{Y}}}\,\mathrm{M}^{-1}_{\underline{\mathsf{X}}\to\underline{\mathsf{X}'}} \tag{3.26}$$

Demonstração. Já observamos que as matrizes de mudança de bases são invertíveis e, consequentemente, todos os termos da equação 3.26 fazem sentido.

Para ver que a equação 3.26 é de fato verdadeira, apenas combinamos o diagramas comutativos 2.16 e 3.21. Considere o seguinte diagrama:

$$\begin{tikzcd}
\mathbb{K}_n \arrow[rrr, "{[T]_{\underline{X},\underline{Y}}}", "\boxed{3}"'] \arrow[dd, bend right, "M_{\underline{X}\to\underline{X}'}"'] & & & \mathbb{K}_m \arrow[dd, bend left, "M_{\underline{Y}\to\underline{Y}'}"] \\
\boxed{1} & V \arrow[ul, "{[\,\cdot\,]_{\underline{X}}}"'] \arrow[r, "T"] \arrow[dl, "{[\,\cdot\,]_{\underline{X}'}}"] & W \arrow[ur, "{[\,\cdot\,]_{\underline{Y}}}"'] \arrow[dr, "{[\,\cdot\,]_{\underline{Y}'}}"'] & \boxed{2} \\
\mathbb{K}_n \arrow[rrr, "\boxed{4}", "{[T]_{\underline{X}',\underline{Y}'}}"'] & & & \mathbb{K}_m
\end{tikzcd} \tag{3.27}$$

O diagrama 3.27 é composto de nossas partes, que rotulamos como $\boxed{1}$, $\boxed{2}$, $\boxed{3}$, $\boxed{4}$. Pelo Teorema 2.6, os diagramas $\boxed{1}$, $\boxed{2}$ são comutativos. Pelo Teorema ??, os diagramas $\boxed{3}$, $\boxed{4}$ são comutativos. Segue-se que todo o diagrama 3.27 é comutativo. Em particular, $M_{\underline{Y}\to\underline{Y}'}[T]_{\underline{X},\underline{Y}} = [T]_{\underline{X}',\underline{Y}'}M_{\underline{X}\to\underline{X}}$ Resolvendo esta equação para $[T]_{\underline{X}',\underline{Y}'}$ temos 3.26. ■

Exemplo 8. *Seja $V = \mathbb{K}_3[x]$ o espaço de todos os polinômios de grau menor ou igual a 3 e D o operador de diferenciação como no exemplo 3. Sejam* $\underline{B} = (p_1, p_2, p_3, p_4) = (1, x, x^2, x^3)$ e $\underline{C} = (q_1, q_2, q_3, q_4) = (1, (x-1), (x-1)^2, (x-1)^3)$ *duas bases ordenadas de D. Então*

$$\begin{aligned}
q_1 &= p_1\\
q_2 &= -1p_1 + p_2\\
q_3 &= p_1 - 2p_2 + p_3\\
q_4 &= -1p_1 + 3p_2 - 3p_3 + p_4.
\end{aligned}$$

Pelo Exemplo 7 as matrizes de mudança de base são

$$M_{\underline{C}\to\underline{B}} = \begin{bmatrix} 1 & -1 & 1 & -1\\ 0 & 1 & -2 & 3\\ 0 & 0 & 1 & -3\\ 0 & 0 & 0 & 1 \end{bmatrix} \quad M_{\underline{B}\to\underline{C}} = \begin{bmatrix} 1 & 1 & 1 & 1\\ 0 & 1 & 2 & 3\\ 0 & 0 & 1 & 3\\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Pelo Exemplo 3 sabemos que:

$$[D]_{\underline{B}} = \begin{bmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A matriz de D na base ordenada $\underline{C}$ será

$$\begin{aligned}
[D]_{\underline{C}} &= M_{\underline{B}\to\underline{C}}[D]_{\underline{B}}M_{\underline{C}\to\underline{B}}\\
&= \begin{bmatrix} 1 & 1 & 1 & -1\\ 0 & 1 & 2 & 3\\ 0 & 0 & 1 & 3\\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 & -1\\ 0 & 1 & -2 & 3\\ 0 & 0 & 1 & -3\\ 0 & 0 & 0 & 1 \end{bmatrix}\\
&= \begin{bmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

Logo D é representado pela mesma matriz nas bases ordenadas $\underline{B}$ e $\underline{C}$. É claro que poderíamos ter visto isso diretamente, pois

$$\begin{aligned} Dq_1 &= 0 \\ Dq_2 &= q_1 \\ Dq_3 &= 2q_2 \\ Dq_4 &= 3q_3. \end{aligned}$$

◁

Exercícios

Ex. 3.35 --- Seja $T : M_2(\mathbb{R}) \to M_2(\mathbb{R})$ uma transformação linear dada por

$$T \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} 0 & x \\ z-w & 0 \end{bmatrix}.$$

1. Determine a matriz de T com relação à base canônica.
2. Determine a matriz de T com relação à base

$$\underline{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \right\}$$

de $M_2(\mathbb{R})$.
3. Exiba a matriz M tal que $[T]^{\underline{B}}_{\underline{B}} = M^{-1}[T]^{\mathrm{Can}}_{\mathrm{Can}} M$.

Ex. 3.36 --- Seja $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Mostre que:

1. Dadas A e B duas matrizes sobre $\mathbb{K}$ é impossível termos $AB - BA = I$.
2. Conclua que não existem transformações lineais T, G sobre $\mathbb{K}$-espaços vetoriais de dimensão finita tais que $T \circ G - G \circ T = \mathrm{Id}$.
3. Mostre que isso não acontece se desconsiderarmos a hipóteses de dimensão finita.

Ex. 3.37 --- Seja $T : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ a função dada por $T[(a, b, c)_{\underline{B}}] = (2c - 2b, a + c, a + b + c)_{\underline{C}}$ onde $\underline{C}$ é a base $\{1, x, x^2\}$ e $\underline{B}$ é a base $\{1, x + x^2, 1 + x^2\}$.

1. Verifique que T é uma transformação linear.
2. Existe um vetor $u \in \mathbb{R}_2[x]$ não nulo tal que $T(u) = u$? Justifique sua resposta.

Ex. 3.38 --- Seja W um subespaço de V com $m = \dim W < \dim V = n$. Seja $Z = \{T \in \mathrm{Hom}(V,V)T\boldsymbol{w} = 0$ para todos os $\boldsymbol{w} \in W\}$. Mostre que Z é um subespaço de $\mathrm{Hom}(V,V)$ e calcule sua dimensão.

Ex. 3.39 --- Prove que

Ex. 3.40 --- Sejam V e W espaços vetoriais de dimensão finita sobre $\mathbb{K}$ com bases $\underline{\mathrm{v}}$ e $\underline{\mathrm{w}}$, respectivamente. Suponha que $T : V \to W$ é uma transformação linear e $A = [T]_{\underline{\mathrm{v}},\underline{\mathrm{w}}}$. Prove que um vetor $\boldsymbol{v} \in \ker T$ se, e somente se, $[\boldsymbol{v}] \in \mathrm{nulo} A$. (O espaço-solução do sistema homogêneo de equações $A\boldsymbol{x} = \mathbf{0}$ é um subespaço de R^n, denominado de espaço nulo de A.)

3.5 Somas Diretas e Projeções

3.5.1 Soma Direta de Transformações Lineares

3.1 Definição Soma Direta de Transformações Lineares

Suponha que $U = V \oplus W$ e $T_1 : V \to V$ e $T_2 : W \to W$ sejam transformações lineares. A soma direta de T_1 e T_2, é a transformação linear $T_1 \oplus T_2$ de U para U definida por

$$T_1 \oplus T_2(\boldsymbol{v} + \boldsymbol{w}) = T_1(\boldsymbol{v}) + T_2(\boldsymbol{v})$$

Observamos que os subespaços V e W são invariantes por $T_1 \oplus T_2$, ou seja, $T_1 \oplus T_2(V) \subseteq V$ e $T_1 \oplus T_2(W) \subseteq W$.

3.2 Teorema

Sejam V_i espaços vetoriais com base $\underline{\mathrm{B}}_i = \left(\boldsymbol{v}_{i_1}, \ldots, \boldsymbol{v}_{i_{k_i}}\right)$ e $\dim V_i = k_i$, para $1 \leq i \leq j$. Sejam $T_i : V_i \to V_i$ transformações lineares. Finalmente seja

$$T_1 \oplus \cdots \oplus T_j : V = V_1 \oplus V_2 \oplus \cdots \oplus V_j \to V = V_1 \oplus V_2 \oplus \cdots \oplus V_j.$$

Então a representação de T na base ordenada obtida pela concatenação

concatenação
Dadas duas listas ordenadas $\underline{\mathrm{L}}_1 = (a_1 \ldots, a_n)$ e $\underline{\mathrm{L}}_2 = (b_1 \ldots, b_m)$ a **concatenação** de $\underline{\mathrm{L}}_1$ e $\underline{\mathrm{L}}_1$ denotada $\underline{\mathrm{L}}_1 + \!\!+ \underline{\mathrm{L}}_1$ é a lista ordenada

$\underline{\mathrm{L}}_1 + \!\!+ \underline{\mathrm{L}}_1 \triangleq (a_1 \ldots, a_n, b_1 \ldots, b_m)$

$\underline{\mathrm{B}} = \underline{\mathrm{B}}_1 \mathbin{+\!\!+} \underline{\mathrm{B}}_2 \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} \underline{\mathrm{B}}_n$ é

$$T_{\underline{\mathrm{B}}} = \begin{bmatrix} T_1 & 0 & 0 \\ 0 & T_2 & 0 \\ \vdots & \dots & \vdots \\ 0 & 0 & T_j \end{bmatrix},$$

com T_i é um bloco de tamanho $k_i \times k_i$:

$$T_i = \begin{bmatrix} c_{11} & c_{21} & c_{1k_i} \\ c_{21} & c_{22} & c_{2k_i} \\ \vdots & \vdots & \vdots \\ c_{k_i 1} & c_{k_i 2} & c_{k_i k_i} \end{bmatrix}.$$

Demonstração. Seja $\underline{\mathrm{B}}_i = \{\boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_{k_i}}\}$ uma base de V_i. Então

$$\underline{\mathrm{B}} = \{\boldsymbol{v}_{1_1}, \dots, \boldsymbol{v}_{1_{k_1}}, \boldsymbol{v}_{2_1}, \dots, \boldsymbol{v}_{2_{k_2}}, \dots, \boldsymbol{v}_{j_1}, \dots, \boldsymbol{v}_{j_{k_j}}\}$$

é uma base de V. Do fato que $T(V_i) \subset V_i$ temos que

$$T(\boldsymbol{v}_{i_m}) = c_{1m}\boldsymbol{v}_{i_1} + c_{2m}\boldsymbol{v}_{i_2} + \cdots + c_{k_i m}\boldsymbol{v}_{k_i}.$$

■

Por outro lado, suponha que $T : U \to U$ seja uma transformação linear que deixe os subespaços V e W invariantes, ou seja, $T(V) \subseteq V$ e $T(W) \subseteq W$. Podemos então definir uma transformação linear restrita $T_1 : V \to V$ por $T_1(\boldsymbol{v}) = T(\boldsymbol{v})$ e uma transformação linear restrita $T_2 : W \to W$ por $T_2(w) = T(w)$. Então podemos escrever $T = T_1 \oplus T_2$. Se pudermos entender as transformações T_1 e T_2, obteremos uma boa imagem de T.

3.3 Definição Subespaço Invariante

Dado $T \in \mathrm{Hom}(V, V)$ dizemos que um subespaço $W \subset V$ é **invariante** por T (ou ainda que W é **T-invariante**) se $T(W) \subset W$.

Dado $T \in \mathrm{Hom}(V, V)$ e $W \subset V$ um espaço T-invariante, podemos restringir T a W obtendo $T\restriction_W \in \mathrm{Hom}(W, W)$.

3.4 Teorema

Dado um operador $T : V \to V$ e uma decomposição de V em subespaços T-invariantes

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_j.$$

com $\dim V_i = n_i$, para $1 \leq i \leq j$. Então a representação de T na base B é

$$T_{\underline{B}} = \begin{bmatrix} T_1 & 0 & 0 \\ 0 & T_2 & 0 \\ \vdots & \dots & \vdots \\ 0 & 0 & T_j \end{bmatrix},$$

com T_i é um bloco de tamanho $k_i \times k_i$:

$$T_i = \begin{bmatrix} c_{11} & c_{21} & c_{1k_i} \\ c_{21} & c_{22} & c_{2k_i} \\ \vdots & \vdots & \vdots \\ c_{k_i 1} & c_{k_i 2} & c_{k_i k_i} \end{bmatrix}.$$

Portanto, se V se dividir como uma soma direta de subespaços invariantes por T, então podemos encontrar uma base de V de modo que a matriz de T seja diagonal por bloco. Da mesma forma, se uma transformação linear tiver uma forma diagonal de bloco em alguma base, então V será dividido como uma soma de subespaços invariantes.

3.5.2 Projeções

3.5 Definição Projeção

Seja $V = V_1 \oplus \ldots \oplus V_n$, então para todo i existe um operador de projeção natural P_i: $V \to V_i$:

$$P_i(\boldsymbol{v}) = P_i\left(\sum_{j=1}^{n} \boldsymbol{v}_j\right) = \boldsymbol{v}_i$$

P_i é denominado projeção de V em V_i ao longo do subespaço complementar $\oplus_{j \neq i} V_j$.

Várias propriedades desses operadores de projeção são facilmente verificadas.

3.6 Proposição

As projeções P_i associadas a uma decomposição de soma direta $V = V_1 \oplus \ldots \oplus V_n$ têm as seguintes propriedades.

a) Linearidade: Cada $P_i : V \to V$ é um operador linear;

b) Propriedade Idempotente: $P_i^2 = P_i \circ P_i = P_i$ para todos os i;

c) $P_i \circ P_j = 0$ se $i \neq j$;

d $\operatorname{im}(P_i) = V_i$ e $\ker P_i = \sum_{j \neq i} V_j$;

e $P_1 + \ldots + P_n = \mathrm{I}_V$.

Se representarmos os vetores $v \in V$ como n-tuplas $(\boldsymbol{v}_1,\ \boldsymbol{v}_n)$ no conjunto de produtos cartesianos $V_1 \times \ldots \times V_n$, a i-ésima projeção assume a forma

$$P_i(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = (0,\ 0,\ \boldsymbol{v}_i,\ 0,\ 0) \in V_i \subseteq V.$$

Não se deixe enganar por esta notação ao pensar que estamos falando de projeções ortogonais nos subespaços ortogonais em $\mathbb{R}^n$.

3.7 Proposição

Seja $V = V_1 \oplus \ldots \oplus V_n$, então a projeção $P_i \colon V \to V_i$ pode ser escrito como a soma direta

$$P_i = 0_{V_1} \oplus 0_{V_2} \oplus \cdots \mathrm{I}_{V_i} \oplus \cdots \oplus 0_{V_n}$$

Se $\underline{\mathrm{B}}_i$ é base de V_1, $r_i = |B_i|$ e $\underline{\mathrm{B}}$ a base ordenada obtida pela concatenação $\underline{\mathrm{B}}_1 +\!\!\!+ \underline{\mathrm{B}}_2 +\!\!\!+ \cdots +\!\!\!+ \underline{\mathrm{B}}_n$ então

$$[P_i]_{\underline{\mathrm{B}},\underline{\mathrm{B}}} = \begin{bmatrix} 0_{r_1} & & & & \\ & \ddots & & & \\ & & I_{r_i} & & \\ & & & \ddots & \\ & & & & 0_{r_n} \end{bmatrix}$$

A propriedade de ser idempotente $P^2 = P$ para um operador linear é característico das projeções associadas a uma decomposição de soma direta $V = V_1 \oplus V_2$. Já vimos que, se $P, Q = (I - P)$ são as projeções associadas a essa decomposição, então

1. $P^2 = P$ e $Q^2 = Q$
2. $PQ = QP = 0$
3. $P + Q = I$ (operador identidade)

A recíproca também é verdadeira.

3.8 Proposição

Se $P : V \to V$, é um operador linear idempotente, i.e., $P^2 = P$, então V é soma direta $V = \operatorname{im} P \oplus \ker P$ e P é a projeção de V em $\operatorname{im} P$, ao longo de

$\ker P$. O operador $Q = \mathrm{I} - P$ também é idempotente, com

$$\operatorname{im} Q = \operatorname{im} \mathrm{I} - P = \ker P \text{ e } \ker Q = \ker(\mathrm{I} - P) = \operatorname{im} P,$$

e projeta V em $\operatorname{im} Q = \ker P$ ao longo de $\ker Q = \operatorname{im} P$.

Demonstração. Primeiro observe que $Q = \mathrm{I} - P$ também é idempotente, pois

$$(\mathrm{I} - P)^2 = I - 2P + P^2 = I - 2P + P = \mathrm{I} - P$$

Em seguida, observe que

$$\boldsymbol{v} \in \ker Q \Leftrightarrow Q\,\mathrm{I}\,\boldsymbol{v} = (\mathrm{I} - P)\boldsymbol{v} = 0 \Leftrightarrow P(\boldsymbol{v}) = \boldsymbol{v} \Leftrightarrow \boldsymbol{v} \in \operatorname{im} P\,.$$

Por outro lado, se $\boldsymbol{v} \in \operatorname{im} P$, então $\boldsymbol{v} = P\boldsymbol{w}$ para algum $\boldsymbol{w}$ e, logo $P\boldsymbol{v} = P^2(\boldsymbol{w}) = P\boldsymbol{w} = \boldsymbol{v}$, provando assim

- □ $\ker Q = \ker(I - P) = \operatorname{im} P$
- □ $\operatorname{im} Q = \operatorname{im}(I - P) = \ker P$,

Obviamente, $P + Q = I$ porque $\boldsymbol{v} = P\boldsymbol{v} + (I - P)\boldsymbol{v}$ implica que $P\boldsymbol{v} \in \operatorname{im} P$, enquanto $(I - P)\boldsymbol{v} \in \operatorname{im}(Q) = \ker P$. Portanto, $\operatorname{im} P + \ker P = \operatorname{im} P + \operatorname{im} Q$ é todo V. Além disso, $\ker P \cap \operatorname{im} P = \{\mathbf{0}\}$, pois se $\boldsymbol{v}$ estiver na intersecção, temos $\boldsymbol{v} \in \ker P \Rightarrow P\boldsymbol{v} = 0$. Mas também temos $v \in \operatorname{im} P$, então $\boldsymbol{v} = Pw$ para algum $\boldsymbol{w}$ e, assim

$$0 = P\boldsymbol{v} = P^2\boldsymbol{w} = P\boldsymbol{w} = \boldsymbol{v}.$$

E logo $V = \operatorname{im} P \oplus \ker P$.

■

3.6 Mudança de Corpo

Lembrando que um espaço vetorial V sobre $\mathbb{K}$ é um conjunto não vazio, com duas funções:

- □ $(\boldsymbol{x}, \boldsymbol{y}) \to \boldsymbol{x} + \boldsymbol{y} : V \times V \to V$
- □ $(\lambda_1, \boldsymbol{x}) \to \lambda_1 \boldsymbol{x} : \mathbb{K} \times V \to V$.

Então se V for um espaço complexo o produto por escalar é uma função $\mathbb{C} \times V \to V$ e podemos considerar suas restrição $\mathbb{R} \times V \to V$. Essa observação ingênua nos permite construir um espaço vetorial real associado a cada espaço vetorial complexo.

3.1 Definição Descomplexificação

Dado V um espaço vetorial sobre $\mathbb{C}$. Então V é um corpo vetorial sobre $\mathbb{R}$

se restringirmos a possibilidade de multiplicar os vetores em V apenas por escalares em $\mathbb{R}$. Denominaremos esse espaço de **descomplexificação** de V e denotaremos por $V_{\mathbb{R}}$.

Exemplo 2. *Seja $\mathbb{C}$ o espaço complexo unidimensional. Se considerarmos a sua descomplexificação $\mathbb{C}_{\mathbb{R}}$ obtemos um espaço vetorial real bidimensional de base* $\{1, i\}$.

$\triangleleft$

Podemos fazer algo similar com as transformações lineares:

3.3 Definição Descomplexificação de uma Transformação

Sejam V e W espaços lineares sobre $\mathbb{C}$, e $T : V \to W$ uma transformação linear. Podemos considerar T como uma transformação linear de $V_{\mathbb{R}} \to W_{\mathbb{R}}$. Essa transformação é denominada **descomplexificação** de T e é denotada por $T^{\mathbb{R}}$.

3.4 Teorema

(a) Se $\underline{\mathsf{E}} = \{\boldsymbol{e}_1, \ \dots \ \boldsymbol{e}_n\}$ é a base de um espaço V sobre $\mathbb{C}$. Então

$$\underline{\mathsf{E}}_{\mathbb{R}} \triangleq \{\boldsymbol{e}_1, \ \dots, \boldsymbol{e}_n, \ i\boldsymbol{e}_1, \ \dots, \ i\boldsymbol{e}_n\}$$

é uma base de $V_{\mathbb{R}}$ sobre $\mathbb{R}$. Em particular, $\dim_{\mathbb{R}} V_{\mathbb{R}} = 2\dim_{\mathbb{C}} V$.

(b) Se $A = B + iC$ é a matriz da transformação linear $T : V \to W$ nas bases $\underline{\mathsf{E}} = \{\boldsymbol{e}_1, \ \dots \ \boldsymbol{e}_m\}$ e $\underline{\mathsf{F}} = \{\boldsymbol{f}_1, \dots, \boldsymbol{f}_n\}$, sobre $\mathbb{C}$, onde B e C são matrizes reais. Então a matriz da transformação linear $T^{\mathbb{R}}$: $V_{\mathbb{R}} \to W_{\mathbb{R}}$ nas bases $\underline{\mathsf{E}}_{\mathbb{R}}$ e $\underline{\mathsf{F}}_{\mathbb{R}}$ é dada por

$$[T^{\mathbb{R}}]_{\underline{\mathsf{E}}_{\mathbb{R}}, \underline{\mathsf{F}}_{\mathbb{R}}} = \begin{bmatrix} B & -C \\ C & B \end{bmatrix}.$$

Demonstração.

(a) Para qualquer elemento $\boldsymbol{x} \in V$, temos

$$\boldsymbol{x} = \sum_{k=1}^{m} a_k \boldsymbol{e}_k = \sum_{k=1}^{n} (b_k + ic_k)\boldsymbol{e}_k = \sum_{k=1}^{n} b_k \boldsymbol{e}_k + \sum_{k=1}^{n} c_k i\boldsymbol{e}_k$$

onde b_k, c_k são as partes reais e imaginárias de a_k. Portanto, $\{\boldsymbol{e}_k, \ i\boldsymbol{e}_k\}$ gera $V_{\mathbb{R}}$.

Se $\sum_{k=1}^{m} b_k \boldsymbol{e}_k + \sum_{k=1}^{m} c_k(i\boldsymbol{e}_k) = 0$, em que $b_k, c_k \in \mathbb{R}$ então $b_k + ic_k = 0$ em virtude da independência linear de $\{\boldsymbol{e}_1,\ \dots\ \boldsymbol{e}_k\}$ sobre $\mathbb{C}$, e consequentemente $b_k = c_k = 0$ para todos os k.

$\boxed{b}$ A definição de A implica que

$$T(\boldsymbol{e}_1,\ \dots\ \boldsymbol{e}_m) = (B + iC)(\boldsymbol{f}_1,\ \dots\ \boldsymbol{f}_n)$$

de onde, por causa da linearidade de T sobre $\mathbb{C}$ temos

$$T(i\boldsymbol{e}_1,\ \dots\ i\boldsymbol{e}_m) = (-C + iB)(\boldsymbol{f}_1,\ \dots,\ \boldsymbol{f}_n)\ .$$

Dessa forma a matriz de T é

$$\begin{bmatrix} B & -C \\ C & B \end{bmatrix},$$

o que completa a prova. ■

3.5 Corolário

Seja $f : V \to V$ um operador linear sobre um espaço complexo finito dimensional V. Então $\det T^{\mathbb{R}} = |\det T|^2$.

Deve ser bastante óbvio agora como estender as definições anteriores. Seja $\mathbb{K}$ um corpo, $\mathbb{F}$ um subcorpo e V um espaço vetorial sobre $\mathbb{K}$. Restringindo a multiplicação a elementos de $\mathbb{F}$, obtemos o espaço vetorial $V_{\mathbb{F}}$ sobre $\mathbb{F}$. Analogamente, a transformação linear $T : V \to W$ sobre $\mathbb{K}$ se transforma na transformação linear $T^{\mathbb{F}} : V_{\mathbb{F}} \to W_{\mathbb{F}}$. O nome para essas operações é restrição do corpo de escalares $\mathbb{K}$ para $\mathbb{F}$.

O próprio corpo $\mathbb{K}$ também pode ser visto como um espaço vetorial sobre $\mathbb{F}$. E se for de dimensão finita, as dimensões $\dim_{\mathbb{K}} V$ e $\dim_{\mathbb{F}} V_{\mathbb{F}}$ serão relacionadas pela fórmula

$$\dim_{\mathbb{F}} V_{\mathbb{F}} = \dim_{\mathbb{F}} \mathbb{K} \cdot \dim_{\mathbb{K}} V.$$

Para a prova, basta verificar que $\{\boldsymbol{e}_1,\ \dots\ \boldsymbol{e}_n\}$ é uma base em V sobre $\mathbb{K}$ e $\{\boldsymbol{k}_1,\ \dots,\ \boldsymbol{k}_m\}$ é base de $\mathbb{K}$ sobre $\mathbb{F}$, então os vetores $\{\boldsymbol{k}_1\boldsymbol{e}_1,\ \dots,\ \boldsymbol{k}_1\boldsymbol{e}_n, \dots, \boldsymbol{k}_m\boldsymbol{e}_1,\ \dots,\ \boldsymbol{k}_m\boldsymbol{e}_n\}$ formam uma base de $V_{\mathbb{F}}$ sobre $\mathbb{F}$.

Complexificação. Dado um espaço vetorial real V introduzimos uma **estrutura complexa** J, definida pela fórmula

$$\mathrm{J}(\boldsymbol{x},\ \boldsymbol{y}) = (-\boldsymbol{y}, \boldsymbol{x})\ .$$

na soma direta externa $V \oplus V$. Observamos que $\mathrm{J}^2 = -\mathrm{I}$.

3.6 Definição

Seja $(V \oplus V,\ \mathrm{J})$ um espaço vetorial com estrutura complexa. Então V pode ser munido de uma operação de multiplicação por números complexos de acordo com o formula

$$(a+bi)\boldsymbol{x} \triangleq a(\boldsymbol{x}) + b\ \mathrm{J}(\boldsymbol{x})\ .$$

3.7 Teorema

O espaço $V \oplus V$ munido da adição usual de vetores e munido da operação de multiplicação por números complexos apresentada na Definição 3.6 é em um espaço vetorial complexo $V^{\mathbb{C}}$, para o qual $V^{\mathbb{C}}_{\mathbb{R}} = V \oplus V$.

Demonstração. Ambos os axiomas da distributividade são facilmente verificados a partir da linearidade de J e das fórmulas a adição de números complexos. Verificamos o axioma da associatividade para multiplicação:

$$\begin{aligned}
(a+bi)[(c+di)\boldsymbol{x}] &= (a+bi)[c\boldsymbol{x} + d\ \mathrm{J}(\boldsymbol{x})] && (3.28)\\
&= a[c\boldsymbol{x} + d\ \mathrm{J}(\boldsymbol{x})] + +b\ \mathrm{J}[c\boldsymbol{x} + d\ \mathrm{J}(\boldsymbol{x})] && (3.29)\\
&= ac\boldsymbol{x} + ad\ \mathrm{J}(\boldsymbol{x}) + bc\ \mathrm{J}(\boldsymbol{x}) - bd\boldsymbol{x} && (3.30)\\
&= (ac-bd)\boldsymbol{x} + (ad+bc)\ \mathrm{J}(\boldsymbol{x}) && (3.31)\\
&= [ac-bd+(ad+bc)i]\boldsymbol{x} && (3.32)\\
&= ((a+bi)(c+di)]\boldsymbol{x}. && (3.33)
\end{aligned}$$

■

3.8 Definição Complexificação

O espaço $V^{\mathbb{C}}$ é dito **complexificação** do espaço V.

Ressaltamos que $V^{\mathbb{C}} = V \oplus V$ e que essa decomposição como soma direta ocorre sobre $\mathbb{R}$, mas não sobre $\mathbb{C}$.

Identificando V com o subconjunto de vetores do formato $(\boldsymbol{v}, \boldsymbol{0})$ em $V^{\mathbb{C}}$ e usando o fato de que $i(\boldsymbol{v},\ 0) = \mathrm{J}(\boldsymbol{v}, 0) = (0,\ \boldsymbol{v})$, podemos escrever qualquer vetor de $V^{\mathbb{C}}$ no formato

$$(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, 0) + (0, \boldsymbol{y}) = (\boldsymbol{x}, 0) + i(\boldsymbol{y}, 0) = \boldsymbol{x} + i\boldsymbol{y}.$$

Dessa forma qualquer base de V sobre $\mathbb{R}$ será uma base de $V^{\mathbb{C}}$ sobre $\mathbb{C}$, de modo que $\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} V^{\mathbb{C}}$.

3.9 Definição

Seja $T : V \to M$ um transformação linear de espaços vetoriais sobre $\mathbb{R}$. Definimos a transformação $T^{\mathbb{C}} : V^{\mathbb{C}} \to W^{\mathbb{C}}$ denominada **complexificação da transformação** T como

$$T(\boldsymbol{x},\ \boldsymbol{y}) = (T(\boldsymbol{x}),\ T(\boldsymbol{y}))$$

A transformação $T^{\mathbb{C}} : V^{\mathbb{C}} \to W^{\mathbb{C}}$ é linear sobre $\mathbb{R}$ e comuta com J, pois

$$T\,\mathrm{J}(\boldsymbol{x},\ \boldsymbol{y}) = T(-\boldsymbol{y},\ \boldsymbol{x}) = (-T(\boldsymbol{y}),\ T(\boldsymbol{x})) = \mathrm{J}\,T(\boldsymbol{x},\boldsymbol{y})\ .$$

Portanto, é linear sobre $\mathbb{C}$.

3.10 Proposição

Sejam $\underline{\mathrm{v}}$ e $\underline{\mathrm{w}}$ bases de V e W respectivamente. Usando a identificação $\boldsymbol{v} \mapsto (\boldsymbol{v}, \boldsymbol{0})$ podemos ver $\underline{\mathrm{v}}$ e $\underline{\mathrm{w}}$ como bases de $V^{\mathbb{C}}$ e $W^{\mathbb{C}}$, respectivamente. Temos que que a matriz de T em relação as bases $\underline{\mathrm{v}}$ e $\underline{\mathrm{w}}$ é igual à matriz de $T^{\mathbb{C}}$ nessas mesmas bases:

$$[T]_{\underline{\mathrm{v}},\underline{\mathrm{w}}} = [T^{\mathbb{C}}]_{\underline{\mathrm{v}},\underline{\mathrm{w}}}$$

3.11 Definição

Uma conjugação em um espaço vetorial complexo V é uma função $c : V \to V$ tal que

- *a* $c(\boldsymbol{u} + \boldsymbol{v}) = c(\boldsymbol{u}) + c(\boldsymbol{v})$ para todos os $\boldsymbol{u}, \boldsymbol{v} \in V$,
- *b* $c(z\boldsymbol{v}) = zc(\boldsymbol{v})$ para todos os $\boldsymbol{z} \in \mathbb{C}$ e $v \in V$,
- *c* $c(c(\boldsymbol{v})) = v$ para todos os $\boldsymbol{v} \in V$.

Uma conjugação é linear sobre $\mathbb{R}$, mas não é linear sobre $\mathbb{C}$:

$$c(i\boldsymbol{v}) = -ic(\boldsymbol{v}).$$

Exemplo 12. *Em $\mathbb{C}^n$, a função $c(z_1, \ldots, z_n) = (\overline{z}_1, \ldots, \overline{z}_n)$, onde cada coordenada é substituída por seu complexo conjugado, é uma conjugação*

$\triangleleft$

Exemplo 13. *Em $V^{\mathbb{C}} = V \oplus V$ a função que leva $(\boldsymbol{u}, \boldsymbol{v}) \mapsto (\boldsymbol{u}, \boldsymbol{v})$ é uma conjugação.*

$\triangleleft$

Exercícios

Ex. 3.41 --- Mostre que as complexificações de $\mathbb{R}^n, \mathcal{M}_{n,n}(\mathbb{R})$ são naturalmente isomorfas a $\mathbb{C}^n, \mathcal{M}_{n,n}(\mathbb{C})$.

Ex. 3.42 --- Mostre que $\mathbb{R}[x]^{\mathbb{C}}$ é naturalmente isomorfo a $\mathbb{C}[x]$.

Ex. 3.43 --- Mostre que se $W = \mathbf{0}$, então, $W_{\mathbb{C}} = 0$. Se $W \neq 0$ e $\{\boldsymbol{e}_j\}$ é uma $\mathbb{R}$-base de W, então $\{(ej, 0)\}$ é uma $\mathbb{C}$-base de $W_{\mathbb{C}}$. Em particular,

$$\dim_{\mathbb{C}}(W_{\mathbb{C}}) = \dim_{\mathbb{R}}(W)$$

para todo W.

Ex. 3.44 --- Se $\phi : W \to W$ é $\mathbb{R}$-linear, sua complexificação$\phi : C : W_{\mathbb{C}} \to W_{\mathbb{C}}$ tem kernel e imagem

$$\ker(\phi_{\mathbb{C}}) = (\ker \phi)_{\mathbb{C}}, \operatorname{im}(\phi_{\mathbb{C}}) = (\operatorname{im} \phi)_{\mathbb{C}}$$

Ex. 3.45 --- Se V é um espaço vetorial complexo, uma $\mathbb{R}$-forma de V é um subespaço real W de V tendo V como sua complexificação. Mostre que duas $\mathbb{R}$-formas de $\mathcal{M}_{n,n}(\mathbb{C})$ são:

$$\mathcal{M}_{n,n}(\mathbb{R}) \text{ e } \left\{ \left(\begin{matrix} a & b+ci \\ b-ci & d \end{matrix} \right] : a,b,c,d \in \mathbb{R} \right\}.$$

Capítulo 4

Espaços Quocientes

Nesse capítulo

- Espaços Afins (p. 128)
- Espaço Quociente (p. 132)
- Teoremas de Isomorfismo (p. 136)

4.1 Espaços Afins

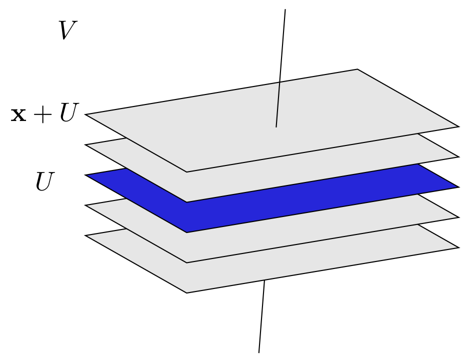
Começamos com uma generalização da noção de subespaço de um espaço vetorial: os subespaços afins. Um subespaço afim de um espaço vetorial é uma translação de um subespaço vetorial.

4.1 Definição Subespaço Afim

Um subconjunto X de V é um **subespaço afim** de V paralelo ao subespaço $U \subset V$ se, e somente se, para algum elemento x de X,

$$X = \boldsymbol{x} + U = \{\boldsymbol{x} + \boldsymbol{u} | \boldsymbol{u} \in U\}.$$

Seja X um subespaço afim de V paralelo ao subespaço U. Então a dimensão de X é $\dim(X) \triangleq \dim(U)$.



4.2 Proposição

Seja X um subespaço afim de V paralelo ao subespaço U de V. Sejam $\boldsymbol{x}_0$

um elemento de X e $\{\boldsymbol{u}_1, \ \boldsymbol{u}_2, \ \ldots \boldsymbol{u}_n\}$ é uma base de U. Então, qualquer elemento $\boldsymbol{x}$ de X pode ser escrito de maneira única como

$$\boldsymbol{x} = x_0 + \sum c_i \boldsymbol{u}_i$$

para $c_i \in \mathbb{K}$.

A demonstração dessa proposição é direta. Se $\boldsymbol{v} \in X$ então $\boldsymbol{v} = \boldsymbol{x}_0 + \boldsymbol{v}$. Como $\boldsymbol{v} \in V$ e $\{\boldsymbol{u}_1, \ \boldsymbol{u}_2, \ \ldots \boldsymbol{u}_n\}$ é base de V temos que $\boldsymbol{v} = \sum c_i \boldsymbol{u}_i$ e logo

$$\boldsymbol{x} = x_0 + \sum c_i \boldsymbol{u}_i.$$

A demonstração da unicidade dessa representação será deixada como exercício.

Exemplo 3. *O conjunto solução de um sistema linear*

$$C = \{\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b}\}$$

é um subespaço afim, com $A \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$.

$\triangleleft$

4.4 Definição

O conjunto de todos os subespaços afins de V será denotado $\mathrm{Afim}[V]$. E o conjunto de todos os subespaços afins de V paralelos ao subespaço U será denotado $\mathrm{Afim}[V, U]$.

Assim, $A \in \mathrm{Afim}[V]$ se, e somente se, $A = \boldsymbol{x} + W$ para algum subespaço $W \subseteq V$ e algum $\boldsymbol{x} \in V$.

4.5 Teorema

Seja V um espaço vetorial sobre $\mathbb{K}$ e $\mathrm{Afim}[V]$ denote o conjunto de todos os subespaços afins de V.

- a Se $\{A_i \,|\, i \in \Delta\}$ for uma coleção indexada de subespaços, então $\bigcap_{i\in\Delta} A_i = \emptyset$ ou $\bigcap_{i\in\Delta} A_i \in \mathrm{Afim}[V]$.
- b Se $A, B \in \mathrm{Afim}[V]$ então $A + B \in \mathrm{Afim}[V]$.
- c Se $A \in \mathrm{Afim}[V]$ e $\lambda_1 \in \mathbb{K}$, então $\lambda_1 A \in \mathrm{Afim}[V]$.
- d Se $A \in \mathrm{Afim}[V]$ e $\boldsymbol{x} \in V$, então $\boldsymbol{x} + A \in \mathrm{Afim}[V]$.
- e Se $A \in \mathrm{Afim}[V]$ e $T \in \mathrm{Hom}(V, V')$ então $T(A) \in \mathrm{Afim}[V']$.

Demonstração. Vamos demonstrar $\boxed{a}$ apenas. Suponha que $A_i = \boldsymbol{x}_i + W_i$ para todo $i \in \Delta$. Aqui W_i é um subespaço de V e $\boldsymbol{x}_i$ um vetor em V. Suponha que $\bigcap_{i\in\Delta} A_i \neq \emptyset$. Seja $\boldsymbol{y} \in \bigcap_{i\in\Delta} A_i$. Então, para todo $i \in \Delta, \boldsymbol{y} = \boldsymbol{x}_i + \boldsymbol{z}_i$ com $\boldsymbol{z}_i \in W_i$. Então $\boldsymbol{y} + W_i = \boldsymbol{x}_i + W_i$ e $\bigcap_{i\in\Delta} A_i = \bigcap_{i\in\Delta}(\boldsymbol{y} + W_i)$.

Afirmamos que $\bigcap_{i\in\Delta}(\boldsymbol{y} + W_i) = \boldsymbol{y} + (\bigcap_{i\in\Delta} W_i)$. Claramente,

$$\boldsymbol{y} + (\bigcap_{i\in\Delta} W_i) \subseteq \bigcap_{i\in\Delta}(\boldsymbol{y} + W_i),$$

Para a inclusão inversa, seja $\boldsymbol{x} \in \bigcap_{i\in\Delta}(\boldsymbol{y} + W_i)$. Então, para $i \neq j, \boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}_i = \boldsymbol{y} + \boldsymbol{z}_j$ com $\boldsymbol{z}_i \in W_i$ e $\boldsymbol{z}_j \in W_j$. Logo $\boldsymbol{z}_i = \boldsymbol{z}_j$ e $\boldsymbol{x} \in \boldsymbol{y} + (W_i \cap W_j)$. Assim, $\bigcap_{i\in\Delta}(\boldsymbol{y} + W_i) \subseteq \boldsymbol{y} + (\bigcap_{i\in\Delta} W_i)$. Portanto, $\bigcap_{i\in\Delta}(\boldsymbol{y} + W_i) = \boldsymbol{y} + (\bigcap_{i\in\Delta} W_i)$. Como $\boldsymbol{y} + (\bigcap_{i\in\Delta} W_i) \in \text{Afim}[V]$, terminamos a demonstração de $\boxed{a}$. ■

De modo análogo, para o conjunto de todos os subespaços afins de V paralelos ao subespaço U temos:

4.6 Teorema

Seja V um espaço vetorial sobre $\mathbb{K}$, U um subespaço de V e $\text{Afim}[V, U]$ o conjunto de todos os subespaços afins de V paralelos a U.

- $\boxed{a}$ Se $A, B \in \text{Afim}[V, U]$ então $A + B \in \text{Afim}[V, U]$.
- $\boxed{b}$ Se $A \in \text{Afim}[V, U]$ e $\lambda_1 \in \mathbb{K}$, então $\lambda_1 A \in \text{Afim}[V, U]$.
- $\boxed{c}$ Se $A \in \text{Afim}[V, U]$ e $\boldsymbol{x} \in V$, então $\boldsymbol{x} + A \in \text{Afim}[V, U]$. []

A maneira mais importante pela qual surgem subespaços afins é a seguinte.

4.7 Teorema Imagem Inversa

Seja $T\colon V \to W$ uma transformação linear e $w_0 \in W$ seja um elemento arbitrário de W. Se $T^{-1}(\boldsymbol{w}_0)$ não for vazio, então $T^{-1}(\boldsymbol{w}_0)$ é um subespaço afim de V paralelo a $\ker T$.

Demonstração. Escolha $\boldsymbol{v}_0 \in V$ com $T(\boldsymbol{v}_0) = \boldsymbol{w}_0$. Se $\boldsymbol{v} \in T^{-1}(\boldsymbol{w}_0)$ for arbitrário, $\boldsymbol{v} = \boldsymbol{v}_0 + (\boldsymbol{v} - \boldsymbol{v}_0) = \boldsymbol{v}_0 + \boldsymbol{u}$ e $T(\boldsymbol{u}) = T(\boldsymbol{v} - \boldsymbol{v}_0) = T(\boldsymbol{v}) - T(\boldsymbol{v}_0) = \boldsymbol{w}_0 - \boldsymbol{w}_0 = 0$, então $\boldsymbol{u} \in \ker(T)$. Por outro lado, se $\boldsymbol{u} \in \ker(T)$ e $\boldsymbol{v} = \boldsymbol{v}_0 + \boldsymbol{u}$, então $T(\boldsymbol{v}) = T(\boldsymbol{v}_0 + \boldsymbol{u}) = T(\boldsymbol{v}_0) + T(\boldsymbol{u}) = \boldsymbol{w}_0 + 0 = \boldsymbol{w}_0$. Assim, vemos que

$$T^{-1}(\boldsymbol{w}_0) = \boldsymbol{v}_0 + \ker T.$$

■

Podemos generalizar o Teorema 4.1 introduzindo o conceito de uma transformação afim entre dois espaços vetoriais. Se $\boldsymbol{x} \in V$, depois pela translação através de $\boldsymbol{x}$, entenderemos a função $S_{\boldsymbol{x}} : V \to V$ fornecido por $S_{\boldsymbol{x}}(\boldsymbol{y}) = \boldsymbol{x}+\boldsymbol{y}$. Qualquer classe $\boldsymbol{x} + W$ é apenas $S_{\boldsymbol{x}}(W)$ f para a translação $S_{\boldsymbol{x}}$. Observe que quando $\boldsymbol{x} \neq 0, S_{\boldsymbol{x}}$ não é uma transformação linear.

4.8 **Definição**

Sejam V e V' dois espaços vetoriais sobre um corpo $\mathbb{K}$. Uma função $\mathbb{K}$: $V \to V'$ é denominada **transformação afim** se $f = S_{\boldsymbol{x}}T$ para alguma $T \in \operatorname{Hom}_{\mathbb{K}}(V, V')$ e alguns $\boldsymbol{x} \in V'$. O conjunto de todas as transformações afins de V para V 'será denotado como $\operatorname{Afim}_{\mathbb{K}}(V, V')$

É imediato observar que $\operatorname{Hom}_{\mathbb{K}}(V, V') \subseteq \operatorname{Afim}_{\mathbb{K}}[V, V'] \subseteq (V')^V$. O Teorema 4.1 $\boxed{d}$ pode ser reformulado da seguinte forma:

4.9 **Teorema**

Se $A \in \operatorname{Afim}[V]$ e $f \in \operatorname{Afim}_{\mathbb{K}}[V, V']$ então $f(A) \in \operatorname{Afim}[V']$.

Agora, definimos a noção importante do espaço vetorial quociente V/W e investigamos algumas de suas propriedades.

Exercícios

Ex. 4.1 --- Mostre que um subconjunto X de V é um subespaço afim de V paralelo ao subespaço $U \subset V$ se, e somente se, for para alguns e, portanto, para todo elemento x de X,

$$X = \boldsymbol{x} + U = \{\boldsymbol{x} + \boldsymbol{u} | \boldsymbol{u} \in U\}.$$

Ex. 4.2 --- Seja X um subespaço afim de V paralelo ao subespaço U de V. Seja $\boldsymbol{x}_0$ um elemento de X e seja $\{\boldsymbol{u}_1, \ \boldsymbol{u}_2, \ \dots \boldsymbol{u}_n\}$ é uma base de U. Então, qualquer elemento $\boldsymbol{x}$ de X pode ser escrito de maneira única como

$$\boldsymbol{x} = x_0 + \sum c_i \boldsymbol{u}_i$$

para $c_i \in \mathbb{K}$.

Ex. 4.3 --- Mostre que

1. Se $A \in \operatorname{Afim}[V]$ e $\lambda_1 \in \mathbb{K}$, $\boldsymbol{x} + A \in \operatorname{Afim}[V]$.

2. Se $A \in \operatorname{Afim}[V]$ e $T \in \operatorname{Hom}(V, V)$ e $T(A) \in \operatorname{Afim}[V']$.

3. Se $A' e \operatorname{Afim}[V']$ e $T \in \operatorname{Hom}(V, V$ então $T^{-1}(A')$ é vazio ou é um subespaço afim de V.

Ex. 4.4 --- O espaço afim gerado pelos vetores $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r$ em V é o conjunto

$$\{\lambda_1 \boldsymbol{v}_+ \ldots + \lambda_r \boldsymbol{v}_r | \lambda_i \in \mathbb{K}, \lambda_1 + \ldots + \lambda_r = 1\}$$

. Prove que é um subespaço afim.

4.2 Espaço Quociente

Seja S um subespaço de um espaço vetorial V. É fácil ver que a relação binária em V definida por

$$\boldsymbol{u} \equiv \boldsymbol{v} \;\Leftrightarrow\; \boldsymbol{u} - \boldsymbol{v} \in S$$

é uma relação de equivalência. Quando $\boldsymbol{u} \equiv \boldsymbol{v}$, dizemos que $\boldsymbol{u}$ e $\boldsymbol{v}$ são congruentes módulo S e $\boldsymbol{u} \equiv \boldsymbol{v}$ é frequentemente escrito

$$\boldsymbol{u} \equiv \boldsymbol{v} \mod S$$

Quando o subespaço em questão estiver claro, escreveremos simplesmente $\boldsymbol{u} \equiv \boldsymbol{v}$.

Para ver como são as classes de equivalência, observe que

$$\begin{aligned}
[\boldsymbol{v}] &= \{\boldsymbol{u} \in V \,|\, \boldsymbol{u} \equiv \boldsymbol{v}\} && (4.1)\\
&= \{\boldsymbol{u} \in V \,|\, \boldsymbol{u} - \boldsymbol{v} \in S\} && (4.2)\\
&= \{\boldsymbol{u} \in V \,|\, \boldsymbol{u} = \boldsymbol{v} + \boldsymbol{s} \text{ para algum } \boldsymbol{s} \in S\} && (4.3)\\
&= \{\boldsymbol{v} + \boldsymbol{s} \,|\, \boldsymbol{s} \in S\} && (4.4)\\
&= \boldsymbol{v} + S && (4.5)
\end{aligned}$$

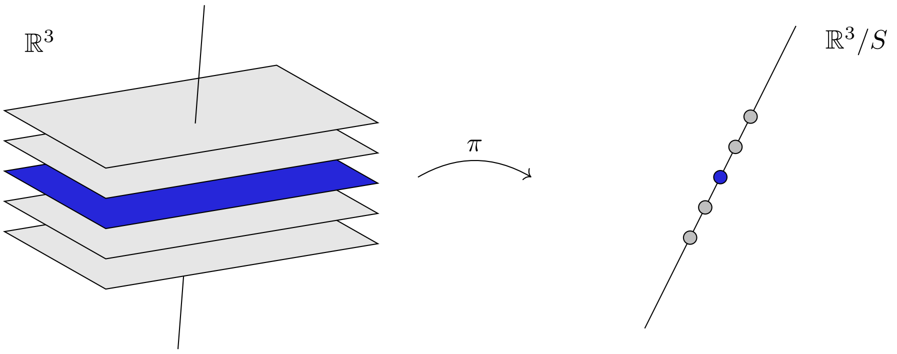
O conjunto

$$[\boldsymbol{v}] = \boldsymbol{v} + S = \{\boldsymbol{v} + \boldsymbol{s} \,|\, \boldsymbol{s} \in S\}$$

é denominado **classe** de S em V e $\boldsymbol{v}$ é denominado **representante** da classe $\boldsymbol{v}+S$.

4.1 Definição Espaço Quociente

O conjunto de todas as classes de S em V é denotado por $V/S = \{\boldsymbol{v}+S \,|\, \boldsymbol{v} \in V\}$ é denominado **espaço quociente** de V modulo S.



Podemos definir uma estrutura de espaço vetorial em V/S.

A escolha natural para essas operações de espaço vetorial é

$$(\boldsymbol{u}+S)+(\boldsymbol{v}+S) \triangleq (\boldsymbol{u}+\boldsymbol{v})+S \quad \text{e} \quad \lambda(\boldsymbol{u}+S) \triangleq (\lambda\boldsymbol{u})+S$$

mas devemos verificar se essas operações estão bem definidas, ou seja,

- ☐ $\boldsymbol{u}_1+S=\boldsymbol{u}_2+S, \boldsymbol{v}_1+S=\boldsymbol{v}_2+S \Rightarrow (\boldsymbol{u}_1+\boldsymbol{v}_1)+S=(\boldsymbol{u}_2+\boldsymbol{v}_2)+S$
- ☐ $\boldsymbol{u}_1+S=\boldsymbol{u}_2+S \Rightarrow r\boldsymbol{u}_1+S=r\boldsymbol{u}_2+S$

De maneira equivalente, a relação de equivalência $\equiv$ deve ser consistente com as operações de espaço vetorial em V, ou seja,

- ☐ $\boldsymbol{u}_1\equiv\boldsymbol{u}_2, \boldsymbol{v}_1\equiv\boldsymbol{v}_2 \Rightarrow (\boldsymbol{u}_1+\boldsymbol{v}_1)\equiv(\boldsymbol{u}_2+\boldsymbol{v}_2)$
- ☐ $\boldsymbol{u}_1\equiv\boldsymbol{u}_2 \Rightarrow \lambda\boldsymbol{u}_1\equiv\lambda\boldsymbol{u}_2$

Esse cenário é recorrente na álgebra. Uma relação de equivalência em uma estrutura algébrica, como um grupo, anel, módulo ou espaço vetorial é denominada relação de congruência se preservar as operações algébricas.

Essas condições decorrem facilmente do fato de que S é um subespaço, pois se $\boldsymbol{u}_1\equiv\boldsymbol{u}_2$ e $\boldsymbol{v}_1\equiv\boldsymbol{v}_2$, então $\boldsymbol{u}_1-\boldsymbol{u}_2\in S,\ \boldsymbol{v}_1-\boldsymbol{v}_2\in S \Rightarrow \lambda_1(\boldsymbol{u}_1-\boldsymbol{u}_2)+\lambda_2(\boldsymbol{v}_1-\boldsymbol{v}_2)\in S$ $\Rightarrow (\lambda_1\boldsymbol{u}_1+\lambda_2\boldsymbol{v}_1)-(\lambda_1\boldsymbol{u}_2+\lambda_2\boldsymbol{v}_2)\in S \Rightarrow \lambda_1\boldsymbol{u}_1+\lambda_2\boldsymbol{v}_1\equiv\lambda_1\boldsymbol{u}_2+\lambda_2\boldsymbol{v}_2$ que verifica as duas condições ao mesmo tempo. Deixamos ao leitor verificar se V/S é realmente um espaço vetorial acima de $\mathbb{K}$ nessas operações bem definidas.

4.2 Teorema

Seja S um subespaço de V. A relação binária $\boldsymbol{u}\equiv\boldsymbol{v} \;\Leftrightarrow\; \boldsymbol{u}-\boldsymbol{v}\in S$ é uma relação de equivalência em V, cujas classes de equivalência são os espaços afins paralelos a S $\boldsymbol{v}+S=\{\boldsymbol{v}+\boldsymbol{s}\,|\,\boldsymbol{s}\in S\}$ de S em V.

O conjunto V/S de todos as classes de S em V, chamado **espaço quociente** de V modulo S, é um espaço vetorial com operações

$$\lambda(\boldsymbol{u}+S)=\lambda\boldsymbol{u}+S \tag{4.6}$$

$$(\boldsymbol{u}+S)+(\boldsymbol{v}+S)=(\boldsymbol{u}+\boldsymbol{v})+S \tag{4.7}$$

O vetor zero em V/S é a classe $\mathbf{0} + S = S$.

4.3 Definição Projeção Canônica

Se S é um subespaço de V, podemos definir a aplicação $\pi_S : V \to V/S$ enviando cada vetor a classe lateral associada a ele: $\pi_S(v) = v + S$ Essa aplicação é denominada **projeção canônica** ou **projeção natural** de V para V/S.

Quando não houver risco de confusão escreveremos simplesmente π

4.4 Proposição

A projeção canônica $\pi_S : V \to V/S$ definida por

$$\pi_S(v) = v + S$$

é uma transformação linear sobrejetiva com $\ker(\pi_S) = S$.

Demonstração. A projeção canônica é linear, pois

$$\pi(\lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y}) = (\lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y}) + S = \lambda_1(\boldsymbol{x} + S) + \lambda_2(\boldsymbol{y} + S) = \lambda_1 \pi \boldsymbol{x} + \lambda_1 \pi \boldsymbol{y}$$

A projeção canônica é claramente sobrejetiva, pela definição de V/S.

Para determinar o núcleo de π, observe que

$$\boldsymbol{v} \in \ker \pi \Leftrightarrow \pi \boldsymbol{v} = \mathbf{0} \Leftrightarrow \boldsymbol{v} + S = S \Leftrightarrow \boldsymbol{v} \in S$$

e logo $\ker \pi = S$.

■

4.5 Teorema Teorema da Correspondência

Seja S um subespaço de V. Então a função que atribui a cada subespaço $S \subseteq T \subseteq V$ o subespaço T/S de V/S é uma correspondência biunívoca que preserva a ordem entre entre o conjunto de todos os subespaços de V contendo S e o conjunto de todos os subespaços de V/S.

Demonstração. Provaremos apenas que a correspondência é sobrejetiva. Para tanto, seja

$$X = \{\boldsymbol{u} + S | \boldsymbol{u} \in U\}$$

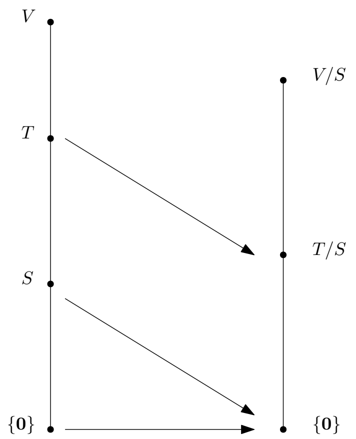


Figura 4.1 Correspondência entre os subespaços de V e V/S.

um subespaço em V/S e seja T a união de todos as classes em X:

$$T = \bigcup_{\boldsymbol{u} \in U} (\boldsymbol{u} + S)$$

Mostraremos que $S \leq T \leq V$ e que $T/S = X$. Se $\boldsymbol{x}, \boldsymbol{y} \in T$, $\boldsymbol{x} + S$ e $\boldsymbol{y} + S$ estão em X e, como X é subespaço de V/S, temos

$$\lambda \boldsymbol{x} + S,\ (\boldsymbol{x} + \boldsymbol{y}) + S \in X$$

o que implica que $\lambda \boldsymbol{x}, \boldsymbol{x} + \boldsymbol{y} \in T$. Portanto, T é um subespaço de V contendo S. Além disso, se $\boldsymbol{v} + S \in T/S$, então $\boldsymbol{v} \in T$ e, portanto, $\boldsymbol{v} + S \in X$.

Por outro lado, se $\boldsymbol{u}+S \in X$, $\boldsymbol{u} \in T$ e, portanto, $\boldsymbol{u}+S \in T/S$. Assim, $X = T/S$. ■

Exercícios

Ex. 4.5 --- Seja V o subespaço de $\mathbb{R}^2$ dado por $V = \langle\{(1,\ -1)\}\rangle$.

1. Desenhe V em $\mathbb{R}^2$.

2. Desenhe as classes

$$S_1 = (1, 1) + \mathrm{V},\ S_2 =\ (2,\ 1) + \mathrm{V}$$

3. Descreva o espaço quociente $\mathbb{R}^2/V$.

4. Desenhe a classe $S_3 = (-2)S_1 + S_2$.

5. Determine se S_3 é igual a $(-1, 0) + \mathrm{V}$.

Ex. 4.6 --- Seja V o subespaço de $\mathbb{K}[x]_2$ satisfazendo

$$\int_{-1}^{1} p(t)\mathrm{d}t = 0,\ p(t) \in \mathbb{K}[x]_2.$$

1. Determine uma base de V.

2. Determine uma base do espaço quociente $\mathbb{K}[x]_2/V$.

Ex. 4.7 --- Descreva o espaço quociente $\mathbb{K}^\infty/U$, onde $U = \{(0, a_2, a_3, \dots) : a_i \in \mathbb{K}\}$.

Ex. 4.8 --- Se V é de dimensão infinita e S é um subespaço de dimensão infinita, a dimensão de V/S deve ser finita? Prove ou dê contra-exemplo.

Ex. 4.9 --- Prove o teorema da correspondência.

Ex. 4.10 --- Seja S um subespaço de V. Começando com uma base $\{s_1,\ \dots,\ s_k\}$ para S, como você encontraria uma base para V/S?

Ex. 4.11 --- Sejam U um espaço vetorial e V e W subespaços de U. Dado $\boldsymbol{u} \in U$, denote por $[\boldsymbol{u}]_V$ e $[\boldsymbol{u}]_W$ as classes laterais $\boldsymbol{u} + V$ e $\boldsymbol{u} + W$, respectivamente.

Suponha que $V \subset W$ e $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k \in U$. Então mostre que se $[u_1]_V,\ ldots, [u_k]_V$ são linearmente dependentes temos que $[u_1]_W, \dots, [u_k]_W$ são linearmente dependentes.

Em particular, se U é finito dimensional, $\dim(U/V) \geq \dim(U/W)$, se $V \subset W$.

Ex. 4.12 --- Suponha que W é um subespaço de dimensão finita de V e que V/W é finito dimensional. Mostre que V deve ser finito dimensional.

4.3 Teoremas de Isomorfismo

Apresentaremos agora os Teoremas de Isomorfismo. Esses teoremas aparecem de múltiplas formas em diversas áreas da matemática e são muito úteis.

4.1 Teorema Propriedade Universal do Quociente

Sejam S um subespaço de V e $T \in \mathrm{Hom}(V,\ W)$ satisfazendo $S \subseteq \ker(T)$. Então, existe uma única transformação linear $\overline{T} : V/S \to W$ que faz o seguinte diagrama comutativo

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\pi"'] & W \\ V/S \arrow[ur, "\overline{T}"'] & \end{tikzcd},$$

ou seja, com a propriedade que $\overline{T} \circ \pi_S = T$.

Além disso, $\ker(\overline{T}) = \ker T/S$ e $\mathrm{im}(\overline{T}) = \mathrm{im}\, T$.

Demonstração. Não temos outra opção senão definir $\overline{T}$ pela condição $\overline{T} \circ \pi_S = T$, ou seja, $\overline{T}(\boldsymbol{v} + S) = T\boldsymbol{v}$. Esta função está bem definida se, e somente se, $\boldsymbol{v} + S = \boldsymbol{u} + S$ implicar que $\overline{T}(\boldsymbol{v} + S) = \overline{T}(\boldsymbol{u} + S)$ que é equivalente a cada uma das seguintes afirmações:

$$\begin{aligned} \boldsymbol{v} + S = \boldsymbol{u} + S &\Rightarrow T\boldsymbol{v} = T\boldsymbol{u} \\ \boldsymbol{v} - \boldsymbol{u} \in S &\Rightarrow T(\boldsymbol{v} - \boldsymbol{u}) = 0 \\ \boldsymbol{x} \in S &\Rightarrow T\boldsymbol{x} = 0 \\ S &\subseteq \ker T \end{aligned}$$

Assim, $\overline{T} : V/S \to W$ está bem definida. Além disso,

$$\mathrm{im}(\overline{T}) = \{\overline{T}(\boldsymbol{v} + S) | v \in V\} = \{T\boldsymbol{v} | \boldsymbol{v} \in V\} = \mathrm{im}\, T \quad \text{e}$$

$$\begin{aligned} \ker(\overline{T}) &= \{\boldsymbol{v} + S | \overline{T}(\boldsymbol{v} + S) = 0\} \\ &= \{\boldsymbol{v} + S | T\boldsymbol{v} = 0\} \\ &= \ker T/S \end{aligned}$$

A unicidade de $\overline{T}$ é direta. Se $T' \in \mathrm{Hom}(V/W, V')$ é outra transformação tal que $T'\pi = T$ e $\overline{T} = T'$ em $\mathrm{im}\, H$. Mas π é sobrejetiva. Portanto, $T = T'$. ■

O teorema 4.3 possui um corolário muito importante, que é frequentemente denominado primeiro teorema do isomorfismo e é obtido tomando $S = \ker(T)$.

4.2 Teorema Primeiro Teorema de Isomorfismo

Suponha que $T \in \mathrm{Hom}(V, V')$. Então $\mathrm{im}\, T \cong V/\ker T$.

Demonstração. Podemos ver T como uma transformação linear sobrejetiva de V para $\mathrm{im}\, T$. Aplicando o Teorema 4.3, obtemos uma única transformação linear $\overline{T} : V/\ker T \to \mathrm{im}\, T$ de modo que o diagrama a seguir é comutativo:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\pi"'] & \mathrm{im}\, T \\ V/\ker T \arrow[ur, "\overline{T}"'] & \end{tikzcd} \tag{4.8}$$

Claramente $\overline{T}$ é um isomorfismo. ■

4.3 Proposição

Existe uma sequência exata $0 \longrightarrow X \xrightarrow{A} Y \xrightarrow{B} Z \longrightarrow 0$ se, e somente se, $Y/X \cong Z$

Demonstração. Se $Y/X \cong Z$. Então $\dim Z = \dim Y - \dim X$ e logo $\dim Y \geq \dim Z$. Logo existe $B : Y \to Z$ uma aplicação sobrejetiva. Os espaços $\ker B$ e X são isomorfos e seja A uma aplicação injetiva que leva X em $\ker B \subset Y$. Então a sequência

$$0 \longrightarrow X \xrightarrow{A} Y \xrightarrow{B} Z \longrightarrow 0$$

é exata.

Se a sequência é exata então a aplicação $B : Y \to Z$ é sobrejetiva, e a aplicação $A : X \to Y$ é injetiva com $\operatorname{im}(A) = \ker B$. Assim, a aplicação $A : X \to \ker B$ é um isomorfismo.

Logo, pelo Primeiro Teorema do Isomorfismo $Z = \operatorname{im} B \cong Y/\ker B \cong Y/X$.

■

O Segundo Teorema do Isomorfismo lida com múltiplos quocientes. Suponha que W seja um subespaço de V e considere a projeção natural $\pi : V \to V/W$. Se W' for um subespaço de V contendo W, $\pi(W')$ será um subespaço de V/W. Por isso, podemos formar o espaço quociente $\dfrac{V/W}{\pi(W')}$.

Pelo Teorema 4.3, $\pi(W')$ é isomorfo a W'/W. Assim,

$$\frac{V/W}{\pi(W')} \cong \frac{(V/W)}{(W'/W)}.$$

4.4 Teorema Segundo Teorema de Isomorfismo

Suponha que $W \subseteq W'$ sejam subespaços de V. Então

$$\frac{V/W}{W'/W} \cong \frac{V}{W'}.$$

Demonstração. Sejam

$$\pi : V \to V/W \quad \text{e} \quad \pi' : V/W \to \frac{(V/W)}{\pi(W')}$$

as projeções naturais. Defina

$$T = \pi' \circ \pi : V \to \frac{V/W}{\pi(W')}.$$

Como π e π' são ambos sobrejetivos, T é uma transformação linear sobrejetiva. Claramente, $W' \subseteq \ker T$. Seja $\boldsymbol{x} \in \ker T$. Então $\pi'\pi(\boldsymbol{x}) = 0$. Assim, $\overline{\boldsymbol{x}} = \pi(\boldsymbol{x}) \in \pi(W')$. Seja $\boldsymbol{y} \in W'$ tal que $\pi(\boldsymbol{y}) = \pi(\boldsymbol{x})$. Então $\pi(\boldsymbol{y} - \boldsymbol{x}) = 0$. Portanto, $\boldsymbol{y} - \boldsymbol{x} \in \ker \pi = W \subseteq W'$. Em particular, $\boldsymbol{x} \in W'$. Agora provamos que $\ker T = W'$. Aplicando o Teorema 4.3, temos $\dfrac{V/W}{\pi(W')} = \operatorname{im} T \cong V/\ker T = V/W'$. ■

O Terceiro Teorema de Isomorfismo lida com a relação entre somas e quocientes.

4.5 Teorema Terceiro Teorema de Isomorfismo

Suponha que W e W' sejam subespaços de V. Então

$$\frac{W + W'}{W} \cong \frac{W'}{W \cap W'}.$$

Demonstração. Seja $\pi : W + W' \to \dfrac{W+W'}{W}$ a projeção natural. A aplicação de inclusão de W' em $W + W'$ quando composto com π nos fornece uma transformação linear $T : W' \to (W + W')/W$. Como o núcleo de π é $W, \ker T = W \cap W'$.

