

# Álgebra Linear: Principais Ideias

Adair Antonio da Silva Neto

1 de setembro de 2022

# Sumário

<b>1</b>	<b>Estruturas Algébricas</b>	<b>1</b>
1.1	O que é um corpo (ou field)?	1
1.2	O que é um corpo ordenado?	2
1.3	Números complexos	2
1.4	Como somar números complexos?	3
1.5	Como multiplicar números complexos?	3
1.6	Representação geométrica de $\mathbb{C}$	3
1.7	Complexo conjugado	3
1.8	O que é uma métrica?	3
1.9	Função Traço (Trace)	4
1.10	Posto de uma Matriz (Matrix Rank)	4
1.11	Matrizes adjuntas	4
1.12	Material Complementar	5
<b>2</b>	<b>Espaço Vetorial</b>	<b>6</b>
2.1	O que é um espaço vetorial?	6
2.2	O que é combinação linear?	7
2.3	O que são subespaços?	7
2.4	O que significa subespaço gerado (subspace spanned)?	7
2.5	O que são base e dimensão?	8
2.6	O que são coordenadas?	8
2.7	Como fazer matriz de mudança de base?	9
2.8	Material Complementar	10
<b>3</b>	<b>Transformações Lineares</b>	<b>11</b>
3.1	O que é uma transformação linear?	11
3.2	O que são Núcleo e Imagem?	12
3.3	O que são Posto e Nulidade?	12
3.4	Teorema do Núcleo e Imagem.	12
3.5	A Álgebra das Transformações Lineares	13
3.6	O que são isomorfismos?	14
3.7	Como representar transformações com matrizes?	15
3.8	Exercícios Resolvidos	18
<b>4</b>	<b>Produto Interno</b>	<b>19</b>
4.1	O que é?	19
4.2	Desigualdade de Cauchy-Schwarz	20
4.3	Norma e Métrica	21
4.4	Ângulo e Ortogonalidade	22

4.5	Processo de Gram-Schmidt	22
4.6	Complemento Ortogonal	23
4.7	Projeção Ortogonal	24
4.8	Operadores auto-adjuntos	24
4.9	Operadores ortogonais	25
<b>5</b>	<b>Autovalores e Autovetores</b>	<b>26</b>
5.1	Motivação	26
5.2	O que são?	27
5.3	Como encontrá-los?	27
5.4	Multiplicidades algébrica e geométrica	28
5.5	Diagonalização	29
5.6	Subespaço Invariante	30
5.7	Matrizes Especiais	31
5.8	Teorema Espectral	33

Aviso: este material está em construção. Ele é escrito por **Adair Antonio da Silva Neto**, aluno do bacharelado em Matemática da Unicamp, para estudo próprio, mas na expectativa de que ajude outras pessoas em seu aprendizado.

## 1 Estruturas Algébricas

Nesta seção veremos os conceitos de corpo, corpo ordenado, números complexos e métrica.

### 1.1 O que é um corpo (ou field)?

Antes de definir o que é um corpo, vamos considerar um exemplo de um corpo.

Seja  $\mathbb{F}$  o conjunto dos números reais ou complexos. Sabemos que:

1. A adição é comutativa:  $x + y = y + x, \forall x, y \in \mathbb{F}$ .
2. A adição é associativa  $x + (y + z) = (x + y) + z, \forall x, y, z \in \mathbb{F}$ .
3. Existe um único elemento  $0 \in \mathbb{F}$  tal que  $x + 0 = x$ , para todo  $x$  em  $\mathbb{F}$  (elemento neutro).
4. Para cada  $x \in \mathbb{F}$ , existe um único elemento correspondente  $(-x) \in \mathbb{F}$  tal que  $x + (-x) = 0$  (elemento simétrico).
5. A multiplicação é comutativa  $xy = yx, \forall x, y \in \mathbb{F}$ .
6. A multiplicação é associativa  $x(yz) = (xy)z, \forall x, y, z \in \mathbb{F}$ .
7. Existe um único elemento  $1 \in \mathbb{F}$  tal que  $x1 = x, \forall x \in \mathbb{F}$  (elemento neutro).
8. Para cada elemento  $x \in \mathbb{F}$  diferente de zero, existe um único elemento correspondente  $x^{-1} = (1/x)$  tal que  $xx^{-1} = 1$  (inverso multiplicativo).
9. Vale a distributividade:  $x(y + z) = xy + xz, \forall x, y, z \in \mathbb{F}$ .

A motivação por trás da definição de corpo é generalizar essas propriedades para outros conjuntos e outras operações. Ou seja, queremos trabalhar com conjuntos, munidos de duas operações que "funcionam" como a adição e a multiplicação dos reais.

Assim, dizemos que um corpo  $\mathbb{F}$  é um conjunto não vazio dotado de duas operações  $+$  e  $\times$  satisfazendo as propriedades 1-9.

Há ainda uma restrição que precisamos fazer. Queremos que toda operação de adição ou de multiplicação nos leve a um elemento dentro do corpo. Isto é, queremos que o corpo seja fechado quanto à adição e multiplicação. Formalmente,

(F1) Para todos  $x, y \in \mathbb{F}$ ,  $x + y \in \mathbb{F}$ . (F2) para todos  $x, y \in \mathbb{F}$ ,  $x \times y \in \mathbb{F}$ .

Esses são chamados axiomas de fechamento. As propriedades 1-4 são chamadas axiomas da operação de adição e as propriedades 5-9, axiomas da operação de multiplicação.

## 1.2 O que é um corpo ordenado?

Um corpo  $\mathbb{F}$  munido de uma relação de ordem  $<$  é chamado de corpo ordenado se ele satisfaz os seguintes axiomas:

1. (O1) Princípio da Comparação: Queremos que apenas um dos seguintes casos aconteça: ou  $x < y$ , ou  $y < x$  ou  $x = y$ .
2. (O2) Transitividade:  $x < y \wedge y < z \implies x < z$ .
3. (O3) Consistência da adição:  $y < z \implies x + y < x + z$ .
4. (O4) Consistência da multiplicação:  $(0 < x \wedge 0 < y) \implies 0 < x \times y$ .

Onde  $x, y, z \in \mathbb{F}$ .

## 1.3 Números complexos

Antes de tudo, como podemos definir o conjunto dos números complexos  $\mathbb{C}$ ?

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$$

## 1.4 Como somar números complexos?

Seja  $z = a + bi$  e  $w = c + di \in \mathbb{C}$ . Definimos a sua soma como

$$z + w = (a + bi) + (c + di) = (a + c) + (b + d)i$$

## 1.5 Como multiplicar números complexos?

E definimos seu produto como

$$z \cdot w = (a + bi) \cdot (c + di) = (ac - bd) + (bc + ad)i$$

Note que  $i^2 = i \cdot i = (0 + 1i) \cdot (0 + 1i) = (0 - 1) + (0 + 0)i = -1$ .

O número  $i$  é chamado imaginário puro.

## 1.6 Representação geométrica de $\mathbb{C}$

Dado  $\mathbb{R}^2$  o plano cartesiano usual, identificamos o número complexo  $z = a + bi$  com o ponto  $(a, b) \in \mathbb{R}^2$ .

Com essa representação geométrica em mente, vem do Teorema de Pitágoras que o valor absoluto (ou módulo) de um número complexo  $z$  é

$$|z| = \sqrt{a^2 + b^2}$$

Em coordenadas polares, temos  $a = r \cos \theta$  e  $b = r \sin \theta$ .

## 1.7 Complexo conjugado

Geometricamente, o conjugado complexo de  $z$  é a reflexão do ponto  $z$  em relação ao eixo real  $Ox$ . Simbolicamente,

$$\bar{z} := a - bi$$

## 1.8 O que é uma métrica?

Como podemos generalizar a noção que temos de distância?

Intuitivamente uma distância exige um ponto inicial e um ponto final, resultando, a partir desses pontos, um número real.

Sabemos que toda distância é simétrica (isto é, têm o mesmo valor independentemente do sentido) e sempre positiva. Além disso, queremos que, dados três pontos  $x, y, z$ , a distância de  $x$  até  $z$  seja menor ou igual a distância de  $x$  até  $y$  mais a distância de  $y$  até  $z$ .

