Fundamentos de Álgebra Linear

Notas de aula PPGEM

Prof. Eduardo Lenz Cardoso

Conteúdo

I Fundamentos de Análise

1 Introdução . . . 7

1.1 Conceitos Básicos 7

1.1.1 Afirmações do tipo Se->Então . . . 9

1.1.2 Afirmações do tipo Se-e-somente-se . . . 10

1.2 Operadores Lógicos 10

1.2.1 E Lógico (AND) . . . 10

1.2.2 Não Lógico (NOT) . . . 11

1.2.3 Ou lógico (OR) . . . 11

1.3 Provas 11

1.3.1 Prova Direta . . . 11

1.3.2 Refutação direta . . . 12

1.3.3 Refutação por Contra-Exemplo . . . 12

2 Teoria de Conjuntos . . . 15

2.1 Introdução 15

2.2 Conceitos fundamentais 16

2.2.1 Negação de quantificadores . . . 17

2.2.2 União de quantificadores . . . 17

2.3 Operações com Conjuntos 18

3 Números Reais . . . 21

3.1 Algumas propriedades de conjuntos de números reais. 23

4 Funções 25
4.1 Continuidade 29
4.2 Composição de Funções 31

5 $\mathbb{R}^n$ 35
5.1 Bases 37
5.2 Conjuntos abertos e fechados em $\mathbb{R}^n$ 40

6 Espaços Vetoriais 45
6.1 Transformações Lineares 47
6.2 Transformação Linear de Espaços Finito Dimensionais 49
6.3 Mudança de Base 56
6.4 Mudança de Base Aplicada a Operadores Lineares 61
6.5 Formas lineares, bilineares e quadráticas 63
6.5.1 Funcional Linear 64
6.5.2 Forma Bilinear 64
6.5.3 Formas Quadráticas 65
6.6 Continuidade de Operadores 66

7 Autovalores e Autovetores 69
7.1 Autovalores e Autovetores Reais 73
7.2 Operador Diagonalizável e Multiplicidade de Autovalores 75
7.3 Problema Generalizado de Autovalores e Autovetores 77

8 Produto Interno 81
8.1 Ortogonalidade 82
8.2 Projeção 85
8.3 Ortogonalização de Gramm-Schmidt 89

9 Sequências 91
9.1 Sequências de Cauchy 92
9.2 Convergência de sequência de funções 93
9.3 Espaço Completo 96

10 Série de Fourier 99

Fundamentos de Análise

1 Introdução **7**
1.1 Conceitos Básicos
1.2 Operadores Lógicos
1.3 Provas

2 Teoria de Conjuntos **15**
2.1 Introdução
2.2 Conceitos fundamentais
2.3 Operações com Conjuntos

3 Números Reais **21**
3.1 Algumas propriedades de conjuntos de números reais.

4 Funções **25**
4.1 Continuidade
4.2 Composição de Funções

5 $\mathbb{R}^n$ **35**
5.1 Bases
5.2 Conjuntos abertos e fechados em $\mathbb{R}^n$

6 Espaços Vetoriais **45**
6.1 Transformações Lineares
6.2 Transformação Linear de Espaços Finito Dimensionais
6.3 Mudança de Base
6.4 Mudança de Base Aplicada a Operadores Lineares
6.5 Formas lineares, bilineares e quadráticas
6.6 Continuidade de Operadores

7 Autovalores e Autovetores **69**
7.1 Autovalores e Autovetores Reais
7.2 Operador Diagonalizável e Multiplicidade de Autovalores
7.3 Problema Generalizado de Autovalores e Autovetores

8 Produto Interno **81**
8.1 Ortogonalidade
8.2 Projeção
8.3 Ortogonalização de Gramm-Schmidt

9 Sequências **91**
9.1 Sequências de Cauchy
9.2 Convergência de sequência de funções
9.3 Espaço Completo

10 Série de Fourier **99**

1. Introdução

Os conceitos apresentados aqui servirão para que o leitor seja capaz de compreender um texto da área de mecânica computacional, bem como entenda a lógica por trás das implementações computacionais que serão discutidas ao longo do curso de mestrado.

O texto é organizado de maneira a introduzir os conceitos fundamentais em uma sequência lógica. Para isto, iremos iniciar com uma revisão de conceitos e nomenclaturas básicas da matemática. Após, será apresentado o conceito de conjunto e todas as propriedades de interesse neste curso, uma discussão sobre os números reais e suas propriedades, conceitos sobre funções e uma discussão sobre o $\mathbb{R}^N$ e suas propriedades. Com esta revisão de conceitos fundamentais, iremos estudar em detalhes os espaços de funções para, após, nos concentrarmos na definição dos problemas a serem estudados na disciplina e na sua implementação numérica.

É importante salientar que este material é de apoio e **não pretende substituir a bibliografia básica sobre os assuntos abordados**. Para tanto, sugiro fortemente que o leitor consulte sempre que possível livros clássicos da área, como por exemplo os livros: *Optimization by Space Methods* de David. G. Luemberger, *Engineering Mathematics* de Stroud and Booth, *Mathematical Methods for Physics and Engineering (3rd edition): A Comprehensive Guide* de K.F. Riley, *Advanced Engineering Mathematics* de Kreyszig e *Introdução à Análise Linear*-1, 2 e 3 de Donald Kreider.

Os quatro primeiros capítulos são fortemente baseados no livro *Matemática Discreta: uma Introdução* de Edward R. Scheinerman. Os demais capítulos são um apanhado de um vasto número de livros e das notas de aula.

1.1 Conceitos Básicos

Para iniciar, vamos relembrar alguns conjuntos básicos

Notação 1.1. *Conjuntos básicos*

$$\begin{array}{lcll}
\mathbb{N} & = & \{0,1,2,3,....\} & \textit{Números naturais} \\
\mathbb{N}^* & = & \{1,2,3,4,....\} & \textit{Números naturais sem o zero} \\
\mathbb{Z} & = & \{0,1,-1,2,-2,...\} & \textit{Números inteiros} \\
\mathbb{Z}^* & = & \{1,-1,2,-2,...\} & \textit{Números inteiros sem o zero} \\
\mathbb{Q} & = & \left\{\frac{m}{n} : m,n \in \mathbb{Z}, n \neq 0\right\} & \textit{Números racionais} \\
\mathbb{Q} & = & \left\{\frac{m}{n} \middle| m \in \mathbb{Z}\, e\, , n \in \mathbb{Z}^*\right\} & \textit{Números racionais sem o zero}
\end{array}$$

em que $\mathbb{N}$ é o conjunto dos números naturais, $\mathbb{Z}$ dos números inteiros (do alemão *Zhalen*) e $\mathbb{Q}$ dos números racionais (de *Quotient*). Com estes conjuntos básicos, podemos construir definições mais avançadas, como por exemplo:

Definição 1.1.1 Número par
Um número par é um número natural divisível por 2.

Observem que a definição de par depende de 3 conceitos prévios: número natural, divisível e 2. Alguns são "óbvios", como o 2 (que já que é necessário para definirmos o conjunto $\mathbb{N}$) mas alguns que parecem óbvios para nós nem sempre são tão claros como parecem. Vamos definir o conceito de divisível:

Definição 1.1.2 Divisível
Sejam a e b inteiros. Dizemos que a é divisível por b se existe um inteiro c tal que $bc = a$. Dizemos também que b divide a, ou que b é um fator de a, ou que b é um divisor de a. A notação correspondente é $b|a$.

Desta forma, sempre partimos de conceitos primitivos (axiomas) para obtermos conceitos derivados, por meio de inferências lógicas. O importante é dominar os conceitos fundamentais e, com algumas ferramentas da matemática, construir o conhecimento na nossa área de interesse. Desta forma, tendo o conceito de divisível e de inteiro, podemos escrever:

Definição 1.1.3 Número par (definição alternativa)
Um número natural a é par se $2|a$

onde lemos que 2 é divisor de a ou que 2 divide a.

Com esta ideia básica sobre como devemos proceder para construirmos conceitos mais avançados, devemos agora definir algumas ferramentas muito utilizadas na bibliografia técnica da nossa área. Uma forma de definirmos conceitos derivados é pelo estabelecimento de um Teorema (ou os seus equivalentes). Um teorema nada mais é do que uma afirmação declarativa sobre matemática, para a qual existe uma PROVA (uma afirmação declarativa é algo como "vai chover"ou "o Internacional é melhor do que o grêmio"). Assim, podemos dizer que um teorema é uma afirmação sobre algo que sabemos ser verdade e sabemos provar. Por sua vez, uma Conjectura não tem prova, como por exemplo, "todo o inteiro par maior do que 2 é a soma de dois primos"(conjectura de Goldbach, 1742). Podemos rodar um programa de computador por 500 anos sem que isto se verifique falso, mas para um matemático isto não é uma prova (para aplicações de engenharia, pode servir, desde que se mostre que trabalharemos na faixa de valores verificados pelo programa). Atualmente, a afirmação é verdadeira para a faixa $[4, 4 \times 10^{18}]$.

Existem alguns termos utilizados em textos matemáticos, como por exemplo (essa é uma leitura minha, que pode variar um pouco de autor para autor)

- Teorema: quando é importante :0)
- Resultado: mesma coisa, só que é menos importante
- Fato: teorema de importância limitada ($1 + 1 = 2$)
- Proposição: Um teorema de importância limitada

- Lema: Um teorema cujo objetivo é ajudar a provar um outro teorema
- Corolário: Resultado com uma prova rápida, onde usamos resultados já provados em outros teoremas
- Alegação: Equivale a um lema.

Como os conceitos mais importantes que devemos utilizar estarão na forma de um teorema (ou um sinônimo), devemos entender como eles são apresentados e quais são os conceitos básicos envolvidos nas provas que deveremos ler. Vamos avaliar dois tipos de apresentação de teoremas na forma se-então e se-e-somente-se.

1.1.1 Afirmações do tipo Se->Então

São afirmações que apresentam a seguinte estrutura: "Se A, então B" ou $A \Rightarrow B$, onde A é uma hipótese e B uma conclusão. A lógica desta afirmação segue a tabela

A	B	$A \Rightarrow B$
V	V	Possível
V	F	Impossível
F	V	Possível
F	F	Possível

Observem que a situação IMPOSSÍVEL basta para mostrar que um teorema é falso (veremos isto mais para frente quando estudarmos os tipos de prova). É interessante, também, verificar que a situação não A mas B, como no caso da soma de dois ímpares resultar em um par e possível e o teorema não afirma que isto não possa ocorrer.

Este tipo de afirmação pode ser escrita de formas alternativas, como por exemplo:

- "A implica em B";
- "Sempre que A, temos B";
- "A é suficiente para B";
- "A é condição suficiente para B"

■ **Exemplo 1.1** Exemplo de teorema com afirmação $A \Rightarrow B$

Considere o teorema

Teorema 1.1.1 Exemplo de teorema

Se x e y forem pares, então $x+y$ é par

e podemos observar que A corresponde a "Se x e y forem pares"e B a "então $x+y$ é par". Desta forma, temos uma implicação em somente um sentido, de tal forma que o simbolismo gráfico para este tipo de estrutura é $A \Rightarrow B$. No entanto, nada é dito sobre, por exemplo, a soma de dois impares também resultar em um par. ■

■ **Exemplo 1.2** Exemplo de afirmação genérica com afirmação $A \Rightarrow B$

Considere a seguinte afirmação: "Se eu ganhar na Mega-Sena, todos os alunos da disciplina serão aprovados". Neste caso, a hipótese A é "ganhar na Mega-Sena"e a conclusão B seria a aprovação de todos os alunos da disciplina. Vamos avaliar a tabela verdade:

- A mas não B: Isto viola a afirmação, pois se eu ganhar os alunos tem que passar. Por isto (lembrem-se que estamos falando de matemática, onde não tem enganação !) esta situação consta como impossível;
- A e B: Situação em que A se verifica e, portanto, B ocorre;
- B mas não A: Pode ser que eu não ganhe, mas os alunos sejam muito bons e, portanto, serão aprovados de qualquer forma;

- Nem A nem B: Pode ser que eu não ganhe, mas os alunos reprovem por falta de competência, sem qualquer relação com o resultado da Mega-Sena.

■

1.1.2 Afirmações do tipo Se-e-somente-se

São afirmações que apresentam a seguinte estrutura: "Se A então B e se B então A". A lógica segue a tabela verdade

A	B	$A \Leftrightarrow B$
V	V	Possível
V	F	Impossível
F	V	Impossível
F	F	Possível

Outras formas utilizadas são

- "A sse B";
- "A iff B";
- "A é equivalente a B".