Afirmamos que T é uma transformação sobrejetiva. Para demonstrar isso, considere um elemento $\overline{\boldsymbol{z}} \in \dfrac{W+W'}{W}$, então $\overline{\boldsymbol{z}}$ é um conjunto de W da forma $\boldsymbol{z} = \boldsymbol{w} + W$ com $\boldsymbol{w} \in W + W'$. Assim, $\boldsymbol{w} = \boldsymbol{x} + \boldsymbol{y}$ com $\boldsymbol{x} \in W$ e $\boldsymbol{y} \in W'$. Mas $\boldsymbol{x} + W = W$. Assim, $\boldsymbol{z} = \boldsymbol{w} + W = (\boldsymbol{y} + \boldsymbol{x}) + W = \boldsymbol{y} + W$. Em particular, $T(\boldsymbol{w}) = \boldsymbol{y} + W = \boldsymbol{z}$ e assim T é sobrejetiva.

Pelo Teorema 4.3, $\dfrac{W+W'}{W} = \operatorname{im} T \cong W'/\ker T = \dfrac{W'}{W \cap W'}$. ■

4.6 Proposição

Suponha que V seja uma soma direta interna dos subespaços $V_1, \ldots, V_n$.

$$\frac{V}{V_i} \cong \frac{V_1 \oplus \cdots \oplus V_n}{V_i} \cong V_1 \oplus \cdots \widehat{V_i} \oplus \cdots \oplus V_n \tag{4.9}$$

Onde $\widehat{V_i}$ significa a omissão de V_i nesta soma.

Demonstração. Suponha que $V = V_1 \oplus \cdots \oplus V_n$. Como $V_i \cap (\sum_{j \neq i} V_j) = (0)$, o

Teorema 4.3 implica

$$V/V_i = (V_i + \sum_{j\neq i} V_j)/V_i \tag{4.10}$$
$$\cong (\sum_{j\neq i} V_j)/(V_i \cap (\sum_{j\neq i} V_j)) \tag{4.11}$$
$$= (\sum_{j\neq i} V_j)/(0) \tag{4.12}$$
$$= V_1 \oplus \cdots \oplus \widehat{V_i} \oplus \cdots \oplus V_n \tag{4.13}$$

■

4.7 Definição Codimensão

Seja W um subespaço de V. Então a codimensão de W em V é

$$\operatorname{codim}_V W = \dim V/W.$$

4.8 Proposição

Seja W_1 um subespaço de V. Seja W_2 qualquer complemento de W_1 em V. Então $\operatorname{codim}_V W_1 = \dim W_2$.

Demonstração. V/W_1 e W_2 são isomorfos. ■

4.9 Proposição

Seja V um espaço vetorial da dimensão n e W seja um subespaço de V da dimensão k. Então $\dim V/W = \operatorname{codim}_V W = n - k$.

Aqui está uma maneira importante pela qual surgem espaços quocientes.

4.10 Definição Conúcleo

Seja $T\colon V \to W$ uma transformação linear. Então o **conúcleo** de T é o espaço quociente

$$\operatorname{coker}(T) = W/\operatorname{im}(T).$$

4.11 Corolário

Sejam V um espaço vetorial finito dimensional e $T\colon V \to V$ uma transformação linear. Então $\dim(\ker(T)) = \dim(\operatorname{coker}(T))$.

Demonstração.

$$\begin{aligned}\dim(\ker(T)) &= \dim(V) - \dim(\operatorname{im}(T)) = \dim(V/\operatorname{im}(T)) \\ &= \dim(\operatorname{coker}(T)).\end{aligned}$$

■

4.3.1 Somas diretas e sequências exatas cindidas

Considere a soma direta

$$Z = X \oplus Y$$

onde X e Y são subespaços do espaço vetorial Z sobre $\mathbb{K}$. Então, para todos os $z \in Z$, temos uma única decomposição

$$z = x + y,\ x \in X,\ y \in Y.$$

Definimos

- $i(x) := x$ para todos os $x \in X$ (inclusão canônica de X),
- $j(y) := y$ para todos os $y \in Y$ (inclusão de Y),
- $\pi(z) := y$ para todos os $z \in Z$ (projeção em Y).

Então a sequência

$$0 \longrightarrow X \xrightarrow{i} X \oplus Y \xrightarrow{\pi} Y \longrightarrow 0$$

é exata. Além disso, temos

$$\begin{tikzcd} 0 \arrow[r] & X \arrow[r, "i"] & X \oplus Y \arrow[r, shift left, "\pi"] & Y \arrow[l, shift left, "j"] \arrow[r] & 0 \end{tikzcd}$$

onde $j\colon Y \to X \oplus Y$ é a aplicação de inclusão. Motivados por essa situação, dizemos que a sequência

$$0 \longrightarrow X \xrightarrow{i} Z \xrightarrow{\pi} Y \longrightarrow 0$$

é uma sequência exata cindida se for uma sequência exata e a aplicação π possuir uma inversa a direita, ou seja, existe uma aplicação injetiva $j : Y \to Z$, de modo que $\pi \circ j = \mathrm{I}_Y$.

$$\begin{tikzcd} 0 \arrow[r] & X \arrow[r, "i"] & X \oplus Y \arrow[r, shift left, "\pi"] & Y \arrow[l, shift left, "j"] \arrow[r] & 0 \end{tikzcd}$$

4.12 Proposição

A sequência

$$0 \longrightarrow X \xrightarrow{i} Z \xrightarrow{\pi} Y \longrightarrow 0 \tag{4.14}$$

é uma sequência exata cindida se, e somente se, $Z \cong X \oplus Y$.

Demonstração. Suponha que 4.14 seja uma sequência exata cindida. Existe um subespaço vetorial $\ker(\pi)^{\perp}$ de Z tal que

$$Z = \ker(\pi) \oplus \ker(\pi)^{\perp}. \tag{4.15}$$

A aplicação π é sobrejetiva e portanto, $\operatorname{im}(\pi) = Y$ e $Y \cong Z/\ker(\pi)$. Como a aplicação i é injetiva, obtemos que $\operatorname{im}(i) \cong X$. Por exatidão, $\operatorname{im}(i) \cong \ker(\pi)$. Finalmente, a Equação 4.15 finaliza a demonstração. ■

Exercícios

Ex. 4.13 --- Use o primeiro teorema do isomorfismo para provar o teorema do rank-nulidade $\operatorname{rk}(T) + \operatorname{nul}(T) = \dim(V)$ para $T \in \operatorname{Hom}(V,\ W)$ e $\dim(V) < \infty$.

Ex. 4.14 --- Seja $T \in \operatorname{Hom}(V, V)$ e suponha que S seja um subespaço de V. Defina um mapa $T' : V/S \to V/S$ por

$$T'(v + S) = Tv + S$$

Quando T' está bem definido? Se T' está bem definido, é uma transformação linear? Determine $\operatorname{im}(T')$ e $\ker(T')$?

Ex. 4.15 --- Dados $M, N \subset L$. Prove que a seguinte aplicação é um isomorfismo linear

$$(M + N)/N \to M/(M \cap N) : m + n + N \mapsto m + M \cap N$$

Ex. 4.16 --- Prove que $\mathbb{R}^n/\mathbb{R}$ é isomorfo a $\mathbb{R}^{n-1}$.

Ex. 4.17 --- Seja V o espaço das funções reais contínuas no intervalo fechado $[a, b]$, $V = C[a, b]$ e seja $W = \operatorname{Const}[a, b]$ o espaço das funções constantes em [a,b]

1. Prove que $\operatorname{Const}[a, b]$ é isomorfo a $\mathbb{R}$

2. Prove que V/W é isomorfo ao espaço vetorial das funções continuas em $[a, b]$ que se anulam em a.

Ex. 4.18 --- Dado $L = M \oplus N$. Então a aplicação canônica

$$M \to L/N : m \mapsto m + N$$

é um isomorfismo.

Ex. 4.19 --- Encontre espaços vetoriais isomorfos V e W com

$$V = S \oplus B \text{ e } W = S \oplus D$$

mas $B \ncong D$. Portanto, $V \cong W$ não implica que $V/S \cong W/S$.

Ex. 4.20 --- Seja V um espaço vetorial com

$$V = S_1 \oplus T_1 = S_2 \oplus T_2$$

Prove que, se S_1 e S_2 possuem codimensão finita em V, o mesmo acontece com $S_1 \cap S_2$ e

$$\operatorname{codim}(S_1 \cap S_2) \leq \dim(T_1) + \dim(T_2)$$

Ex. 4.21 --- Seja V um espaço vetorial com

$$V = S_1 \oplus T_1 = S_2 \oplus T_2$$

Suponha que S_1 e S_2 tenham codimensão finita. Portanto, pelo exercício anterior, o mesmo acontece com $S_1 \cap S_2$. Encontre uma decomposição de soma direta $V = W \oplus X$ para a qual

1. W tenha codimensão finita,
2. $W \subseteq S_1 \cap S_2$ e
3. $X \supseteq T_1 + T_2$.

Ex. 4.22 --- Suponha $f \in \operatorname{Hom}_{\mathbb{K}}(V, \mathbb{K})$. Se $f \neq 0$, mostre que $V/\ker f \cong \mathbb{K}$.

Ex. 4.23 --- Sejam $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{R}[X]$ e $W = \{pf | p \in \mathbb{R}[X]\}$.

1. Mostre que W é um subespaço de $\mathbb{R}[X]$.
2. Mostre que $\dim(\mathbb{R}[X]/W) = n$. (Dica Use o algoritmo de divisão em $\mathbb{R}[X]$.)

Capítulo 5

Espaços Duais

O conceito de funcional linear, como um importante caso especial do conceito de transformação linear, é um dos principais conceitos da álgebra linear e desempenha também um papel significativo na análise. Neste capítulo introduzimos o conceito de funcional linear e de espaço dual, que é o espaço vetorial de todos os funcionais. No final do capítulo apresentamos a relação entre os funcionais lineares e os sistemas de equações lineares.

Nesse capítulo

- Espaços Duais (p. 144)
- Aniquiladores (p. 153)
- Transposta de uma Transformação Linear (p. 156)
- Sistemas de Equações (p. 161)

5.1 Espaços Duais

5.1 Definição Funcional Linear

Seja V um espaço vetorial sobre $\mathbb{K}$. Uma transformação linear $f \in \mathrm{Hom}(V,\ \mathbb{K})$ é dita **funcional linear** (ou simplesmente **funcional**) em V.

Os funcionais lineares também são denominados **covetores**, particularmente na literatura de Física.

5.2 Definição Espaço Dual

O espaço vetorial de todos os funcionais lineares em V é denotado por $V^* = \mathrm{Hom}_{\mathbb{K}}(V,\ \mathbb{K})$ e é denominado **espaço dual** de V.

Exemplo 3. *Os vetores colunas $f \in \mathbb{K}_n$ podem ser vistos como funcionais lineares no espaço $\mathbb{K}^n$.*

Dado $f \in \mathbb{K}_n$ e $\boldsymbol{v} \in \mathbb{K}^n$ então definimos

$$f(\boldsymbol{v}) = \boldsymbol{v}.f = [v_1, \dots v_n].\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Um exemplo interessante nesse contexto é o funcional média dado por:

$$\text{média}(v) = [1/n, 1/n, \ldots, 1/n] \cdot v.$$

$\triangleleft$

Exemplo 4. *A aplicação $E_0 : \mathbb{K}[x] \to \mathbb{K}$ definida por $E_0(p(x)) = p(0)$ é funcional linear, conhecido como avaliação em 0. Esse exemplo pode ser generalizado para todo $c \in \mathbb{K}$ definindo o funcional E_c como $E_c(p(x)) = p(c)$.*

$\triangleleft$

Exemplo 5. *Seja $C[a,\ b]$ o espaço vetorial de todas as funções contínuas em $[a,\ b] \subseteq \mathbb{R}$. Seja $f : C[a,\ b] \to \mathbb{R}$ definido por $f(\alpha(x)) = \displaystyle\int_a^b \alpha(x)dx$ Então $f \in C[a,\ b]^*$.*

$\triangleleft$

Exemplo 6. *Seja $\mathcal{M}_{n,n}(\mathbb{K})$ o espaço vetorial das matrizes quadradas $n \times n$. Então o traço de uma matriz é um funcional linear.*

$\triangleleft$

Como $\operatorname{im} f \subseteq \mathbb{K}$, temos dois casos:

- $\operatorname{im} f = \{\mathbf{0}\}$, nesse caso f é a função linear nula;
- $\operatorname{im} f = \mathbb{K}$, e nesse caso f é sobrejetiva.

Em outras palavras, um funcional linear não nulo é sobrejetivo.

Além disso, se $f \neq 0$ e se $\dim(V) < \infty$, então temos que

$$\dim(\ker f) = \dim(V) - 1.$$

5.7 Teorema

a Para qualquer vetor diferente de zero $\boldsymbol{v} \in V$, existe um $f \in V^*$ funcional linear para que $f(\boldsymbol{v}) \neq \mathbf{0}$.

b Um vetor $\boldsymbol{v} \in V$ é zero se, e somente se, $f(\boldsymbol{v}) = \mathbf{0}$ para todos os funcionais $f \in V^*$

c Dado $f \in V^*$ tal que $f(\boldsymbol{x}) \neq 0$ então ,

$$V = \langle \boldsymbol{x} \rangle \oplus \ker f$$

d Dois funcionais lineares não nulos $f, g \in V^*$ têm o mesmo núcleo se, e somente se, existir um escalar diferente de zero λ, de modo que $f = \lambda g$.

Demonstração.

$\boxed{c}$ Se $\mathbf{0} \neq \boldsymbol{v} \in \langle \boldsymbol{x} \rangle \cap \ker f$, então $f(\boldsymbol{v}) = 0$ e $\boldsymbol{v} = a\boldsymbol{x}$ para $0 \neq a \in \mathbb{K}$, e logo $f(\boldsymbol{x}) = 0$, o que é impossível. Portanto, $\langle \boldsymbol{x} \rangle \cap \ker f = \{\mathbf{0}\}$ e logo podemos considerar a soma direta $S = \langle x \rangle \oplus \ker f$. Além disso, para qualquer $\boldsymbol{v} \in V$, temos

$$\boldsymbol{v} = \frac{f(\boldsymbol{v})}{f(\boldsymbol{x})}\boldsymbol{x} + (\boldsymbol{v} - \frac{f(\boldsymbol{v})}{f(\boldsymbol{x})}\boldsymbol{x}) \in \langle \boldsymbol{x} \rangle + \ker f$$

e então $V = \langle \boldsymbol{x} \rangle \oplus \ker f$.

$\boxed{d}$ Se $f = \lambda g$ com $\lambda \neq 0$, $\ker f = \ker g$. Por outro lado, se $\mathbb{K} = \ker f = \ker(g)$, então para $\boldsymbol{x} \notin \mathbb{K}$, temos $V = \langle \boldsymbol{x} \rangle \oplus K$ Obviamente, $f\restriction_K = \lambda g \restriction_K$ para qualquer λ. Portanto, escolhendo $\lambda = f(\boldsymbol{x})/g(\boldsymbol{x})$, temos que $\lambda g(\boldsymbol{x}) = f(\boldsymbol{x})$ e, portanto, $f = \lambda g$.

■

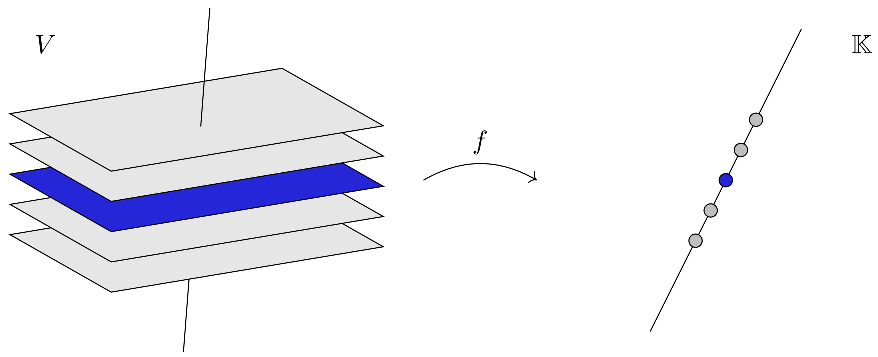


Figura 5.1 O Teorema 5.1 mostra que todo funcional linear "folheia" o espaço V em espaços afins de dimensão $n-1$, as classes $\lambda\boldsymbol{x} + \ker f$ com $\lambda \in \mathbb{K}$. Cada folha é levada num único ponto de $\mathbb{K}$.

5.1.1 Bases duais

Seja V um espaço vetorial com base $\underline{\mathrm{B}} = \{\boldsymbol{x}_i \,|\, i \in \Delta\}$. Para cada $i \in \Delta$, podemos definir um funcional linear $\boldsymbol{x}_i^* \in V^*$ pela condição de ortogonalidade

$$\boldsymbol{x}_i^*(\boldsymbol{x}_j) \triangleq \delta_{i,j}$$

onde $\delta_{i,j}$ é a função delta de Kronecker, definida como

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

e estendendo por linearidade.

Então o conjunto $\underline{\mathrm{B}}^* = \{\boldsymbol{x}_i^* \,|\, i \in \Delta\}$ é linearmente independente, pois se tivermos uma combinação que é igual ao funcional nulo

$$\mathbf{0}^* = a_{i_1}\boldsymbol{x}_i^* + \cdots + a_{i_n}\boldsymbol{x}_{i_n}^*.$$

aplicando essa equação no vetor da base $\boldsymbol{x}_{i_k}$ temos

$$\mathbf{0} = \sum_{j=1}^{n} a_{i_j} \boldsymbol{x}_{i_j}^{*}(\boldsymbol{x}_{i_k}) = \sum_{j=1}^{n} a_{i_j} \delta_{i_j, i_k} = a_{i_k}$$

para todos os i_k.

5.8 Teorema

Seja V um espaço vetorial com base $\underline{B} = \{\boldsymbol{x}_i \,|\, i \in \Delta\}$.

a O conjunto $\underline{B}^* = \{\boldsymbol{x}_i^* \,|\, i \in \Delta\}$ é linearmente independente.

b Se V é de dimensão finita, então $\underline{B}^*$ é uma base para V^*, denominada **base dual** de V^*.

Demonstração.

a Já foi demonstrada acima.

b Para qualquer $f \in V^*$, vamos provar que $f = \sum_j f(\boldsymbol{x}_j)\boldsymbol{x}_j^*$. Vamos provar isso calculando o lado direito da expressão anterior nos vetores da base:

$$\sum_j f(\boldsymbol{x}_j)\boldsymbol{x}_j^*(\boldsymbol{x}_i) = \sum_j f(\boldsymbol{x}_j)\delta_{i,j} = f(\boldsymbol{x}_i).$$

Portanto, $\underline{B}^*$ é uma base para V^*. ■

Segue-se do teorema anterior que se $\dim(V) < \infty$, então

$$\dim(V^*) = \dim(V)$$

já que os vetores duais também formam uma base para V^*. Nosso objetivo agora é mostrar que a recíproca também vale. Mas primeiro, vamos considerar um exemplo.

Exemplo 9. *Seja V um espaço vetorial de dimensão infinita sobre o corpo $\mathbb{K} = \mathbb{Z}_2 = \{0, 1\}$, com base $\underline{B}$. Como os únicos coeficientes em $\mathbb{K}$ são 0 e 1, uma combinação linear finita sobre $\mathbb{K}$ é apenas uma soma finita.*

Portanto, V é o conjunto de todas as somas finitas de vetores em $\underline{B}$ e, assim, podemos ver a escolha de um vetor $\boldsymbol{v} \in V$ como a escolha de um número finito de vetores em $\underline{B}$ (os vetores da base que aparecem com coeficiente 1 quando escrevemos $\boldsymbol{v}$ na base $\underline{B}$) e assim

$$|V| \leq |\mathscr{P}_0(\underline{B})| = |\underline{B}|$$

Por outro lado, cada funcional linear $f \in V^*$ é definido de maneira única especificando seus valores em $\underline{B}$. Como esses valores devem ser 0 ou 1, especificar um funcional linear é equivalente a especificar o subconjunto de $\underline{B}$ no qual f assume o valor 1. Em outras palavras, há uma correspondência biunívoca entre os *funcionais lineares em V e todos os subconjuntos de $\underline{B}$. Consequentemente,*

$$|V^*| = |\mathscr{P}(\underline{B})| > |\underline{B}| \geq |V|$$

Isso mostra que V^* não pode ser isomorfo a V, nem a qualquer subconjunto próprio de V. Portanto, $\dim(V^*) > \dim(V)$.

◁

O comportamento no exemplo anterior é típico, ou seja,

$$\dim(V) \leq \dim(V^*)$$

com igualdade se, e somente se, V for de dimensão finita.

5.10 Teorema

Seja V um espaço vetorial. Então $\dim(V) \leq \dim(V^*)$ com igualdade se, e somente se, V for de dimensão finita.

Demonstração. Para qualquer espaço vetorial V, temos

$$\dim(V) \leq \dim(V^*)$$

i.e., já que os vetores duais para uma base $\underline{B}$ para V são linearmente independentes em V^*. Já vimos que se V é de dimensão finita, então $\dim(V) = \dim(V^*)$.

A demonstração de que se V é de dimensão infinita, então $\dim(V) < \dim(V^*)$ é bastante elaborada e não será feita nesse texto.

■

Exemplo 11. *Se $V = \mathbb{K}^n$ então $V^* = \mathbb{K}_n$.*

Pelo Teorema 3.1.1 todo funcional linear em $\mathbb{K}^n$ é da forma $\mathbb{K}_n$, e assim temos o resultado.

◁

5.12 Definição Hiperplano

Seja V um espaço vetorial de dimensão finita n, um **hiperplano** H é um subespaço de V de dimensão $n-1$.

O seguinte resultado descreve a relação entre os funcionais lineares e os hiperplanos e sua demonstração é simples e será deixada como exercício para o leitor.

5.13 Teorema

Seja V um espaço vetorial de dimensão finita

1. Seja $0 \neq f \in V^*$. Então $\ker f$ é um hiperplano.
2. Todo hiperplano é da forma $\ker f$ para algum $0 \neq f \in V^*$

5.1.2 Bidual

Se V é um espaço vetorial, com espaço dual V^*. Podemos formar o espaço bidual V^{**}, que consiste em todos os funcionais lineares $S : V^* \to \mathbb{K}$. Em outras palavras, um elemento S de V^{**} é um funcional linear que atribui um escalar a cada funcional linear em V.

5.14 Definição Espaço Bidual

Definimos o **espaço bidual** de V, denotado por V^{**} como

$$V^{**} \triangleq \mathrm{Hom}(V^*, \mathbb{K})$$

Com isso em mente, há uma maneira bastante óbvia de obter um elemento de V^{**}. Ou seja, se $\boldsymbol{v} \in V$, considere a aplicação $\overline{\boldsymbol{v}} : V^* \to \mathbb{K}$ definida por

$$\overline{\boldsymbol{v}}(f) \triangleq f(\boldsymbol{v})$$

que envia o f funcional linear para o escalar $f(\boldsymbol{v})$. A aplicação $\overline{\boldsymbol{v}}$ é denominada **avaliação** em $\boldsymbol{v}$. Para ver que $\overline{\boldsymbol{v}} \in V^{**}$, sejam $f, g \in V^*$ e $a, b \in \mathbb{K}$, então

$$\overline{\boldsymbol{v}}(af + bg) = (af + bg)(\boldsymbol{v}) = af(\boldsymbol{v}) + bg(\boldsymbol{v}) = a\overline{\boldsymbol{v}}(f) + b\overline{\boldsymbol{v}}(g)$$

e assim $\overline{\boldsymbol{v}}$ é de fato um funcional linear.

Agora podemos definir uma aplicação $\tau : V \to V^{**}$ por $\tau\boldsymbol{v} = \overline{\boldsymbol{v}}$ Essa aplicação é denominada **aplicação natural** de V em V^{**}. Essa aplicação é injetiva e no caso de dimensão finita, também é sobrejetiva.

5.15 Teorema

A aplicação natural $\tau : V \to V^{**}$ definida por $\tau\boldsymbol{v} = \overline{\boldsymbol{v}}$, em que $\overline{\boldsymbol{v}}$ é a avaliação em $\boldsymbol{v}$, é injetiva.

Se V é de dimensão finita, então τ é um isomorfismo.

Demonstração. A aplicação τ é linear, pois

$$\overline{a\boldsymbol{u}+b\boldsymbol{v}}(f)=f(a\boldsymbol{u}+b\boldsymbol{v})=af(\boldsymbol{u})+bf(\boldsymbol{v})=(a\overline{\boldsymbol{u}}+b\overline{\boldsymbol{v}})(f)$$

para todos os $f\in V^*$. Para determinar o núcleo de τ, observe que

$$\tau\boldsymbol{v}=0\Rightarrow\overline{\boldsymbol{v}}=0 \tag{5.1}$$
$$\Rightarrow\overline{\boldsymbol{v}}(f)=0 \text{ para todos os } f\in V^* \tag{5.2}$$
$$\Rightarrow f(\boldsymbol{v})=0 \text{ para todos os } f\in V^* \tag{5.3}$$
$$\Rightarrow\boldsymbol{v}=0 \tag{5.4}$$

pelo Teorema 5.1 e, portanto, $\ker(\tau)=\{\mathbf{0}\}$. No caso de dimensão finita, como

$$\dim(V^{**})=\dim(V^*)=\dim(V)$$

temos que τ também é sobrejetivo, e logo um isomorfismo. ∎

Observe que se $\dim(V)<\infty$, como as dimensões de V e V^{**} são as mesmas, deduzimos imediatamente que $V\cong V^{**}$. Este não é o ponto do Teorema 5.1.2. O ponto é que a aplicação natural $\boldsymbol{v}\to\overline{\boldsymbol{v}}$ é um isomorfismo.

Quando $V\cong V^{**}$, V é dito algebricamente reflexivo. O Teorema 5.1.2 e o Teorema 5.1.1 juntos implicam que um espaço vetorial é algebricamente reflexivo se, e somente se, for finito-dimensional.

Se V for de dimensão finita, é habitual identificar o espaço bi-dual V^{**} com V e pensar nos elementos de V^{**} simplesmente como vetores em V. Vamos considerar um exemplo específico para mostrar como a reflexividade algébrica falha no caso de dimensão infinita.

Exemplo 16. *Seja $V=(\mathbb{Z}_2)^\infty$ espaço vetorial sobre $\mathbb{Z}_2$ e seja a base*

$$\boldsymbol{e}_k=(0,\ \ldots,\ 0,1,0,\ \ldots)$$

onde o 1 está na k-ésima posição. Portanto, V é o conjunto de todas as sequências binárias infinitas com um número finito de 1s. Defina a ordem $o(\boldsymbol{v})$ de qualquer $\boldsymbol{v}\in V$ como a maior coordenada de $\boldsymbol{v}$ com o valor 1. Então $o(\boldsymbol{v})<\infty$ para todos os $\boldsymbol{v}\in V$.

Considere os vetores duais $\boldsymbol{e}_k^*$, definidos (como de costume) por

$$\boldsymbol{e}_k^*(\boldsymbol{e}_j)=\delta_{k,j}$$

Para qualquer $\boldsymbol{v}\in V$, a avaliação funcional $\overline{\boldsymbol{v}}$ tem a propriedade que

$\overline{\boldsymbol{v}}(\boldsymbol{e}_k^*)=\boldsymbol{e}_k^*(\boldsymbol{v})=0$ se $k>o(\boldsymbol{v})$

No entanto, como os vetores duais $\boldsymbol{e}_k^*$ são linearmente independentes, Pelo Teorema 3.1 existe um f funcional linear em V^{**} para o qual

$$f(\boldsymbol{e}_k^*)=1$$

para todos os $k \geq 1$. *Portanto,* f *não possui o formato* $\overline{\boldsymbol{v}}$ *para qualquer* $\boldsymbol{v} \in V$. *Isso mostra que a aplicação natural não é sobrejetiva e, portanto,* V *não é algebricamente reflexivo.*

$\triangleleft$

Exercícios

Ex. 5.1 --- Seja $V = \mathbb{R}_2[x]$ o espaço vetorial dos polinomiais de grau menor ou igual a dois.

1. Mostre que os funcionais $f_1(p) = \int_0^1 p(x)dx$, $f_2(p) = \int_0^2 p(x)dx$ e $f_3(p) = \int_0^3 p(x)dx$ formam uma base de V^*

2. Encontre sua sua base dual em V.

Ex. 5.2 --- Seja V um espaço vetorial finito dimensional e sejam $\boldsymbol{v}_1 \neq \boldsymbol{v}_2$ em V. Prove que existe $f \in V^*$ com $f(\boldsymbol{v}_1) \neq f(\boldsymbol{v}_2)$.

Ex. 5.3 --- Seja V um espaço vetorial e sejam $f, g \in V^*$ tais que que $f(\boldsymbol{v}) = 0$ se, e somente se, $g(\boldsymbol{v}) = 0$. Prove que $f = \lambda g$ para algum$\lambda \in \mathbb{K}$.

Ex. 5.4 --- Seja $0 \neq \boldsymbol{y} \in V$ e $f \in V^* \setminus \{\boldsymbol{0}\}$. Defina $T : V \to V$ por $T(\boldsymbol{x}) = f(\boldsymbol{x})\, \boldsymbol{y}$. Uma função definida desse modo é denominado **díade**.

1. Mostre que $T \in \mathrm{Hom}(V,\ V)$ satisfaz $\dim(\mathrm{im}\, T) = 1$.

2. Se $S \in \mathrm{Hom}(V,\ V)$ satisfaz $\dim(\mathrm{im}\, S) = 1$, mostre que S é uma díade.

Ex. 5.5 --- Seja $V = C([0, 1])$ o espaço vetorial sobre $\mathbb{R}$ das funções contínuas a valores reais no intervalo $[0, 1]$. Defina a transformação

$$\phi\ \ V \to V^*,\ f \mapsto (g \mapsto \int_0^1 f(x)g(x)dx)\ .$$

Mostre que ϕ é uma transformação linear cujo núcleo é trivial.

Ex. 5.6 --- Seja V um espaço vetorial de dimensão finita sobre $\mathbb{K}$. Seja $W = V \oplus V^*$. Mostre que a aplicação $(\boldsymbol{x},\ \boldsymbol{y}) \to (\boldsymbol{y},\ \boldsymbol{x})$ é um isomorfismo entre W e W^*.

Ex. 5.7 --- Prove que $(V_1 \oplus \cdots \oplus V_n)^* \cong V_1^* \oplus \cdots \oplus V_n^*$.

Ex. 5.8 --- Seja $v \in V$ um vetor fixo. Prove que a transformação de avaliação:

$$\overline{\boldsymbol{v}} : V^* \to \mathbb{R},\ \overline{\boldsymbol{v}}(f) := f(v)$$

é um elemento de V^{**}.

Ex. 5.9 --- Seja $V = \mathbb{R}[x]$. Quais das seguintes funções em V são elementos em V^*:

1. $T(p) = \int_0^1 p(x)dx.$
2. $T(p) = \int_0^1 p(x)^2 dx.$
3. $T(p) = \int_0^1 x^2 p(x)dx.$
4. $T(p) = dp/dx.$
5. $T(p) = dp/dx|_{x=0}.$

Ex. 5.10 --- Suponha que $\mathbb{K}$ é um corpo finito. Seja V um espaço vetorial sobre $\mathbb{K}$ de dimensão n. Para cada $m \leq n$, mostre o número de subespaços de V de dimensão m é exatamente o mesmo que o número de subespaços de V de dimensão $n - m$.

Ex. 5.11 --- Seja $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$ onde $A = (a_{ij})$. Mostre que

$$\operatorname{tr} \in \left(\mathcal{M}_{n,n}(\mathbb{K})\right)^* .$$

Ex. 5.12 --- Mostre que $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ para todo $A, B \in \mathcal{M}_{n,n}(\mathbb{K})$.

Ex. 5.13 --- Seja $m,\ n \in \mathbb{N}$. Seja $f_1,\ \ldots,\ f_m \in (\mathbb{K}^n)^*$. Defina $T : \mathbb{K}^n \to \mathbb{K}^m$ por $T(\boldsymbol{x}) = (f_1(\boldsymbol{x}),\ \ldots,\ f_m(\boldsymbol{x}))$

1. Mostre que $T \in \operatorname{Hom}_{\mathbb{K}}(\mathbb{K}^n,\ \mathbb{K}^m)$.
2. Mostre que todo $T \in \operatorname{Hom}_{\mathbb{K}}(\mathbb{K}^n,\ \mathbb{K}^m)$ é dado desta maneira para alguns $f_1, \ldots, f_m$.

Ex. 5.14 --- Seja V um espaço vetorial de dimensão finita sobre $\mathbb{C}$. Suponha que $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ sejam vetores distintos e diferentes de zero em V. Mostre que existe um $T \in V^*$ tal que $T(\boldsymbol{x}_k) \neq 0$ para todos os $k = 1, \ldots, n$.

5.2 Aniquiladores

Os funcionais lineares $f \in V^*$ são definidas para vetores em V, mas também podemos definir f para subconjuntos M de V fazendo

$$f(M) = \{f(v)|v \in M\}$$

5.1 Definição Aniquilador

Sejam V um espaço vetorial vetorial sobre $\mathbb{K}$ e V^* seu espaço dual. Seja M um subconjunto de V. Definimos

$$M^\perp = \operatorname{Aniq}_{V^*} M \triangleq \{f \in V^* \mid f(M) = 0\} = \{f \in V^* \mid f(\boldsymbol{u}) = 0, \text{ para todo } \boldsymbol{u} \in M\}$$

o aniquilador de M em V^*.

A razão do termo aniquilador é intuitiva, pois $M^\perp$ consiste em todos os funcionais lineares que aniquilam, iso é, enviam para 0 todos os vetores em M. Ressaltamos que $M^\perp$ é um subespaço de V^*, mesmo quando M não é um subespaço de V.

As propriedades básicas dos aniquiladores estão contidas no seguinte teorema.

5.2 Teorema

a $M^\perp = (\langle M \rangle)^\perp$

b Se M e N forem subconjuntos de V,

$$M \subseteq N \Rightarrow N^\perp \subseteq M^\perp$$

c Se $\dim(V) < \infty$, então, para qualquer subconjunto M de V a aplicação natural

$$\tau : \langle M \rangle \to M^{\perp\perp}$$

é um isomorfismo de $\langle M \rangle$ para $M^{\perp\perp}$. Em particular, se S for um subespaço de V, $S^{\perp\perp} \cong S$.

d Se S e T são subespaços de V, então

$$(S \cap T)^\perp = S^\perp + T^\perp \text{ e } (S + T)^\perp = S^\perp \cap T^\perp$$

Demonstração. Deixamos a prova da parte *a* e *b* para o leitor.

Para a parte *c*, como $M^{\perp\perp} = (\langle M \rangle)^{\perp\perp}$ é suficiente provar que $\tau : S \to S^{\perp\perp}$ é um isomorfismo, onde S é um subespaço de V. Observamos inicialmente que

τ é injetiva, logo resta demonstrar que $\tau S = S^{\perp\perp}$. Se $\boldsymbol{s} \in S$, $\tau \boldsymbol{s} = \overline{\boldsymbol{s}}$ possui a propriedade que, para todos os $f \in S^{\perp}$, $\overline{\boldsymbol{s}}(f) = f\boldsymbol{s} = 0$ e então $\tau\boldsymbol{s} = \overline{\boldsymbol{s}} \in S^{\perp\perp}$, o que implica que $\tau S \subseteq S^{\perp\perp}$. Além disso, se $\overline{\boldsymbol{v}} \in S^{\perp\perp}$, então, para todos os $f \in S^{\perp}$, temos $f(\boldsymbol{v}) = \overline{\boldsymbol{v}}(f) = 0$ e assim, todo funcional linear que aniquila S também aniquila $\boldsymbol{v}$. Mas se $\boldsymbol{v} \notin S$, existe um funcional linear $g \in V$ para o qual $g(S) = \{\mathbf{0}\}$ e $g(\boldsymbol{v}) \neq 0$.

Portanto, $\boldsymbol{v} \in S$ e, portanto, $\overline{\boldsymbol{v}} = \tau\boldsymbol{v} \in \tau S$ e então $S^{\perp\perp} \subseteq \tau S$.

Para a parte d, é claro que f aniquila $S+T$ se, e somente se, f aniquila S e T. Portanto, $(S+T)^{\perp} = S^{\perp} \cap T^{\perp}$. Além disso, se $f = g+h \in S^{\perp}+T^{\perp}$ em que $g \in S^{\perp}$ e $h \in T^{\perp}$, $g, h \in (S\cap T)^{\perp}$ e, portanto, $f \in (S\cap T)^{\perp}$. Portanto,

$$S^{\perp}+T^{\perp} \subseteq (S\cap T)^{\perp}$$

Para a inclusão inversa, suponha que $f \in (S\cap T)^{\perp}$. Escreva

$$V = S' \oplus (S\cap T) \oplus T' \oplus U$$

onde $S = S' \oplus (S\cap T)$ e $T = (S\cap T) \oplus T'$. Defina $g \in V^*$ por

$$g\restriction_{S'} = f,\ g\restriction_{S\cap T} = f\restriction_{S\cap T} = 0,\ g\restriction_{T'} = 0,\ g\restriction_{U} = f$$

e defina $h \in V^*$ por

$$h\restriction_{S'} = 0,\ h\restriction_{S\cap T} = f\restriction_{S\cap T} = 0,\ h\restriction_{T'} = f,\ h\restriction_{U} = 0$$

Dessa forma temos que $g \in T^{\perp}, h \in S^{\perp}$ e $g+h = f$. ■

5.2.1 Aniquiladores e Somas Diretas

Considere uma decomposição de soma direta $V = S \oplus T$.

Então qualquer funcional linear $f \in T^*$ pode ser estendido a um funcional linear $\overline{f}$ em V fazendo $\overline{f}(S) = 0$. Vamos chamar essa extensão de extensão por 0. Claramente, $\overline{f} \in S^{\perp}$ e é fácil ver que a extensão por 0 leva $f \to \overline{f}$ e é um isomorfismo de T^* a $S^{\perp}$, cuja inversa é a restrição a T.

5.3 Teorema

Seja $V = S \oplus T$.

- a A extensão por 0 é um isomorfismo entre T^* e $S^{\perp}$, i.e., $T^* \cong S^{\perp}$.
- b Se V é de dimensão finita, então

$$\dim(S^{\perp}) = \operatorname{codim}_V(S) = \dim(V) - \dim(S)\ \square$$

Exemplo 4. *A parte b do Teorema 5.2.1 pode falhar no caso de dimensão infinita, pois pode acontecer facilmente que $S^{\perp} \cong V^*$. Como exemplo, seja V o espaço* *vetorial sobre* $\mathbb{Z}_2$ *com uma base ordenada enumerável* $\underline{B} = (\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \ldots)$*. Seja* $S = \langle \boldsymbol{e}_1 \rangle$ *e* $T = \langle \boldsymbol{e}_2, \boldsymbol{e}_3, \ldots \rangle$*. É fácil ver que* $S^{\perp} \cong T^* \cong V^*$ *e que* $\dim(V^*) > \dim(V)$.

◁

O aniquilador fornece uma maneira de descrever o espaço dual de uma soma direta.

5.5 Teorema

Um funcional linear na soma direta $V = S \oplus T$ pode ser escrito como a soma de uma funcional linear que aniquila S e de uma funcional linear que aniquila T, ou seja,

$$(S \oplus T)^* = S^{\perp} \oplus T^{\perp}$$

Demonstração. Claramente $S^{\perp} \cap T^{\perp} = \{\mathbf{0}\}$, pois qualquer funcional que aniquile S e T deve aniquilar $S \oplus T = V$. Portanto, a soma $S^{\perp} + T^{\perp}$ é direta. O restante segue do Teorema 5.2, já que $V^* = \{\mathbf{0}\}^{\perp} = (S \cap T)^{\perp} = S^{\perp} + T^{\perp} = S^{\perp} \oplus T^{\perp}$ Como alternativa, como $\rho_T + \rho_S = \mathrm{I}$ é a aplicação identidade, se $f \in V^*$, podemos escrever $f = f \circ (\rho_T + \rho_S) = (f \circ \rho_T) + (f \circ \rho_S) \in S^{\perp} \oplus T^{\perp}$ e então $V^* = S^{\perp} \oplus T^{\perp}$. ■

Exercícios

Ex. 5.15 --- Seja $A = \{f_1,\ \ldots,\ f_n\} \subseteq V^*$. Mostre $A^{\perp} = \bigcap_{i=1}^{n} \ker f_i$.

Ex. 5.16 --- Seja $A = \{f_1,\ \ldots,\ f_n\} \subseteq V^*$ e suponha que $g \in V^*$ tal que g se anula em $A^{\perp}$. Mostre que $g \in \langle A \rangle$.

Ex. 5.17 --- Seja V um espaço vetorial finito dimensional com espaço dual V^*. Se W_1, W_2 forem subespaços de V, mostre que

1. $\mathrm{Aniq}(W_1 \cap W_2) = \mathrm{Aniq}(W_1) + \mathrm{Aniq}(W_2)$.
2. $\mathrm{Aniq}(W_1 + W_2) = \mathrm{Aniq}(W_1) \cap \mathrm{Aniq}(W_2)$.

Ex. 5.18 --- Seja V um espaço vetorial com um subespaço $U \subseteq V$. Mostre que o aniquilador $U^{\perp}$ é realmente um espaço vetorial.

Ex. 5.19 --- Seja V um espaço vetorial. Quais são os aniquiladores de $\{\mathbf{0}\}$ e V respectivamente?

Ex. 5.20 --- Seja V um espaço vetorial de dimensão finita. Suponha que $U \subseteq V$ seja um subespaço diferente de $\{\mathbf{0}\}$. Mostre que $U^{\perp} \neq V^*$

Ex. 5.21 --- Seja V um espaço vetorial de dimensão finita e $U \subseteq V$ seja um subespaço. Mostre que $\dim V = \dim U + \dim U^{\perp}$.

Ex. 5.22 --- Seja V um espaço vetorial de dimensão finita e U e W sejam subespaços de V. Mostre que se $U \subseteq W$, $W^{\perp} \subseteq U^{\perp}$.

Ex. 5.23 --- Prove que se $m < n$, e se $f_1 \dots, f_m$ são funcionais lineares em um espaço vetorial n-dimensional V, então existe um vetor diferente de zero x em V tal que $f_j(x) = 0$ para $j = 1, \dots, m$. O que esse resultado diz sobre as soluções de equações lineares?

Ex. 5.24 --- Suponha que $m < n$ e $f_1, \dots, f_m$ sejam funcionais lineares em um espaço vetorial n-dimensional V. Em que condições dos escalares $\alpha_1, \dots, \alpha_m$ é verdade que existe um vetor x em V tal que $f_j(x) = \alpha_j$ para $j = 1, \dots, m$. O que esse resultado diz sobre as soluções de equações lineares?

Ex. 5.25 --- Seja $V_0 \subset V_1 \subset \dots \subset V_n$ uma bandeira maximal em V. Mostre que $V_n^{\perp} \subset V_{n-1}^{\perp} \subset \dots \subset V_0^{\perp}$ é uma bandeira maximal em V^*.

Ex. 5.26 --- Sejam V um espaço vetorial de dimensão finita sobre $\mathbb{K}$ e seja W um subespaço próprio de V de dimensão k.

1. Mostre que existem $n - k$ funcionais lineares $\{f_i\}$ de modo que

$$W = \bigcap_{i=1}^{m} \ker f_i$$

2. Conclua usando o Teorema 5.1.1 que W é a intersecção de $n - k$ hiperplanos.

5.3 Transposta de uma Transformação Linear

5.1 Definição Transposta de uma Transformação

Se $T \in \mathrm{Hom}(V,\ W)$, poderemos definir uma transformação $T^* : W^* \to V^*$

por

$$T^*(f) = f \circ T = fT$$

para $f \in W^*$. Assim, para qualquer $v \in V$,

$$[T^*(f)](v) = f(Tv)$$

A transformação T^* é chamada de **transposta da transformação** de T. . A transposta da transformação será também denotada por T^{t}.

5.2 Teorema Propriedades da Tranposta de uma Transformação

a Para $T, S \in \mathrm{Hom}(V,\ W)$ e $a, b \in \mathbb{K}$, $(aT + bS)^* = aT^* + bS^*$

b Para $S \in \mathrm{Hom}(V,\ W)$ e $T \in \mathrm{Hom}(W,\ U)$, $(TS)^* = S^*T^*$

c Para qualquer $T \in \mathrm{Hom}(V, V)$ invertível , $(T^{-1})^* = (T^*)^{-1}$

Demonstração. A prova 1 é deixada para o leitor.

2 Para todos os $f \in U^*$, temos que

$$(TS)^*(f) = f(TS) = S^*(fT) = S^*(T^*(f)) = (S^*T^*)(f)$$

4

$$T^*(T^{-1})^* = (T^{-1}T)^* = \mathrm{I}^* = \mathrm{I}$$

e analogamente , $(T^{-1})^*T^* = \mathrm{I}$. Portanto, $(T^{-1})^* = (T^*)^{-1}$. ■

Se $T \in \mathrm{Hom}(V,\ W)$, $T^* \in \mathrm{Hom}(W^*,\ V^*)$ e então $T^{**} \in \mathrm{Hom}(V^{**},\ W^{**})$. Obviamente, T^{**} não é igual a T. No entanto, no caso de dimensão finita, se usarmos as aplicações naturais para identificar V^{**} com V e W^{**} com W, poderemos pensar em T^{**} como estando em $\mathrm{Hom}(V,\ W)$. Usando essas identificações, temos igualdade no caso de dimensão finita.

5.3 Teorema

Seja V e W de dimensão finita e seja $T \in \mathrm{Hom}(V,\ W)$. Se identificarmos V^{**} com V e W^{**} com W usando as aplicações naturais, então T^{**} será identificado com T.

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, "\tau"'] & W \arrow[d, "\tau"] \\
V^{**} \arrow[r, "T^{**}"'] & W^{**}
\end{tikzcd}$$

Demonstração. Antes de fazer qualquer identificação, temos para $\boldsymbol{v} \in V$,

$$T^{**}(\overline{\boldsymbol{v}})(f) = \overline{\boldsymbol{v}}[T^*(f)] = \overline{\boldsymbol{v}}(fT) = f(T\boldsymbol{v}) = \overline{T\boldsymbol{v}}(f)$$

para todos os $f \in W^*$ e assim

$$T^{**}(\overline{\boldsymbol{v}}) = \overline{T\boldsymbol{v}} \in W^{**}$$

Portanto, usando as identificações canônicas para V^{**} e W^{**}, temos

$$T^{**}(\boldsymbol{v}) = T\boldsymbol{v}$$

para todos os $\boldsymbol{v} \in V$. ■

O próximo resultado descreve o núcleo e a imagem da transposta de uma transformação linear.

5.4 Teorema

Seja $T \in \operatorname{Hom}(V, W)$. Então

a) $\ker(T^*) = \operatorname{im}(T)^{\perp}$

b) $\operatorname{im}(T^*) = \ker(T)^{\perp}$

Demonstração. Para a parte a),

$$\begin{aligned}
\ker(T^*) &= \{f \in W^* \mid T^*(f) = 0\} && (5.5)\\
&= \{f \in W^* \mid f(TV) = \{\mathbf{0}\}\} && (5.6)\\
&= \{f \in W^* \mid f(\operatorname{im}(T)) = \{\mathbf{0}\}\} && (5.7)\\
&= \operatorname{im}(T)^{\perp} && (5.8)
\end{aligned}$$

Para a parte b), se $f = gT = T^*g \in \operatorname{im}(T^*)$, então $\ker(T) \subseteq \ker(f)$ e assim $f \in \ker(T)^{\perp}$.

Para inclusão inversa, seja $f \in \ker(T)^{\perp} \subseteq V^*$. Queremos mostrar que $f = T^*g = gT$ para algum $g \in W^*$. Em $K = \ker(T)$, não há problema, já que f e $T^*g = gT$ concordam em K para qualquer $g \in W^*$. Seja S um complemento de $\ker(T)$. Então T leva uma base $\underline{\mathrm{B}} = \{\boldsymbol{e}_i \mid i \in \Delta\}$ de S para um conjunto linearmente independente

$$T\underline{\mathrm{B}} = \{T\boldsymbol{e}_i \mid i \in \Delta\}$$

em W e, assim, podemos definir $g \in W^*$ em $T\underline{\mathrm{B}}$ fazendo

$$g(T\boldsymbol{e}_i) = f\boldsymbol{e}_i$$

e estendendo a W. Então $f = gT = T^*g$ em $\underline{\mathrm{B}}$ e, portanto, em S. Assim, $f = T^*g \in \operatorname{im}(T^*)$. ■

5.5 Corolário

Seja $T \in \operatorname{Hom}(V,\ W)$, onde V e W são de dimensões finitas. Então $\operatorname{posto}(T) = \operatorname{posto}(T^*)$.

No caso de dimensão finita, T e T^* podem ser representados por matrizes. Sejam $\underline{B} = (\boldsymbol{b}_1,\ \ldots,\ \boldsymbol{b}_n)$ e $\underline{C} = (\boldsymbol{c}_1,\ \ldots,\ \boldsymbol{c}_m)$ bases ordenadas de V e W, respectivamente, e sejam $\underline{B}^* = (\boldsymbol{b}_1^*,\ \ldots,\ \boldsymbol{b}_n^*)$ e $\underline{C}^* = (\boldsymbol{c}_1^*,\ \ldots,\ \boldsymbol{c}_m^*)$ as bases duais correspondentes. Então

$$([T]_{\underline{B},\underline{C}})_{i,j} = ([T\boldsymbol{b}_j]_{\underline{C}})_i = \boldsymbol{c}_i^*[T\boldsymbol{b}_j]$$

e

$$([T^*]_{\underline{C}^*,\underline{B}^*})_{i,j} = ([T^*(\boldsymbol{c}_j^*)]_{\underline{B}^*})_i = \boldsymbol{b}_i^{**}[T^*(\boldsymbol{c}_j^*)] = T^*(\boldsymbol{c}_j^*)(\boldsymbol{b}_i) = \boldsymbol{c}_j^*(T\boldsymbol{b}_i)$$

Comparando as duas últimas expressões, vemos que elas são as mesmas, exceto que os papéis de i e j estão invertidos. Portanto, as matrizes em questão são transpostas uma da outra.

5.6 Teorema

Seja $T \in \operatorname{Hom}(V,\ W)$, onde V e W são espaços vetoriais de dimensões finitas. Se $\underline{B}$ e $\underline{C}$ são bases ordenadas para V e W, respectivamente, e $\underline{B}^*$ e $\underline{C}^*$ são as bases duais correspondentes, então

$$[T^*]_{\underline{C}^*,\underline{B}^*} = ([T]_{\underline{B},\underline{C}})^{\mathrm{t}}$$

Ou seja, as matrizes de T e de T^* são transpostas uma da outra, o que justifica a nomenclatura de transposta de uma transformação.

Exercícios

Ex. 5.27 --- Prove que, se S e T são subespaços de V, $(S \oplus T)^* \cong S^\perp \oplus T^\perp$.

Ex. 5.28 --- Prove que $\{\mathbf{0}\}^* = \{\mathbf{0}\}$ e $\mathrm{I}^* = \mathrm{I}$ em que 0 é a transformação linear nula e ι é a identidade.

Ex. 5.29 --- Seja S um subespaço de V. Prove que $(V/S)^* \cong S^\perp$.

Ex. 5.30 --- Verifique que

1. $(T+S)^* = T^* + S^*$ por $T, S \in \operatorname{Hom}(V,\ W)$.

2. $(\lambda T)^* = \lambda T^*$ para qualquer $\lambda \in \mathbb{K}$ e $T \in \operatorname{Hom}(V,\ W)$

Ex. 5.31 --- Seja $T \in \operatorname{Hom}(V,\ W)$, onde V e W são de dimensão finita. Prove que $\operatorname{posto}(T) = \operatorname{posto}(T^*)$.

Ex. 5.32 --- Se $U \to^T V \to^S W$ são aplicações lineares entre espaços vetoriais finito dimensionais , prove que

$$(S \circ T)^* = T^* \circ S^*$$

Observe a reversão da ordem quando calcularmos a transposta de um produto.

Ex. 5.33 --- . Se V, W são finito dimensionais e $T : V \to W$ é uma transformação linear invertível , prove que $T^* : W^* \to V^*$ também é invertível e $(T^{-1})^* = (T^*)^{-1}$ como mapas de $V^* \to W^*$.

Ex. 5.34 --- Se T é uma díade em V, mostre que T^* é uma díade em V^*.

Ex. 5.35 --- Sejam V e W espaços vetoriais de dimensão finita sobre $\mathbb{K}$. Sejam $\tau_V : V \to V^{**}$ e $\tau_W : W \to W^{**}$ os isomorfismos dados no Teorema 5.1.2. Mostre que para todo $T \in \operatorname{Hom}(V,\ W)$ o seguinte diagrama é comutativo:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\tau_V"'] & W \arrow[d, "\tau_W"'] \\ V^{**} \arrow[r, "T^{**}"] & W^{**} \end{tikzcd}$$

Ex. 5.36 --- Se

$$0 \longrightarrow V \xrightarrow{T} W \xrightarrow{S} Z \longrightarrow 0$$

é uma sequência exata curta de espaços vetoriais sobre $\mathbb{K}$. Mostre que

$$0 \longrightarrow Z^* \xrightarrow{T^*} W^* \xrightarrow{S^*} V^* \longrightarrow 0$$

é exata.

Ex. 5.37 --- Sejam V, W espaços vetoriais de dimensão finita sobre $\mathbb{K}$, e $f \in \operatorname{Hom}(V, W)$. Mostre que

1. f é injetiva se, e somente se f^* for sobrejetiva.

2. f é sobrejetiva se, e somente se, f^* é injetiva.

3. f é bijetiva se, e somente se, f^* é bijetiva.

5.4 Sistemas de Equações

Sabemos que um sistema de m equações lineares em n variáveis é o mesmo que uma equação matricial

$$A\boldsymbol{x} = \boldsymbol{b},$$

onde A é a matriz coeficiente do sistema e $\boldsymbol{x}$ é o vetor de variáveis. Vamos tentar considerar sistemas de equações lineares usando funcionais lineares.

Seja $\underline{E} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ a base padrão de $\mathbb{K}^n$, e $\underline{E}^* = (\boldsymbol{e}_1^*, \ldots, \boldsymbol{e}_n^*)$ a base dual. Então, se f é um funcional linear em $(\mathbb{K}^n)^*$, temos que f pode ser escrito como

$$f = a_1\boldsymbol{e}_1^* + \cdots + a_n\boldsymbol{e}_n^*.$$

Podemos expressar $\boldsymbol{x} \in \mathbb{K}^n$ na base como $\boldsymbol{x} = x_1\boldsymbol{e}_1 + +x_n\boldsymbol{e}_n$, e assim

$$f(\boldsymbol{x}) = 0 \Leftrightarrow a_1x_1 + \cdots + a_nx_n = 0.$$

Portanto, $\ker f = \{\boldsymbol{x} \in \mathbb{K}^n | a_1x_1 + \cdots + a_nx_n = 0\}$.

Agora, suponha que $f_1, f_2, \ldots, f_m \in \mathbb{K}^{n*}$ sejam funcionais lineares. Então,

$$\bigcap_{i=1}^{m} \ker f_j = \{\boldsymbol{x} \in \mathbb{K}^n | f_i(\boldsymbol{x}) = 0, \text{ para todo } i\}.$$

Portanto, se $f_i = a_{i1}\boldsymbol{e}_1^* + \ldots + a_{in}\boldsymbol{e}_n^*$, então

$$\bigcap_{i=1}^{m} \ker f_i = \left\{ \boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{K}^n \,\middle|\, \begin{cases} a_{11}x_1 + \cdots + a_{n1}x_x & = 0 \\ \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = 0 \end{cases} \right\}$$

Provamos o seguinte teorema:

5.1 Teorema

Seja $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$. Defina os funcionais $f_i = a_{i1}\boldsymbol{e}_1^* + \ldots + a_{in}\boldsymbol{e}_n^*$ para $i = 1, \ldots, m$. Então

$$\{\boldsymbol{x} \in \mathbb{K}^n \mid A\boldsymbol{x} = \boldsymbol{0}\} = \bigcap_{i=1}^{m} \ker f_i$$

Portanto, traduzimos o problema de resolver um sistema de equações lineares para um problema envolvendo funcionais lineares: queremos tentar determinar $\bigcap_j \ker f_j$.

Iremos interpretar as operações elementares sobre as linhas nesta nova linguagem. Considere a lista dos funcionais $(f_1, \ldots, f_m)$

- trocar duas linhas é o mesmo que reordenar os funcionais f_j. Ou seja,

$$(f_1, \dots, f_i, \dots, f_j, \dots, f_m) \mapsto (f_1, \dots, f_j, \dots, f_i, \dots, f_m)$$

- multiplicar uma linha por $\lambda \in \mathbb{K}^n$ é o mesmo que considerar o funcional linear $\lambda f \in \mathbb{K}^{n*}$. Ou seja

$$(f_1, \dots, f_i, \dots, f_m) \mapsto (f_1, \dots, \lambda f_i, \dots, f_m)$$

- Da mesma forma, adicionar a linha i à linha j é o mesmo que adicionar f_i a f_j para obter o funcional linear $f_i + f_j$.

$$(f_1, \dots, f_i, \dots, f_j, \dots f_m) \mapsto (f_1, \dots, f_i, \dots, f_j + f_i, \dots, f_m)$$

Em resumo, executar operações elementares sobre as linhas é o mesmo que formar combinações lineares das funcionais lineares f_j.

O motivo pelo qual fazemos a redução em linha uma matriz A para uma escalonada reduzida U é porque os conjuntos de soluções de $A\boldsymbol{x} = 0$ e $U\boldsymbol{x} = 0$ são os mesmos (veja Proposição 5.4) e é mais fácil determinar soluções para a equação da matriz U. Como obtemos U aplicando operações elementares sobre as linhas a A, isso é o mesmo que fazer cálculos em $\langle\{f_j\}\rangle \subset \mathbb{K}^{n*}$, conforme discutimos acima.

Além disso, como U está na forma escalonada reduzida, isso significa que as linhas de U são linearmente independentes (isso é fácil de ver, pela definição de forma escalonada reduzida) e porque as operações elementares sobre as linhas correspondem à formação de combinações lineares de funcionais lineares temos que os funcionais lineares que correspondem às linhas de U devem formar uma base do subespaço $\langle\{f_j\}\rangle \subset \mathbb{K}^{n*}$.

5.2 Teorema

Seja $T \in \operatorname{Hom}(V, W)$. Suponha que $A = [T]_{\underline{E},\underline{F}} = [a_{ij}]$ é a matriz de T com relação às bases $\underline{E} = (\boldsymbol{e}_i)$ de V e $\underline{F} = (f_j)$ de W. Então,

$$\operatorname{posto}_l T = \operatorname{posto}_c T$$

Demonstração. O $\operatorname{posto}_l T$ é igual a $\dim\langle\{f_j\}\rangle$, com $f_j = a_{i1}\boldsymbol{e}_1^* + \cdots + a_{in}\boldsymbol{e}_n^* \in V^*$

Agora, temos que

$$\begin{aligned}\operatorname{Aniq}_V\langle\{f_j\}\rangle &= \{\boldsymbol{v} \in V \mid \overline{\boldsymbol{v}}(f_j) = 0, \text{ para todo } i\} \\ &= \{\boldsymbol{v} \in V \mid f_j(\boldsymbol{v}) = 0, \text{ para todo } i\},\end{aligned}$$

e esse último conjunto não é outro senão $\ker T$. Assim, pelo Teorema do Núcleo-imagem, temos que $\dim V = \dim \operatorname{im} T + \dim\ker T = \operatorname{posto} T + \dim \operatorname{Aniq}_{V^*}\langle\{f_j\}\rangle =$

$\operatorname{posto} T + (\dim V - \dim \langle \{f_j\} \rangle)$, usando a Proposição 5.2.1 na última igualdade. Temos

$$\operatorname{posto} T = \dim \langle \{f_j\}, \rangle$$

que é o que queríamos mostrar. ■

5.3 Proposição

Seja $A \in \mathcal{M}_{n,n}(\mathbb{K})$ e seja U sua forma escalonada reduzida. Então, $\boldsymbol{v}$ satisfaz $U\boldsymbol{v} = 0$ se, e somente se, satisfaz $A\boldsymbol{v} = 0$.

Demonstração. Sejam $f_1, \ldots, f_m \in \mathbb{K}^{n*}$ os funcionais lineares correspondentes às linhas de A e $g_1, \ldots, g_r$ os funcionais lineares correspondentes às linhas (diferentes de zero) de U (acabamos de ver que $r = \operatorname{posto} A$). Então, pelas discussões acima, sabemos que (g_j) é uma base de $W = \langle \{f_j\}_{j=1}^m \rangle \subset \mathbb{K}^{n*}$. Em particular, $\langle \{g_j\} \rangle = W$. Agora, temos

$$\operatorname{Aniq}_{\mathbb{K}^n} W = \{X \in \mathbb{K}^n | f_j(X) = 0, \text{ para todo } i\} = \{X \in \mathbb{K}^n | g_j(X) = 0, \text{ para todo } j\}.$$

O resultado segue essa igualdade de conjuntos, pois esse conjunto comum é o conjunto de soluções de $A\boldsymbol{v} = \boldsymbol{0}$ e $U\boldsymbol{v} = \boldsymbol{0}$. ■

Capítulo

6

Espaços com Produto Interno

Nesse capítulo

- Produto Interno (p. 165)
- Normas (p. 170)
- Bases Ortonormais (p. 176)
- Melhor Aproximação (p. 181)
- Projeção (p. 183)
- Adjuntas (p. 186)
- ♯Espaços de Hilbert (p. 190)

Começamos recordando a definição do produto escalar:

6.1 Definição Produto Escalar

Sejam $\boldsymbol{u} = [u_1, u_2, \cdots, u_n]^t, \boldsymbol{v} = [v_1, v_2, \cdots, v_n]^t$ dois vetores de $\mathbb{R}^n$. Então o produto escalar de $\boldsymbol{u}$ e $\boldsymbol{v}$ é dado por $\boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$.

É o produto escalar que permite introduzir noções como o comprimento (norma, magnitude) de um vetor, bem como o ângulo entre dois vetores.

Lembre-se de que se $\boldsymbol{u}$ e $\boldsymbol{v}$ são dois vetores diferentes de zero, então o ângulo $\theta(\boldsymbol{u}, \boldsymbol{v})$ entre $\boldsymbol{u}$ e $\boldsymbol{v}$ é dado por

$$\cos\theta(\boldsymbol{u}, \boldsymbol{v}) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|},$$

onde para um vetor $\boldsymbol{u} = (u_1, \dots, u_n) \in \mathbb{R}^n$, $\|u\|$ denota a norma do vetor $\boldsymbol{u}$, isto é,

$$\|\boldsymbol{u}\| = \sqrt{u_1^2 + \cdots + u_n^2} = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}.$$

Esta é a distância do ponto $(u_1, \dots, u_n) \in \mathbb{R}^n$ até a origem.

As propriedades básicas do produto escalar são enumeradas a seguir:

6.2 Teorema

Seja $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ vetores de $\mathbb{R}^n$ e λ em qualquer escalar. Em seguida, segure o seguinte:

1. $\boldsymbol{u} \cdot \boldsymbol{u} \geq 0$ e $\boldsymbol{u} \cdot \boldsymbol{u} = 0$ se, e somente se, $\boldsymbol{u} = \boldsymbol{0}$. Dizemos que o produto escalar é positivo definido.
2. $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$. Dizemos que o produto escalar é simétrico.
3. $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u}\boldsymbol{w} + \boldsymbol{v}\boldsymbol{w}$. Dizemos que o produto escalar é aditivo no primeiro argumento.

4 Para todos os $(\lambda \boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (\lambda \boldsymbol{v}) = \lambda(\boldsymbol{u}\ cdot \boldsymbol{v})$. Dizemos que o produto escalar é homogêneo em relação aos escalares.

Tomaremos as propriedades do produto escalar como base para nossa definição de um espaço com produto interno. Como a definição abrange espaços reais e complexos, as condições são ligeiramente modificadas. Em especial queremos que a norma seja um número real, isso se manifesta na propriedade de conjugação.

6.1 Produto Interno

Passamos agora a uma discussão de espaços vetoriais reais e complexos que possuem uma estrutura adicional, denominado produto interno, conforme descrito na definição a seguir. Os produtos internos permitem introduzir e definir diversos conceitos geométricos nos espaços vetoriais. Em espaços munidos de produtos internos podemos definir o comprimento de um vetor ou o ângulo entre dois vetores. Eles também fornecem os meios para definir a ortogonalidade entre vetores e mesmo a projeção num subespaço.

Neste capítulo, $\mathbb{K}$ denotará o corpo real ou complexo. Além disso, o complexo conjugado de $\lambda \in \mathbb{C}$ será denotado por $\overline{\lambda}$.

6.1 Definição Produto Interno

Se V é um espaço vetorial sobre $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Um **produto interno** em V é uma função $\langle\ ,\ \rangle : V \times V \to \mathbb{K}$ com as seguintes propriedades:

a **Linear na primeira coordenada**: para todos $\boldsymbol{u}, \boldsymbol{v} \in V$ e $\lambda_1, \lambda_2 \in \mathbb{K}$

$$\langle \lambda_1 \boldsymbol{u} + \lambda_2 \boldsymbol{v},\ \boldsymbol{w}\rangle = \lambda_1 \langle \boldsymbol{u},\ \boldsymbol{w}\rangle + \lambda_2 \langle \boldsymbol{v},\ \boldsymbol{w}\rangle$$

b **Simétrico**:

$$\langle \boldsymbol{u},\ \boldsymbol{v}\rangle = \overline{\langle \boldsymbol{v}, \boldsymbol{u}\rangle} \quad \text{(simetria Hermitiana)}$$

c **Positivo definido**: Para todo $\boldsymbol{v} \in V$,

$$\langle \boldsymbol{v}, \boldsymbol{v}\rangle \geq 0 \text{ e } \langle \boldsymbol{v}, \boldsymbol{v}\rangle = 0 \text{ se, e somente se, } \boldsymbol{v} = 0.$$

Quando $\mathbb{K} = \mathbb{R}$ a condição de simetria se reduz a

$$\langle \boldsymbol{u},\ \boldsymbol{v}\rangle = \langle \boldsymbol{v},\ \boldsymbol{u}\rangle \quad \text{(simetria)}\ .$$

Um espaço vetorial V, munido de um produto interno, é denominado espaço com produto interno.

Observe que um subespaço vetorial S de um espaço com produto interno V também é um espaço com produto interno com a restrição do produto interno de V a S.