Assim, dado um conjunto não vazio qualquer  $\mathbb{X}$ , podemos definir uma métrica (ou distância) como sendo uma função  $d(\cdot, \cdot) : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$  satisfazendo:

1. Simetria:  $d(x, y) = d(y, x)$ . 2. Positividade:  $d(x, y) \geq 0$ , sendo que  $d(x, y) = 0 \iff x = y$ . 3. Desigualdade Triangular:  $d(x, z) \leq d(x, y) + d(y, z)$ .

Onde  $x, y, z$  são elementos quaisquer do conjunto  $\mathbb{X}$ .

Vamos utilizar a notação  $(\mathbb{X}, d)$  para indicar que o conjunto  $\mathbb{X}$  possui a métrica  $d(\cdot, \cdot)$ . Dizemos que  $(\mathbb{X}, d)$  é um **espaço métrico**.

## 1.9 Função Traço (Trace)

É a soma dos elementos da diagonal principal de uma matriz quadrada. Isto é,

$$\text{tr} A = \sum_{i=1}^n a_{ii}$$

onde  $A = (a_{ij})_{i,j} \in \mathbb{M}_n(\mathbb{K})$ .

## 1.10 Posto de uma Matriz (Matrix Rank)

É o número de linhas não nulas de uma matriz em sua forma escalonada.

## 1.11 Matrizes adjuntas

É a transposta da matriz (quadrada!) dos cofatores.

## **1.12 Material Complementar**

O vídeo [Vetores, o que são eles afinal?](#) é uma ótima introdução ao que vamos estudar nesta disciplina.



## 2 Espaço Vetorial

Vamos introduzir aqui o objeto matemático que será o centro de nosso estudo em Álgebra Linear. A ideia é trabalhar com um sistema algébrico que generalize a noção de combinação linear num dado conjunto.

### 2.1 O que é um espaço vetorial?

Um **espaço vetorial** consiste de:

1. Um conjunto  $\mathbb{F}$  de escalares.
2. Um conjunto  $\mathbb{V}$  de objetos chamados vetores.
3. Uma regra (i.e. operação) chamada **adição de vetores**, que associa cada par de vetores  $\alpha, \beta \in \mathbb{V}$  um vetor  $\alpha + \beta \in \mathbb{V}$  tal que as seguintes propriedades são satisfeitas:
  - (a) **Comutatividade:**  $\alpha + \beta = \beta + \alpha$ .
  - (b) **Associatividade:**  $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ .
  - (c) **Elemento neutro:** Existe um único vetor  $0 \in \mathbb{V}$  tal que  $\alpha + 0 = \alpha$ ,  $\forall \alpha \in \mathbb{V}$ .
  - (d) **Elemento simétrico:** Para cada vetor  $\alpha$ , existe um único vetor  $-\alpha$  tal que  $\alpha + (-\alpha) = 0$ .
4. Uma regra chamada **multiplicação por escalar** que associa o par  $c \in \mathbb{F}$  e  $\alpha \in \mathbb{V}$  ao vetor  $c\alpha \in \mathbb{V}$ , satisfazendo:
  - (a) **Elemento identidade:**  $1\alpha = \alpha$ ,  $\forall \alpha \in \mathbb{V}$ .
  - (b) **Associatividade:**  $(c_1c_2)\alpha = c_1(c_2\alpha)$ .
  - (c) **Distributividade para adição de vetores:**  $c(\alpha + \beta) = c\alpha + c\beta$ .
  - (d) **Distributividade para a multiplicação por escalar:**  $(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha$ .

Os objetos que chamamos de vetores num espaço vetorial podem não ser os vetores aos quais estamos acostumados do Ensino Médio. Podem ser matrizes, funções, polinômios etc.

## 2.2 O que é combinação linear?

Falamos em combinação linear anteriormente remetendo à noção intuitiva que temos de Geometria Analítica. Formalmente, dizemos que um vetor  $\beta$  é dito **combinação linear** dos vetores  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  se existirem escalares  $c_1, \dots, c_n$  tais que  $\beta$  pode ser escrito como

$$\beta = c_1\alpha_1 + \dots + c_n\alpha_n$$

Ou, em notação mais compacta,

$$\beta = \sum_{i=1}^n c_i\alpha_i$$

## 2.3 O que são subespaços?

Considere  $V$  um espaço vetorial sobre o corpo  $F$ . Um subconjunto  $W \subset V$  é dito um **subespaço** de  $V$  se  $W$  for um espaço vetorial sobre  $F$  com as operações de soma de vetores e multiplicação por escalar.

### Exemplo 2.1: Subespaços

Como verificar se  $S_A$  é subespaço vetorial de  $S$ ?

**Solução:** Basta verificar se  $c\alpha + d\beta \in S_A$ , onde  $c, d \in \mathbb{F}$  e  $\alpha, \beta \in \mathbb{V}$ .

## 2.4 O que significa subespaço gerado (subspace spanned)?

Dado um conjunto de vetores  $S$  em um espaço vetorial  $V$ , dizemos que o **subespaço gerado** por  $S$  é a interseção de todos os subespaços de  $V$  que contêm  $S$ .

Caso  $S$  seja um conjunto não nulo, então o subespaço gerado por ele (vamos chamá-lo de  $W$ ) é o conjunto de todas as combinações lineares de vetores em  $S$ . Isto é, a partir dos vetores de  $S$  eu consigo escrever qualquer vetor em  $W$ .

## 2.5 O que são base e dimensão?

Uma **base** para o espaço vetorial  $V$  é um conjunto de vetores linearmente independentes que geram o espaço (\*span the space\*)  $V$ . Vamos dizer que o espaço  $V$  tem **dimensão finita** se ele possui uma base finita.

Por isso, os vetores  $e_1 = (1, 0, 0, \dots, 0)$ ,  $e_2 = (0, 1, 0, \dots, 0)$ ,  $\dots$ ,  $e_n = (0, 0, 0, \dots, 1)$  formam uma base do espaço  $\mathbb{R}^n$  e são chamados de **base canônica** de  $\mathbb{R}^n$ .

Observe que a definição de base implica que se  $V$  é um espaço vetorial gerado por um conjunto de  $m$  vetores, então qualquer conjunto de vetores linearmente independentes em  $V$  é finito e contém no máximo  $m$  elementos. Consequentemente, quaisquer duas bases de  $V$  têm o mesmo número  $m$  de elementos.

Com isso em mente, vamos definir **dimensão** de um espaço vetorial finito  $V$  como sendo o número de elementos da base de  $V$  e denotamos  $\dim(V)$ .

Essas definições têm como consequência dois fatos importantes:

1. Qualquer subconjunto de  $V$  com mais de  $\dim(V)$  elementos é linearmente dependente. 2. Nenhum subconjunto de  $V$  que tem menos que  $\dim(V)$  elementos pode gerar  $V$ .

Caso  $W_1$  e  $W_2$  sejam dois subespaços finitos de  $V$ , então  $W_1 + W_2$  tem dimensão finita e

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

## 2.6 O que são coordenadas?

As coordenadas de um vetor  $v \in V$ , relativas à base  $\beta$ , são os coeficientes que permitem expressar  $v$  como uma combinação linear dos vetores de  $\beta$ . As coordenadas mais naturais são aquelas que utilizam a base canônica do corpo  $\mathbb{F}$  no qual estamos trabalhando. Ou seja,

$$v = (x_1, \dots, x_n) = \sum x_i e_i$$

onde  $e_i$  é o  $i$ -ésimo elemento da base canônica.

Para trabalhar com mudança de coordenadas, precisamos antes entender o que é uma **base ordenada** de um espaço de dimensão finita  $V$ .

Dizemos que  $\beta$  é uma base ordenada de  $V$  se  $\beta$  é uma sequência finita de vetores que é linearmente independente e gera  $V$ .

Note que uma base é um conjunto, um objeto no qual a ordem não faz diferença. Porém, uma base ordenada é uma sequência, o que nos permite distinguir quem é seu  $i$ -ésimo elemento. Assim, se  $a_1, \dots, a_n$  é uma base ordenada de  $V$ , então  $\{a_1, \dots, a_n\}$  é uma base de  $V$ .