■ **Exemplo 1.3** Exemplo de teorema com afirmação $A \iff B$

Teorema 1.1.2 Um inteiro x é par se e somente se $x+1$ é impar.

■

que poderia ser lido como "Se um inteiro x é par, então $x+1$ é impar, e se $x+1$ é impar, então x é par (observem como fica redundante..). Neste teorema, temos então que A seria "x é par"e B seria "$x+1$ é impar". Para indicar este caminho em duas vias, utilizamos a notação $A \iff B$.

■ **Exemplo 1.4** Exemplo de afirmação $A \iff B$

A frase "Um ponto x de uma função contínua é um máximo local se e somente se a primeira derivada neste ponto for nula e a segunda derivada for negativa". Observem que nesta frase temos as seguintes partes: A é "Um ponto x de uma função contínua é um máximo local"e a parte B é "primeira derivada neste ponto for nula **e** a segunda derivada for negativa". Assim, do cálculo I, sabemos que estas condições só ocorrem juntas e, portanto, chamamos B de "condição necessária e suficiente"para A.

■

1.2 Operadores Lógicos

A leitura de afirmações que utilizem lógica requer o conhecimento dos operadores lógicos fundamentais.

1.2.1 E Lógico (AND)

Observem que no exemplo da seção anterior utilizamos duas condições para montar o B: "primeira derivada neste ponto for nula **e** a segunda derivada for negativa". O **e** que está unindo as condições "primeira derivada neste ponto for nula"com "segunda derivada for negativa"é um operador lógico (E, AND ou $\wedge$) e tem a tabela verdade

A	B	$A \wedge B$
V	V	V
V	F	F
F	V	F
F	F	F

■ **Exemplo 1.5** Exemplo com operador lógico **e**

"Todo o inteiro cujo algarismo das unidades é zero é divisível por 2 **e** por 5". ■

1.2.2 Não Lógico (NOT)

Simplesmente inverte o significado lógico do operador. Por exemplo "todo número par não é impar" pode ser visto como $A \Rightarrow not(B)$ ou, SE par ENTÃO NÃO impar. A tabela verdade é simplesmente

A	$\neg A$
V	F
F	V

1.2.3 Ou lógico (OR)

é um operador que admite a possibilidade de ao menos uma afirmação ser verdadeira, ou em outras palavras, se um elemento for satisfeito então a condição conjunta é satisfeita. A tabela verdade para este operador é

A	B	$A \vee B$
V	V	V
V	F	V
F	V	V
F	F	F

■ **Exemplo 1.6** Teorema com lógica **ou**

Teorema 1.2.1 Extremos de uma função

Considere uma função $f(x)$. Se a primeira derivada da função em relação a x for nula em um ponto x^*, então este ponto é de mínimo **ou** de máximo **ou** uma inflexão.

Observem que nesta afirmação temos: A como "se a primeira derivada for nula em um ponto x"e B como "este ponto é de mínimo **ou** de máximo **ou** uma inflexão". Assim, B será válido se ao menos uma condição for satisfeita (não precisamos que todas sejam satisfeitas ao mesmo tempo). ■

1.3 Provas

Uma prova é uma argumentação que mostra, de maneira irrefutável, que uma afirmação é verdadeira. Para entendermos uma prova, precisamos compreender a linguagem utilizada para escrever um texto matemático e as implicações da lógica discutida nas seções anteriores. Existem várias maneiras de provar ou de refutar uma afirmação.

1.3.1 Prova Direta

A prova direta é uma forma de mostrar que uma afirmação é verdadeira ou falsa através de uma combinação de axiomas e teoremas (nas suas mais variadas formas) já conhecidos.

■ **Exemplo 1.7** Prova de afirmação do tipo $A \Rightarrow B$
Considere a afirmação

Teorema 1.3.1 A soma de dois inteiros pares é par.

Demonstração. Sejam x e y inteiros pares: como x é par, sabemos que $2|x$. De forma análoga, $2|y$. Portanto, existem inteiros a e b que satisfazem $x = 2a$ e $y = 2b$, tal que $x + y = 2a + 2b = 2(a + b)$. Assim, existe um inteiro c tal que $x + y = 2c$ e, portanto, $c = a + b$, tal que $2|c$. ■

■

■ **Exemplo 1.8** Prova de uma afirmação $A \iff B$
Considere a afirmação

Teorema 1.3.2 Seja x um inteiro. Então x é par se e somente se $x + 1$ é impar

Demonstração. Para provarmos uma afirmação do tipo Se-e-Somente-Se, temos que realizar o raciocínio nos dois sentidos:

$A \Rightarrow B$: supondo que x é par, temos que $2|x$. Logo, existe um inteiro a tal que $x = 2a$ e, somando 1 em ambos os lados obtemos $x + 1 = 2a + 1$, que satisfaz a condição de um número ímpar.

$A \Leftarrow B$: supondo que $x + 1$ é ímpar, pela definição temos que existe um inteiro b tal que $x + 1 = 2b + 1$. Subtraindo 1 de ambos os lados, obtemos $x = 2b$, que é a condição para x ser par (por definição). ■

■

1.3.2 Refutação direta

A mesma lógica utilizada para provar uma afirmação como correta pode ser usada para provar que uma afirmação é incorreta (refutar uma afirmação).

■ **Exemplo 1.9** Refutação direta

Teorema 1.3.3 Sejam a e b inteiros. Se $a|b$ e $b|a$, então $a = b$.

Demonstração. Se $a|b$ então existe um inteiro x tal que $b = ax$. Da mesma forma, existe um inteiro y tal que $a = by$. Assim, substituindo a segunda expressão na primeira, obtemos $b = (by)x$ e considerando que b é diferente de zero, podemos dividir ambos os lados por b, tal que $xy = 1$. Este resultados pode ser obtido com $x = -1$ e $y = -1$, no entanto, nesta situação teremos $b = -1a$ e $a = -1b$, tal que $b = -a$ ou $a = -b$. ■

■

1.3.3 Refutação por Contra-Exemplo

Mais fácil do que construir é destruir. Assim, podemos refutar uma afirmação se mostrarmos que ela falha em uma determinada situação. Para a lógica matemática, basta um caso inválido para refutar a validade da afirmação como um todo (ao contrário de mostrar que funciona em uma situação específica, que não é suficiente).

■ **Exemplo 1.10** Refutação por contra-exemplo

Teorema 1.3.4 Sejam a e b inteiros. Se $a|b$ e $b|a$, então $a = b$.

Demonstração. Se escolhermos $a = -6$ e $b = 6$, temos que $-6|6$ e $6|-6$ mas obviamente $-6 \neq 6$. Assim, a afirmação é incorreta ∎

∎

2. Teoria de Conjuntos

2.1 Introdução

Um conjunto é uma coleção de objetos distintos e bem definidos. O conceito de conjuntos (*sets*) foi introduzido por Georg Cantor, que utilizou a definição

Definição 2.1.1 Conjunto (Cantor)
Um conjunto é um ajuntamento de objetos definidos e distintos (de nossa percepção e do nosso pensamento), que são chamados de elementos do conjunto. Dois conjuntos são iguais se e somente se contém os mesmos objetos.

No entanto, esta definição carece de uma formalidade matemática e pode ser considerada um axioma. O mais importante do axioma é que um conjunto tem elementos e que dois conjuntos são iguais se e somente se contém os mesmos objetos. Vamos formalizar alguns conceitos

Definição 2.1.2 Definição extensiva de um conjunto
Um conjunto A pode ser definido de forma extensiva se listarmos todos os elementos do conjunto, na forma

$$A = \{1,1,2,3,4\}.$$

Esse tipo de definição não é muito prático, servindo para conjunto com poucos elementos.

Com isso, podemos apresentar outra definição básica, como a cardinalidade de um conjunto:

Definição 2.1.3 Cardinalidade de um Conjunto (*cardinality*)
Seja um conjunto A. Definimos como cardinalidade de A, ou $|A|$, o número de membros distintos que o conjunto contém.

■ **Exemplo 2.1** Cardinalidade
O conjunto $A = \{1,1,2,3,4\}$ tem cardinalidade $|A| = 4$, pois contamos somente os elementos distintos.

■

No que segue iremos formalizar alguns conceitos fundamentais para operações com conjuntos.

2.2 Conceitos fundamentais

Assumindo que o conceito de conjunto é primitivo, assim como o de objeto e o conceito de "pertencer", indicamos que um elemento x pertence (faz parte de) um conjunto A com a notação

Definição 2.2.1 $\in$: Pertence a
Indicamos que um conjunto A contém um elemento x com a notação $x \in A$ (x pertence à A).

Com isso, podemos construir conjuntos e definir algumas operações básicas.

Definição 2.2.2 Definição compreensiva de um conjunto
Podemos definir um conjunto a partir de uma regra (ou lei)

$$A = \{x \in \mathbb{N}, x \geqslant 1 \land x \leq 1000\},$$

onde a primeira parte da definição é chamada de domínio, $x \in \mathbb{N}$, e a segunda parte é chamada de condição (observem o uso do operador lógico **e**). Esta definição é equivalente a dizer que A contém todos os números naturais entre 1 e 1000.

Outras definições muito importantes são as dos **quantificadores**, que são usados para indicar grupos de elementos ou elementos individuais de um conjunto.

Definição 2.2.3 Quantificador universal: $\forall$
Para todo, $\forall$, implica em considerarmos todos os possíveis valores de um dado conjunto, na forma

$$\forall x \in A, \text{ afirmação sobre } x$$

O quantificador universal $\forall$ pode assumir variações, como por exemplo:
- "todo inteiro é par ou ímpar"
- "todos os inteiros são pares ou ímpares"
- "cada inteiro ou é par ou é ímpar"
- "seja x um inteiro qualquer. Então x é par ou é ímpar"

Para provarmos uma afirmação com "para todo", temos que mostrar que a afirmação é sempre válida, para qualquer ocorrência de x.

Definição 2.2.4 Quantificador existencial: $\exists$
O quantificador *existe*, $\exists$, implica na existência de **ao menos uma** ocorrência do elemento, na forma

$$\exists x \in A, \text{ afirmação sobre } x$$

Este quantificador pode ser escrito nas seguintes variações:
- "Existe um número natural que é primo e é par"
- "Existe um x, membro de $\mathbb{N}$, tal que x é primo e par"

É interessante notar que para provarmos uma afirmação com Existe, basta mostrar que uma ocorrência da afirmação é verdadeira (neste caso, o número 2 é par e primo e, portanto, a afirmação é válida).

Com as definições acima, podemos então definir "estar contido"

Definição 2.2.5 $\subseteq$: Estar contido

Sendo dois conjuntos A e B, dizemos que A está contido em B ($A \subseteq B$)

$$A \subseteq B\, sse\, \forall x \in U\,, x \in A \Rightarrow x \in B$$

onde se lê: A está contido em B se e somente se, para todo x (pertencente ao universo), x pertence a A implica em x pertencer a B. O conceito de universo é um artifício para considerarmos todos os possíveis elementos existentes (no escopo do problema).

Observem que esta definição é de conhecimento geral, mas novamente utilizamos uma série de símbolos e conceitos que devem ser bem definidos.