6.2 Proposição

a) Se $\mathbb{K} = \mathbb{R}$, então o produto interno é linear em ambas as coordenadas, ou seja, o produto interno é bilinear.

b) No entanto, se $\mathbb{K} = \mathbb{C}$, então o produto interno é linear na primeira coordenada e linear conjugado na segunda, isto é

$$\langle \boldsymbol{w}, \lambda_1 \boldsymbol{u} + \lambda_2 \boldsymbol{v} \rangle = \overline{\lambda_1} \langle \boldsymbol{w},\ \boldsymbol{u} \rangle + \overline{\lambda_1} \langle \boldsymbol{w},\ \boldsymbol{v} \rangle$$

c) $\langle \mathbf{0}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \mathbf{0} \rangle = 0.$

Demonstração. Vamos provar apenas b). Nesse caso $\mathbb{K} = \mathbb{C}$, e

$$\langle \boldsymbol{w}, \lambda_1 \boldsymbol{u} + \lambda_2 \boldsymbol{v} \rangle = \overline{\langle \lambda_1 \boldsymbol{u} + \lambda_2 \boldsymbol{v}, \boldsymbol{w} \rangle} = \overline{\lambda_1} \overline{\langle \boldsymbol{u}, \boldsymbol{w} \rangle} + \overline{\lambda_2} \overline{\langle \boldsymbol{v}, \boldsymbol{w} \rangle} = \overline{\lambda_1} \langle \boldsymbol{w},\ \boldsymbol{u} \rangle + \overline{\lambda_1} \langle \boldsymbol{w},\ \boldsymbol{v} \rangle$$

■

Assim, um produto interno complexo é linear na primeira coordenada e linear conjugado na segunda coordenada. Nesse caso dizemos que um produto interno complexo é sesquilinear (sesqui significa uma vez e meia).

Exemplo 3. *O espaço vetorial $\mathbb{R}^n$ é um espaço com produto interno com o produto escalar, definido por*

$$\langle (x_1,\ \ldots,\ x_n),\ (y_1,\ \ldots,\ y_n) \rangle = x_1 y_1 + \cdots + x_n y_n$$

O espaço com produto interno $\mathbb{R}^n$ é geralmente denominado espaço euclidiano n-dimensional.

◁

Exemplo 4. *O espaço vetorial $\mathbb{C}^n$ é um espaço com produto interno com o produto escalar definido por*

$$\langle (x_1,\ \ldots,\ x_n),\ (x_1,\ \ldots,\ x_n) \rangle = x_1 \overline{y}_1 + \cdots + x_n \overline{y}_n$$

Esse espaço com produto interno costuma ser denominado espaço hermitiano n-dimensional.

◁

Exemplo 5. *Se o espaço vetorial V possui uma base $\underline{B} = \{\boldsymbol{v}_1,\ \ldots,\ \boldsymbol{v}_n\}$, então*

$$\langle \boldsymbol{x},\ \boldsymbol{y} \rangle = [\boldsymbol{x}]_{\underline{B}}^T [\overline{\boldsymbol{y}}]_{\underline{B}}$$

define um produto interno em V.

$\triangleleft$

Exemplo 6. *Para* $A, B \in \mathcal{M}_{n,n}(\mathbb{K})$ *definimos*

$$\langle A,\ B\rangle = \mathrm{tr}(A^{\mathrm{t}}\overline{B})\ .$$

Então $\langle\cdot,\cdot\rangle$ *é um produto interno em* $\mathcal{M}_{n,n}(\mathbb{K})$ *conhecido como produto interno da Frobenius.*

$\triangleleft$

Exemplo 7. *O espaço vetorial* $C[a,\ b]$ *de todas as funções contínuas reais no intervalo fechado* $[a,\ b]$ *é um espaço de produto interno com*

$$\langle f,\ g\rangle = \int_a^b f(x)g(x)dx$$

Vamos provar apenas que é positivo definido.

Ou seja, vamos provar que se

- f *é contínuo no intervalo fechado* $[a, b]$.
- $f \geq 0$ *em* (a, b).

Então, $f = 0$ *em* $[a, b]$ *ou* $\displaystyle\int_a^b f(t)dt > 0$.

Sejam m *e* M *os valores mínimo e máximo de* f *no intervalo* $[a, b]$*. Portanto, temos* $0 \leq m \leq f(t) \leq M$ *para todos os* t *em* $[a, b]$.

Se $M = 0$*, então* f *é identicamente nula.*

Então agora assuma que $M > 0$*. Se* $m = M$*, então* $f(t) = M$ *para todos os* t *e* $\displaystyle\int_a^b f(t)dt = M(b-a) > 0$ *e terminamos a demonstração.*

Assim podemos assumir que $m < M$*. Então o número* $A = \frac{1}{2}(M+m)$ *é positivo e estritamente menor que* M*. Seja* $f(c) = M$*, com* c *em* $[a, b]$*. Como* f *é contínua em* $[a, b]$*, temos que* f *é maior que* A *se estivermos perto o suficiente de* $t = c$*. Logo existe uma vizinhança* $[p, q]$ *contendo* c*, de modo que* $f(t) \geq A$ *nesse intervalo, onde temos:* $a \leq p < q \leq b$ *e* $p \leq c \leq q$*. Então, como* $f(t) \geq 0$ *em* $[a, b]$*, temos* $\displaystyle\int_a^p f(t)dt \geq 0$ *e* $\displaystyle\int_q^b f(t)dt \geq 0$*, então obtemos:*

$$\begin{aligned}\int_a^b f(t)dt &= \int_a^p f(t)dt + \int_p^q f(t)dt + \int_q^b f(t)dt\\ &\geq \int_p^q f(t)dt \geq \int_p^q Adt = A(q-p) > 0.\end{aligned}$$

Então $\displaystyle\int_a^b f(t)dt > 0$ *e obtemos o resultado desejado.*

$\triangleleft$

Exemplo 8. *O espaço vetorial $C^1[a,\ b]$ de todas as funções com a primeira derivada contínua de valor complexo no intervalo fechado $[a,\ b]$ é um espaço complexo de produto interno com*

$$\langle f,\ g\rangle = \int_a^b f(x)\overline{g(x)} + f'(x)\overline{g'(x)}dx$$

$\triangleleft$

Exemplo 9. *No espaço vetorial real $\mathcal{M}_{n,n}(\mathbb{R})$ o produto interno usual é definido da seguinte forma:*

$$\langle A, B\rangle = \operatorname{tr}(B^t A)$$

$\triangleleft$

Exemplo 10 -ℓ^2. *Um dos exemplos de espaços de produto interno mais fundamentais é o espaço vetorial ℓ^2 de todas as sequências reais (ou complexas) (s_n) com a propriedade que*

$$\sum |s_n|^2 < \infty$$

munido do produto interno

$$\langle (s_n),\ (t_n)\rangle = \sum_{n=0}^{\infty} s_n\overline{t_n} \tag{6.1}$$

Tais sequências são denominadas quadrado somáveis. Para que o produto interno esteja bem definido a soma à direita na Equação 6.1 deve convergir. Para provar isso, observamos que, se $(s_n), (t_n) \in \ell^2$, então

$$0 \leq (|s_n| - |t_n|)^2 = |s_n|^2 - 2|s_n||t_n| + |t_n|^2$$

e consequentemente

$$2|s_n t_n| \leq |s_n|^2 + |t_n|^2$$

o que implica que $(s_n t_n) \in \ell^2$. Deixamos ao leitor verificar que ℓ^2 é um espaço com produto interno.

$\triangleleft$

Exemplo 11 -Matriz de Gram. *Seja V um espaço vetorial de dimensão finita e base $\underline{B} = \{\boldsymbol{v}_1,\ \dots,\ \boldsymbol{v}_n\}$,. Sejam $\boldsymbol{v} = \sum_{i=1}^n a_i\boldsymbol{v}_i$ e $\boldsymbol{u} = \sum_{j=1}^n b_j\boldsymbol{v}_j$ vetores em V. Então*

$$\langle \boldsymbol{v}, \boldsymbol{u}\rangle = \sum_{i=1}^n a_i\bar{b}_j\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \begin{bmatrix} a_1, \cdots, a_n \end{bmatrix} G \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

onde $G = [g_{ij}]$ é a matriz definida por $g_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle$ para $1 \leq i, j \leq n$. A matriz G é conhecida como matriz de Gram na base $\underline{B}$.

$\triangleleft$

6.12 Definição Ângulo

Dados $\boldsymbol{u}, \boldsymbol{v} \in V$ dois vetores não nulos. O **ângulo** entre $\boldsymbol{u}$ e $\boldsymbol{v}$ é definido como

$$] \cos\theta(\boldsymbol{u}, \boldsymbol{v}) = \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}.$$

Figura 6.1 Ângulo entre vetores

6.13 Definição Vetores Ortogonais

Sejam $\boldsymbol{u}, \boldsymbol{v}$ vetores no espaço com produto interno V. Esses vetores são ditos **ortogonais** ou **perpendiculares** se $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$. Nesse caso escrevemos $\boldsymbol{u} \perp \boldsymbol{v}$.

Figura 6.2 Vetores ortogonais

O seguinte resultado simples é bastante útil.

6.14 Proposição

Se V é um espaço com produto interno $\langle \boldsymbol{u}, \boldsymbol{x} \rangle = \langle \boldsymbol{v}, \boldsymbol{x} \rangle$ para todo $\boldsymbol{x} \in V$, então $\boldsymbol{u} = \boldsymbol{v}$.

6.2 Normas

Se V for um espaço com produto interno, a **norma** ou **comprimento** (associada ao produto interno) de $\boldsymbol{v} \in V$ é definida como

$$\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$$

Um vetor $\boldsymbol{v}$ é dito **unitário** se $\|\boldsymbol{v}\| = 1$.

6.1 Teorema

a $\|\boldsymbol{v}\| \geq 0$ e $\|\boldsymbol{v}\| = 0$ se, e somente se, $\boldsymbol{v} = 0$.

b Para todos $\lambda \in \mathbb{K}$ e $\boldsymbol{v} \in V$,

$$\|\lambda \boldsymbol{v}\| = |\lambda| \|\boldsymbol{v}\|$$

c (Pitágoras) Se $\boldsymbol{v} \perp \boldsymbol{u}$, temos

$$\|\boldsymbol{v} + \boldsymbol{u}\|^2 = \|\boldsymbol{v}\|^2 + \|\boldsymbol{u}\|^2$$

d (Desigualdade de Cauchy-Schwarz-Bunyakovsky) Para todos $\boldsymbol{u}, \boldsymbol{v} \in V$,

$$|\langle \boldsymbol{u},\ \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\| \|\boldsymbol{v}\|$$

com igualdade se, e somente se, $\boldsymbol{u}$ e $\boldsymbol{v}$ forem múltiplo escalar um do do outro.

e (Desigualdade triangular) Para todos $\boldsymbol{u}, \boldsymbol{v} \in V$,

$$\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$$

com igualdade se, e somente se, $\boldsymbol{u}$ e $\boldsymbol{v}$ forem um múltiplo escalar um do do outro.

f Para todos $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$,

$$\|\boldsymbol{u} - \boldsymbol{v}\| \leq \|\boldsymbol{u} - \boldsymbol{w}\| + \|\boldsymbol{w} - \boldsymbol{v}\|$$

g Para todos $\boldsymbol{u}, \boldsymbol{v} \in V$,

$$|\|\boldsymbol{u}\| - \|\boldsymbol{v}\|| \leq \|\boldsymbol{u} - \boldsymbol{v}\|$$

h (Regra do Paralelogramo) Para todos $\boldsymbol{u}, \boldsymbol{v} \in V$,

$$\|\boldsymbol{u} + \boldsymbol{v}\|^2 + \|\boldsymbol{u} - \boldsymbol{v}\|^2 = 2\|\boldsymbol{u}\|^2 + 2\|\boldsymbol{v}\|^2$$

Demonstração. Para demonstrar Pitágoras observamos que

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

pois $\boldsymbol{v}$ e $\boldsymbol{u}$ são ortogonais.

Para demonstrar Cauchy-Schwarz-Bunyakovsky, se $\boldsymbol{u}$ ou $\boldsymbol{v}$ for o vetor zero, o resultado é imediato, assim podemos assumir que $\boldsymbol{u}, \boldsymbol{v} \neq 0$. Então, para qualquer $\lambda \in \mathbb{K}$,

$$0 \leq \|\boldsymbol{u}-\lambda\boldsymbol{v}\|^2 \quad = \langle \boldsymbol{u}-\lambda\boldsymbol{v},\ \boldsymbol{u}-\lambda\boldsymbol{v}\rangle \ = \langle \boldsymbol{u},\ \boldsymbol{u}\rangle - \overline{\lambda}\langle \boldsymbol{u},\ \boldsymbol{v}\rangle - \lambda[\langle \boldsymbol{v},\ \boldsymbol{u}\rangle - \overline{\lambda}\langle \boldsymbol{v},\ \boldsymbol{v}\rangle]$$

Escolhendo $\overline{\lambda} = \langle \boldsymbol{v}, \boldsymbol{u}\rangle / \langle \boldsymbol{v}, \boldsymbol{v}\rangle$ fazemos a expressão entre colchetes se anular e logo

$$0 \leq \langle \boldsymbol{u},\ \boldsymbol{u}\rangle - \frac{\langle \boldsymbol{v}, \boldsymbol{u}\rangle\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\langle \boldsymbol{v}, \boldsymbol{v}\rangle} = \|\boldsymbol{u}\|^2 - \frac{|\langle \boldsymbol{u}, \boldsymbol{v}\rangle|^2}{\|\boldsymbol{v}\|^2}$$

o que equivale à desigualdade de Cauchy-Schwarz-Bunyakovsky. Além disso, a igualdade vale se, e somente se, $\|\boldsymbol{u}-\lambda\boldsymbol{v}\|^2 = 0$, isto é, se, e somente se, $\boldsymbol{u}-\lambda\boldsymbol{v} = 0$, o que equivale a $\boldsymbol{u}$ e $\boldsymbol{v}$ serem múltiplos escalares um do outro.

Para provar a desigualdade triangular, utilizaremos a desigualdade de Cauchy--Schwarz-Bunyakovsky

$$\begin{aligned}
\|\boldsymbol{u}+\boldsymbol{v}\|^2 &= \langle \boldsymbol{u}+\boldsymbol{v},\ \boldsymbol{u}+\boldsymbol{v}\rangle \\
&= \langle \boldsymbol{u},\ \boldsymbol{u}\rangle + \langle \boldsymbol{u},\ \boldsymbol{v}\rangle + \langle \boldsymbol{v},\ \boldsymbol{u}\rangle + \langle \boldsymbol{v},\ \boldsymbol{v}\rangle \\
&\leq \|\boldsymbol{u}\|^2 + 2\|\boldsymbol{u}\|\|\boldsymbol{v}\| + \|\boldsymbol{v}\|^2 \\
&= (\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)^2
\end{aligned}$$

e assim temos a desigualdade triangular. ■

Podemos generalizar o conceito de norma de modo a torna-lo independente do produto interno.

6.2 Definição Norma

Seja V um espaço vetorial sobre o corpo $\mathbb{K}$. Uma **norma** em V é uma aplicação $\|\ \| : V \to [0,\ \infty)$ satisfazendo às seguintes propriedades:

- *a* $\|\boldsymbol{v}\| > 0$ se $\boldsymbol{v} \neq 0$;
- *b* $\|\lambda\boldsymbol{v}\| = |\lambda|\|\boldsymbol{v}\|$, para $\lambda \in \mathbb{K}$;
- *c* $\|\boldsymbol{v}+\boldsymbol{u}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{u}\|$.

Se V possui uma norma, dizemos que V é um espaço normado.

Exemplo 3. *Seja V um espaço com produto interno. Então $\|\boldsymbol{v}\| := \langle \boldsymbol{v}, \boldsymbol{v}\rangle^{1/2}$ r define uma norma em V.*

$\triangleleft$

Exemplos 4.

1. *No espaço vetorial real* $\mathbb{R}^n$ *temos as seguintes normas*

   a) Norma do Máximo: $\|x\|_\infty = \max\{|x_i| \; ; \; 1 \leq i \leq n\}$

   b) Norma do Táxi: $\|x\|_1 = \sum_{i=1}^{n} |x_i|$

2. *No espaço vetorial real* $\mathcal{M}_{n,n}(\mathbb{K})(\mathbb{R})$. *temos as seguintes normas*

   a) $\|A\|_\infty = \max\{\sum_{j=1}^{n} |a_{ij}| \; ; \; 1 \leq i \leq n\}$

   b) $\|A\|_1 = \max\{\sum_{i=1}^{n} |a_{ij}| \; ; \; 1 \leq j \leq n\}$

$\triangleleft$

Observamos que nem todas as normas provem de um produto interno.

É interessante observar que o produto interno em V pode ser recuperado da norma associada. Assim, conhecer o comprimento de todos os vetores em V é equivalente a conhecer todos os produtos internos dos vetores em V.

6.5 Teorema Identidades de Polarização

a Se V é um espaço real com produto interno, então

$$\langle \boldsymbol{u},\ \boldsymbol{v}\rangle = \frac{1}{4}(\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2)$$

b Se V é um espaço complexo com produto interno, então

$$\langle \boldsymbol{u},\ \boldsymbol{v}\rangle = \frac{1}{4}(\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2) + \frac{1}{4}i(\|\boldsymbol{u}+i\boldsymbol{v}\|^2 - \|\boldsymbol{u}-i\boldsymbol{v}\|^2)$$

Demonstração. Primeiramente observamos que $\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2 = 4\,\mathrm{Re}\langle \boldsymbol{u},\boldsymbol{v}\rangle$ e

$$\begin{aligned} i\|\boldsymbol{u}+i\boldsymbol{v}\|^2 - i\|\boldsymbol{u}-i\boldsymbol{v}\|^2 &= i(\langle \boldsymbol{u}+i\boldsymbol{v},\boldsymbol{u}+i\boldsymbol{v}\rangle - \langle \boldsymbol{u}-i\boldsymbol{v},\boldsymbol{u}-i\boldsymbol{v}\rangle) \quad = i(\langle \boldsymbol{u},\boldsymbol{u}\rangle + \langle \boldsymbol{u},i\boldsymbol{v}\rangle + \langle i\boldsymbol{v},\boldsymbol{u}\rangle + \langle i\boldsymbol{v},i\boldsymbol{v}\rangle - \langle \boldsymbol{u},\boldsymbol{u}\rangle + \langle \boldsymbol{u},i\boldsymbol{v}\rangle \\ &= 2i(\langle \boldsymbol{u},i\boldsymbol{v}\rangle \\ &+ \langle i\boldsymbol{v},\boldsymbol{u}\rangle) \\ &= 2i(-i\langle \boldsymbol{u},\boldsymbol{v}\rangle + i\langle \boldsymbol{v},\boldsymbol{u}\rangle) \qquad = - \end{aligned}$$

Logo

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2 + i\|\boldsymbol{u}+i\boldsymbol{v}\|^2 - i\|\boldsymbol{u}-i\boldsymbol{v}\|^2 = 4\,\mathrm{Re}\langle \boldsymbol{u},\boldsymbol{v}\rangle + 4\,\mathrm{Im}\langle \boldsymbol{u},\boldsymbol{v}\rangle = 4\langle \boldsymbol{u},\boldsymbol{v}\rangle$$

■

6.6 Teorema von Neumann

Seja V um espaço vetorial. Uma norma em V é induzida por um produto interno em V se, e somente se, satisfizer a identidade de polarização. Além disso. se uma norma em V satisfaz a a identidade de polarização, então o produto interno único que a induz é dado pela identidade da polarização.

Demonstração. Provaremos o caso real. Começaremos definindo o produto interno sugerido pela identidade de polarização

$$\langle \boldsymbol{x},\ \boldsymbol{y}\rangle = \frac{1}{4}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2\right)\ \forall\, \boldsymbol{x},\boldsymbol{y}\in V\ .$$

Devemos mostrar que é um produto interno.

Provaremos apenas a linearidade $\langle \lambda\boldsymbol{x}+\boldsymbol{z},\boldsymbol{y}\rangle = \lambda\langle \boldsymbol{x},\boldsymbol{y}\rangle + \langle \boldsymbol{z},\boldsymbol{y}\rangle,\quad \boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in V$ deixando as outras propriedades como exercício ao leitor. Começaremos demonstrando que $\langle -\boldsymbol{x},\boldsymbol{y}\rangle = -\langle \boldsymbol{x},\boldsymbol{y}\rangle$

$$\langle -\boldsymbol{x},\boldsymbol{y}\rangle = \frac{1}{4}\left(\|-\boldsymbol{x}+\boldsymbol{y}\|^2 - \|-\boldsymbol{x}-\boldsymbol{y}\|^2\right) = \frac{1}{4}\left(\|\boldsymbol{x}-\boldsymbol{y}\|^2 - \|\boldsymbol{x}+\boldsymbol{y}\|^2\right) = -\langle \boldsymbol{x},\boldsymbol{y}\rangle$$

Agora usando a polarização temos:

$$
\begin{aligned}
\langle \boldsymbol{x},\ \boldsymbol{y}\rangle &= \frac{1}{4}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2\right)\\
&= \frac{1}{2}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\right)\\
&= \frac{1}{2}\left(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2\right) \forall\, \boldsymbol{x},\boldsymbol{y} \in V
\end{aligned}
$$

Tomando $\lambda \geq 0$ temos $\langle \lambda\boldsymbol{x}, \boldsymbol{y}\rangle - \lambda\langle \boldsymbol{x}, \boldsymbol{y}\rangle \leq 0$

$$\langle \lambda\boldsymbol{x}, \boldsymbol{y}\rangle - \lambda\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \frac{1}{2}\left(\|\lambda\boldsymbol{x}+\boldsymbol{y}\|^2 - \lambda^2\|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\right.$$

$$\leq \frac{1}{2}\left((\lambda\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2 - \lambda^2\|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2 + \lambda\|\boldsymbol{x}-\boldsymbol{y}\|^2 - \lambda\|\boldsymbol{x}\|^2 - \lambda\|\boldsymbol{y}\|^2\right)$$

$$= \frac{1}{2}\lambda\left(2\|\boldsymbol{x}\|\|\boldsymbol{y}\| + \|\boldsymbol{x}-\boldsymbol{y}\|^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\right)[4pt] = \frac{1}{2}\lambda\left(\|\boldsymbol{x}-\boldsymbol{y}\|^2 - (\|\boldsymbol{x}\| - \|\boldsymbol{y}\|)^2\right)[4pt] \leq 0$$

Como $(\|\boldsymbol{x}\| - \|\boldsymbol{y}\|)^2 \geq \|\boldsymbol{x}-\boldsymbol{y}\|^2$

$$
\begin{aligned}
0 &\geq \langle \lambda\boldsymbol{x}, -\boldsymbol{y}\rangle - \lambda\langle \boldsymbol{x}, -\boldsymbol{y}\rangle\\
&= \langle -\lambda\boldsymbol{x}, \boldsymbol{y}\rangle - \lambda\langle -\boldsymbol{x}, \boldsymbol{y}\rangle\\
&= -\left(\langle \lambda\boldsymbol{x}, \boldsymbol{y}\rangle - \lambda\langle \boldsymbol{x}, \boldsymbol{y}\rangle\right)
\end{aligned}
$$

Assim $\langle \lambda\boldsymbol{x}, \boldsymbol{y}\rangle - \lambda\langle \boldsymbol{x}, \boldsymbol{y}\rangle \geq 0$. Logo

$$\langle \lambda\boldsymbol{x}, \boldsymbol{y}\rangle - \lambda\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0 \Rightarrow \langle \lambda\boldsymbol{x}, \boldsymbol{y}\rangle = \lambda\langle \boldsymbol{x}, \boldsymbol{y}\rangle$$

Se $\lambda < 0$ então fazendo $\beta = -\lambda$ (com $\beta > 0$):

$$\langle \lambda\boldsymbol{x}, \boldsymbol{y}\rangle = \langle -\beta\boldsymbol{x}, \boldsymbol{y}\rangle = -\langle \beta\boldsymbol{x}, \boldsymbol{y}\rangle = -\beta\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \lambda\langle \boldsymbol{x}, \boldsymbol{y}\rangle$$

Agora podemos mostrar que $\langle \boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}\rangle - \langle \boldsymbol{x}, \boldsymbol{y}\rangle - \langle \boldsymbol{z}, \boldsymbol{y}\rangle \leq 0$

$$
\begin{gathered}
\langle \boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}\rangle - \langle \boldsymbol{x}, \boldsymbol{y}\rangle - \langle \boldsymbol{z}, \boldsymbol{y}\rangle =\\
= \frac{1}{2}\left(\|\boldsymbol{x}+\boldsymbol{y}+\boldsymbol{z}\|^2 - \|\boldsymbol{x}+\boldsymbol{z}\|^2 - \|\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2 + \|\boldsymbol{z}-\boldsymbol{y}\|^2 - \|\boldsymbol{z}\|^2 - \|\boldsymbol{y}\|^2\right)\\
\leq \frac{1}{2}\left((\|\boldsymbol{x}+\boldsymbol{z}\| + \|\boldsymbol{y}\|)^2 - \|\boldsymbol{x}+\boldsymbol{z}\|^2 - \|\boldsymbol{y}\|^2 + (\|\boldsymbol{x}\| - \|\boldsymbol{y}\|)^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2 + (\|\boldsymbol{z}\| - \|\boldsymbol{y}\|)^2 - \|\boldsymbol{z}\|^2 - \|\boldsymbol{y}\|^2\right)\\
= \|\boldsymbol{y}\|\left(\|\boldsymbol{x}+\boldsymbol{z}\| - \|\boldsymbol{x}\| - \|\boldsymbol{z}\|\right)\\
\leq 0
\end{gathered}
$$

Pois $\|\boldsymbol{x}+\boldsymbol{z}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{z}\|$

Novamente, considerando a relação recém-encontrada com $\boldsymbol{x}, \boldsymbol{z}, -\boldsymbol{y}$ no lugar de $\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{y}$, podemos provar que

$$\langle \boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}\rangle - \langle \boldsymbol{x}, \boldsymbol{y}\rangle - \langle \boldsymbol{z}, \boldsymbol{y}\rangle \geq 0.$$

Logo $\langle \boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle + \langle \boldsymbol{z}, \boldsymbol{y}\rangle$

Combinando as equações que encontramos, podemos concluir que

$$\langle \lambda \boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} \rangle = \langle \lambda \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{z}, \boldsymbol{y} \rangle = \lambda \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{z}, \boldsymbol{y} \rangle, \quad \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$$

Logo $\langle \cdot, \cdot \rangle$ é linear.

Agora demonstraremos que $\langle \cdot, \cdot \rangle$ induz $\| \cdot \|$

$$\sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \sqrt{\frac{1}{4} \left(\|\boldsymbol{x} + \boldsymbol{x}\|^2 + \|\boldsymbol{x} - \boldsymbol{x}\|^2 \right)} = \sqrt{\frac{1}{4} \|2\boldsymbol{x}\|^2} = \|\boldsymbol{x}\|$$

■

A norma pode ser usada para definir a distância entre dois vetores em um espaço com produto interno .

6.7 Definição Distância

Se V é um espaço com produto interno. Definimos a **distância** $d(\boldsymbol{u},\ \boldsymbol{v})$ entre os vetores $\boldsymbol{u}$ e $\boldsymbol{v}$ em V como

$$d(\boldsymbol{u},\ \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|$$

6.8 Proposição

Se V é um espaço munido da distância $d(\boldsymbol{u},\ \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|$ então:

- *a* $d(\boldsymbol{u},\ \boldsymbol{v}) \geq 0$ e $d(\boldsymbol{u},\ \boldsymbol{v}) = 0$ se, e somente se, $\boldsymbol{u} = \boldsymbol{v}$.
- *b* $d(\boldsymbol{u},\ \boldsymbol{v}) = d(\boldsymbol{v},\ \boldsymbol{u})$. Nesse caso dizemos que d é simétrica.
- *c* $d(\boldsymbol{u},\ \boldsymbol{v}) \leq d(\boldsymbol{u},\ \boldsymbol{w}) + d(\boldsymbol{w},\ \boldsymbol{v})$. Nesse caso dizemos que d satisfaz a desigualdade triangular.

Qualquer conjunto não vazio V, juntamente com uma função $d : V \times V \to \mathbb{R}$ que satisfaça as propriedades da Proposição 6.2, é denominado **espaço métrico** e a função d é dita **métrica** em V. Assim, qualquer espaço com produto interno é um espaço métrico com a métrica induzida pelo produto interno.

A presença de um produto interno e, portanto, de uma métrica, permite a definição de diversos conceitos analíticos em V como convergência de sequências e séries infinitas, funções contínuas e conjuntos abertos, fechados e compactos.

6.3 Bases Ortonormais

6.1 Definição Ortogonalização

Seja V um espaço com produto interno. Um conjunto não vazio $O = \{\boldsymbol{u}_i, i \in \Delta\} \subset V$ é **ortogonal** se $\boldsymbol{u}_i \perp \boldsymbol{u}_j$ para quaisquer $\boldsymbol{u}_i, \boldsymbol{u}_j \in X$. Se, além disso, todos os seus vetores são unitários, então O é **ortonormal**. Ou seja, O é ortonormal se

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \delta_{i,j}$$

Para todo $i, j \in \Delta$.

6.2 Definição Subespaço Ortogonal

Seja V um espaço com produto interno e $A \subset V$. Então definimos o **subespaço ortogonal** a A como

$$A^{\perp} = \{\boldsymbol{v} \in V \text{ tais que } \boldsymbol{v} \perp \boldsymbol{a}, \forall \boldsymbol{a} \in A\}$$

6.3 Lema

O conjunto $S^{\perp}$ é um subespaço de V, mesmo que S não o seja. Se S é um subespaço de V então $S \cap S^{\perp} = \{\mathbf{0}\}$.

Exemplo 4. *Seja $C[-1,\ 1]$ o espaço vetorial de todas as funções contínuas reais no intervalo fechado $[-1,\ 1]$ munido do produto interno*

$$\langle f,\ g\rangle = \int_a^b f(x)\overline{g(x)}dx$$

Então o conjunto $\{\operatorname{sen}(n\pi x)\}_{n=1}^{\infty}$ é ortonornal.

$$\int_{-1}^{1} \operatorname{sen}(n\pi x)\operatorname{sen}(m\pi x)dx = \begin{cases} 0 & \text{se } n \neq m \\ 1 & \text{se } m = n \end{cases}$$

$\mathbf{m} = \mathbf{n}$:

$$\int_{-1}^{1} \frac{1}{2}\left(\cos(n\pi x - n\pi x) - \cos(n\pi x + n\pi x)\right) \tag{6.2}$$

$$= \frac{1}{2}\int_{-1}^{1} \cos(0) - \frac{1}{2}\int_{-1}^{1} \cos(2n\pi x)dx \tag{6.3}$$

$$= \frac{1}{2}[1-(-1)] - \frac{1}{2}[0-0] \tag{6.4}$$

$$= 1 \tag{6.5}$$

$\mathbf{m} \neq \mathbf{n}$:

$$\int_{-1}^{1} \frac{1}{2}\left(\cos(m\pi x - n\pi x) - \cos(m\pi x + n\pi x)\right) dx \tag{6.6}$$

$$= \frac{1}{2}\int_{-1}^{1} \cos((m-n)\pi x)dx - \frac{1}{2}\int_{-1}^{1} \cos((m+n)\pi x)dx \tag{6.7}$$

$$= 0 \tag{6.8}$$

◁

A ortogonalidade é mais forte que a independência linear.

6.5 Lema

Todo conjunto ortogonal formado por vetores não nulos é linearmente independente.

Demonstração. Sejam $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in X$ tais que

$$\alpha_1\boldsymbol{x}_1 + \ldots + \alpha_m\boldsymbol{x}_m = \mathbf{0}.$$

Então

$$0 = \langle \mathbf{0}, \boldsymbol{x}_i\rangle = \langle \alpha_1\boldsymbol{x}_1 + \ldots + \alpha_m\boldsymbol{x}_m, \boldsymbol{x}_i\rangle = \alpha_1\langle \boldsymbol{x}_1, \boldsymbol{x}_i\rangle + \ldots + \alpha_m\langle \boldsymbol{x}_m, \boldsymbol{x}_i\rangle = \alpha_i\langle \boldsymbol{x}_i, \boldsymbol{x}_i\rangle.$$

Como $\langle \boldsymbol{x}_i, \boldsymbol{x}_i\rangle = ||\boldsymbol{x}_i||^2 \neq 0$, temos que $\alpha_i = 0$. ■

Seja O a coleção de todos os conjuntos ortonormais em um espaço com produto interno V. Observe que qualquer conjunto unitário $\{\boldsymbol{v}\}$ com $||\boldsymbol{v}|| = 1$ é um conjunto ortonormal em V. Como uma subfamília do conjunto das partes $\mathscr{P}(V)$ a família O é parcialmente ordenada pela ordem dada pela inclusão de conjuntos.

6.6 Definição Base de Hilbert

Um conjunto ortonormal maximal A em um espaço com produto interno V é chamado de **base de Hilbert** para V.

Se V não for um espaço vetorial de dimensão finita, uma base no sentido acima nunca será uma base de V como espaço vetorial. Para diferenciar denominaremos uma base de V no sentido de espaço vetorial como base de Hamel.

6.7 Teorema

Seja V um espaço com produto interno e A seja um conjunto ortonormal em V. As seguintes asserções são equivalentes

$\boxed{a}$ A é um conjunto ortonormal maximal em V.

$\boxed{b}$ Não há vetor unitário $\boldsymbol{v}$ para o qual $A \bigcup \{\boldsymbol{v}\}$ seja um conjunto ortonormal.

$\boxed{c}$ Se $\boldsymbol{v} \perp A$, $\boldsymbol{v} = 0$ (isto é, $A^{\perp} = \{\mathbf{0}\}$).

Demonstração. Seja A um conjunto ortonormal em um espaço com produto interno V.

$\boxed{a} \Rightarrow \boxed{b}$ Se existe um vetor unitário $\boldsymbol{v}$ em V para o qual $A \bigcup \{\boldsymbol{v}\}$ é um conjunto ortonormal, $A \bigcup \{\boldsymbol{v}\}$ é um conjunto ortonormal que inclui propriamente A (pois $\boldsymbol{v} \perp A$). Portanto, A não é um conjunto ortonormal maximal em V.

$\boxed{b} \Rightarrow \boxed{c}$ Se houver um vetor diferente de zero $\boldsymbol{v}$ em V, de modo que $\boldsymbol{v} \perp A$, existe um vetor unitário $\boldsymbol{v}' = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}$ em V de modo que $A \bigcup \{\boldsymbol{v}'\}$ é um conjunto ortonormal.

$\boxed{c} \Rightarrow \boxed{a}$ Se $\boxed{a}$ não for verdadeiro, então existe um conjunto ortonormal A' em V que contém propriamente A de modo que $A' \backslash A \neq \emptyset$. Seja $\boldsymbol{v}$ em $A' \backslash A$, que é um vetor diferente de zero (na verdade, $\boldsymbol{v}$ é um vetor unitário) ortogonal a A (porque $\boldsymbol{v} \in A'$, $A \subset A'$ e A' é um conjunto ortonormal). Assim $\boxed{c}$ não é verdadeiro.

∎

O Lema de Zorn pode ser usado para mostrar que qualquer espaço com produto interno possui uma base de Hilbert.

6.8 Teorema

> Se A é um conjunto ortonormal em um espaço com produto interno V então existe uma base de Hilbert B em V tal que $A \subseteq B$.

Demonstração. Seja A um conjunto ortonormal em um espaço com produto interno V. Definimos

$$O_A = \{S \in \mathscr{P}(V) : S \text{ é um conjunto ortonormal em } V \text{ e } A \subseteq S\}$$

a família de todos os conjuntos ortonormais em V que incluem A.

Como O_A é uma sub-família não vazia (pois $A \in O_A$) do conjunto das partes $\mathscr{P}(V)$, temos que O_A é parcialmente ordenado pela inclusão de conjuntos. Considere uma cadeia arbitrária $C = \{C_i, i \in \Delta\}$ em O_A e considere a união $\bigcup_{i\in\Delta} C_i$ de todos os conjuntos em C. Se $\boldsymbol{v}$ e $\boldsymbol{u}$ são vetores distintos em $\bigcup_{i\in\Delta} C_i$, $\boldsymbol{v} \in C_i$ e $\boldsymbol{u} \in C_j$, em que $C_i, C_j \in C \subseteq O_A$. Como C é uma cadeia, segue-se que $C_i \subseteq C_j$ ou $C_j \subseteq C_i$. Suponha (sem perda de generalidade) que $C_i \subseteq C_j$. (portanto $\boldsymbol{v}, \boldsymbol{u} \in C_j$, e então $\bigcup_{i\in\Delta} C_i$ é um conjunto ortonormal (pois $C_j \in O_A$) . Além disso, $A \subseteq \bigcup_{i\in\Delta} C_i$.

Logo $\bigcup_{i\in\Delta} C_i \in O_A$. Como $\bigcup_{i\in\Delta} C_i$ é um limite superior para C, podemos concluir que toda cadeia em O_A tem um limite superior em O_A. Portanto, O_A possui um elemento maximal pelo Lema de Zorn. Portanto, seja B ser um elemento maximal de O_A, que claramente é um conjunto ortonormal em V que inclui A. Se houver um vetor unitário $\boldsymbol{v}$ em V de modo que $B \bigcup \{\boldsymbol{v}\}$ seja um conjunto ortonormal, então $B \bigcup \{\boldsymbol{v}\}$ estará em O_A e contém propriamente B, o que contradiz o fato de que B é um elemento maximal de O_A. Portanto, não há vetor unitário $\boldsymbol{v}$ em V, para o qual $B \bigcup \{\boldsymbol{v}\}$ é um conjunto ortonormal e, portanto, B é um conjunto ortonormal maximal em V

■

O próximo resultado nos fornece um algoritmo para a construção de sequências de vetores linearmente independentes.

> **6.9 Teorema** Processo de Ortogonalização de Gram-Schmidt
>
> Seja $\underline{\mathrm{B}} = (\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ \ldots)$ uma sequência de vetores linearmente independentes em um espaço com produto interno V, então a sequência $O = (\boldsymbol{u}_1,\ \boldsymbol{u}_2,\ \ldots)$ definida por
>
> $$\boldsymbol{u}_k = \boldsymbol{v}_k - \sum_{i=1}^{k-1} \frac{\langle \boldsymbol{v}_k, \boldsymbol{u}_i \rangle}{\langle \boldsymbol{u}_i, \boldsymbol{u}_i \rangle} \boldsymbol{u}_i$$
>
> é uma sequência ortogonal em V com a propriedade que
>
> $$\langle \boldsymbol{u}_1,\ \ldots,\ \boldsymbol{u}_k \rangle = \langle \boldsymbol{v}_{,1}\ ,\ \ldots,\ \boldsymbol{v}_k \rangle$$
>
> para todos os $k > 0$.

> Obviamente, da partir da sequência ortogonal $(\boldsymbol{u}_i)$, obtemos a sequência ortonornal $(\boldsymbol{w}_i)$, onde $\boldsymbol{w}_i = \boldsymbol{u}_i / \|\boldsymbol{u}_i\|$.

Demonstração. A demonstração é por indução. Seja $\boldsymbol{u}_1 = \boldsymbol{v}_1$. Para o passo indutivo suponha que o conjunto ortogonal $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ de vetores diferentes de zero foi escolhido de modo que

$$\langle \boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \rangle = \langle \boldsymbol{v},_1, \ldots, \boldsymbol{v}_k \rangle$$

o próximo vetor $\boldsymbol{u}_{k+1}$ é escolhido definindo

$$\boldsymbol{u}_k = \boldsymbol{v}_k + \lambda_1 \boldsymbol{u}_1 + \cdots + \lambda_{k-1} \boldsymbol{u}_{k-1}$$

e exigindo que $\boldsymbol{u}_{k+1}$ seja ortogonal a cada vetor $\boldsymbol{u}_i$ para $i < k$, ou seja,

$$0 = \langle \boldsymbol{u}_k, \boldsymbol{u}_i \rangle = \langle \boldsymbol{v},_k + \lambda_1 \boldsymbol{u}_1 + \cdots + \lambda_{k-1} \boldsymbol{u}_{k-1}, \boldsymbol{u}_i \rangle = \langle \boldsymbol{v},_k, \boldsymbol{u}_i \rangle + \lambda_i \langle \boldsymbol{u}_i, \boldsymbol{u}_i \rangle$$

ou, finalmente,

$$\lambda_i = -\frac{\langle \boldsymbol{v}_k, \boldsymbol{u}_i \rangle}{\langle \boldsymbol{u}_i, \boldsymbol{u}_i \rangle}$$

para todos os $i = 1, \ldots, k$.

■

Exemplo 10. *Considere o espaço vetorial $P_3(R)$ munido do produto interno*

$$\langle p, q \rangle = \int_{-1}^{1} p(x) q(x) dx.$$

Podemos a partir da base $\underline{B} = \{1, x, x^2, x^3\}$ obter uma base ortogonal utilizando o processo de ortogonalização de Gram-Schmidt:

$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = x^2 - \frac{1}{3}$ *e* $P_3(x) = x^3 - \frac{3}{5} x.$

Os polinômios P_0, P_1, P_2, P_3 são denominados polinômios ortogonais de Legendre.

◁

Exemplo 11. *Considere o espaço vetorial $P_3(\mathbb{R})$ munido do produto interno*

$$\langle p, q \rangle = \int_{0}^{\infty} \exp(-x) p(x) q(x) dx.$$

Utilizando o processo de ortogonalização de Gram-Schmidt na base $\underline{B} = \{1, x, x^2, x^3\}$, obtemos os polinômios

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= x - 1 \\ L_2(x) &= x^2 - 4x + 2 \\ L_3(x) &= x^3 - 9x^2 + 18x - 6 \end{aligned}$$

que são denominados polinômios ortogonais de Laguerre.

Na construção dos polinômios de Laguerre utilizamos que $\int_0^\infty \exp(-x) x^n dx = n!$ para $n \in \mathbb{N}$

◁

6.12 Teorema

Seja V um espaço com produto interno.

a) Se $\dim V = n$ for finita, então V terá uma base Hilbert de tamanho n e todas as bases Hilbert para V terão tamanho n.

b) Se V possuir uma base de Hilbert finita com tamanho n, $\dim V = n$.

Demonstração. Para a parte a), a aplicação do processo de ortogonalização de Gram-Schmidt a uma base de Hamel fornece uma base de Hilbert de mesmo tamanho, n. Além disso, se V possuir uma base Hilbert de tamanho maior que n, também deverá ter uma base Hamel de tamanho maior que n, o que não é possível. Finalmente, se V possuir uma base de Hilbert, B de tamanho menor que n, então B poderá ser estendido para um conjunto próprio C que também é linearmente independente. O processo Gram-Schmidt aplicado a C fornece um superconjunto próprio de B ortonormal, o que não é possível. Portanto, todas as bases de Hilbert têm tamanho n.

Para a parte b), suponha que $\dim V > n$. Como uma base de Hilbert $\mathcal{H}$ de tamanho n é linearmente independente, podemos adicionar um novo vetor a $\mathcal{H}$ para obter um conjunto linearmente independente do tamanho $n + 1$. A aplicação do processo Gram-Schmidt a esse conjunto fornece um conjunto ortonormal que contém propriamente $\mathcal{H}$, o que não é possível.

■

6.4 Melhor Aproximação

Bases ortonormais têm uma grande vantagem sobre bases arbitrárias. Do ponto de vista computacional, se $B = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ é uma base para V, então cada $\boldsymbol{v} \in V$ pode ser escrito na forma

$$\boldsymbol{v} = \lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n$$

Em geral, no entanto, determinar as coordenadas λ_i requer a resolução de um sistema de equações lineares de tamanho $n \times n$.

Por outro lado, se $O = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ é uma base ortonormal para V e

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \cdots + \lambda_n \boldsymbol{u}_n$$

então os coeficientes são facilmente calculados:

$$\langle \boldsymbol{v}, \boldsymbol{u}_i \rangle = \langle \lambda_1 \boldsymbol{u}_1 + \cdots + \lambda_n \boldsymbol{u}_n, \boldsymbol{u}_i \rangle = \lambda_i \langle \boldsymbol{u}_i, \boldsymbol{u}_i \rangle = \lambda_i$$

Mesmo quando $O = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ não é uma base (mas apenas um conjunto ortonormal), ainda podemos considerar a expansão

$$\hat{\boldsymbol{v}} = \langle \boldsymbol{v}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \cdots + \langle \boldsymbol{v}, \boldsymbol{u}_n \rangle \boldsymbol{u}_n$$

6.1 Definição Coeficiente de Fourier

Seja V um espaço com produto interno. Seja $O = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \ldots\}$ uma sequência ortonormal de vetores de V. Para qualquer $\boldsymbol{v} \in V$, o k **coeficiente de Fourier** v em relação a O é definido como

$$f_k \triangleq \langle \boldsymbol{v}, \boldsymbol{u}_k \rangle$$

Exemplo 2. *Seja $C[-1, 1]$ o espaço vetorial de todas as funções contínuas reais no intervalo fechado $[-1, 1]$ munido do produto interno*

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx$$

Já vimos que conjunto $\underline{B} = \{\operatorname{sen}(n\pi x)\}_{n=1}^{\infty}$ *é ortogonal.*

Os coeficientes de Fourier da função $f(x) = x$, para $x \in [-\pi, \pi]$, com relação ao conjunto ortogonal B, são dados por:

$$f_k = \begin{cases} \frac{2}{k\pi} & \text{se } k \text{ ímpar} \\ -\frac{2}{k\pi} & \text{se } k \text{ par} \end{cases}$$

◁

6.3 Teorema Expansão de Fourier

Seja V um espaço com produto interno de dimensão finita. Seja $O = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ um conjunto ortonormal de vetores de V. Para qualquer $\boldsymbol{v} \in V$, a expansão de Fourier de v em relação a O é

$$\hat{\boldsymbol{v}} \triangleq \langle \boldsymbol{v}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \cdots + \langle \boldsymbol{v}, \boldsymbol{u}_k \rangle \boldsymbol{u}_k$$

Nesse caso, a desigualdade de Bessel vale para todos os $\boldsymbol{v} \in V$, isto é,

$$\|\hat{\boldsymbol{v}}\| \leq \|\boldsymbol{v}\|$$

Além disso, são equivalentes:

a) O conjunto O é uma base ortonormal para V.

b) Todo vetor é igual à sua expansão de Fourier, ou seja, para todos os $\boldsymbol{v} \in V$

$$\hat{\boldsymbol{v}} = \boldsymbol{v}$$

c) A identidade de Bessel vale para todos os $\boldsymbol{v} \in V$, isto é,

$$\|\hat{\boldsymbol{v}}\| = \|\boldsymbol{v}\|$$

d) A identidade de Parseval vale para todos os $\boldsymbol{v}, \boldsymbol{w} \in V$, ou seja

$$\langle \boldsymbol{v},\ \boldsymbol{w}\rangle = \langle \boldsymbol{v},\ \boldsymbol{u}_1\rangle\overline{\langle \boldsymbol{w}, \boldsymbol{u}_1\rangle} + \cdots + \langle \boldsymbol{v},\ \boldsymbol{u}_k\rangle\overline{\langle \boldsymbol{w}, \boldsymbol{u}_k\rangle}$$

Vimos que, se S é um subespaço de um espaço com produto interno V, então $S \cap S^{\perp} = \{\mathbf{0}\}$. Isso levanta a questão de saber se o complemento ortogonal $S^{\perp}$ é um complemento de espaço vetorial de S, ou seja, se $V = S \oplus S^{\perp}$.

Se S é um subespaço de dimensão finita de V, a resposta é sim, mas para subespaços de dimensão infinita em geral, $V \neq S \oplus S^{\perp}$.

Exemplo 4. *Seja $V = \ell^2$ e seja S o subespaço de todas as sequências de suporte finito, ou seja, S seja o subespaço gerado pelos vetores*

$$\boldsymbol{e}_i = (0,\ \ldots,\ 0, 1, 0,\ \ldots)$$

onde $\boldsymbol{e}_i$ tem um 1 na i-ésima coordenada e 0s em outros lugares. Se $x = (x_n) \in S^{\perp}$, então

$x_i = \langle x, \boldsymbol{e}_i\rangle = 0$ para todos os i e, portanto, $x = 0$. Portanto, $S^{\perp} = \{\mathbf{0}\}$. No entanto,

$$S \oplus S^{\perp} = S \neq \ell^2$$

◁

6.5 Projeção

Como mostra o próximo teorema, no caso de dimensão finita, os complementos ortogonais também são complementos de espaço vetorial.

6.1 Teorema Projeção

Se S for um subespaço finito-dimensional de um espaço com produto interno V (que não precisa ser finito-dimensional), então

$$V = S \oplus S^{\perp}$$

Demonstração. Seja $O = \{\boldsymbol{u}_1,\ \ldots,\ \boldsymbol{u}_k\}$ uma base ortonormal por S. Para cada $\boldsymbol{v} \in V$, considere a expansão de Fourier

$$\hat{\boldsymbol{v}} = \langle \boldsymbol{v},\ \boldsymbol{u}_1\rangle\boldsymbol{u}_1 + \cdots + \langle \boldsymbol{v},\ \boldsymbol{u}_k\rangle\boldsymbol{u}_k$$

em relação a O. Podemos escrever

$$\boldsymbol{v} = \hat{\boldsymbol{v}} + (\boldsymbol{v} - \hat{\boldsymbol{v}})$$

onde $\hat{\boldsymbol{v}} \in S$. Além disso, $\boldsymbol{v} - \hat{\boldsymbol{v}} \in S^{\perp}$, já que

$$\langle \boldsymbol{v} - \hat{\boldsymbol{v}},\ \boldsymbol{u}_i \rangle = \langle \boldsymbol{v},\ \boldsymbol{u}_i \rangle - \langle \hat{\boldsymbol{v}},\ \boldsymbol{u}_i \rangle = \mathbf{0}$$

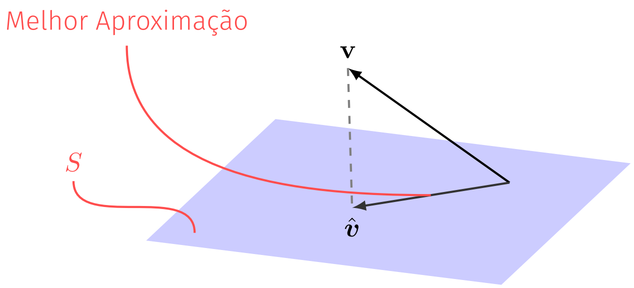
Portanto, $V = S + S^{\perp}$. Já observamos que $S \cap S^{\perp} = \{\mathbf{0}\}$ e, portanto, $V = S \oplus S^{\perp}$. ■

De acordo com a prova do teorema da projeção, o componente de v que está em S é apenas a expansão de Fourier de v em relação a qualquer base ortonormal O por S.

O teorema da projeção implica que, se $\boldsymbol{v} = \hat{\boldsymbol{v}} + \boldsymbol{s}^{\perp}$ em que $\hat{\boldsymbol{v}} \in S$ e $\boldsymbol{s}^{\perp} \in S^{\perp}$ então $\hat{\boldsymbol{v}}$ é o elemento de S mais próximo de v, ou seja, $\hat{\boldsymbol{v}}$ é a melhor aproximação para v de S. Pois se $\boldsymbol{t} \in S$, como $\boldsymbol{v} - \hat{\boldsymbol{v}} \in S^{\perp}$, temos $(\boldsymbol{v} - \hat{\boldsymbol{v}}) \perp (\hat{\boldsymbol{v}} - \boldsymbol{t})$ e assim

$$\|\boldsymbol{v} - \boldsymbol{t}\|^2 = \|\boldsymbol{v} - \hat{\boldsymbol{v}} + \hat{\boldsymbol{v}} - \boldsymbol{t}\|^2 = \|v - \hat{\boldsymbol{v}}\|^2 + \|\hat{\boldsymbol{v}} - \boldsymbol{t}\|^2$$

Desse fato resulta que $\|\boldsymbol{v} - \boldsymbol{t}\|$ é menor quando $\boldsymbol{t} = \hat{\boldsymbol{v}}$. Além disso, observamos que $\hat{\boldsymbol{v}}$ é o único vetor em S para o qual $\boldsymbol{v} - \hat{\boldsymbol{v}} \perp S$. Assim, podemos dizer que a melhor aproximação de $\boldsymbol{v}$ a partir de S é o único vetor $\boldsymbol{s} \in S$ para o qual $(\boldsymbol{v} - \boldsymbol{s}) \perp S$ e que esse vetor é a expansão de Fourier $\hat{\boldsymbol{v}}$ de $\boldsymbol{v}$.



6.2 Definição Soma Direta Ortogonal

Seja V um espaço com produto interno e sejam $S_1, \ldots, S_n$ subespaços de V. Então V é dito **soma direta ortogonal** de $S_1, \ldots, S_n$, escrita

$$S = S_1 \odot S_n \qquad \text{se}$$

se

- a) $V = S_1 \oplus \cdots \oplus S_n$
- b) $S_i \perp S_j$ por $i \neq j$

O teorema 6.5 afirma que $V = S \odot S^{\perp}$, para qualquer subespaço de dimensão finita S de um espaço vetorial V. O seguinte resultado simples é muito útil.

6.3 Teorema

Seja V um espaço com produto interno. Então são equivalentes

a $V = S \odot T$

b $V = S \oplus T$ e $T = S^{\perp}$

c $V = S \oplus T$ e $T \subseteq S^{\perp}$

Demonstração. Suponha que a seja válido. Então $V = S \oplus T$ e $S \perp T$, o que implica que $T \subseteq S^{\perp}$. Mas se $w \in S^{\perp}$ então $w = s + t$ para $s \in S, t \in T$ e assim

$$0 = \langle s,\ w\rangle = \langle s,\ s\rangle + \langle s,\ t\rangle = \langle s,\ s\rangle$$

Portanto, $s = 0$ e $w \in T$, o que implica que $S^{\perp} \subseteq T$. Portanto, $S^{\perp} = T$, o que fornece b. Obviamente, b implica c. Finalmente, se c contém $T \subseteq S^{\perp}$, o que implica que $S \perp T$ e, portanto, a são válidos. ■

6.4 Teorema

Seja V um espaço com produto interno.

a Se $\dim V < \infty$ e S for um subespaço de V,

$$\dim(S^{\perp}) = \dim V - \dim(S)$$

b Se S é um subespaço de dimensão finita de V, então

$$S^{\perp\perp} = S$$

c Se X for um subconjunto de V e $\dim(\langle X\rangle) < \infty$, então

$$X^{\perp\perp} = \langle X\rangle$$

Demonstração. Como $V = S \oplus S^{\perp}$, temos $\dim V = \dim(S) + \dim(S^{\perp})$, o que prova a parte a. Quanto à parte b, está claro que $S \subseteq S^{\perp\perp}$. Por outro lado, se $\boldsymbol{v} \in S^{\perp\perp}$ então pelo teorema da projeção

$$\boldsymbol{v} = \boldsymbol{s} + \boldsymbol{s}'$$

onde $\boldsymbol{s} \in S$ e $\boldsymbol{s}' \in S^{\perp}$. Mas $\boldsymbol{v} \in S^{\perp\perp}$ implica que $0 = \langle \boldsymbol{v}, \boldsymbol{s}'\rangle = \langle \boldsymbol{s}', \boldsymbol{s}'\rangle$ e assim $\boldsymbol{s}' = 0$, mostrando que $\boldsymbol{v} \in S$. Portanto, $S^{\perp\perp} \subseteq S$ e $S^{\perp\perp} = S$. Deixamos a prova da parte c como um exercício. ■

6.6 Adjuntas

Em nosso primeiro resultado, mostramos que em um espaço com produto interno existe um isomorfismo canônico entre os vetores no espaço dual V^* e os vetores em V.

6.1 Teorema de Representação de Riesz

Todo funcional linear $f : V \to \mathbb{K}$ num espaços com produto interno V pode ser escrita como um produto interno. Mais precisamente, existe um único $\boldsymbol{y} \in V$ tal que

$$f(\boldsymbol{x}) = \langle \boldsymbol{x},\ \boldsymbol{y} \rangle \qquad \text{para todo } x \in V.$$

Demonstração. Considere uma base ortonormal $\underline{\mathrm{x}} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ de V. Se $\boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{x}_1 \rangle \boldsymbol{x}_1 + \ldots + \langle \boldsymbol{x}, \boldsymbol{x}_n \rangle \boldsymbol{x}$, então

$$\begin{aligned} f(\boldsymbol{x}) &= \langle \boldsymbol{x},\ \boldsymbol{x}_1 \rangle f(\boldsymbol{x}_1) + \ldots + \langle \boldsymbol{x},\ \boldsymbol{x}_n \rangle f(\boldsymbol{x}_n) \\ &= \langle \boldsymbol{x},\ \overline{f(\boldsymbol{x}_1)} \boldsymbol{x}_1 \rangle + \ldots + \langle \boldsymbol{x}, \overline{f(\boldsymbol{x}_n)} \boldsymbol{x}_n \rangle \\ &= \langle \boldsymbol{x},\ \overline{f(\boldsymbol{x}_1)} \boldsymbol{x}_1 + \ldots + \overline{f(\boldsymbol{x}_n)} \boldsymbol{x}_n \rangle. \end{aligned}$$

Defina $\boldsymbol{y} = \overline{f(\boldsymbol{x}_1)} \boldsymbol{x}_1 + \ldots + \overline{f(\boldsymbol{x}_n)} \boldsymbol{x}_n$. A unicidade do vetor $\boldsymbol{y}$ é consequência de $\underline{\mathrm{x}}$ ser uma base de V.

■

Para $f \in V^*$, seja f^* o vetor $\boldsymbol{y}$ em V de modo que $f(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$. A função $f \to f^*$ de V^* a V é aditiva. Se o corpo for os reais, o mapa $f \to f^*$ é linear. No caso, em que o corpo é os complexos o mapa não será linear, mas linear conjugada pois $(\lambda f)^* = \overline{\lambda} f^*$.

6.2 Corolário

Se V é um espaços com produto interno real, a aplicação $f \mapsto f^*$ é um isomorfismo entre $V*$ e V.

6.3 Teorema

Sejam V, W espaços com produto interno. Dada uma aplicação linear $T : V \to W$, existe uma única aplicação linear $T^* : W \to V$, denominada adjunta de T, satisfazendo

$$\langle T\boldsymbol{v},\ \boldsymbol{w}\rangle = \langle \boldsymbol{v},\ T^*\boldsymbol{w}\rangle \ \text{ para todo } \boldsymbol{v} \in V,\ \boldsymbol{w} \in W.$$

Demonstração. Para $\boldsymbol{w} \in W$ fixo, a aplicação $\boldsymbol{v} \mapsto \langle T\boldsymbol{v}, \boldsymbol{w}\rangle$ pertence ao dual de V. O Teorema de Representação de Riesz garante então que existe um único $\boldsymbol{y} \in W$ tal que

$$\langle T\boldsymbol{v},\ \boldsymbol{w}\rangle = \langle \boldsymbol{v},\ \boldsymbol{y}\rangle.$$

Definimos $T^*\boldsymbol{w} = \boldsymbol{y}$. Temos assim uma aplicação $T^* : W \to V$.

Para demonstrar a linearidade de T^* sejam $\boldsymbol{u}, \boldsymbol{w} \in W$ e $\lambda \in \mathbb{K}$. Então

$$\langle \boldsymbol{v},\ T^*(\boldsymbol{u} + \lambda \boldsymbol{v})\rangle = \langle T\boldsymbol{v},\ \boldsymbol{u} + \lambda\boldsymbol{w}\rangle = \langle T\boldsymbol{v},\ \boldsymbol{u}\rangle + \overline{\lambda}\langle T\boldsymbol{v},\ \boldsymbol{w}\rangle = \langle \boldsymbol{v},\ T^*\boldsymbol{u}\rangle + \langle \boldsymbol{v},\ \lambda T^*\boldsymbol{w}\rangle.$$

Se $S^* : W \to V$ fosse outra aplicação linear tal que $\langle T\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{v}, S^*\boldsymbol{w}\rangle$, então $\langle \boldsymbol{v}, T^*\boldsymbol{w} - S^*\boldsymbol{w}\rangle = 0$ para todo $\boldsymbol{v} \in V$. Escolhendo $\boldsymbol{v} = T^*\boldsymbol{w} - S^*\boldsymbol{w}$, concluímos que $T^* = S^*$ ∎

6.4 Definição Adjunta

Sejam V, W espaços com produto interno. A aplicação linear $T^* : W \to V$ que satisfaz

$$\langle T\boldsymbol{v},\ \boldsymbol{w}\rangle = \langle \boldsymbol{v},\ T^*\boldsymbol{w}\rangle \text{ para todos } \boldsymbol{v} \in V,\ \boldsymbol{w} \in W$$

é denominada **adjunta** de T.

O resultado a seguir enumera algumas propriedades do mapa $T \to T^*$ de $\mathrm{Hom}(V,\ W)$ a $\mathrm{Hom}(W^*,\ V^*)$.

6.5 Teorema

Sejam V, W e Z espaços de produto interno de dimensão finita. Então:

1. Se $S, T \in \mathrm{Hom}(V,\ W)$ então $(S + T)^* = S^* + T^*$;
2. Se $T \in \mathrm{Hom}(V,\ W)$ e $\lambda \in \mathbb{K}$, então $(\lambda T)^* = \overline{\lambda} T^*$.;
3. Se $S \in \mathrm{Hom}(V,\ W)$ e $T \in \mathrm{Hom}(W,\ Z)$, então $(TS)^* = S^*T^*$;
4. Se $T \in \mathrm{Hom}(V,\ W)$ então $(T^*)^* = T$; e
5. $I_V^* = I_{V^*}$.

Demonstração. Sejam $(V,\ \langle,\ \rangle_V), (W,\ \langle,\ \rangle_W), (Z,\ \langle,\ \rangle_Z)$

1 Seja $\boldsymbol{v} \in V, \boldsymbol{w} \in W$. Então

$$\begin{aligned}
\langle \boldsymbol{v},\ (S+T)^*(\boldsymbol{w})\rangle_V &= \langle (S+T)(\boldsymbol{v}), \boldsymbol{w}\rangle_W \\
&= \langle S(\boldsymbol{v}) + T(\boldsymbol{v}), \boldsymbol{w}\rangle_W \\
&= \langle S(\boldsymbol{v}), \boldsymbol{w}\rangle_W + \langle T(\boldsymbol{v}), \boldsymbol{w}\rangle_W \\
&= \langle \boldsymbol{v}, S^*(\boldsymbol{w})\rangle_V + \langle \boldsymbol{v}, T^*(\boldsymbol{w})\rangle_V \\
&= \langle \boldsymbol{v}, S^*(\boldsymbol{w}) + T^*(\boldsymbol{w})\rangle_V \\
&= \langle \boldsymbol{v},\ (S^* + T^*)(\boldsymbol{w})\rangle_V .
\end{aligned}$$

Consequentemente, $(S+T)^*(\boldsymbol{w}) = S^*(\boldsymbol{w}) + T^*(\boldsymbol{w})$ para todos os $\boldsymbol{w} \in W$ e, portanto, $(S+T)^* = S^* + T^*$.

2 Sejam $\boldsymbol{v} \in V, \boldsymbol{w} \in W$ e λ um escalar. Então

$$\begin{aligned}
\langle \boldsymbol{v},\ (\lambda T)^*(\boldsymbol{w})\rangle_V &= \langle (\lambda T)(\boldsymbol{v}), \boldsymbol{w}\rangle_W \\
&= \langle \lambda T(\boldsymbol{v}), \boldsymbol{w}\rangle_W \\
&= \lambda\langle (T(\boldsymbol{v}), \boldsymbol{w}\rangle_W \\
&= \lambda\langle \boldsymbol{v}, T^*(\boldsymbol{w})\rangle_V \\
&= \langle \boldsymbol{v}, \overline{\lambda} T^*(\boldsymbol{w})\rangle_V .
\end{aligned}$$

Podemos então concluir que $(\lambda T)^* = \overline{\lambda} T^*$.

3 Seja $\boldsymbol{v} \in V, \boldsymbol{z} \in Z$. Então $ST(\boldsymbol{v}) \in Z$ e pela equação fundamental que define $(ST)^*$ temos

$$\begin{aligned}
\langle \boldsymbol{v},\ (ST)^*(\boldsymbol{z})\rangle_V &= \langle (ST)(\boldsymbol{v}),\ \boldsymbol{z}\rangle_Z \\
&= \langle S(T(\boldsymbol{v})),\ \boldsymbol{z}\rangle_Z .
\end{aligned}$$

Como $T(v) \in W$, pela equação fundamental que define S^*, temos

$$\langle S(T(\boldsymbol{v})),\ \boldsymbol{z}\rangle_Z = \langle T(\boldsymbol{v}),\ S^*(\boldsymbol{z})\rangle_W .$$

Por sua vez, como $\boldsymbol{v} \in V$ e $S^*(\boldsymbol{x}) \in W$, temos pela equação fundamental aplicada a T

$$\begin{aligned}
\langle T(\boldsymbol{v}), S^*(\boldsymbol{z})\rangle_W &= \langle \boldsymbol{v}, T^*(S^*(\boldsymbol{z}))\rangle_V \\
&= \langle \boldsymbol{v}, (T^*S^*)(\boldsymbol{z})\rangle_V
\end{aligned}$$

Então para todos os vetores $\boldsymbol{z} \in Z$ que $(ST)^*(\boldsymbol{z}) = T^*S^*(\boldsymbol{z})$.

As duas últimas partes são diretas e as deixamos como exercícios. ■

Em seguida, apresentaremos algumas relações entre a imagem e o núcleo de $T \in \operatorname{Hom}(V,\ W)$ e de $T^* \in \operatorname{Hom}(W,\ V)$.

6.6 Teorema

Sejam V, W espaço com produto interno com dimensões finitas e $T \in \operatorname{Hom}(V,\ W)$. Então

1 $\ker(T^*) = \operatorname{im}(T)^{\perp}$;

2 $\operatorname{im}(T^*) = \ker(T)^{\perp}$;

3 $\ker(T) = \operatorname{im}(T^*)^{\perp}$; e

4 $\operatorname{im}(T) = \ker(T^*)^{\perp}$.

Demonstração. Sejam $(V,\ \langle\cdot,\cdot\rangle_V)$, $(W,\ \langle\cdot,\cdot\rangle_W)$.