Vamos denotar  $[v]_\beta$  as coordenadas do vetor  $v$  em relação à base ordenada  $\beta$ . Além disso, vamos fazer um abuso de notação utilizando  $\beta = \{a_1, \dots, a_n\}$  para denotar uma base ordenada.

## 2.7 Como fazer matriz de mudança de base?

Vamos tomar  $\beta = \{\beta_1, \dots, \beta_n\}$  e  $\gamma = \{\gamma_1, \dots, \gamma_n\}$  duas bases ordenadas de um espaço finito  $V$ .

Note que podemos escrever cada vetor da base  $\gamma$  como combinação linear dos vetores de  $\beta$  da seguinte forma:

$$\begin{aligned}\gamma_1 &= a_{11} \cdot \beta_1 + a_{21} \cdot \beta_2 + \dots + a_{n1} \cdot \beta_n \\ \gamma_2 &= a_{12} \cdot \beta_1 + a_{22} \cdot \beta_2 + \dots + a_{n2} \cdot \beta_n \\ &\vdots \\ \gamma_n &= a_{1n} \cdot \beta_1 + a_{2n} \cdot \beta_2 + \dots + a_{nn} \cdot \beta_n\end{aligned}$$

onde cada  $a_{ij}$  é um escalar.

Assim, para cada  $i \in \{1, 2, \dots, n\}$ , as coordenadas do vetor  $\gamma_i$  na base  $\beta$  são dadas por

$$[\gamma_i]_\beta = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix}$$

Dessa maneira conseguimos obter as coordenadas de cada vetor da base  $\gamma$  em relação à base  $\beta$ . Com isso, montamos a **matriz de mudança de base** de  $\beta$  para  $\gamma$ :

$$P_{\beta \rightarrow \gamma} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

Note que cada coluna é formada pelas coordenadas de  $\beta_1, \dots, \beta_n$  em relação à base  $\beta$ .

### Exemplo 2.2: Mudança de base

Considere  $\beta$  a base canônica de  $\mathbb{R}^3$  e  $\gamma = \{(1, 0, 1), (1, 1, 1), (1, 1, 2)\}$ . Encontre a matriz de mudança de base  $P_{\gamma \rightarrow \beta}$ .

#### Solução:

O primeiro passo é escrever cada vetor de  $\beta$  como combinação dos vetores de  $\gamma$ . Isto é,

$$\begin{aligned}(1, 0, 0) &= a_{11} \cdot (1, 0, 1) + a_{21} \cdot (1, 1, 1) + a_{31} \cdot (1, 1, 2) \\ &= 1 \cdot (1, 0, 1) + 1 \cdot (1, 1, 1) - 1 \cdot (1, 1, 2)\end{aligned}$$

$$\begin{aligned}(0, 1, 0) &= a_{12} \cdot (1, 0, 1) + a_{22} \cdot (1, 1, 1) + a_{32} \cdot (1, 1, 2) \\ &= -1 \cdot (1, 0, 1) + 1 \cdot (1, 1, 1) + 0 \cdot (1, 1, 2)\end{aligned}$$

$$\begin{aligned}(0, 0, 1) &= a_{13} \cdot (1, 0, 1) + a_{23} \cdot (1, 1, 1) + a_{33} \cdot (1, 1, 2) \\ &= 0 \cdot (1, 0, 1) - 1 \cdot (1, 1, 1) + 1 \cdot (1, 1, 2)\end{aligned}$$

E com esses valores montamos a matriz mudança de base:

$$P_{\gamma \rightarrow \beta} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}$$

## 2.8 Material Complementar

Combinações lineares, subespaços gerados, e bases

### 3 Transformações Lineares

Nesta seção veremos Teorema do Núcleo e Imagem, Posto e Nulidade, Isomorfismo, Transformação Inversa, Matriz de Representação.

#### 3.1 O que é uma transformação linear?

É um mapa que leva um vetor do espaço vetorial  $V$  em um elemento de do espaço  $W$ . Ou seja, uma **transformação linear** de  $V$  em  $W$  é uma função

$$T(cv + w) = cT(v) + T(w)$$

onde  $v, w \in V$  e  $c$  é um escalar em  $\mathbb{F}$ .

##### Exemplo 3.1: Linearidade

Dado o corpo  $\mathbb{R}$  e  $V$  o espaço formado pelas funções de  $\mathbb{R} \rightarrow \mathbb{R}$  contínuas, podemos definir uma transformação linear

$$T(f(x)) = \int_0^x f(t)dt$$

Lembre-se que a linearidade da integração é uma de suas principais propriedades e aparece de maneira análoga nas transformações lineares.

É importante notar que para toda transformação linear  $T$  é verdade que  $T(0) = 0$ . Isto é,  $T$  sempre passa pela origem.

Outro fato importante, que vem direto da definição, é que transformações lineares preservam combinações lineares. Ou seja,

$$T(c_1v_1 + \dots + c_nv_n) = c_1T(v_1) + \dots + c_nT(v_n)$$

Muitas vezes encontramos o problema inverso de ter que encontrar a transformação  $T$  a partir de suas aplicações em vetores  $e_1, \dots, e_n$ . Um resultado útil nesses casos é o seguinte:

Dada uma base ordenada para  $V$  então existe uma única transformação linear de  $V$  em  $W$  que leva cada vetor da base ordenada de  $V$  em um vetor de  $W$ .

### 3.2 O que são Núcleo e Imagem?

Considere uma transformação linear  $T : V \rightarrow W$ . Temos que a **imagem** de  $T$  é um subespaço de  $W$  chamado de imagem de  $T$  e denotamos  $Im(T)$ . Pois bem, quais são os elementos de  $Im(T)$ ? São todos os vetores em  $W$  tais que algum vetor  $v \in V$  é levado até ele pela transformação  $T$ . Isto é,

$$Im(T) = \{w \in W : T(v) = w, v \in V\}$$

Outro subespaço bastante útil associado à  $T$  é aquele que contém todos os vetores  $v \in V$  que levam ao vetor nulo. Ou seja, todos os  $v$  tais que  $T(v) = 0$ . Esse subespaço é chamado **núcleo** (ou **kernel**) de  $T$  e é definido por

$$Ker(T) = \{v \in V : T(v) = 0\}$$

### 3.3 O que são Posto e Nulidade?

Dado um espaço de dimensão finita  $V$ , dizemos que o **posto** (ou **rank**) de uma transformação  $T$  é a dimensão da imagem de  $T$ . Também dizemos que a **nulidade** (ou **nullity**) de  $T$  é a dimensão do núcleo de  $T$ .

Essas duas definições são relacionadas pelo seguinte

### 3.4 Teorema do Núcleo e Imagem.

#### Teorema 3.1: Núcleo e Imagem

Se  $V$  e  $W$  são dois espaços vetoriais sobre o corpo  $\mathbb{F}$  e  $T$  é uma transformação linear de  $V$  em  $W$ . Supondo que  $V$  tem dimensão finita, então

$$\text{posto}(T) + \text{nulidade}(T) = \dim(V)$$

Em outras palavras,

$$\dim(V) = \dim(Ker(T)) + \dim(Im(T))$$

Intuitivamente, o **posto** de uma matriz é o número de linhas ou colunas linearmente independentes da matriz. É um resultado importante

da Álgebra Linear que o posto das linhas de uma matriz é igual ao posto de suas colunas.

### 3.5 A Álgebra das Transformações Lineares

É importante e bonito notar que o conjunto das transformações lineares herda uma estrutura natural do espaço vetorial. Isso ficará claro com a generalização que faremos agora.

Considerando novamente  $V, W$  espaços vetoriais sobre um corpo  $\mathbb{F}$  e  $T$  e  $U$  transformações lineares. Temos que

$$(T + U)(v) = T(v) + U(v) \quad (cT)(v) = cT(v)$$

são ambas transformações lineares.

Com isso, vem que o conjunto de todas as transformações lineares de  $V$  em  $W$ , com as operações de adição e multiplicação por escalar conforme definidos acima, são um espaço vetorial sobre  $\mathbb{F}$ , que denotaremos  $L(V, W)$ .

Um importante resultado sobre esse espaço é que se  $V$  tem dimensão  $n$  e  $W$  tem dimensão  $m$ , então a dimensão de  $L(V, W) = mn$ .