2.2.1 Negação de quantificadores

A definição de condições mais elaboradas geralmente necessita de negação de quantificadores. Vamos estudar dois exemplos

■ **Exemplo 2.2** Negação de quantificadores
Vamos considerar a afirmação

Proposition 2.2.1 "Não existe inteiro que seja simultaneamente par e ímpar"

Podemos representar esta afirmação como

$$\neg(\exists x \in \mathbb{Z}, x \text{ é par e é ímpar})$$

que se lê como: *não* (existe x inteiro que é par e é impar). Da mesma forma, podemos representar essa mesma proposição com

$$\forall x \in \mathbb{Z}, \neg(\text{x é par e é ímpar})$$

que se lê como: para todo x inteiro *não* existe x par e ímpar. As duas afirmações são equivalentes. ■

■ **Exemplo 2.3** Negação de quantificadores

Proposition 2.2.2 "Nem todos os inteiros são primos"

Podemos representar essa proposição como

$$\neg(\forall x \in \mathbb{Z}, x \text{é primo})$$

que pode ser lida como: *não* (todo x inteiro é primo). Da mesma forma, podemos representar a mesma proposição com

$$\exists x \in \mathbb{Z}, \neg(x \text{ é primo})$$

que se lê como: existe x inteiro *não* primo. ■

2.2.2 União de quantificadores

É possível agrupar quantificadores para fins de brevidade de notação. No entanto, devemos ter muito cuidado com a sequência dos quantificadores, pois isso pode alterar completamente a lógica do resultado final. Considere os dois exemplos com cuidado:

■ **Exemplo 2.4**

Proposition 2.2.3 "Para todo o x, existe um y tal que $x+y=0$"
Podemos representar essa afirmação como

$$\forall x, \exists y, x+y=0$$

Demonstração. Seja um x arbitrário e seja $y = -x$.Então $x + y = x - x = 0$. ■

■

No entanto, uma pequena alteração pode mudar totalmente o sentido da afirmação

■ Exemplo 2.5

Proposition 2.2.4 "Existe um y, tal que para todo o x, temos x+y=0"
Podemos representar essa afirmação como

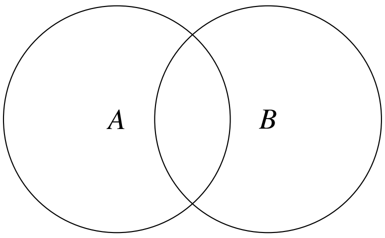
$$\exists y, \forall x, x + y = 0$$

Demonstração. Dado um $y = 2$ podemos escolher $x = 10$, que viola a condição e, portanto, invalida a afirmação ■

■

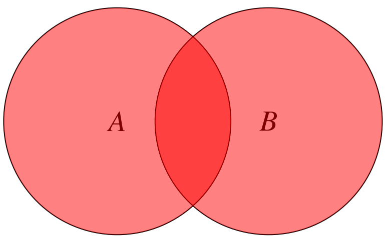
2.3 Operações com Conjuntos

Com as definições das seções anteriores, podemos focar com mais profundidade em suas operações. Para isso, vamos considerar dois conjuntos A e B:



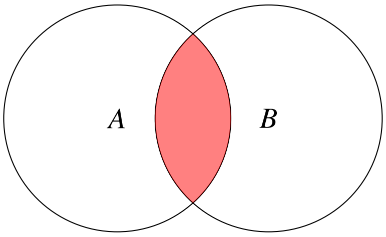
Definição 2.3.1 União de conjuntos
Sejam dois conjuntos A e B. A união de A e B é o conjunto de todos os elementos que estão em A ou em B, tal que

$$A \cup B = \{x : x \in A \lor x \in B\}$$



Definição 2.3.2 Intersecção de conjuntos
Sejam dois conjuntos A e B. A interseção de A e B é o conjunto de todos os elementos que estão em A e em B, tal que

$$A \cap B = \{x : x \in A \land x \in B\}$$



Com as seguintes propriedades (comutativas, associativas e distributivas):

1. $A \cup B = B \cup A$
2. $A \cap B = B \cap A$
3. $A \cup (B \cup C) = (A \cup B) \cup C$
4. $A \cap (B \cap C) = (A \cap B) \cap C$
5. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Em relação ao conjunto vazio, observamos que

1. $A \cup \emptyset = A$
2. $A \cap \emptyset = \emptyset$.

Para provarmos estas propriedades, podemos utilizar os conceitos de união e interseção.

■ **Exemplo 2.6** Demonstração de que $A \cup (B \cup C) = (A \cup B) \cup C$

Da definição de união entre conjuntos, temos que $B \cup C$ pode ser escrita como

$$B \cup C = \{x : x \in B \vee x \in C\}$$

tal que o lado esquerdo da afirmação que estamos tentando provar pode ser escrito como

$$A \cup (B \cup C) = \{x : (x \in A) \vee (x \in B \vee x \in C)\}.$$

Usando a propriedade associativa do operador *ou* podemos escrever

$$\{x : (x \in A) \vee (x \in B) \vee (x \in C)\}$$

que equivale a

$$\{x : (x \in A \cup B) \vee (x \in C)\}$$

ou

$$(A \cup B) \cup C.$$

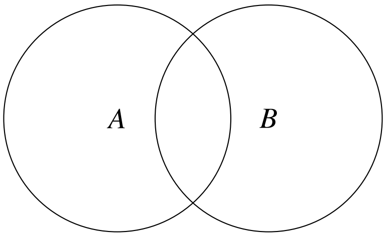
■

Observem que com este tipo de construção podemos provar várias propriedades de conjuntos. Continuando, podemos ainda definir:

Definição 2.3.3 Diferença de Conjuntos

A diferença entre dois conjuntos é denotada por $A - B$ e significa que queremos todos os elementos que pertencem a A mas não pertencem a B, tal que

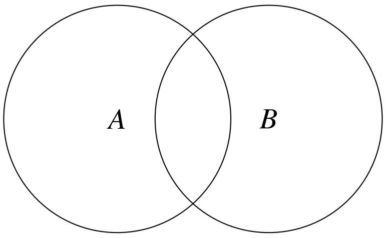
$$A - B = \{x : x \in A \wedge x \notin B\}.$$



Definição 2.3.4 Diferença Simétrica de Conjuntos

A diferença simétrica entre dois conjuntos é denotada por $A\Delta B$ e significa que queremos todos os elementos que estão em A e não em B E em B mas não em A, tal que

$$A\Delta B = (A - B) \cup (B - A).$$



Definição 2.3.5 Produto Cartesiano de Conjuntos

Sejam dois conjuntos A e B. O produto cartesiano de A e B, denotado por $A \times B$, é um novo conjunto C com todos os pares ordenados (listas de dois elementos) formados tomando-se um elemento de A juntamente com um elemento de B de todas as maneiras possíveis, tal que

$$C = A \times B = \{(a,b) : a \in A, b \in B\}$$

Com isto, verificamos que o produto cartesiano entre dois conjuntos não é uma operação comutativa ($A \times B \neq B \times A$).

■ **Exemplo 2.7** Produto cartesiano de dois conjuntos

Vamos considerar $A = \{1,2\}$ e $B = \{3,4\}$. Assim, pela definição acima, temos que

$$\begin{aligned} A \times B &= \{(1,3),(1,4),(2,3),(2,4)\} \\ B \times A &= \{(3,1),(3,2),(4,1),(4,2)\} \end{aligned} \tag{2.1}$$

Ainda, observamos que a cardinalidade do produto cartesiano é igual ao produto das cardinalidades: $|A \times B| = |A||B|$, pois $|A| = 2$ e $|B| = 2$, tal que $|A \times B| = 4$, pois o conjunto $\{(1,3),(1,4),(2,3),(2,4)\}$ tem quatro elementos (pares).O mesmo se observa para $B \times A$ ■

3. Números Reais

Os números reais, $\mathbb{R}$, são utilizados para representar quantidades contínuas e são a extensão do conjunto dos números racionais, pela consideração dos números irracionais (frações que não podem ser obtidas pela divisão de dois inteiros). Assim, temos que

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}...$$

O conjunto dos números reais é obtido por meio de alguns axiomas, conhecidos como axiomas de corpo.

Definição 3.0.1 $\mathbb{R}$

$\mathbb{R}$ é um conjunto não vazio onde podemos definir duas operações fechadas (isto é, o resultado das operações gera um elemento que também pertence ao $\mathbb{R}$):

- Adição: $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ ou $(x,y) \to x+y$;
- Multiplicação: $* : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ ou $(x,y) \to x*y$;

Observe que utilizamos duas definições que não estão formalizadas: soma e multiplicação. Vamos formalizar por meio de axiomas.

Definição 3.0.2 Axiomas da operação de soma (+) em $\mathbb{R}$

A operação de soma, +, no conjunto $\mathbb{R}$ deve respeitar:

- Associatividade: $(x+y)+z = x+y+z \quad x,y,z \in \mathbb{R}$;
- Comutatividade: $x+y = y+x \quad x,y \in \mathbb{R}$;
- Elemento Neutro: $x+0 = x \quad x,0 \in \mathbb{R}$;
- Simétrico: $x+(-x) = 0 \quad x,0 \in \mathbb{R}$;

Definição 3.0.3 Axiomas da operação de multiplicação (*) em $\mathbb{R}$

A operação de multiplicação, *, no conjunto $\mathbb{R}$ deve respeitar:

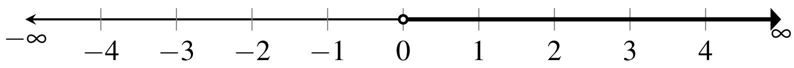
- Associatividade: $(x*y)*z = x*y*z \quad x,y,z \in \mathbb{R}$;
- Comutatividade: $x*y = y*x \quad x,y \in \mathbb{R}$;
- Distributividade: $x*(y+z) = x*y+x*z \quad x,y,z \in \mathbb{R}$;
- Elemento Neutro: $x*1 = x \quad 1,0 \in \mathbb{R}$;

- Inverso Multiplicativo: $x * x^{-1} = 1 \quad x, 1 \in \mathbb{R} \quad x \neq 0$;

Por fim, mas não menos importante, temos um conjunto de axiomas referentes ao conceito de Conjunto Ordenado (ou Corpo Ordenado), que dá origem ao conceito de inequações. Para este fim, definimos um subconjunto $P \subseteq \mathbb{R}$, que contém os elementos **positivos** de $\mathbb{R}$, atendendo aos seguintes axiomas:

Definição 3.0.4 Elementos positivos de $\mathbb{R}$

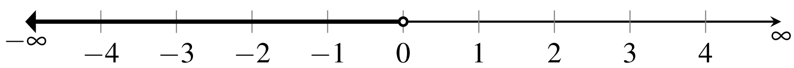
Existe um subconjunto $P \subseteq \mathbb{R}$, ou $\mathbb{R}_>$, chamado de conjunto dos números reais positivos, tal que a soma, o produto e a divisão de elementos de P está contido em P (iremos ver que isso gera um conjunto fechado).



Com isso, podemos definir o conjunto dos números negativos, $N \subseteq \mathbf{R}$

Definição 3.0.5 Elementos negativos de $\mathbb{R}$

Existe um subconjunto $N \subseteq \mathbb{R}$, ou $\mathbb{R}_<$, chamado de números reais negativos. Seja o conjunto P dos números positivos. Assim, $\mathbb{R} = 0 \cup P \cup (-P)$, onde o conjunto $N = -P$ é conhecido como conjunto dos números (reais) negativos.



É interessante notar que não temos o 0 em nenhum destes conjunto. No entanto, podemos definir os conjuntos dos números não negativos e dos números não positivos

Definição 3.0.6 Conjunto dos números não negativos e dos números não positivos

O conjunto dos números reais não negativos é definido como $\mathbb{R}_+ = 0 \cup \mathbb{R}_>$

O conjunto dos números reais não positivos é definido como $\mathbb{R}_- = 0 \cup \mathbb{R}_<$

Uma vez que ordenamos os números reais, podemos observar que

Definição 3.0.7 Desigualdade

Após definirmos os conjuntos P e $N = -P$, podemos afirmar que:

- $x > 0 \quad \text{se} \quad x \in P$;
- $x < 0 \quad \text{se} \quad x \in N$;
- $x \geq 0 \quad \text{se} \quad x \in P \cup 0$;
- $x \leq 0 \quad \text{se} \quad x \in N \cup 0$;

tal que

- $x > y \Rightarrow x - y > 0$;
- $x < y \Rightarrow x - y < 0$;
- $x \geq y \Rightarrow x - y \geq 0$;
- $x \leq y \Rightarrow x - y \leq 0$;

Um intervalo nada mais é do que um subconjunto de $\mathbb{R}$, definido entre dois valores (finitos ou infinitos). Desta forma, temos as seguintes possibilidades:

- $(a,b) = \{x \in R : x > a \wedge x < b\}$;
- $[a,b] = \{x \in R : x \geq a \wedge x \leq b\}$;
- $(a,b] = \{x \in R : x > a \wedge x \leq b\}$;
- $[a,b) = \{x \in R : x \geq a \wedge x < b\}$;

para intervalos finitos e

- $(a,+\infty) = \{x \in R : x > a\}$;
- $[a,+\infty) = \{x \in R : x \geq a\}$;

- $(-\infty, a) = \{x \in R : x < a\}$;
- $(-\infty, a] = \{x \in R : x \leq a\}$;

para intervalos infinitos.