1 Suponha $\boldsymbol{w} \in \ker(T^*)$. Então $\langle \boldsymbol{v}, T^*(\boldsymbol{w})\rangle_V = \langle \boldsymbol{v}, \boldsymbol{0}\rangle_V = 0$ para todos os $\boldsymbol{v} \in V$. Pela definição de T^*, $\langle \boldsymbol{v}, T^*(\boldsymbol{w})\rangle_V = \langle T(\boldsymbol{v}), \boldsymbol{w}\rangle_W$. Isso implica que $\boldsymbol{w} \perp T(\boldsymbol{v})$ para todos os $\boldsymbol{v} \in V$ e, portanto, $\boldsymbol{w} \in \operatorname{im}(T)^{\perp}$. Portanto, $\ker(T^*) \subset \operatorname{im}(\perp)^{\perp}$.

Seja $\boldsymbol{w} \in \operatorname{im}(T)^{\perp}$. Então, para todos os $\boldsymbol{v} \in V$, $\langle T(\boldsymbol{v}), \boldsymbol{w}\rangle_W = 0$. Mas então, pela definição de T^*, $\langle \boldsymbol{v}, T^*(\boldsymbol{w})\rangle_V = 0$. Em particular, $\langle T^*(\boldsymbol{w}), T^*(\boldsymbol{w})\rangle_V = 0$, e logo , $T^*(\boldsymbol{w}) = \boldsymbol{0}$ e $\boldsymbol{w} \in \ker(T^*)$.

Como $(T^*)^* = T$, temos que 3 é consequência de 1.

De 1 também deduzimos que $\ker(T^*)^{\perp} = [\operatorname{im}(T)^{\perp}]^{\perp} = \operatorname{im}(T)$ e, consequentemente, 4 também é verdadeiro.

Finalmente, como $\ker(T) = \operatorname{im}(T^*)^{\perp}$, temos que $\ker(T)^{\perp} = [\operatorname{im}(T^*)^{\perp}]^{\perp} = \operatorname{im}(T^*)$, e assim 2 é válido. ∎

Chegamos ao nosso teorema final, que relaciona a matriz de T e $T*$ quando eles são computados em relação às bases ortonormais de V e W. Esse teorema justifica chamarmos o operador $T*$ de transposta de T

6.7 Teorema

Sejam V e W espaço com produto interno com bases ortonormais $\underline{\mathrm{v}} = (\boldsymbol{v}_1,\ \ldots,\ \boldsymbol{v}_n)$ e $\underline{\mathrm{w}} = (\boldsymbol{w}_1,\ \ldots,\ \boldsymbol{w}_m)$ por V e W, respectivamente. Seja $A = [T]_{\underline{\mathrm{v}},\underline{\mathrm{w}}}$ e $B = [T^*]_{\underline{\mathrm{w}},\underline{\mathrm{v}}}$. Então $B = \overline{A}^{\mathrm{t}}$

Demonstração. Sejam $(V,\ \langle,\ \rangle_V)$ e $(W,\ \langle,\ \rangle_W)$. Então

$$T(\boldsymbol{v}_j) = a_{1j}\boldsymbol{w}_1 + + \cdots + a_{mj}\boldsymbol{w}_m. \tag{6.9}$$

$$T^*(\boldsymbol{w}_i) = b_{1i}\boldsymbol{v}_1 + + \cdots + b_{ni}\boldsymbol{v}_n. \tag{6.10}$$

Precisamos provar que $b_{ji} = \overline{a_{ij}}$ ou equivalente, que $a_{ij} = \overline{b_{ji}}$. Fazemos isso calculando $\langle T(\boldsymbol{v}_j), \boldsymbol{w}_i\rangle_W = \langle \boldsymbol{v}_j, T^*(\boldsymbol{w}_i)\rangle_V$ usando as Equações 6.9 e 6.10.

Por um lado

$$\langle T(\boldsymbol{v}_j),\ \boldsymbol{w}_i\rangle_W = \sum_{k=1}^{m} a_{kj}\langle \boldsymbol{w}_k,\ \boldsymbol{w}_i\rangle_W = a_{ij},$$

na última igualdade, utilizamos que $\underline{w}$ é uma base ortonormal de W. Por outro lado,

$$\langle \boldsymbol{v}_j,\ T^*(\boldsymbol{w}_i)\rangle_V = \sum_{l=1}^{m} \overline{b_{li}} \langle \boldsymbol{v}_j,\ \boldsymbol{v}_l\rangle_V = \overline{b_{ji}}.$$

Assim, $a_{ij} = \overline{b_{ji}}$. ∎

6.7 ⚡Espaços de Hilbert

Seja V um espaço de produto interno. Dizemos que uma sequência $(\boldsymbol{v}^n)$ de vetores em V **converge** para $\boldsymbol{v} \in V$ se

$$\lim_{n\to\infty} \|\boldsymbol{v}^n - \boldsymbol{v}\| = 0$$

6.1 Definição Sequência de Cauchy

Seja (M, d) um espaço métrico. Uma sequência (x^n) em M é uma **sequência de Cauchy** se, para todo $\epsilon > 0$, existir um número inteiro n_0 tal que $n, m \geq n_0$ implica que $d(x^n, x^m) < \epsilon$.

6.2 Definição Espaço Completo

Dizemos que o espaço métrico (M, d) é **completo** o se todas as sequências de Cauchy em M convergirem para um limite em M.

Exemplo 3. *O corpo dos reais $\mathbb{R}$ é um espaço métrico completo em relação a sua métrica natural. Também é $\mathbb{C}$, pois se (z_n) é uma sequência Cauchy em $\mathbb{C}$, então $(\operatorname{Re} z_n)$ e $(\operatorname{Im} z_n)$ são sequências de Cauchy em $\mathbb{R}$. Essas sequências possuem limites $x, y \in \mathbb{R}$, e assim $z_k \to x + \mathrm{i}y \in \mathbb{C}$*

◁

6.4 Definição Espaço de Hilbert

Um **espaço de Hilbert** é um espaço com produto interno que é um espaço métrico completo em relação à métrica induzida por seu produto interno.

No caso de dimensão finita, a situação é muito simples: todos os espaço com produto interno são completos, todos os subespaços são fechados e todos os operadores e funcionais lineares são contínuos. No entanto, no caso de dimensão infinita, as coisas não são tão simples.

6.5 Teorema

$\mathbb{R}^n, \mathbb{C}^n$ e ℓ^2 são espaços de Hilbert.

Demonstração. Os espaços $\mathbb{R}^n, \mathbb{C}^n$ e ℓ^2 são espaços com produto interno. Assim temos que provar somente a completeza. Provamos a completeza de ℓ^2 já que as outras completezas são similares e mais simples.

Seja $(\boldsymbol{x}^k)$ uma sequência Cauchy em ℓ^2. Cada elemento dessa sequência $(\boldsymbol{x}^k)$ é em si uma sequência real:

$$\boldsymbol{x}^k \triangleq (x_n^k)_{n=1}^{\infty}.$$

Precisamos mostrar que $(\boldsymbol{x}^k)$ converge em ℓ^2. As três etapas a seguir são padrão nas provas de completeza:

1. Encontrar um possível limite $\boldsymbol{a}$;
2. Mostrar que $\boldsymbol{a}$ pertence ao espaço em questão;
3. Mostrar que $\boldsymbol{x}^k \to \boldsymbol{a}$.

Etapa 1. Encontrar um possível limite $\boldsymbol{a}$. Estamos procurando por um vetor $\boldsymbol{a} = (a_n)$ tal que

$$\begin{aligned} \boldsymbol{x}^1 &= (x_1^1, x_2^1,\ x_3^1,\ \dots,\ x_n^1,\ \dots) \\ \boldsymbol{x}^2 &= (x_1^2, x_2^2,\ x_3^2,\ \dots, x_n^2,\ \dots) \\ \vdots\ \boldsymbol{x}^k &= (x_1^k,\ x_2^k, x_3^k,\ \dots,\ x_n^k,\ \dots) \end{aligned}$$

converge para

$$\boldsymbol{a} = (a_1, a_2, a_3,\ \dots,\ a_n,\ \dots).$$

Considere uma coluna fixa: ou seja, fixe n. A sequência $(x_n^k)_{k=1}^{\infty}$ é uma sequência de Cauchy em $\mathbb{C}$. Como $\mathbb{C}$ é um espaço métrico completo, existe $a_n \in \mathbb{C}$ de modo que $x_n^k \to a_n$ quando $k \to \infty$.

Seja $\boldsymbol{a} = (a_1, a_2, \dots) \in \mathbb{C}^{\mathbb{N}}$.

Etapa 2. Vamos provar que possível limite o pertence a ℓ^2. Mostraremos que $\boldsymbol{x}^k - \boldsymbol{a} \in \ell^2$ para algum k. Como ℓ^2 é fechado em relação subtração, seguirá que

$$\boldsymbol{a} = \boldsymbol{x}^k - (\boldsymbol{x}^k - \boldsymbol{a}) \in \ell^2.$$

Seja $\epsilon > 0$. Por hipótese, existe $k_0 \in \mathbb{N}$ de tal forma que $k, l \geq k_0$ implica $\|\boldsymbol{x}^k - \boldsymbol{x}^l\| < \epsilon$. Escolha $N \in \mathbb{N}$: temos, por $k, l \geq k_0$,

$$\sum_{n=1}^{N} |x_n^k - x_n^l|^2 \leq \sum_{n=1}^{\infty} |x_n^k - x_n^l|^2 < \epsilon^2.$$

Fazendo $l \to \infty$ na soma finita do lado esquerdo: obtemos

$$\sum_{n=1}^{N} |x_n^k - a_n|^2 \leq \epsilon^2.$$

Como isso vale para todos os $N \in \mathbb{N}$ que temos, fazendo $N \to \infty$,

$$\sum_{n=1}^{\infty} |x_n^k - a_n|^2 \leq \epsilon^2,$$

isto é, $\boldsymbol{x}^k - \boldsymbol{a} \in \ell^2$ e

$$\|\mathrm{x}^k - \boldsymbol{a}\| \leq \epsilon.$$

Etapa 3. $\boldsymbol{x}^k \to \boldsymbol{a}$.

De fato, mostramos na Etapa 2 que, para qualquer $\epsilon > 0$, existe $k_0 \in \mathbb{N}$ de tal forma que $k \geq k_0$ implique $\|\boldsymbol{x}^k - \boldsymbol{a}\| \leq \epsilon$. Portanto, $\boldsymbol{x}^k \to \boldsymbol{a}$.

Assim, toda sequência Cauchy em ℓ^2 converge para um limite em ℓ^2 e, portanto, ℓ^2 é completo.

■

O próximo teorema mostra que os conjuntos ortonormais maximais em um espaço de Hilbert $\mathcal{H}$ são precisamente aqueles conjuntos ortonormais que geram $\mathcal{H}$.

6.6 Proposição

Seja A um conjunto ortonormal em um espaço com produto interno V.

a) Se $\langle A \rangle = V$, A é um conjunto ortonormal maximal .

b) Se V é um espaço de Hilbert e A é um conjunto ortonormal maximal , então $\langle A \rangle = V$.

Demonstração. Seja A qualquer conjunto ortonormal em V. De acordo com a Proposição 6.3, A é um conjunto ortonormal maximal se, e somente se, $A^\perp = \{\mathbf{0}\}$. Se $\langle A \rangle = V$, $(\langle A \rangle)^\perp = V^\perp = \{\mathbf{0}\}$. Mas $A^\perp = (\langle A \rangle)^\perp$ e, portanto, $A^\perp = \{\mathbf{0}\}$. Se V for um espaço de Hilbert e $A^\perp = \{\mathbf{0}\}$, $\langle A \rangle = V$. ■

Um conjunto ortonormal em um espaço com produto interno V que gera V é chamado de base ortonormal (ou uma base de Hilbert) por V. Em outras palavras, um subconjunto B de um espaço com produto interno V é uma base ortonormal se

- B é um conjunto ortonormal
- $\langle B \rangle = V$.

Cuidado extremo deve ser tomado aqui para não confundir os conceitos de uma base para um espaço vetorial e uma base de Hilbert para um espaço vetorial com produto interno. Para evitar confusão, uma base de espaço vetorial, ou seja, um conjunto de vetores linearmente independente maximal , é denominada base de Hamel. Uma base ortonormal de Hamel será chamada de base ortonormal, para distingui-la de uma base de Hilbert.

O exemplo a seguir mostra que, em geral, os dois conceitos de base não são os mesmos.

Exemplo 7. *Seja $V = \ell^2$ e seja M o conjunto de todos os vetores da forma*

$$\boldsymbol{e}_i = (0, \dots, 0, 1, 0, \dots)$$

onde $\boldsymbol{e}_i$ tem um 1 na ésima coordenada e 0 em outro lugar. Claramente, M é um conjunto ortonormal. Além disso, é máximal. Pois se $v = (x_n) \in \ell^2$ possui a propriedade que $v \perp M$ então

$$x_i = \langle \boldsymbol{v}, \boldsymbol{e}_i \rangle = 0$$

para todos os i e então $\boldsymbol{v} = 0$. Portanto, nenhum vetor diferente de zero $\boldsymbol{v} \notin M$ é ortogonal a M. Isso mostra que M é uma base de Hilbert para o espaço com produto interno ℓ^2.

Por outro lado, o espaço gerado por M é o subespaço S de todas as sequências em ℓ^2 que possuem suporte finito, ou seja, possuem apenas um número finito de termos diferentes de zero e assim $\langle M \rangle = S \neq \ell^2$, vemos que M não é uma base de Hamel para o espaço vetorial ℓ^2.

◁

6.7.1 Espaços de Hilbert Separáveis e Séries de Fourier

6.8 Definição Separável

Um espaço de Hilbert é dito **separável** se possuir uma base ortonormal de cardinalidade no máximo enumerável.

6.9 Proposição

Seja V um espaço de Hilbert separável. Se V tem dimensão infinita, então H é isometricamente isomorfo a ℓ^2..

Em espaços de dimensão infinita temos outro conceito de base que é útil as bases de Schauder. Essas bases são semelhante às bases usuais (Hamel) de um espaço vetorial; a diferença é que as bases de Hamel usam combinações lineares que são somas finitas, enquanto nas bases de Schauder consideramos somas infinitas.

6.10 Definição Base de Schauder

Seja V um espaço de produto interno. Uma base de Schauder para V é uma sequência $\boldsymbol{v}_n$ de vetores de V, de modo que para todo elemento $v \in V$ existe uma sequência única v_n de escalares em $\mathbb{K}$, de modo que

$$\boldsymbol{v} = \sum_{n=0}^{\infty} v_n \boldsymbol{v}_n$$

onde a convergência é entendida com relação à norma associada ao produto interno, ou seja,

$$\lim_{n\to\infty} \left\| \boldsymbol{v} - \sum_{k=0}^{n} v_k \boldsymbol{v}_k \right\| = 0.$$

6.11 Teorema

Seja V um espaço de Hilbert separável. As seguintes afirmações são equivalentes:

1. $O = \{e_n\}_{n\in\mathbb{N}}$ é um sistema ortonormal completo para V.
2. Se $x \perp e_n$ para todo $n \in \mathbb{N}$,então $x = 0$.
3. Expansão em série de Fourier: Se $x \in V$, então

$$x = \sum_{n\in\mathbb{N}} \langle x,\ e_n \rangle e_n.$$

   ou seja, O é base de Schauder de V.

4. Identidade de Parseval Se $x \in V$, então

$$\|x\|^2 = \sum_{n\in\mathbb{N}} |x_n|^2$$

6.12 Definição

Definimos $L^2[a,b]$ como o menor espaço de Hilbert que contém $C[a,b]$ o espaço das funções contínuas a valores reais em $[a,b]$.

$L^2[a,b]$ é um espaço de Hilbert separável.

6.13 Teorema

São sistemas ortogonais completos

1. Série de Fourier. As funções

$$\left\{\frac{1}{2\pi}, \frac{1}{\pi}\sqrt{2}\cos(2\pi nx), \ldots, \frac{1}{\pi}\sqrt{2}\,\text{sen}(2\pi nx), \ldots\right\} \text{ para } n = 1, 2, 3, \ldots$$

formam um sistema ortonormal completo com relação ao produto interno $\langle f, g\rangle = \int_{-L}^{L} f(x)g(x)dx$ no intervalo $[-\pi, \pi]$.

2. Polinômios Legendre , $P_k(x)$: definidos através da fórmula de Rodrigues:

$$P_0(x) = 1 \; P_k(x) = \frac{1}{2^k k!}\frac{d^k}{dx^k}(x^2-1)^k, \; k = 1, 2, \; \ldots$$

formam um sistema ortonormal completo com relação ao produto interno $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$ no intervalo $[-1, 1]$.

3. Função Bessel de Primeiro Tipo e de Ordem l,

$$J_l(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(m+l+1)}\left(\frac{x}{2}\right)^{2m+l}$$

com $\Gamma(l)$ a função gama, Dado um $l \geq 0$ fixo, as funções $\{\sqrt{x}J_l(\lambda_{lk}x)\}_{k=1}^{\infty}$ formam um sistema ortonormal completo com relação ao produto interno $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$ no intervalo $[0, 1]$.

4. polinômios de Hermite,

$$H_k(x) = (-1)^k \exp\left(x^2\right)\frac{d^k}{dx^k}(\exp\left(-x^2\right)), k = 0, 1, \ldots$$

formam um conjunto completo de funções com relação ao produto interno $\langle f, g\rangle = \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}x^2\right) f(x)g(x)dx$ no intervalo $[-\infty, \infty]$.

5. polinômios de Chebyshev do primeiro tipo, $T_k(x)$:

$$T_k(x) = \frac{k}{2}\sum_{r=0}^{\lfloor k/2\rfloor} \frac{(-1)^r}{kr}\begin{bmatrix} k-r \\ r \end{bmatrix}(2x)^{k-2r}, k = 0, 1, \ldots$$

formam um conjunto completo de funções com relação ao produto interno $\langle f, g\rangle = \int_{-\infty}^{\infty} \frac{1}{1-x^2} f(x)g(x)dx$ no intervalo $[-1, 1]$.

Exercícios

Ex. 6.1 --- Seja $T : W \to V$ uma transformação linear *injetiva* entre $\mathbb{K}$-espaços vetoriais. Suponha que $\langle , \rangle$, é um produto interno em V. Mostre que $\langle w_1, w_2\rangle =$

$\langle T(w_1), T(w_2)\rangle$ para todos $w_1, w_2 \in W$ define um produto interno em W. [Em particular, se V é um espaço vetorial com produto interno $\langle , \rangle$, W um subespaço de V e $T : W \to V$ é a inclusão natural, teremos como consequência que $\langle , \rangle$ restrito aos elementos de W é um produto interno em W.]

Ex. 6.2 --- Dizemos que uma matriz $A \in M_n(\mathbb{K})$ hermitiana (i.e. $A^* = A$) é *positiva definida* se para toda matriz coluna não nula $\mathrm{X} \in M_{n\times 1}(\mathbb{K})$ tem-se que $\mathrm{X}^*A\mathrm{X} > 0$.

1. Mostre que se V é um espaço vetorial sobre $\mathbb{K}$ com produto interno $\langle , \rangle$ e se $\underline{\mathrm{B}} = \{v_1, v_2, \ldots, v_n\}$ é uma base de V então $A = (a_{ij}) \in M_n(\mathbb{K})$, onde $a_{ij} = \langle v_i, v_j\rangle$ para todo $1 \le i, j \le n$ é uma matriz positiva definida.

2. Mostre que se $A = (a_{ij}) \in M_n(\mathbb{K})$ é uma matriz positiva definida, $\underline{\mathrm{B}}$ é uma base de um espaço vetorial sobre $\mathbb{K}$ V então

$$\langle u, v\rangle := [v]^*_{\underline{\mathrm{B}}} A[u]_{\underline{\mathrm{B}}} \quad \text{para todo } u, v \in V$$

define um produto interno em V.

Ex. 6.3 --- Seja V um espaço com produto interno $\langle , \rangle$ de dimensão finita n. Mostre que um conjunto de vetores $\{v_1, v_2, \ldots, v_n\}$ é linearmente independente se, e só se, a matriz $A = (\langle v_i, v_j\rangle) \in M_n(\mathbb{K})$ é não singular.

Ex. 6.4 --- Seja V um $\mathbb{K}$-e.v. com produto interno.

1. Se $\mathbb{K} = \mathbb{R}$, mostre que $\langle u, v\rangle = 0 \Leftrightarrow \|u + v\|^2 = \|u\|^2 + \|v\|^2$

2. Mostre que isso não é verdade se $\mathbb{K} = \mathbb{C}$.

3. Se $\mathbb{K} = \mathbb{C}$, mostre que para $u, v \in V$ $\langle u, v\rangle = 0 \Leftrightarrow \|\alpha u + \beta v\|^2 = \|\alpha u\|^2 + \|\beta v\|^2$, para todos $\alpha, \beta \in \mathbb{C}$.

Ex. 6.5 --- Seja V um $\mathbb{K}$-e.v. com produto interno. Sejam $u, v \in V$, prove:

1. **Lei do Paralelogramo**: $\|u + v\|^2 + \|u - v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$.

2. **Desigualdade de Bessel**: seja $\{v_1, \ldots, v_n\}$ um conjunto ortonormal de V, então $\displaystyle\sum_{k=1}^{n} |\langle v, v_k\rangle|^2 \le \|v\|^2$.

3. **Identidade de Parseval**: seja $\{v_1, \ldots, v_n\}$ uma base ortonormal de V. Mostre que $\langle v, u\rangle = \sum_{k=1}^{n}\langle v, v_k\rangle\langle v_k, u\rangle$.

Ex. 6.6 --- Seja V um $\mathbb{C}$-e.v. com produto interno

1. Fixe vetores $w, u \in V$. Mostre que o operador $T : V \to V$ definido por $T(v) = \langle v, w \rangle u$, admite adjunta e a descreva explicitamente.

2. Suponha $\dim V < \infty$ e que $\langle T(v), v \rangle = 0$, para todo $v \in V$. Mostre que $T = 0$. Mostre que o resultado não é necessariamente verdadeiro se V é um $\mathbb{R}$-e.v.

Ex. 6.7 --- Seja V um $\mathbb{R}$-espaço vetorial de dimensão finita com um produto interno. Seja $W \subseteq V$ um subespaço, mostre que cada classe de equivalência do quociente V/W possui exatamente um vetor ortogonal a W.

Ex. 6.8 --- Determine o mínimo do conjunto : $\left\{ \int_0^1 (t^2 - at - b)^2 \mathrm{d}t : a, b \in \mathbb{R} \right\}$.

Ex. 6.9 --- Seja $V = C([0,1], \mathbb{R})$ o espaço das funções contínuas no intervalo $[0,1]$ com o produto interno $\langle f, g \rangle = \displaystyle\int_0^1 f(t)g(t)dt$.

1. Exiba uma base ortonormal do subespaço de V gerado pelos polinômios 1, t e t^2.

2. Encontre o polinômio de grau menor ou igual a 2 que melhor aproxima $f(t) = \cos t$ no intervalo $[0,1]$.

Ex. 6.10 --- Seja V um $\mathbb{K}$-e.v. com produto interno. Suponha que U e W sejam subespaços de V. Mostre que:

1. $W \subseteq (W^\perp)^\perp$ e se $V = W \oplus W^\perp$ então $W = (W^\perp)^\perp$.

2. $(U + W)^\perp = U^\perp \cap W^\perp$.

3. $U^\perp + W^\perp \subseteq (U \cap W)^\perp$.

4. Se $V = U \oplus U^\perp = W \oplus W^\perp$, então $U^\perp + W^\perp = (U \cap W)^\perp$.

Ex. 6.11 --- Considere $\mathbb{R}[x]$ com o produto interno $\langle p, q \rangle = \displaystyle\int_0^1 p(x)q(x)dx$ e seja W o subespaço constituído por todos os polinômios com termo constante igual a zero. Mostre que $W^\perp = \{0\}$ e então $W^{\perp\perp} = \mathbb{R}[x] \neq W$. Também $\mathbb{R}[x] \neq W \oplus W^\perp$.

Ex. 6.12 --- Seja W um subespaço de dimensão finita de V. Existem, em geral, vários operadores projeções cuja imagem é W. Prove que se $E \in \mathrm{Hom}(V, V)$ é uma projeção cuja imagem é W e $\|E(v)\| \leq \|v\|$, para todo $v \in V$, então E é a projeção ortogonal em W, i.e. $E = E_W$.

Ex. 6.13 --- Seja W um subespaço vetorial de dimensão finita de um espaço V com produto interno. Seja E_W a projeção ortogonal de V em W. Mostre que $\mathrm{I} - E_W$ é a projeção ortogonal de V em $W^\perp$, é um operador projeção e $\ker(\mathrm{I} - E_W) = W$. (Obs: $W^\perp$ não necessariamente terá dimensão finita).

Ex. 6.14 --- Seja V um $\mathbb{K}$-e.v. com produto interno $\langle , \rangle$, $\alpha \in \mathbb{K}$ e sejam $T, S \in \mathrm{Hom}(V, V)$ que admitem adjuntos. Mostre que os seguintes operadores admitem adjuntos:

1. $T + S$ e $(T + S)^* = T^* + S^*$.
2. αT e $(\alpha T)^* = \overline{\alpha} T^*$.
3. $T \circ S$ e $(T \circ S)^* = S^* \circ T^*$.
4. T^* e $(T^*)^* = T$.

   Se $T, S \in \mathrm{Hom}(V, V)$ são operadores normais então é verdade que αT, $T + S$ e $T \circ S$ são normais?

Ex. 6.15 --- Seja T um operador linear em V que possui um adjunto T^*. Mostre que:

1. $\ker T^* = (\mathrm{im}\, T)^\perp$ e que $\mathrm{im}\, T^* \subset (\ker T)^\perp$.
2. Se $\dim V < \infty$, então $\mathrm{im}\, T^* = (\ker T)^\perp$.
3. Se $\dim V < \infty$, T é injetiva se, e somente se, T^*é sobrejetiva e vice-versa.
4. Seja $V = C([0, 1], \mathbb{R})$, com o produto interno $\langle f, g\rangle = \displaystyle\int_0^1 f(t)g(t)dt$, e considere o operador

$$\begin{aligned} T : V &\to V \\ f &\mapsto T(f) : [0, 1] \mapsto \mathbb{R} \\ &\qquad\qquad\qquad t \mapsto tf(t) \end{aligned}$$

   Determine T^* e mostre que $\mathrm{im}\, T^* \neq (\ker T)^\perp$.

Ex. 6.16 --- Seja $T \in L(V)$. Prove que se T é invertível, então T^* é invertível e $(T^*)^{-1} = (T^{-1})^*$.

Ex. 6.17 --- Suponha que V é de dimensão finita. Sejam $T : V \to V$ um operador linear e $\lambda \in K$. Mostre que λ é um autovalor de T se, e somente se, $\overline{\lambda}$ é um autovalor de T^*.

Ex. 6.18 --- Seja V um espaço com produto interno complexo de dimensão finita. Seja $T : V \to V$ um operador linear. Definimos

$$T_1 = \frac{1}{2}(T + T^*), \ T_2 = \frac{1}{2i}(T - T^*).$$

Mostre que

1. T_1 e T_2 são autoadjuntos.
2. $T = T_1 + iT_2$.
3. O que pode dizer sobre a unicidade dessa decomposição?
4. Escreva T^* em função de T_1 e T_2.
5. Encontre condições em T_1 e T_2 para que o operador T seja: i-autoadjunto; ii-unitário; iii-normal.

Ex. 6.19 --- Seja E um operador linear em V tal que $E^2 = E$ e tal que E possui um adjunto E^*. Prove que E é autoadjunto se, e somente se é normal. Prove também que, neste caso, E é a projeção ortogonal em $W = \operatorname{im} E$.

Ex. 6.20 --- Sejam V um espaço vetorial sobre $\mathbb{C}$ e $T \in \operatorname{Hom}(V, V)$. Prove que T é autoadjunto se, e somente se $\langle T(v), v\rangle \in \mathbb{R}$, para todo $v \in V$. Conclua que os autovalores de um operador autoadjunto são reais.

Ex. 6.21 --- Sejam $S, T \in \operatorname{Hom}(V, V)$. Suponha que S e T admitem adjunto. Mostre que:

1. Se S e T são autoadjuntos, então ST é autoadjunto se, e somente se, $ST = TS$.
2. T^*T é autoadjunto.
3. Se T é autoadjunto, então S^*TS é autoadjunto.

Ex. 6.22 --- Suponha que $\dim V$ é finita e seja $T : V \to V$ um operador normal. Mostre que:

1. Se T é nilpotente então $T = 0$.
2. Se T é um operador projeção então $T^* = T$.
3. Se $T^3 = T^2$ então T é um operador projeção.

Ex. 6.23 --- Sejam V um $\mathbb{C}$-e.v. com produto interno de dimensão finita e $T \in \mathrm{Hom}(V, V)$ um operador normal. Prove que:

1. T é autoadjunto se, e somente se, todo autovalor de T é um número real.

2. T é unitário se, e somente se, todo autovalor de T é um número complexo de módulo igual a 1.

3. T é o operador nulo se, e somente se todos os autovalores de T são nulos.

Ex. 6.24 --- Sejam T e S dois operadores autoadjuntos em um espaço vetorial V de dimensão finita. Prove que se T e S comutam, então existe uma base ortonormal de V que diagonaliza estes dois operadores simultaneamente.

Ex. 6.25 --- Um operador linear $T \in \mathrm{Hom}(V, V)$ é *não negativo* se T é autoadjunto e $\langle T(v), v\rangle \geq 0$ para todo $v \in V$. Mostre que se V é um $\mathbb{C}$-e.v. de dimensão finita:

1. T é não negativo se, e somente se, T é normal e os autovalores de T são números reais não negativos.

2. Se T é não negativo então existe um único operador linear não-negativo $R \in \mathrm{Hom}(V, V)$ que é raiz quadrada de T, isto é, que satisfaz $R^2 = T$.

Ex. 6.26 --- Seja $\mathbb{R}_2[x]$ com o produto interno dado por $\langle f, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x$. Considere o operador $T : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ dado por $T(ax^2 + bx + c) = bx$.

1. Mostre que T não é autoadjunto.

2. A matriz de T na base canônica é

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

   que é simétrica. Explique por que isso não contradiz o item 1.

Ex. 6.27 --- Sejam V um $\mathbb{C}$-e.v. com produto interno de dimensão finita e $T \in \mathrm{Hom}(V, V)$ um operador normal. Mostre que:

1. $T = \lambda_1 P_1 + \ldots + \lambda_k P_k$ onde $\mathrm{I} = P_1 + \ldots + P_k$, $P_i P_j = 0$ se $i \neq j$ e $E_i^2 = E_i = E_i^*$.

2. Existe $g \in \mathbb{C}[x]$ tal que $T^* = g(T)$.

3. Todo subespaço de V T-invariante é também T^*-invariante.

Ex. 6.28 --- Seja V um $\mathbb{C}$-espaço vetorial com produto de interno complexo de dimensão finita. Considere $T \in \mathrm{Hom}(V, V)$ um operador qualquer e $\lambda_1, \ldots, \lambda_n$ os autovalores de T (aparecendo tantas vezes quanto sua multiplicidade algébrica). Mostre que:

1. $\sum_{i=1}^{n} |\lambda_i|^2 \leq \mathrm{tr}(T^* \circ T)$
2. A igualdade vale se, e somente se, T é normal.

Ex. 6.29 --- Seja

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \in M_3(\mathbb{R}).$$

Ache uma matriz ortogonal $P \in M_3(\mathbb{R})$ tal que $P^T AP$ é diagonal.

Ex. 6.30 --- Considere um $\mathbb{C}$-espaço vetorial V com o produto interno e dimensão finita. Dizemos que um operador $T : V \to V$ é *anti-hermitiano* ou *anti-autoadjunto* se para quaisquer vetores $u, v \in V$ tem-se que $\langle T(u), v \rangle = - \langle u, T(v) \rangle$. Assuma que T é anti-hermitiano e mostre que:

1. Os autovalores de T são imaginários puros.
2. Se $v_1, v_2 \in V$ são autovetores associados a autovalores distintos $\lambda_1, \lambda_2 \in \mathbb{C}$ então eles são ortogonais.
3. Existe uma base ortonormal de V formada por autovetores de T.

Ex. 6.31 --- Determine a forma polar da matriz

$$A = \begin{bmatrix} 1+i & -1 & 1 \\ 0 & i & i \\ 1 & 0 & 1+i \end{bmatrix} \in M_3(\mathbb{C}).$$

Capítulo 7

Determinantes

Neste capítulo, apresentaremos alguns resultados sobre o determinante de uma matriz quadrada. Nossa abordagem partirá de uma descrição geométrica para o determinante da qual derivaremos as propriedades algébricas.

Nesse capítulo

- Geometria dos Volumes (p. 202)
- Determinantes e Permutações (p. 208)
- Propriedades (p. 210)

7.1 Geometria dos Volumes

O determinante de uma matriz A tem um significado geométrico simples. É o volume orientado da imagem do cubo unitário sob a transformação linear T_A.

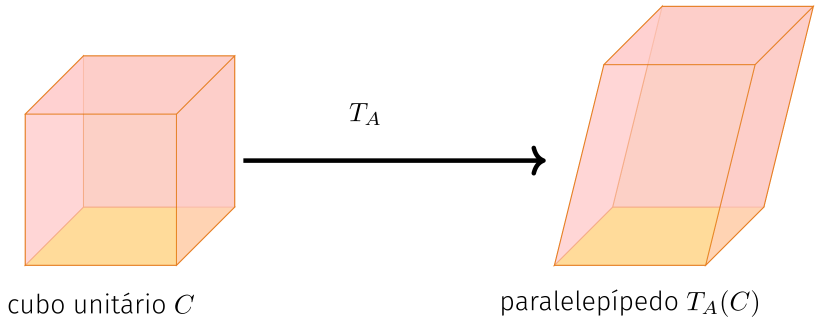


Figura 7.1 Determinante é o volume orientado.

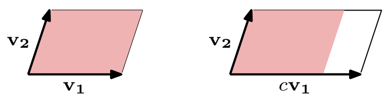
Nesse sentido, começaremos descrevendo quais propriedades o volume orientado deve ter e posteriormente tomaremos essas propriedades como base para a definição algébrica.

Doravante, por preguiça, diremos simplesmente volume no lugar volume orientado.

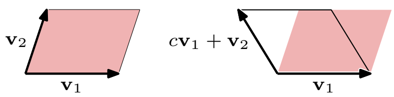
Ao considerar as propriedades que o volume deve ter, suponha que estamos trabalhando em $\mathbb{R}^2$, onde o volume se reduz à área. Seja A a matriz $A = [\boldsymbol{v}_1 \mid \boldsymbol{v}_2]$. O quadrado unitário em $\mathbb{R}^2$ é determinado pelos vetores da base padrão $\boldsymbol{e}_1$ e $\boldsymbol{e}_2$. Assim, nesse caso teremos que $T_A(\boldsymbol{e}_1) = \boldsymbol{v}_1$ e $T_A(\boldsymbol{e}_2) = \boldsymbol{v}_2$, por isso nesse caso queremos determinar a área do paralelogramo P determinado pelos vetores $\boldsymbol{v}_1$ e $\boldsymbol{v}_2$, que são as duas colunas de A.

A área de um paralelogramo certamente deve ter as seguintes duas propriedades:

A1 Se multiplicarmos um lado de P por um número c, por exemplo, se substituirmos o paralelogramo P pelo paralelogramo P' determinado por $c\boldsymbol{v}_1$ e $\boldsymbol{v}_2$, a área de P' deve ser c vezes a área de P.



A2 Se adicionarmos um múltiplo de um lado de P a outro, por exemplo, se substituirmos P pelo paralelogramo P' determinado por $\boldsymbol{v}_1$ e $\boldsymbol{v}_2 + c\boldsymbol{v}_1$, a área de P' deve ser igual à área de P. Podemos compreender esse fato observando que a área de um paralelogramo é base vezes a altura e, enquanto a operação descrita apesar de alterar a forma do paralelogramo, não altera nem sua base nem sua altura.



A propriedade A1, em particular, permanece válida $c = 0$, quando um dos lados se torna o vetor zero; nesse caso, o paralelogramo degenera em uma segmento de reta (ou em um ponto se os dois lados forem o vetor zero) e reta ou um ponto tem área 0.

Agora, consideramos um corpo arbitrário $\mathbb{K}$ e consideramos as matrizes $n \times n$. Ainda somos guiados pelas propriedades A1 e A2, estendendo-as para matrizes $n \times n$ usando a ideia de que se apenas uma ou duas colunas forem alteradas como em A1 ou A2 e a outras não forem alteradas, o volume deverá mudar como descrito em A1 ou A2. Assim, somos levados à seguinte definição.

7.1 Definição Função Volume

Uma função de volume $\mathrm{vol} : \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ é uma função que satisfaz as propriedades:

V1 Para qualquer escalar c e qualquer i,

$$\begin{aligned} &\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_{i-1} \mid c\boldsymbol{v}_i \mid \boldsymbol{v}_{i+1} \mid \cdots \mid \boldsymbol{v}_n\right]\right) && (7.1)\\ &= c\,\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_{i-1} \mid \boldsymbol{v}_i \mid \boldsymbol{v}_{i+1} \mid \cdots \mid \boldsymbol{v}_n\right]\right) && (7.2)\end{aligned}$$

V2 Para qualquer escalar c e qualquer $j \neq i$,

$$\begin{aligned} &\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_{i-1} \mid \boldsymbol{v}_i + c\boldsymbol{v}_j \mid \boldsymbol{v}_{i+1} \mid \cdots \mid \boldsymbol{v}_n\right]\right) && (7.3)\\ &= \mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_{i-1} \mid \boldsymbol{v}_i \mid \boldsymbol{v}_{i+1} \mid \cdots \mid \boldsymbol{v}_n\right]\right) && (7.4)\end{aligned}$$

Observamos que ainda não mostramos que alguma função vol existe, mas prosseguiremos com a suposição de que existe tal função e derivaremos propriedades que essa função deveria ter e finalmente utilizaremos essas propriedades para demonstrar a existência de tal função.

Observamos também que a função vol não é única, pois podemos multiplicá-la por um fator arbitrário obtendo uma nova função volume. Depois que especificarmos uma escala, obteremos uma função única que indicaremos por $\det$, o determinante. Mas é conveniente trabalhar com funções volume arbitrárias e normalizar o resultado somente ao final. A condição de escala é que as funções vol serão redimensionados de modo que o volume orientado cubo unitário n-dimensional, com as colunas dispostas na ordem padrão, seja 1.

7.2 Proposição

Seja $\mathrm{vol} : \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ uma função de volume. Então

1. Se alguma coluna de A for zero, $\mathrm{vol}(A) = 0$.
2. Se as colunas de A forem linearmente dependentes então $\mathrm{vol}(A) = 0$.
   Em particular, se duas colunas de A forem iguais, $\mathrm{vol}(A) = 0$.
3. $$\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_j \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) = -\,\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_j \mid \cdots \mid \boldsymbol{v}_n\right]\right)$$
4. $$\begin{aligned}\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid a\boldsymbol{u} + b\boldsymbol{w} \mid \cdots \mid \boldsymbol{v}_n\right]\right) &= a\,\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{u} \mid \cdots \mid \boldsymbol{v}_n\right]\right) \\ &\quad + b\,\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{w} \mid \cdots \mid \boldsymbol{v}_n\right]\right)\end{aligned}$$

Demonstração.

1. Seja $\boldsymbol{v}_i = 0$. Então $\boldsymbol{v}_i = 0\boldsymbol{v}_i$, então pela propriedade V1

$$\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) = 0 \cdot \mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) = 0.$$

2. Sejam

$$\begin{aligned}\boldsymbol{v}_i &= a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_{i-1}\boldsymbol{v}_{i-1} + a_{i+1}\boldsymbol{v}_{i+1} + \cdots + a_n\boldsymbol{v}_n \\ \boldsymbol{v}_i' &= \widehat{a_1\boldsymbol{v}_1} + a_2\boldsymbol{v}_2 + \cdots + a_{i-1}\boldsymbol{v}_{i-1} + a_{i+1}\boldsymbol{v}_{i+1} + \cdots + a_n\boldsymbol{v}_n\end{aligned}$$

Assim $\boldsymbol{v}_i = a_1\boldsymbol{v}_1 + \boldsymbol{v}_i'$, e aplicando a propriedade V2,

$$\begin{aligned}\mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) &= \mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid a_1\boldsymbol{v}_1 + \boldsymbol{v}_i' \mid \cdots \mid \boldsymbol{v}_n\right]\right) \\ &= \mathrm{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_i' | \cdots \mid \boldsymbol{v}_n\right]\right)\end{aligned}$$

Procedendo da mesma maneira, aplicando a propriedade $\boxed{V2}$ repetidamente, obtemos

$$\operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) = \operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid 0 \mid \cdots \mid \boldsymbol{v}_n\right]\right) = 0.$$

$\boxed{3}$

$$\begin{aligned}
\operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_j \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) &= \operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_j \mid \cdots \mid \boldsymbol{v}_j + \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) \\
&= \operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid -\boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_j + \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) \\
&= \operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid -\boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_j \mid \cdots \mid \boldsymbol{v}_n\right]\right) \\
&= -\operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_j \mid \cdots \mid \boldsymbol{v}_n\right]\right)
\end{aligned}$$

$\boxed{4}$ Se $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}_n\}$ não for linearmente independente. Então, a afirmação é direta.

Suponha que $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}_n\}$ seja linearmente independente. Podemos estender esse conjunto para uma base $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}_n, \boldsymbol{z}\}$ de $\mathbb{K}^n$. Então podemos escrever

$$\begin{aligned}
\boldsymbol{u} &= c_1\boldsymbol{v}_1 + \cdots + c_{i-1}\boldsymbol{v}_{i-1} + c_{i+1}\boldsymbol{v}_{i+1} + \cdots + c_n\boldsymbol{v}_n + c'\boldsymbol{z}, \\
\boldsymbol{w} &= d_1\boldsymbol{v}_1 + \cdots + d_{i-1}\boldsymbol{v}_{i-1} + d_{i+1}\boldsymbol{v}_{i+1} + \cdots + d_n\boldsymbol{v}_n + \boldsymbol{z}.
\end{aligned}$$

Seja $\boldsymbol{v} = a\boldsymbol{u} + b\boldsymbol{w}$. Então

$$\boldsymbol{v} = e_1\boldsymbol{v}_1 + \cdots + e_{i-1}\boldsymbol{v}_{i-1} + e_{i+1}\boldsymbol{v}_{i+1} + \cdots + e_n\boldsymbol{v}_n + e'\boldsymbol{z}$$

onde $e' = ac' + bd'$.

Aplicando a propriedade $\boxed{V2}$ repetidamente e a propriedade $\boxed{V1}$, vemos que

$$\begin{aligned}
\operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v} \mid \cdots \mid \boldsymbol{v}_n\right]\right) &= e' \operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{z} \mid \cdots \mid \boldsymbol{v}_n\right]\right) \\
\operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{u} \mid \cdots \mid \boldsymbol{v}_n\right]\right) &= c' \operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{z} \mid \cdots \mid \boldsymbol{v}_n\right]\right)
\end{aligned}$$

$$\operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{w} \mid \cdots \mid \boldsymbol{v}_n\right]\right) = d' \operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{z} \mid \cdots \mid \boldsymbol{v}_n\right]\right)$$

o que conclui a demonstração desse item. ∎

Observação 3. *O item 3 da Proposição 7.1 implica que*

$$2\operatorname{vol}\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{z} \mid \cdots \mid \boldsymbol{z} \mid \cdots \mid \boldsymbol{v}_n\right]\right) = 0$$

e, portanto,

$$\operatorname{vol}\left(\left[\boldsymbol{v}_1 \cdots \mid \boldsymbol{z} \mid \cdots \mid \boldsymbol{z} \mid \cdots \mid \boldsymbol{v}_n\right]\right) = 0$$

se $\mathbb{K}$ não tiver característica 2.

7.4 Teorema

Uma função $f\colon \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ é uma função de volume se, e somente se, satisfizer:

1. Multilinearidade: Se $A = [\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_n]$ com $\boldsymbol{v}_i = a\boldsymbol{u} + b\boldsymbol{w}$ para algum i,,

$$\begin{aligned} f\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) = a f\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{u} \mid \cdots \mid \boldsymbol{v}_n\right]\right) \\ + b f\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{w} \mid \cdots \mid \boldsymbol{v}_n\right]\right) \end{aligned}$$

2. Alternatividade: Se $A = [\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_n]$ com $\boldsymbol{v}_i = \boldsymbol{v}_j$ para $i \neq j$, então

$$f\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_n\right]\right) = 0.$$

Demonstração. Já provamos que qualquer função de volume satisfaz o Lema 7.1 3 e 4, o que fornece alternatividade e multilinearidade. Por outro lado, é fácil ver que a multilinearidade e a alternatividade fornecem as propriedades *V1* e *V2* na Definição 7.1. ■

As condições do Teorema 7.1 são geralmente consideradas como a definição de uma função de volume.

Observação 5. *Em característica 2, a função* $f\left(\left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]\right) = ac$ *é multilinear e satisfaz* $f\left(\left[\boldsymbol{v}_2 \mid \boldsymbol{v}_1\right]\right) = f\left(\left[\boldsymbol{v}_1 \mid \boldsymbol{v}_2\right]\right) = -f\left(\left[\boldsymbol{v}_1 \mid \boldsymbol{v}_2\right]\right)$*, mas não satisfaz a alternatividade.*

7.6 Definição Determinante

Uma função $\det\colon \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ é dita determinante se satisfizer:

D1 Multilinearidade: Se $A = [\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_n]$ com $\boldsymbol{v}_i = a\boldsymbol{u} + b\boldsymbol{w}$ para algum i,,

$$\begin{aligned} \det\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_i \mid \cdots \mid \boldsymbol{v}_n\right]\right) = a \det\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{u} \mid \cdots \mid \boldsymbol{v}_n\right]\right) \\ + b \det\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{w} \mid \cdots \mid \boldsymbol{v}_n\right]\right) \end{aligned}$$

D2 Alternatividade: Se $A = [\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_n]$ com $\boldsymbol{v}_i = \boldsymbol{v}_j$ para $i \neq j$, então

$$\det\left(\left[\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_n\right]\right) = 0.$$

D3 Volume do cubo unitário

$$\det\left(\left[\boldsymbol{e}_1 \mid \cdots \mid \boldsymbol{e}_n\right]\right) = 1.$$

Primeiramente provaremos a unicidade da função determinante.

7.7 Teorema Unicidade do Determinante

Suponha que exista uma função de volume não trivial $\mathrm{vol} : \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$. Então existe uma única função determinante. Além disso, qualquer função de volume vol é da forma $\mathrm{vol} = a \det$ para algum $a \in \mathbb{K}$

Demonstração. Seja A uma matriz com $\mathrm{vol}(A) \neq 0$. Então, pela Proposição 7.1 $\boxed{2}$, A deve ser não singular. Logo temos que existe uma sequência de operações elementares sobre as colunas que levam A a I. Pela definição 7.1 $\boxed{V1}$ e $\boxed{V2}$ e pela Proposição 7.1 $\boxed{4}$, cada uma dessas operações tem o efeito de multiplicar $\mathrm{vol}(A)$ por um escalar diferente de zero, então $\mathrm{vol}(\mathrm{I}) \neq 0$.

Por outro lado temos que qualquer múltiplo escalar de uma função de volume é uma função de volume; portanto, podemos obter uma função de volume $\det$ fazendo $\det(A) = (1/\mathrm{vol}(\mathrm{I}))\,\mathrm{vol}(A)$ e claramente $\det(\mathrm{I}) = 1$. Em seguida, definimos uma coleção de funções volumes

$$\mathrm{vol}_a(A) = a \det(A).$$

Agora seja f qualquer função de volume. Defina $a = f(\mathrm{I})$. Se A for singular, $f(A) = 0$. Suponha que A não seja singular. Então existe uma sequência de operações de coluna que levam I a A, e cada uma dessas operações de coluna tem o efeito de multiplicar o valor de qualquer função de volume por uma constante diferente de zero, independentemente da escolha da função de volume. Assim, se b é o produto dessas constantes, teremos

$$f(A) = bf(\mathrm{I}) = ba = b\,\mathrm{vol}_a(\mathrm{I}) = \mathrm{vol}_a(A)\ ,$$

e logo $f = \mathrm{vol}_a$. Em particular, se f é alguma função de volume com $f(\mathrm{I}) = 1$, então $f = \det$, o que mostra que $\det$ é único. ■

Observe que a prova desse teorema não mostra que $\det$ existe, pois a priori, poderíamos escolher duas sequências diferentes de operações elementares sobre as colunas para ir de I a A e dessa forma obter dois valores diferentes para $\mathrm{vol}(A)$. De fato, vol existe, como veremos na próxima seção.

7.2 Determinantes e Permutações

Seja $A = [\boldsymbol{v}_1 \mid \ldots \mid \boldsymbol{v}_n]$. Escrevendo cada um desses vetores em termos da base canônica de $\mathbb{R}^n$, obtemos

$$\begin{aligned}
\boldsymbol{v}_1 &= v_{11}\boldsymbol{e}_1 + \ldots + v_{n1}\boldsymbol{e}_n, \\
\boldsymbol{v}_2 &= v_{12}\boldsymbol{e}_1 + \ldots + v_{n2}\boldsymbol{e}_n, \\
\vdots \quad & \quad \vdots \\
\boldsymbol{v}_n &= v_{1n}\boldsymbol{e}_1 + \ldots + v_{nn}\boldsymbol{e}_n
\end{aligned}$$

Dessa forma

$$\det(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = \det(v_{11}\boldsymbol{e}_1 + \ldots + v_{n1}\boldsymbol{e}_n, \ldots, v_{1n}\boldsymbol{e}_1 + \ldots + v_{nn}\boldsymbol{e}_n) \tag{7.5}$$

$$= \sum_{i_1, \ldots, i_n = 1}^{n} v_{i_1 1} v_{i_2 2} \cdots v_{i_n n} \det(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_n}) \tag{7.6}$$

Observamos inicialmente que no somatório 7.6 se anulam todos os termos nos quais ocorre a repetição de algum dos índices $i_1, \ldots, i_n$. Pois nesse caso, $\boldsymbol{e}_{i_k} = \boldsymbol{e}_{i_j}$ e pela propriedade $\boxed{D1}$ do determinante temos que $\det(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_j}, \ldots, \boldsymbol{e}_{i_k}, \ldots, \boldsymbol{e}_{i_n}) = 0$.

Como todos os índices $i_1, \ldots, i_n$ são diferentes entre si no somatório 7.6 precisamos apenas considerar

$$\det(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = \sum_{\sigma \in S_n} v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n} \det(\boldsymbol{e}_{\sigma(1)}, \ldots, \boldsymbol{e}_{\sigma(n)}),$$

em que σ percorre o grupo de permutações de $S = \{1, \ldots, n\}$. Se σ é uma permutação , podemos considerar sua decomposição como produto de transposição. Pela propriedade do determinante, uma transposição altera o valor de $\det$ pelo fator -1. Se σ é o produto de k transposições, então o sinal de $\det$ será alterado por $(-1)^k$. Dessa forma temos que

$$\det(\boldsymbol{e}_{\sigma(1)}, \ldots, \boldsymbol{e}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \det(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = \operatorname{sgn}(\sigma).$$

E logo

$$\det(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n},$$

que é a expressão do determinante em termos de permutações.

Podemos agora demonstrar da existência do determinante.

7.1 Teorema Existência do Determinante

Existe uma única função determinante. Ou seja, uma função de volume $\det$:

$\mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ com $\det(\mathrm{I}) = 1$.

Demonstração. Já demonstramos a unicidade. Mostraremos agora que

$$\det(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n}, \tag{7.7}$$

satisfaz às propriedades D1 - D3 da função determinante. No que se segue σ denotará uma permutação do conjunto $S = \{1, \dots, n\}$

D1 Suponhamos que os vetores $\boldsymbol{v}_i$ e $\boldsymbol{v}_j$ sejam iguais. Seja τ a transposição entre $\boldsymbol{v}_i$ e $\boldsymbol{v}_j$. Então $v_{\tau\sigma(1)1} v_{\tau\sigma(2)2} \cdots v_{\tau\sigma(n)n} = v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n}$, pois τ transpõe os vetores $\boldsymbol{v}_i$ e $\boldsymbol{v}_j$, que são iguais, e mantém fixos os outros vetores. Assim,

$$\begin{aligned}
\det(\dots,\ \boldsymbol{v}_i,\ \dots,\ \boldsymbol{v}_j,\ \dots) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) v_{\sigma(1)1} v_{\sigma(2)2} \cdots v_{\sigma(n)n} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) v_{\tau\sigma(1)1} v_{\tau\sigma(2)2} \cdots v_{\tau\sigma(n)n} \\
&= -\sum_{\sigma \in S_n} \operatorname{sgn}(\tau\sigma) v_{\tau\sigma(1)1} v_{\tau\sigma(2)2} \cdots v_{\tau\sigma(n)n} \\
&= -\det(\dots,\ \boldsymbol{v}_j, \dots,\ \boldsymbol{v}_i, \dots)\ ,
\end{aligned}$$

Como $\det(\dots,\ \boldsymbol{v}_i,\ \dots,\ \boldsymbol{v}_j,\ \dots) = \det(\dots,\ \boldsymbol{v}_j,\ \dots,\ \boldsymbol{v}_i,\ \dots)$, segue que $\det(\dots,\ \boldsymbol{v}_i,\ \dots,\ \boldsymbol{v}_j,\ \dots) = 0$.

D1 A linearidade é imediata, notando que cada parcela de 7.7 contém exatamente uma coordenada do vetor $\boldsymbol{v}_i + k\boldsymbol{u}_i$, de modo que

$$\begin{aligned}
&\det(\boldsymbol{v}_1,\ \dots,\ \boldsymbol{v}_i + k\boldsymbol{u}_i,\ \dots,\ \boldsymbol{v}_n) \\
&= \det(\boldsymbol{v}_1,\ \dots,\ \boldsymbol{v}_i,\ \dots,\ \boldsymbol{v}_n) + k \det(\boldsymbol{v}_1,\ \dots,\ \boldsymbol{u}_i,\ \dots,\ \boldsymbol{v}_n).
\end{aligned}$$

D3 Vamos mostrar que $\det(\boldsymbol{e}_1, \dots,\ \boldsymbol{e}_n) = 1$.

Se $\sigma(i) \neq i$, então $v_{\sigma(i)i} = 0$. Assim, apenas a permutação identidade, produz um termo não-nulo. No caso da permutação identidade, temos que o sinal é 1 e todos os termos $v_{\sigma(i)i}$ são iguais a 1, e logo $\det(\boldsymbol{e}_1, \dots,\ \boldsymbol{e}_n) = 1$. ∎

7.2 Proposição

Seja $A \in \mathcal{M}_{n,n}(\mathbb{K})$. Então $\det(A) \neq 0$ se, e somente se, A não for singular.

Demonstração. Pela Proposição 7.1, para qualquer função de volume $\mathrm{vol}_a, \mathrm{vol}_a(A) = 0$ se A for singular. Por outro lado, para qualquer função de volume não trivial, ou seja, para qualquer função vol_a com $a \neq 0$, observamos no decorrer da prova do Teorema 7.1 que, para qualquer matriz não singular $A, \mathrm{vol}_a(A) = c\,\mathrm{vol}_a(\mathrm{I}) = ca$ para $c \neq 0$. ∎

Notação 3. *O determinante de A, denotado por* $\det A$ *ou* $|A|$*, pode ser denotado diretamente em termos de entradas da matriz escrevendo barras em vez de colchetes:*

$$\det A = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix}.$$

7.3 Propriedades

Agora derivamos algumas propriedades importantes do determinante.

7.1 Teorema

Sejam $A, B \in \mathcal{M}_{n,n}(\mathbb{K})$. Então

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

Demonstração. Defina uma função $f\colon \mathcal{M}_{n,n}(\mathbb{K}) \to \mathrm{V}$ por $f(B) = \det(AB)$. É simples verificar se f é multilinear e alternado; portanto, f é uma função de volume $f(B) = \mathrm{vol}_a(B) = a\det(B)$ onde $a = f(\mathrm{I}) = \det(A\,\mathrm{I}) = \det(A)$. ■

7.2 Proposição

1. $\det(A) \neq 0$ se, e somente se, A for invertível.
2. Se A for invertível, $\det(A^{-1}) = 1/\det(A)$.
3. Se A for invertível, então para qualquer matriz $B, \det(ABA^{-1}) = \det(B)$.

Demonstração. Já vimos no Lema 7.1 que, para qualquer função de volume, $f, f(A) = 0$, se A não for invertível. Se A é invertível, temos $1 = \det(\mathrm{I}) = \det(AA^{-1}) = \det(A)\det(A^{-1})$ a partir do qual o corolário segue. ■

7.3 Proposição

1. Seja A uma matriz diagonal. Então $\det(A)$ é o produto de suas entradas diagonais.
2. De maneira mais geral, seja A uma matriz triangular superior ou uma

triangular inferior. Então $\det(A)$ é o produto de suas entradas diagonais.

Demonstração.

1 Se A é diagonal, existe apenas um termo diferente de zero na Equação 7.7, o termo correspondente à permutação identidade ($\sigma(i) = i$ para todo i), que possui sinal $+$ l.

2 Se σ não for a identidade, haverá um j com $\sigma(j) < j$ e um k com $\sigma(k) > k$, portanto, para um triangular matriz há novamente apenas o termo diagonal. □ ■

7.4 **Teorema**

1 Seja M uma matriz diagonal por bloco,

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

Então $\det(M) = \det(A)\det(D)$.

2 Em geral, seja M uma matriz triangular superior por bloco ou uma matriz triangular inferior por bloco,

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \quad \text{ou} \quad M = \begin{bmatrix} A & 0 \\ C & D \end{bmatrix}$$

Então $\det(M) = \det(A)\det(D)$.

Demonstração. 1 Defina uma função $f \colon \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ por

$$f(D) = \det\left(\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}\right).$$

Então f é multilinear e alternado, então $f(D) = f(\mathrm{I})\det(D)$. Mas $f(\mathrm{I}) = \det\left(\begin{bmatrix} A & 0 \\ \mathrm{I} & 0 \end{bmatrix}\right) = \det(A)$.(É fácil ver esta última igualdade como qualquer permutação que contribua com zero para $\det\left(\begin{bmatrix} A & 0 \\ \mathrm{I} & 0 \end{bmatrix}\right)$ deve corrigir todos, exceto (possivelmente) as primeiras entradas de n.)

2 Suponha que M seja triangular superior (a caixa triangular inferior é semelhante). Se A for singular, haverá um vetor $v \neq 0$ com $Av = 0$. Então seja w o vetor cujas primeiras entradas n são as de v e as entradas restantes são 0. Então $Mw = 0$. Assim, M também é singular e $0 = 0 \cdot \det(D)$.

Suponha que A não seja singular. Então

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} \mathrm{I} & A^{-1}B \\ 0 & \mathrm{I} \end{bmatrix}.$$

■

A primeira matriz do lado direito possui o determinante $\det(A)\det(D)$, e a segunda matriz do lado direito tem o determinante 1, pois é triangular superior, e o teorema segue. □

7.5 Proposição

Seja A^{t} a matriz transposta de A. Então $\det(A^{\mathrm{t}}) = \det(A)$.

Demonstração. Se $\sigma \in S_n, \operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$. Seja $B = (b_{ij}) = A^{\mathrm{t}}$. Então

$$\begin{aligned}
\det(A) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)} \\
&= \sum_{\sigma^{-1} \in S_n} \operatorname{sgn}(\sigma^{-1}) b_{\sigma^{-1}(1),1} \cdots b_{\sigma^{-1}(n),n} \\
&= \det(A^{\mathrm{t}})
\end{aligned}$$

■

Denote por A_{ij} o menor $(i,\ j)$ da matriz A, i.e., a submatriz obtida pela exclusão da linha i e da coluna j de A.

7.6 Teorema Expansão de Laplace

Seja A uma matriz n por n, $A = (a_{ij})$.

1. Para qualquer i,

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

2. Para qualquer j,

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

3. Para qualquer i e para qualquer $k \neq i$,

$$0 = \sum_{j=1}^{n} (-1)^{i+j} a_{kj} \det(A_{ij})\ .$$

$\boxed{4}$ Para qualquer j e para qualquer $k \neq j$,

$$0 = \sum_{i=1}^{n} (-1)^{i+j} a_{ik} \det(A_{ij}) \ .$$

Demonstração. Demonstraremos $\boxed{1}$ e $\boxed{3}$ simultaneamente, portanto fixamos k (que pode ou não ser igual a i).

A soma do lado direito é a soma das funções multilineares e, portanto, é multilinear.(Isso também é fácil de ver diretamente.)

Agora mostramos que é alternado. Seja A uma matriz com as colunas p e q iguais, onde $1 \leq pqn$. Se $j \neq p, q$, então A_{ij} for uma matriz com duas colunas iguais, então $\det(A_{ij}) = 0$. Assim, os únicos dois termos que contribuem para a soma são

$$(-1)^{i+p} a_{kp} \det(A_{ip}) + (-1)^{i+q} a_{kq} \det(A_{iq}) \ .$$

Por hipótese, $a_{kq} = a_{kp}$. Agora

$$A_{ip} = [\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_{p-1} \mid \boldsymbol{v}_{p+1} \mid \cdots \mid \boldsymbol{v}_{q-1} \mid \boldsymbol{v}_q \mid \boldsymbol{v}_{q+1} \mid \cdots \mid \boldsymbol{v}_n],$$

$$A_{iq} = [\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_{p-1} \mid \boldsymbol{v}_p \mid \boldsymbol{v}_{p+1} \mid \cdots \mid \boldsymbol{v}_{q-1} \mid \boldsymbol{v}_{q+1} \mid \cdots \mid \boldsymbol{v}_n].$$

onde $\boldsymbol{v}_m$ indica a coluna m da matriz obtida de A excluindo a linha i de A. Por hipótese, $\boldsymbol{v}_p = \boldsymbol{v}_q$, portanto, essas duas matrizes têm as mesmas colunas, mas em uma ordem diferente. Passamos do primeiro para o segundo, executando sucessivamente trocas de coluna $qp - 1$ (primeiro alternando $\boldsymbol{v}_q$ e $\boldsymbol{v}_{q-1}$ e depois alternando $\boldsymbol{v}_q$ e $\boldsymbol{v}_{q-2}, \ldots$ e, finalmente, alternando $\boldsymbol{v}_q$ e $\boldsymbol{v}_{p+1}$), então $\det(A_{iq}) = (-1)^{qp-1} \det(A_{ip})$. Assim, vemos que a contribuição desses dois termos para a soma é

$$(-1)^{i+p} a_{kp} \det(A_{ip}) + (-1)^{i+q} a_{kp} (-1)^{qp-1} \det(A_{ip})$$

e como $(-1)^{i+p}$ e $(-1)^{i+2q-p-1}$ sempre têm sinais opostos, eles cancelam.

Pela unicidade, o lado direito é um múltiplo $a \det(A)$ para a. Uma computação mostra que, se $A = \mathrm{I}$, o lado direito fornece 1 se $k = i$ e 0 se $k \neq i$, comprovando o teorema nesses casos. ■

Exemplo 7. *Para $A \in \mathcal{M}_{3,3}(\mathbb{K})$ temos*

$$\det A = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} \square & \square & \square \\ \square & e & f \\ \square & h & i \end{vmatrix} - b \begin{vmatrix} \square & \square & \square \\ d & \square & f \\ g & \square & i \end{vmatrix} + c \begin{vmatrix} \square & \square & \square \\ d & e & \square \\ g & h & \square \end{vmatrix} \tag{7.8}$$

$$= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \tag{7.9}$$

$$= aei + bfg + cdh - ceg - bdi - afh. \tag{7.10}$$

◁

7.8 Definição

Dada $A \in \mathcal{M}_{n,n}(\mathbb{K})$

- O **cofator** de um elemento $a_{i,j}$ da matriz quadrada A é o número $C_{i,j} = (-1)^{i+j}\ A_{i,j}$, sendo $A_{i,j}$ o determinante da matriz obtida a partir da matriz original A eliminando-se a linha e a coluna que contenham o elemento $a_{i,j}$.
- A **matriz dos cofatores** C é a matriz formada pelos respectivos cofatores dos elementos da matriz A.
- A **matriz adjunta** da matriz A é a transposta de sua matriz de cofatores.

7.9 Proposição

Para qualquer matriz A,

1. $(\mathrm{Adj}(A)) = A(\mathrm{Adj}(A)) = \det(A)\,\mathrm{I}$.
2. Se A for invertível,
$$A^{-1} = \frac{1}{\det(A)}\mathrm{Adj}(A).$$

Demonstração. A prova de 1 pode ser verificada por uma computação que segue diretamente do Teorema 7.3. Então 2 segue imediatamente. ∎

7.10 Teorema Regra de Cramer

Seja $A \in \mathcal{M}_{n,n}(\mathbb{K})$ uma matriz invertível e $\boldsymbol{b}$ seja um vetor em $\mathbb{K}^n$. Considere o sistema linear $A\boldsymbol{x} = \boldsymbol{b}$. Então:

1. Existe um único vetor $\boldsymbol{x}$ em $\mathbb{K}^n$ que resolve o sistema linear $A\boldsymbol{x} = \boldsymbol{b}$.
2. Se $\boldsymbol{x} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^t$. Então:
$$x_i = \frac{\det(A_i(b))}{\det(A)}, \quad 1 \le i \le n$$
onde $A_i(b)$ é a matriz obtida de A substituindo sua $i-$ésima coluna por b.

Demonstração. Denotaremos as colunas de A por $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. Por linearidade, basta provar o corolário para todos os elementos de qualquer base, $\underline{\mathrm{B}}$ de $\mathbb{K}^n$. Escolhemos a base $\underline{\mathrm{B}} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$.

Corrija j e considere $Ax = \boldsymbol{a}_i$. Então $A_i(\boldsymbol{a}_i) = A$, então a fórmula acima fornece $x_i = 1$. Para $j \neq i, A_i(\boldsymbol{a}_j)$ é uma matriz com duas colunas idênticas; portanto, a fórmula acima fornece $x_j = 0$. Assim, $x = \boldsymbol{e}_i$, o i-ésimo vetor da base canônica, e de fato $A\boldsymbol{e}_i = \boldsymbol{a}_i$. ■

Esta fórmula é de interesse teórico, mas quase nunca deve ser usada na prática.

7.11 Definição Posto por Determinante

Se a matriz A possuir uma submatriz k por k com determinante diferente de zero, mas não possuir uma submatriz $(k+1)$ por $(k+1)$ com determinante diferente de zero, então dizemos que o posto por determinante de A é k. O posto por determinante de A será denotado por $\text{posto}_{\text{det}}(A)$.