Além disso, se definirmos  $T : V \rightarrow W$  e  $U : W \rightarrow Z$ , ambas sobre o mesmo corpo  $\mathbb{F}$ , então a composição  $U(T(v))$  é uma transformação linear de  $V$  em  $Z$ .

Por simplicidade, vamos definir um **operador linear** como sendo uma transformação linear de um espaço vetorial  $V$  sobre ele mesmo. Isto é, de  $V$  em  $V$ .

Para os operadores  $U, T_1, T_2$  valem as seguintes propriedades:

1.  $IU = UI = U$ ;
2.  $U(T_1 + T_2) = UT_1 + UT_2$  e  $(T_1 + T_2)U = T_1U + T_2U$ ;
3.  $\alpha(UT_1) = (\alpha U)T_1 = U(\alpha T_1)$ , onde  $\alpha \in \mathbb{F}$ .

Uma transformação  $T : V \rightarrow W$  é dita **invertível** se existe  $U : W \rightarrow V$  tal que  $UT$  é a identidade em  $V$  e  $TU$  é a identidade em  $W$ . Se  $T$  é invertível, então a função  $U$  é única e denotada por  $T^{-1}$ .

$T$  é invertível se, e somente se,



1.  $T$  é injetora. Ou seja  $T(v) = T(w) \implies v = w$ .
2.  $T$  é sobrejetora. Isto é,  $Im(T) = W$ .

Uma propriedade importante, análoga a matrizes, é que  $(UT)^{-1} = T^{-1}U^{-1}$ .

Note que se pela linearidade de  $T$  vem que  $T(v - w) = T(v) - T(w)$ . Portanto,  $T(v) = T(w)$  se, e somente se,  $T(v - w) = 0$ . Esse resultado é bastante útil para verificar se uma transformação é injetora.

Dizemos que  $T$  é **não singular** se  $T(v) = 0$  implica que  $v = 0$ . Ou seja, o núcleo de  $T$  é  $\{0\}$ . Assim,  $T$  é injetora se, e somente se,  $T$  é não singular.

Uma extensão do fato acima é que a não singularidade de  $T$  é equivalente a dizer que  $T$  leva cada subconjunto linearmente independente de  $V$  em um subconjunto linearmente independente de  $W$ .

Caso  $\dim(V) = \dim(W)$ , então o Teorema do Núcleo e Imagem nos garante que as seguintes afirmações são equivalentes:

1.  $T$  é invertível.
2.  $T$  é não singular (ou seja, é injetora).
3.  $T$  é sobrejetora.

Assim, caso  $T$  seja um operador linear, é suficiente verificar que  $T$  é injetora.

Observação: o conjunto dos operadores lineares em um espaço  $V$ , munido da operação de composição, é um **grupo**, conceito importante da Álgebra. Além disso, o conjunto de vetores em um espaço vetorial com a operação de soma de vetores é um **grupo comutativo**.

### 3.6 O que são isomorfismos?

Dizemos que uma transformação linear  $T : V \rightarrow W$  bijetora é um **isomorfismo de  $V$  em  $W$** . E se existe um isomorfismo de  $V$  em  $W$ , dizemos que  $V$  é **isomorfo** a  $W$ .

Observe que:

1.  $V$  é trivialmente isomorfo a  $V$  (reflexividade).

2. Se  $V$  é isomorfo a  $W$  através do isomorfismo  $T$ , então  $W$  é isomorfo a  $V$  via  $T^{-1}$  (simetria).
3. Se  $V$  é isomorfo a  $W$  e  $W$  é isomorfo a  $Z$ , então  $V$  é isomorfo a  $Z$  (transitividade).

Ou seja, o isomorfismo é uma relação de equivalência.

Um importante resultado sobre isomorfismos é que todo espaço vetorial de dimensão  $n$  sobre o corpo  $\mathbb{F}$  é isomorfo ao espaço  $\mathbb{F}^n$ . Ou seja, dois espaços vetoriais finitos, definidos sobre um mesmo corpo, são isomórficos se, e somente se, eles possuem a mesma dimensão.

Em certo sentido, espaços vetoriais isomorfos são o "mesmo". Um exemplo disso é que isomorfismos preservam a dimensão, isto é, qualquer subespaço finito de  $V$  tem a mesma dimensão que sua imagem sobre o isomorfismo  $T$ .

### 3.7 Como representar transformações com matrizes?

Para facilitar a notação, definimos  $V$  um espaço vetorial de dimensão  $n$  e  $W$  um espaço vetorial de dimensão  $m$ , ambos sobre o corpo  $\mathbb{F}$ . Definimos também  $\beta = \{v_1, \dots, v_n\}$  uma base ordenada para  $V$  e  $\gamma = \{w_1, \dots, w_m\}$  uma base ordenada de  $W$ .

Se  $T$  é uma transformação linear de  $V$  em  $W$ , então  $T$  é definido de maneira única pela sua ação nos vetores  $v_j$  de  $\beta$ . Isto é, cada um dos  $n$  vetores  $T(v_j)$  pode ser escrito de maneira única como uma combinação linear dos vetores de  $w_i$  de  $\gamma$ :

$$T(v_j) = c_{1j}w_1 + c_{2j}w_2 + \dots + c_{mj}w_m = \sum_{i=1}^m c_{ij}w_i$$

onde os escalares  $c_{1j}, \dots, c_{mj}$  são as coordenadas de  $T(v_j)$  na base  $\gamma$ . Dessa forma, determinamos  $T$  a partir de  $mn$  escalares  $c_{ij}$ .

Com isso, podemos definir uma matriz  $m \times n$  a partir de  $A(i, j) = c_{ij}$ . Dizemos que  $A$  é a **matriz de  $T$  relativa às bases  $\beta$  e  $\gamma$** . Por simplicidade, denotaremos  $A = [T]_{\beta, \gamma}$ .

### Exemplo 3.2: Representação matricial

Se  $T$  é transformação de  $\mathbb{R}^2$  em  $\mathbb{R}^3$  definida por

$$T(x, y) = (x + 3y, 2x + 5y, 7x + 9y)$$

Encontre a matriz de  $T$  com respeito às bases canônicas  $\beta$ , de  $\mathbb{R}^2$ , e  $\gamma$ , de  $\mathbb{R}^3$ .

#### Solução:

Primeiro, vamos calcular a ação de  $T$  sobre cada vetor de  $\beta$ :

$$T(1, 0) = (1, 2, 7)$$

$$T(0, 1) = (3, 5, 9)$$

Como escrever cada vetor encontrado na base  $\gamma$  é simplesmente

$$(1, 2, 7) = 1 \cdot (1, 0, 0) + 2 \cdot (0, 1, 0) + 7 \cdot (0, 0, 1)$$

$$(3, 5, 9) = 3 \cdot (1, 0, 0) + 5 \cdot (0, 1, 0) + 9 \cdot (0, 0, 1)$$

podemos montar a matriz

$$[T]_{\beta, \gamma} = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{bmatrix}$$

Note que  $T(v_j)$  pode ser encontrado ao multiplicar cada elemento da  $j$ -ésima coluna da matriz de representação pelo elemento da base  $\gamma$  correspondente e somando os resultados.

Se  $j = 2$ , por exemplo, temos  $v_2 = (0, 1)$  e  $T(0, 1)$  é dado por

$$3 \cdot (1, 0, 0) + 5 \cdot (0, 1, 0) + 9 \cdot (0, 0, 1) = (3, 5, 9)$$

Um caso especial do que acabamos de ver é quanto  $T$  é um operador linear, isto é,  $V = W$ . Nesse caso, só precisamos da base  $\beta$ . Assim, a matriz de  $T$  em relação a  $\beta$  é a matriz  $n \times n$ , denotada por  $[T]_{\beta}$  cujas entradas são dadas por

$$T(v_j) = c_{1j}v_1 + c_{2j}v_2 + \dots + c_{nj}v_n = \sum_{i=1}^n c_{ij}v_i$$

Um importante resultado sobre representação matricial é que se  $V$  é um espaço vetorial de dimensão finita com duas bases ordenadas  $\beta$  e  $\gamma$  e  $T$  é um operador linear em  $V$ , então

$$[T]_{\gamma} = P^{-1}[T]_{\beta}P$$

onde  $P$  é a matriz de mudança de base de  $\gamma$  para  $\beta$ .