Outra definição muito importante quando lidamos com números reais é o conceito de valor absoluto:

Definição 3.0.8 Valor Absoluto de um número real
Seja um número $a \in \mathbb{R}$. O valor absoluto (ou em módulo) de a é dado por

$$|a| = \begin{cases} +a & se \quad a > 0 \\ -a & se \quad a < 0 \end{cases}$$

Com propriedades:

1. $|-a| = |a|, \forall a \in \mathbb{R}$
2. $|ab| = |a|\,|b|, \forall a, b \in \mathbb{R}$
3. $|a| \leq k \Leftrightarrow a \in [-k, k], \forall a, k \in \mathbb{R}$

■ **Exemplo 3.1** Módulo
Se $a = -1$, temos que $|a| = -(-1) = 1$. ■

Com isso, podemos introduzir uma das definições mais importantes deste texto

Teorema 3.0.1 Desigualdade Triangular
Sejam $a, b \in \mathbb{R}$,

$$|a + b| \leq |a| + |b|$$

Demonstração. Sabendo que $-|a| \leq a \leq |a|$ e que $-|b| \leq b \leq |b|$, podemos somar estas duas definições, obtendo $-|a| - |b| \leq a + b \leq |a| + |b|$. Da definição de que $|a| \leq k \Leftrightarrow a \in [-k, k]$, verificamos que k, nesta expressão, assume a forma de $|a| + |b|$, e, portanto, a afirmação pode ser escrita como $|a + b| \leq |a| + |b|$. ■

■ **Exemplo 3.2** Desigualdade triangular
Sejam dois números reais $a = 1$ e $b = -2$. Neste caso $|1 - 2| = |-1| = 1$ e $|1| + |-2| = 1 + 2 = 3$. Como $1 < 3$, verificamos a desigualdade triangular. ■

3.1 Algumas propriedades de conjuntos de números reais.

Vamos apresentar algumas definições importantes ao trabalharmos com números reais:

Definição 3.1.1 Cota superior e cota inferior
Seja um conjunto $S \subseteq \mathbb{R}$. Dizemos que $u \in \mathbb{R}$ é cota superior de S se $\forall s \in S$, $s \leq u$. Analogamente, dizemos que $v \in \mathbb{R}$ é cota inferior de S se $\forall s \in S$, $v \leq s$.

■ **Exemplo 3.3** Cota inferior
Considere o conjunto dos números naturais. Este conjunto tem 0 como cota inferior. No entanto, não tem cota superior. Aqui, podemos argumentar que todos os reais na faixa $(-\infty, 0]$ são cotas inferiores, mas se nos restringirmos aos naturais, restaria apenas o 0. ■

■ **Exemplo 3.4** Cota superior
Considere o conjunto $A = \{x \in \mathbb{R}, x \leq 0\}$. Esse conjunto tem todos os números no intervalo $[0, \infty)$ como cota superior, mas não tem cota inferior. ■

Com a definição acima, podemos definir

Definição 3.1.2 Conjunto Limitado por baixo e Conjunto Limitado por Cima
Se um conjunto S tem cota inferior, dizemos que ele é limitado por baixo ou limitado inferiormente. Analogamente, se o conjunto tem cota superior, dizemos que ele é limitado por cima ou limitado superiormente.

Devemos ter cuidado com um detalhe da definição de Cotas: estamos trabalhando com elementos s de um conjunto $S \subseteq \mathbb{R}$ e avaliando as cotas no $\mathbb{R}$. Assim, um conjunto limitado por cima pode ter diversas cotas superiores e um conjunto limitado por baixo pode ter diversas cotas inferiores.

Exemplo 3.5 Multiplas cotas
Seja $A = [0,1]$. Qualquer número real maior ou igual a 1 será uma cota superior para o conjunto e qualquer número menor ou igual a 0 será cota inferior ■

Com isto em mente, podemos definir os conceitos de Supremo e de Ínfimo:

Definição 3.1.3 Supremo de um conjunto / Ínfimo de um conjunto
Se um conjunto S é limitado por cima, chamamos de supremo de S ou $\sup(S)$ a menor de suas cotas superiores. Analogamente, se S é limitado por baixo, chamamos de ínfimo ou $\inf(S)$ a maior de suas cotas inferiores.

Exemplo 3.6 Supremo e Ínfimo - intervalo fechado
Seja $A = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$. O conjunto A contém infinitas cotas superiores, pois qualquer número real na faixa $[1,\infty)$ satisfaz a definição de cota superior. No entanto, 1 é a menor das cotas superiores e, portanto, $\sup(A) = 1$. A também contém infinitas cotas inferiores na faixa $(-\infty,0]$, tal que $\inf(A) = 0$. Neste exemplo, podemos afirmar que $\sup(A) \in A$ e $\inf(A) \in A$.

■

Exemplo 3.7 Supremo e Ínfimo - intervalo aberto
Seja $B = \{x \in \mathbb{R} : 0 \leq x < 1\}$. O que muda neste exemplo, em relação ao anterior, é o fato de o limite superior do intervalo ser definido por um intervalo aberto. Neste caso, as cotas superiores estarão na faixa $(1,\infty)$ e o supremo será o menor valor deste intervalo. $sup(B) = 1$, mas com uma diferença que é muito importante: $sup(B) \notin B$. ■

4. Funções

Antes de definirmos o conceito de função, vamos definir uma relação:

Definição 4.0.1 Relação
Uma relação R é um conjunto de pares ordenados

$$R = \{(a,b), a \in A \wedge b \in B\}$$

Com isso, podemos definir uma função como:

Definição 4.0.2 Função
Uma relação f é chamada de função desde que $(a,b) \in f$ e $(a,c) \in f$ impliquem em $b = c$.

■ **Exemplo 4.1** Relação que é uma função
Seja uma relação $R = \{(1,2),(3,4),(5,6)\}$. Essa relação é uma função, e podemos ver cada par como sendo $(entrada, saida)$ da função. ■

Assim, o que a definição de função está dizendo, é que se aplicarmos as entradas da função e observarmos as saídas, não poderemos ver um resultado diferente para uma mesma entrada.

■ **Exemplo 4.2** Relação que não é uma função
A relação $R = \{(1,2),(1,4),(2,7)\}$ não é uma função, pois para uma mesma entrada, 1, tivemos duas saídas diferentes: 2 e 4.

■

Assim, de forma alternativa, poderíamos escrever:

Definição 4.0.3 Função (definição alternativa)
Uma relação f é chamada de função desde que $f(a) = b$ e $f(a) = c$ impliquem em $b = c$.

Como uma função está associada a conjuntos, podemos definir uma função compreensivamente

$$f = \{(x,y) : x, y \in \mathbb{Z}, y = x^2\}$$

que, em termos de pares ordenados teria a forma

$$f = \{(1,1),(2,4),(3,9),(4,16),...\}.$$

Definição 4.0.4 Funções Multi-valoradas (Polídromas)
Algumas "funções" como por exemplo raiz quadrada ou funções trigonométricas inversas apresentam mais de um valor para o mesmo argumento de entrada. Neste caso, mesmo sendo muito comuns, não podemos utilizar a nomenclatura de função. O correto seria chamar de relação, conforme vimos anteriormente. Contribuição de *Barbara Haensch Schneider*.

É importante entender o conceito de função e como ela se relaciona com a teoria de conjuntos que revisamos anteriormente. Para isto, vamos estudar as seguintes definições:

Definição 4.0.5 Domínio de uma função
Seja uma função f. O conjunto de todos os primeiros elementos possíveis dos pares ordenados de f é chamado de domínio de f ou $dom\,f$

$$dom\,f = \{a : \exists b, (a,b) \in f\}$$

ou, de forma mais direta:

$$dom\,f = \{a : f(a)\,\text{definido}\}$$

■ **Exemplo 4.3** Domínio de uma função
Na definição $f = \{(x,y) : x,y \in \mathbb{Z}, y = x^2\}$ o domínio é $\mathbb{Z}$. ■

Definição 4.0.6 Imagem e contradomínio de uma função
Seja uma função f. O conjunto de todos os possíveis segundos elementos dos pares ordenados de f é chamada de imagem de f ou $im\,f$

$$im\,f = \{b : \exists a, (a,b) \in f\}$$

ou, de forma mais direta:

$$im\,f = \{b : b = f(a)\,\text{para algum}\ a \in dom\,f\},$$

ou mais diretamente ainda, a imagem é a própria saída da função. O contradomínio da função, por sua vez, é um conjunto mais amplo, que contém a imagem.

■ **Exemplo 4.4** Imagem e contradomínio
A imagem de $f = \{(x,y) : x,y \in \mathbb{Z}, y = x^2\}$ são os quadrados perfeitos e o contradomínio é $\mathbb{Z}$. ■

Com isso, podemos então definir:

Definição 4.0.7 $f : A \to B$
Seja f uma função e sejam A e B conjuntos. Dizemos que f é uma função de A para B se $dom\,f = A$ e se $im\,f \subseteq B$, onde B é o contradomínio. Alternativamente, dizemos que f é uma aplicação de A em B.

■ **Exemplo 4.5** $f : A \to B$
A função seno é uma função que mapeia o conjunto dos reais no conjunto dos reais, pois $f = \{(x,y) : x,y \in \mathbb{R}, y = sen(x)\}$, e, portanto, temos que $dom\,f = \mathbb{R}$ e $im\,f = [-1,1] \subseteq \mathbb{R}$. ■

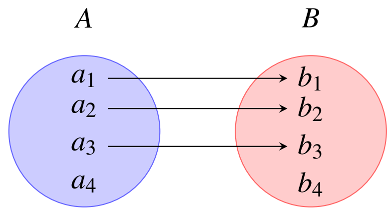
Definição 4.0.8 Inversa de uma função
Seja uma função $f : A \to B = \{(a_1,b_1),(a_2,b_2),....\}$. A inversa da função, denotada por f^{-1} é obtida pela inversão dos termos de cada um dos pares que formam a relação, tal que

$$f^{-1} = \{(b_1,a_1),(b_2,a_2),...\}$$

É importante salientar que a inversa da função $f : A \to B$ não é necessariamente uma função $f^{-1} : B \to A$. Para entendermos isto, vamos considerar os conjuntos $A = \{1,2,3,6,8\}$ e $B = \{2,4,5,4,7\}$ e a função $f : A \to B = \{(1,4),(2,5),(3,4),(6,2),(8,5)\}$. Neste caso, a inversa de f será dada por $f^{-1} = \{(4,1),(5,2),(4,3),(2,6),(5,8)\}$ que não é uma função pois temos os pares $(4,1)$ e $(4,3)$, bem como $(5,2)$ e $(5,8)$. Ainda, temos que $dom\, f^{-1} = \{4,5,2\} \neq B$.

A existência de uma inversa que é uma função pode ser estudada com o auxílio das seguintes definições:

Definição 4.0.9 Função injetiva (um a um)
Uma função $f : A \to B$ é injetiva se mapeia cada valor distinto de A em um valor distinto de B, isso é, se $(x,b),(y,b) \in f \Leftrightarrow x = y$ ou, de forma equivalente, $f(x) = f(y)$ sse $x = y$.



■ **Exemplo 4.6** Função injetiva
Um exemplo bastante simples de uma função um-a-um é um polinômio linear com forma $ax + b$. Se a função for injetiva, temos que satisfazer a condição

$$ax + b = ay + b$$

e, se $a \neq 0$, podemos subtrair b dos dois lados e depois dividir por a. Neste caso, provamos que $x = y$. ■

■ **Exemplo 4.7** Função não injetiva
Um exemplo de uma função que não é injetiva é a função $f = \{(x,y) : x,y \in \mathbb{R}, y = x^2\}$. Neste caso, teríamos que verificar se igualdade

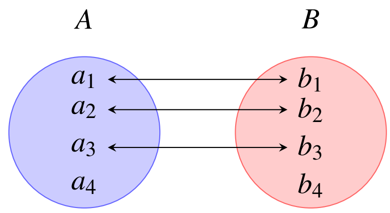
$$a^2 = b^2$$

implica em $a = b$, o que obviamente não ocorre para $\forall a,b \in \mathbb{R}$ (podemos pensar no caso óbvio onde $a = -1$ e $b = 1$). ■

Exercício 4.1 Verifique se $f = \{(x,y) : x,y \in \mathbb{N}, y = x^2\}$ é injetiva. ■

Com esta definição, podemos então afirmar que:

Definição 4.0.10 Funções injetivas possuem função inversa
Seja f uma função. A relação inversa f^{-1} é uma função se e somente se f for injetiva. Neste caso, temos que $dom\, f = im\, f^{-1}$e $im\, f = dom\, f^{-1}$.

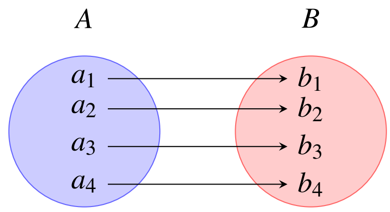


A seguinte definição também é muito importante para entendermos vários resultados futuros:

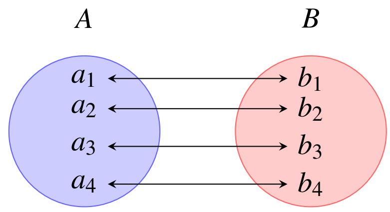
Definição 4.0.11 Função Sobre (B)
Uma função $f : A \to B$ é dita sobre B desde que, para todo $b \in B$, exista um $a \in A$ tal que $f(a) = b$, ou seja, $im\, f = B$

$$\forall b \in B, \exists a \in A, f(a) = b$$

Portanto, para que a função seja sobre, todo elemento de B deve ter um correspondente em A, ou de forma mais direta, temos que acessar todos os elementos de B.