7.12 Teorema

Seja A uma matriz. Então o posto por linha, o posto por coluna e o posto por determinante de A são iguais.

$$\text{posto}(A) = \text{posto}_{\text{det}}(A) = \text{posto}_l(A) = \text{posto}_c(A)$$

Demonstração. Já demonstramos que o posto por linha e por coluna de A são iguais no Teorema 2.6.1 Agora, mostraremos que o posto por da coluna de A é igual à posto por determinante de A.

Escreva $A = [\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_n]$, onde A é m por n. Suponha que A tenha uma submatriz k por k B com determinante diferente de zero. Por uma questão de simplicidade, assumimos que B é o canto superior esquerdo de A. Suponha que B seja k por k. Seja $\pi \colon \mathbb{K}^m \to \mathbb{K}^k$ definido por

$$\pi\left(\begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}\right) = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$$

Então $B = [\pi(\boldsymbol{v}_1) \mid \ldots \| \pi(\boldsymbol{v}_k)]$. Como $\det(B) \neq 0, B$ não é singular, então $\{\pi(\boldsymbol{v}_1), \ \ldots \ , i(\boldsymbol{v}_k)\}$ são linearmente independentes, e, portanto, $\{\boldsymbol{v}_1, \ \ldots \ , \ \boldsymbol{v}_k\}$ são linearmente independente. Mas esse conjunto gera um subespaço dimensional de k do espaço da coluna de A, portanto A tem umo posto por de coluna de pelo menos k.

Por outro lado, suponha que A tenha k colunas linearmente independentes. Novamente, por simplicidade, suponha que essas sejam as colunas k mais à esquerda de A. Agora $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ são linearmente independentes e $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$ geram $\mathbb{K}^m$, então $\{\boldsymbol{v}_1, \ \ldots \ , \ \boldsymbol{v}_k, \ \boldsymbol{e}_1, \ \ldots \ , \ \boldsymbol{e}_m\}$ geram $\mathbb{K}^m$ também. Então, existe uma base, $\underline{B}$ de $\mathbb{K}^m$ com $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \underline{B} \subseteq \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$. Escreva $\underline{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_m\}$ e observe que, para todo $i \geq k+1, \boldsymbol{v}_i = \boldsymbol{e}_j$ para algum j. Forme a matriz $B' = [\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_k \mid \boldsymbol{v}_{k+1} \mid \cdots \mid \boldsymbol{v}_n]$ e observe que $\det(B') \neq 0$. Expanda por menores de colunas $n, n-1, \ldots, k+1$ para obter $0 \neq \det(B') = \pm\det(B)$ em que B é sub-matriz k por k de A, então A tem umo posto por determinante de pelo menos k.

■

Definimos o determinante para matrizes. Podemos definir o determinante para transformações lineares $T \colon V \to V$, em que V é um espaço vetorial dimensional limitado.

7.13 Definição Determinante de um Operador

Seja $T \colon V \to V$ uma transformação linear com V um espaço vetorial de dimensão finita. O **determinante** $\det(T)$ é definido como $\det(T) = \det\left([(T)_{\underline{B}}]\right)$ onde $\underline{B}$ é uma base de V.

Para verificar que o determinante está bem definido, precisamos saber que é independente da escolha de a base, $\underline{B}$. Isso segue imediatamente do Teorema de Mudança de Base 3.4.1 e da Proposição 7.3.

Exercícios

Ex. 7.1 --- Calcule o seguinte determinante, onde $a_i := \sum_{k=1}^{i} k$:

$$\begin{vmatrix} a_1 & a_1 & \cdots & \cdots & a_1 \\ a_1 & a_2 & a_2 & \cdots & a_2 \\ \vdots & a_2 & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & a_{n-1} & a_{n-1} \\ a_1 & a_2 & \cdots & a_{n-1} & a_n \end{vmatrix}$$

Ex. 7.2 --- Calcule o seguinte determinante $n \times n$, onde $a, b \in \mathbb{K}$:

$$D_n(a,b) = \begin{vmatrix} a+b & ab & 0 & \cdots & 0 \\ 1 & a+b & ab & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & a+b & ab \\ 0 & \cdots & 0 & 1 & a+b \end{vmatrix}$$

Ex. 7.3 --- Seja $V = V_1 \oplus V_2$ uma soma direta de dois subespaços V_1 e V_2. Sejam $f_i \in \mathcal{L}(V_i)$ para $i = 1, 2$. Defina o endomorfismo $f \in \mathrm{Hom}(V, V)$ dado por $f(v) = f_1(v_1) + f_2(v_2)$ onde $v = v_1 + v_2$ com $v_i \in V_i$ para $i = 1, 2$. Prove que $\det(f) = \det(f_1) \cdot \det(f_2)$.

Ex. 7.4 --- Prove que o determinante da matriz $n \times n$ de entradas $a_{ij} = 1 - \delta_{ij}$ é igual a $(n-1)(-1)^{n-1}$.

Ex. 7.5 --- **[Determinante de Gram]** Sejam $f_1, \ldots, f_n : X \to \mathbb{K}$ funções definidas em um conjunto arbitrário X. Mostre que $\{f_1, \ldots, f_n\}$ é linearmente independenteem $\mathcal{F}(X, \mathbb{K})$ se, e somente se, existem $x_1, \ldots, x_n \in X$ tais que:

$$\begin{vmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_n(x_n) \end{vmatrix} \neq 0.$$

Ex. 7.6 --- Seja $A \in M_n(\mathbb{Z})$ uma matriz com coeficientes inteiros. Mostre que A é invertível se, e somente se, $\det(A) = \pm 1$ (i.e. se, e somente se, $\det(A)$ é invertível em $\mathbb{Z}$).

Ex. 7.7 --- **[Determinante de Vandermonde]** Dado um inteiro $n \geq 2$ e escalares $a_1, \ldots, a_n \in \mathbb{K}$ denote por:

$$\mathcal{V}_n(a_1, \ldots, a_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{vmatrix}$$

ao chamado determinante de Vandermonde. Denote também por $p \in \mathbb{K}[x]$ ao polinômio $p = \mathcal{V}_n(a_1, \ldots, a_{n-1}, x)$.

1. Determine $\mathrm{grau}(p)$ e suas raízes.

2. Deduza uma expressão de p em função de $\mathcal{V}_{n-1}(a_1, \ldots, a_{n-1})$.

3. Deduza uma expressão explicita para $\mathcal{V}_n(a_1, \ldots, a_{n-1}, a_n)$.

Ex. 7.8 --- **[Determinante das multiplicações à esquerda e à direita]** Dada uma matriz $A \in M_n(\mathbb{K})$, denote por

$$\begin{aligned} L_A, R_A : M_n(\mathbb{K}) &\to M_n(\mathbb{K}) \\ L_A(X) &= A \cdot X \\ R_A(X) &= X \cdot A \end{aligned}$$

as multiplicações à esquerda e à direita por A, respectivamente.

1. Verifique que L_A e R_A são lineares.

2. Mostre que $\det(L_A) = \det(R_A) = (\det A)^n$. **Dica**: escolha cuidadosamente uma base de $M_n(\mathbb{K})$.

Ex. 7.9 --- [Wronskiano] Dado o conjunto $\{f_1, \ldots, f_n\}$ de funções no $\mathbb{K}$-e.v. $\mathcal{F}(U, \mathbb{K})$ onde $U \subseteq \mathbb{R}$ é um aberto. Suponha ainda que as funções f_i são $(n-1)$-vezes diferenciáveis, então para todo $x \in U$, define-se o *Wronskiano* como sendo a aplicação $\mathcal{W}(f_1, \ldots, f_n) : U \to \mathbb{K}$ dada por:

$$\mathcal{W}(f_1, \ldots, f_n)(x) := \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ f_1''(x) & f_2''(x) & \cdots & f_n''(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}.$$

Mostre que se existe $x \in U$ tal que $\mathcal{W}(f_1, \ldots, f_n)(x) \neq 0$ então $\{f_1, \ldots, f_n\}$ é linearmente independente. Vale a recíproca?

Ex. 7.10 --- Seja V um $\mathbb{K}$-e.v. de dimensão n e V^* seu espaço dual.

1. Sejam Δ e Δ^* duas funções determinantes em V e V^*, respectivamente. Mostre que a aplicação $(2n)$-linear $\Delta^* \boxtimes \Delta$, assume o mesmo valor em quaisquer dois sistemas de bases duais, i.e.

$$\Delta^* \boxtimes \Delta(v_1^*, \ldots, v_n^*, v_1, \ldots, v_n) = \Delta^* \boxtimes \Delta(w_1^*, \ldots, w_n^*, w_1, \ldots, w_n).$$

2. Duas funções determinante $\Delta \neq 0$ e $\Delta^* \neq 0$ em V e V^*, respectivamente, satisfazendo

$$\Delta^*(v_1^*, \ldots, v_n^*) \cdot \Delta(v_1, \ldots, v_n) = 1 \text{ sempre que } v_i^*(v_j) = \delta_{ij}.$$

   são chamadas *duais*.

   Seja $\Delta \neq 0$ uma função determinante em V. Defina a função $\Delta^* : \underbrace{V^* \times \cdots \times V^*}_{n\text{-vezes}} \to \mathbb{K}$ como segue:

3. Se os vetores $v_1^*, \ldots, v_n^* \in V^*$ são linearmente dependente então $\Delta^*(v_1^*, \ldots, v_n^*) = 0$.

4. Se os vetores $v_1^*, \ldots, v_n^* \in V^*$ sejam linearmente independentes(logo eles formam uma base $\underline{C}$ de V^*) então $\Delta^*(v_1^*, \ldots, v_n^*) = \frac{1}{\Delta(v_1, \ldots, v_n)}$ onde $\underline{B} = \{v_1, \ldots, v_n\}$ é a base de V tal que $\underline{B}^* = \underline{C}$. Prove que Δ^* é uma função

determinante em V^* e que as funções determinantes Δ e Δ^* são duais. Conclua que para todo função determinante não trivial em V existe exatamente uma função determinante dual em V^*.

5. Sejam Δ e Δ^* funções determinantes duais, mostre que

$$\Delta^* \boxtimes \Delta(v_1^*, \dots, v_n^*, u_1, \dots, u_n) = \begin{vmatrix} v_1^*(u_1) & v_1^*(u_2) & \cdots & v_1^*(u_n) \\ v_2^*(u_1) & v_2^*(u_2) & \cdots & v_2^*(u_n) \\ \vdots & \vdots & \ddots & \vdots \\ v_n^*(u_1) & v_n^*(u_2) & \cdots & v_n^*(u_n) \end{vmatrix}.$$

Capítulo 8

Estrutura dos Operadores Lineares

O objetivo deste capítulo e dos próximos é entender a estrutura dos operadores lineares. Boa parte dessa tarefa é o de encontrar bases apropriadas nas quais o operador linear possua uma representação ``simples'' ou de outra forma é encontrar bases nas quais a imagem geométrica da transformação linear $T : V \to W$ é mais compreensível.

Quando os espaços V e W não estão relacionados, ou seja, podemos escolher de modo independente uma base para V e uma base para W então a resposta é muito simples e é dada pelo Teorema 8.1. Uma imagem muito mais interessante e variada é obtida quando $W = V$ e nesse caso queremos escolher a mesma base no domínio e contradomínio. Nesse caso apresentaremos múltiplas respostas: forma de Schur, forma normal de Jordan e forma canônica racional.

Em linguagem matricial, estamos falando em colocar a matriz de T em uma representação matricial mais simples, com a ajuda de uma base apropriada, especialmente adaptada à estrutura de T. No primeiro caso, as bases de V e W podem ser selecionadas independentemente; no segundo caso, estamos falando de uma base em V o menor grau de liberdade de escolha leva a uma maior complexidade e variedade de respostas.

Essa forma matricial mais simples é dita forma canônica ou normal de uma matriz. Uma forma canônica divide as matrizes para as quais está definidas em conjuntos de matrizes, cada uma com o mesma forma canônica e essa matriz da forma canônica serve como representante.

Nesse capítulo

- Equivalência de matrizes (p. 221)
- Operadores Similares (p. 223)
- Autovalores e Autovetores (p. 226)
- Teoremas de Schur e Cayley-Hamilton (p. 234)
- Teorema da Decomposição Primária (p. 239)
- Diagonalizabilidade (p. 243)

Forma Canônica/Normal
Em geral, uma forma canônica especifica uma representação única para todo objeto, enquanto uma forma normal não temos a exigência de unicidade.
Uma forma canônica resolve um problema de classificação e nos fornece ainda mais dados: ela não apenas classifica todas as classes, mas fornece um elemento distinto (canônico) de cada classe.

8.1 Equivalência de matrizes

No Teorema 3.4.1 demonstramos que a mudança de bases é dada por

$$[T]_{\underline{B}',\underline{C}'} = M_{\underline{C}\to\underline{C}'}\,[T]_{\underline{B},\underline{C}}\,M^{-1}_{\underline{B}\to\underline{B}'} \text{ e logo}$$
$$[T]_{\underline{B}',\underline{C}'} = P[T]_{\underline{B},\underline{C}}Q^{-1}$$

onde P e Q são matrizes invertíveis. Isso motiva a seguinte definição.

8.1 Definição Matrizes Equivalentes

Duas matrizes A e B são **equivalentes** se existirem matrizes inversíveis P e Q para as quais

$$B = PAQ^{-1}$$

Observamos que B é equivalente a A se, e somente se, B puder ser obtido de A por uma série de operações elementares de linha e coluna. Executar as operações da linha é equivalente a multiplicar a matriz A à esquerda por P e executar as operações da coluna é equivalente a multiplicar A à direita por Q^{-1}.

De acordo com o Teorema 3.4.1, se A e B são matrizes que representam T em relação a possíveis bases ordenadas diferentes, então A e B são equivalentes. A recíproca também vale.

8.2 Teorema

Sejam V e W espaços vetoriais com $\dim(V) = n$ e $\dim(W) = m$. Então as matrizes A e $B \in \mathcal{M}_{n,m}(\mathbb{K})$ são equivalentes se, e somente se, representam a mesma transformação linear $T \in \mathrm{Hom}(V,\ W)$, mas possivelmente em relação a diferentes bases ordenadas.

Demonstração. Se A e B representam T, isto é, se

$$A = [T]_{\underline{B},\underline{C}} \quad \text{e} \quad B = [T]_{\underline{B}',\underline{C}'}$$

para bases ordenadas $\underline{B}, \underline{C}, \underline{B}'$ e $\underline{C}'$, então A e B são equivalentes. Agora, suponha que A e B sejam equivalentes, digamos

$$B = PAQ^{-1}$$

onde P e Q são invertíveis. Suponha também que A represente uma transformação linear $T \in \mathrm{Hom}(V,\ W)$ para algumas bases ordenadas $\underline{B}$ e $\underline{C}$, ou seja,

$$A = [T]_{\underline{B},\underline{C}}$$

O Teorema 3.2 implica que existe uma única base ordenada $\underline{B}'$ para V para a qual $Q = M_{\underline{B},\underline{B}'}$ e uma única base ordenada $\underline{C}'$ por W para a qual $P = M_{\underline{C},\underline{C}'}$. Consequentemente

$$B = M_{\underline{C},\underline{C}'}[T]_{\underline{B},\underline{C}}M_{\underline{B}',\underline{B}} = [T]_{\underline{B}',\underline{C}'}$$

Portanto, B também representa T. Por simetria, vemos que A e B representam o mesmo conjunto de transformações lineares. Isso completa a demonstração. ■

Vamos provar agora que toda matriz é equivalente a exatamente uma matriz da forma de bloco

$$J_k = \begin{bmatrix} \mathrm{I}_k & 0_{k,n-k} \\ 0_{m-k,k} & 0_{m-k,n-k} \end{bmatrix}$$

8.3 Teorema da Forma Normal Zero-Um

Seja $T : V \to W$ uma transformação linear de espaços dimensionais de dimensão finita. Então:

1. Existe uma decomposição direta $V = V_0 \oplus V_1, W = W_1 \oplus W_2$ tal que $\ker T = V_0$ e T induz um isomorfismo de V_1 em W_1.

2. Existem bases em V e W tais que a matriz de T nessas bases tem a forma (a_{ij}), com $a_{ii} = 1$ para $1 \leq i \leq r$ e $a_{ij} = 0$ para os valores restantes de i, j.

3. Seja $A \in \mathcal{M}_{m,n}(\mathbb{K})$. Então, existem matrizes quadradas não singulares $B \in \mathcal{M}_{m,m}(\mathbb{K})$ e $C \in \mathcal{M}_{n,n}(\mathbb{K})$ e um número $r \leq \min(m,n)$ tal que a matriz BAC tenha o formato descrito no item anterior. O número T é único e igual ao posto de A.

Demonstração.

1. Defina $V_0 = \ker T$ e faça V_1 um complemento direto de V_0. Em seguida, defina $W_1 = \operatorname{im} T$ e faça W_2 um complemento direto de W_1. Precisamos apenas verificar se T determina um isomorfismo de V_1 a W_1.

   A aplicação $T : V_1 \to W_1$ é injetiva, porque o núcleo de T, ou seja, V_0, intercepta V_1 apenas na origem. É sobrejetiva, porque se $\boldsymbol{v} = \boldsymbol{v}_0 + \boldsymbol{v}_1 \in V, \boldsymbol{v}_0 \in V_0, \boldsymbol{v}_1 \in V_1$, então $T(\boldsymbol{v}) = T(\boldsymbol{v}_1)$.

2. Faremos $r = \dim V_1 = \dim W_1$ e escolheremos a base $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r, \boldsymbol{e}_{r+1}, \ldots, \boldsymbol{e}_n\}$ de V, onde os primeiros vetores r formam um base de V_1 e os vetores restantes formam uma base de V_0. Além disso, os vetores $\boldsymbol{e}'_i = T(\boldsymbol{e}_i)$, $1 \leq i \leq r$ formam uma base de $W_1 = \operatorname{im} T$. Estendemos esse conjunto de modo a obter uma base de W com os vetores $\{\boldsymbol{e}'_{r+1}, \ldots \boldsymbol{e}'_m\}$. Obviamente

$$T(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r, \boldsymbol{e}_{r+1}, \ldots, \boldsymbol{e}_n) = (\boldsymbol{e}'_1, \ldots, \boldsymbol{e}_r, 0, \ldots, 0)$$

E dessa forma nas bases $\underline{B} = \{e\}$ e $\underline{B}' = \{e\}$

$$[T]_{\underline{B},\underline{B}'} = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix}$$

3 Com base na matriz A, construímos uma transformação linear T do espaço de coordenadas $\mathbb{K}^n \to \mathbb{K}^m$ com essa matriz e depois utilizamos a afirmação 2. Nas novas bases, a matriz de T terá o formato necessário e será expressa em termos de A no formato BAC^{-1}, onde B e C são as matrizes de mudança de base. Por fim, $\operatorname{posto} A = \operatorname{posto} BAC = \operatorname{posto} T = \dim \operatorname{im} T$, o completa a prova.

■

Portanto, o conjunto dessas matrizes é um conjunto de formas canônicas pela relação de equivalência. Além disso, o posto é uma invariante completo para equivalência. Em outras palavras, duas matrizes são equivalentes se, e somente se, tiverem o mesmo posto.

Conjunto completo de invariantes

Um conjunto completo de invariantes para um problema de classificação é uma coleção de aplicações de invariantes

$$f_i : X \to Y_i$$

onde X é a coleção de objetos sendo classificados, a menos da relação de equivalência e Y_i são os conjuntos de invariantes, de modo que $x \sim x'$ se, e somente se, $f_i(x) = f_i(x')$ para todos os i. Ou seja, dois objetos são equivalentes se, e somente se, todos os invariantes forem iguais

8.2 Operadores Similares

Quando um operador linear $T \in \operatorname{Hom}(V)$ é representado em diferentes bases temos que:

$$[T]_{\underline{B}'} = P[T]_{\underline{B}}P^{-1}$$

onde P é uma matriz invertível. Isso motiva a seguinte definição.

8.1 Definição Operadores Semelhantes

a Dois operadores lineares $T, S \in \operatorname{Hom}(V)$ são ditos **similares** ou **semelhantes**, denotados por $T \sim S$, se existir um automorfismo $P \in \operatorname{Hom}(V)$ para os quais

$$S = PTP^{-1}$$

b Suas matrizes A e B são ditas **similares** ou **semelhantes**, indicadas por $A \sim B$, se existir uma matriz invertível P para a qual

$$B = PAP^{-1}$$

As classes de equivalência associadas à similaridade são chamadas classes de similaridade

Similaridade

As classes de similaridade não são todas iguais. Alguns são triviais, consistindo em um único ponto: por exemplo, se $A = 0$ ou $A = \lambda \mathrm{I}$ temos

$$PAP^{-1} = \lambda S \, \mathrm{I} \, S^{-1} = \lambda SS^{-1} = \lambda \mathrm{I} = A$$

para todas as matrizes invertíveis P.

Dessa forma a classe de similaridade $[A]$ consiste no único ponto A. Em particular,

$$[0] = \{0\}, [\mathrm{I}] = \{\mathrm{I}\} \text{ e } [-\mathrm{I}] = \{-\mathrm{I}\}.$$

Ressaltamos que em geral as classes podem ser conjuntos bem mais complicados.

são **semelhantes**

8.2 Teorema

Seja V um espaço vetorial de dimensão n. Então, dois operadores lineares T e S em V são semelhantes se, e somente se, houver uma matriz $A \in \mathcal{M}_{n,n}(\mathbb{K})$ que represente os dois operadores, mas com relação a bases ordenadas possivelmente diferentes.

Demonstração. Se T e S forem representados por $A \in \mathcal{M}_{n,n}(\mathbb{K})$, ou seja, se

$$[T]_{\underline{B}} = A = [S]_{\underline{C}}$$

para bases ordenadas $\underline{B} = \{\boldsymbol{b}_i\}$ e $\underline{C} = \{\boldsymbol{c}_i\}$,

$$[S]_{\underline{C}} = [T]_{\underline{B}} = M_{\underline{C},\underline{B}}[T]_{\underline{C}}M_{\underline{B},\underline{C}}$$

Se $P : V \to V$ for definido por $P(\boldsymbol{c}_i) = \boldsymbol{b}_i$, então

$$[P]_{\underline{C}} = M_{\underline{B},\underline{C}}$$

e então

$$[S]_{\underline{C}} = [P]_{\underline{C}}^{-1}[T]_{\underline{C}}[P]_{\underline{C}} = [P^{-1}TP]_{\underline{C}}$$

e logo S e T são semelhantes. Por outro lado, suponha que T e S sejam semelhantes, digamos

$$S = PTP^{-1}$$

onde P é um automorfismo de V. Suponha também que T seja representado pela matriz $A \in \mathcal{M}_{n,n}(\mathbb{K})$, ou seja,

$$A = [T]_{\underline{B}}$$

para alguma base ordenada $\underline{B}$. Então $[P]_{\underline{B}} = M_{\underline{C},\underline{B}}$ e assim

$$[S]_{\underline{B}} = [PTP^{-1}]_{\underline{B}} = [P]_{\underline{B}}[T]_{\underline{B}}[P]_{\underline{B}}^{-1} = M_{\underline{C},\underline{B}}[T]_{\underline{B}}M_{\underline{C},\underline{B}}^{-1}$$

Segue que

$$A = [T]_{\underline{B}} = M_{\underline{B},\underline{C}}[S]_{\underline{B}}M_{\underline{B},\underline{C}}^{-1} = [S]_{\underline{C}}$$

e então A também representa S. Por simetria, vemos que T e S são representados pelo mesmo conjunto de matrizes.Isso completa a prova. ■

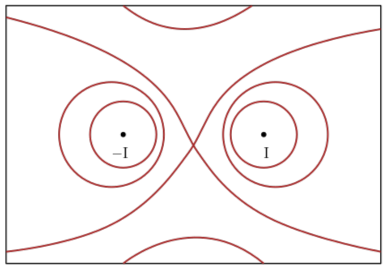


Figura 8.1 As classes de similaridade em $\mathcal{M}_{n,n}(\mathbb{K})$ particionam o espaço das matrizes em subconjuntos disjuntos $[A]$. Algumas classes são pontos, por exemplo $[-\mathrm{I}]$, $[\mathrm{I}]$ e $[0]$. Algumas classes são hiper-superfícies complicadas em $\mathcal{M}_{n,n}(\mathbb{K}) \cong \mathbb{K}^{n^2}$. Uma classe de similaridade pode ter várias componentes conexas.

8.3 Teorema

Seja V um espaço vetorial de dimensão n. Então, duas matrizes A e $B \in \mathcal{M}_{n,n}(\mathbb{K})$ são semelhantes se, e somente se, representam o mesmo operador linear $T \in \mathrm{Hom}(V)$, mas possivelmente em relação a diferentes bases ordenadas.

Demonstração. Deixaremos a demonstração do teorema como exercício.

■

Dedicaremos muito esforço no que se segue a encontrar formas canônicas por similaridade.

Exercícios

Ex. 8.1 --- Se $A \in \mathcal{M}_{n,n}(\mathbb{K})$ prove que

1. A matriz A comuta com todas as matrizes $\mathcal{M}_{n,n}(\mathbb{K})$ se, e somente se, $A = \mathrm{I}$, um múltiplo escalar se a matriz identidade para algum $\lambda \in \mathbb{K}$.

2. A matriz A comuta com todas as matrizes invertíveis

$$GL(n, K) = \{A : det(A) \neq 0\}$$

se, e somente se, A comuta com todas as matrizes $\mathcal{M}_{n,n}(\mathbb{K})$.

Dica: Lembre-se das matrizes E_{ij} são uma base para o espaço das matrizes. E se $i \neq j$então $\mathrm{I} + E_{ij}$ é invertível (verifique que $(\mathrm{I} - E_{ij})$ é a inversa), e comuta com A. Logo, E_{ij} comuta com A; deixamos você descobrir o que fazer quando $i = j$. Finalmente se A comuta com todos os vetores de base E_{ij}, ela obviamente comuta com todas as matrizes em $\mathcal{M}_{n,n}(\mathbb{K})$.

Isso mostra que uma classe de similaridade $[A]$ em $\mathcal{M}_{n,n}(\mathbb{K})$ consiste em um único ponto se, e somente se, $A = \lambda \mathrm{I}$

Ex. 8.2 --- Quando identificamos $\mathcal{M}_{2,2}(\mathbb{R}) \cong \mathbb{R}^4$ através do isomorfismo linear

$$\phi(A) = (a_{11}, a_{12}, a_{21}, a_{22}),$$

então a classe de similaridade $[A]$ da matriz

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

é a superfície bidimensional em $\mathbb{R}^4$, cuja descrição na forma paramétrica, é a imagem do um mapa polinomial $\phi : \mathbb{R}^2 \to \mathcal{M}_{2,2}(\mathbb{R})$, é

$$[A(s,t) = \left\{ \begin{bmatrix} 1 - st & s^2 \\ -t^2 & 1 + st \end{bmatrix} \mid s, t \in \mathbb{R} \text{ e } (s,t) \neq (0,0) \right\}$$

Ex. 8.3 --- Mostre que a classe de similaridade da matriz

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

não contem nenhuma matriz diagonal. Prove que em nenhuma base $\underline{B} = \boldsymbol{f}_1, \boldsymbol{f}_2$ o operador $[T_N]_{\underline{BB}}$ é diagonal.

8.3 Autovalores e Autovetores

Um caso particular extremamente importante de espaço invariante ocorre quando um subespaço invariante possui dimensão 1, isto é, quando W é T-invariante e $\dim W = 1$. Nesse caso temos que W é gerado por um vetor $\boldsymbol{w} \in W$, com $\boldsymbol{w} \neq 0$, e claramente existe λ tal que

$$T(\boldsymbol{w}) = \lambda \boldsymbol{w} \tag{8.1}$$

Um vetor $\boldsymbol{w}$ não nulo que satisfaz a Equação 8.1 é dito autovetor de T. Nesse caso como $T(\boldsymbol{v}) = \lambda \boldsymbol{v}$ então $(T - \lambda\,\mathrm{I})(\boldsymbol{v}) = 0$ e consequentemente $\boldsymbol{v} \in \ker(T - \lambda\,\mathrm{I})$ e claramente vale a reciproca.

Autovalores e Autovetores
Um autovetor de uma transformação linear é um vetor diferente de zero que muda no máximo por um fator escalar quando essa transformação linear é aplicada a ele. O autovalor correspondente é o fator pelo qual o autovetor é multiplicado.

8.1 Definição Autovalor e Autovetor

Sejam $T : V \to V$ uma transformação linear e $\lambda \in \mathbb{K}$.

a) Se $\ker(T - \lambda\,\mathrm{I}) \neq \{\mathbf{0}\}$, λ é dito **autovalor** de T.

b) Nesse caso, qualquer $\boldsymbol{v}$ não nulo em $\ker(T - \lambda\,\mathrm{I})$ é dito **autovetor** de T.

c) O subespaço $\ker(T - \lambda\,\mathrm{I})$ de V é denominado **autoespaço** de T.

d) Neste caso, dizemos que $\lambda, \boldsymbol{v}$ e $\ker(T - \lambda\,\mathrm{I})$ estão associados.

O conjunto de todos os autovalores de T com as respectivas multiplicidades é denominado **espectro** de T e denotado por $\operatorname{espec} T$.

Exemplos 2.

1. *Seja $T : V \to V$ que é a multiplicação por um escalar $T\boldsymbol{v} = \lambda\boldsymbol{v}$. Neste caso todo vetor não nulo em V é um autovetor associado ao autovalor λ.*

2. *Reflexão no plano xy. A reflexão no plano xy é a transformação R_{xy} que na base canônica é representada pela matriz:*

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

Nesse caso $\operatorname{espec} R_{xy} = \{1, -1\}$. Os autovetores associados ao 1 são $\boldsymbol{e}_1, \boldsymbol{e}_2$ e associado ao -1 temos o autovetor $\boldsymbol{e}_3$.

3 *Uma rotação em $\mathbb{R}^2$ pelo ângulo θ denotada R_θ é a transformação que na base canônica é representada pela matriz*

$$R_\theta = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

Para determinar os autovalores que queremos encontrar $\lambda, x, y \in \mathbb{R}$ tais que

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \lambda x \\ \lambda y \end{bmatrix}$$

Simplificando obtemos

$$x\cos(\theta) + \lambda(-x) - y\operatorname{sen}(\theta) = 0 \quad x\operatorname{sen}(\theta) + y\cos(\theta) + \lambda(-y) = 0$$

Elevando as duas equações ao quadrado e somando obtemos que é necessário que $\left(x^2+y^2\right)\left(a^2 - 2a\cos(\theta) + 1\right) = 0$. Para que essa equação de segundo grau tenha soluções reais é necessário que o discriminante seja positivo: $\Delta = 4\cos^2(\theta) - 4 \geq 0$. E assim $\theta = 0$ ou $\theta = \pi$. Se $\theta = 0$ então $R_\theta = \mathrm{I}$ e se Se $\theta = \pi$ então $R_\theta = -\mathrm{I}$. No primeiro caso temos um único autovalor 1 e no segundo um único autovalor -1. Em ambos os casos todos os vetores não nulos de $\mathbb{R}^2$ são autovetores.

Se $\theta \neq 0, \pi$ então R_θ não possui nem autovalores nem autovetores.

4 *Se T numa base $\underline{B} = \{\boldsymbol{v}_1, \dots \boldsymbol{v}_n\}$ possui representação matricial por uma matriz diagonal com os valores na diagonal principal distintos, isto é:*

$$[T]_{\underline{B}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \text{ com } \lambda_i \neq \lambda_j \text{ se } i \neq j.$$

Então $\lambda_1, \dots \lambda_n$ são autovalores com autovetores associados $\{\boldsymbol{v}_1, \dots \boldsymbol{v}_n\}$ respectivamente.

5 *Seja $C^\infty[a,b]$ o espaço vetorial das funções suaves em $[a,b]$. Seja $D : C^\infty[a,b] \to C^\infty[a,b]$ o operador derivação. Então a função $f(x)$ é um autovetor se satisfizer*

$$Df(x) = \lambda f(x) \to f(x) = Ce^{\lambda x}.$$

No caso particular em que $\lambda = 0$ a função f é constante.

6 *Seja $C^\infty[a,b]$ o espaço vetorial das funções suaves em $[a,b]$. Seja $I : C^\infty[a,b] \to C^\infty[a,b]$ o operador integração $I(f)(x) = \displaystyle\int_a^x f(u)du$. Então a função $f(x)$ é um autovetor se satisfizer $\displaystyle\int_a^x f(u)du = f(x)$, Como f é contínua, o Teorema Fundamental do Cálculo implica que,*

$$f(x) = \lambda f'(x)$$

cujas soluções são dadas por $f(x) = Ce^{\frac{x}{\lambda}}$ com $C \in \mathbb{R}$. Logo todo $\lambda \in \mathbb{R}$ é autovetor com autovalores associados $f(x) = Ce^{\frac{x}{\lambda}}$ para $C \in \mathbb{R}$

$\triangleleft$

8.3 Definição Autoespaço Generalizado

Seja $T : V \to V$ uma transformação linear e $\lambda \in \mathbb{K}$ seja um autovalor de T.

a O **autoespaço generalizado** de T associado a λ, denotado E_λ^∞, é o subespaço de V definido como

$$E_\lambda^\infty = \{\boldsymbol{v} \mid (T - \lambda\,\mathrm{I})^k(\boldsymbol{v}) = 0 \text{ para algum } k \in \mathbb{N}\}.$$

b Se $\boldsymbol{v}$ for um vetor diferente de zero nesse autoespaço generalizado, então $\boldsymbol{v}$ é dito **autovetor generalizado** associado ao autovalor λ.

c Para cada um desses vetores $\boldsymbol{v}$, o menor número inteiro positivo k para o qual $(T - \lambda\,\mathrm{I})^k(\boldsymbol{v}) = 0$ é dito o **índice** de $\boldsymbol{v}$.

Exemplo 4. 1 *Sejam $\lambda_1, \ldots, \lambda_n$ elementos distintos de $\mathbb{K}$ e $A \in \mathcal{M}_{n,n}(\mathbb{K})$ a matriz diagonal*

$$A = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

Para cada $i = 1, \ldots, n, \lambda_i$ é um autovalor de A com autoespaço unidimensional E_{λ_i} com base $\{\boldsymbol{e}_i\}$.

2 *Seja λ um elemento de $\mathbb{K}$ e $A \in \mathcal{M}_{n,n}(\mathbb{K})$ dada por*

$$A = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

com entradas λ na diagonal e 1 imediatamente acima da diagonal e 0 em qualquer outra entrada. Para cada $k = 1, \ldots, n, \boldsymbol{e}_k$ é um autovetor generalizado do índice k, e o autoespaço generalizado E_λ^k é k-dimensional com base $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$.

Nessa base a transformação associada a A tem a seguinte descrição:

$$e_k \mapsto \begin{cases} \lambda e_k & \text{se } k = 1 \\ \lambda e_k + e_{k-1} & \text{se } 1 < k \leq n. \end{cases}$$

◁

Para uma transformação linear T e um autovalor λ de T,

- um autovetor generalizado de índice 1 é um autovetor
- E_λ denotará o autoespaço $E_\lambda = \ker(T - \lambda\,\mathrm{I})$.

- Para um número inteiro positivo k, E_λ^k denotará o subespaço $E_\lambda^k = \ker(T - \lambda\,\mathrm{I})^k$.
- Temos que $E_\lambda^1 \subseteq E_\lambda^2 \subseteq \cdots$ e que a união desses subespaços é E_λ^∞:

$$E_\lambda^\infty = \bigcup E_\lambda^k$$

Exemplos 5. $\boxed{1}$ *Seja $\mathbb{K}$ um corpo de característica 0 e seja $V = \mathbb{K}[x]$, o espaço de todos os polinômios com coeficientes em $\mathbb{K}$. Seja $\mathrm{D} : V \to V$ o operador diferenciação, $\mathrm{D}(p(x)) = p'(x)$. Então D tem um único autovalor 0 e o autoespaço correspondente E_0 é 1-dimensional consistindo nos polinômios constantes. De maneira mais geral, E_0^k é k-dimensional, consistindo em todos os polinômios de grau no máximo $k-1$.*

$\boxed{2}$ *Seja $V = \mathbb{K}[x]$ o espaço de todos os polinômios com coeficientes em um corpo de característica 0 e seja $T : V \to V$ definida por $T(p(x)) = xp'(x)$. Então, os autovalores de T são os números inteiros não negativos e, para todo número inteiro não negativo m, o autoespaço E_m é unidimensional com base $\{x^m\}$.*

$\boxed{3}$ *Seja V o espaço das funções $C^\infty(\mathbb{R})$, e seja $\mathrm{D} : V \to V$ o operador de diferenciação, $\mathrm{D}(f(x)) = f'(x)$. Para qualquer número real λ, E_λ é unidimensional com base $f(x) = e^{\lambda x}$. Além disso, E_λ^k é k-dimensional com base $\{e^{\lambda x},\ xe^{\lambda x},\ \ldots\ ,\ x^{k-1}e^{\lambda x}\}$.*

$\boxed{4}$ *Seja $V = (\mathbb{K}^\infty)_0$ e seja $\mathrm{L} : V \to V$ o shift para a esquerda. Então L tem um único autovalor $\lambda = 0$ e o autoespaço E_0 é unidimensional, $E_0 = \{(a_1,\ a_2,\ \ldots) \in V | a_i = 0$ por $i > 1\}$. De maneira mais geral, $E_0^k = \{(a_1,\ a_2,\ \ldots) \in V \,|\, a_i = 0$ para $i > k\}$, então $\dim E_0^k = k$ para todo k e, finalmente, $V = E_0^\infty$. Por outro lado, $\mathbb{R} : V \to V$ não possui nenhum autovalor.*

$\boxed{5}$ *Seja $V = \mathbb{K}^\infty$ e seja $\mathrm{L} : V \to V$ o shift para a esquerda. Então, para qualquer $\lambda \in \mathbb{K}, E_\lambda$ é unidimensional com base $\{(1, \lambda, \lambda^2, \ldots,)\}$. Então E_λ^k é k-dimensional para todo $\lambda \in \mathbb{K}$ e todo número inteiro positivo k. Por outro lado, $\mathbb{R} : V \to V$ não possui nenhum autovalor.*

$\triangleleft$

8.6 Definição Polinômio Característico

Seja $A \in \mathcal{M}_{n,n}(\mathbb{K})$. O **polinômio característico** $c_A(x)$ de A é o polinômio

$$c_A(x) = \det(x\,\mathrm{I} - A).$$

Pelas propriedades do determinante, fica claro que $c_A(x)$ é um polinômio mônico de grau n.

8.7 Lema

Sejam A e B matrizes semelhantes. Então $c_A(x) = c_B(x)$.

Demonstração. Se $B = PAP^{-1}$, então $c_B(x) = \det(x\,\mathrm{I} - B) = \det(x\,\mathrm{I} - PAP^{-1}) = \det(P(x\,\mathrm{I} - A)P^{-1}) = \det(x\,\mathrm{I} - A) = c_A(x)$ ■

8.8 Definição Polinômio Característico

Sejam V um espaço vetorial de dimensão finita e $T : V \to V$ uma transformação linear. Seja $\underline{\mathrm{B}}$ uma base de V e seja $A = [T]_{\underline{\mathrm{B}}}$. O **polinômio característico** $c_T(x)$ é o polinômio

$$c_T(x) = c_A(x) = \det(x\,\mathrm{I} - A)\ .$$

Pelo Lema 8.3, $c_T(x)$ está bem definido, i.e., independe da escolha da base, $\underline{\mathrm{B}}$ de V.

8.9 Teorema

Seja V um espaço vetorial de dimensão finita e $T : V \to V$ seja uma transformação linear. Então λ é um autovalor de T se, e somente se, λ for uma raiz do polinômio característico $c_T(x)$, ou seja, se, e somente se, $c_T(\lambda) = 0$.

Demonstração. Seja $\underline{\mathrm{B}}$ uma base de V e faça $A = [T]_{\underline{\mathrm{B}}}$. Então, por definição, λ é um autovalor de T se, e somente se, houver um vetor diferente de zero $\boldsymbol{v}$ em $\ker(T - \lambda\,\mathrm{I})$, ou seja, se, e somente se, $(A - \lambda\,\mathrm{I})\boldsymbol{u} = 0$ para algum vetor diferente de zero $\boldsymbol{u}$ em $\mathbb{K}^n$ (onde $\boldsymbol{u} = [\boldsymbol{v}]_{\underline{\mathrm{B}}}$). Este é o caso se, e somente se, $A - \lambda\,\mathrm{I}$ for singular, que é o caso se, e somente se, $\det(A - \lambda\,\mathrm{I}) = 0$. Mas $\det(A - \lambda\,\mathrm{I}) = (-1)^n \det(\lambda - \mathrm{I}\,A)$, onde $n = \dim(V)$, então esse é o caso se, e somente se, $c_T(\lambda) = c_A(\lambda) = \det(\lambda - \mathrm{I}\,A) = 0$. ■

Definimos $c_A(x) = \det(x\,\mathrm{I} - A)$ e esta é a definição correta, pois queremos que $c_A(x)$ seja um polinômio mônico. Mas na tarefa de encontrar autoespaços, geralmente é mais conveniente trabalhar com $A - \lambda\,\mathrm{I}$ em vez de $\lambda\,\mathrm{I} - A$.

8.10 Teorema

Sejam $A, B \in \mathcal{M}_{n,n}(\mathbb{K})$ duas matrizes semelhantes, ou seja,

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

Então

1. A e B possuem os mesmos autovalores λ_i;
2. Os espaços $\ker(A-\lambda_i\,\mathrm{I})^j$ e $\ker(B-\lambda_i\,\mathrm{I})^j$ possuem a mesma dimensão para todo $j \in \mathbb{N}$ e todo autovalor λ_i.

Para o restante desta seção, assumimos que V é finito dimensional.

8.11 Definição Multiplicidade

Sejam $T : V \to V$ e λ um autovalor de T.

- A **multiplicidade algébrica** de $\lambda, \operatorname{algmult}(\lambda)$, é a multiplicidade de λ como raiz do polinômio característico $c_T(x)$.
- A **multiplicidade geométrica** de λ, $\operatorname{geomult}(\lambda)$, é a dimensão do auto-espaço associado $E_\lambda = \ker(T - \lambda\,\mathrm{I})$.

Exemplo 12. *Sejam*

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \quad e\ A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$$

então 3 é um autovalor de multiplicidade algébrica e geométrica 2 para A. Por outro lado, 3 é um autovalor de multiplicidade algébrica 2 e geométrica 1 para B.

◁

Quando não qualificado, usaremos multiplicidade para significar multiplicidade algébrica, como é o padrão na literatura.

8.13 Proposição

Sejam $T : V \to V$ e λ seja um autovalor de T. Então $1 \leq \operatorname{geomult}(\lambda) \leq \operatorname{algmult}(\lambda)$.

Demonstração. Por definição, se λ é um autovalor de T, existe um autovetor $\boldsymbol{v}$ (diferente de zero), portanto, $1 \leq \dim(E_\lambda)$.

Suponha que $\dim(E_\lambda) = d$ e seja $\underline{\mathrm{B}}_1 = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d\}$ uma base para E_λ. Estenderemos $\underline{\mathrm{B}}_1$ para uma base, $\underline{\mathrm{B}} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ de V. Então

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} \lambda\,\mathrm{I} & B \\ 0 & D \end{bmatrix} = A,$$

uma matriz de bloco com o bloco superior esquerdo de tamanho $d \times d$. Então

$$[x\,\mathrm{I} - T]_{\underline{\mathrm{B}}} = x\,\mathrm{I} - A = \begin{bmatrix} x\,\mathrm{I} - \lambda\,\mathrm{I} & -B \\ 0 & x\,\mathrm{I} - D \end{bmatrix} = \begin{bmatrix} (x-\lambda)\,\mathrm{I} & -B \\ 0 & x\,\mathrm{I} - D \end{bmatrix} \text{ logo}$$

$$c_T(x) = \det(x\,\mathrm{I} - A) = \det((x-\lambda)\,\mathrm{I})\det(x\,\mathrm{I} - D)$$

$$= (x - \lambda)^d \det(x\,\mathrm{I} - D)$$

e, portanto, $d \leq \operatorname{algmult}(\lambda)$. ■

8.14 Corolário

Sejam $T : V \to V$ e λ um autovalor de T com $\operatorname{algmult}(\lambda) = 1$. Então $\operatorname{geomult}(\lambda) = 1$.

É importante observar que a existência de autovalores e autovetores depende do corpo $\mathbb{K}$, como veremos no próximo exemplo.

Exemplo 15. *Para qualquer número racional diferente de zero t, seja A_t a matriz*

$$A_t = \begin{bmatrix} 0 & 1 \\ t & 0 \end{bmatrix}$$

dessa forma temos:

$$A_t^2 = \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} = t\,\mathrm{I}\,.$$

Seja λ um autovalor de A_t com o autovetor associado v. Então, por um lado,

$$A_t^2(\boldsymbol{v}) = A_t(A_t(\boldsymbol{v})) = A_t(\lambda\boldsymbol{v}) = \lambda A_t(\boldsymbol{v}) = \lambda^2 v,$$

mas por outro lado,

$$A_t^2(\boldsymbol{v}) = t\,\mathrm{I}(\boldsymbol{v}) = t\boldsymbol{v},$$

então $\lambda^2 = t$.

a) *Suponha $t = 1$. Então $\lambda^2 = 1, \lambda = \pm 1$, e temos o autovalor $\lambda = 1$ com o autovetor associado $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ e o autovalor $\lambda = -1$ com o autovetor associado $v = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.*

b) *Suponha $t = 2$. Se considerarmos A definido como $\mathbb{Q}$, não haverá $\lambda \in \mathbb{Q}$ com $\lambda^2 = 2$, portanto A tem sem autovalores. Se considerarmos A definido como $\mathbb{R}$, $\lambda = \pm\sqrt{2}$ e $\lambda = \sqrt{2}$ serão um autovalor com o autovetor associado $\begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$ e $\lambda = -\sqrt{2}$ são um autovalor com o autovetor associado $\begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix}$*

◁

Agora, apresentamos o polinômio minimal .

8.3.1 Ação dos Polinômios em $\mathrm{Hom}(V,V)$

Agora veremos como podemos "calcular" um polinômio numa transformação linear, ou seja, como podemos interpretar cada polinômio $p(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{K}[x]$ como uma função de $\mathrm{Hom}(V,V)$ em $\mathrm{Hom}(V,V)$ definida como:

$$\begin{aligned} p(T) : \mathrm{Hom}(V,V) &\longrightarrow \mathrm{Hom}(V,V) \\ T &\longrightarrow \textstyle\sum_{i=0}^{n} a_i T^i \end{aligned}$$

sendo $T^0 = \mathrm{I}$ o operador identidade e T^n a composta iterada n vezes de T.

De maneira análoga ao caso em que vemos como $p(x)$ como uma função de $\mathbb{K}$ em $\mathbb{K}$ estamos interessados nas "raízes":

8.16 Definição

Dizemos que um polinômio $p(x)$ **anula** uma transformação T se $p(T) = 0$.

8.17 Lema

Se V é um espaço vetorial de dimensão n e $T \in \mathrm{Hom}(V,V)$ então existe um polinômio $g(x) \in \mathbb{K}[x]$ tal que $g(T) = 0$.

Demonstração. Seja $C = \left\{\mathrm{I}, T, T^2, \dots T^{n^2}\right\}$. Como $\dim \mathrm{Hom}(V,V) = n^2$, temos que C é um conjunto linearmente dependente, logo existem $a_i \in \mathbb{K}$ tais que

$$a_0\,\mathrm{I} + a_1 T + \cdots + a_n T^{n^2} = 0$$

Logo $g(x) = \sum_{i=0}^{n} a_i x^i$ é um polinômio que anula T. ■

8.18 Corolário

Qualquer operador linear $T : V \to V$ sobre um corpo algebricamente fechado possui um autovetor.

Demonstração. Como $g(T)v = 0$. Fatorando esse polinômio temos que $(T - \lambda_1)\cdots(T-\lambda_n)v = 0$. Segue que $\ker(T - \lambda_i)$ é não trivial para algum i. ■

O **aniquilador** de T é o conjunto de todos os polinômios que anulam T:

$$\mathrm{Aniq}(T) = \{p(x) \,|\, p(T) = 0\}$$

$\mathrm{Aniq}(T)$ é um ideal de $\mathbb{K}[x]$. O lema anterior prova que esse ideal é não trivial. Como $\mathbb{K}[x]$ é um domínio de ideais principais, temos que existe um único polinômio mônico $m(x)$ tal que:

$$\mathrm{Aniq}(T) = \langle m(x) \rangle = \{a(x)m(x) \text{ com } a(x) \in \mathbb{K}[x]\}$$

8.19 Definição Polinômio Minimal

O gerador $m_T(x)$ do ideal $\mathrm{Aniq}(T)$ é dito **polinômio minimal** de T.

O polinômio minimal também pode ser caracterizado como o único polinômio mônico $m_T(x)$ de menor grau com $m_T(T) = 0$. Além disso, temos que $m_T(x)$ divide todo polinômio $p(x)$ com $p(A) = 0$.

8.4 Teoremas de Schur e Cayley-Hamilton

Nesta seção, provamos alguns resultados estruturais básicos, mas importantes, sobre uma transformação linear, obtendo informações sobre autoespaços generalizados, decomposições diretas de soma e a relação entre os polinômios característicos e minimais. Como aplicação, derivamos o famoso teorema de Cayley-Hamilton.

Embora provemos resultados muito mais fortes posteriormente, o resultado a seguir é tão fácil que faremos uma pausa para obtê-lo aqui.

8.1 Definição Triangularizável

Seja V um espaço vetorial de dimensão finita e $T : V \to V$ seja uma transformação linear.

1. T é **triangularizável superior** se houver uma base, $\underline{\mathrm{B}}$ de V na qual a matriz $A = [T]_{\underline{\mathrm{B}}}$ é triangular superior, i.e., $A(i,j) = 0$ para todos os $i > j$.
2. T é **triangularizável estritamente superior** se houver uma base, $\underline{\mathrm{B}}$ de V na qual a matriz $A = [T]_{\underline{\mathrm{B}}}$ é triangular estritamente superior, i.e., $A(i,j) = 0$ para todos os $i \geq j$.

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & \mathrm{O} & & \times & \times \\ & & & & \times \end{bmatrix} \qquad \begin{bmatrix} & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & \mathrm{O} & & & \times \\ & & & & \end{bmatrix} \tag{8.2}$$

8.2 Definição

Dizemos que um mapa linear $T \in \mathrm{Hom}(V, V)$ **estabiliza a bandeira** de subespaços se $T(V_i) \subseteq V_i$ para todos os i com $0 \leq i \leq n$. Dizemos que T **estabiliza estritamente a bandeira** se $T(V_i) \subseteq V_{i-1}$ para todos os i com $0 \leq i \leq n$.

8.3 Proposição

Suponha $T \in \mathrm{Hom}(V, V)$ e $\underline{\mathrm{B}} = \{\boldsymbol{v}_1, \ldots \boldsymbol{v}_n\}$ base de V. Então são equivalentes:

1. A matriz de T na base $\underline{\mathrm{B}}$ é triangular superior; (estritamente superior)
2. $T\boldsymbol{v}_k \in \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$ para todo $k = 1, \ldots n$ $(T(\boldsymbol{v}_1) = 0$, $T\boldsymbol{v}_k \in \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1} \rangle$ para todo $k = 2, \ldots n$;)
3. $[T]_{\underline{\mathrm{B}}}$ é (estritamente) triangular superior se T (estritamente) estabiliza a bandeira dos subespaços associados a $\underline{\mathrm{B}}$.

Demonstração. Exercício ■

8.4 Teorema de Schur

Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e $T : V \to V$ uma transformação linear. Então T é triangularizável se, e somente se, o seu polinômio característico $c_T(x)$ for um produto de fatores lineares.

Em particular, se $\mathbb{K}$ for fechado algebricamente, então todo operador linear $T : V \to V$ é triangularizável.

Demonstração. **Indução sobre a imagem**

Faremos a demonstração por indução sobre $\dim(V)$. A afirmação é clara se $\dim(V) = 1$.

Para o caso geral, seja um autovetor $\boldsymbol{v} \in V$ e um autovalor λ associado, e defina $U := \mathrm{im}(T - \lambda\,\mathrm{I})$.

Então pelo Teorema do Núcleo-Imagem temos que U é um subespaço invariante de V com $\dim(U) < \dim(V)$. Pela hipótese de indução, existe uma base $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ de U tal que $T\restriction_U$ é representado por uma matriz triangular superior. Estenda esse conjunto a uma base $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ de V.

Então

$$T\boldsymbol{v}_k = (T - \lambda \mathrm{I})\boldsymbol{v}_k + \lambda \boldsymbol{v}_k \in \langle(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)\rangle \text{ para todo } k,$$

Logo a matriz de T com relação a essa base é triangular superior. ■

Demonstração. Indução e quociente

Se $[T]_{\underline{\mathrm{B}}} = A$ é uma matriz triangular superior com entradas diagonais $d_1, \ldots, d_n$ e $c_T(x) = c_A(x) = \det(x\,\mathrm{I} - A) = (x - d_1)\cdots(x - d_n)$ é um produto de fatores lineares.

Provamos a recíproca por indução em $n = \dim(V)$. Seja $c_T(x) = (x - d_1)\cdots(x - d_n)$. Então d_1 é um autovalor de T; escolha um autovetor $\boldsymbol{v}_1$ e seja V_1 o subespaço de V gerado por $\boldsymbol{v}_1$. Definimos $\overline{V} = V/V_1$. Então T induz: $\overline{T} : \overline{V} \to \overline{V}$ com $c_{\overline{T}}(x) = (x - d_2)\cdots(x - d_n)$. Por indução, $\overline{V}$ tem uma base $\overline{B} = \{\overline{v}_2, \ \ldots \ , \overline{v}_n\}$ com $\overline{T}_{\overline{B}}$ triangular superior. Escolhemos $\boldsymbol{v}_i \in V$ com $\pi(\boldsymbol{v}_i) = \overline{v}_i$ por $i = 2, \ldots, n$ e fazemos $\underline{\mathrm{B}} = \{\boldsymbol{v}_1, \ \boldsymbol{v}_2, \ \ldots \ , \ \boldsymbol{v}_n\}$.

Observe que $\overline{T}(\overline{\boldsymbol{v}}_1) = T\boldsymbol{v}_1 + V_1$.

Então

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} d_1 & C \\ 0 & D \end{bmatrix}$$

para uma matriz $C \in \mathcal{M}_{1,n-1}(\mathbb{K})$. Independentemente da forma da matriz C, essa matriz é triangular superior.

■

8.5 Proposição

Seja v um autovetor de T com o autovalor associado λ e seja $p(x) \in \mathbb{K}[x]$ um polinômio. Então $p(T)(\boldsymbol{v}) = p(\lambda)v$. Assim, se $p(\lambda) \neq 0$, então $p(T)(\boldsymbol{v}) \neq 0$.

Demonstração. Observamos inicialmente que podemos fatorar qualquer polinômio $p(x) \in \mathbb{K}[x]$ em termos de $x - \lambda$:

$$p(x) = a_n(x - \lambda)^n + a_{n-1}(x - \lambda)^{n-1} + \cdots + a_1(x - \lambda) + a_0.$$

Substituindo $x = \lambda$, vemos que $a_0 = p(\lambda)$.

Se $\boldsymbol{v}$ for um autovetor de T com o autovalor associado λ,

$$\begin{aligned} p(T)(\boldsymbol{v}) &= (a_n(T - \lambda \mathrm{I})^n + \cdots + a_1(T - \lambda \mathrm{I}) + p(\lambda)\,\mathrm{I})(\boldsymbol{v}) \\ &= p(\lambda)\,\mathrm{I}(\boldsymbol{v}) = p(\lambda)\boldsymbol{v} \end{aligned}$$

como todos os termos, exceto o último se anulam. ■

Exercícios

Ex. 8.4 --- Suponha que W seja um subespaço de V invariante sob uma transformação linear $T : V \to V$. Prove que T induz uma aplicação linear $\overline{T} : V/W \to V/W$ dada por $\overline{T}(\boldsymbol{v}+W) = T(\boldsymbol{v})+W$. Prove que o polinômio mínimo de $\overline{T}$ divide o polinômio mínimo de T.

8.4.1 Teorema de Cayley-Hamilton

8.6 Teorema de Cayley-Hamilton

Seja V um espaço vetorial de dimensão finita e seja $T : V \to V$ uma transformação linear. Então

$$c_T(T) = 0.$$

Para demonstrarmos o teorema de Cayley-Hamilton usaremos a seguinte identidade envolvendo a matriz A e sua matriz de cofatores. Seja $A \in \mathcal{M}_{n,n}(\mathbb{K})$ com $n \geq 2$ então:

$$A(\operatorname{cof} A)^{\mathrm{t}} = (\det A)\,\mathrm{I}$$

Demonstração. Sejam A a matriz de T e $p(x)$ o polinômio característico de A:

$$c_A(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + a_n x^n$$

seja $B(x) = (b_{ij}(x))$ a matriz adjunta de $A - x\,\mathrm{I}$, ou seja, a transposta da matriz de cofatores. Como $b_{ij}(x)$ são os cofatores da matriz $A - x\,\mathrm{I}$ eles são polinômio em x de grau menor ou igual que $n-1$. Assim

$$b_{ij}(x) = b_{ij_0} + b_{1j_1}x + \cdots + b_{nj_{n-1}}x^{n-1}$$

Considere as matrizes $B_k = (b_{i,j_k})$, para $k = 0, 1, \ldots, n-1$. Então temos que

$$B(x) = B_0 + B_1 x + \cdots + B_{n-1}x^{n-1}$$

Pela igualdade

$$(A - x\,\mathrm{I})[\operatorname{adj}(A - x\,\mathrm{I})] = [\operatorname{adj}(A - x\,\mathrm{I})](A - x\,\mathrm{I}) = \det(A - x\,\mathrm{I})1$$

temos que $(A - x\,\mathrm{I})B(x) = [\det(A - x\,\mathrm{I})]1$. Assim

$$(A - x\,\mathrm{I})[B_0 + B_1 x + \cdots + B_{n-1}x^{n-1}] = (a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + a_n x^n)1$$

Expandindo o lado esquerdo desta equação e igualando as potências de mesmo grau, temos que

$$-B_{n-1} = a_n 1,\; AB_{n-1} - B_{n-2} = a_{n-1}1, \ldots, AB_1 - B_0 = a_1 1,\; AB_0 = a_0 1.$$

Multiplicando as equações matriciais acima por $A^n, A^{n-1}, \ldots, A, \mathrm{I}$, respectivamente, temos

$$-A^n B_{n-1} = a_n A^n, \quad A^n B_{n-1} - A^{n-1} B_{n-2} = a_{n-1} A^{n-1}$$

$$, \ldots, A^2 B_1 - AB_0 = a_1 A, \quad AB_0 = a_0 1$$

Somando as equações matriciais acima temos que $c_A(A) = 0$.

■

8.7 Corolário

Sejam V um espaço de dimensão finita e $T \in \mathrm{Hom}(V, V)$. Então o polinômio minimal $m_T(x)$ divide o polinômio característico $c_T(x)$.

8.8 Teorema

Sejam V um espaço de dimensão n e $T \in \mathrm{Hom}(V, V)$. Então o polinômio característico $c_T(x)$ divide a n potência do polinômio minimal: $(m_T(x))^n$.

Demonstração. Para demonstrar esse fato, fixamos uma base de V e seja M a matriz de T nessa base. Nesse contexto, $c_T = \det(x\,\mathrm{I} - M)$.

Escreva $m_T = \sum_{k=0}^{d} a_k x^k$. Então

$$m_T(x\,\mathrm{I}) = m_T(x\,\mathrm{I}) - m_T(M) \tag{8.3}$$

$$= \sum_{k=0}^{d} a_k (x^k\,\mathrm{I} - M^k) \tag{8.4}$$

$$= \sum_{k=1}^{d} a_k (x^k\,\mathrm{I} - M^k) \tag{8.5}$$

$$= (x\,\mathrm{I} - M) \sum_{k=1}^{d} a_k \sum_{p=0}^{k-1} x^p M^{k-1-p} \tag{8.6}$$

$$= (x\,\mathrm{I} - M) B \tag{8.7}$$

Onde $B := \sum_{k=1}^{d} a_k \sum_{p=0}^{k-1} x^p M^{k-1-p}$.

Em seguida, calculamos o determinante dessas matrizes: $m_T^n = m_T^n \det(\mathrm{I}) = \det(m_T(x\,\mathrm{I})) = \det((x\,\mathrm{I} - M)B) = \det(x\,\mathrm{I} - M)\det(B) = c_T \det(B)$, e o resultado segue. ■

Como consequências do Corolário 8.4.1 e do Teorema 8.4.1 temos:

8.9 Corolário

Seja $T : V \to V$ uma aplicação linear e $c_T(x) \in \mathbb{K}[x]$ seu polinômio característico. Se

$$c_T(x) = [p_1(x)]^{s_1} \cdots [p_j(x)]^{s_j}$$

é a decomposição de $c_T(x)$ em fatores irredutíveis, com $p_i \neq p_k$ para $i \neq k$, então, o polinômio minimal de T é

$$m_T(x) = [p_1(x)]^{d_1} \cdots [p_j(x)]^{d_j},$$

em que d_i é o índice de $p_i(T)$ e $0 < d_i \leq s_i$ para $i = 1, \ldots, j$. Em outras palavras, o polinômio minimal possui todos os fatores irredutíveis do polinômio característico de T.

8.10 Definição Autoespaço Generalizado Associado ao Polinômio

Sejam V um espaço vetorial de dimensão finita e $T : V \to V$ uma transformação linear. Seja $c_T \in \mathbb{K}[t]$ o polinômio característico de T. Se

$$c_T(t) = [p_1(t)]^{s_1} \ldots [p_j(t)]^{s^j}$$

é a decomposição de c_T em fatores irredutíveis, com $p_i \neq p_k$ para $i \neq k$. Definimos, o **autoespaço generalizado associado ao polinômio** p_i como o conjunto de todos os vetores $\boldsymbol{v} \in V$ para os quais existe um inteiro positivo k tal que

$$(p_i(T))^k \boldsymbol{v} = \mathbf{0}.$$

Se V for de dimensão finita a cadeia

$$\mathbf{0} \subset \ker\left(p_i(T)\right)^1 \subset \ker\left(p_i(T)\right)^2 \subset \cdots \ker\left(p_i(T)\right)^k \cdots$$

estabiliza. Seja d_i o menor inteiro positivo com a propriedade que $\ker\left(p_i(T)\right)^{d_i} = \ker\left(p_i(T)\right)^{d_i+1}$. O inteiro positivo d_i é dito o **índice** de $p_i(T)$.

8.5 Teorema da Decomposição Primária

Se $p(x)$ e $q(x)$ são polinômios coprimos tais que $p(T)q(T) = 0$, então $V = \ker p(T) \oplus \ker q(T)$.

8.1 Proposição

Sejam $p, q \in \mathbb{K}[x]$ coprimos e $0 \neq T \in \operatorname{Hom}(V)$. Sejam N_p, N_q e N_{pq} os núcleos das operadores $p(T), q(T)$ e $p(T)q(T)$, respectivamente. Então

$$N_{pq} = N_p \oplus N_q \tag{8.8}$$

Demonstração. Pela Identidade de Bézout 1.4.2 existem polinômios $a, b \in \mathbb{K}[x]$ tais que $a(x)p(x) + b(x)q(x) = 1$, temos que

$$a(T)p(T) = +b(T)q(T)\,\mathrm{I}\,.$$

Se $\boldsymbol{v} \in N_{pq}$, então $b(T)q(T)\boldsymbol{v} \in N_p$. De fato, aplicando $p(T)$ a esse vetor, temos $p(T)b(T)q(T)\boldsymbol{v} = b(T)p(T)q(T)\boldsymbol{v} = 0$. Da mesma forma temos $a(T)p(T)\boldsymbol{v} \in N_q$, se $\boldsymbol{v} \in N_{pq}$. Como $b(T)q(T)\boldsymbol{v} + a(T)p(T)\boldsymbol{v} = \boldsymbol{v}$, mostramos que $\boldsymbol{v} = \boldsymbol{v}_p + \boldsymbol{v}_q$, com $\boldsymbol{v}_p \in N_p$ e $\boldsymbol{v}_q \in N_q$.

Para mostrar que essa decomposição é única, suponhamos que $\boldsymbol{v} = \boldsymbol{v}_p + \boldsymbol{v}_q = \boldsymbol{v}'_p + \boldsymbol{v}'_q$. Mas então $\boldsymbol{w} := \boldsymbol{v}_p - \boldsymbol{v}'_p = \boldsymbol{v}'_q - \boldsymbol{v}_q$ pertence, simultaneamente, a N_p e N_q. Aplicando $b(T)q(T) + a(T)p(T) = \mathrm{I}$ em $\boldsymbol{w}$, temos

$$b(T)q(T)\boldsymbol{w} + a(T)p(T)\boldsymbol{w} = \boldsymbol{w}.$$

Mas $b(T)q(T)\boldsymbol{w} = 0 = a(T)p(T)\boldsymbol{w}$, de modo que $\boldsymbol{w} = 0$, o que implica $\boldsymbol{v} = \boldsymbol{v}'_p$ e $\boldsymbol{v}_q = \boldsymbol{v}'_q$, mostrando a unicidade da decomposição.

■

Dessa forma temos que o polinômio característico de uma transformação linear T está no aniquilador de T e assim temos o seguinte corolário.

8.2 Corolário

Seja $0 \neq T \in \operatorname{Hom}(V, V)$. Se $p_1, p_2, \ldots, p_k \in \mathbb{K}[x]$ são coprimos, se N_{p_i} denota o núcleo de $p_i(T)$ e $N_{p_1 \ldots p_k}$ o núcleo de $p_1(T) \ldots p_k(T)$, então

$$N_{p_1 \ldots p_k} = N_{p_1} \oplus \cdots \oplus N_{p_k}.$$

Um caso particularmente interessante ocorre quando aplicamos o Coroloário anterior a uma fatoração em termos irredutíveis do polinômio minimal.

8.3 Teorema Decomposição Primária

Sejam $T : V \to V$ uma aplicação linear e $m_T(x) \in \mathbb{K}[x]$ seu polinômio minimal. Se

$$m_T(x) = [p_1(x)]^{d_1} \cdots [p_j(x)]^{d_j}$$

é uma decomposição desse polinômio em fatores irredutíveis, com $p_i \neq p_k$

para $i \neq k$, então

$$V = W_1 \oplus \cdots \oplus W_j,$$

onde $W_i = \ker(p_i(T))^{d_i}$, são subespaços T-invariantes e o polinômio minimal de $T\restriction_{W_i}$ é $p_i(x))^{d_i}$.