Por fim, dadas duas matrizes  $A$  e  $B$ ,  $n \times n$  sobre o corpo  $\mathbb{F}$ , dizemos que  $B$  é **similar** a  $A$  sobre  $\mathbb{F}$  se existe uma matriz  $P$  invertível,  $n \times n$ , tal que  $B = P^{-1}AP$ . Observamos que a similaridade é uma relação de equivalência no conjunto  $\mathbb{M}_n(\mathbb{F})$ .

Um resultado que ilustra a importância de matrizes de representação e a facilidade que ela oferece para os cálculos é o seguinte.

Suponha que  $S$  e  $T$  são transformações lineares e  $\beta$  é uma base ordenada para  $S, T, S + T$ . Então a matriz da soma das duas transformações é igual à soma das matrizes das transformações. Isto é,

$$[S + T]_{\beta} = [S]_{\beta} + [T]_{\beta}$$

Similarmente, se  $\lambda$  é um escalar, então a matriz do escalar vezes uma transformação é igual ao escalar vezes a matriz da transformação. Em símbolos,

$$[\lambda T]_{\beta} = \lambda[T]_{\beta}$$

Assim, podemos perceber que transformações lineares “agem” como multiplicação matricial.

### Exemplo 3.3

Seja  $D$  uma transformação linear de  $\mathcal{P}_3(\mathbb{R})$  para  $\mathcal{P}_2(\mathbb{R})$  definida como  $D(p) = p'$ . Então a matriz de representação de  $D$  nas bases canônicas de  $\mathcal{P}_3(\mathbb{R})$  e  $\mathcal{P}_2(\mathbb{R})$ , denotada por  $\mathcal{M}(D)$ , é dada por:

$$\mathcal{M}(D) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

### 3.8 Exercícios Resolvidos

#### Exemplo 3.4: Encontrando a inversa

Encontre a **inversa** de  $T(x_1, x_2) = (x_1 + x_2, x_1)$ , onde  $T$  é um operador linear em  $\mathbb{F}^2$ .

**Solução:**

Primeiro verificamos que  $T$  é injetora.

Note que se  $T(x_1, x_2) = 0$ , então temos

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

Ou seja,  $x_1 = x_2 = 0$ . Portanto,  $\text{Ker}(T) = \{0\}$ .

Para verificar a sobrejetividade, considere  $(z_1, z_2) \in \mathbb{F}^2$ . Queremos demonstrar que  $(z_1, z_2)$ , que são vetores arbitrários, estão em  $\text{Im}(T)$ . Ou seja, queremos encontrar escalares  $x_1, x_2$  tais que

$$x_1 + x_2 = z_1 \text{ e } x_1 = z_2$$

De onde temos  $x_1 = z_2$ ,  $x_2 = z_1 - z_2$ .

Portanto, como  $T$  é bijetora, temos que  $T$  é invertível. O processo para verificar a sobrejetividade nos dá a seguinte fórmula para  $T^{-1}$ :

$$T^{-1}(z_1, z_2) = (z_2, z_1 - z_2)$$

## 4 Produto Interno

Assuntos abordados: Desigualdade de Cauchy-Schwarz, Norma, Ângulo e Ortogonalidade, Base Ortogonal, Processo de Gram-Schmidt.

### 4.1 O que é?

Começamos lembrando da definição de produto interno (ou produto escalar) da Geometria Analítica. Dados dos vetores de  $\mathbb{R}^3$   $u = (x_1, x_2, x_3)$  e  $v = (y_1, y_2, y_3)$ , sabemos que o produto escalar entre eles é o número real

$$\langle u, v \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3$$

que também é denotado por  $u \cdot v$ . Geometricamente, esse produto escalar é o produto do comprimento de  $u, v$  e o cosseno do ângulo entre eles. Ou seja, é possível definir conceitos geométricos bastante intuitivos como comprimento e ângulo de vetores em  $\mathbb{R}^3$  a partir da definição algébrica de produto escalar.

Nossa tarefa é generalizar esse conceito, permitindo estudar espaços vetoriais em que faz sentido falar do comprimento de um vetor e do ângulo entre dois vetores.

Isso será possível a partir de uma função que irá levar um par de vetores a um escalar. Essa função será chamada de produto interno se satisfizer algumas condições similares às que temos para o produto escalar ao qual estamos acostumados. Que condições devem ser essas? Como podemos generalizar a noção de ângulo?

Dado o corpo  $\mathbb{R}$  ou  $\mathbb{C}$  e um espaço vetorial  $V$  sobre esse corpo, diremos que o **produto interno** em  $V$  é uma função que mapeia cada par ordenado de vetores  $u, v \in V$  a um escalar  $\langle u, v \rangle$ , isto é,

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$$

satisfazendo as seguintes propriedades:

1. Simetria:  $\langle u, v \rangle = \overline{\langle v, u \rangle}$
2. Positividade:  $\langle u, u \rangle \geq 0$  sendo que  $\langle u, u \rangle = 0$  sse.  $u$  for o vetor nulo.

3. Distributividade:  $\langle u + w, v \rangle = \langle u, v \rangle + \langle w, v \rangle$ .

4. Homogeneidade:  $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$ , onde  $\lambda \in \mathbb{F}$ .

A partir dessas propriedades também segue que

- $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$
- $\langle u, \lambda v \rangle = \bar{\lambda} \langle u, v \rangle$

Note que tomar o complexo conjugado só faz sentido no caso de  $\mathbb{C}$ , porém é necessário para garantir a positividade.

Para ilustrar, daremos alguns exemplos de produtos internos. O mais natural é o **produto interno usual**, chamado de produto interno euclidiano para  $\mathbb{R}^n$  e produto interno hermitiano para  $\mathbb{C}^n$  e é definido por

$$\langle u, v \rangle = \sum_{i=1}^n x_i \bar{y}_i$$

No espaço vetorial real das funções contínuas no intervalo  $[a, b]$ , i.e.  $\mathcal{C}([a, b])$ , o produto interno usual é definido como sendo

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx; \forall f, g \in \mathcal{C}([a, b])$$

Já no espaço das matrizes com entradas reais  $\mathbb{M}_n(\mathbb{R})$ , o produto interno usual é dado por

$$\langle A, B \rangle = \text{tr}(B^t A) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ij}$$

Caso as entradas das matrizes  $A$  e  $B$  sejam complexas, basta tomar a **transposta conjugada**  $B^*$ .

## 4.2 Desigualdade de Cauchy-Schwarz

Dado um espaço vetorial real  $V$  munido de um produto interno, para todos  $u, v \in V$  temos que

$$\langle u, v \rangle^2 \leq \langle u, u \rangle \langle v, v \rangle$$

E a igualdade é válida sse. os elementos  $u$  e  $v$  são linearmente dependentes.

### 4.3 Norma e Métrica

A partir do conceito de produto interno conseguimos formalizar nossa intuição do que é um “comprimento” ou “magnitude”, o que chamaremos de norma.

Uma **norma** em um espaço vetorial  $V$  é uma função  $\|\cdot\|$  que leva cada elemento  $u \in V$  a um número real  $\|u\|$  satisfazendo:

1. Positividade:  $\|u\| > 0$  para  $u \neq 0$  e  $\|u\| = 0 \iff u = 0$ .
2. Homogeneidade:  $\|\lambda u\| = |\lambda| \|u\|$ ,  $\lambda \in \mathbb{F}$ .
3. Desigualdade triangular:  $\|u + v\| \leq \|u\| + \|v\|$ .

Um espaço vetorial  $V$  munido de uma norma  $\|\cdot\|$  é chamado de **espaço normado** e denotamos  $(V, \|\cdot\|)$ .

Além da norma euclidiana, com a qual estamos acostumados, podemos também definir:

Norma do Máximo:  $\|x\|_\infty = \max |x_i|; 1 \leq i \leq n$

Norma-1 (Norma do Táxi ou de Manhattan):  $\|x\|_1 = \sum_{i=1}^n |x_i|$

De modo geral, podemos definir a **norma induzida do produto interno** em um espaço vetorial  $V$  sobre o corpo  $\mathbb{F}$  munido do produto interno  $\langle \cdot, \cdot \rangle$ .