Definição 4.0.12 Bijeção
Seja $f : A \to B$. Se f for injetiva e sobre B, dizemos que é uma bijeção. Com isso, podemos afirmar que $f^{-1} : B \to A$.



■ **Exemplo 4.8** Bijeção
Seja A o conjunto dos números inteiros pares e B o conjunto dos inteiros ímpares. Vamos mostrar que $f(x) = x + 1$ é uma bijeção. Primeiro, vamos mostrar que f é um-a-um. Para isto, basta mostrarmos que

$$x + 1 = y + 1$$

e, descontando 1 dos dois lados, mostramos que

$$x = y.$$

Para verificarmos se é sobre, temos que selecionar um elemento de B, que pode ser escrito na forma genérica como $y = 2k + 1$. Por definição, um elemento de A pode ser escrito na forma $x = 2k$, tal que a função a ser estudada pode ser escrita na forma

$$f(x) = x + 1 = 2k + 1$$

que mostra que f é sobre e, portanto f é uma bijeção. ■

Por fim, podemos definir um princípio muito importante e que facilita muito essas análises quando os conjuntos A e B são finito-dimensionais:

Definição 4.0.13 Princípio da casa do Pombo
Sejam A e B conjuntos finitos e seja $f : A \to B$.

- Se $|A| > |B|$, então f não é injetiva (um-a-um).
- Se $|A| < |B|$, então f não é sobre.

Demonstração. $|A| > |B| \Rightarrow$ não é injetiva
Se $|A| > |B|$ então os primeiros $|A|$ elementos de A serão levados para diferentes elementos de B. Após isto, não há elementos em B aos quais podemos aplicar os elementos restantes de A. Em outras palavras, se os elementos de A são pombos e os de B são casas, teremos mais pombos do que casas e mais de um pombo em algumas casas. ■

Demonstração. $|A| < |B| \to$, f não é sobre
Se $|A| < |B|$, então sabemos que não existem elementos suficientes em A para cobrir todos os elementos em B. Novamente em se tratando de pombos, não teremos elementos suficientes em A para ocupar todas as casas em B. ■

Exercício 4.2 Preencha a tabela e justifique cada uma das suas respostas.

Domínio	Contradomínio	f	um a um ?	sobre ?	Bijeção ?
$\mathbb{Z}$	$\mathbb{Z}$	$x \to x^3$			
$\mathbb{R}$	$\mathbb{R}$	$x \to x^3$			
$\mathbb{Z}$	$\mathbb{N}$	$x \to \|x\|$			
$\mathbb{Z}$	$\mathbb{Z}$	$x \to x^2$			

■

4.1 Continuidade

O conceito de continuidade é fundamental para estudarmos procedimentos de solução utilizados na mecânica computacional.

Definição 4.1.1 Continuidade de uma Função
Seja $f : X \to Y$. Esta função é dita contínua em $x_0 \in X$ se para cada $\varepsilon > 0$ existe $\delta > 0$ tal que

$$|f(x_0) - f(\mathbf{x})| < \varepsilon$$

onde

$$|x_0 - \mathbf{x}| < \delta.$$

Ou seja, se nos deslocarmos um valor δ no entorno do ponto $\mathbf{x}_0$ (mas que não depende de x_0), então o valor da função não terá variação maior do que ε.
A função f é dita contínua se for satisfizer esse requisito em todos os pontos de X.

■ **Exemplo 4.9** Seja a função $f : \mathbb{R} \to [-1, 1]$ dada por $f(x) = \cos(x)$. Essa função é contínua se satisfizer

$$|\cos(x_0) - \cos(x_0 \pm \delta)| < \varepsilon.$$

Podemos reescrever $\cos(x_0+\delta) = \cos(x_0)\cos(\delta) - \sin(x_0)\sin(\delta)$ e sabemos que $\delta \to 0 \implies \sin(\delta) \to 0$ e que $\delta \to 0 \implies \cos(\delta) \to 1$. Assim, podemos reescrever a condição como

$$|\cos(x_0) - \cos(x_0)| = 0 < \varepsilon, \quad \forall x_0 \in \mathbb{R}, \varepsilon > 0.$$

■

■ **Exemplo 4.10** Seja a função $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = \frac{1}{x}$.

Essa função será contínua no domínio $\mathbb{R}$ se

$$|f(x_0) - f(x_0 \pm \delta)| < \varepsilon.$$

No entanto, podemos ver que em $x_0 = 0$ pela esquerda, $x_0 - \delta$, teremos

$$f(0^-) = \frac{1}{0^-} = -\infty$$

e em $x_0 = 0$ pela direita, $x_0 + \delta$,

$$f(0^+) = \frac{1}{0^+} = \infty$$

tal que em $x_0 = 0$ temos uma descontinuidade, pois o lado da esquerda da nossa condição irá para $|\infty|$ que sempre será maior do que ε. Basta um ponto para que a função seja dita descontínua em seu domínio, que no caso é todo o $\mathbb{R}$, de tal forma que podemos afirmar que essa função é descontínua no domínio estipulado.

De maneira mais formal, podemos escolher um ponto x no entorno de x_0 tal que

$$\left|\frac{1}{x_0} - \frac{1}{x}\right| = \frac{|x-x_0|}{|x_0||x|} < \varepsilon$$

com $|x-x_0| < \delta$ tal que

$$\frac{\delta}{|x_0||x|} < \varepsilon$$

ou

$$\delta < \varepsilon|x_0||x|$$

de onde já podemos verificar que se $x_0 = 0$, violamos a condição de que $\delta > 0$, mostrando que a função não é contínua em 0. Para outros pontos que não o 0, podemos afirmar que

$$\delta < \varepsilon|x_0||x_0+\delta|$$

pode ser "aproximado" por (isso não está formalmente correto, pois deveríamos trabalhar com as desigualdades provocadas pelos valores absolutos)

$$\delta < \varepsilon\left(\pm|x_0|^2 \pm |x_0|\delta\right)$$

e, dividindo por $\pm\delta$

$$|1| < \pm\frac{\varepsilon}{\delta}|x_0|^2 \pm \varepsilon|x_0|.$$

Aqui vemos que essa expressão ainda mostra a descontinuidade em $x_0 = 0$, pois a expressão indica que

$$|1| < 0,$$

o que não é possível. No entanto, para $x_0 \neq 0$, podemos escolher valores de δ suficientemente pequenos para que o lado direito da desigualdade seja sempre maior (em módulo) do que 1, de tal forma que a função é contínua para $x_0 \neq 0$.

■

■ **Exemplo 4.11** Seja a função $f:(0,\infty)\to\mathbb{R}$ dada por $f(x)=\frac{1}{x}$.

Essa é a mesma função discutida no exemplo anterior, mas com um domínio restrito a valores maiores do que zero. Desta forma, conforme discutido no exemplo anterior, a função é contínua no domínio considerado.

■

4.2 Composição de Funções

Sejam duas funções $f:A\to B$ e $g:B\to C$. É possível definirmos uma função que mapeie diretamente de A para C, com notação

$$(g\circ f)(a)=g[f(a)],\, a\in A$$

e que chamamos de composição de g e f. Observem que a notação indica que aplicamos $f(a)$ (f está mais próximo de a) e depois aplicamos g no resultado desta operação. Ainda, temos que

$$dom\,(g\circ f)(a)=dom\,f$$

pois não estamos modificando a entrada do processo. O mais importante diz respeito ao acoplamento entre f e g, pois a saída de f deve ser uma entrada válida para g. Isto pode ser garantido com

$$im\,f\subseteq dom\,g$$

e, muito importante, devemos lembrar que a composição de funções não satisfaz a propriedade comutativa (isso implicará, em alguns capítulos, na mesma propriedade no produto de matrizes)

$$(g\circ f)\neq(f\circ g).$$

No entanto, a composição satisfaz a propriedade associativa:

Definição 4.2.1 Composição de funções: associatividade

Sejam $f:A\to B$, $g:B\to C$ e $h:C\to D$, então

$$h\circ(g\circ f)=(h\circ g)\circ f$$

■ **Exemplo 4.12** Composição de funções

Sejam $f:\mathbb{Z}\to\mathbb{Z}$ e $g:\mathbb{Z}\to\mathbb{Z}$, com $f(x)=x^2+1$ e $g(x)=2x-3$. Para calcularmos $(g\circ f)(4)$, procedemos da seguinte forma:

$$f(4)=(4)^2+1=17$$

$$g(f(4))=2*17-3=31$$

ou

$$(g\circ f)=2(x^2+1)-3=2x^2+2-3=2x^2-1.$$

■

■ **Exemplo 4.13** Composição de funções Com as mesmas funções do exemplo anterior, podemos ver que

$$(f\circ g)=(2x-3)^2+1=4x^2-12x+10$$

de modo que a composição de funções não satisfaz a propriedade comutativa. ■

■ **Exemplo 4.14** Sejam $A = \{1,2,3,4,5\}$, $B = \{6,7,8,9\}$ e $C = \{10,11,12,13,14\}$. As funções $f : A \to B$ e $g : B \to C$ são definidas por

$$f : \{(1,6),(2,6),(3,9),(4,7),(5,7)\}$$

$$g : \{(6,10),(7,11),(8,12),(9,13)\}$$

então

$$(g \circ f) = \{(1,10),(2,10),(3,13),(4,11),(5,11)\}$$

onde utilizamos o seguinte raciocínio: $(1,6)$ é um par de f na forma (a,b) que serve como entrada para g que tem a forma (b,c). Assim, a composição retorna um par na forma $(a,(,c))$ onde o b deve ser igual. Assim, $b = 6$ está presente em f nos pares $(1,6)$ e $(2,6)$, tal que isto gera duas saídas em g, pois temos $(1,(,10))$ e $(2,(,10))$.

■

Definição 4.2.2 Função Identidade em um conjunto
Seja A um conjunto. A função identidade em A é a função id_A cujo domínio é A e satisfaz

$$id_A = \{(a,a), \forall a \in A\}$$

os seja, a imagem é sempre igual ao domínio. As seguintes propriedades são observadas:

1. Sejam dois conjuntos A e B e uma função $f : A \to B$, então

$$f \circ id_A = id_B \circ f = f$$

2. Sejam dois conjuntos A e B e uma bijeção $f : A \to B$, então

$$f \circ f^{-1} = id_B$$

$$f^{-1} \circ f = id_A$$

■ **Exemplo 4.15** Identidade
Considere os conjuntos $A = \{1,2,3\}$ e $B = \{2,4,6\}$ e a função

$$f = \{(x,y) : x \in A, y \in B, y = 2x\} = \{(1,2),(2,4),(3,6)\}.$$

Neste caso, temos que

$$id_A = \{(1,1),(2,2),(3,3)\}$$

$$id_B = \{(2,2),(4,4),(6,6)\}$$

e a inversa da função é tal que

$$f^{-1} = \{(2,1),(4,2),(6,3)\} = \{(y,x) : y \in B, x \in A, x = \frac{y}{2}\}.$$

Assim, ao aplicarmos as definições acima, observamos que

$$f^{-1} \circ f = \{(1,1),(2,2),(3,3)\} = id_A$$

$$f \circ f^{-1} = \{(2,2),(4,4),(6,6)\} = id_B$$

e

$$f \circ id_A = \{(1,2),(2,4),(3,6)\}$$

$$id_B \circ f = \{(1,2),(2,4),(3,6)\}.$$

Devemos observar que as funções identidade podem ser definidas como

$$id_A = \{(x,x) : x \in A\}$$

$$id_B = \{(y,y) : y \in B\}$$

tal que

$$f \circ id_A = f(x) = 2x$$

$$id_B \circ f = y \circ (2x) = 2x$$

e

$$f^{-1} \circ f = \frac{1}{2}(2x) = id_A$$

$$f \circ f^{-1} = 2(\frac{y}{2}) = id_B.$$

■

5. $\mathbb{R}^n$

Para iniciarmos os nossos estudos sobre espaços vetoriais, iremos abordar o $\mathbb{R}^n$. No $\mathbb{R}^n$ os elementos são chamados de vetores e são formados por uma coleção de números reais, como por exemplo $(2,1)$ no $\mathbb{R}^2$ ou $(1,2,3)$ no $\mathbb{R}^3$ (que em duas e três dimensões permitem uma representação gráfica que já é de conhecimento de qualquer aluno de graduação). É importante que os conceitos aprendidos aqui sejam bem fixados, pois iremos re-utilizar estes conceitos quando formos estudar os espaços de funções, mais adiante. De maneira mais formal

Definição 5.0.1 $\mathbb{R}^n$

$\mathbb{R}^N$ é o conjunto das n-úplas ordenadas de números reais, na forma

$$\mathbb{R}^n = \{\mathbf{u} = (u_1, u_2, u_3, \ldots.., u_n) : u_i \in \mathbb{R}, i = 1, \ldots, n\}$$

Cada elemento $\mathbf{u} \in \mathbb{R}^n$ é chamado de um **vetor**. Neste texto tratamos vetores como colunas, tal que a dimensão é $n \times 1$ (n linhas e 1 coluna).