Demonstração. Como $m_T(T) = 0$ e como os polinômios na fatoração $m_1(x) = [p_1(x)]^{d_1}, \ldots, m_j(x) = [p_j(x)]^{d_j}$ são coprimos, podemos aplicar o corolário 8.5 e concluir que

$$V = N_{m_1 \cdots m_j} = W_1 \oplus \cdots \oplus W_j. \tag{8.9}$$

Como $W_i = \ker(p_i(T))^{d_i}$, temos que o polinômio minimal de $T\restriction_{W_i}$ é da forma $p_i(x))^{r_i}$ com $r_i \leq d_i$. Se $r_i < d_i$ então $[p_1(x)]^{d_1} \cdots [p_i(x)]^{r_i} \cdots [p_j(x)]^{d_j}$ seria um polinômio de menor grau que anula T, contradizendo a minimalidade de $m_T(x)$. ■

Para o caso no qual o polinômio característico de T se fatore como produto de termos lineares temos como corolário a seguinte decomposição.

8.4 Teorema Decomposição em Autoespaços Generalizados

Suponha que o polinômio característico de T se fatore como produto de termos lineares sobre $\mathbb{K}$ e sejam $\lambda_1, \ldots, \lambda_m$ os autovalores distintos de T. Então

$$V = \bigoplus_{j=1}^{m} E^{\infty}_{\lambda_j}(T).$$

Exemplo 5. *O operador T associado a matriz na base canônica B*

$$[T]_{\underline{B}} = \begin{bmatrix} -4 & 5 & 7 & -4 \\ -3 & 4 & 4 & -2 \\ -3 & 3 & 5 & -2 \\ -1 & 1 & 1 & 1 \end{bmatrix}$$

tem polinômio característico $c_T(x) = (x-2)^2(x-1)^2$ e logo os autovalores são $2, 1$. O autoespaço generalizado $\ker(T - 2\,\mathrm{I})^2 = \{[x, y, z, w] \mid w = -x + 2y, z = y\}$. E logo $\{[1, 0, 0, -1], [1, 1, 1, 1]\}$ é uma base de $\ker(T - 2\,\mathrm{I})^2$. Por outro lado temos dois autovetores $\{[2, 0, 2, 1], [1, 1, 0, 0]\}$ associados ao autovalor 1. E assim $\{[2, 0, 2, 1], [1, 1, 0, 0]\}$ é uma base de $\ker(T - \mathrm{I})$. Podemos assim concluir também que o polinômio minimal é $m_T(x) = (x-2)^2(x-1)$. Seja a base

$$\underline{E} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4) = ([1, 0, 0, -1], [1, 1, 1, 1], [2, 0, 2, 1], [1, 1, 0, 0])$$

É fácil ver que

$$\begin{aligned}
\boldsymbol{T}\boldsymbol{e}_1 &= \boldsymbol{e}_1 - \boldsymbol{e}_2 \\
\boldsymbol{T}\boldsymbol{e}_2 &= \boldsymbol{e}_1 + 3\boldsymbol{e}_2 \\
\boldsymbol{T}\boldsymbol{e}_3 &= \boldsymbol{e}_3 \\
\boldsymbol{T}\boldsymbol{e}_4 &= \boldsymbol{e}_4
\end{aligned}$$

E assim o operador T nessa base é

$$[T]_{\underline{E}} = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Essa é uma decomposição em blocos associada a Decomposição Primária. Essa decomposição não é única e dependendo da escolha dos vetores nas bases dos autoespaços generalizados podemos simplificar os blocos. É o que faremos no próximo capítulo com a Forma de Jordan.

$\triangleleft$

Exemplo 6. *Seja V o espaço de todas as funções infinitamente diferenciáveis e considere o operador linear*

$$E[f] \triangleq (D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)f$$

onde D é a diferenciação. Associado temos a equação diferencial

$$(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)f = 0$$

com coeficientes constantes (complexos). Resolver a equação é determinar o núcleo de $E(f)$, que denotaremos por V.

A partir da teoria das equações diferenciais, temos que V é de dimensão finita com $\dim V = n$. Considere o polinômio

$$m = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Em $\mathbb{C}$, esse polinômio fatora como

$$m = (x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2}\cdots(x-\lambda_k)^{r_k}$$

Então o polinômio minimal de E é m. Pelo Corolário 8.5, V se decompõe como soma direta dos espaços de solução V_i das equações diferenciais

$$V_i = \{f \mid (D - \lambda_i \mathrm{I})^{r_j} f = 0\}.$$

Agora, as soluções de $(D - \lambda_i \mathrm{I})^{r_i} f = 0$ podem ser determinadas usando o fato de que, por um simples argumento indutivo, temos que

$$(D - \lambda_i \mathrm{I})^{r_i} f = e^{\lambda_i t} D^{r_i}(e^{-\lambda_i t} f)\ .$$

Portanto, f é uma solução se, e somente se, $D^{r_i}(e^{-\lambda_i t} f) = 0$, que é o caso se, e somente se, $e^{-\lambda_i t} f$ é um polinômio de grau no máximo $r-1$. Uma base para o espaço de solução de $(D - \lambda_i \mathrm{I})^{r_i} f = 0$ é então $\{e^{\lambda_i t},\ te^{\lambda_i t},\ \ldots,\ t^{r-1}e^{\lambda_i t}\}$.

◁

Exercícios

Ex. 8.5 --- Seja $\{T_i : i \in I\}$ um subconjunto de $\mathrm{Hom}(V, V)$ onde V é um e.v. de dimensão finita sobre um corpo algebricamente fechado $\mathbb{K}$. Suponha que $T_iT_j = T_jT_i$ para todo $i, j \in I$. Mostre que V pode ser escrito como soma direta de autoespaços generalizados comuns a todos os $T_i, i \in I$.

Ex. 8.6 --- Considere o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ cuja matriz na base canônica é

$$\begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}.$$

Ache a decomposição primária de $\mathbb{R}^3$ e encontre bases para todo um desses subespaços T-invariantes.

8.6 Diagonalizabilidade

Antes de continuarmos com nossa análise de transformações lineares gerais, consideramos um caso particular, mas muito útil.

8.1 Definição Diagonalizável

a) Dado V um espaço vetorial de dimensão finita, seja $T : V \to V$ uma transformação linear. Então T é **diagonalizável** se existir uma base $\underline{B}$ de V onde $[T]_{\underline{B}}$ é uma matriz diagonal.

b) Uma matriz $A \in \mathcal{M}_{n,n}(\mathbb{K})$ é **diagonalizável** se $T_A : \mathbb{K}^n \to \mathbb{K}^n$ é diagonalizável.

Ou seja, a matriz A é diagonalizável se for semelhante a uma matriz diagonal.

8.2 Proposição

Seja V um espaço vetorial de dimensão finita e $T : V \to V$ seja uma transformação linear. Então T é diagonalizável se, e somente se, V possuir uma base, $\underline{B}$ consistindo de autovetores de T.

Demonstração. Seja $\underline{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ uma base e seja $D = [T]_{\underline{B}}$ uma matriz diagonal com entradas diagonais $\mu_1, \ldots, \mu_n$. Para cada i,

$$[T(\boldsymbol{v}_i)]_{\mathbb{B}} = [T]_{\mathbb{B}}[\boldsymbol{v}_i]_{\mathbb{B}} = D\boldsymbol{u}_i = \mu_i\boldsymbol{u}_i = \mu_i[\boldsymbol{v}_i]_{\mathbb{B}},$$

então $T(\boldsymbol{v}_i) = \mu_i\boldsymbol{v}_i$ e $\boldsymbol{v}_i$ é um autovetor.

Por outro lado, se $B = \{\boldsymbol{v}_1, \ \ldots \ , \ \boldsymbol{v}_n\}$ é uma base de autovetores, então $T(\boldsymbol{v}_i) = \mu_i\boldsymbol{v}_i$ para todo i, e assim

$$\begin{aligned}[T]_{\mathbb{B}} &= \left[[T(\boldsymbol{v}_1)]_{\mathbb{B}} \mid [T(\boldsymbol{v}_2)]_{\mathbb{B}} \mid \cdots \mid [T(\boldsymbol{v}_n)]_{\mathbb{B}}\right] \\ &= \left[[\mu_1\boldsymbol{v}_1]_{\mathbb{B}} \mid [\mu_2\boldsymbol{v}_2]_{\mathbb{B}} \mid \cdots \mid [][\mu_n\boldsymbol{v}_n]_{\mathbb{B}}\right] \\ &= \left[\mu_1\boldsymbol{e}_1 \mid \mu_2\boldsymbol{e}_2 \mid \cdots \mid \mu_n\boldsymbol{e}_n\right] \quad = D\end{aligned}$$

é uma matriz diagonal.

■

8.3 Teorema

Sejam V um espaço vetorial de dimensão finita e $T : V \to V$ uma transformação linear. Se $c_T(x)$ não se decompor como um produto de fatores lineares sobre $\mathbb{K}$, então T não será diagonalizável. Se $c_T(x)$ se decompor como um produto de fatores lineares (o que é sempre o caso se $\mathbb{K}$ for algebricamente fechado), então as seguintes afirmações são equivalentes:

a T é diagonalizável.

b $m_T(x)$ se fatora em um produto de termos lineares distintos.

c Se $\lambda_1, \ldots, \lambda_m$ são os autovalores distintos de T, então

$$V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}.$$

d A soma das multiplicidades geométricas dos autovalores é igual à dimensão de V.

e Para cada autovalor λ de T, $\mathrm{geomult}(\lambda) = \mathrm{algmult}(\lambda)$.

f Para cada autovalor λ de T, $E_\lambda = E_\lambda^\infty$ (ou seja, todo autovetor generalizado de T é um autovetor de T).

Demonstração.

a implica *b*

Suponha que T é diagonalizável. Isso significa que tem uma base de autovetores, cujos autovalores são $\lambda_1, \ldots, \lambda_k$, nesse caso é fácil verificar que

$$(x - \lambda_1) \cdots (x - \lambda_k)$$

é um polinômio aniquilador para T. E portanto, este é o polinômio minimal.

$\boxed{b}$ implica $\boxed{a}$

Se $m_T(x)$ se fatora em um produto de termos lineares distintos então a composta

$$V \xrightarrow{T-\lambda_k \mathrm{I}} V \xrightarrow{T-\lambda_{k-1} \mathrm{I}} \dots \xrightarrow{T-\lambda_1 \mathrm{I}} V$$

é 0. Logo

$$\begin{aligned} \dim(V) &= \dim\ker((T-\lambda_1 \mathrm{I})\circ\cdots\circ(T-\lambda_k \mathrm{I})) && (8.10)\\ &\leq \dim\ker(T-\lambda_1 \mathrm{I}) + \cdots + \dim\ker(T-\lambda_k \mathrm{I}) && (8.11)\\ &= \dim(\ker(T-\lambda_1 \mathrm{I})\oplus\cdots\oplus\ker(T-\lambda_k \mathrm{I})) && (8.12)\end{aligned}$$

onde a desigualdade vem da Proposição 3.3, e a segunda igualdade é justificada pelo fato de que a soma dos autoespaços é uma soma direta. Portanto, a soma dos autoespaços tem a mesma dimensão que V, ou seja, essa soma é V e T é diagonalizável.

$\boxed{b}$ implica $\boxed{c}$

A Equação 8.12 implica que $V = E_{\lambda_1}\oplus\cdots\oplus E_{\lambda_m}$.

$\boxed{c}$ implica $\boxed{a}$

Se $V = E_{\lambda_1}\oplus\cdots\oplus E_{\lambda_m}$, seja $\underline{\mathrm{B}}_i$ uma base para E_{λ_i} e faça $\underline{\mathrm{B}} = \underline{\mathrm{B}}_1\cup\cdots\cup\underline{\mathrm{B}}_m$. Seja T_i a restrição de T a E_{λ_i}. Então $\underline{\mathrm{B}}$ é uma base para V e

$$[T]_{\underline{\mathrm{B}}} = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{bmatrix} = A,$$

uma matriz diagonal de bloco com $A_i = [T_i]_{\underline{\mathrm{B}}_i}$. Mas, neste caso, A_i é a matriz $\lambda_i \mathrm{I}$ (um múltiplo escalar da matriz identidade).

$\boxed{c}$ se, e somente se, $\boxed{d}$

Por definição.

$\boxed{d}$ se, e somente se, $\boxed{e}$

Suponha que $c_T(x) = (x-\mu_1)\cdots(x-\mu_n)$. Os escalares $\mu_1,\ldots,\mu_n$ podem não ser todos distintos, por isso os agrupamos. Sejam $\lambda_1,\ldots,\lambda_m$ os autovalores distintos logo $c_T(x) = (x-\lambda_1)^{r_1}\ldots(x-\lambda_m)^{r_m}$ para números inteiros positivos $r_1,\ldots,r_m$.

Seja $n = \dim(V)$. Claramente, r_i é a multiplicidade algébrica de λ_i e $r_1+\cdots+r_m = n$. Seja f_i a multiplicidade geométrica de λ_i. Então, pela Proposição 8.3, sabemos que $1\leq f_i\leq r_i$, portanto, $f_1+\cdots+f_m = n$ se, e somente se, $f_i = r_i$ para todo i, então $\boxed{d}$ e $\boxed{e}$ são equivalentes.

$\boxed{c}$ se, e somente se, $\boxed{f}$

Pelo Teorema 8.5, $V = E^{\infty}_{\lambda_1}\oplus\cdots\oplus E^{\infty}_{\lambda_k}$, então $V = E_{\lambda_1}\oplus\cdots\oplus E_{\lambda_k}$ se, e somente se, $E_{\lambda_1} = E^{\infty}_{\lambda_1}$ para todo i, então $\boxed{c}$ e $\boxed{f}$ são equivalentes.

■

8.4 Corolário

Seja V um espaço vetorial de dimensão finita e $T: V \to V$ uma transformação linear. Suponha que $c_T(x) = (x - \lambda_1)\cdots(x-\lambda_n)$ seja um produto de fatores lineares distintos. Então T é diagonalizável.

Demonstração. Pelo Corolário 8.3 temos que $\mathrm{algmult}(\lambda_i) = 1$ implica que $\mathrm{geomult}(\lambda_i) = 1$. ■

Exercícios

Ex. 8.7 --- Seja $A \in M_n(\mathbb{K})$

1. Mostre que P_A é um polinômio mônico de grau n.
2. Prove ainda que $P_A(x) = x^n - \mathrm{tr}(A)x^{n-1} + \cdots + (-1)^n \det(A)$.

Ex. 8.8 --- Decida se o operador linear $T : \mathbb{K}^n \to \mathbb{K}^n$ é diagonalizável. Em caso positivo, calcule uma base de autovetores e a sua forma diagonal.

1. $[T]_{\underline{\mathsf{B}}}^{\underline{\mathsf{B}}} = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}$ com $\mathbb{K} = \mathbb{C}$
2. $[T]_{\underline{\mathsf{B}}}^{\underline{\mathsf{B}}} = \begin{bmatrix} -4 & -1 \\ 4 & 0 \end{bmatrix}$ com $\mathbb{K} = \mathbb{R}$
3. $[T]_{\underline{\mathsf{B}}}^{\underline{\mathsf{B}}} = \begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}$ com $\mathbb{K} = \mathbb{R}, \mathbb{C}$

Ex. 8.9 --- Prove que

1. Se $\lambda_1, \ldots, \lambda_n$ são autovalores de T então $\lambda_1^k, \ldots, \lambda_n^k$ são os autovalores de T^k.
2. Seja V um e.v. sobre um corpo $\mathbb{K}$ algebricamente fechado e seja $T : V \to V$ um operador linear. Prove que α é um autovalor de $p(T)$ se, e somente se, $\alpha = p(\lambda)$ para algum λ autovalor de T.

Ex. 8.10 --- Prove que o polinômio característico da transposta de um operador T^{t} coincide com o polinômio característico de T.

Ex. 8.11 --- Seja $T : V \to V$ um operador linear. Mostre que se $\dim \operatorname{Im}(T) = m$, então T tem no máximo $m+1$ autovalores.

Ex. 8.12 --- Sejam $T, S : V \to V$ operadores lineares. Suponha que v é autovetor de T e de S associado aos autovalores λ_1, λ_2 de T e S, respectivamente. Ache um autovetor e um autovalor de:

1. $\alpha S + \beta T$ onde $\alpha, \beta \in \mathbb{R}$
2. $S \circ T$

Ex. 8.13 ---

1. Mostre que se $B, M \in M_n(\mathbb{K})$ com M invertível, então $(M^{-1}BM)^n = (M^{-1}B^nM)$ para todo $n \in \mathbb{N}$.
2. Calcule A^n, $n \in \mathbb{N}$ onde $A = \begin{bmatrix} 2 & 4 \\ 3 & 13 \end{bmatrix}$.
3. Seja $A = \begin{bmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{bmatrix} \in M_3(\mathbb{C})$. Dado $n \in \mathbb{N}$ determine $B \in M_3(\mathbb{C})$ tal que $B^n = A$

Ex. 8.14 --- Seja $A = \begin{bmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 3 & 0 & -3 \end{bmatrix}$ calcule A^{2020}. Agora seja $B = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 2 & 0 \\ 3 & -2 & 0 \end{bmatrix}$ consegue calcular B^{2020}? **Dica:** Pela divisão euclideana de x^k por P_B temos que $x^k = P_B(x)q_k(x) + r_k(x)$ onde $r_k = 0$ ou $\operatorname{grau}(r_k) < \operatorname{grau}(P_B)$. Encontre r_k.

Ex. 8.15 --- Seja $T : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ tal que

$$[T]^{\underline{B}}_{\underline{C}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

onde $\underline{B} = \{\frac{1}{2}x^2, -\frac{1}{2}, \frac{1}{2}x\}$ e $\underline{C} = \{x^2, x, 1\}$. Mostre que T é diagonalizável.

Ex. 8.16 --- Em $\mathcal{F}(\mathbb{R}, \mathbb{R})$ considere as funções $f_1(x) = e^{2x}\operatorname{sen}(x)$, $f_2(x) = e^{2x}\cos(x)$ e $f_3(x) = e^{2x}$, o subespaço $S = \langle f_1, f_2, f_3 \rangle$ e o operador linear $D : S \to S$ definido por $D(f) = f'$. Determine:

1. A matriz de D em relação à base $\underline{B} = \{f_1, f_2, f_3\}$ de S.

2. Os autovalores de D e as funções de S que são autovetores de D.

Ex. 8.17 --- Seja $T : V \to V$ $(\dim V < \infty)$ um operador diagonalizável cujos autovalores têm multiplicidade algébrica 1.

1. Prove que qualquer operador $G : V \to V$ tal que $GT = TG$ pode ser representado como um polinômio em T.

2. Prove que a dimensão do espaço vetorial formado por tais operadores (operadores que comutam com T) é igual a dimensão de V.

Ex. 8.18 --- Sejam $S, T : V \to V$ operadores diagonalizáveis que comutam. Então eles são simultaneamente diagonalizáveis, i.e. existe uma base $\underline{B}$ de V tal que $\underline{B}$ consiste em autovetores de S e de T.

Ex. 8.19 --- Sejam V um $\mathbb{K}$-e.v. de $\dim V < \infty$ e $T : V \to V$ um operador linear. Suponha que T comuta com todo operador diagonalizável. Prove que T é um múltiplo escalar da identidade.

Ex. 8.20 --- Seja $m \in \mathbb{R}$ e A a matriz dada por

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

1. Encontre os autovalores de A.

2. Para que valores de m a matriz A é diagonalizável?

3. Determine, de acordo com os diferentes valores de m, o polinômio minimal de A.

Ex. 8.21 --- Fixemos um vetor não nulo $a \in \mathbb{R}^3$ e defina a aplicação $T : \mathbb{R}^3 \to \mathbb{R}^3$ por $T(v) = a \times v$ (produto vetorial).

1. Prove que T é uma transformação linear.

2. Determine os autovalores e autovetores de T.

Ex. 8.22 --- Considere uma matriz real simétrica A de ordem 3 com determinante igual a 6. Suponha que $u = (4, 8, -1)$ e $v = (1, 0, 4)$ sejam autovetores desta matriz associados aos autovalores 1 e 2, respectivamente. Decida se as seguintes afirmações são verdadeiras ou falsas:

1. Os autovalores de A são apenas 1 e 2.

2. O produto vetorial $u \times v$ é autovetor de A.

3. O vetor $(5, 8, 3)$ é autovetor de A.

4. A pode não ser diagonalizável.

Ex. 8.23 --- Prove que se $T : V \to V$ é um operador linear, então $\ker T$, $\mathrm{Im}(T)$ são subespaços T-invariantes. Se λ for um autovalor de T então $\mathrm{Aut}_T(\lambda)$ é um subespaço T-invariante.

Ex. 8.24 --- Prove que a soma e a intersecção de subespaços T-invariantes é T-invariante.

Ex. 8.25 --- Prove que se um operador $T \in \mathrm{Hom}(V, V)$ com $\dim V < \infty$ é um isomorfismo então T e T^{-1} possuem os mesmos subespaços invariantes. Vale em dimensão infinita?

Ex. 8.26 --- Mostre que se todo subespaço de V for T-invariante então $T = \lambda \mathrm{I}$ para algum $\lambda \in \mathbb{K}$.

Ex. 8.27 --- Mostre que $W \subseteq V$ é um subespaço invariante para $T \in \mathrm{Hom}(V, V)$ se e só se $W^\perp$ é T^t-invariante.

Ex. 8.28 --- Sejam $T, G : V \to V$ operadores que comutam. Mostre que $\ker T$, $\mathrm{Im}(T)$ e $\mathrm{Aut}_T(\lambda)$ são G-invariantes.

Ex. 8.29 --- Sejam V um $\mathbb{K}$-e.v. de $\dim V = n$, $T : V \to V$ um operador linear, $W \subseteq V$ um subespaço T-invariante. Mostre que $c_{T|_W} \mid c_T$ e $m_{T|_W} \mid m_T$. **Dica**: lembre que se $A \in M_n(\mathbb{K})$ é diagonal por blocos

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

onde $B \in M_k(\mathbb{K})$ e $D \in M_{n-k}(\mathbb{K})$, então $A^n = \begin{bmatrix} B^n & \widetilde{C} \\ 0 & D^n \end{bmatrix}$.

Ex. 8.30 --- Sejam $T : V \to V$ um operador linear, $W \subseteq V$ um subespaço. Então W é $p(T)$-invariante para todo polinômio $p \in \mathbb{K}[x]$ se, e somente se, W for T-invariante.

Ex. 8.31 --- Seja $\pi : V \to V$ um operador projeção não trivial (i.e., $\pi \neq 0_{\mathrm{Hom}(V,V)}$ e $\pi \neq \mathrm{I}_V$) mostre que $m_\pi = x^2 - x$.

Ex. 8.32 --- Seja V um $\mathbb{C}$-espaço vetorial. Se o polinômio característico de uma transformação linear $T : V \to V$ é $x^2 - x - 1$, então é correto afirmar que:

1. T não é necessariamente invertível.

2. T é invertível e $T^{-1} = T + I$.

3. Não existe T com tal polinômio característico.

4. T é invertível e $T^{-1} = T - I$.

5. T é invertível, mas nenhuma das fórmulas para a inversa de T nos outros itens é válida.

Ex. 8.33 --- Seja $T \in \mathrm{Hom}(V, V)$.

1. Mostre que $T : V \to V$ é um operador linear não injetor se e somente se 0 é um autovalor de T.

2. Mostre que T é invertível se, e somente se, o termo independente de seu polinômio minimal é não-nulo.

3. Nestas circunstâncias, T^{-1} é um polinômio em T, i.e., existe $p(x) \in \mathbb{K}[x]$ tal que $T^{-1} = p(T)$.

Ex. 8.34 --- Seja A uma matriz complexa tal que $A^k = I$ para algum inteiro k. Prove que A é diagonalizável.

Ex. 8.35 --- Prove que uma matriz $n \times n$ sobre $\mathbb{C}$ que satisfaz $A^3 = A$ pode ser diagonalizada.

Ex. 8.36 --- Encontre todas as possibilidades para o polinômio minimal de um operador $T : \mathbb{R}^5 \to \mathbb{R}^5$ com polinômio característico c_T dado. É possível concluir que algum deles é necessariamente diagonalizável?

1. $c_T(x) = (x - 3)^3(x - 2)^2$

2. $c_T(x) = (x - 1)(x - 2)(x - 3)(x - 4)(x - 5)$

3. $c_T(x) = (x - 1)^m$, $m \geq 1$

Ex. 8.37 --- Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ algebricamente fechado e seja $T : V \to V$ um operador. Prove que T é nilpotente se, e somente se, todos os seus autovalores forem iguais a zero.

Ex. 8.38 --- Seja V um e.v. sobre um corpo $\mathbb{K}$ algebricamente fechado e $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ os autovalores de T e $q(x) = (x - \lambda_1) \cdots (x - \lambda_r)$. Mostre que $q(T)$ é nilpotente. Qual o índice de nilpotência?

Ex. 8.39 --- Dado T, G operadores lineares num espaço n-dimensional sobre um corpo de característica zero. Assuma que $T^n = 0$, $\dim \ker T = 1$ e que $GT - TG = T$. Prove que os autovalores de G são da forma $\alpha, \alpha - 1, \alpha - 2, \ldots, \alpha - (n-1)$ para algum $\alpha \in \mathbb{K}$.

Ex. 8.40 --- Seja V um $\mathbb{K}$-e.v. de $\dim V < \infty$ e $T : V \to V$ um operador linear. Então $T = T_1 \oplus T_2$ (de forma única) onde T_1 é nilpotente e T_2 é um isomorfismo.

Ex. 8.41 --- Seja $T : \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$ dada por $T(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$. Mostre que T não se escreve como soma direta de um operador nilpotente com um isomorfismo.

Ex. 8.42 --- Seja $T : \mathbb{K}^n \to \mathbb{K}^n$ um operador com polinômio característico $c_T(x) = (x - \lambda)^n$. Mostre que o operador $T' = \lambda \operatorname{Id} - T$ é nilpotente.

Exemplo 5. *Seja V um $\mathbb{K}$-e.v. de dimensão finita e $T : V \to V$ um operador linear. Se $\dim \operatorname{im}(T) = 1$ mostre que ou T é diagonalizável ou T é nilpotente.*

$\triangleleft$

Ex. 8.43 --- Se V é um $\mathbb{C}$-e.v. de dimensão finita e $T : V \to V$ é um operador linear mostre que $T^{n+1} = T$ se, e somente se, $T^{2n} = T^n$ e T diagonalizável.

Ex. 8.44 --- Para V espaço vetorial real de dimensão finita, uma involução em V é um operador linear $\varphi : V \to V$ tal que $\varphi^2 = \mathrm{I}$. Dado um operador linear $T : V \to V$, sejam

$$\operatorname{Fix}(T) = \{x \in V : T(x) = x\} \quad \text{e} \quad A(T) = \{x \in V : T(x) = -x\}$$

chamados subespaço de pontos fixos de T e subespaço antipodal de T, respectivamente. Decida se as seguintes afirmações são verdadeiras ou falsas:

1. Se φ é uma involução, então $\det(\varphi) = 1$.

2. Se φ_1, φ_2 são involuções então $\operatorname{Fix}(\varphi_1 \circ \varphi_2) = \ker(\varphi_1 - \varphi_2)$.

3. Se λ é autovalor de uma involução, então $\lambda = \pm 1$.

4. Se φ_1, φ_2 são involuções, então $\varphi_1 \circ \varphi_2$ é também uma involução.

5. Se φ é uma involução, então $A(\varphi) = \operatorname{im}(\mathrm{I} - \varphi)$.

Capítulo

9

Forma Normal de Jordan

Nesse capítulo

▸ Operadores Nilpotentes (p. 253)

▸ Forma Normal de Jordan (p. 261)

▸ Cálculo da Forma de Jordan (p. 263)

▸ Forma de Jordan Real (p. 270)

Neste capítulo e no próximo discutimos as duas formas canônicas mais importantes para operadores lineares e matrizes quadradas associadas, a forma normal de Jordan, particularmente interessante para corpos algebricamente fechados e $\mathbb{R}$, e a forma canônica racional. Essas formas canônicas capturam profundamente a estrutura de uma transformação linear e desempenham papéis importantes em muitas áreas. Alguns desses exemplos serão apresentados no final do capítulo.

A forma normal de Jordan em geral exige que todos os autovalores da matriz $A \in \mathcal{M}_{n,n}(\mathbb{K})$ estejam no corpo $\mathbb{K}$, o que ocorre o corpo é algebricamente fechado (por exemplo, $\mathbb{C}$) ou que pode ser contornado através da complexificação quando o corpo é $\mathbb{R}$. Nas outras situações a forma canônica racional é usada (por exemplo, para $\mathbb{Q}$ ou $\mathbb{Z}_p$).

Passaremos agora a uma descrição da enunciado do Teorema da Forma Normal de Jordan.

Seja $T : V \to V$ um operador linear em um espaço vetorial de dimensão finita tal que o polinômio característico do operador se decompõe em fatores lineares sobre $\mathbb{K}$ (o que sempre ocorre se $\mathbb{K}$ é algebricamente fechado). Mostraremos que existe uma base $\underline{\mathrm{B}}$ de V, na qual $[T]_{\underline{\mathrm{B}}}$ é uma matriz na forma normal de Jordan:

$$\begin{bmatrix} J_{t_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{t_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{t_r}(\lambda_r) \end{bmatrix} \qquad \text{com} \qquad J_t(\lambda) = \begin{bmatrix} \lambda & & & \\ 1 & \lambda & & \\ & & & \\ & & 1 & \lambda \end{bmatrix}.$$

As matrizes $J_t(\lambda)$ para algum $t \in \mathbb{N}^*$ e $\lambda \in \mathbb{K}$ são denominadas blocos de Jordan.

A base no qual o operador está na forma de Jordan é denominada base de Jordan.

Como a demonstração do Teorema de Jordan é envolvente começaremos discutindo o demonstração deste teorema para operadores nilpotentes.

9.1 Operadores Nilpotentes

9.1 Definição Nilpotente

Um operador linear $T : V \to V$ é dito

1. **nilpotente** se $T^k = 0$ para algum $k \in \mathbb{N}$
2. **unipotente** se $T = \mathrm{I} + N$ com N nilpotente.

Obviamente, T é unipotente se, e somente se, $T - \mathrm{I}$ é nilpotente. Matrizes nilpotentes e unipotentes $A \in \mathcal{M}_{n,n}(\mathbb{K})$ são definidas de maneira análoga.

Exemplo 2. *Seja* $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ *em* $\mathcal{M}_{2,2}(\mathbb{K})$*. Esta é uma matriz nilpotente e, em qualquer corpo, a única raiz de seu polinômio característico* $c_A(\lambda) = \det(A - x\,\mathrm{I}) = x^2$ *é* $x = 0$*. Existe um autovetor não trivial* $\boldsymbol{e}_1 = (1, 0)$*, correspondente ao autovalor* $x = 0$*. Os múltiplos escalares de* $\boldsymbol{e}_1$ *são os únicos autovetores de* A*, e portanto, não existe uma base de autovetores e assim a matriz* A *não pode ser diagonalizada, independentemente do corpo* $\mathbb{K}$*.*

$\triangleleft$

De modo geral, os operadores nilpotentes não podem ser diagonalizados, a menos que sejam o operador nulo. Assim a nossa análise deve examinar esses operadores em detalhes e é o que faremos nessa seção.

9.3 Proposição

Seja V um espaço vetorial sobre $\mathbb{K}$. Se $T : V \to V$ for nilpotente então $c_T(x) = \det(x\,\mathrm{I} - T) = x^n$ onde $n = \dim V$ e assim 0 é o único autovalor e o autoespaço associado a 0 é $\ker T$.

Como consequência direta do Teorema de Schur temos:

9.4 Proposição

Dado T um operador nilpotente então existe uma base na qual a matriz de T é triangular estritamente superior.

Decomposição de Fitting Se $T\colon V \to V$ for um operador linear em um espaço vetorial finito dimensional, definimos $K_i = \ker(T^i)$ e $R_i = \operatorname{im}(T^i)$ para $i \in \mathbb{N}$. Esses espaços formam duas cadeias de espaços vetoriais

$$\{0\} \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_i \subseteq K_{i+1} \subseteq \cdots \tag{9.1}$$

$$V \supseteq R_1 \supseteq R_2 \supseteq \cdots \supseteq R_i \supseteq R_{i+1} \supseteq \cdots, \tag{9.2}$$

e se $\dim(V) < \infty$, pelo Teorema 2.7 cada uma dessas cadeias se estabiliza em algum momento, digamos com $K_r = K_{r+1} = \cdots$ e $R_s = R_{s+1} = \cdots$ para inteiros r e s. De fato, se r é o menor índice tal que $K_r = K_{r+1} = \cdots$, a sequência de imagens também deve se estabilizar no mesmo índice pois $\dim V = \dim K_i + \dim R_i$ em cada etapa.

Com isso em mente, definimos

- a **imagem estável** de T: $R_\infty = \bigcap_{i=1}^{\infty} R_i = R_r = R_{r+1} = \cdots$
- o **núcleo estável** de T: $K_\infty = \bigcup_{i=1}^{\infty} K_i = K_r = K_{r+1} = \cdots$

9.5 Teorema Decomposição de Fitting

Sejam V um espaço de dimensão finita e $T : V \to V$. Então V se decompõe de maneira única como soma direta $V = R \oplus K$ satisfazendo

1. os espaços R, K são T-invariantes;
2. $T\restriction_K$ é um operador linear nilpotente em K;
3. $T\restriction_R$ é um operador linear bijetivo em R.

E portanto, todo operador linear T em um espaço finito dimensional V, sobre qualquer corpo, tem uma decomposição como soma direta

$$T = (T\restriction_R) \oplus (T\restriction_K)$$

com $T\restriction_K$ nilpotente e $T\restriction_R$ é bijetivo em R.

Demonstração. Escolhemos $R = R_\infty$ e $K = K_\infty$.

Afirmamos que T é nilpotente em K_∞ e T é invertível em R_∞. De fato, é fácil ver que $T(R_\infty) = R_\infty$, o que implica que T é invertível em R_∞. Também temos que $K_\infty = \ker(T^r)$ para algum $r \geq 1$ e, portanto, T é nilpotente em K_∞.

Afirmamos que $V = K_\infty \oplus R_\infty$. É direto que $K_\infty \cap R_\infty = \{\mathbf{0}\}$. Portanto, basta mostrar que todo $\boldsymbol{v} \in V$ pode ser decomposto como $k + r$ com $k \in K_\infty$ e $r \in R_\infty$. Sem perda de generalidade, podemos assumir que $K_\infty = \ker(T^r)$ e $R_\infty = T^r V$. Como $\mathrm{im}(T^{2r}) = \mathrm{im}(T^r)$, temos $T^r(\boldsymbol{v}) = T^{2r}(\boldsymbol{w})$ para algum $\boldsymbol{w} \in V$. Logo $T^r(\boldsymbol{v} - T^r(\boldsymbol{w})) = T^r(\boldsymbol{v}) - T^{2r}(\boldsymbol{w}) = 0$. Então definimos

$$k := \boldsymbol{v} - T^r(\boldsymbol{w}) \in K_\infty \qquad r := T^r(\boldsymbol{w}) \in R_\infty$$

e temos a decomposição desejada de $\boldsymbol{v}$.

Finalmente, mostramos que a decomposição $V = K_\infty \oplus R_\infty$ é única. Suponha $V = A \oplus B$ com T nilpotente em A e invertível em B. Então $A \subseteq \ker(T^k)$ para algum número inteiro positivo k e $B \subseteq \mathrm{im}(T^k)$ para todo número inteiro positivo

k. Assim,

$$A \subseteq K_\infty \text{ e } B \subseteq R_\infty.$$

Por considerações de dimensão, devemos ter $A = K_\infty$ e $B = R_\infty$.

■

9.6 **Definição** Índice de Nilpotência

Seja $T : V \to V$ nilpotente, o **índice de nilpotência**, denotado por $\operatorname{ind} T$, é o menor expoente r tal que $T^r = 0$.

Se V é um espaço vetorial de dimensão finita então existe o menor expoente tal que $T^r = 0$ pois a cadeia

$$\{0\} \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_i \subseteq K_{i+1} \subseteq \cdots$$

se estabiliza em $K_r = V$, sendo $K_i = \ker(T^i)$.

Exemplos 7. *Sejam*

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

então é direto ver que $\operatorname{ind} A = 1, \operatorname{ind} B = 2$ *e* $\operatorname{ind} C = 1$.

◁

9.8 **Lema**

Se, para algum vetor $\boldsymbol{v} \in V$ e algum inteiro m, tivermos:

$$T^{m-1}\boldsymbol{v} \neq 0 \text{ mas } T^m\boldsymbol{v} = 0$$

então o conjunto $\{\boldsymbol{v}, T\boldsymbol{v}, \ldots, T^{m-1}\boldsymbol{v}\}$ é linearmente independente.

Demonstração. A demonstração deste lema será deixada como exercício ao leitor.

■

9.9 **Definição**

Uma transformação linear $T \in \operatorname{Hom}(V, V)$ é dita **cíclica** se houver um vetor $\boldsymbol{v} \in V$ de modo que $\{\boldsymbol{v}, T\boldsymbol{v}, \ldots, T^{n-1}\boldsymbol{v}\}$ é uma base de V. Nesse caso diremos que $\boldsymbol{v}$ é um **vetor cíclico** para T e que o espaço V é **cíclico** para T ou T-cíclico. Se $\boldsymbol{v}$ é um vetor cíclico para T, diremos que a base correspondente

$\{\boldsymbol{v}, T\boldsymbol{v}, \ldots, T^{n-1}\boldsymbol{v}\}$ é uma **base cíclica** para V.

9.10 Proposição

Se V é um espaço vetorial finito dimensional e T é um operador nilpotente com índice de nilpotência igual à dimensão de V, então T é cíclico.

Demonstração. Seja $n = \dim V$. Nesse caso como o índice de nilpotência é n existe $\boldsymbol{v}$ tal que $T^{n-1}\boldsymbol{v} \neq \boldsymbol{0}$. Logo pelo Lema o conjunto $\{\boldsymbol{v}, T\boldsymbol{v}, \ldots, T^{n-1}\boldsymbol{v}\}$ é linearmente independente e por argumento de dimensão é base de V. ■

Exemplo 11. *O operador linear associado a matriz é nilpotente de índice de nilpotência 4*

$$\begin{bmatrix} -2 & 6 & -5 & 6 \\ -2 & 4 & -1 & 2 \\ 0 & 2 & -4 & 4 \\ 1 & 0 & -3 & 2 \end{bmatrix}$$

O vetor $\boldsymbol{e}_1 = [1, 0, 0, 0] \notin \ker T^3$. *Logo o conjunto*

$$\underline{B} = (\boldsymbol{e}_1, T\boldsymbol{e}_1, T^2\boldsymbol{e}_1, T^3\boldsymbol{e}_1) = ([1, 0, 0, 0], [-2, -2, 0, 1], [-2, -2, 0, 0], [-8, -4, -4, -2])$$

é uma base de V *e nessa base*

$$[T]_{\underline{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Por outro lado se escolhermos a base na ordem reversa

$$\underline{B}_2 = (T^3\boldsymbol{e}_1, T^2\boldsymbol{e}_1, T\boldsymbol{e}_1, \boldsymbol{e}_1) = ([-8, -4, -4, -2], [-2, -2, 0, 0], [-2, -2, 0, 1], [1, 0, 0, 0])$$

temos que

$$[T]_{\underline{B}_2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

◁

Na próxima seção trataremos o caso geral no qual T é nilpotente com índice de nilpotência for menor que a dimensão de V.

Operadores nilpotentes e vetores cíclicos Começaremos caracterizando as bases de Jordan definidas no inicio do capítulo.

9.12 Proposição

Seja V um espaço de dimensão finita e $T \in \mathrm{Hom}(V, V)$ nilpotente então são equivalentes:

1. $(\boldsymbol{u}_1^1, \dots, \boldsymbol{u}_{k_1}^1, \boldsymbol{u}_1^2, \dots, \boldsymbol{u}_{k_2}^2, \dots, \boldsymbol{u}_1^m, \dots \boldsymbol{u}_{k_m}^m)$ é uma base de Jordan para T;
2. $(\boldsymbol{u}_1^1, \dots, \boldsymbol{u}_{k_1}^1, \boldsymbol{u}_1^2, \dots, \boldsymbol{u}_{k_2}^2, \dots, \boldsymbol{u}_1^m, \dots \boldsymbol{u}_{k_m}^m)$ é uma base tal que $T\boldsymbol{u}_j^l = \boldsymbol{u}_{j+1}^l$, onde interpretamos $\boldsymbol{u}_j^l = 0$ se $j > k_l$.
3. Se $\boldsymbol{u}_j = \boldsymbol{u}_1^j$ então

$$\{\boldsymbol{u}_1, T\boldsymbol{u}_1, \dots, T^{k_1}\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, T^{k_2}\boldsymbol{u}_2, \dots, \boldsymbol{u}_m, \dots, T^{k_m}\boldsymbol{u}_m\}$$

é base de V.

Demonstração. A demonstração será deixada como exercício. ■

Uma representação diagramática da base de Jordan é

$$\begin{array}{cccc}
\boldsymbol{u}_1 & & & \\
\downarrow & & & \\
T\boldsymbol{u}_1 & \boldsymbol{u}_2 & & \\
\downarrow & \downarrow & & \\
T^2\boldsymbol{u}_1 & T\boldsymbol{u}_2 & & \boldsymbol{u}_m \\
\downarrow & \downarrow & & \downarrow \\
\vdots & \vdots & & \vdots \\
\downarrow & \downarrow & & \downarrow \\
T^{k_1}\boldsymbol{u}_1 & T^{k_2}\boldsymbol{u}_2 & \dots & T^{k_m}\boldsymbol{u}_m \\
\downarrow & \downarrow & & \downarrow \\
\mathbf{0} & \mathbf{0} & \dots & \mathbf{0}
\end{array}$$

9.13 Teorema Decomposição Cíclica para Operadores Nilpotentes

Todo operador T nilpotente agindo num espaço vetorial finito dimensional possui uma base de Jordan $\underline{\mathrm{J}}$. Nessa base a matriz de T é uma combinação de blocos da forma

$$[T]_{\underline{\mathrm{J}}} = \mathrm{diag}(J_{a_1}(0), \dots, J_{a_k}(0))$$

com $a_1 \geq \dots \geq a_k \geq 1$ e $a_1 + \dots + a_k = n$.

Se denotarmos por n_i o número de blocos de Jordan $J_i(0)$ de ordem i que aparecem em $[T]_{\underline{\mathrm{J}}}$, então

$$n_i = 2\dim\ker T^i - \dim\ker T^{i-1} - \dim\ker T^{i+1}.$$

> Além disso, tal representação de T como uma matriz diagonal de blocos com blocos Jordan é única a menos de permutação dos blocos.

Demonstração. Faremos a demonstração por indução sobre $\dim V$. Como T é nilpotente, $\dim \operatorname{im} T < \dim V$. Se $\operatorname{im} T = 0$, $T = 0$ e o resultado é trivial, por isso, podemos assumir que $\operatorname{im} T \neq 0$.

Por indução, existem $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in \operatorname{im} T$, de modo que

$$\{\boldsymbol{u}_1, T\boldsymbol{u}_1, \ldots, T^{a_1-1}\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, T\boldsymbol{u}_k, \ldots, T^{a_k-1}\boldsymbol{u}_k\}$$

é uma base de Jordan para T.

Para $1 \leq i \leq k$ escolha $\boldsymbol{v}_i \in V$ tal que $\boldsymbol{u}_i = T\boldsymbol{v}_i$.

Claramente

$$\left\langle T^{a_1-1}\boldsymbol{u}_1, \ldots, T^{a_k-1}\boldsymbol{u}_k \right\rangle \subseteq \ker T.$$

Estenderemos essa base a uma base de $\ker T$, adicionando os vetores $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_l$.

Afirmamos que os vetores

$$\{\boldsymbol{v}_1, T\boldsymbol{v}_1, \ldots, T^{a_1}\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, T\boldsymbol{v}_k, \ldots, T^{a_k}\boldsymbol{v}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_l\}$$

formam uma base para V.

A independência linear pode ser verificada facilmente. Suponha que exista uma combinação linear não trivial dos vetores dando 0

$$\sum_{i=1}^{k}\sum_{j=1}^{a_i} c_{ij}T^j\boldsymbol{v}_i + \sum_{m=1}^{l} b_m\boldsymbol{w}_m = 0$$

aplicando T teremos que os coeficientes dos vetores:

$$\sum_{i=1}^{k}\sum_{j=1}^{a_i-1} c_{ij}T^j\boldsymbol{u}_i = 0$$

e como esses vetores são linearmente independentes, temos que $c_{ij} = 0$ se $1 \leq i \leq k$ e $1 \leq j \leq a_i - 1$. Logo a combinação inicial se reduz a uma combinação dos vetores $T^{a_1}\boldsymbol{v}_1, \ldots, T^{a_k}\boldsymbol{v}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_l$. Mas esses vetores formam uma base para o núcleo de T, logo os coeficientes desses vetores também são nulos. E assim temos que são linearmente independentes.

Para mostrar que este vetores geram V, usamos um argumento dimensional. Sabemos que $\dim \ker T = k + l$ e $\dim \operatorname{im} T = a_1 + \ldots + a_k$. Por isso $\dim T = (a_1 + 1) + \ldots + (a_k + 1) + l$, que é o número de vetores acima.

Portanto, construímos uma base para V na qual $T : V \longrightarrow V$ está na forma normal de Jordan.

$$
\begin{array}{ccccccc}
\boldsymbol{v}_1 & & & & & & \\
\downarrow & & & & & & \\
T\boldsymbol{v}_1 & \boldsymbol{v}_2 & & & & & \\
\downarrow & \downarrow & & & & & \\
T\boldsymbol{v}_1 & T\boldsymbol{v}_2 & & \boldsymbol{v}_k & & & \\
\downarrow & \downarrow & & \downarrow & & & \\
\vdots & \vdots & \cdots & \vdots & & & \\
\downarrow & \downarrow & & \downarrow & & & \\
T^{a_1}v & T^{a_2}\boldsymbol{v}_2 & & T^{a_m-1}\boldsymbol{v}_k & \boldsymbol{w}_1 & \cdots & \boldsymbol{w}_l
\end{array}
$$

Tabela 9.1 Base de Jordan para V. Em vermelho a base de Jordan para $\operatorname{im} T$..

No diagrama anterior podemos ter várias colunas de mesma altura e a altura da maior coluna é o índice de nilpotência da transformação.

Para calcular a fórmula para n_i, observe a maneira como escrevemos a base $\underline{\mathrm{J}}$ como um diagrama bidimensional. Cada bloco de Jordan em $[T]_{\underline{\mathrm{J}}}$ corresponde a uma coluna de vetores nessa matriz.

$$
\begin{array}{cccccccccc}
\bullet & & & & & \in & & & & \ker T^n \\
\downarrow & & & & & & & & & \\
\bullet & \bullet & & & & \in & & & & \cdots \\
\downarrow & \downarrow & & & & & & & & \\
\vdots & \vdots & \cdots & \vdots & & & & & & \vdots \\
\downarrow & \downarrow & & \downarrow & & & & & & \\
\bullet & \bullet & \cdots & \bullet & & \in & & \ker T^2 & \cdots & \ker T^n \\
\downarrow & \downarrow & & \downarrow & & & & & & \\
\bullet & \bullet & & \bullet & \bullet & \in & \ker T & \ker T^2 & \cdots & \ker T^n
\end{array}
$$

Os vetores nas últimas k linhas correspondem a uma base de $\ker T^k$. Assim o número de vetores na k-ésima linha é $\dim \ker T^k - \dim \ker T^{k-1}$. E na $k+1$-ésima linha é $\dim \ker T^{k+1} - \dim \ker T^k$. A diferença desses dois números é exatamente o número de colunas de altura exatamente k.

$$
\begin{aligned}
n_i &= \dim \ker T^k - \dim \ker T^{k-1} - (\dim \ker T^{k+1} - \dim \ker T^k) && (9.3)\\
&= 2\dim \ker T^k - \dim \ker T^{k-1} - \dim \ker T^{k+1} && (9.4)
\end{aligned}
$$

Observe que os números n_i não dependem da base específica. Portanto, a representação do bloco Jordan de T é única a menos de uma permutação dos blocos.

■

Exemplo 14. *O operador linear associado a matriz é nilpotente com índice de nilpotência* 2

$$
\begin{bmatrix}
-4 & 6 & 4 & -4 \\
0 & 2 & -2 & 0 \\
0 & 2 & -2 & 0 \\
4 & -1 & -9 & 4
\end{bmatrix}
$$

Escalonando a matriz temos a seguinte forma escalonada:

$$\begin{bmatrix} 1 & 0 & -\frac{5}{2} & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A base da imagem são as colunas da matriz original que correspondem às colunas pivôs da matriz reduzida por linha e assim os vetores $\{\boldsymbol{e}_1, \boldsymbol{e}_2\} = \{[1,0,0,0],[0,1,0,0]\}$ *geram a imagem de* T *e assim não estão em* $\ker T$. *Do fato de que esses vetores são a primeira e a terceira coluna da matriz temos que esses vetores são as imagens de* $[1,0,0,0]$ *e* $[0,1,0,0]$ *respectivamente. Logo o conjunto*

$$\underline{C} = (\boldsymbol{e}_1, T\boldsymbol{e}_1, \boldsymbol{e}_2, T\boldsymbol{e}_2) = ([1,0,0,0],[-4,0,0,4],[0,1,0,0],[6,2,2,-1])$$

é uma base de Jordan para V. *O diagrama representando esse operador é:*

$$\begin{tikzcd} \bullet \arrow[d] & \bullet \arrow[d] \\ \bullet & \bullet \end{tikzcd}$$

E na base $\underline{C}$ *temos que:*

$$[T]_{\underline{C}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

As matrizes de mudança de base são

$$\mathrm{M}_{\underline{C}\to\underline{B}} = \begin{bmatrix} 1 & -4 & 0 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \\ 0 & 4 & 0 & -1 \end{bmatrix} \qquad \mathrm{M}_{\underline{B}\to\underline{C}} = \begin{bmatrix} 1 & 0 & -\frac{5}{2} & 1 \\ 0 & 0 & \frac{1}{8} & \frac{1}{4} \\ 0 & 1 & -1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

E assim

$$[T]_{\underline{C}} = \mathrm{M}_{\underline{B}\to\underline{C}}[T]_{\underline{B}}\mathrm{M}_{\underline{C}\to\underline{B}}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\frac{5}{2} & 1 \\ 0 & 0 & \frac{1}{8} & \frac{1}{4} \\ 0 & 1 & -1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} -4 & 6 & 4 & -4 \\ 0 & 2 & -2 & 0 \\ 0 & 2 & -2 & 0 \\ 4 & -1 & -9 & 4 \end{bmatrix} \begin{bmatrix} 1 & -4 & 0 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \\ 0 & 4 & 0 & -1 \end{bmatrix}.$$

◁

9.2 Forma Normal de Jordan

Antes de demonstrarmos o Teorema da Forma Normal de Jordan observamos que para $T \in \mathrm{Hom}(V,V)$ são equivalentes:

1. $(\boldsymbol{u}_1^1, \ldots, \boldsymbol{u}_{k_1}^1, \boldsymbol{u}_1^2, \ldots, \boldsymbol{u}_{k_2}^2, \ldots, \boldsymbol{u}_1^m, \ldots \boldsymbol{u}_{k_m}^m)$ é uma base de Jordan para T;
2. Existem $\lambda_j \in \overline{\mathbb{K}}$ tal que $T\boldsymbol{u}_j^l = \lambda_j \boldsymbol{u}_j^l + \boldsymbol{u}_{j+1}^l$, onde $\boldsymbol{u}_j^l = 0$ se $j > k_l$.

9.1 Teorema Forma Normal de Jordan

Seja $T: V \to V$ um operador linear em um espaço vetorial de dimensão finita tal que o polinômio característico do operador se decompõe em fatores lineares sobre $\mathbb{K}$. Então:

1. Existe uma base de Jordan para T, isto é existe uma base de V na qual a matriz de T está na forma normal de Jordan, i.e., existe uma matriz mudança de base M tal que a matriz do operador na base original A pode ser reduzida a forma de Jordan

$$M^{-1}AM = J$$

2. Se denotarmos por n_i^λ o número de blocos de Jordan $J_i(\lambda)$ de ordem i que aparecem na matriz de Jordan para T, então

$$n_i^\lambda = 2\dim\ker\left(T - \lambda\,\mathrm{I}\right)^i - \dim\ker\left(T - \lambda\,\mathrm{I}\right)^{i-1} - \dim\ker\left(T - \lambda\,\mathrm{I}\right)^{i+1}.$$

3. A matriz J é única, a menos de permutação dos blocos de Jordan.

Demonstração. Seja $T: V \to V$ um operador linear. Pelo Teorema da Decomposição em Autoespaços Generalizados (8.5) existem $E_{\lambda_j}^\infty$ autoespaços generalizados associados a T de modo que

$$V = \bigoplus_{j=1}^{m} E_{\lambda_j}^\infty(T)$$

Restrito a cada $E_{\lambda_j}^\infty$, a transformação $T - \lambda_i\,\mathrm{I}$ é nilpotente e assim pelo Teorema da Decomposição Cíclica para Operadores Nilpotentes existe uma base $\underline{\mathrm{B}}_{\lambda_i} = \left(e_k^{\lambda_i}\right)_{k=i}^{r_i}$ e blocos de Jordan $J_{t_k}(0)$ com $k = 1, \ldots, r_k$ de modo que a matriz de $T - \lambda_i\,\mathrm{I}$ nessa base é:

$$[T - \lambda_i\,\mathrm{I}]_B = \begin{bmatrix} J_{t_1^{\lambda_i}}(0) & & \\ & \ddots & \\ & & J_{t_{r_i}^{\lambda_i}}(0) \end{bmatrix} \text{ e assim } [T]_{\underline{\mathrm{B}}} = \begin{bmatrix} J_{t_1^{\lambda_i}}(\lambda_i) & & \\ & \ddots & \\ & & J_{t_{r_k}^{\lambda_i}}(\lambda_i) \end{bmatrix}$$

Seja $\bigcup_{\lambda_i \in \operatorname{espec} T} \underline{\mathrm{B}}_{\lambda_i} = \bigcup_{\lambda_i \in \operatorname{espec} T} \{e_k^{\lambda_i}\}_{k=i}^{r_{\lambda_i}}$ a base de V obtida pela concatenação das bases de $V(\lambda_i)$. Nessa base a matriz de T é:

$$\begin{bmatrix} J_{t_1^{\lambda_1}}(\lambda_1) & & & & & \\ & \ddots & & & & \\ & & J_{t_{r_1}^{\lambda_1}}(\lambda_1) & & & \\ & & & \ddots & & \\ & & & & J_{t_1^{\lambda_m}}(\lambda_m) & & \\ & & & & & \ddots & \\ & & & & & & J_{t_{r_m}^{\lambda_m}}(\lambda_m) \end{bmatrix}$$

O que termina a parte de existência demonstração do Teorema de Jordan.

Unicidade da forma de Jordan

Seja $\underline{\mathrm{B}}$ uma base de Jordan do operador T. Nessa base os elemento diagonais da matriz $[T]_{\underline{\mathrm{B}}}$ são os autovalores λ desse operador. Fixemos um autovalor λ e sejam os blocos correspondentes a esse autovalor e denote por J_λ o subespaço gerado. Isto é, J_λ é o espaço gerado pelos vetores $\boldsymbol{u}_j \in \underline{\mathrm{B}}$ tais que $T\boldsymbol{u}_j = \lambda \boldsymbol{u}_j + \boldsymbol{u}_{j+1}$.

Como $(J_r(\lambda) - \lambda\,\mathrm{I})^r = 0$, temos $J_\lambda \subset E_\lambda^\infty$.

Como $V = \oplus J_{\lambda_i}$ pela definição da base da Jordan e $V = \oplus E_{\lambda_i}^\infty(T)$ Teorema da Decomposição em Autoespaços Generalizados temos que $\dim J_{\lambda_i} = \dim E_{\lambda_i}^\infty(T)$ e consequentemente $J_{\lambda_i} = E_{\lambda_i}^\infty(T)$. Portanto, a soma das dimensões dos blocos do Jordan, correspondentes a cada autovalor λ_{i_1}, é independente da escolha da base de Jordan e, além disso, o espaço gerado pelos subconjuntos correspondentes da base J_{λ_i} são independentes da base. Portanto, é suficiente verificar a unicidade para o caso $E_{\lambda_i}^\infty(T)$ ou sem perda de generalidade para $E_0^\infty(T)$, o que já fizemos.

■

9.2 Corolário

Todo operador linear em um espaço vetorial de dimensão finita sobre um corpo $\overline{\mathbb{K}}$ algebricamente fechado possui uma forma de Jordan.

9.3 Cálculo da Forma de Jordan

Dividiremos o processo de obtenção da forma de Jordan de um operador em duas etapas:

1. Primeiramente calcularemos os diagramas de Jordan para todo autovalor obtendo assim a matriz de Jordan desse operador;
2. Finalmente calcularemos a base na qual o operador é descrito por essa matriz.

Cálculo dos Diagramas de Jordan Dada uma transformação T cuja matriz numa certa base é A. Considere o diagrama de Jordan correspondente aum autovalor λ de T. Como já observamos, a linha inferior desse diagrama consiste numa base para $\ker(T - \lambda\,\mathrm{I})$ e assim o número de vetores na linha inferior é $\dim\ker(T - \lambda\,\mathrm{I})$. De modo análogo temos que as duas linhas inferiores formam uma base para

$\ker (T - \lambda \mathrm{I})^2$ e assim sucessivamente.

$$
\begin{array}{ccccccccccc}
\bullet & & & & & \in & & & & \ker T^k \\
\downarrow & & & & & & & & & \\
\bullet & \bullet & & & & \in & & & & \cdots \\
\downarrow & \downarrow & & & & & & & & \\
\vdots & \vdots & \cdots & \vdots & & & & & & \vdots \\
\downarrow & \downarrow & & \downarrow & & & & & & \\
\bullet & \bullet & \cdots & \bullet & & \in & & \ker T^2 & \cdots & \ker T^k \\
\downarrow & \downarrow & & \downarrow & & & & & & \\
\bullet & \bullet & & \bullet & \bullet & \in & \ker T & \ker T^2 & \cdots & \ker T^k
\end{array}
$$

Logo para determinarmos esse diagrama é suficiente calcularmos:

$$\dim \ker (T - \lambda \mathrm{I}), \dim \ker (T - \lambda \mathrm{I})^2 - \dim \ker (T - \lambda \mathrm{I}), \dim \ker (T - \lambda \mathrm{I})^3 - \dim \ker (T - \lambda \mathrm{I})^2, \ldots.$$

Cada um desses números nos fornece quantos vetores aparecem em cada linha do diagrama.

Cálculo da Matriz Mudança de Base Para calcularmos a matriz mudança de base M é suficiente resolvermos a equação linear:

$$AM = MJ.$$

O sistema linear resultante pode ser indeterminado, mas nesse caso qualquer solução do sistema linear servirá.

Exemplo 1. *Determine a forma de Jordan de*

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

O polinômio característico de A é $(x-2)^3$e assim seu único autovalor é 2. Logo

$$A - 2\,\mathrm{I} = \begin{bmatrix} -1 & -\frac{1}{2} & 0 \\ 2 & 1 & 0 \\ 3 & \frac{3}{2} & 0 \end{bmatrix}.$$

Assim $\dim \ker (A - 2\ \mathrm{I}) = 2$. Como $(A - 2\,\mathrm{I})^2 = 0$. Temos que $\dim \ker (A - 2\,\mathrm{I})^2 = 3$. Logo o diagrama associado a essa matriz é:

$$
\begin{array}{cc}
\bullet & \\
\downarrow & \\
\bullet & \bullet
\end{array}
$$

E assim sua forma de Jordan é

$$\begin{bmatrix} 2 & & \\ 1 & 2 & \\ & & 2 \end{bmatrix}$$

Para calcularmos a matriz mudança de base resolvemos o sistema linear:

$$\begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 2 & 3 & 0 \\ 3 & \frac{3}{2} & 2 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

E assim

$$-a-b-\frac{d}{2}=0 \quad -b-\frac{e}{2}=0 \quad -c-\frac{f}{2}=0 \quad 2a+d-e=0 \quad 2b+e=0 \quad 2c+f=0$$

$$a+\frac{3d}{2}-h=0 \quad 3b+\frac{3e}{2}=0 \quad 3c+\frac{3f}{2}=0$$

Cuja solução é $d=-2a-2b, e=-2b, f=-2c, h=-3b$. *E tomando* $a=1, b=1, c=1, g=1$ *e* $i=1$ *temos: Resolvendo o sistema obtemos*

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

◁

Exemplo 2. *Ache a forma de Jordan para*

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

Nesse caso o polinômio característico é: $(x-3)^4$ *e assim o único autovalor é* 3.

Temos que $\dim\ker(C-3\,\mathrm{I}) = 1$ *e* $\dim\ker(B-3\,\mathrm{I})^2 = 2$, $\dim\ker(B-3\,\mathrm{I})^3 = 3$ *e logo* $\dim\ker(B-3\,\mathrm{I})^3 = 4$.

Logo o diagrama associado a essa matriz é:

$$\begin{tikzcd} \bullet \arrow[d] \\ \bullet \arrow[d] \\ \bullet \arrow[d] \\ \bullet \end{tikzcd}$$

E assim sua forma de Jordan é

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

◁

Exemplo 3. *Ache a forma de Jordan para*

$$B = \begin{bmatrix} 7 & 4 & 4 & -4 & 0 \\ -2 & 1 & -2 & 2 & 0 \\ -5 & 0 & -4 & 6 & 1 \\ -2 & 0 & -3 & 6 & 1 \\ 1 & -1 & 2 & -2 & 2 \end{bmatrix}$$

Nesse caso o polinômio característico é: $(x-3)^2(x-2)^3$ *e assim seus autovalores são:* $3, 2$.

Temos que $\dim\ker(B - 3\,\mathrm{I}) = 2$ *e* $\dim\ker(B - 3\,\mathrm{I})^2 = 2$ *e que*

$\dim\ker(B - 2\,\mathrm{I}) = 1$, $\dim\ker(B - 2\,\mathrm{I})^2 = 2$ *e* $\dim\ker(B - 2\,\mathrm{I})^3 = 3$

$\triangleleft$

Logo o diagrama de Jordan é

$$\begin{tikzcd}
\bullet \arrow[d] & & \\
\bullet \arrow[d] & & \\
\bullet & \bullet & \bullet \\
2 & 3 & 3
\end{tikzcd}$$

E assim a forma de Jordan é:

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

Exemplo 4.

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

Nesse caso $(x-2)^5$, *logo* 2 *é autovalor. Também temos que:* $\dim\ker(C - 2\,\mathrm{I}) = 3$ *e* $\dim\ker(C - 2\,\mathrm{I})^2 = 5$ *e logo o diagrama para* C *é*

$$\begin{tikzcd}
\bullet \arrow[d] & \bullet \arrow[d] & \\
\bullet & \bullet & \bullet \\
2 & 2 & 2
\end{tikzcd}$$

e consequentemente a forma de Jordan de C é:

$$\begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

◁

Exemplo 5. *Suponha $A \in \mathcal{M}_{6,6}(\mathbb{R})$ com polinômio minimal*

$$(x-1)^2(x+1)^2$$

Encontre todas as formas normais de Jordan possíveis e não semelhantes para A.

Nenhum dos blocos de Jordan pode ter tamanho maior que 3 porque corresponderia a um fator $(x \pm 1)^{3 \text{ ou mais}}$ no polinômio minimal. No entanto, para chegar ao polinômio minimal com os fatores $(x \pm 1)^2$, deve haver pelo menos um bloco Jordan de tamanho 2.

Assim necessariamente temos os blocos $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$ e $\left[\begin{smallmatrix} -1 & 0 \\ 1 & -1 \end{smallmatrix}\right]$. Temos assim a liberdade em um espaço bidimensional no qual podemos escolher um dos seguintes blocos correspondentes aos autovalores -1 e 1:

- □ $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$
- □ $\left[\begin{smallmatrix} -1 & 0 \\ 1 & -1 \end{smallmatrix}\right]$
- □ $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$
- □ $\left[\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$
- □ $\left[\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$

Portanto, as soluções a menos de permutações dos blocos são:

$$\left[\begin{array}{cc|cc|cc} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right] \quad \left[\begin{array}{cc|c|c|cc} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right] \quad \left[\begin{array}{cc|cc|cc} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

$$\left[\begin{array}{cc|cc|c|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right] \quad \left[\begin{array}{cc|c|cc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right]$$

◁

Exercícios

Seja $A \in M_6(\mathbb{R})$ tal que $A^4 - 8A^2 + 16I_6 = 0$. Quais são as possíveis formas de Jordan não semelhantes para A?

Ex. 9.1 --- Seja $T : \mathbb{R}_n[x] \to \mathbb{R}_n[x]$ dado por $T(p(x)) = p(x+1)$.

1. Determine a forma de Jordan de T.
2. Para $n = 4$, encontre uma base $\underline{\mathrm{B}}$ de $\mathbb{R}_4$ tal que $[T]_{\underline{\mathrm{B}}}^{\underline{\mathrm{B}}} = J$.

Ex. 9.2 --- Sejam as matrizes

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

veja que $c_A = pc_B = c_C = c_D = (x-2)^4$. Determine a forma de Jordan de cada uma.

Ex. 9.3 --- Dados um polinômio $f(x) = (x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\dots(x-\lambda_k)^{d_k}$ e um espaço vetorial V com $\dim V = n = d_1 + d_2 + \dots + d_k$. Quantos operadores $T : V \to V$ com $c_T = f$ existem?

Ex. 9.4 --- Ache todas as formas de Jordan possíveis para uma transformação linear com polinômio característico $c_T = (x-2)^3(x-1)^2(x-5)$. Ache o polinômio minimal correspondente a cada uma dessas formas de Jordan.

Ex. 9.5 --- Considere a transformação linear $T : \mathbb{R}^4 \to \mathbb{R}^4$ dada por

$$T(x, y, z, w) = (3z + w, 2x + 2y - 4z + 2w, z + w, 3w - z)$$

1. Encontre a base e a forma de Jordan de T.
2. Descreva todos os subespaços T-invariantes.

Ex. 9.6 --- Seja T um operador linear no $\mathbb{K}$-e.v. V de dimensão finita. Suponha que $m_T = (x-\lambda_1)^{m_1}\dots(x-\lambda_k)^{m_k}$, mostre que existe um operador diagonalizável $D \in \mathrm{Hom}(V,V)$ e um operador nilpotente $N \in \mathrm{Hom}(V,V)$ tal que

1. $T = D + N$,
2. $DN = ND$.

Prove ainda que os operadores D e N são univocamente determinados por 1 e 2 e cada um deles é um polinômio em T.