Para isso, definimos a função  $q(\cdot) : V \rightarrow \mathbb{R}$  como

$$q(u) = \sqrt{\langle u, u \rangle}$$

E  $q(\cdot)$  satisfaz as propriedades de norma.

Também podemos definir uma **métrica** ou **distância** a partir da função

$$\begin{aligned} d(\cdot, \cdot) : V \times V &\rightarrow \mathbb{R} \\ (u, v) &\rightarrow d(u, v) \end{aligned}$$



satisfazendo as propriedades de positividade, simetria e desigualdade triangular.

Um espaço vetorial munido de uma métrica é chamado de **espaço métrico**.

Um exemplo imediato de métrica é  $d(u, v) = \|u - v\|$ .

## 4.4 Ângulo e Ortogonalidade

Utilizando a desigualdade de Cauchy-Schwarz, definimos o **ângulo** entre dois vetores não nulos  $u, v \in V$  como sendo o valor  $\theta \in [0, \pi]$  que satisfaz

$$\cos(\theta) = \frac{\langle u, v \rangle}{\|u\|_2 \|v\|_2}$$

Considere dois vetores  $u, v$  em um espaço com produto interno  $V$ .  $u$  é dito **ortogonal** a  $v$  se  $\langle u, v \rangle = 0$ . Pela simetria do produto interno, temos que  $v$  é ortogonal a  $u$  e assim dizemos que  $u$  e  $v$  são ortogonais.

Um conjunto  $S$  de vetores em  $V$  é dito **conjunto ortogonal** se todos os pares de vetores distintos em  $S$  são ortogonais. Caso  $\|u\| = 1$  para todo vetor  $u \in S$ , dizemos que  $S$  é um **conjunto ortonormal**.

É vantajoso trabalhar com bases ortonormais porque elas facilitam os cálculos com coordenadas.

Um importante resultado é que qualquer conjunto de vetores ortogonais não nulos é linearmente independente. Note que o número de vetores em um conjunto ortogonal sem vetores nulos é menor ou igual à dimensão do espaço  $V$ .

## 4.5 Processo de Gram-Schmidt

### Teorema 4.1: Processo de Ortogonalização de Gram-Schmidt

Dado um espaço vetorial  $V$  com produto interno e  $v_1, \dots, v_n$  vetores linearmente independentes em  $V$ . Então é possível construir vetores ortogonais  $u_1, \dots, u_n \in V$  tal que para todo  $k = 1, 2, \dots, n$

o conjunto

$$\{u_1, \dots, u_k\}$$

é uma base para o subespaço gerado por  $v_1, \dots, v_k$ .

**Demonstração:** Primeiro vamos definir  $u_1 = v_1$ . Encontraremos os outros vetores de maneira recursiva.

Suponha que  $u_1, \dots, u_m$ , onde  $1 \leq m < n$  foram escolhidos satisfazendo que para todo  $k$

$$\{u_1, \dots, u_k\}, 1 \leq k \leq m$$

é uma base ortogonal para o subespaço de  $V$  gerado por  $v_1, \dots, v_k$ .

Para construir o próximo vetor  $u_{m+1}$ , seja

$$u_{m+1} = v_{m+1} - \sum_{k=1}^m \frac{\langle v_{m+1}, u_k \rangle}{\|u_k\|^2} u_k$$

Sabemos que  $u_{m+1} \neq 0$  pois caso contrário  $v_{m+1}$  seria uma combinação linear de  $u_1, \dots, u_m$  e, portanto, combinação linear de  $v_1, \dots, v_m$ .

Logo, se  $1 \leq j \leq m$  então

$$\begin{aligned} \langle u_{m+1}, u_j \rangle &= \langle v_{m+1}, u_j \rangle - \sum_{k=1}^m \frac{\langle v_{m+1}, u_k \rangle}{\|u_k\|^2} \langle u_k, u_j \rangle \\ &= \langle v_{m+1}, u_j \rangle - \langle v_{m+1}, u_j \rangle \\ &= 0 \end{aligned}$$

Dessa forma,  $\{u_1, \dots, u_{m+1}\}$  é um conjunto ortogonal com  $m+1$  vetores não nulos no subespaço gerado por  $v_1, \dots, v_{m+1}$ . Portanto, é uma base para esse subespaço.

E os vetores  $u_1, \dots, u_n$  podem ser construídos recursivamente como definimos. ■

## 4.6 Complemento Ortogonal

Seja  $V$  um espaço vetorial com produto interno e  $S$  um conjunto de vetores em  $V$ . O **complemento ortogonal** de  $S$  é o conjunto  $S^\perp$  de todos os vetores em  $V$  que são ortogonais a todo vetor em  $S$ . Isto é,

$$S^\perp = \{v \in V : \langle v, u \rangle = 0, \forall u \in S\}$$

## 4.7 Projeção Ortogonal

Dado um conjunto vetorial  $V$  munido de um produto interno,  $S$  um subespaço de  $V$  com dimensão finita,  $\beta = \{q_1, \dots, q_n\}$  uma base ortonormal de  $S$  e um vetor  $v \in V$ , dizemos que o vetor  $s \in S$  dado por

$$s = \sum_{j=1}^n \langle v, q_j \rangle q_j$$

é **projeção ortogonal** do elemento  $v$  sobre o subespaço  $S$ . E o elemento  $w = v - s$  é a projeção ortogonal do elemento  $v$  sobre o espaço  $S^\perp$ .

Com isso, se  $\beta$  for uma base ortogonal (não ortonormal) para  $S$ , então a projeção de  $v \in V$  sobre  $S$  é dado por

$$s = \sum_{j=1}^n \frac{\langle v, q_j \rangle}{\langle q_j, q_j \rangle} q_j$$

Se todo vetor em  $V$  tem projeção ortogonal em  $S$ , a função que relaciona cada vetor em  $V$  com sua projeção em  $S$  é dito **projeção ortogonal de  $V$  em  $S$** .

## 4.8 Operadores auto-adjuntos

Seja  $T$  um operador linear sobre um espaço  $V$  com produto interno. Dizemos que  $T$  tem uma **adjunta** em  $V$  se existe um operador linear  $T^*$  em  $V$  tal que  $\langle Tu, v \rangle = \langle u, T^*v \rangle$  para todo  $u$  e  $v$  em  $V$ .

Notemos que a adjunta é semelhante ao conjugado em números complexos. E valem as seguintes propriedades:

1.  $(T + U)^* = T^* + U^*$
2.  $(cT)^* = \bar{c}T^*$

$$3. (TU)^* = U^*T^*$$

$$4. (T^*)^* = T$$

Assim como um número complexo  $z$  é real sse.  $z = \bar{z}$ , queremos definir algo semelhante para o caso em que  $T = T^*$ . Caso essa igualdade seja satisfeita, dizemos que o operador linear  $T$  é **auto-adjunto** (ou **hermitiano**, no caso complexo, e **simétrico**, no caso real).

Caso  $T$  comute com sua adjunta, i.e.,  $TT^* = T^*T$ , então  $T$  é dito operador **normal**.

## 4.9 Operadores ortogonais

Análogo ao que fizemos para operadores auto-adjuntos, temos que  $T$  é um **operador ortogonal** se

$$\langle T(u), T(v) \rangle = \langle u, v \rangle \forall u, v \in V$$

Note que se  $T$  é ortogonal, então  $T$  preserva ângulos.

Dizemos que  $T$  é uma **isometria** se:

$$\|T(u)\| = \|u\|, \forall u \in V$$

Com um operador ortogonal  $T$ , podemos enunciar alguns resultados importantes:

1.  $T$  é um isomorfismo (i.e.  $T$  é bijetora)
2. A inversa  $T^{-1}$  é ortogonal
3.  $T$  é ortogonal sse.  $T$  é isometria.

## 5 Autovalores e Autovetores

Nesta seção veremos o que são Autovalores e Autovetores de Operadores e Matrizes, Multiplicidade Geométrica e Algébrica, Matrizes Especiais, Diagonalização e Teorema Espectral.

### 5.1 Motivação

A questão que motiva esta seção é encontrar uma base ordenada do espaço vetorial  $V$  na qual o operador linear  $T$  seja representado (de forma matricial) da maneira mais simples possível.