Observe que a definição acima nada mais é do que a definição de um conjunto, não definindo nenhum tipo de operação entre os elementos deste conjunto. Chamamos de um **espaço vetorial** um conjunto que é equipado, com no mínimo, um conjunto de operações binárias. Vamos definir um espaço vetorial sobre os reais:

Definição 5.0.2 Espaço vetorial sobre os reais

Um espaço vetorial V sobre os reais é um conjunto cujos elementos chamamos de vetores, com duas operações binárias: soma e multiplicação por um escalar, tais que:

- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}, \forall \mathbf{u}, \mathbf{v} \in V$
- $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}), \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$
- $\exists \mathbf{0} \in \mathbf{V} : \mathbf{0} + \mathbf{u} = \mathbf{u}, \forall \mathbf{u} \in V$
- $\forall \mathbf{u} \in V, \exists \mathbf{v}, \mathbf{u} + \mathbf{v} = \mathbf{0}$
- $\exists \mathbf{I} \in V : \mathbf{I}\mathbf{u} = \mathbf{u}, \forall \mathbf{u} \in V$
- $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}, \forall \alpha, \beta \in \mathbb{R}, \forall \mathbf{u} \in V$
- $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}\ \forall \alpha, \beta \in \mathbb{R}, \forall \mathbf{u} \in V$

- $\alpha(\mathbf{u}+\mathbf{v}) = \alpha\mathbf{u}+\alpha\mathbf{v}, \forall\alpha \in \mathbb{R}, \forall \mathbf{u},\mathbf{v} \in \mathbf{V}$

É importante salientar que $\mathbf{0}$ é um vetor com todos os termos nulos e $\mathbf{I}$ é uma matriz (Identidade) com dimensão $n \times n$.

Exercício 5.1 Verifique todas as afirmações acima considerando o $\mathbb{R}^2$ ■

Essas operações implicam em um espaço vetorial em que podemos somar elementos e também aplicar um fator de escala, que é a multiplicação por um escalar real. No entanto, diversas medidas que são usuais para nós, como ângulos e comprimentos (tamanho), requerem a definição de operações adicionais. Nem todos os espaços possuem essas operações, mas o $\mathbb{R}^n$ é um espaço "bem comportado".

Uma das operações mais importantes na álgebra vetorial é o produto interno, ou produto escalar:

Definição 5.0.3 Produto interno

Seja V espaço vetorial sobre os reais. Um produto interno é uma função de $V \times V \to \mathbb{R}$, com notação $\mathbf{u},\mathbf{v} \to \mathbf{u}\cdot\mathbf{v}$ ou $< \mathbf{u},\mathbf{v} >$ com as seguintes propriedades:

- $\mathbf{u}\cdot\mathbf{u} > \mathbf{0}, \forall \mathbf{u} \in V, \mathbf{u} \neq \mathbf{0}$
- $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}, \forall \mathbf{u},\mathbf{v} \in \mathrm{V}$
- $(\alpha\mathbf{u})\cdot\mathbf{v} = \alpha(\mathbf{u}\cdot\mathbf{v}), \forall\alpha \in \mathbb{R}, \forall \mathbf{u},\mathbf{v} \in V$
- $(\mathbf{u}+\mathbf{v})\cdot\mathbf{w} = \mathbf{u}\cdot\mathbf{w}+\mathbf{v}\cdot\mathbf{w}, \forall \mathbf{u},\mathbf{v},\mathbf{w} \in V$

Sendo assim, é importante verificar que os axiomas que definem um produto interno são bastante gerais e podem ser atendidos por operações diferentes. A mais conhecida é chamada de *produto interno canônico* no $\mathbb{R}^n$:

Definição 5.0.4 Produto interno canônico no $\mathbb{R}^n$

Sejam dois vetores $\mathbf{u}$ e $\mathbf{v} \in \mathbb{R}^n$. O produto interno canônico é definido como

$$\mathbf{u}\cdot\mathbf{v} = \sum_{i=1}^{n} u_i v_i$$

Exercício 5.2 Produto interno canônico

Verifique se a definição acima atende aos axiomas de produto interno. ■

Outro conceito importante é o de norma. A norma nada mais é do que uma “medida” de um elemento do espaço (vetor), sendo que, assim como o produto interno, tem uma definição geral:

Definição 5.0.5 Norma

Dado um espaço vetorial V, uma norma é uma função de $V \to \mathbb{R}$, denotada por $\mathbf{u} \to ||\mathbf{u}||$ e que satisfaz as seguintes condições:

- $||\mathbf{u}+\mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||, \forall \mathbf{u},\mathbf{v} \in V$ (desigualdade triangular)
- $||\alpha\mathbf{u}|| = |\alpha|\,||\mathbf{u}|| \forall \mathbf{u} \in V, \forall\alpha \in \mathbb{R}$
- $||\mathbf{u}|| > 0, \forall \mathbf{u} \in V, \mathbf{u} \neq \mathbf{0}$

Quando um espaço vetorial tem uma norma associada, este é dito espaço normado (veremos isto em mais detalhes no capítulo sobre espaço de funções). Relembrando o $\mathbb{R}^2$, temos como norma usual (Euclidiana)

$$||\mathbf{u}|| = \sqrt{u_1^2 + u_2^2}$$

que satisfaz os requisitos listados na definição da norma. No entanto, podemos estender o conceito para

$$\|\mathbf{u}\|_P = \left(\sum_{i=1}^{N} |u_i|^P\right)^{\frac{1}{P}}, p \in \mathbb{N}^*, p\, par$$

que gera a norma usual (também chamada de norma-2 ou $\|\|_2$) quando $p = 2$ e infinitas normas. Um caso interessante é a norma com $p \to \infty$, também chamada de max, pois

$$\|\mathbf{u}\|_\infty = \max_{1 \leq i \leq N} |u_i|$$

retornando o maior valor (em módulo) do vetor.

É interessante notar que com o conceito de norma podemos definir o conceito de distância entre dois vetores:

Definição 5.0.6 Distância entre dois vetores

Seja V um espaço vetorial normado e $\mathbf{u}$,$\mathbf{v}$ dois elementos deste espaço. Definimos como a distância entre $\mathbf{u}$ e $\mathbf{v}$ o escalar

$$d(\mathbf{u},\mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Um caso interessante é quando podemos associar a norma ao produto interno. Considerando a norma 2 no $\mathbb{R}^n$, podemos escrever

$$\|\mathbf{u}\|_2 = \sqrt{\mathbf{u} \cdot \mathbf{u}}$$

o que permite definir a **desigualdade de Cauchy-Schwartz**

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|_2 \|\mathbf{v}\|_2$$

5.1 Bases

Um conceito fundamental quando lidamos com espaços vetoriais é o conceito de Base, que está relacionado ao conceito de vetores linearmente independentes.

Assim, é interessante iniciarmos com alguns conceitos básicos.

Definição 5.1.1 Combinação Linear

Sejam $V \subseteq \mathbb{R}^n$ um espaço vetorial e $\mathbf{u}$ um membro deste espaço. Se $\mathbf{u}$ puder ser escrito na forma

$$\mathbf{u} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$

onde $\alpha_i \in \mathbb{R}$ e $\mathbf{v}_i \in V$, então dizemos que $\mathbf{u}$ é obtido por meio de uma combinação linear de outros membros do espaço.

■ **Exemplo 5.1** Combinação linear

Sejam $\mathbf{v}_1 = (1,2)$ e $\mathbf{v}_2 = (-1,1)$. O vetor $\mathbf{u} = (-1,7)$ é uma combinação linear de $\mathbf{v}_1$ e $\mathbf{v}_2$ pois

$$(-1,7) = 2(1,2) + 3(-1,1) = (2-3, 4+3)$$

tal que $\alpha_1 = 2$ e $\alpha_2 = 3$. ■

Definição 5.1.2 Vetores Linearmente Dependentes e Linearmente Independentes
Sejam dois elementos $\mathbf{u}$ e $\mathbf{v} \in V \subseteq \mathbb{R}^n$. Se $\mathbf{u}$ não puder ser escrito por meio de uma combinação linear de $\mathbf{v}$ então dizemos que $\mathbf{u}$ e $\mathbf{v}$ são *linearmente independentes*. Alternativamente, se $\mathbf{u}$ puder ser escrito por meio de uma combinação linear de $\mathbf{v}$, dizemos que $\mathbf{u}$ e $\mathbf{v}$ são *linearmente dependentes*.

■ **Exemplo 5.2** Vetores linearmente independentes
Sejam $\mathbf{v}_1 = (1,2)$ e $\mathbf{v}_2 = (-1,1)$. Não existe $\alpha \in \mathbb{R}$ que satisfaça

$$(-1,2) = \alpha(-1,1)$$

tal que os vetores são linearmente independentes. ■

■ **Exemplo 5.3** Vetores linearmente dependentes
Sejam $\mathbf{v}_1 = (2,2)$ e $\mathbf{v}_2 = (-1,-1)$. É possível escrever

$$(2,2) = -2(-1,-1)$$

tal que os vetores são linearmente dependentes. ■

Com isso, podemos definir

Definição 5.1.3 Conjunto Linearmente Dependente e Linearmente Independente
Seja $V \subseteq \mathbb{R}^n$ um espaço vetorial e S um conjunto que contém m elementos de V. Se nenhum elemento de S puder ser escrito por meio de uma combinação linear dos outros elementos de S, então o conjunto é dito *linearmente independente*. Neste caso, temos que

$$\alpha_1\mathbf{u}_1 + ... + \alpha_m\mathbf{u}_m = \mathbf{0}$$

se e somente se $\alpha_1 = \alpha_2 = ... = \alpha_m = 0$. Do contrário, o conjunto é dito *linearmente dependente*.

■ **Exemplo 5.4** Conjunto linearmente independente
Seja o conjunto $V = \{(1,1),(1,-1),(1,2)\}$. Observamos que

$$\alpha_1(1,1) + \alpha_2(1,-1) + \alpha_3(1,2) = (0,0)$$

só é verificado se $\alpha_1 = \alpha_2 = \alpha_3 = 0$, tal que o conjunto é linearmente independente.

■

■ **Exemplo 5.5** Conjunto linearmente dependente
Seja o conjunto $V = \{(1,1),(2,2)\}$. Observamos que

$$\alpha_1(1,1) + \alpha_2(2,2) = (0,0)$$

é verificado se $\alpha_1 = -2$, $\alpha_2 = 1$, tal que o conjunto é linearmente dependente.

■

Definição 5.1.4 Conjunto gerador ou *span*
Seja $V \subseteq \mathbb{R}^n$ um espaço vetorial e seja $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ um conjunto linearmente independente em V. Se cada vetor em V puder ser escrito como uma combinação linear dos elementos de S, então dizemos que S *spans* V, ou S gera V. Assim,

$$\mathbf{v} = \sum_{\mathbf{i=1}}^{\mathbf{n}} \alpha_\mathbf{i}\mathbf{u}_\mathbf{i}$$

onde $\mathbf{v} \in V$.

Tendo em mente esses conceitos, podemos selecionar um subconjunto B de V, que seja linearmente independente. Por exemplo, no $\mathbb{R}^2$ temos como vetores linearmente independentes o conjunto $B = \{(1,0),(0,1)\}$. Estes vetores são L.I, pois não é possível gerar $(1,0) = \alpha(0,1)$ para quaisquer valores de $\alpha \in \mathbb{R}$. No entanto, outros vetores de $\mathbb{R}^2$ podem ser obtidos a partir de uma combinação linear de elementos de B, como por exemplo:

- $(2,4) = 2(1,0)+4(0,1)$
- $(33.5,-100) = 33.5(1,0)-100(0,1)$.