Ex. 9.7 --- A seguinte afirmação é falsa ou verdadeira: se $A \in M_n(\mathbb{C})$ é uma matriz com entradas complexas de ordem n é tal que $A^4 = I_n$ então

$$\begin{bmatrix} i & 0 \\ 1 & i \end{bmatrix}$$

pode ser um bloco de Jordan de A.

Ex. 9.8 --- Seja $T : \mathbb{R}^6 \to \mathbb{R}^6$ um operador linear com polinômios característico e minimal dados, respectivamente, por $c_T = (x-2)^4(x-1)^2$ e $m_T = (x-2)^2(x-1)^2$. Além disso suponha que $\dim \ker (T - 2\,\mathrm{I}) = 2$. Nestas condições encontre a forma de Jordan de T. Com os polinômios característico e minimal acima é possível supormos que $\dim \ker (T - 2\,\mathrm{I}) = 1$?

Ex. 9.9 ---

1. Seja $A \in M_n(\mathbb{C})$ uma matriz complexa invertível, definimos suas partes *real* e *imaginária* $R, J \in M_n(\mathbb{R})$ como sendo

$$R = \mathrm{Re}(A) \text{ e } J = \mathrm{im}(A)$$

   de forma que $A = R + iJ$. Mostre que existe $\lambda_0 \in \mathbb{R}$ tal que $R + \lambda_0 J$ é invertível.

2. Deduza que se duas matrizes reais $A, B \in M_n(\mathbb{R})$ são semelhantes em $M_n(\mathbb{C})$ então elas são semelhantes em $M_n(\mathbb{R})$.

Ex. 9.10 --- Seja $A \in M_n(\mathbb{R})$ uma matriz real satisfazendo a seguinte condição $A^3 = I_n$. Mostre que $\mathrm{tr}(A) \in \mathbb{Z}$.

Ex. 9.11 --- Sejam N_1 e N_2 matrizes de ordem 6 nilpotentes. Suponha que elas têm o mesmo polinômio minimal e o mesmo posto. Prove que elas são semelhantes. Mostre que o mesmo resultado não é verdadeiro para matrizes de ordem 7.

Ex. 9.12 --- Dê a forma de Jordan de um operador linear $T : \mathbb{R}^7 \to \mathbb{R}^7$ com polinômio característico $c_T(x) = (x-1)^2(x-2)^4(x-3)$ e tal que $\dim(\ker(T - 2\,\mathrm{I}) = 2, \dim(\ker(T - \mathrm{I})) = 1$ e $\ker(T - 2\,\mathrm{I})^3 \neq \ker(T - 2\,\mathrm{I})^2$.

Ex. 9.13 --- Seja $N \in M_n(\mathbb{R})$ uma matriz nilpotente tal que $\dim \ker(N) = k$, $0 < k < n$.

1. Mostre que $\dim \ker(N^l) \leq kl$, para todo $l \geq 1$.

2. Prove que $n \leq kr$, onde r e o é grau do polinômio minimal de N.

9.4 Forma de Jordan Real

Seja V um espaço vetorial sobre os reais e seja A a matriz do operador $T : V \to V$. O Teorema 1.4.2 diz que sobre os reais o polinômio característico $c_A(x)$ de A pode ser fatorado como produto de termos lineares e termos de grau dois que não possuem raízes em $\mathbb{R}$:

$$c_A(x) = (X - c_1)^{m_1} \cdots (X - c_k)^{m_k}((X - a_1)^2 + b_1^2)^{n^1} \cdots ((X - a_l)^2 + b_l^2)^{n_k}$$

que, por outro lado, pode ser fatorado como

$$c_A(x)(X - c_1)^{m_1} \cdots (X - c_k)^{m_k}(X - \alpha_1)^{n_1}(X - \overline{\alpha}_1)^{n_1} \cdots (X - \alpha_l)^{n_l}(X - \overline{\alpha}_l)^{n_l}$$

sobre $\mathbb{C}$ com $\alpha_j = a_j + b_j i$ e $\overline{\alpha_j} = a_j - b_j i$.

Nesta seção dado um operador T utilizaremos a forma de Jordan de T sobre os complexos para obter uma forma similar para T sobre os reais. Para obtermos uma "forma de Jordan" para operadores reais utilizaremos a técnica de complexificar e descomplexificar o operador T aliada ao fato que os termos de grau dois irredutíveis de $c_T(x)$ possuem duas raízes conjugadas.

Começamos estendendo a noção de conjugação para matrizes.

9.1 Definição

Sejam $A \in \mathcal{M}_{n,n}(\mathbb{C})$ e $\boldsymbol{v} \in \mathbb{C}^n$ um vetor qualquer. Definimos $\overline{A} \in \mathcal{M}_{n,n}(\mathbb{C})$ como a matriz obtida ao se tomar o conjugado em cada uma das entradas de A e $\overline{\boldsymbol{v}} \in \mathbb{C}^n$ como o vetor obtido ao se tomar o conjugado em cada uma das coordenadas de $\boldsymbol{v}$.

É imediato provar que $\overline{A + \lambda B} = \overline{A} + \overline{\lambda B}, \overline{AB} = \overline{AB}$ e $\overline{A\boldsymbol{v}} = \overline{A}\overline{\boldsymbol{v}}$ para quaisquer matrizes $A, B \in \mathcal{M}_{n,n}(\mathbb{C})$, $\lambda \in \mathbb{C}$ e $\boldsymbol{v} \in \mathbb{C}^n$.

As próximas proposições relacionam os autovalores, os autoespaços invariantes, etc. para um operador linear T e sua complexificação $T^{\mathbb{C}}$.

9.2 Proposição

Sejam V um espaço vetorial real de dimensão finita, $T : V \to V$ um operador linear e e $T^{\mathbb{C}}$ sua complexificação. Então:

a) os polinômios característicos de T e $T^{\mathbb{C}}$ são iguais;

b se λ é um autovalor de $T^{\mathbb{C}}$, então $\overline{\lambda}$ é também um autovalor de $T^{\mathbb{C}}$;

c as multiplicidades algébricas dos autovalores λ e $\overline{\lambda}$ são iguais;

d se W' é um subespaço fechado por conjugação, então W' possui uma base formada por vetores reais. Um vetor $\boldsymbol{v}$ é dito **real** se $\overline{\boldsymbol{v}} = \boldsymbol{v}$.

Demonstração. *a* Pela Proposição 3.6 as matrizes de T e $T^{\mathbb{C}}$ numa base de V são iguais.

b Sejam λ um autovalor de $T^{\mathbb{C}}$ e $c_T(z)$ o polinômio característico de $T^{\mathbb{C}}$. Como $c_T(z)$ também é o polinômio característico de T, os coeficientes de $c_T(z)$ são reais.

Tomando o conjugado na equação $c_T(\lambda) = 0$, obtemos $c_T(\overline{\lambda}) = 0$, o que mostra que $\overline{\lambda}$ também é uma raiz do polinômio característico de $T^{\mathbb{C}}$.

c Se λ é raiz de multiplicidade d do polinômio característico, então

$$c_T'(\lambda) = \ldots = c_T^{(d-1)}(\lambda) = 0 \text{ e } c_T^{(d)}(\lambda) \neq 0$$

tomando o conjugado em cada uma dessas equações

$$c_T'(\overline{\lambda}) = \ldots = c_T^{(d-1)}(\overline{\lambda}) = 0 \text{ e } c_T^{(d)}(\overline{\lambda}) \neq 0,$$

mostrando que $\overline{\lambda}$ também tem multiplicidade d.

d Seja $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ uma base de W', com $\boldsymbol{w}_j = \boldsymbol{u}_j + i\boldsymbol{v}_j, j = 1, \ldots, k$. Então considerando os vetores $\boldsymbol{w}_j$+$\overline{\boldsymbol{w}}_j$ e $i(\boldsymbol{w}_j$ - $\overline{\boldsymbol{w}}_j)$, obtemos que os vetores $\boldsymbol{u}_j = \boldsymbol{u}_j + i\boldsymbol{0}$ e $\boldsymbol{v}_j = \boldsymbol{v}_j + i\boldsymbol{0}$ estão em W'.

Assim, o conjunto $S = \{\boldsymbol{u}_1, \boldsymbol{v}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_k\}$ é um conjunto de vetores reais que gera W'. Uma base formada de vetores reais é obtida ao se tomar um subconjunto de S com k elementos que seja linearmente independente em $V_{\mathbb{C}}$.

■

9.3 Lema

Sejam $T : V \to V$ um operador linear e $T^{\mathbb{C}}$ sua complexificação. Se o subespaço $W' \subset V_{\mathbb{C}}$ possui uma base formada por vetores reais, então ele é a complexificação de um subespaço $W \subset V$.

Demonstração. Todo vetor de $\boldsymbol{w} \in W'$ pode ser escrito como $\boldsymbol{w} = \boldsymbol{u} + i\boldsymbol{v}$, sendo $\boldsymbol{u}$ e $\boldsymbol{v}$ vetores reais. Escrevendo $\boldsymbol{u}$ e $\boldsymbol{v}$ em termos dos vetores da base real, segue imediatamente que W' é a complexificação do espaço real W gerado pelos vetores dessa base. ■

9.4 Proposição

Seja V um espaço vetorial real e $T \in \mathrm{Hom}(V,V)$. Sejam $\alpha_1,\dots,\alpha_k$ os autovalores reais de T e $\lambda_1,\overline{\lambda_1},\dots,\lambda_m,\overline{\lambda_m}$ os autovalores não-reais de T. Então

1. V admite a decomposição

$$V = W_1 \oplus \dots W_k \oplus Z_1 \oplus \overline{Z_1} \oplus \dots Z_l \oplus \overline{Z_l}$$

onde $W_j = \ker{(T-\alpha_j\,\mathrm{I})^{q_j}}$, $j = 1,\dots,\ k$ e $Z_j = \ker{(T-\lambda_j\,\mathrm{I})^{r_j}}, j = 1,\dots,m$.

2. Se denotarmos $W = W_1 \oplus \dots \oplus W_k$ e $Z = Z_1 \oplus \overline{Z_1} \oplus \dots Z_l \oplus \overline{Z_l}$, então se $\underline{\mathrm{B}}_1,\dots,\underline{\mathrm{B}}_l$ são bases de $Z_1,\dots,Z_l$, respectivamente, então $\underline{\mathrm{B}}_Z = \underline{\mathrm{B}}_1\cup\overline{\underline{\mathrm{B}}}_1\cup\dots\cup\underline{\mathrm{B}}_l\cup\overline{\underline{\mathrm{B}}}_l$ é uma base de Z, onde $\overline{\underline{\mathrm{B}}}_j$ denota a base de Z_j formada pelos conjugados dos elementos de $\underline{\mathrm{B}}_j$, para todo $j = 1,\dots,l$.

3. Se $\{\boldsymbol{z}_1,\overline{\boldsymbol{z}_1},\dots,\boldsymbol{z}_m,\overline{\boldsymbol{z}}_m\}$ é uma base para $Z_i \oplus \overline{Z_i}$ construída no item anterior. E se $\boldsymbol{z}_i = \boldsymbol{u}_i + \boldsymbol{z}_i i$. Então $\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_m,\boldsymbol{v}_i,\dots,\boldsymbol{v}_m\}$ é uma base para $Z_i \oplus \overline{Z_i}$ constituída só de vetores reais.

Demonstração. 1 Sejam $\alpha_1,\dots,\alpha_k$ os autovalores reais de T e $\lambda_1,\overline{\lambda_1},\dots,\lambda_m,\overline{\lambda_m}$ os autovalores não-reais de T. Então o Teorema da Decomposição em Autoespaços Generalizados nos fornece a decomposição:

$$V = W_{\alpha_1} \oplus \dots W_{\alpha_k} \oplus Z_{\lambda_1} \oplus Z_{\overline{\lambda_1}} \oplus \dots Z_{\lambda_l} \oplus Z_{\overline{\lambda_l}}$$

Como $\overline{(T^{\mathbb{C}} - \lambda\,\mathrm{I})^j\boldsymbol{z}} = (T^{\mathbb{C}} - \overline{\lambda}\,\mathrm{I})^j\overline{z}$

Temos que a conjugação é um isomorfismo entre $Z_j = \ker{(T-\lambda_j\,\mathrm{I})^{r_j}}$ e $\overline{Z}_j$ e a decomposição pedida segue.

2 Se $\underline{\mathrm{B}}_i$ é uma base para Z_i. Então $\overline{\underline{\mathrm{B}}_i}$ é uma base para $\overline{Z}_i$.

3 É consequência da demonstração do item 3 da Proposição .

■

9.5 Teorema Forma de Jordan Real

Se V é um espaço vetorial real e $T : V \to V$ é um operador linear, existe uma base $\underline{\mathrm{B}}$ de V em relação à qual a matriz de T tem, ao longo da diagonal, blocos de Jordan (correspondentes aos autovalores reais) e blocos de Jordan aumentados (correspondentes aos autovalores complexos). A soma das ordens dos blocos de Jordan correspondentes a um mesmo autovalor λ é igual à multiplicidade algébrica de λ, se $\lambda \in \mathbb{R}$ e é igual ao dobro da multiplicidade algébrica de λ se $\lambda \in \mathbb{C}\setminus\mathbb{R}$.

Os blocos de Jordan aumentados são da forma

$$J_{a,b} = \begin{bmatrix} C_{a,b} & 0 & \cdots & 0 & 0 \\ \mathrm{I}_2 & C_{a,b} & \cdots & \cdots & 0 \\ \vdots & \mathrm{I}_2 & \ddots & & \\ 0 & & \cdots & C_{a,b} & \vdots \\ 0 & 0 & 0 & \mathrm{I}_2 & C_{a,b}, \end{bmatrix}$$

onde $C_{a,b} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, sendo $a + ib$ um autovalor complexo de $T^{\mathbb{C}}$ e I_2 a matriz identidade 2×2.

Demonstração. Sejam $\alpha_1, \ldots, \alpha_k$ os autovalores reais de T e $\lambda_1, \overline{\lambda_1}, \ldots, \lambda_m, \overline{\lambda_m}$ os autovalores não-reais de T. Então V admite a decomposição

$$V = W_1 \oplus \ldots W_k \oplus Z_1 \oplus \overline{Z_1} \oplus \ldots Z_l \oplus \overline{Z_l}$$

onde $W_j = \ker{(T - \alpha_j)^{q_j}},\ j = 1, \ldots,\ k$ e $Z_j = \ker{(T - \lambda_j)^{r_j}}, j = 1, \ldots, m$.

Para $\alpha_1 \in \mathbb{R}$ temos que $T - \alpha_j\,\mathrm{I}$ restrito a W_j é um operador nilpotente e a sua base de Jordan pode ser construída como na demonstração do Teorema de Jordan.

Suponhamos agora que $T^{\mathbb{C}}$ possua um autovalor $\lambda \in \mathbb{C} \backslash \mathbb{R}$.

Considere os espaços Z_λ e $Z_{\overline{\lambda}}$. Então $\{\boldsymbol{u}_1,\ \boldsymbol{v}_1,\ \boldsymbol{u}_2,\ \boldsymbol{v}_2,\ \ldots,\ \boldsymbol{u}_k,\ \boldsymbol{v}_k\}$ é uma base de $V_\lambda \oplus V_{\overline{\lambda}}$ formada por vetores reais.

Finalmente, se $\boldsymbol{w}_j = \boldsymbol{u}_j + i\boldsymbol{v}_j$, para $j \in \{1, 2,\ \ldots,\ r\}$, satisfaz $T^{\mathbb{C}}\boldsymbol{w}_j = \lambda\boldsymbol{w}_j + \boldsymbol{w}_{j+1}$, para $\lambda = a + ib \in \mathbb{C} \backslash \mathbb{R}$, então

$$T\boldsymbol{u}_j + iT\boldsymbol{v}_j = (a\boldsymbol{u}_j - b\boldsymbol{v}_j + \boldsymbol{u}_{j+1}) + i(b\boldsymbol{u}_j + a\boldsymbol{v}_j + \boldsymbol{v}_{j+1})$$

e logo

$$\begin{aligned} T\boldsymbol{u}_j &= a\boldsymbol{u}_j - b\boldsymbol{v}_j + \boldsymbol{u}_{j+1} \\ T\boldsymbol{v}_j &= b\boldsymbol{u}_j + a\boldsymbol{v}_j + \boldsymbol{v}_{j+1} \end{aligned}$$

de onde segue que, na base $\{\boldsymbol{u}_1,\ \boldsymbol{v}_1,\ \boldsymbol{u}_2,\ \boldsymbol{v}_2,\ \ldots,\ \boldsymbol{u}_k,\ \boldsymbol{v}_k\}$ de $V_\lambda \oplus V_{\overline{\lambda}}$, o operador linear $T^{\mathbb{C}}$ é representado por bloco(s) da forma descrita no enunciado do teorema. Como as matrizes de $T^{\mathbb{C}}$ e de T são iguais nessa base, a demonstração está completa.

■

Exemplo 6. *Seja*

$$A = \begin{bmatrix} 0 & -6 & 10 & -5 \\ -1 & -4 & 10 & -5 \\ -1 & -3 & 8 & -4 \\ -2 & -2 & 9 & -4 \end{bmatrix} \in \mathcal{M}_{4,4}(\mathbb{R})$$

O polinômio característico de A é $c_A = (1 + x^2)^2 = (x - i)^2(x + 1)^2$ e é igual ao minimal de A. Os autovalores sobre os complexos são $i, -i$. Consequentemente *temos dois blocos de Jordan, um associado a cada autovalor. Logo a forma normal de Jordan obtida complexificando A é.*

$$\left[\begin{array}{cc|cc} -i & 0 & 0 & 0 \\ 1 & -i & 0 & 0 \\ \hline 0 & 0 & i & \\ 0 & 0 & 1 & i \end{array}\right]$$

e pelo Teorema 9.4 temos que a forma normal de Jordan sobre $\mathbb{R}$ é:

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

◁

Exercícios

Ex. 9.14 --- Encontre uma base e a forma de Jordan real para o operador $T : \mathbb{R}^4 \to \mathbb{R}^4$ tal que

$$[T]_{\mathsf{B}}^{\mathsf{B}} = \begin{bmatrix} 2 & 0 & 0 & -2 \\ 5 & 2 & -2 & 2 \\ 3 & 1 & 0 & -3 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Ex. 9.15 --- Seja $A \in M_n(\mathbb{R})$, tal que $A^2 + I_n = 0$. Prove que $n = 2k$ e que A é semelhante à matriz

$$A = \begin{bmatrix} 0 & -I_k \\ I_k & 0 \end{bmatrix},$$

onde $I_k \in M_k(\mathbb{R})$ é a matriz identidade.

Ex. 9.16 --- Sejam $A, B \in M_n(\mathbb{R})$. Prove que se A e B são semelhantes sobre o corpo dos números complexos, então elas são também semelhantes sobre $\mathbb{R}$.

Ex. 9.17 --- Mostre que toda matriz $A \in M_n(\mathbb{C})$ é semelhante à sua transposta.

Capítulo 10

Forma Canônica Racional

Nesse capítulo

- Decomposição Cíclica (p. 275)
- Forma Canônica Racional (p. 281)
- Forma Normal de Jordan Generalizada (p. 287)
- Calculando a Forma Canônica Racional (p. 290)

Embora, em geral, o polinômio característico de transformação linear $T\colon V \to V$ possa ser decomposto como um produto de potências de polinômios irredutíveis sobre $\mathbb{K}[x]$, digamos $c_T = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$, os polinômios irredutíveis p_j não precisam ser lineares. Em outras palavras, os autovalores de T não precisam pertencer ao corpo $\mathbb{K}$. Portanto, é natural procurar uma forma canônica para o operador T nesse caso geral. Nesta seção apresentamos a forma canônica racional e na próxima mostramos como essa forma se reduz à representação da Jordan quando todos os autovalores de T pertencerem a $\mathbb{K}$.

Seja $T : V \to V$ um operador linear em um espaço vetorial de dimensão finita. Mostraremos que existe uma base $\underline{B}$ de V, na qual $[T]_{\underline{B}}$ é uma matriz na forma canônica racional:

$$\begin{bmatrix} C_{p_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & C_{p_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & C_{p_r} \end{bmatrix} \quad \text{com} \quad C_{p_1} = \begin{bmatrix} 0 & 0 & 0 & 0 & -a_{i0} \\ 1 & 0 & 0 & 0 & -a_{i1} \\ 0 & 1 & 0 & 0 & -a_{i2} \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & 1 & -a_{i\,k-1} \end{bmatrix}$$

onde C_{p_i} é uma matriz companheira para algum polinômio $p_i(x) = a_{ik}x^k + \cdots + a_{i1}x + a_{i0} \in \mathbb{K}[x]$.

10.1 Decomposição Cíclica

Vamos agora focar em um tipo particular de subespaço invariante. Suponha que V seja um espaço vetorial de dimensão finita e que $T\colon V \to V$ seja um operador linear. É fácil ver que a interseção de qualquer família de subespaços de V invariantes por T também é um subespaço de V invariante por T. Disso segue imediatamente que para todo subconjunto A de V existe um menor subespaço invariante por T que contém A, ou seja, a interseção de todos os subespaços invariante por T que contêm A. Denotaremos esse subespaço por Z_A^T. No caso em que $A = \{\boldsymbol{v}\}$, escreveremos $Z_{\boldsymbol{v}}^T$ ou simplesmente $Z_{\boldsymbol{v}}$ quando T for claramente subentendido.

10.1 Definição

Seja $T\colon V \to V$ um operador linear então o **subespaço T-cíclico** de V gerado por $\boldsymbol{v}$, e denotado por $Z_{\boldsymbol{v}}$, é o menor subespaço invariante por T que contém $\boldsymbol{v}$.

Exemplo 2. *Seja a transformação linear $T\colon \mathbb{K}^3 \to \mathbb{K}^3$ definida por*

$$T(x,\ y,\ z) = (y + 2z,\ -2z,\ x)$$

Considere o vetor $(1,0,0)$. *Temos* $T(1,0,0) = (0,0,1)$ *e* $T^2(1,0,0) = T(0,0,1 = (2,0,1)$, *e assim temos que*

$$Z_{(1,0,0)} = \{(x,\ 0,\ z); x,\ z \in \mathbb{K}\}.$$

◁

O subespaço $Z_{\boldsymbol{v}}$ pode ser caracterizado do seguinte modo.

10.3 Teorema

Sejam V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e $T\colon V \to V$ um operador linear. Então, para todo $\boldsymbol{v} \in V$,

$$Z_{\boldsymbol{v}} = \{p(T)(\boldsymbol{v}); p \in \mathbb{K}[x]\}.$$

Demonstração. É fácil ver que o conjunto $W = \{p(T)(\boldsymbol{v}) \quad p \in \mathbb{K}[x]\}$ é um subespaço de V que contém $\boldsymbol{v}$. Como T comuta com $p(T)$, esse subespaço é T-invariante.

Suponha agora que U seja um subespaço invariante por T que contenha $\boldsymbol{v}$. Então, claramente, U contém $T^k(\boldsymbol{v})$ para todos os inteiros não negativos k e, consequentemente, $p(T)(\boldsymbol{v})$ para todo polinômio $p \in \mathbb{K}[x]$. Assim, U contém W. Portanto, W é o menor subespaço invariante por T que contém $\boldsymbol{v}$ e, portanto, coincide com $Z_{\boldsymbol{v}}$. ■

Nosso objetivo imediato é descobrir uma base para o subespaço $Z_{\boldsymbol{v}}$. Para esse propósito, considere a sequência

$$\boldsymbol{v},\ T(\boldsymbol{v}),\ T^2(\boldsymbol{v}),\ \ldots,\ T^r(\boldsymbol{v}),\ \ldots$$

de elementos em $Z_{\boldsymbol{v}}$. Claramente, existe um menor número inteiro positivo k, de modo que $T^k(\boldsymbol{v})$ é uma combinação linear dos elementos que o precedem nesta lista, digamos

$$T^k(\boldsymbol{v}) = \lambda_0\boldsymbol{v} + \lambda_1 T(\boldsymbol{v}) + \cdots + \lambda_{k-1}T^{k-1}(\boldsymbol{v})$$

e $\{\boldsymbol{v}, T(\boldsymbol{v}),\ \ldots, T^{k-1}(\boldsymbol{v})\}$ é então um subconjunto linearmente independente de $Z_{\boldsymbol{v}}$.

Escrevendo $a_j = -\lambda_j$ para $i = 0, \ldots, k-1$ deduzimos que o polinômio

$$m_{\boldsymbol{v}} = a_0 + a_1 x + \cdots + a_{k-1} x^{k-1} + x^k$$

é o polinômio mônico de menor grau tal que $m_{\boldsymbol{v}}(T)(\boldsymbol{v}) = \mathbf{0}$.

10.4 Definição

Seja $\boldsymbol{v} \in V$. Então o conjunto

$$\{p(X) \in \mathbb{K}[X] : p(T)(v) = 0\}$$

é um ideal de $\mathbb{K}[X]$ e seu gerador mônico $m_{\boldsymbol{v}}$ é dito **T-aniquilador** de $\boldsymbol{v}$.

O **T-aniquilador** de $\boldsymbol{v}$ também será denotado por $\operatorname{Aniq}(v; T)$

Exemplo 5. *Seja T a transformação apresentada no Exemplo 2. Nesse exemplo tínhamos que se $\boldsymbol{u} = (0, 0, 1)$ então $T^2\boldsymbol{u} = 2\,\mathrm{I}\,\boldsymbol{u} + T\boldsymbol{u}$ e logo o T-aniquilador de $\boldsymbol{u}$ é o polinômio $m_{\boldsymbol{v}} = x^2 - x - 2 = (x-2)(x+1)$.*

◁

Com a notação acima, temos o seguinte resultado.

10.6 Teorema

Sejam V um espaço vetorial de dimensão finita e $T : V \to V$ uma transformação linear. Se $\boldsymbol{v} \in V$ tem T-aniquilador

$$m_{\boldsymbol{v}} = a_0 + a_1 x + \cdots + a_{k-1} x^{k-1} + x^k$$

Então o conjunto

$$\underline{\mathsf{B}}_{\boldsymbol{v}} = \{\boldsymbol{v},\ T(\boldsymbol{v}),\ \ldots,\ T^{k-1}(\boldsymbol{v})\}$$

é uma base de $Z_{\boldsymbol{v}}$ e, portanto, $\dim Z_{\boldsymbol{v}} = \operatorname{grau} m_{\boldsymbol{v}}$. Além disso, se $T\restriction_{Z_{\boldsymbol{v}}} : Z_{\boldsymbol{v}} \to Z_{\boldsymbol{v}}$, é a transformação linear induzida no T-subespaço invariante $Z_{\boldsymbol{v}}$, então a matriz de $T\restriction_{Z_{\boldsymbol{v}}}$ em relação à base ordenada $\underline{\mathsf{B}}_{\boldsymbol{v}}$ é

$$C_{m_{\boldsymbol{v}}} = \begin{bmatrix} 0 & 0 & 0 & 0 & -a_0 \\ 1 & 0 & 0 & 0 & -a_1 \\ 0 & 1 & 0 & 0 & -a_2 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & 1 & -a_{k-l} \end{bmatrix}$$

Finalmente, o polinômio minimal de $T\restriction_{Z_{\boldsymbol{v}}}$ é $m_{\boldsymbol{v}}$.

Demonstração. Claramente, $\underline{\mathsf{B}}_{\boldsymbol{v}}$ é linearmente independente e $T^k(\boldsymbol{v}) \in \langle \underline{\mathsf{B}}_{\boldsymbol{v}} \rangle$. Provaremos por indução que, de fato, $T^n(\boldsymbol{v}) \in \langle B \rangle$ para todo n. A afirmação é verdadeira para $n = 1, \ldots, k$ suponha que $n > k$ e $T^{n-i}(\boldsymbol{v}) \in \langle \underline{\mathrm{B}}_{\boldsymbol{v}} \rangle$. Então $T^{n-i}(\boldsymbol{v})$ é uma combinação linear de $\boldsymbol{v}, T(\boldsymbol{v}), \ldots, T^{k-i}(\boldsymbol{v})$ e assim $T^n(\boldsymbol{v})$ é combinação linear de $T(\boldsymbol{v}), T^2(\boldsymbol{v}), \ldots, T^k(\boldsymbol{v})$, e logo $T^n(\boldsymbol{v}) \in \langle B \rangle$. É imediato a partir dessa observação que $p(T)(\boldsymbol{v}) \in \langle \underline{\mathrm{B}}_{\boldsymbol{v}} \rangle$ para todo polinômio p. Portanto, $Z_{\boldsymbol{v}} \subseteq \langle \underline{\mathrm{B}}_{\boldsymbol{v}} \rangle$ de onde temos igualdade, pois a demonstração da inclusão inversa é direta. Consequentemente, $\underline{\mathrm{B}}_{\boldsymbol{v}}$ é uma base de $Z_{\boldsymbol{v}}$.

Agora como

$$\begin{aligned}
T\restriction_{Z_{\boldsymbol{v}}}(\boldsymbol{v}) &= T(\boldsymbol{v}) \\
T\restriction_{Z_{\boldsymbol{v}}}[T(\boldsymbol{v})] &= T^2(\boldsymbol{v}) \\
&\vdots \\
T\restriction_{Z_{\boldsymbol{v}}}[T^{k-2}(\boldsymbol{v})] &= T^{k-1}(\boldsymbol{v}) \\
T\restriction_{Z_{\boldsymbol{v}}}[T^{k-1}(\boldsymbol{v})] &= T^k(\boldsymbol{v}) = -a_0\boldsymbol{v} - a_1T(\boldsymbol{v}) - \cdots - a_{k-1}T^{k-1}(\boldsymbol{v})
\end{aligned}$$

é claro que a matriz de $T\restriction_{Z_{\boldsymbol{v}}}$ em relação à base $\underline{\mathrm{B}}_{\boldsymbol{v}}$ é a matriz acima C_{m_λ} acima.

Por fim, suponha que o polinômio minimal de $T\restriction_{Z_{\boldsymbol{v}}}$ seja

$$m_{T\restriction_{Z_{\boldsymbol{v}}}} = b_0 + b_1x + \cdots + b_{r-1}x^{r-1} + x^r$$

Então

$$0 = m_{T\restriction_{Z_{\boldsymbol{v}}}}(T)(\boldsymbol{v}) = b_0\boldsymbol{v} + b_1T(\boldsymbol{v}) + \cdots + b_{r-1}T^{r-1}(\boldsymbol{v}) + T^r(\boldsymbol{v})$$

e logo $T^r(\boldsymbol{v})$ é combinação linear de $\boldsymbol{v}, T(\boldsymbol{v}), \ldots, T^{r-1}(\boldsymbol{v})$ e, portanto, $k \leq R$. Mas $m_{T\restriction_{Z_{\boldsymbol{v}}}}$ é a transformação nula em $Z_{\boldsymbol{v}}$, bem como o é $m_{\boldsymbol{v}}(T\restriction_{Z_{\boldsymbol{v}}})$. Consequentemente, temos $m_{T\restriction_{Z_{\boldsymbol{v}}}} | m_{\boldsymbol{v}}$ e, portanto, $r \leq k$. Assim, $r = k$ e $m_{T\restriction_{Z_{\boldsymbol{v}}}} = m_{\boldsymbol{v}}$. ■

Exemplo 7. *Seja T a transformação apresentada no Exemplo 2. No Exemplo 2 mostramos que o T-aniquilador de $\boldsymbol{v}$ é o polinômio $m_{\boldsymbol{v}} = (x-2)(x+1)$. Nesse caso o polinômio minimal de T é $m_T = -(x-2)(x+1)(x+2)$ e logo $m_{\boldsymbol{v}}$ divide m_T*

◁

Exemplo 8. *Seja T o operador que na base canônica $\underline{B}$ é representado pela matriz*

$$[T]_{\underline{B}} = \begin{bmatrix} -2 & 3 & 6 \\ -2 & 2 & 5 \\ -1 & 1 & 3 \end{bmatrix}$$

Vamos calcular o aniquilador de $\boldsymbol{v}_1 = [1,1,0]$. Observamos que $\boldsymbol{v}_2 = T\boldsymbol{v}_1 = [1,0,0]$ $\boldsymbol{v}_3 = T^2\boldsymbol{v}_1 = T\boldsymbol{v}_2 = [-2,-2,-1]$ $\boldsymbol{v}_4 = T^3\boldsymbol{v} = [-8,-5,-3]$. Os vetores $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ são linearmente independentes e formam uma base de $Z_{\boldsymbol{v}} = V$.

Temos para $\boldsymbol{v}_4$ a seguinte relação linear $1\boldsymbol{v}_1 + 3T\boldsymbol{v}_1 - 3T^2\boldsymbol{v}_1 + T^3\boldsymbol{v}_1 = 0$. Logo $m_{\boldsymbol{v}} = -1 + 3x - 3x^2 + x^3$. Nessa base temos

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

$\triangleleft$

10.9 Definição

Se V é um espaço vetorial de dimensão finita e $T\colon V \to V$ é linear, então um subespaço W de V é dito T-cíclico se é T-invariante e possui uma base no formato $\{\boldsymbol{v}, T(\boldsymbol{v}), \ldots, T^m(\boldsymbol{v})\}$. Essa base é denominada **base cíclica** e $\boldsymbol{v}$ é denominado **vetor T-cíclico** para W.

Em particular, o Teorema 10.1 mostra que $\boldsymbol{v} \in V$ é um vetor cíclico para o subespaço $Z_{\boldsymbol{v}}$ com base cíclica $\underline{\mathrm{B}}_{\boldsymbol{v}}$. O subespaço $Z_{\boldsymbol{v}}$ é denominado **subespaço cíclico de** T gerado por $\{\boldsymbol{v}\}$. A matriz $C_{m_{\boldsymbol{v}}}$ do Teorema 10.1 é denominada **matriz companheira** do T-aniquilador $m_{\boldsymbol{v}}$.

Exemplo 10. *Um operador nilpotente N com índice de nilpotência $n = \dim V$ sempre possui um vetor cíclico, conforme a Proposição 9.1.*

$\triangleleft$

Nosso primeiro objetivo principal agora pode ser revelado. Reunindo os resultados acima, mostraremos que se $T\colon V \to V$ tiver polinômio minimal da forma p^k onde p é irredutível, então V pode ser expresso como uma soma direta dos subespaços cíclicos de T. A principal consequência disso é que T possui uma representação diagonal por bloco por matrizes companheiras.

10.1 Teorema Decomposição Cíclica

Sejam V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e $T : V \to V$ um operador linear. Então existem vetores T-cíclicos $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ tais que

$$V = \bigoplus_{j=1}^{k} Z_{\boldsymbol{v}_j},$$

Em particular, existe uma base de V na qual a matriz de T é da forma

$$\bigoplus_{i=1}^{k} C_i = \begin{bmatrix} C_{p_1} & & & \\ & C_{p_2} & & \\ & & \ddots & \\ & & & C_{p_k} \end{bmatrix}$$

e que $c_T = p_1 \cdots p_k$.

Demonstração. A prova será feita por indução sobre $n = \dim V$. O teorema claramente vale se $n = 1$, portanto, assuma que o teorema vale para todos os operadores lineares em espaços vetoriais de dimensão menor que n. Nosso objetivo é mostrar que $V = Z_{\boldsymbol{v}_1}$ para algum $\boldsymbol{v}_1 \in V$ ou que $V = Z_{\boldsymbol{v}_1} \oplus M$ para algum subespaço T-invariante M.

Seja $m \leq n$ a maior dimensão de um subespaço cíclico, ou seja, $\dim Z_{\boldsymbol{v}} \leq m$ para todos os $\boldsymbol{v} \in V$, e seja $\boldsymbol{v}_1 \in V$, de modo que $\dim Z_{\boldsymbol{v}_1} = m$.

Se $m = n$, então $Z_{\boldsymbol{v}_1} = V$ temos o que queríamos demonstrar. Caso contrário, devemos mostrar que existe um complemento T-invariante para

$$Z_{\boldsymbol{v}_1} = \langle\{\boldsymbol{v}_1,\ T(\boldsymbol{v}_1),\ \ldots,\ T^{m-1}(\boldsymbol{v}_1)\rangle\}$$

em V. Para construir esse complemento, consideramos a aplicação linear $\tau : V \to \mathbb{K}^m$ definida como

$$\tau(\boldsymbol{v}) = [f(\boldsymbol{v}), f(T(\boldsymbol{v})), \ldots, f(T^{m-1}(\boldsymbol{v}))]^t$$

onde $f\colon V \to \mathbb{K}$ é um funcional linear escolhido de modo que

$$f(\boldsymbol{v}_1) = 0, \quad f(T(\boldsymbol{v}_1)) = 0, \quad \ldots\ f(T^{m-2}(\boldsymbol{v}_1)) = 0, \quad f(T^{m-1}(\boldsymbol{v}_1)) = 1.$$

Observe que é possível escolher tal funcional pois os vetores $\boldsymbol{v}_1, T(\boldsymbol{v}_1), \ldots$, $T^{m-1}(\boldsymbol{v}_1)$ são linearmente independentes e, portanto, parte de uma base para V.

Afirmamos agora que $\tau\restriction_{Z_{\boldsymbol{v}_1}} : Z_{\boldsymbol{v}_1} \to \mathbb{K}^m$ é um isomorfismo. Para demonstrar isso, encontramos a representação matricial para $\tau\restriction_{Z_{\boldsymbol{v}_1}}$. Usando a base $\underline{\mathrm{B}} = \{\boldsymbol{v}_1, T(\boldsymbol{v}_1), \ldots, T^{m-1}(\boldsymbol{v}_1)\}$ para $Z_{\boldsymbol{v}_1}$ e a base canônica $\underline{\mathrm{C}} = \boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ para $\mathbb{K}^m$, vemos que:

$$[\tau]_{\underline{\mathrm{B}},\underline{\mathrm{C}}} = \begin{bmatrix} 0 & 0 & \ldots & 1 \\ \vdots & & \iddots & * \\ 0 & 1 & \ldots & * \\ 1 & * & * & * \end{bmatrix}$$

onde $*$ indica que não nos importamos com o valor da entrada. Como a matriz é invertível, temos que $\tau\restriction_{Z_{\boldsymbol{v}_1}} : Z_{\boldsymbol{v}_1} \to \mathbb{K}^m$ é um isomorfismo.

Em seguida, precisamos mostrar que $\ker \tau$ é T-invariante. Seja $\boldsymbol{v} \in \ker \tau$ logo por definição

$$\tau(\boldsymbol{v}) = [f(\boldsymbol{v}), f(T(\boldsymbol{v})), \ldots, f(T^{m-1}(\boldsymbol{v}))] = [0, 0, \ldots, 0]$$

e consequentemente

$$\tau(T(\boldsymbol{v})) = [f(T(\boldsymbol{v})), f(T^2(\boldsymbol{v})), \ldots, f(T^{m-1}(\boldsymbol{v})), f(T^m(\boldsymbol{v}))] = [0, 0, \ldots, f(T^m(\boldsymbol{v}))]$$

Agora, pela definição de m, temos que $T^m(\boldsymbol{v})$ é uma combinação linear de $(\boldsymbol{v}), T(\boldsymbol{v}), \ldots$, $T^{m-1}(\boldsymbol{v})$. Isso mostra que $f(T^m(\boldsymbol{v})) = 0$ e, consequentemente, $T(\boldsymbol{v}) \in \ker \tau$.

Finalmente, mostramos que $V = Z_{\boldsymbol{v}_1} \oplus \ker \tau$. Vimos que $\tau\restriction_{Z_{\boldsymbol{v}_1}}\colon Z_{\boldsymbol{v}_1} \to \mathbb{K}^m$ é um isomorfismo. Isso implica que $Z_{\boldsymbol{v}_1} \cap \ker \tau = \{\boldsymbol{0}\}$. Logo

$$\begin{aligned} \dim(V) &= \dim(\ker \tau) + \dim(\operatorname{im} \tau) && (10.1) \\ &= \dim(\ker \tau) + m && (10.2) \\ &= \dim(\ker \tau) + \dim(Z_{\boldsymbol{v}_1}) && (10.3) \\ &= \dim(\ker \tau + Z_{\boldsymbol{v}_1}). && (10.4) \end{aligned}$$

Assim, $V = Z_{\boldsymbol{v}_1} + \ker \tau = Z_{\boldsymbol{v}_1} \oplus \ker \tau$.

Logo temos a decomposição $V = \bigoplus_{j=1}^{k} Z_{\boldsymbol{v}_j}$. Como $m_T = p^{\mathrm{t}}$ temos que existe um vetor $\boldsymbol{v}_1 \neq \mathbf{0}$ em V com $p^{t-1}(T)(\boldsymbol{v}_1) \neq \mathbf{0}$. O T-aniquilador de $\boldsymbol{v}_1$ é então $m_{\boldsymbol{v}_1} = p^{\mathrm{t}}$.

■

Exercícios

Ex. 10.1 --- Se $\boldsymbol{u}_j = \boldsymbol{v}_{1j} + \cdots + \boldsymbol{v}_{kj}$ então $m_{\boldsymbol{u}_j}(x) = f_i(x)$.

10.2 Forma Canônica Racional

Pelo Teorema da Decomposição Cíclica temos que sempre existe uma decomposição de soma direta em subespaços cíclicos. No entanto, é possível que os subespaços cíclicos ainda possam ser decompostos como soma direta de subespaços cíclicos menores. Portanto, apenas conhecer para ambas as matrizes alguma decomposição em subespaços cíclicos, e conhecer os polinômios mínimos correspondentes, não é suficiente para decidir sua similaridade.

Dessa forma Uma condição adicional deve ser imposta para garantir a unicidade: na lista de polinômios mínimos associados, cada um deve dividir o próximo.

A lista de polinômios resultante é chamada de fatores invariantes da matriz, e duas matrizes são semelhantes se, e somente se, tiverem listas idênticas de fatores invariantes.

10.1 Proposição

Sejam V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e $T : V \to V$ um operador linear com polinômio minimal $m_T = p^{\mathrm{t}}$ onde p é um polinômio irredutível sobre $\mathbb{K}$. Então existem vetores T-cíclicos $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ e números inteiros positivos $n_1, \ldots, n_k$ com cada $n_j \leq t$ de modo que

1. $V = \bigoplus_{j=1}^{k} Z_{\boldsymbol{v}_j}$;
2. T-aniquilador de $\boldsymbol{v}_j$ é p^{n_j}.
3. $\dim V = (n_1 + \cdots + n_k) \operatorname{grau} p$.
4. p^{n_j} divide $p^{n_{j+1}}$.
5. Além disso, os polinômios mônicos $p^{n_1}, \cdots, p^{n_k}$ são únicos e determinados por T.

Demonstração. As três primeiras afirmações seguem diretamente da Decomposição Cíclica.

A quarta afirmação é direta pois sem perda de generalidade, podemos assumir que os vetores T-cíclicos $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ do Teorema 10.1 estejam ordenados de modo que os números inteiros correspondentes n_i satisfaçam

$$t = n_1 \geq n_2 \geq \cdots \geq n_k \geq 1.$$

Afirmação: Os números inteiros $n_1, \ldots, n_k$ são determinados de maneira única por T. Essa é o ponto central que precisa ser demonstrado e faremos isso associando esses números inteiros as dimensões dos espaços $\dim \operatorname{im} p(T)^j$.

Temos, para todo i,

$$\dim Z_{\boldsymbol{v}_i} = \operatorname{grau} m_{\boldsymbol{v}_i} = \operatorname{grau} p^{n_i} = dn_i.$$

Precisaremos de dois fatos cuja demonstração deixaremos como exercício:

- Para cada j a imagem de $Z_{\boldsymbol{v}_i}$ por $p(T)^j$ é o subespaço T-cíclico $Z_{p(T)^j(\boldsymbol{v}_i)}$.
- Como o T-aniquilador de $\boldsymbol{v}_i$ é p^{n_i}, de grau dn_i, temos que

$$\dim Z_{p(T)^j(\boldsymbol{v}_i)} = \begin{cases} 0 & \text{se } j \geq n_i; \\ d(n_i - j) & \text{se } j < n_i. \end{cases}$$

Todos os vetores $\boldsymbol{v} \in V$ podem ser escritos de maneira única na forma

$$\boldsymbol{v} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k \ (\boldsymbol{v}_i \in Z_{\boldsymbol{v}_j})$$

e assim todos os elementos de $\operatorname{im} p(T)^j$ podem ser escritos de maneira única como

$$p(T)^j(\boldsymbol{v}) = p(T)^j(\boldsymbol{v}_1) + \cdots + p(T)^j(\boldsymbol{v}_k) \ .$$

Portanto, se r é o número inteiro tal que $n_1, \ldots, n_r > j$ e $n_{r+1} \leq j$, então vemos que

$$\operatorname{im} p(T)^j = \bigoplus_{i=1}^{r} Z_{p(T)^j(\boldsymbol{v}_i)}$$

e consequentemente

$$\dim \operatorname{im} p(T)^j = d \sum_{i=1}^{r} (n_i - j) = d \sum_{n_j > j} (n_i - j) \ .$$

E logo

$$\begin{aligned}
\dim \operatorname{im} p(T)^{j-1} - \dim \operatorname{im} p(T)^j &= d \left(\sum_{n_j > j-1} (n_j - j + 1) - \sum_{n_j > j} (n_i - j) \right) && (10.5) \\
&= d \left(\sum_{n_j \geq j} (n_j - j + 1) - \sum_{n_j \geq j} (n_i - j) \right) && (10.6) \\
&= d \sum_{n_i \geq j} (n_i - j + 1 - n_i + j) && (10.7) \\
&= d \sum_{n_i \geq j} 1 && (10.8)
\end{aligned}$$

Agora, as dimensões à esquerda são determinadas por T, de modo que a expressão acima fornece, para todo j, o número de n_i que é maior ou igual a j. Isso determina a sequência

$$t = n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$$

completamente. ■

10.2 Definição

Quando o polinômio minimal de T é da forma p^t, em que p é irredutível, em relação à cadeia de números inteiros determinada unicamente $t = n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$, conforme descrito acima, os polinômios $p^t = p^{n_1}, p^{n_2}, \ldots, p^{n_k}$ são denominados **divisores elementares** de T.

Destacamos que o primeiro divisor elementar na sequência é o polinômio minimal de T.

10.3 Teorema Forma Canônica Racional por Divisores Elementares

Sejam V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e $T : V \to V$ um operador linear cujos polinômios característico e minimal são

$$c_T = p_1^{d_i} p_2^{d_2} \cdots p_k^{d_k}, \qquad m_T = p_1^{e_j} p_2^{e_2} \cdots p_k^{e_k}$$

onde $p_1, \ldots, p_k$ são polinômios irredutíveis distintos.

Pelo Teorema da Decomposição Primária que existe uma base ordenada de V com relação à qual a matriz de T é uma matriz diagonal por bloco

$$\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix}$$

em que cada A_i é a matriz (de tamanho $d_i \operatorname{grau} p_j \times d_j \operatorname{grau} p_j$) que representa a aplicação induzido T_i em $V_i = \ker p_j(T)^{e_j}$. Agora, o polinômio minimal de T_i é $p_i^{e_i}$ e, portanto, pela Proposição 10.2 e pelo Teorema da Decomposição Cíclica, existe uma base para V_j com relação ao qual A_j é a matriz diagonal por bloco

$$\begin{bmatrix} C_{i1} & & & \\ & C_{i2} & & \\ & & \ddots & \\ & & & C_{it} \end{bmatrix}$$

em que C_{ij} são as matrizes associadas aos divisores elementares de T_i.

Pela discussão anterior, esse formato diagonal por bloco, no qual cada bloco A_j é ele próprio uma diagonal por bloco de matrizes companheiras, é único (a menos de permutações de A_j). É denominada **matriz canônica racional por divisores elementares** de T.

É importante notar que na sequência de divisores elementares pode haver repetições, pois alguns dos n_i podem ser iguais. O resultado disso é que algumas matrizes companheiras podem aparecer mais de uma vez na forma racional.

Exemplo 4. *Seja*

$$A = \begin{bmatrix} -3 & 0 & 12 \\ -2 & 1 & 6 \\ -2 & 0 & 7 \end{bmatrix} \in \mathcal{M}_{3,3}(\mathbb{Q}).$$

Nesse caso $p_T = -(x-3)(x-1)^2$ *e* $m_T = (3-x)(-1+x)$. *Como a matriz companheira de qualquer polinômio linear* $X - a$ *é uma matriz* 1×1. *Concluímos que a forma canônica racional de* A *deve ser a matriz diagonal* $D = \operatorname{diag}(1,1,3)$

◁

Exemplo 5. *Seja*

$$A = \begin{bmatrix} 8 & -4 & -12 & 7 \\ 5 & -5 & -4 & 4 \\ 3 & -3 & -2 & 2 \\ 0 & -3 & 5 & -1 \end{bmatrix} \in \mathcal{M}_{4,4}(\mathbb{Q})$$

◁

Então $p_T(x) = m_T(x) = (x^2-2)(x^2+1)$ e logo a forma racional é

$$C_{x^2-2} \oplus C_{x^2+1} \begin{bmatrix} 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Exemplo 6. *Suponha agora que* $T\colon \mathbb{Q}^6 \to \mathbb{Q}^6$ *possua um polinômio minimal*

$$m_T = (X^2+1)(X-2)^2$$

Consequentemente o polinômio característico de T *é um dos*

$$c_1 = (X^2+1)^2(X-2)^2, \quad c_2 = (X^2+1)(X-2)^4$$

Suponha primeiro que $c_T = C_1$. *Nesse caso, temos* $\mathbb{Q}^6 = V_1 \oplus V_2$ *com* $\dim V_1 = 4$ *e* $\dim V_2 = 2$. *A aplicação linear induzida* T_1 *em* V_1 *possui o polinômio minimal* $m_1 = x^2 + 1$ *e a aplicação induzida* T_2 *em* V_2 *possui polinômio minimal* $m_2 = (X-2)^2$.

A situação para V_1 *é que* $C_{X^2+1} \oplus C_{X^2+1}$.

Quanto a V_2, *logo pelo Teorema da Forma Canônica Racional temos que* $2 = n_1 + \cdots + n_k$ *de onde necessariamente* $k = 1$ *pois* $n_1 = \operatorname{grau} p_2 = 2$. *Portanto, o único divisor elementar de* T_2 *é* $(X-2)^2$. *Combinando essas observações, vemos que, neste caso, a matriz canônica racional de* T *é*

$$C_{X^2+1} \oplus C_{X^2+1} \oplus C_{(X-2)^2}.$$

Suponha agora que $c_T = C_2$. *Nesse caso, temos* $\mathbb{Q}^6 = V_1 \oplus V_2$ *com* $\dim V_1 = 2$ *e* $\dim V_2 = 4$. *Além disso, a aplicação induzido* T_2 *em* V_2 *possui o polinômio minimal* $m_2 = (X-2)^2$. *Pelo Teorema 10.2, aplicado a* T_2, *temos* $4 = n_1 + \cdots + n_k$ *com* $n_1 = 2$. *Existem, portanto, duas possibilidades, a saber:*

- $k = 2$ *com* $n_1 = n_2 = 2$;
- $k = 3$ *com* $n_1 = 2, n_2 = n_3 = 1$.

A matriz canônica racional de T *neste caso é, portanto, uma das formas*

$$C_{X^2+1} \oplus C_{(X-2)^2} \oplus C_{(X-2)^2};$$

$$C_{X^2+1} \oplus C_{(X-2)^2} \oplus C_{X-2} \oplus C_{X-2}.$$

Observe no exemplo acima que o conhecimento dos polinômios característico e minimal geralmente não é suficiente para determinar completamente a forma racional.

◁

'

Podemos também obter unicidade da Decomposição Cíclica se impormos que $f_{i+1}(x) | f_i(x)$. Desse modo obtemos a Forma Canônica Racional por Fatores Invariantes.

10.7 Teorema Forma Canônica Racional por Fatores Invariantes

Seja $T \in L(V)$. Existe um inteiro $r > 0$ e vetores u_1, u_r e polinômios $f_1(x), \ldots, f_r(x)$ tais que:

1. $V = Z_{\boldsymbol{v}_1} \oplus \cdots \oplus Z_{\boldsymbol{v}_r}$, com $m_{\boldsymbol{v}_i}(x) = f_i(x)$,
2. $f_1(x) = m_T(x)$ e $f_{i+1}(x) | f_i(x)$ para todo $i = 1, \ldots, r-1$.

Além disso o inteiro r e os polinômios $f_1, \ldots, f_r$ são unicamente determinados por T.

10.8 Definição

Os polinômios $f_1, \ldots, f_r$ são denominados **fatores invariantes** de T.

Demonstração. **Existência dos fatores invariantes:** Para obter os fatores invariantes a partir divisores elementares alinhamos em uma matriz os divisores elementares do maior grau ao menor grau, com cada linha correspondendo a um fator irredutível do polinômio minimal. Os fatores invariantes são os produtos ao longo das colunas dessa matriz, completando com $1 = p_i(x)^0$ as células vazias, quando necessário.

p_1	$p_1^{n_{11}}$	$p_1^{n_{12}}$	$\cdots$	$p_1^{n_{1j}}$	$\cdots$
p_2	$p_1^{n_{11}}$	$p_2^{n_{12}}$	$\cdots$	$p_2^{n_{1j}}$	$\cdots$
$\vdots$					$\vdots$
p_k	$p_k^{n_{k1}}$	$p_k^{n_{k2}}$	$\cdots$	$p_k^{n_{kj}}$	$\cdots$
fatores elementares	$f_1 = p_1^{n_{11}} \cdots p_k^{n_{k1}}$	$f_2 = p_1^{n_{21}} \cdots p_k^{n_{k2}}$	$\cdots$	$f_j = p_j^{n_{1j}} \cdots p_k^{n_{kj}}$	$\cdots$

De maneira formal. Sejam $p_i(x)^{n_{ij}}$ com $i = 1, \ldots, k$ e $j = 1, \ldots r_i$ os divisores elementares de T. Seja $r = \max\{r_i,\ i = 1, \ldots, k\}$.

Se $r_i < r$ e $j = r_i + 1, \ldots, r$, redefiniremos $n_{ij} = 0,$. Definimos finalmente os fatores invariantes

$$f_i(x) \triangleq p_1(x)^{n_{1j}} \ldots p_k(x)^{n_{kj}} \text{ para } j = 1, \ldots r.$$

Temos que

- $f_1(x) = n_T(x)$
- $f_{i+1}(x) | f_i(x)$ para todo $j = 1, \ldots r - 1$

A primeira afirmação segue do Teorema 10.2. Como $n_{ij} \geq n_{ij+1}$ então $f_{i+1}(x) | f_i(x)$.

Pelo Teorema 10.2. Temos que

$$V = \bigoplus Z_{\boldsymbol{v}_{1j}}$$

como $m_{\boldsymbol{v}_{1j}} = pij$ e com

Se $\boldsymbol{u}_j = \boldsymbol{v}_{1j} + \cdots + \boldsymbol{v}_{kj}$ então $n_{\boldsymbol{u}_j}(x) = f_i(x)$ pelo exercício 10.1.

Também é fácil ver que

$$V = Z_{\boldsymbol{u}_1} \oplus \cdots \oplus Z_{\boldsymbol{u}_r}.$$

Unicidade dos fatores invariantes:

Suponha que existem vetores $\boldsymbol{w}_1, \boldsymbol{w}_s$ com T-anuladores $n_{\boldsymbol{w}_j(X)} = g_j(x)$, $g_{j+1}(x)|g(x)$ para todo $j = 1, \ldots, s-1$ e que $V = Z_{\boldsymbol{w}_1,} \oplus \oplus Z_{\boldsymbol{w}_s,}$.

Observe que $n_T(x) = g_1(x)$ (por que?). Assim, $g_j(x) = p_1(x)^{l_i} i \ldots p_k(x)^{l_k} i$ e como $g_{j+1}(x)|g(x)$ vale que para todo $i = 1, \ldots, k$. Agora, aplique o Teorema da decomposição Primária em cada $Z_{\boldsymbol{u}_j}$. Como $Z_{\boldsymbol{u}_j}$ e é cíclico, seus subespaços são cíclicos, e então existem vetores $\boldsymbol{y}_{ij}$ com $i = 1, \ldots, s$ com T-anuladores $n_{\boldsymbol{y}_{ij}}(x) = p_i(x)^{l_i} i$ tais que

$$Z_{\boldsymbol{u}_j,} = Z_{\boldsymbol{y}_{1j},} \oplus \cdots \oplus Z_{\boldsymbol{y}_{kj},} .$$

Observe que $W_i = \ker p_i(T)^{n_i} = Z_{\boldsymbol{y}_{i1},} \oplus \cdots \oplus Z(y_{is})$ (prove isso!) Seja, para todo $i = 1, \ldots, k, s_i$ o maior inteiro j tal que $l_{ij} \neq 0$. Pelo Teorema 10.2, temos que $r_i = s_i$ e $n_{ij} = l_{ij}$. Assim, vale que $f_i(x) = g_j(x)$ para todo j. ■

10.3 Forma Normal de Jordan Generalizada

Nesta seção modificaremos a forma canônica racional de forma a obter uma nova forma canônica a forma normal de Jordan generalizada. Essa forma se reduz à forma de Jordan quando todos os autovalores pertencerem ao corpo.

Obteremos essa forma modificando as bases cíclicas que foram usadas para obter a forma racional. Ao fazer isso, obteremos uma representação matricial que é construída a partir da matriz companheira de p_j, em vez da matriz companheira de $p_i^{n_i}$

10.1 Teorema

Seja $\boldsymbol{v}$ um vetor cíclico de V e seja $T : V \to V$ um operador linear com o polinômio minimal p^n Então existe uma base de V na qual a matriz de T é a matriz $kn \times kn$

$$\begin{bmatrix} C_p & N & & & \\ & C_p & N & & \\ & & \ddots & \ddots & \\ & & & C_p & N \\ & & & & C_p \end{bmatrix}$$

onde C_p é a matriz companheira de p e $N = (n_{ij})$ é a matriz $k \times k$ com todas as entradas nulas exceto a entrada $(n_{nn}) = 1$:

$$\begin{bmatrix} \text{O} & ^{1} \end{bmatrix}$$

Demonstração. Considere o conjunto de $n \times k$ vetores

$$\underline{\mathrm{B}} = \left\{ \begin{array}{ccccc} T^{k-1}(\boldsymbol{v}) & T^{k-2}(\boldsymbol{v}) & \cdots & T(\boldsymbol{v}) & \boldsymbol{v} \\ p(T)T^{k-1}(\boldsymbol{v}) & p(T)T^{k-2}(\boldsymbol{v}) & \cdots & p(T)T(\boldsymbol{v}) & p(T)\boldsymbol{v} \\ & & \vdots & & \\ p(T)^{n-1}T^{k-1}(\boldsymbol{v}) & p(T)^{n-1}T^{k-2}(\boldsymbol{v}) & \cdots & p(T)^{n-1}T(\boldsymbol{v}) & p(T)^{n-1}\boldsymbol{v} \end{array} \right\}$$

Primeiro demonstraremos que $\underline{\mathrm{B}}$ é uma base de V. Para esse propósito, basta mostrar que os vetores em $\underline{\mathrm{B}}$ são linearmente independentes. Suponhamos, de modo a obter uma contradição, que não fosse o caso. Então, uma combinação linear não trivial desses elementos seria $\mathbf{0}$ e, consequentemente, existiria um polinômio q tal que $q(T)(\boldsymbol{v}) = \mathbf{0}$ com $\operatorname{grau} q < kn = \operatorname{grau} p^n$. Como $\boldsymbol{v}$ é cíclico, isso contraria a hipótese de que p^n é o polinômio minimal de T. Portanto, a matriz acima constitui uma base para V.

Ordenamos essa base linha por linha, como normalmente lemos.

Observe agora que T leva cada elemento da matriz para seu predecessor na mesma linha, exceto aqueles no início de uma linha. Para esses elementos, temos

$$T[T^{k-1}(\boldsymbol{v})] = T^k(\boldsymbol{v}) = -\alpha_{k-1}T^{k-1}(\boldsymbol{v}) - \cdots - \alpha_0 x + p(T)(\boldsymbol{v})$$

e da mesma forma

$$\begin{aligned} T[p(T)^{nm}(T^{k-1})(\boldsymbol{v})] &= p(T)^{nm}[T^k(\boldsymbol{v})] && (10.9)\\ &= p(T)^{nm}[-\alpha_{k-1}T^{k-1}(\boldsymbol{v}) - \cdots - \alpha_0\boldsymbol{v} + p(T)(\boldsymbol{v})] && (10.10)\\ &= -\alpha_{k-1}p(T)^{nm}[T^{k-1}(\boldsymbol{v})] - \cdots - \alpha_0 p(T)^{nm}(\boldsymbol{v}) + p(T)^{n-m+1}(\boldsymbol{v})\,. && (10.11) \end{aligned}$$

Agora é simples verificar se a matriz de T em relação à base ordenada acima é da forma descrita. ■

10.2 Definição

A matriz de blocos descrita no 10.3 é denominada matriz de Jordan generalizada associada à matriz companheira C_p.

Passando agora para o caso geral, considere uma aplicação linear $T\colon V \to V$. Com a notação usual, a aplicação induzida T_i em V_i possui polinômio minimal $p_i^{e_i}$. Se agora aplicarmos o 10.3 aos subespaços cíclicos que aparecem no Teorema da Decomposição Cíclica para T_i, veremos que na matriz canônica racional de T_i, podemos substituir cada bloco diagonal de matrizes companheiras C_{ij} associado aos divisores elementares $p_i^{n_j}$ pela matriz de Jordan generalizada associada à matriz companheira. A matriz que surge dessa maneira é denominada matriz de jordan generalizada f.

Exemplo 3. *Seja $T: \mathbb{R}^6 \to \mathbb{R}^6$ seja linear e tal que*

$$c_T = m_T = (X^2 - X + 1)^2 (X + 1)^2$$

Pelo Teorema da Decomposição Primária, temos que o polinômio minimal da aplicação induzido T_1 é $(X^2 - X + 1)^2$ enquanto o polinômio minimal da aplicação induzida T_2 é $(X+1)^2$. Usando o Teorema 10.2, vemos que o único divisor elementar de T_1 é $(X^2 - X + 1)^2$ e o único divisor elementar de T_2 é $(X+1)^2$. Consequentemente, a matriz canônica racional de T é

$$C_{(X^2-X+1)^2} \oplus C_{(X+1)^2} = \left[\begin{array}{cccc|cc} 0 & 0 & 0 & -1 & & \\ 1 & 0 & 0 & 2 & & \\ 0 & 1 & 0 & -3 & & \\ 0 & 0 & 1 & 2 & & \\ \hline & & & & 0 & -1 \\ & & & & 1 & -2 \end{array}\right]$$

A matriz de Jordan generalizada de T pode ser obtida a partir da seguinte forma. Com $p_1 = x^2 - X + 1$, substituímos a matriz companheira associada a p_1^2 pela de Jordan generalizada associada à matriz companheira de p_1. Realizamos uma substituição semelhante em relação a $p_2 = x + 1$.Obtendo

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

Finalmente, vamos observar o caso particular do 10.3 que ocorre quando $k = 1$. Nessa situação, temos $p = X - \alpha_0$ e $f - \alpha_0 \mathrm{id}_y$ é nilpotente do índice n. Consequentemente, C_p reduz para a matriz 1×1 $[\alpha_0]$ e a de Jordan generalizada associada a C_p reduz para $n \times n$ matriz Jordan elementar associada ao valor próprio α_0. Assim, vemos que quando todos os autovalores pertencem ao corpo, a forma de Jordan generalizada se reduz à forma de Jordan.

$\triangleleft$

Exemplo 4. *Seja $T: \mathbb{R}^7 \to \mathbb{R}^7$ seja linear e tal que*

$$c_T = (X-2)^3 (X-3)^4,$$

A matriz canônica racional é

$\triangleleft$

10.4 Calculando a Forma Canônica Racional

Nesta seção apresentamos algorítmos para calcular formas canônicas racionais e de Jordan Generalizada, continue lendo.

Seja $T : V \to V$ uma transformação linear. Suponha que $p(x)$ seja um fator irredutível do polinômio minimal de T do grau d. Construiremos o diagrama de pontos correspondente a $p(x)$, que indicará os divisores elementares correspondentes a $p(x)$. Esse cálculo é baseado na Equação 10.8 da Proposição 10.2.

Para esse fim calcularemos recursivamente os seguintes números até obtermos 0:

$$
\begin{aligned}
r_1 &= \frac{1}{d}\left(\dim(V) - \text{posto}(p(T))\right) \\
r_k &= \frac{1}{d}\left(\text{posto}(p(T)^{k-1}) - \text{posto}(p(T)^k)\right) \text{ para } k \geq 2
\end{aligned}
$$

A partir desses dados vamos construir um diagrama com r_k pontos na k-ésima linha. Os divisores elementares são $p(x)^{s_i}$, onde s_i é o número de pontos na i-ésima coluna do diagrama de pontos.

Exemplo 1. *Suponha que V é um espaço vetorial. Suponha que $p(x) = x^2 + 1$ e calculemos os números r como $r_1 = 3, r_2 = 3, r_3 = 1, r_4 = 1, r_5 = 0$.*

O diagrama de pontos seria:

$$
\begin{array}{ccc}
\bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet \\
\bullet & & \\
\bullet & &
\end{array}
$$

Logo os divisores elementares correspondentes a $x^2 + 1$ seriam então $(x^2 + 1)^4, (x^2 + 1)^2, (x^2 + 1)^2$.

$\triangleleft$

Para obter a forma canônica racional por meio de divisores elementares, faça isso para todo fator irredutível do polinômio minimal e, em seguida, preencha uma matriz com as matrizes companheiras dos divisores elementares ao longo da diagonal.

Para obter os fatores invariantes a partir divisores elementares: Em uma matriz, alinhe os divisores elementares do maior grau ao menor grau, com cada linha correspondendo a um fator irredutível do polinômio minimal . Os fatores invariantes são os produtos ao longo das colunas dessa matriz. O maior fator invariante é sempre o polinômio minimal.

Por exemplo: Suponha que os divisores elementares correspondentes a $x^2 + 1$ sejam como acima, e os divisores elementares correspondentes a $x - 2$ sejam:

$(x-2)^3, (x-2)$. E que esses são os únicos fatores irredutíveis do polinômio minimal .

Então os fatores invariantes são

$$(x^2+1)^4(x-2)^3, (x^2+1)^2(x-2), (x^2+1)^2.$$

Para obter a forma canônica racional por meio de fatores invariantes, preencha uma matriz com matrizes companheiras dos fatores invariantes ao longo da diagonal.