Para exemplificar, considere a matriz diagonal

$$D = \begin{bmatrix} c_1 & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ 0 & 0 & c_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{bmatrix}$$

E suponha que  $T$  é um operador linear num espaço finito  $V$ . Se existe uma base ordenada  $\beta = \{v_1, v_2, \dots, v_n\}$  de  $V$  na qual  $T$  é representada matricialmente como a matriz diagonal  $D$ , então é possível retirar informações sobre o operador linear  $T$ , como seu posto ou determinante, de maneira mais simples e direta.

Como

$$[T]_{\beta} = D \iff T(v_k) = c_k v_k, k = 1, 2, \dots, n$$

A imagem de  $T$  é simplesmente o subespaço gerado pelos vetores  $v_k$  nos quais  $c_k \neq 0$ . Analogamente, o núcleo de  $T$  é gerado pelos  $v_k$  restantes.

Com isso em mente, levantam-se as seguintes questões. É sempre possível representar um operador linear  $T$  como uma matriz diagonal? Se não, qual é a forma mais simples de representar matricialmente esse operador?

## 5.2 O que são?

Vimos que no caso em que  $T$  pode ser representado como uma matriz diagonal, temos que

$$[T]_{\beta} = D \iff T(v_k) = c_k v_k, k = 1, 2, \dots, n$$

Assim, vamos estudar quais vetores são levados por  $T$  em múltiplos escalares de si mesmos.

Dado um espaço vetorial  $V$  sobre um corpo  $\mathbb{F}$  e  $T$  um operador linear em  $V$ , vamos definir o **autovalor** (ou valor característico, ou **eigenvalue**) de  $T$  como sendo o escalar  $\lambda \in \mathbb{F}$  tal que existe um valor  $v \in V, v \neq 0$  satisfazendo  $T(v) = \lambda v$ .

Se  $\lambda$  é um autovalor de  $T$ , então:

- Qualquer vetor  $v$  satisfazendo  $T(v) = \lambda v$  é dito **autovetor** (ou vetor característico, ou **eigenvector**) de  $T$  associado ao autovalor  $\lambda$ .
- O conjunto de todos os autovetores  $v$  é chamado **autoespaço** (ou espaço característico, ou **eigenspace**) associado a  $\lambda$ .

Além dos nomes citados acima, autovalores também são conhecidos como valores próprios, valores espectrais, raízes características ou raízes latentes.

## 5.3 Como encontrá-los?

Note que o autoespaço associado a  $\lambda$  é um subespaço de  $V$  e é precisamente o núcleo da transformação linear  $(T - \lambda I)$ . Dizemos que  $\lambda$  é autovalor de  $T$  quando o autoespaço é diferente do espaço nulo, isto é, se  $(T - \lambda I)$  não é isomorfismo. Se  $V$  for um espaço de dimensão finita, então  $(T - \lambda I)$  não é isomorfismo justamente quando seu determinante é diferente de zero. Resumindo, temos o seguinte

### Teorema 5.1

Seja  $T$  um operador linear em um espaço de dimensão finita  $V$  e seja  $\lambda$  um escalar. As seguintes afirmações são equivalentes:

1.  $\lambda$  é autovalor de  $T$ .

2. O operador  $(T - \lambda I)$  é singular (i.e. não invertível).
3.  $\det(T - \lambda I) = 0$ .

A partir do terceiro critério temos um caminho para encontrar os autovalores de  $T$ . Como  $\det(T - \lambda I)$  é um polinômio de grau  $n$  na variável  $\lambda$ , podemos encontrar os autovalores como sendo as raízes desse polinômio.

Se  $A$  é a representação matricial de  $T$  na base ordenada  $\beta$  (i.e.  $A = [T]_\beta$ ), então  $(T - \lambda I)$  é invertível sse.  $(A - \lambda I)$  for invertível. O que podemos resumir na seguinte definição:

Se  $A$  é uma matriz  $n \times n$  sobre o corpo  $\mathbb{F}$ , um autovalor de  $A$  em  $\mathbb{F}$  é um escalar  $\lambda \in \mathbb{F}$  tal que a matriz  $(A - \lambda I)$  é singular.

Ou seja,  $\lambda$  é um **autovalor da matriz**  $A$  sse.  $\det(A - \lambda I) = 0$ .

Isso nos motiva a definir o **polinômio característico** de  $A$  como sendo

$$f(x) = \det(A - xI)$$

Um importante resultado é que matrizes similares têm o mesmo polinômio característico. Isso implica que elas também possuem os mesmos autovalores.

## 5.4 Multiplicidades algébrica e geométrica

Considere um operador linear  $T : V \rightarrow V$  sobre  $\mathbb{F}$  e uma base  $\beta$  qualquer de  $V$ . Seu polinômio característico é dado por

$$p_T(\lambda) = \det([T]_\beta^\beta - \lambda I)$$

Se  $\lambda_1, \lambda_2, \dots, \lambda_k$  são as raízes de  $p_T(\lambda)$ , então pelo Teorema Fundamental da Álgebra temos que

$$p_T(\lambda) = a(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_k)^{m_k}$$

Escolhendo um autovalor  $\lambda_i$ , definimos a **multiplicidade**

- **algébrica** de  $\lambda_i$  como o expoente do termo  $(\lambda - \lambda_i)$  em  $p_T(\lambda)$ .
- **geométrica** de  $\lambda_i$  como  $\dim \text{Ker}(T - \lambda_i I)$ .

É importante notar que a multiplicidade geométrica é sempre menor ou igual à multiplicidade algébrica.

## 5.5 Diagonalização

Dado  $T \in \mathcal{L}(V)$ , dizemos que  $T$  é **diagonalizável** se existe uma base  $\beta = \{v_1, v_2, \dots, v_n\}$  para  $V$  formada pelos autovetores de  $T$ . Isto é, o operador linear tem uma matriz diagonal com respeito a alguma base de  $V$ .

Como  $T(v_i) = \lambda_i v_i$ , a representação de  $T$  na base  $\beta$  é dada por:

$$[T]_{\beta} = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

Alguns resultados importantes:

1.  $T$  é diagonalizável sse. existir uma base de  $V$  formada por autovetores de  $T$ .
2. Se  $f$  é um polinômio qualquer e  $T(v) = \lambda v$ , então  $f(T(v)) = f(\lambda)v$ .
3. Se  $\lambda_1, \dots, \lambda_k$  são autovalores distintos e  $v_1, \dots, v_k$  são autovetores associados a  $\lambda_1, \dots, \lambda_k$  respectivamente, então  $\{v_1, \dots, v_k\}$  é linearmente independente.
4. Se  $\dim(V) = n$  e  $\lambda_1, \dots, \lambda_n$  são autovalores distintos de  $T$ , então  $T$  é diagonalizável. Em outras palavras, se  $T$  possui todos os autovalores distintos, então  $T$  é diagonalizável.
5. Se  $W_i$  é o autoespaço associado ao autovalor  $\lambda_i$  e  $W = W_1 + W_2 + \dots + W_k$ , então

$$\dim(W) = \dim(W_1) + \dots + \dim(W_k)$$

Além disso, se  $\beta_i$  é uma base ordenada para  $W_i$ , então  $\beta = \{\beta_1, \dots, \beta_n\}$  é base ordenada para  $W$ . Note que isso significa que a soma dos autoespaços é uma soma direta.



Com essas conclusões é possível desconfiar que existem mais equivalências entre transformações diagonalizáveis e seus autovalores e autoespaços. De fato, temos o seguinte

### Teorema 5.2

Suponha  $V$  um espaço vetorial de dimensão finita. Seja  $T \in \mathcal{L}(V)$ ,  $\lambda_1, \dots, \lambda_k$  autovalores distintos de  $T$  e  $W_i = \text{Ker}(T - \lambda_i I)$ . As seguintes afirmações são equivalentes:

1.  $T$  é diagonalizável.
2. O polinômio característico de  $T$  é

$$p_T(\lambda) = (\lambda - \lambda_1)^{d_1} (\lambda - \lambda_2)^{d_2} \dots (\lambda - \lambda_k)^{d_k}$$

e  $\dim(W_i) = d_i$  para  $i = 1, 2, \dots, k$ .