Assim, podemos verificar que o conjunto $B \subseteq \mathbb{R}^2$ contém os elementos que servem para construir quaisquer outros elementos de $\mathbb{R}^2$, de tal forma que B é conhecido como base do espaço $\mathbb{R}^2$(neste caso particular, também chamada de base canônica)

Definição 5.1.5 Conjunto de Base

Seja $V \subseteq \mathbb{R}^N$ um espaço vetorial e B um conjunto de V . B é dito base de V se:

- B é linearmente independente
- B spans V

É importante verificar que:

- um espaço pode ter mais de uma base
- a cardinalidade de B define a dimensão do espaço
- diferentes bases de um mesmo espaço terão a mesma cardinalidade

■ **Exemplo 5.6** Base

Utilizando o $\mathbb{R}^2$ por uma questão de simplicidade, podemos verificar que os conjuntos

$$
\begin{aligned}
B_1 &= \{(1,0),(0,1)\} \\
B_2 &= \{(3,1),(-2,1)\}
\end{aligned} \tag{5.1}
$$

são linearmente independentes, pois

$$(1,0) \neq \alpha(0,1), \forall \alpha \in \mathbb{R}$$

e

$$(3,1) \neq \beta(-2,1), \forall \alpha \in \mathbb{R}$$

e podemos gerar um dado elemento de $\mathbb{R}^2$ por meio de uma combinação linear dos elementos de B_1 ou de B_2. Por exemplo, $\mathbf{u} = (7,9)$ pode ser gerado por

$$
\begin{aligned}
(7,9) &= \alpha_1(1,0)+\alpha_2(0,1) \\
(7,9) &= \beta_1(3,1)+\beta_2(-2,1)
\end{aligned} \tag{5.2}
$$

tal que, para a primeira base podemos obter os coeficientes diretamente, pois

$$7 = \alpha_1 \text{ e } 9 = \alpha_2$$

e, para a segunda base temos

$$
\begin{aligned}
7 &= 3\beta_1 - 2\beta_2 \\
9 &= \beta_1 + \beta_2
\end{aligned} \tag{5.3}
$$

cuja solução é $\beta_1 = 5$, $\beta_2 = 4$.

■

■ **Exemplo 5.7** Qual a cardinalidade do espaço gerado por $(2,3,4)$, $(1,-1,1)$ e $(5,5,9)$?
Começamos observando que os vetores são linearmente dependentes

$$(2,3,4) = -0.5(1,-1,1) + 0.5(5,5,9) = (0.5+2.5, 0.5+2.5, -0.5+4.5). \tag{5.4}$$

Isso quer dizer que esses três vetores não formam uma base para o $\mathbb{R}^3$, como poderíamos pensar em um primeiro momento. Observando bem, notamos que os vetores $(2,3,4)$ e $(1,-1,1)$ são linearmente independentes pois

$$(2,3,4) = \alpha(1,-1,1) \tag{5.5}$$

implica em

$$\begin{aligned} 2 &= \alpha \\ -3 &= \alpha \\ 4 &= \alpha \end{aligned} \tag{5.6}$$

que não tem solução (não existe um α que satisfaça as igualdades ao mesmo tempo). Isso quer dizer que podemos gerar, no máximo, um espaço de dimensão 2.

■

5.2 Conjuntos abertos e fechados em $\mathbb{R}^n$

Antes de iniciarmos a definição de conjuntos abertos e conjuntos fechados, precisamos definir uma bola em $\mathbb{R}^n$. Vamos apresentar três definições distintas:

Definição 5.2.1 Bola aberta em $\mathbb{R}^n$
Uma bola aberta de raio r e centro $\mathbf{u}$ é definida por

$$B_r(\mathbf{u}) = \{\mathbf{v} \in \mathbb{R}^{\mathbb{N}} : d(\mathbf{u},\mathbf{v}) < r\},$$

ou seja, é o conjunto de todos os elemento $\mathbf{v}$ cuja distância em relação ao centro $\mathbf{u}$ é menor do que o raio r.

Definição 5.2.2 Bola fechada em $\mathbb{R}^n$
Uma bola fechada de raio r e centro $\mathbf{u}$ é definida por

$$B_r(\mathbf{u}) = \{\mathbf{v} \in \mathbb{R}^{\mathbb{N}} : d(\mathbf{u},\mathbf{v}) \leq r\},$$

ou seja, é o conjunto de todos os elemento $\mathbf{v}$ cuja distância em relação ao centro $\mathbf{u}$ é menor ou igual ao raio r.

e, finalmente

Definição 5.2.3 Esfera em $\mathbb{R}^N$
Uma esfera de raio r e centro $\mathbf{u}$ é definida por

$$B_r(\mathbf{u}) = \{\mathbf{v} \in \mathbb{R}^{\mathbb{N}} : d(\mathbf{u},\mathbf{v}) = r\},$$

ou seja, é o conjunto de todos os elemento $\mathbf{v}$ cuja distância em relação ao centro $\mathbf{u}$ é exatamente igual ao raio r.

Essas definições nada mais são do que a especificação de uma região no entorno de um ponto $\mathbf{u}$, onde iremos considerar elementos “vizinhos” $\mathbf{v}$, de acordo com alguma norma pré-definida. No que segue, iremos utilizar o conceito de bola aberta.

O conceito de conjunto aberto e de fechado causa muita confusão. Não devemos pensar em uma caixa aberta ou fechada, mas sim respeitar as definições que serão apresentadas a seguir:

Definição 5.2.4 Conjunto Aberto

Um conjunto $G \subseteq \mathbb{R}^n$ é aberto em $\mathbb{R}^n$ se para todo o $\mathbf{u} \in G$ existe $\varepsilon > 0$ real tal que a bola aberta $B_\varepsilon(\mathbf{u}) \subseteq G$. Ou seja, um conjunto é aberto se qualquer ponto do conjunto pode ser perturbado por uma quantidade infinitesimal ε em qualquer direção e ainda continuar no conjunto (a bola toda deve estar contida no conjunto).

Um conjunto é aberto se e somente se coincidir com seu interior.

■ **Exemplo 5.8** Conjunto aberto

Seja $A = \{x : x \in (0,1)\}$. Se definirmos uma bola de raio ε em uma posição genérica x teremos $B_\varepsilon(x) = (x-\varepsilon, x+\varepsilon)$. Agora vamos imaginar que ε tem valores na faixa $min\{\frac{x}{2}, 1-\frac{x}{2}\}$ então, se nos aproximarmos de zero, teremos $B_\varepsilon(x \to 0) = (x-\frac{x}{2}, x+1-\frac{x}{2}) \to (0,1) \in A$. Assim, A contém todos os seus pontos interiores e é aberto.

■

■ **Exemplo 5.9** Conjunto não aberto

Seja $B = \{x : x \in [0,1]\}$. Se definirmos uma bola de raio ε em uma posição genérica x teremos $B_\varepsilon(x) = (x-\varepsilon, x+\varepsilon)$. Agora vamos imaginar que estamos em $x = 0$ e que perturbamos em ε genérico. Termos um intervalo $(-\varepsilon, \varepsilon)$ sendo que o extremo esquerdo deste intervalo não está contido em B. ■

■ **Exemplo 5.10** Conjunto aberto

O conjunto $\mathbb{R}^n$ é aberto pois, como se estende ao infinito, sempre podemos colocar uma bola que conterá pontos no $\mathbb{R}^n$ ■

Assim, para intervalos unidimensionais é fácil de ver que se o intervalo for aberto, então é possível chegar arbitrariamente perto do extremo aberto e ainda definir uma bola infinitesimal que esteja contida no intervalo. Se o intervalo for fechado, então podemos nos posicionar exatamente sobre o extremo e mostrar que uma bola centrada neste ponto terá pontos fora do intervalo.É importante pensar em limites pela esquerda e pela direita.

É interessante notar que o conjunto vazio $\emptyset$ é aberto (por vacuidade, pois se não tem pontos, então não podemos definir a noção de uma bola).

Definição 5.2.5 Complemento de um Conjunto

O complemento $\complement(V)$, ou $V^\complement$, de um conjunto $V \subseteq \mathbb{R}^n$ é definido por

$$\complement(V) = \mathbb{R}^n \backslash V = \{\mathbf{u} \in \mathbb{R}^n : \mathbf{u} \notin V\}$$

ou seja, todos os elementos de $\mathbb{R}^n$ que não estão em V.

■ **Exemplo 5.11** Complemento de um conjunto

Seja $B = [0,1]$: $\complement(B) = (-\infty, 0) \cup (1, \infty)$

Seja $A = (0,1)$: $\complement(A) = (-\infty, 0] \cup [1, \infty)$

Dois complementos interessantes são: $\complement(\emptyset) = \mathbb{R}^N$ e $\complement(\mathbb{R}^N) = \emptyset$. ■

Definição 5.2.6 Conjunto Fechado

Um conjunto $F \subseteq \mathbb{R}^N$ é fechado em $\mathbb{R}^n$ se seu complemento é aberto.

■ **Exemplo 5.12** Conjunto fechado

Seja $B = [0,1]$. O complemento de B é aberto, pois é possível posicionar a bola em qualquer ponto do complemento e garantir que a bola estará contida no complemento (não possuirá pontos em B). Assim, B é fechado.

■

■ **Exemplo 5.13** Conjunto não fechado
Seja $A = (0,1)$. O complemento de A não é aberto, pois se posicionarmos a bola em 0] ou em [1 teremos um pedaço da bola aberta fora do complemento (dentro de A). Assim, A não é fechado; ■

■ **Exemplo 5.14** Conjunto aberto e fechado ao mesmo tempo
É interessante notar que um conjunto pode ser aberto e fechado ao mesmo tempo. Por exemplo, $\emptyset$ é aberto e é fechado (pois o seu complemento é aberto). Da mesma forma, $\mathbb{R}^n$ também tem complemento aberto, sendo portanto fechado. ■

■ **Exemplo 5.15** Conjunto que não é aberto e nem fechado
Seja $E = \{x : x \in (0,1]\}$. Este conjunto não é aberto, pois podemos posicionar a bola aberta em 1 e, com isto, ter pontos da bola fora de E. O interessante é que o complemento de E, $\complement(E) = (-\infty,0] \cup (1,\infty)$ também não é aberto, pois podemos colocar uma bola em 0 e ela conterá pontos de E. Assim, E não é aberto e nem fechado. ■

Definição 5.2.7 Vizinhança, Ponto interior, ponto de fronteira e exterior
Sejam $\mathbf{u} \in \mathbb{R}^n$ e $A \subseteq \mathbb{R}^n$:

- Uma vizinhança de $\mathbf{u}$ é um conjunto que contém um subconjunto aberto que contenha $\mathbf{u}$
- $\mathbf{u}$ é um ponto interior de A se existe uma vizinhança de $\mathbf{u}$ contida em A
- $\mathbf{u}$ é um ponto de fronteira de A se toda a vizinhança de $\mathbf{u}$ está contida em A e em seu complemento
- $\mathbf{u}$ é um ponto exterior de A se existe uma vizinhança de $\mathbf{u}$ contida em $\complement(A)$

■ **Exemplo 5.16** Conjunto aberto
Seja um conjunto aberto I. Podemos afirmar que:

- todos os seus pontos são interiores
- I é uma vizinhança de todos os seus pontos
- I não contém pontos de fronteira

■

■ **Exemplo 5.17** Conjunto fechado
Se F é um conjunto fechado, então F contém todos os seus pontos de fronteira. ■

Definição 5.2.8 Ponto de Acumulação
Um ponto $\mathbf{u} \in \mathbb{R}^n$ é um ponto de acumulação de $S \subseteq \mathbb{R}^n$ se toda vizinhança $B_\varepsilon(\mathbf{u})$ contém pelo menos um ponto de S diferente de $\mathbf{u}$.

■ **Exemplo 5.18** Um conjunto $S \subseteq \mathbb{R}$ limitado superiormente com $\mathbf{u} = sup\, S \notin S$.
Neste caso, $\mathbf{u}$ é ponto de acumulação de S, uma vez que existem $\varepsilon > 0$ e $\mathbf{v} \in S$ tal que $\mathbf{v} \in (\mathbf{u} - \varepsilon, \mathbf{u} + \varepsilon)$. ■

■ **Exemplo 5.19** $A = (0,1)$
Neste caso, qualquer ponto no intervalo $[0,1]$ é ponto de acumulação de A (observe que um ponto fora do conjunto pode ser de acumulação); ■

■ **Exemplo 5.20** $\mathbb{N} = \{0,1,2,3,4,5,....\}$
Esse conjunto não tem ponto de acumulação, pois podemos considerar uma bola com centro em qualquer número e colocar um raio pequeno o suficiente para conter apenas o centro da bola (lembrem-se que estamos trabalhando com números inteiros positivos e, portanto, 0.00001 não faz parte de $\mathbb{N}$. ■

■ **Exemplo 5.21** Conjunto finito

De modo geral, podemos afirmar que conjuntos finitos (com número limitado de elementos) não possuem ponto de acumulação, pois podemos definir uma vizinhança (real) que não contenha pontos do conjunto. ■

■ **Exemplo 5.22** $A = (2,3) \cup \{4\}$

O conjunto tem como pontos de acumulação $[2,3]$, pois podemos colocar uma vizinhança em torno de $\{4\}$ que não intercepte pontos do conjunto (além do 4). ■

Desta forma, podemos definir o conceito de Conjunto Fechado de uma forma alternativa

Definição 5.2.9 Conjunto Fechado (definição alternativa)

Um conjunto $F \subseteq \mathbb{R}^N$ é fechado se e somente se contém todos os seus pontos de acumulação.