Exemplo 2. *Vamos encontrar a forma racional de*

$$A = \begin{bmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{bmatrix}$$

Logo o polinômio característico de A é $p_A = (t-1)^3$ e o polinômio minimal é $m_A = (t-1)^2$, o único fator irredutível do polinômio minimal é $(t-1)$. Vamos calcular seu diagrama de pontos.

$$\begin{aligned} r_1 &= (3 - \operatorname{posto}(A-I)) = (3-1) = 2 \\ r_2 &= (\operatorname{posto}(A-I) - \operatorname{posto}((A-I)^2)) = (1-0) = 1 \\ r_3 &= (\operatorname{posto}((A-I)^2) - \operatorname{posto}((A-I)^3) = (0-0) = 0 \end{aligned}$$

Portanto, o diagrama de pontos é

$$\begin{array}{cc} \bullet & \bullet \\ \bullet & \end{array}$$

Portanto, os divisores elementares são $(t-1)^2, (t-1)$. Nesse caso, como existe apenas um fator irredutível do polinômio minimal, os fatores invariantes são os mesmos que os divisores elementares. Enfim, entendemos que a forma canônica racional (via divisores elementares e fatores invariantes) é:

$$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\triangleleft$

Exercícios

Mostre que toda matriz $A \in M_n(\mathbb{C})$ é semelhante à sua transposta.

Ex. 10.2 --- Seja V um $\mathbb{K}$-e.v. de $\dim V < \infty$, $v \in V$ e $T : V \to V$ um operador linear, seja $Z(v;T) = \bigcap_{W \in C(v;T)} W$ onde $C(v;T) = \{W \subseteq V \mid W$ é subespaço T-invariante e $v \in W\}$. Mostre que:

1. $Z(v;T)$ é T-invariante

2. $Z(v;T)$ é o menor subespaço de V, T-invariante que contem v.

3. $Z(v;T) = \{g(T)(v) \in V \mid g \in \mathbb{K}[x]\}$

4. $\dim Z(v;T) = 1$ se, e somente se, v é um autovetor de T.

Ex. 10.3 --- Seja $\mathbb{K}$ um corpo e seja $B \in M_n(\mathbb{K})$. Mostre que B é semelhante sobre $\mathbb{K}$ a uma e somente uma matriz que está na forma racional.

Ex. 10.4 --- Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathrm{Hom}(V,V)$. Prove que se existe um vetor cíclico para T^2 então existe um vetor cíclico para T. Vale a recíproca?

Ex. 10.5 --- Seja V um espaço vetorial sobre $\mathbb{K}$ e $T : V \to V$ um operador linear com polinômio minimal $m_T(x) = p_1(x)^{m_1} \cdots p_k(x)^{m_k}$ onde os $p_i(x)$ são distintos e irredutíveis em $\mathbb{K}[x]$. Seja $V = W_1 \oplus \cdots \oplus W_k$ a decomposição de V dada pelo teorema da decomposição primária. Mostre que para todo subespaço T-invariante W de V temos que

$$W = \bigoplus_{i=1}^{k} (W \cap W_i).$$

Ex. 10.6 --- Seja V um espaço vetorial sobre $\mathbb{K}$ e $T : V \to V$ um operador linear. Sejam $p \in \mathbb{K}[x]$ um polinômio irredutível e m um inteiro positivo. Suponha que existe $v \in V$ tal que $p_v = p^m$. Seja $w \in Z(v;T)$.

1. Mostre que $w = q(T)p(T)^l(v)$ onde $q \in \mathbb{K}[x]$ é coprimo com p e $0 \le l \le m$.

2. Mostre que $Z(w;T) = Z(p(T)^l(v);T)$. Logo os únicos subespaços T-cíclicos de $Z(v;T)$ são

$$0 = Z(p(T)^m(v);T) \subset Z(p(T)^{m-1}(v);T) \subset \cdots \subset Z(p(T)(v);T) \subset Z(v;T).$$

3. Mostre que todo subespaço T-invariante de $Z(v;T)$ é T-cíclico, ou seja, é da forma $Z(p(T)^l(v);T)$ com $0 \le l \le m$.

Ex. 10.7 --- Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \mathrm{Hom}(V,V)$.

1. Prove que se existe um vetor cíclico para T então todo subespaço próprio T-invariante de V também tem um vetor cíclico.

2. Vale a recíproca do item 1?

Ex. 10.8 --- Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão n e $T \in \operatorname{Hom}(V,V)$ um operador diagonalizável.

1. Mostre que existe um vetor cíclico para T se, e somente se, T tem n autovalores distintos.

2. Mostre que se T tem n autovalores distintos e se $\underline{B} = \{v_1, \dots, v_n\}$ é uma base de autovetores de T, então $v = v_1 + v_2 + \cdots + v_n$ é um vetor cíclico de T.

Ex. 10.9 --- Sejam V um K-espaço vetorial de dimensão n e $T \in \operatorname{Hom}(V,V)$.

1. Prove que $\operatorname{im} T$ tem um complementar T-invariante se, e somente se, $\operatorname{im} T \cap \ker T = 0$.

2. Se $\operatorname{im} T \cap \ker T = 0$, prove que $\ker T$ é o único complementar de $\operatorname{im} T$ que é T-invariante.

Ex. 10.10 --- Seja $A \in M_3(\mathbb{R})$ a matriz

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

Encontre uma matriz invertível P tal que $P^{-1}AP$ esteja na forma racional.

Ex. 10.11 --- Seja $A \in M_3(\mathbb{R})$ a matriz

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

Encontre vetores $v_1, \cdots, v_r$ que satisfazem as condições do Teorema da Decomposição Cíclica.

Ex. 10.12 --- Sejam V um K-espaço vetorial de dimensão finita e $T \in \operatorname{Hom}(V,V)$. Prove que T tem um vetor cíclico se, e somente se, a seguinte afirmação é verdadeira: ``Todo operador linear que comuta com T é um polinômio em T.''

Ex. 10.13 --- Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \operatorname{Hom}(V,V)$. Prove que todo vetor $0 \neq v \in V$ é um vetor cíclico para T se, e somente se, o polinômio característico de T é irredutível em $\mathbb{K}[x]$.

Ex. 10.14 --- Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão finita e $T \in \operatorname{Hom}(V,V)$. Suponha que $c_T = m_T = q^m$ é uma potência de um polinômio irredutível. Prove que nenhum subespaço não trivial T-invariante de V tem um complementar que também é T-invariante.

Ex. 10.15 --- Seja $T \in \mathrm{Hom}(\mathbb{R}^4)$ um operador linear tal que x^2+3 é um divisor do polinômio minimal de T e 1 é o único autovalor de T. Quais são as possíveis formas racionais de T?

Ex. 10.16 --- Seja $T \in \mathrm{Hom}(\mathbb{K}^6)$ com polinômio minimal $m_T(x) = (x^2-2)(x^2+1)$. Ache as possíveis formas racionais de T para $\mathbb{K} = \mathbb{R}$ e $\mathbb{K} = \mathbb{C}$.

Ex. 10.17 --- Um operador linear T é dito *semissimples* se todo subespaço T-invariante de V tem um complemento que é também T-invariante. Seja V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e T um operador linear em V.

1. Mostre que se o polinômio minimal de T for irredutível em $\mathbb{K}$ então T é semissimples.

2. Se $\mathbb{K}$ for algebricamente fechado, prove que T é diagonalizável se, e somente se, é semissimples.

3. Mostre que o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que

$$[T]_{\underline{\mathsf{B}}}^{\underline{\mathsf{B}}} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

   não é semissimples.

Appendices

Capítulo

A

Teoria de Conjuntos

Dados dois conjuntos A e B, diremos que A é **subconjunto** de B, e escreveremos $A \subset B$, se cada elemento de A pertence a B, ou seja, se $x \in A$ implica $x \in B$.

Diremos que A é **igual** a B, e escreveremos $A = B$, se A e B possuem os mesmos elementos, ou seja, se $A \subset B$ e $B \subset A$.

O **conjunto vazio** será denotado por $\emptyset$.

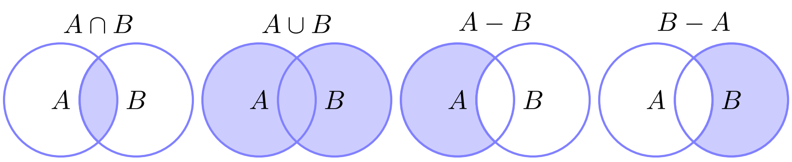
A.1 Definição

Definimos a **união**, a **intersecção** , e a **diferenca** de dois conjuntos A e B é definida por

$$A \cup B = \{x : x \in A \text{ ou } x \in B\},$$

$$A \cap B = \{x : x \in A \text{ e } x \in B\},$$

$$A \backslash B = \{x : x \in A \text{ e } x \notin B\}.$$



Se estamos considerando subconjuntos de um conjunto fixo X, então conjunto $X \backslash A$ é denominado de **complementar** de A em X, e é denotado por $A^{\complement}$.

Dado um conjunto X o conjunto das partes $\mathcal{P}(X)$ é o conjunto formado pelos subconjuntos de X, ou seja

$$\mathcal{P}(X) = \{A : A \subset X\}.$$

Nesse capítulo

- Operações com famílias de conjuntos (p. 298)
- Relações (p. 302)
- Cardinalidade (p. 308)

A.2 Definição Produto Cartesiano

O **produto cartesiano** $X \times Y$ de dois conjuntos X e Y é o conjunto dos pares ordenados $(x,\ y)$ tais que $x \in X$ e $y \in Y$.

O produto cartesiano $X_1 \times \times X_n$ de n conjuntos $X_1, \ldots, X_n$ é o conjunto das ênuplas $(x_1,\ x_n)$ tais que $x_i \in X_i$ para $i = 1, n$.

Escreveremos X^n em lugar de $X \times \cdots \times X$ (n vezes).

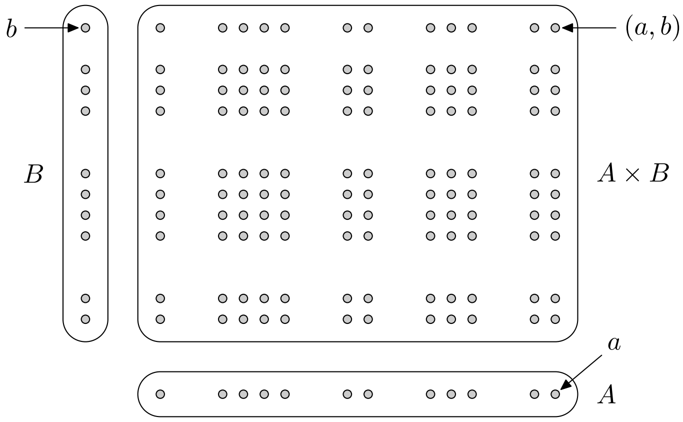


Figura A.1 Produto Cartesiano de dois conjuntos

A.3 Definição

Uma **função** ou **aplicação** f de X em Y, denotada por $f : X \to Y$, é uma regra que associa a cada elemento $x \in X$ um único elemento $f(x) \in$ Y. Nesse caso o conjunto X é denominado de domínio de f e o conjunto Y é denominado de contradomínio de f.

A.4 Definição

Uma função f é dita

a) **injetiva** se $f(x_1) = f(x)$ implica $x_1 = x_2$.

b) **sobrejetiva** se para todo $y \in Y$ existe $x \in X$ tal que $f(x) = y$.

c) **bijetiva** se é injetiva e sobrejetiva.

Se $f : X \to Y$ é bijetiva, a **função inversa** $f^{-1} : Y \to X$ é definida por $f^{-1}(y) = x$ se $f(x) = y$.

Dados $A \subset X$ e $B \subset Y$, a **imagem** de A e a **imagem inversa** de B são respectivamente os conjuntos

$$f(A) = \{y \in Y : y = f(x) \text{ para algum } x \in A\}, \tag{A.1}$$
$$f^{-1}(B) = \{x \in X : f(x) \in B\}. \tag{A.2}$$

Dadas duas aplicações $f : X \to Y$ e $g : Y \to Z$, a **aplicação composta** $g \circ f : X \to Z$ é definida por $g \circ f(x) = d(f(x))$ para todo $x \in X$.

A.5 Definição

- $g : B \to A$ é uma inversa à esquerda de $f : A \to B$ se, e somente se, $g(f(a)) = a$ para todos $a \in A$
- $h : B \to A$ é uma inversa à direita de $f : A \to B$ se, e somente se, $f(h(b)) = b$ para todos os $b \in B$

A.6 Teorema

- Uma função é injetiva se, e somente se, possuir uma inversa à esquerda.
- Uma função é sobrejetiva se, e somente se, possuir uma inversa à direita.

A.1 Operações com famílias de conjuntos

Nesta seção, lidaremos com *famílias (ou classes) de conjuntos*, isto é, conjuntos cujos elementos são, por sua vez, também conjuntos. Queremos estender a essa situação algumas operações entre conjuntos, assim como descrever algumas propriedades.

Seja $\mathcal{F}$ uma família de conjuntos, i.e. $\mathcal{F} = \{A_\imath\}_{\imath \in J}$ onde J é um qualquer conjunto de índices e cada $A_\imath$ é um conjunto.

Exemplo 1.

$$A_i = (0, i] \subset \mathbb{R}$$

com $i \in \mathbb{N}$

◁

Exemplo 2.

$$B_x = (x, x^2] \subset \mathbb{R}$$

com $x \in \mathbb{R}, x > 2$

◁

A **união** dos conjuntos da família $\mathcal{F}$ é o conjunto formado pelos elementos que pertencem a *ao menos um* dos conjuntos de $\mathcal{F}$, i.e.

$$\bigcup_{\imath \in J} A_\imath = \{x \,|\, x \in A_\jmath \text{ para algum } \jmath \in J\}$$

A **intersecção** dos conjuntos da família $\mathcal{F}$ é o conjunto formado pelos elementos que pertencem a *todos* os conjuntos de $\mathcal{F}$, i.e.

$$\bigcap_{\imath \in J} A_\imath = \{x \,|\, x \in A_\jmath \text{ para todo } \jmath \in J\}$$

Dentre as propriedades mais importantes, destacamos as seguintes: dada uma família $\mathcal{F} = \{A_\imath\}_{\imath \in J}$ de conjuntos e dado um conjunto qualquer B, tem-se:

$$B \cap \left(\bigcup_{\imath \in J} A_\imath \right) = \bigcup_{\imath \in J} (B \cap A_\imath)$$

$$B \cup \left(\bigcap_{\imath \in J} A_\imath \right) = \bigcap_{\imath \in J} (B \cup A_\imath)$$

Além disso, se $\mathbb{U}$ é um conjunto que contém todos os conjuntos $A_\imath$, então, tomando o complementar relativamente a $\mathbb{U}$, tem-se:

$$(\bigcup_{\imath \in J} A_\imath)^\complement = \bigcap_{\imath \in J} A_\imath^\complement$$

$$(\bigcap_{\imath \in J} A_\imath)^\complement = \bigcup_{\imath \in J} A_\imath^\complement$$

Utilizaremos a seguinte convenção:

$$\bigcup_{A \in \emptyset} A = \emptyset$$

$$\bigcap_{A \in \emptyset} A = X.$$

Produto Cartesiano

Como primeiro passo, vejamos como definir o produto cartesiano de uma quantidade qualquer (mas finita) de conjuntos. Dados n conjuntos não vazios $A_1, A_2, \ldots, A_n$, o **produto cartesiano** $A_1 \times A_2 \times \cdots \times A_n$ é o conjunto dos elementos na forma $(x_1, x_2, \ldots, x_n)$, onde para todo $1 \leq \imath \leq n$ tem-se que $x_\imath \in A_\imath$. Em símbolos:

$$A_1 \times A_2 \times \cdots \times A_n = \{(x_1, x_2, \ldots, x_n) \,|\, x_\imath \in A_\imath, 1 \leq \imath \leq n\}.$$

Os elementos na forma $(x_1, x_2, \ldots, x_n)$ são denominados de *n-upla ordenada* (que se lê "ênupla" ordenada).

Note-se que o produto cartesiano de n conjuntos é muito semelhante ao produto cartesiano de dois conjuntos, só diferindo, de fato, pelo número de conjuntos envolvidos.

Nosso propósito, agora, é contemplar famílias quaisquer de conjuntos, eventualmente infinitas. Para tanto, não é difícil perceber que a descrição acima não é adequada. Para chegar a um outro modo de tratar o produto cartesiano, pode ser útil revermos, sob outro olhar, o produto cartesiano que nos é já conhecido (vamos considerar o caso mais simples, com somente dois conjuntos). Dados dois conjuntos não vazios A_1 e A_2 (o uso de índices aqui é proposital), podemos identificar um par ordenado (x_1, x_2) do produto cartesiano $A_1 \times A_2$ com a função $f : \{1, 2\} \to (A_1 \cup A_2)$ dada por

$$f(1) = x_1 \quad \text{e} \quad f(2) = x_2$$

Pode parecer um modo exageradamente complicado para descrever um par ordenado e, se fosse esse o único objetivo dessa descrição, seria realmente algo despropositado. Mas essa linguagem apenas traduz a ideia de que um par ordenado nada mais é do que uma escolha particular, simultânea, de um elemento de um conjunto e um de outro. E cada função f como aquela acima descreve exatamente uma escolha desse tipo.

A vantagem dessa linguagem, porém, está no fato de permitir que se defina o produto cartesiano para uma família qualquer de conjuntos. De fato, seja dada uma família de conjuntos

$$\mathcal{F} = \{A_\imath\}_{\imath \in J}$$

onde J é um qualquer conjunto de índices. O **produto cartesiano** dos conjuntos da família $\mathcal{F}$ é o conjunto das funções

$$x : J \to \bigcup_{\imath \in J} A_\imath$$

tais que $x(\jmath) \in A_\jmath$ para todo $\jmath \in J$. Em símbolos:

$$\prod_{\imath \in J} A_\imath = \{x : J \to \bigcup_{\imath \in J} A_\imath \,|\, x(\jmath) \in A_\jmath, \ \forall\, \jmath \in J\}.$$

Escreveremos x_i em lugar de $x(i)$ para todo $x \in \prod_{i \in I}$ e $i \in I$. Para cada $j \in I$ a projeção π_j é definida por

$$\pi_j : x \in \prod_{i \in I} X_i \to x_j \in X_j.$$

Cada $x \in \prod_{i \in I} X_i$ é usualmente denotado por $(x_i)_{i \in I}$.

Mesmo que todo X_i seja não vazio, não é óbvio que o produto $\prod_{i \in I} X_i$ seja não vazio. Isto é consequência do seguinte axioma.

A.3 Axioma Axioma da Escolha

Seja $\{X_i : i \in I\}$ uma família não vazia de conjuntos disjuntos não vazios. Então existe uma função $f : I \to \bigcup_{i \in I} X_i$ tal que $f(i) \in X_i$ para todo $i \in I$. A função f é denominada função escolha.

O Axioma da Escolha é um princípio fundamental na Teoria dos Conjuntos. Até o final do século XIX, o Axioma da Escolha era frequentemente usado implicitamente nas demonstrações. Seu usou parece ter passado despercebido até o desenvolvimento formal da Teoria dos Conjuntos durante os esforços para estabelecer uma base sólida para a matemática. Nesse contexto, o matemático alemão Ernst Zermelo foi um dos primeiros a enunciar formalmente o Axioma.

No entanto, o Axioma da Escolha gerou controvérsia na época de sua formulação. Alguns matemáticos, como Brouwer, eram céticos em relação ao Axioma da Escolha, pois esse axioma levava a resultados paradoxais.

A controvérsia em torno do Axioma da Escolha culminou na "crise dos fundamentos" na matemática, que levou ao desenvolvimento de diferentes sistemas de axiomas e à investigação mais profunda sobre os fundamentos da matemática. Eventualmente, a aceitação do Axioma da Escolha como um dos axiomas padrão da Teoria dos Conjuntos prevaleceu, e ele é amplamente utilizado em matemática moderna.

A.4 Proposição

Seja $\{X_i : i \in I\}$ uma família não vazia de conjuntos não vazios. Então o produto cartesiano $\prod_{i \in I} X_i$ é não vazio.

Demonstração. Se os conjuntos X_i fossem disjuntos, a conclusão seria consequência imediata do Axioma da Escolha. No caso geral definamos $Y_i = X_i \times \{i\}$ para todo $i \in I$. É claro que $\{Y_i \ : \ i \in I\}$ é uma família não vazia de conjuntos disjuntos não vazios. Pelo Axioma da Escolha existe uma função $f : I \to \bigcup_{i \in I} Y_i$ tal que $f(i) \in Y_i$ para todo $i \in I$. Podemos escrever $f(i) = (x_i,\ i)$, com $x_i \in X_i$ para todo $i \in I$. Se definimos $x(i) = x_i$ para todo $i \in I$, então $x \in \prod_{i \in I} X_i$. ■

A.2 Relações

A.1 Definição

Uma **relação** R num conjunto X é um subconjunto R de $X \times X$. Denotaremos por xRy se $(x,\ y) \in R$.

Uma relação R em X é dita:

a) reflexiva se xRx para todo $x \in X$.

b) simétrica se xRy implica yRx.

c) transitiva se xRy e yRz implicam xRz.

Diremos que R é uma relação de equivalência se R é reflexiva, simétrica e transitiva.

A.2.1 Relação de Equivalência

Usualmente escrevemos

$$a \sim b$$

ao invés de dizer que (a, b) pertence à relação $\sim$.

A.2 Definição

Uma relação $\sim$ definida em um conjunto X é uma *relação de equivalência* se satisfizer, para todo $a, b, c \in X$,

1. $a \sim a$.
2. $a \sim b \Rightarrow b \sim a$.
3. $a \sim b$ e $b \sim c \Rightarrow a \sim c$.

A.3 Definição

Dada uma relação de equivalência $\sim$ em um conjunto X, para todo $x \in X$, definimos a **classe de equivalência** de x por $\sim$ como

$$[x] = \{y \in X \ : \ y \sim x\}.$$

Observe que, pela reflexividade de $\sim$, $x \in [x]$ para todo $x \in X$.

A.4 Proposição

Seja $\sim$ uma relação de equivalência no conjunto X. Então, para $x, y \in X$, são equivalentes:

a. $y \in [x]$;
b. $[x] = [y]$;

Demonstração. Sejam $x, y \in X$. Se $y \in [x], y \sim x$, então para todo $z \in [y]$, i.e., $z \sim y$, temos, por transitividade, $z \sim x$, então $z \in [x]$. Por outro lado, pela simetria de $\sim$, temos $x \sim y$, portanto, para todo $z \in [x]$, temos, novamente por transitividade, que $z \in [y]$. Assim, $[x] = [y]$. Por outro lado, se $[x] = [y]$, então $y \in [y] = [x]$. ■

A.5 Definição

O elemento $y \in [x]$ é denominado **representante da classe**.

Exemplo 6. *Seja $A = \{a, b, c\}$, e defina $\sim$ no conjunto das partes de A $\mathscr{P}(A)$ por $X \sim Y$se, e somente se, $|X| = |Y|$. É simples mostrar que $\sim$ é uma relação de* *equivalência em $\mathscr{P}(A)$, sob a qual $\mathscr{P}(A)$ tem exatamente 4 classes de equivalência distintas:*

$$\begin{aligned}
[\emptyset] &= \{\emptyset\}, \\
[\{a\}] = [\{b\}] = [\{c\}] &= \{\{a\},\{b\},\{c\}\}, \\
[\{a,b\}] = [\{a,c\}] = [\{b,c\}] &= \{\{a,b\},\{a,c\},\{b,c\}\}, \text{ e} \\
[A] &= A.
\end{aligned}$$

◁

A.7 Definição

Seja X um conjunto. Então uma coleção de subconjuntos $P = \{X_i\}_{i\in I}$ é uma partição de X se cada $X_i \neq \emptyset$ e cada elemento de X estiver em exatamente um X_i. Em outras palavras, $P = \{X_i\}_{i\in I}$ é uma partição de X se, e somente se,:

a) $X_i \neq \emptyset$;

b) X_i são mutuamente disjuntos: $X_i \cap X_j = \emptyset$ para $i \neq j \in I$); e

c) $X = \bigcup_{i\in I} X_i$.

Os conjuntos X_i são chamados de células da partição.

Exemplo 8. *O conjunto {1,2,3 } possui 5 partições: a saber,*

$$\{\{1,2,3\}\}, \{\{1,2\},\{3\}\}, \{\{1,3\},\{2\}\}, \{\{2,3\},\{1\}\} \text{ e } \{\{1\},\{2\},\{3\}\}.$$

A primeira partição que mencionamos tem uma célula, as próximas três têm duas células e a última tem três células.

◁

A.9 Teorema

Seja X um conjunto. Então:

1) Se $\sim$ é uma relação de equivalência em X, então o conjunto de todas as classes de equivalência de X por $\sim$ é uma partição de X;

2) Se P for uma partição de X, então a relação em X definida por $x \sim y$ se e somente x estiver na mesma célula de P que y é uma relação de

equivalência em X.

Observe que, em cada caso, as células da partição são as classes de equivalência do conjunto na relação de equivalência correspondente.

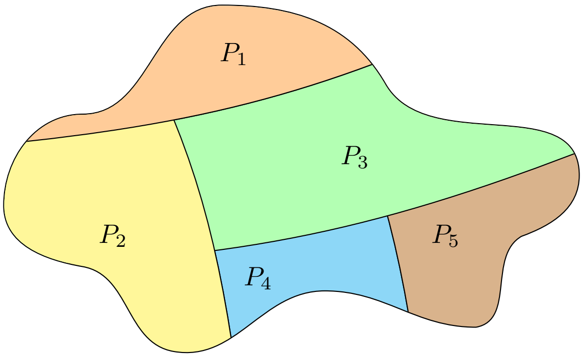


Figura A.2 Uma partição induz uma relação de equivalência e uma relação de equivalência induz uma partição.

Demonstração. Seja $\sim$ uma relação de equivalência em X. Claramente, as classes de equivalência de X por $\sim$ são conjuntos não vazios cuja união é X. Portanto, basta mostrar $Y \cap Y' = \emptyset$ para todo par de classes de equivalência $Y \neq Y'$ de X por $\sim$.

Sejam Y, Y' classes de equivalência de X por $\sim$ que não são disjuntas. Existe um elemento $z \in Y \cap Y'$. Então, $[z] = Y$ e $[z] = Y'$ e então $Y = Y'$. Portanto, se $Y \neq Y'$, $Y \cap Y' = \emptyset$.

Finalmente, a volta é direta. ■

Notação 10. *Dada uma relação de equivalência $\sim$ em X, denotamos por $X/\!\sim$ o conjunto das classes de equivalência de $\sim$. A* ***projeção natural*** *de X em $X/\!\sim$ é a aplicação dada por*

$$\begin{array}{rccc} \pi: & X & \to & X/\!\sim \\ & x & \mapsto & [x] \end{array}.$$

A.2.2 Relações de Ordem e o Lema de Zorn

A.11 Definição

Uma **relação de ordem parcial** num conjunto X é uma relação $\leq$ em X com as seguintes propriedades:

a) $x \leq x$ para todo $x \in X$($\leq$ é reflexiva);

b) se $x \leq y$ e $y \leq x$, então $x = y$($\leq$ é antissimétrica);

c) se $x \leq y$ e $y \leq z$, então $x \leq z$($\leq$ é transitiva).

Neste caso diremos que X é um **conjunto parcialmente ordenado**.

Diremos que $\leq$ é uma relação de ordem total se além de verificar a), b) e c), também verifica

d) dados $x, y \in X$, tem-se que $x \leq y$ ou $y \leq x$.

Neste caso diremos que X é um **conjunto totalmente ordenado**.

Exemplos 12.

a) *Se X é um conjunto, então a relação de inclusão é uma relação de ordem parcial em $\mathcal{P}(X)$.*

b) *A relação de ordem usual em $\mathbb{R}$ é uma relação de ordem total.*

$\triangleleft$

A.13 Definição

Seja X um conjunto parcialmente ordenado, e seja $A \subset$ X.

a) Se existir $a_0 \in A$ tal que $a_0 \leq a$ para todo $a \in A$, diremos que a_0 é 0 **elemento mínimo** de A. De maneira análoga definimos **elemento máximo** .

b) Se existir $a_0 \in A$ tal que $a = a_0$ sempre que $a \in A$ e $a \leq a_0$, diremos que a_0 é um **elemento minimal** de A. De maneira análoga definimos **elemento maximal**.

c) Se existir $c \in X$ tal que $c \leq a$ para todo $a \in A$, diremos que A é **limitado inferiormente** e que c é uma **cota inferior** de A. De maneira análoga definimos **conjunto limitado superiormente** e **cota superior**.

d) Diremos que A é uma **cadeia** em X se A é totalmente ordenado sob a relação de ordem parcial induzida por X.

e) Diremos que A é **bem ordenado** se cada subconjunto não vazio de A possui um elemento mínimo.

Exemplos 14.

a) *$\mathbb{N}$, com a ordem usual, é um conjunto bem ordenado.*

b) *$\mathbb{R}$, com a ordem usual, é um conjunto totalmente ordenado, que não é bem ordenado: o intervalo aberto (a, b) não possui elemento mínimo.*

◁

> If you are building a mathematical object in stages and find that (i) you have not finished even after infinitely many stages, and (ii) there seems to be nothing to stop you continuing to build, then Zorn's lemma may well be able to help you.
>
> — William Timothy Gowers, "How to use Zorn's lemma"

A seguir apresentaremos dois resultados equivalentes ao Axioma da Escolha.

A.15 Lema de Zorn

Seja X um conjunto parcialmente ordenado não vazio tal que cada cadeia em X é limitada superiormente. Então X possui pelo menos um elemento maximal.

Nesse texto utilizaremos o Lema de Zorn para demonstrar que todo espaço vetorial possui base.

A.16 Teorema Teorema de Zermelo

Todo conjunto não vazio pode ser bem ordenado.

A.17 Teorema

As seguintes afirmações são equivalentes:

a) O Axioma da escolha.

b) O Lema de Zorn.

c) O Teorema de Zermelo.

The Axiom of Choice is obviously true, the well-ordering principle obviously false, and who can tell about Zorn's lemma?
Jerry Bona

A.3 Cardinalidade

A.1 Definição

Dizemos que dois conjuntos A e B têm a mesma cardinalidade, denotado por $|A| = |B|$, se existe uma função bijetiva $f\colon A \to B$.

Um conjunto é dito **finito** se existe uma função bijetiva entre seus elementos e um subconjunto dos números naturais da forma $\{1, \ldots, n\}$ para algum $n \in \mathbb{N}$. Um conjunto é dito **infinito** caso contrário.

Se A é um conjunto finito denotaremos tal fato por $|A| < \infty$. Se A não for finito, escreveremos $|A| = \infty$.

Exemplo 2. *Os conjuntos* $A = \{n \in \mathbb{Z} : 0 \leq n \leq 5\}$ *e* $B = \{n \in \mathbb{Z} : -5 \leq n \leq 0\}$ *tem a mesma cardinalidade porque existe uma função bijetiva* $f : A \to B$ *dada por* $f(n) = -n$.

◁

A.3 Teorema

Existe uma bijeção $f\colon \mathbb{N} \to \mathbb{Z}$. Portanto $|\mathbb{N}| = |\mathbb{Z}|$.

O fato de $\mathbb{N}$ e $\mathbb{Z}$ terem a mesma cardinalidade nos levar a querer comparar as cardinalidades de outros conjuntos infinitos. Como, por exemplo, $\mathbb{N}$ e $\mathbb{R}$ se comparam?

De fato, $|\mathbb{N}| \neq |\mathbb{R}|$. Esse fato foi descoberto vez por Georg Cantor (1845-1918), que inventou um argumento engenhoso para mostrar que não há funções sobrejetivas $f : \mathbb{N} \to \mathbb{R}$. Isto por sua vez implica que não pode haver bijeções $f : \mathbb{N} \to \mathbb{R}$, então $|\mathbb{N}| \neq |\mathbb{R}|$

A.4 Teorema

Não existe nenhuma bijeção $f : \mathbb{N} \to \mathbb{R}$. Portanto $|\mathbb{N}| \neq |\mathbb{R}|$.

Esta é o nosso primeiro exemplo da existência de diferentes tipos de infinitos.

A.3.1 Conjuntos Enumeráveis e não Enumeráveis

A.5 Definição

Seja A um conjunto. Dizemos que A é infinito enumerável se $|\mathbb{N}| = |A|$, isto é, se existe uma bijeção $\mathbb{N} \to A$. O conjunto A é não enumerável se A for infinito e $|\mathbb{N}| \neq |A|$, ou seja, se A for infinito e não existir bijeção $\mathbb{N} \to A$.

Assim $\mathbb{Z}$ é infinito enumerável, mas $\mathbb{R}$ é não enumerável.

A relação de ter a mesma cardinalidade é uma relação de equivalência na classe de todos os conjuntos. A classe de equivalência de um conjunto A sob essa relação consiste, então, em todos aqueles conjuntos que têm a mesma cardinalidade que A.

A **cardinalidade** de um conjunto A é definida como sua classe de equivalência sob essa relação de equivalência.

A.6 Definição

A cardinalidade dos números naturais é denotada como $\aleph_0$. Ou seja, $|\mathbb{N}| = \aleph_0$. Assim, qualquer conjunto infinito enumerável tem cardinalidade $\aleph_0$.

Uma consequência simples de ser enuméravel é a seguinte:

A.7 Proposição

Um conjunto A é infinito enumerável se, e somente se, seus elementos puderem ser organizados em uma lista infinita $a_1, a_2, a_3, a_4, \ldots$

A.8 Teorema

O conjunto $\mathbb{Q}$ dos números racionais é infinito enumerável.

A.9 Corolário

Dado n conjuntos infinitos enumeráveis $A_1, A_2, A_3, \ldots, A_n$, com $n \geq 2$, o produto cartesiano $A_1 \times A_2 \times A_3 \times \cdots \times A_n$ também é infinito enumerável.

A.10 Teorema

Se A e B são ambos infinitos enumeráveis, então $A \cup B$ é infinito enumerável.

Um resultado fundamental para nós é que se trocarmos cada ponto de A por um conjunto finito de pontos a cardinalidade não é aumentada.

A.11 Lema

Sejam A e B conjuntos, e suponha $|A| = \infty$. Se, para todo $x \in A$, tivermos um conjunto finito de $I_x \subseteq B$ então $|A| \geq |\bigcup_{x \in A} I_x|$.

Capítulo

B

Matrizes e Sistemas Lineares

Nesse capítulo

- Matrizes (p. 311)
- Sistemas de Equações Lineares (p. 315)
- Operações Elementares por Linhas (p. 317)

B.1 Matrizes

Sejam m, n dois inteiros positivos. Uma **matriz** $m \times n$ A sobre $\mathbb{K}$ é uma tabela de $m \times n$ valores $a_{ij} \in \mathbb{K}$, com $1 < i < m$, $1 < j < n$ agrupados em m linha e n colunas, e será representada em uma das duas formas

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

onde a_{ij} é a entrada na interseção da i-ésima linha e da j-ésima coluna. Escreveremos $A = [a_{ij}]$ para dizer que a matriz de entradas a_{ij}. Escrevemos $A[i, \,] = \mathrm{Lin}_i(A)$ para indicar a i-ésima linha de A, e $A[\,, j] = \mathrm{Col}_j(A)$ para indicar a j-ésima coluna de A.

O conjunto das matrizes $m \times n$ com entradas em um corpo $\mathbb{K}$ é denotado por $\mathcal{M}_{m,n}(\mathbb{K})$ ou por $\mathcal{M}_{m,n}$ quando o corpo estiver subentendido. Se $A \in \mathcal{M}_{m,n}(\mathbb{K})$ é uma matriz, a entrada (i, j) de A será indicada por $A_{i,j}$ ou a_{ij}. A matriz identidade $n \times n$ será denotada por I_n. Os elementos do corpo $\mathbb{K}$ serão denominados escalares.

Os elementos em $\mathcal{M}_{m,1}(\mathbb{K})$ e em $\mathcal{M}_{1,n}(\mathbb{K})$ serão ditos vetores linhas e colunas respectivamente e serão denotados por letras em negrito $\boldsymbol{v}, \boldsymbol{a}, \ldots$.

A **diagonal principal** de uma matriz $m \times n$ A é a sequência de entradas

$$A_{1,1},\ A_{2,2},\ \ldots,\ A_{k,k}$$

onde $k = \min\{m,\ n\}$.

B.1 Definição

Diz-se que duas matrizes $A = [a_{ij}], B = [b_{ij}] \in \mathcal{M}_{m,n}(\mathbb{K})$ são iguais se $a_{ij} = b_{ij}$, para todo $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$.

Em outras palavras, duas matrizes são consideradas iguais se tiverem a mesma ordem e suas entradas correspondentes forem iguais.

B.2 Definição

Seja $A = [a_{ij}], B = [b_{ij}] \in \mathcal{M}_{n,n}(\mathbb{K})$ e $k \in \mathbb{K}$.

1. A soma de A e B, denominada $A + B$, é definida como a matriz $C = [c_{ij}] \in \mathcal{M}_{n,n}(\mathbb{K})$ com $c_{ij} = a_{ij} + b_{ij}$ para todos os i, j.
2. O produto de $k \in \mathbb{K}$ com A, denotado kA, é igual a $kA = [ka_{ij}] = [a_{ij}k] = Ak$.

B.3 Teorema

Sejam $A, B, C \in \mathcal{M}_{n,n}(\mathbb{K})$ e sejam $k, l \in \mathbb{K}$. Então

1. $A + B = B + A$ (comutatividade).
2. $(A + B) + C = A + (B + C)$ (associatividade).
3. $k(lA) = (kl)A$.
4. $(k + l)A = kA + lA$.

Demonstração. 1 Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$. Então, por definição

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = B + A$$

como elementos do corpo comutam. As outras partes são deixadas para o leitor. ■

B.4 Teorema

Seja $A \in \mathcal{M}_{m,n}(\mathbb{K})$. Então

1. a matriz $0_{m\times n}$ satisfazendo $A + 0 = 0 + A = A$ é denominada identidade aditiva.
2. a matriz B com $A+B = 0$ é denominada inversa aditiva de A, denotada $-A = (-1)A$. Nesse caso $[b_{ij}] = [-a_{ij}]$

B.5 Definição

Seja $A = [a_{ij}] \in \mathcal{M}_{n,n}(\mathbb{K})$ e $B = [b_{ij}] \in \mathcal{M}_{n,r}(\mathbb{K})$. Então, o produto de A e B, denominado AB, é uma matriz $C = [c_{ij}] \in \mathcal{M}_{m,r}(\mathbb{K})$ com

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}, \; 1 \leq i \leq m, \; 1 \leq j \leq r.$$

Ressaltamos que o produto AB é definido se, e somente se, o número de colunas de A for igual ao número de linhas de B.

B.6 Definição

A **transposta** de $A \in \mathcal{M}_{m,n}(\mathbb{K})$ é a matriz $A^{\mathrm{t}} \in \mathcal{M}_{n,m}(\mathbb{K})$ definida por

$$(A^{\mathrm{t}})_{i,j} = A_{j,i}$$

Uma matriz $A \in \mathcal{M}_{n,n}(\mathbb{K})$ é dita **simétrica** se $A = A^{\mathrm{t}}$ e **anti-simétrica se** $A^{\mathrm{t}} = -A$.

B.7 Teorema

Seja $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$. Então

a. $(A^{\mathrm{t}})^{\mathrm{t}} = A$
b. $(A+B)^{\mathrm{t}} = A^{\mathrm{t}} + B^{\mathrm{t}}$
c. $(rA)^{\mathrm{t}} = rA^{\mathrm{t}}$ para todos os $r \in \mathbb{K}$
d. $(AB)^{\mathrm{t}} = B^{\mathrm{t}}A^{\mathrm{t}}$ desde que o produto AB seja definido

Particionamento e multiplicação de matrizes Seja M uma matriz de tamanho $m \times n$. Se $B \subseteq \{1, \ldots, m\}$ e $C \subseteq \{1, \ldots, n\}$, a submatriz $M[B, C]$ é a matriz obtida de M mantendo apenas as linhas com índice em B e as colunas com índice em C. Assim, todas as outras linhas e colunas são descartadas e $M[B, C]$ tem tamanho $|B| \times |C|$.

Suponha que $M \in \mathcal{M}_{m,n}(\mathbb{K})$ e $N \in \mathcal{M}_{n,k}(\mathbb{K})$. Seja $\mathcal{P} = \{B_1, \ldots, B_p\}$ seja uma partição de $\{1, \ldots, m\}$, $\mathcal{Q} = \{C_1, \ldots, C_q\}$ uma partição de $\{1, \ldots, n\}$ e $\mathcal{R} = \{D_1, \ldots, D_r\}$ uma partição de $\{1, \ldots, k\}$

Então, é um fato útil que a multiplicação de matrizes pode ser realizada no nível do bloco e também no nível de entrada. Em particular, temos

$$[MN][B_i, D_j] = \sum_{C_h \in Q} M[B_i, C_h] N[C_h, D_j]$$

Quando as partições em questão contêm apenas blocos de elemento único, essa é precisamente a fórmula usual para multiplicação de matrizes

$$[MN]_{i,j} = \sum_{h=1}^{m} M_{i,h} N_{h,j}$$

Matrizes de bloco Será conveniente introduzir o dispositivo notacional de uma matriz de blocos. Se $B_{i,j}$ forem matrizes dos tamanhos apropriados, então por uma matriz de blocos

$$M = \begin{bmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,n} \\ \vdots & \vdots & & \vdots \\ B_{m,1} & B_{m,2} & \cdots & B_{m,n} \end{bmatrix}$$

queremos dizer a matriz cuja submatriz superior esquerda é $B_{1,1}$ e assim por diante. Portanto, os B_{i,j^1} são submatrizes de M e não entradas.

Exemplo 8.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left[\begin{array}{cc|ccc} 1 & 2 & 3 & 2 & 1 \\ 3 & 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 1 & 2 \\ \hline 1 & 2 & 1 & 1 & 2 \\ 2 & 3 & 0 & 2 & 3 \end{array}\right]$$

Nesse caso

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

$\triangleleft$

Um caso particularmente importante é escrever a matriz como blocos de suas colunas. Para esse fim introduzimos a seguinte notação:

$$A = \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array}\right] = [A[,1] \mid A[,2] \mid \cdots \mid A[,n]] = [A\boldsymbol{e}_1 \mid A\boldsymbol{e}_2 \mid \cdots \mid A\boldsymbol{e}_n]$$

sendo $\boldsymbol{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ o vetor com 0 em todas as entradas exceto a i-ésima que é 1.

Uma matriz da forma

$$M = \begin{bmatrix} B_{1,1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & B_{m,m} \end{bmatrix}$$

onde cada B_i é quadrada e 0 é uma submatriz zero, é dita um matriz diagonal por blocos.

B.2 Sistemas de Equações Lineares

B.1 Definição

Um **sistema de** m **equações lineares** em n variáveis $x_1, x_2, \dots, x_n$ é um conjunto de equações da forma

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

onde $1 \leq i \leq m$ e $1 \leq j \leq n$; $a_{ij}, b_i \in \mathbb{R}$. Um sistema linear é dita **homogêneo** se $b_1 = b_2 = \cdots = b_m = 0$ e **não-homogêneo**, caso contrário.

B.2 Definição

Sejam

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e } \boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Então, o sistema B.1 pode ser reescrito como $A\boldsymbol{x} = \boldsymbol{b}$. Neste contexto, a matriz A é dita a matriz de coeficientes e a matriz $[A \mid \boldsymbol{b}]$ é dita a **matriz aumentada** do sistema linear apresentado na Equação B.1 .

Considere o sistema linear $A\boldsymbol{x} = \boldsymbol{b}$, onde $A \in \mathcal{M}_{m,n}(\mathbb{K}), \boldsymbol{b} \in \mathcal{M}_{m,1}(\mathbb{K})$ e $\boldsymbol{x} \in \mathcal{M}_{n,1}(\mathbb{K})$. Se $[A \mid \boldsymbol{b}]$ é a matriz aumentada e $\boldsymbol{x}^T = [x_1, \ \ldots, \ x_n]$ então,

- para $j = 1, 2, \ldots, n$, a variável x_j corresponde a coluna $\mathrm{Col}_j([A \mid \boldsymbol{b}])j$.
- o vetor $\boldsymbol{b}$ a coluna $\mathrm{Col}_{n+1}([A \mid \boldsymbol{b}])$].

B.3 Definição

A **solução** de $A\boldsymbol{x} = \boldsymbol{b}$ é o vetor $\boldsymbol{x}_0$ para o qual $A\boldsymbol{x}_0$ é igual a $\boldsymbol{b}$. O conjunto de tais vetores é dito o **conjunto solução** do sistema.

Por exemplo, o conjunto solução de $A\boldsymbol{x} = \boldsymbol{b}$, com $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 4 & 2 \\ 4 & 1 & 2 \end{bmatrix}$ e $\boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

é $\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$.

B.4 Definição

Considere o sistema linear $A\boldsymbol{x} = \boldsymbol{b}$. Então, este sistema linear é dito **consistente** se admite uma solução e é dito **inconsistente** se não admite solução.

Por exemplo, o sistema homogêneo $A\boldsymbol{x} = 0$ é sempre consistente pois 0 é uma solução, o sistema $x + y = 2, 2x + 2y = 3$ é inconsistente.

B.5 Definição

Considere o sistema linear $\boldsymbol{Ax} = \boldsymbol{b}$. Então, o correspondente sistema linear $\boldsymbol{Ax} = 0$ é dito **sistema homogêneo associado**. O vetor 0 é sempre uma solução do sistema homogêneo associado.

B.6 Teorema

Considere o sistema linear homogêneo $\boldsymbol{Ax} = 0$.

1. Então, $\boldsymbol{x} = 0$, o vetor zero, é sempre uma solução, denominada solução trivial.
2. Seja $\boldsymbol{u} \neq 0$ uma solução de $\boldsymbol{Ax} = 0$. Então, $\boldsymbol{y} = c\boldsymbol{u}$ também é uma solução, para todo $c \in \mathbb{K}$. Uma solução não nula é dita a solução não--trivial. Observe que nesse caso se o sistemas admitir uma solução não-trivial (e o corpo for infinito) o sistema $\boldsymbol{Ax} = 0$ possui um número infinito de soluções.
3. Seja $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ soluções de $\boldsymbol{Ax} = 0$. Então, $\sum_{i=1}^{k} a_i \boldsymbol{u}_i$ também é a solução de $\boldsymbol{Ax} = 0$, para todo escolha de $a_i \in \mathbb{K}, 1 \leq i \leq k$.

Exemplo 7. *Seja* $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. *Então,* $\boldsymbol{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ *é a solução não-trivial de* $\boldsymbol{Ax} = 0$.

◁

Sejam $\boldsymbol{u} \neq \boldsymbol{v}$ soluções do sistema não-homogêneo $\boldsymbol{Ax} = \boldsymbol{b}$. Então, $\boldsymbol{x}_h = \boldsymbol{u} - \boldsymbol{v}$ é uma solução do sistema homogêneo associado $\boldsymbol{Ax} = 0$. Ou seja, , quaisquer duas soluções distintas de $\boldsymbol{Ax} = \boldsymbol{b}$ diferem por uma solução do sistema homogêneo associado $\boldsymbol{Ax} = 0$. Ou equivalentemente, o conjunto solução de $\boldsymbol{Ax} = \boldsymbol{b}$ é da forma , $\{\boldsymbol{x}_0 + \boldsymbol{x}_h\}$, onde $\boldsymbol{x}_0$ é uma solução particular de $\boldsymbol{Ax} = \boldsymbol{b}$ e $\boldsymbol{x}_h$ é uma solução do sistema homogêneo associado $\boldsymbol{Ax} = 0$.

B.3 Operações Elementares por Linhas

B.1 Definição

Seja $A \in \mathcal{M}_{m,n}(\mathbb{K})$. Então, as operações elementares por linha são

1. E_{ij}: Troca a i-ésima e a j-ésima linhas, ou seja, troca $A[i, :]$ e $A[j, :]$.
2. $E_k(c)$ para $c \neq 0$: Multiplica a k-ésima linha por c, ou seja, multiplica $A[k, :]$ por c.
3. $E_{ij}(c)$ para $c \neq 0$: Troca a i-ésima linha pela i-ésima linha mais c-vezes a j-ésima linha, ou seja, troca $A[i, :]$ por $A[i, :] + cA[j,$

B.2 Definição

Duas matrizes são ditas equivalentes por linha se uma pode ser obtida da outra por um número finito de operações elementares por linha.

B.3 Definição

Os sistema lineares $A\boldsymbol{x} = \boldsymbol{b}$ e $C\boldsymbol{x} = \boldsymbol{d}$ são ditos equivalentes por linha se suas respectivas matrizes aumentadas, $[A \mid \boldsymbol{b}]$ e $[C \mid \boldsymbol{d}]$, são equivalentes por linha.

B.4 Proposição

Seja $C\boldsymbol{x} = \boldsymbol{d}$ o sistema linear obtido de $A\boldsymbol{x} = \boldsymbol{b}$ aplicando uma única operação elementar por linha. Então, $A\boldsymbol{x} = \boldsymbol{b}$ e $C\boldsymbol{x} = \boldsymbol{d}$ possuem o mesmo conjunto solução .

Demonstração. Demonstraremos o resultado para a operação elementar por linha $E_{jk}(c)$ com $c \neq 0$.

Nesse caso, o sistema $A\boldsymbol{x} = \boldsymbol{b}$ e $C\boldsymbol{x} = \boldsymbol{d}$ difere somente na j-ésima equação. Devemos mostrar que $\boldsymbol{y}$ satisfaz a j-ésima equação de $A\boldsymbol{x} = \boldsymbol{b}$ se e somente se $\boldsymbol{y}$ satisfaz a j-ésima equação de $C = \boldsymbol{d}$. Seja $\boldsymbol{y}^T = [\alpha_1, \ldots, \alpha_n]$.T Então, a j-ésima e a k-ésima equações de $A\boldsymbol{x} = \boldsymbol{b}$ são $a_{j1}\alpha_1 + \cdots + a_{jn}\alpha_n = b_j$ e $a_{k1}\alpha_1 + \cdots + a_{kn}\alpha_n = b_k$. Consequentemente, temos que os α_i satisfazem

$$(a_{j1} + ca_{k1})\alpha_1 + \cdots + (a_{jn} + ca_{kn})\alpha_n = b_j + cb_k. \tag{B.2}$$

Então por definição a j-ésima equação de $C\boldsymbol{x} = \boldsymbol{d}$ é igual

$$(a_{j1} + ca_{k1})x_1 + \cdots + (a_{jn} + ca_{kn})x_n = b_j + cb_k. \tag{B.3}$$

Consequentemente, usando a Equação B.2, temos que $\boldsymbol{y}^T = [\alpha_1, \ \ldots, \ \alpha_n]$ também é uma solução para a Equação B.3. Agora, usando um argumento similar temos que se $\boldsymbol{z}^T = [\beta_1, \ \ldots, \ \beta_n]$ é a solução de $C\boldsymbol{x} = \boldsymbol{d}$ então também é a solução de $A\boldsymbol{x} = \boldsymbol{b}$. ■

Por indução temos

B.5 Teorema

Sejam $A\boldsymbol{x} = \boldsymbol{b}$ e $C\boldsymbol{x} = \boldsymbol{d}$ dois sistema lineares equivalentes por linha. Então, eles possuem o mesmo conjunto solução.

B.6 Definição

Uma matriz $E \in \mathcal{M}_{n,n}(\mathbb{K})$ é dita matriz elementar se for for obtida aplicando exatamente uma operação elementar por linha a matriz identidade I_n.

As matrizes elementares são de três tipos e correspondem as operações elementares por linha.

1. $E_{ij} = I_n - \boldsymbol{e}_i\boldsymbol{e}_i^T - \boldsymbol{e}_j\boldsymbol{e}_j^T + \boldsymbol{e}_i\boldsymbol{e}_j^T + \boldsymbol{e}_j\boldsymbol{e}_i^T$: Matriz obtida aplicando a operação elementar por linha E_{ij} a I_n.
2. $E_k(c) = I_n + (c-1)\boldsymbol{e}_k\boldsymbol{e}_k^T$ para $c \neq 0$: Matriz obtida aplicando a operação elementar por linha $E_k(c)$ a I_n.
3. $E_{ij}(c) = I_n + c\boldsymbol{e}_i\boldsymbol{e}_j^T$ para $c \neq 0$: Matriz obtida aplicando a operação elementar por linha $E_{ij}(c)$ a I_n.

Logo, quando uma matriz elementar é multiplicada a esquerda por A, obtemos o mesmo resultado que quando realizamos a operação elementar por linha correspondente em A.

B.7 Proposição

Seja A e B duas matrizes equivalentes por linha. Então, existem matrizes elementares $E_1, \ldots, E_k$ tais que $B = E_1 \cdots E_k A$.

Demonstração. Pela definição de linha equivalência, B pode ser obtida de A por um número finito de operações elementares por linha. Mas cada operação elementar por linha corresponde a multiplicação a esquerda por uma matriz elementar. ■

B.8 Definição

Seja A uma matriz não nula. Então, em cada linha não nula de A, a entrada não nula mais a esquerda é dita a entrada pivô. A coluna contendo o pivô é dita coluna pivotal.

Por exemplo, as entrada a_{12} e a_{23} são pivôs em $A = \begin{bmatrix} 0 & 3 & 4 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix}$. Logo, as colunas 2 e 3 são colunas pivotais.

B.9 Definição

Dizemos que a matriz está na forma escalonada por linha

1. se as linhas nulas estão na parte inferior da matriz;
2. se o pivô da $(i+1)$-ésima linha, se existir, está a direita do pivot da i-ésima linha.
3. se as entrada abaixo do pivô numa coluna pivotal são 0.

B.10 Definição

Dizemos que a matriz C está na forma escalonada reduzida por linhas

1. se C é está na forma escalonada ,
2. se o pivô de cada linha não nula é 1,
3. se toda outra entrada em cada coluna pivotal for zero.

Aqui estão os fatos básicos sobre a formas escalonadas reduzidas por linhas.

B.11 Teorema

As matrizes $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$ são equivalentes de linha, denotadas por $A \sim B$, se uma pode ser obtida uma da outra por uma série de operações de linha elementares .

a) A equivalência de linha é uma relação de equivalência.

b) Uma matriz A é equivalente a uma e apenas uma matriz R que está na formas escalonadas reduzidas por linhas. A matriz R é denominada formas escalonadas reduzidas por linhas de A.

$$A = E_1 \cdots E_k R$$

onde E_i são as matrizes elementares necessárias para reduzir A para a formas escalonadas reduzidas por linhas.

c) A é invertível se, e somente se, sua formas escalonadas reduzidas por linhas for uma matriz identidade. Portanto, uma matriz é invertível se, e somente se, for o produto de matrizes elementares .

Seja $A \in \mathcal{M}_{n,n}(\mathbb{K})$.

B.12 Definição

1. As entradas $a_{11}, a_{22}, \ldots, a_{nn}$ são chamadas de entradas diagonais de A. Eles constituem a diagonal principal de A.
2. A é dita uma matriz diagonal, denotado $\operatorname{diag}(a_{11},\ \ldots,\ a_{nn})$, se $a_{ij} = 0$ por $i \neq j$.
3. Então, $A = \operatorname{diag}(1,\ \ldots,\ 1)$ é denominada matriz identidade, denominada I_n ou, em suma, I.. Por exemplo, $I_2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ e $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
4. Se $A = \alpha I$, para $\alpha \in \mathbb{K}$, A é denominada matriz escalar.
5. A é dita matriz triangular superior se $a_{ij} = 0$ para $i > j$.
6. A é dita matriz triangular inferior se $a_{ij} = 0$ para $i < j$.
7. A é dita triangular se for uma matriz triangular superior ou inferior.

Exemplo 13. *Por exemplo,* $\begin{bmatrix} 0 & 1 & 4 \\ 0 & 3 & -1 \\ 0 & 0 & -2 \end{bmatrix}$ *é triangular superior,* $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ *é triangular inferior e as matrizes* $0, I$ *são matrizes triangulares superiores e inferiores.*

$\triangleleft$

Referências

APOSTOL, Tom M. **Calculus, volume II: multi-variable calculus and linear algebra, with applications to differential equations and probability**. New Jersey: John Wiley & Sons, 1969.

AXLER, Sheldon Jay. **Linear algebra done right**. New York: Springer, 1997. v. 2.

BLYTH, Thomas Scott; ROBERTSON, E F. **Further linear algebra**. New York: Springer Science & Business Media, 2013.

BROWN, William Clough; BROWN, William C. **A second course in linear algebra**. New Jersey: Wiley, 1988.

CHURCHILL, Ruel Vance et al. **Fourier series and boundary value problems**. New York: McGraw-Hill, 1941.

COELHO, Flávio Ulhoa; LOURENCO, Mary Lílian. **Curso de Álgebra Linear**. São Paulo: Edusp, 2001.

CURTIS, Morton L. **Abstract linear algebra**. New York: Springer Science & Business Media, 2012.

FARENICK, Douglas R. **Algebras of linear transformations**. New York: Springer Science & Business Media, 2012.

GOLAN, Jonathan S. **The Linear Algebra a Beginning Graduate Student Ought to Know**. New York: Springer, 2007.

GREENLEAF, Frederick P; MARQUES, Sophie. **Linear Algebra I**. Providence, Rhode Island: American Mathematical Soc., 2020. v. 30.

________. **Linear Algebra II**. Providence, Rhode Island: American Mathematical Soc., 2020. v. 30.

GREUB, Werner. **Multilinear Algebra**. New York: Springer-Verlag, 1978.

GREUB, Werner H. **Linear algebra**. New York: Springer Science & Business Media, 2012. v. 23.

HALMOS, Paul R. **Finite-dimensional vector spaces**. New York: Courier Dover Publications, 2017.

________. **Linear algebra problem book**. Washington: MAA, 1995.

HOFFMAN, Kenneth; KUNZE, Ray. **Linear Algebra**. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

HOGBEN, Leslie. **Handbook of linear algebra**. Boca Raton, Florida: CRC press, 2006.

JACOBSON, Nathan. **Lectures in Abstract Algebra: II. Linear Algebra**. New York: Springer Science & Business Media, 2013. v. 31.

KNAPP, Anthony W. **Basic algebra**. New York: Springer Science & Business Media, 2007.

KOSTRIKIN, Alexandra I. **Exercises in Algebra: A Collection of Exercises, in Algebra, Linear Algebra and Geometry**. Boca Raton, Florida: CRC Press, 1996. v. 6.

KOSTRIKIN, Alexei I; MANIN, Yu I. **Linear Algebra and Geometry. 1989**. New York, London, Paris: Gordon e Breach Science Publishers.

KREYSZIG, Erwin. **Introductory functional analysis with applications**. New York: Wiley, 1978. v. 1.

KUBRUSLY, Carlos S. **Elements of operator theory**. New York: Springer, 2011.

LAL, Ramji. **Algebra 2: Linear Algebra, Galois Theory, Representation Theory, Group Extensions and Schur Multiplier**. New York: Springer, 2017.

LIMA, Elon Lages. **Álgebra Linear**. Rio de Janeiro: [s.n.], 2006.

MACKIW, George. A note on the equality of the column, and row rank of a matrix. **Mathematics Magazine**, Mathematical Association of America, Washington, v. 68, n. 4, p. 285, 1995.

MOORE, Gregory H. The axiomatization of linear algebra: 1875-1940. **Historia Mathematica**, Elsevier, v. 22, n. 3, p. 262–303, 1995.

NAIR, M Thamban; SINGH, Arindama. **Linear Algebra**. New York: Springer, 2018.

ORTEGA, James M. **Matrix theory: A second course**. New York: Springer Science & Business Media, 2013.

PETERSEN, Peter. **Linear Algebras**. New York: Springer Science & Business Media, 2012.

ROMAN, Steven; AXLER, S; GEHRING, FW. **Advanced linear algebra**. New York: Springer, 2005. v. 3.

ROSE, Harvey E. **Linear algebra: a pure mathematical approach**. New York: Springer Science & Business Media, 2002.

SHAFAREVICH, Igor R; REMIZOV, Alexey O. **Linear algebra and geometry**. New York: Springer Science & Business Media, 2012.

SHILOV, G. **Linear Algebra**. New York: Dover Books on Advanced Mathematics, 1977.

SZYMICZEK, Kazimierz. **Bilinear algebra: An introduction to the algebraic theory of quadratic forms**. [S.l.]: CRC Press, 1997. v. 7.

WEINTRAUB, Steven H. **A guide to advanced linear algebra**. Washington: MAA, 2011.

________. **Jordan canonical form: theory and practice**. San Rafael, California: Morgan & Claypool Publishers, 2009.

WILDON, Mark. A short proof of the existence of Jordan normal form. Citeseer, 2011.

YANG, Yisong. **A concise text on advanced linear algebra**. Cambridge: Cambridge University Press, 2014.

Índice

T-aniquilador, 277
T-invariante, 119

abeliano, 15
acidental, 102
adjunta, 187
aniquilador, 233, 277
anti-simétrica se, 313
anula, 233
aplicação, 297
    composta, 298
aplicação natural, 149
associativa, 12
autoespaço, 226
autoespaço generalizado, 228
autoespaço generalizado associado ao polinômio, 239
automorfismo, 89
autovalor, 226
autovetor, 226
autovetor generalizado, 228
avaliação, 149

bandeira, 85
    maximal, 85
base, 59
base cíclica, 257
base de Hilbert, 178
base de Jordan, 253
base dual, 147
base ordenada, 79
bem ordenado, 307
bijetiva, 297

cadeia, 306
cadeia ascendente, 85
canônico, 102
cardinalidade, 309
ciclo, 17
classe, 132
classe de equivalência, 303
coeficiente de Fourier, 182
coeficientes, 45
coeficientes do polinômio, 30
cofator, 214
combinação linear, 45
commutativo, 28
complementar, 296
complemento, 73
completo, 190
complexificação, 125
complexificação da transformação, 126
complexo de cadeias, 108
    exato, 108
componentes, 80
composta, 298
comprimento, 170
comprimento da bandeira, 85
comutativo, 15
concatenação, 118
conjunto limitado superiormente, 306
conjunto parcialmente ordenado, 306
conjunto solução, 316
conjunto totalmente ordenado, 306

conjunto vazio, 296
consistente, 316
converge, 190
conúcleo, 140
coordenadas, 80
corpo, 23
cota inferior, 306
cota superior, 306
covetores, 144
cíclica, 256
cíclico, 256

descomplexificação, 123
determinante, 216
diagonal principal, 311
diagonalizável, 243
diferenca, 296
dimensão, 64
distância, 175
divisor de zero, 28
divisores de zero, 25
divisores elementares, 283
domínio de ideais principais, 30
domínio de integridade, 29
díade, 151

elemento maximal, 306
elemento minimal, 306
elemento máximo, 306
elemento mínimo, 306
elemento neutro, 12
elemento neutro , 14
endomorfismo, 89
epimorfismo, 89
equivalentes, 221
escalares, 45
espaço bidual, 149
espaço de Hilbert, 190
espaço dual, 144
espaço métrico, 176
espaço quociente, 132
espaços
  independentes, 71
espectro, 226
estabiliza a bandeira, 235
estabiliza estritamente a bandeira, 235
estrutura algébrica, 13
estrutura complexa, 124
exato, 108

finito, 308
funcional, 144
funcional linear, 88, 144
função, 297
  composta, 298
  inversa, 298

grau, 30
grupo, 14
grupo de permutações, 16

hiperplano, 148
homogêneo, 315
homomorfismo, 13, 88

ideal, 29
ideal gerado, 29
identidade, 28, 89
identidade , 14
igual, 296
imagem, 93, 298
imagem estável, 255
imagem inversa, 298
inconsistente, 316
independente, 71
infinito, 308
injetiva, 297
inteiros, 29
intersecção, 296, 299
invariante, 119
inverso, 13
isomorfismo, 89, 101

limitado inferiormente, 306
linearmente dependente, 58
linearmente independente, 58

matriz, 311
matriz adjunta, 214
matriz aumentada, 316
matriz canônica racional por divisores elementares, 284

matriz companheira, 279
matriz da transformação linear, 113
matriz do sistema, 53
matriz dos cofatores, 214
matriz mudança de base, 81
maximal, 85
monomorfismo, 89
multiplicação, 23
multiplicidade algébrica, 231
multiplicidade geométrica, 231
máximo divisor comum, 33
métrica, 176
mônico, 30

natural, 102
nilpotente, 98, 254
norma, 170, 171
nulidade, 93
não-homogêneo, 315
núcleo, 92
núcleo estável, 255

operador linear, 88
ortogonais, 169
ortogonal, 176
ortonormal, 176

permutações pares, 19
permutações ímpares., 19
perpendiculares, 169
polinômio, 30
polinômio característico, 229, 230
polinômio minimal, 234
posto, 83, 93
produto
    cartesino, 300
produto cartesiano, 297, 300
produto direto, 76
produto interno, 165
projeção canônica, 134
projeção natural, 134, 305

raiz de multiplicidade, 35
raiz múltipla, 35
raiz simples, 35
real, 271
relação, 302
relação de ordem parcial, 305
representante, 132
representante da classe, 303

semelhantes, 223
separável, 193
sequência de Cauchy, 190
sequência exata curta, 109
similares, 223
simétrica, 313
sistema
    equações lineares, 315
sistema homogêneo associado, 317
sobrejetiva, 297
solução, 316
soma, 69
soma direta externa, 77
soma direta ortogonal, 184
subanel, 29
subconjunto, 296
subcorpo, 25
subespaço T-cíclico, 276
subespaço afim, 128
subespaço cíclico de T, 279
subespaço gerado, 55
subespaço ortogonal, 176
subespaço vetorial, 51
subgrupo, 16
subgrupo normal, 16

tabela de Cayley, 10
transformação
    nilpotente, 98
transformação afim , 131
transformação linear, 88
transformação nula, 89
transposta, 313
transposta da transformação, 157
triangularizável estritamente
    superior, 234
triangularizável superior, 234

unipotente, 254
unitário, 170
união, 296, 299

vetor

cíclico, 279
vetor coluna, 48
vetor cíclico, 256
vetores, 45
vetores linhas, 47
zero, 14, 28

ângulo, 169
índice, 228, 239
índice de nilpotência, 256

Espaços Vetoriais

$\mathcal{R}(B)$, funções integráveis, 49
$\overline{V}$, 49
$C(I)$, funções contínuas em I, 48
V^A, espaço de funções vetoriais, 47
$\mathbb{K}[x]$, polinômios em uma variável, 48
$\mathbb{K}$, corpo, 46
$\mathbb{K}^S$, espaço de funções, 46
$\mathbb{K}^n$, espaço de coordenadas, 46
$\mathcal{M}_{m,n}(\mathbb{K})$, espaço das matrizes $m \times n$, 47
$\mathrm{Seq}(\mathbb{K})$, sequências, 48

solução de um sistema de equações lineares homogêneo, 53
soluções de EDO's, 53