3.  $\dim(W_1) + \dots + \dim(W_k) = \dim(V)$ .

A partir desse resultado, dada uma matriz diagonalizável  $A$ , conseguimos encontrar uma matriz diagonal  $\Lambda$ , similar à  $A$ , tal que

$$A = P\Lambda P^{-1} \text{ e } \Lambda = P^{-1}AP$$

onde  $\Lambda$  é construída a partir dos autovalores de  $A$  e  $P$ , a partir dos autovetores de  $A$ .

## 5.6 Subespaço Invariante

Suponha  $T \in \mathcal{L}(V)$ . Se decompormos  $V$  em somas diretas

$$V = U_1 \oplus \dots \oplus U_m$$

onde cada  $U_j$  é um subespaço próprio de  $V$ , então para entender o comportamento de  $T$  basta analisar o comportamento de  $T$  em cada  $U_j$ . Para facilitar a leitura, vamos denotar  $T|_{U_j}$  para nos referirmos a  $T$  restrito a um subespaço  $U_j$ .

Porém, nem sempre  $T|_{U_j}$  terá a imagem no próprio subespaço  $U_j$ . Por isso, iremos nos munir da seguinte definição.

Um subespaço  $U$  de  $V$  é dito **subespaço invariante** sobre  $T$  se  $u \in U$  implica que  $T(u) \in U$ . Isto é,  $U$  é invariante sobre  $T$  se  $T|_U$  é um operador linear em  $U$ .

## 5.7 Matrizes Especiais

Uma matriz quadrada  $A$ , sobre o corpo  $\mathbb{F}$ , tal que  $A = \overline{A^T}$ , i.e. cada  $a_{ij} = \overline{a_{ji}}$ , é dita **matriz simétrica** se  $\mathbb{F} = \mathbb{R}$  e é dita **matriz hermitiana** se  $\mathbb{F} = \mathbb{C}$ .

Para uma matriz  $2 \times 2$ , uma matriz é hermitiana sse. for da forma

$$\begin{bmatrix} z & x + iy \\ x - iy & w \end{bmatrix}$$

onde  $x, y, z, w \in \mathbb{R}$ .

Observe que se  $A \in \mathbb{M}_n(\mathbb{F})$  é matriz simétrica/hermitiana e  $X, Y \in \mathbb{F}^n$ , então

$$\langle AX, Y \rangle = \langle X, A^*Y \rangle = \langle X, AY \rangle$$

Outro resultado importante é que se  $A$  é hermitiana, então os autovalores de  $A$  são reais e os autovetores associados a autovalores distintos são ortogonais entre si.

Dada uma matriz  $A \in \mathbb{M}_n(\mathbb{R})$  simétrica, então  $A$  é dita **matriz positiva definida** caso

$$\langle Ax, x \rangle = x^T Ax > 0 \forall x \neq 0, x \in \mathbb{R}^n$$

Caso  $A \in \mathbb{M}_n(\mathbb{C})$  seja uma matriz hermitiana, então  $A$  é dita **matriz positiva definida** caso

$$\langle Ax, x \rangle = x^* Ax > 0 \forall x \neq 0, x \in \mathbb{C}^n$$

Quando é o caso que uma matriz simétrica ou hermitiana é positiva definida?

### Teorema 5.3

Seja  $A \in \mathbb{M}_n(\mathbb{C})$  hermitiana. Então  $A$  é positiva definida sse. todos seus autovalores são positivos.

Uma matriz  $A$ , com entradas reais ou complexas,  $n \times n$ , é dita **matriz ortogonal** se  $A^T A = I$ . Caso  $A$  tenha entradas complexas e  $A^* A = I$ , então  $A$  é dita **matriz unitária**.

Dois resultados importantes sobre matrizes unitárias são:

1. Se  $A$  é matriz unitária e  $\lambda$  é autovalor de  $A$ , então  $|\lambda| = 1$ .

2. Se  $A$  é matriz unitária, então  $|\det(A)| = 1$ .

### Exemplo 5.1: Diagonalização de operador linear

Seja  $T \in \mathcal{L}(\mathbb{R}^3)$  definido por

$$T(x, y, z) = (-9x + 4y + 4z, -8x + 3y + 4z, -16x + 8y + 7z)$$

Mostre que  $T$  é diagonalizável e encontre os autovetores que formam uma base para  $\mathbb{R}^3$ .

**Solução:** Note que a matriz de  $T$  na base canônica  $\beta$  é:

$$[T]_{\beta} = \begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}$$

O primeiro passo é encontrar os autovalores de  $[T]_{\beta}$ . Calculando  $\det([T]_{\beta} - \lambda I)$ :

$$\begin{vmatrix} -9 - \lambda & 4 & 4 \\ -8 & 3 - \lambda & 4 \\ -16 & 8 & 7 - \lambda \end{vmatrix} = -\lambda^3 + \lambda^2 + 5\lambda + 3 = 0 \iff (\lambda + 1)^2(\lambda - 3) = 0$$

Assim, temos dois autovalores  $\lambda_1 = -1$ , com multiplicidade algébrica igual a dois, e  $\lambda_2 = 3$ .

Calculando o autovetor associado a  $\lambda_1 = -1$ :

$$\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \iff \begin{cases} -8x_1 + 4x_2 + 4x_3 = 0 \\ -8x_1 + 4x_2 + 4x_3 = 0 \\ -16x_1 + 8x_2 + 8x_3 = 0 \end{cases}$$

Note que temos apenas uma linha linearmente independente. Ou seja, o núcleo da matriz dos coeficientes tem posto igual a dois. Isso significa que podemos extrair dois autovetores linearmente independentes.

De fato, podemos tomar  $x_1 = 1, x_2 = 2, x_3 = 0$  e  $x_1 = 1, x_2 = 0, x_3 = 2$ , obtendo os autovetores  $(1, 2, 0)$  e  $(1, 0, 2)$ .

Para  $\lambda_2 = 3$ , temos o sistema:

$$\begin{cases} -12x_1 + 4x_2 + 4x_3 = 0 \\ -8x_1 + 0x_2 + 4x_3 = 0 \\ -16x_1 + 8x_2 + 4x_3 = 0 \end{cases} \iff \begin{cases} x_1 & = \frac{1}{2}x_3 \\ x_2 & = \frac{1}{2}x_3 \\ x_3 & = x_3 \end{cases}$$

Portanto, podemos escolher o vetor  $(1, 1, 2)$ . Como obtivemos três autovetores linearmente independentes, temos que  $T$  é um operador diagonalizável. Além disso, temos a seguinte base para  $\mathbb{R}^3$ :

$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 1 & 1 & 2 \end{bmatrix}$$

## 5.8 Teorema Espectral

### Teorema 5.4: Teorema Espectral

Suponha  $T$  um operador linear em um espaço vetorial finito  $V$ . Caso  $V$  esteja definido sobre  $\mathbb{C}$ , considere  $T$  normal. Caso  $V$  esteja definido sobre  $\mathbb{R}$ , considere  $T$  auto-adjunto.

Seja  $\lambda_1, \dots, \lambda_k$  autovalores distintos de  $T$ ,  $W_j$  o autoespaço associado a  $\lambda_j$  e  $E_j$  a projeção ortogonal de  $V$  em  $W_j$ . Então, valem as seguintes afirmações:

1.  $W_j$  é ortogonal a  $W_i$  quando  $i \neq j$ .
2.  $V$  é a soma direta de  $W_1, \dots, W_k$ .
3.  $T$  pode ser decomposto da forma

$$T = \lambda_1 E_1 + \dots + \lambda_k E_k$$

denominada **resolução espectral**.

## Referências

- [Alg71] Linear Algebra. *Kenneth Hoffman and Ray Kunze*. Prentice Hall, Englewood Cliffs, New Jersey, 1971.
- [Axl14] Sheldon Axler. *Linear algebra done right*. Springer, 2014.
- [Coe01] Flávio Ulhoa Coelho. *Curso de Álgebra Linear, Um Vol. 34*. Edusp, 2001.

[Pul12] Petronio Pulino. Algebra linear e suas aplicações notas de aula.  
*Campinas: Universidade Estadual de Campinas, 2012.*