Por fim, é muito importante definirmos o teorema de Bolzano-Weierstrass:

Teorema 5.2.1 Bolzano-Weierstrass no $\mathbb{R}^n$

Todo subconjunto infinito e limitado (inferiormente e superiormente) de $\mathbb{R}^n$ tem pelo menos um ponto de acumulação.

Este teorema parece meio óbvio se olharmos o exemplo que foi discutido anteriormente: um conjunto $S \subseteq \mathbb{R}$ limitado superiormente com $\mathbf{u} = sup\,S \notin S$. Neste caso, $\mathbf{u}$ é ponto de acumulação de S. O mesmo pode ser dito para o $inf\,S$.

Vamos analisar novamente os nossos exemplos:

- $\mathbb{Z}$ é infinito, mas é ilimitado (não tem limite superior e nem inferior);
- $A = (0,1) \cup \{2,3,4,5,.....\}$ é infinito mas não é limitado superiormente. O teorema não pode afirmar nada sobre este caso, mas podemos observar que $[0,1]$ são pontos de acumulação !!
- $A = \{1,2,3\}$ é finito e, portanto, não tem pontos de acumulação.

6. Espaços Vetoriais

Todos os conceitos aprendidos com o $\mathbb{R}^n$ podem ser estendidos para espaços mais gerais, onde os membros (vetores) são funções. Embora possa parecer algo muito vago em um primeiro momento, pois estamos acostumados a pensar em vetores como flechas no $\mathbb{R}^2$ ou no $\mathbb{R}^3$, este conceito é muito útil para todas as nossas aplicações em mecânica computacional.

Lembrando, um espaço é um conjunto que é equipado com operações binárias de soma e multiplicação (vimos isso no capítulo do $\mathbb{R}^n$). Todas as definições que vimos para o $\mathbb{R}^n$ são definidas por axiomas gerais, que podem ser aplicados para casos mais gerais do que os vetores que estamos acostumados. Vamos a dois exemplos.

■ **Exemplo 6.1** Espaço dos polinômios de segundo grau

Seja o conjunto P_2, com polinômios da forma $p(x) = a + bx + cx^2$, com $a, b, c \in \mathbb{R}$.

A soma de dois polinômios deste conjunto $p_1(x) = a_1 + b_1x + c_1x^2$ e $p_2(x) = a_2 + b_2x + c_2x^2$ é definida como

$$p_1(x) + p_2(x) = (a_1 + a_2) + (b_1 + b_2)x + (c_1 + c_2)x^2$$

e a multiplicação por um escalar $\alpha \in \mathbb{R}$ é definida por

$$\alpha p_1(x) = (\alpha a_1) + (\alpha b_1)x + (\alpha c_1)x^2.$$

Esse espaço tem como conjunto de base (por exemplo) o conjunto de monoides

$$B = \{1, x, x^2\}$$

e, portanto, é um espaço de cardinalidade 3. Com isso, podemos representar elementos deste espaço como "vetores"

$$\mathbf{p} = p(x) = a + bx + cx^2 = a(1,0,0) + b(0,x,0) + c(0,0,x^2)$$

ou, simplesmente, $\mathbf{p} = (a, b, c)$.

■

Exercício 6.1 Prove que $B = \{1, x, x^2\}$ é uma base válida para o P_2 ■

■ **Exemplo 6.2** Solução de equações diferenciais
Seja a equação diferencial

$$\frac{d^2u(x)}{dx^2} - k^2u(x) = 0, k > 0$$

onde $x \in [0,1]$. Como sabemos, a solução da equação diferencial é uma família de funções $u(x)$ que satisfaz a equação diferencial para $\forall x \in [0,1]$. Esta família de funções gera um espaço de funções, que podem ser escritas a partir de um conjunto de funções base. Da literatura, sabemos que $u(x)$ deve fazer parte de um espaço V, com base $B = \{e^{kx}, e^{-kx}\}$ tal que qualquer elemento de V (solução da eq. diferencial) será uma combinação linear na forma $u(x) = \alpha_1 e^{kx} + \alpha_2 e^{-kx}$.

■

Da mesma forma que no $\mathbb{R}^n$, temos que a definição básica de um espaço requer somente os conceitos de soma e de multiplicação por um escalar. Isso já basta para definirmos a combinação linear, independência linear e, portanto, o conjunto de base. No entanto, podemos equipar espaços com mais medidas. Uma das mais importantes é o conceito de norma, e podemos estender o conceito de norma P para espaços de funções, na forma

Definição 6.0.1 Norma P no espaço de funções
Seja $\mathbf{u}$ uma função integrável em um domínio Ω e $p \geq 1$. A norma P de $\mathbf{u}$ é definida como

$$\|\mathbf{u}\|_p = \left(\int_\Omega |\mathbf{u}|^p\right)^{\frac{1}{p}}.$$

Estas normas são conhecidas como Normas de Lebesgue. Outas normas são possíveis, como por exemplo a família de normas de Sobolev:

Definição 6.0.2 Norma de Sobolev Seja $\mathbf{u}$ uma função integrável em um domínio Ω, $p \geq 1$, $m > 0$. A norma m, p de $\mathbf{u}$ é definida como

$$\|\mathbf{u}\|_{m,p} = \left[\int_\Omega \sum_{|\alpha| \leq m} |D^\alpha \mathbf{u}|^p\right]^{\frac{1}{p}}.$$

■ **Exemplo 6.3 — Norma de Sobolev.** Podemos definir a norma $\|\mathbf{u}\|_{1,2}$ em $\Omega \in [0,1] \times [0,1]$

$$\|\mathbf{u}\|_{m,p} = \left[\int_0^1 \int_0^1 \left[|\mathbf{u}|^2 + \left|\frac{d\mathbf{u}}{dx}\right|^2 + \left|\frac{d\mathbf{u}}{dy}\right|^2\right] dxdy\right]^{\frac{1}{2}}$$

e podemos verificar que se $m = 0$ a norma de Sobolev se iguala a norma de Lebesgue.

■

Deve-se enfatizar que estas normas são definidas para aplicações distintas. Assim, normas de Sobolev são necessárias para descrevermos espaços de funções e derivadas suficientemente contínuas para alguma aplicação, como por exemplo na solução de equações diferenciais parciais.

Exercício 6.2 — Efeito da escolha da norma na definição de um círculo. Como exercício, imagine o $\mathbb{R}^2$ equipado com normas L^p para p igual a 1, 2 e ∞ e desenhe um "círculo" de raio 1 em torno da origem (lembrem-se que $p = \infty$ implica em selecionar o maior valor). Lembre-se que o raio é sempre uma distância (associada a norma) que depende das coordenadas do ponto em relação a origem (centro do círculo). ■

Desta forma, conhecendo o espaço vetorial de solução do nosso problema e as propriedades dos operadores envolvidos, saberemos tudo sobre as suas propriedades (se tem solução, quantas tem, estrategias de solução, etc...).

6.1 Transformações Lineares

A relação entre elementos de dois conjuntos foi definida como uma função. Observem que desde o início deste texto, o conceito de elemento tem sido estendido. Atualmente, estamos considerando conjuntos cujos elementos são funções e iremos agora definir relações entre funções. Um exemplo seria a operação

$$-\frac{d}{dx}\left(a\frac{du(x)}{dx}\right) = f(x)$$

onde $a \in \mathbb{R}$. Uma outra maneira de escrever esta expressão é na forma

$$T(u) = f$$

ou

$$T : U \to V$$

que pode ser lida como: $T(\cdot) = -\frac{d}{dx}\left(a\frac{d(\cdot)}{dx}\right)$ é um operador que relaciona o espaço U de funções que contém $\mathbf{u}$ com o espaço V de funções que contém f. Assim, se conhecermos as propriedades do operador, podemos estudar a solução de equações de forma mais geral. Vamos estudar algumas definições pertinentes.

Definição 6.1.1 Núcleo de uma transformação (*Kernel*)
O núcleo de uma transformação (também conhecido como espaço nulo) é dado por

$$\aleph(T) = \{\mathbf{u} : \mathbf{u} \in U, T(\mathbf{u}) = \mathbf{0}\}$$

■ **Exemplo 6.4** Núcleo de uma transformação
Vamos considerar a transformação $T : U \to V$ tal que $T(\mathbf{u}) = u_1 + u_2$. O núcleo desta transformação é definido por

$$\aleph(T) = \{\mathbf{u} \in U \ : \ u_1 + u_2 = 0\}$$

tal que elementos que tenham a estrutura

$$(-a, a)$$

compõe o núcleo desta transformação. Desta forma, $\mathbf{v} = (-1, 1)$ é base do núcleo de T.
■

É interessante notar que podemos re-utilizar todos os conceitos que aprendemos com funções e estender para o conceito de operador. Por exemplo, os conceitos de domínio e de imagem são análogos, bem como a de um-a-um (injetivo). No entanto, costumamos chamar um operador injetivo de um *monomorfismo* (para cada $\mathbf{u}$ temos um $\mathbf{v}$ distinto) e "sobre", ou sobrejetivo, assume a conotação de *epimorfismo* (todo V é utilizado). Neste contexto, se o operador for um *monomorfismo* E um *epimorfismo*, então ele é dito um *isomorfismo* (bijeção) e, portanto, é possível definir um operador inverso.

Dentre as transformações, dedicamos especial atenção a uma classe muito importante e utilizada: as transformações lineares:

Definição 6.1.2 Transformação Linear

T é uma transformação linear de um elemento no espaço U em um elemento do espaço V se cada elemento $\mathbf{u} \in U$ corresponde a um elemento $\mathbf{v} \in V$ e se as seguintes propriedades são observadas:

- homogêneo: $T(\alpha\mathbf{u}) = \alpha T(\mathbf{u}) \forall \mathbf{u} \in U, \forall \alpha \in \mathbb{R}$;
- aditivo: $T(\mathbf{u}_1 + \mathbf{u}_2) = T(\mathbf{u}_1) + T(\mathbf{u}_2)$, $\forall \mathbf{u}_1, \mathbf{u}_2 \in U$;
- se T_1 e T_2 são duas transformações lineares, $T_1 T_2$ também será uma transformação linear (pense em composição de funções).

Com as definições anteriores, podemos então definir

Definição 6.1.3 Monomorfismo e Transformação Linear

Uma transformação Linear $T : U \to V$ é um-a-um (monomorfismo) se e somente se o espaço nulo é trivial, isto é, $\aleph(T) = \{\mathbf{0}\}$.

Demonstração. T Injetivo significa que $T(\mathbf{u}_1) = T(\mathbf{u}_2) \implies \mathbf{u}_1 = \mathbf{u}_2$. Se T for linear, então podemos escrever $T(\mathbf{u}_1 - \mathbf{u}_2) = \mathbf{0} \implies \mathbf{u}_1 - \mathbf{u}_2 = \mathbf{u} \in \aleph(T), \forall \mathbf{u}_1, \mathbf{u}_2 \in U$. Isso só será verdadeiro se $\mathbf{u} = \mathbf{0}$, ou seja, o espaço nulo é trivial. ■

■ **Exemplo 6.5** Operador que é um monomorfismo

Vamos considerar a transformação linear $T : U \to V$ tal que $T(\mathbf{u}) = 7\mathbf{u}$. Neste caso,

$$\aleph(T) = \{\mathbf{u} \in U \, : \, \mathbf{u} = \mathbf{0}\}$$

tal que T é um monomorfismo. ■

■ **Exemplo 6.6** Operador que não é um monomorfismo

Vamos considerar a transformação $T : U \to V$ tal que $T(\mathbf{u}) = u_1 + u_2$. O núcleo desta transformação é $(-1, 1)$ que não é trivial. Desta forma, o operador não é um monomorfismo.

■

Definição 6.1.4 *Rank* de um operador

A dimensão do subespaço $\mathscr{R}(T) = Im(T) \subseteq V$ é chamada de *Rank* da transformação $T : U \to V$. Em outras palavras, é a cardinalidade da imagem de T, $|Im(T)|$.

■ **Exemplo 6.7** Seja o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ com $T(\mathbf{u}) = 7\mathbf{u}$. A dimensão da imagem de T, $Im(T)$ é $n = 3$, tal que $\mathscr{R}(T) = 3$.

■

■ **Exemplo 6.8** Seja o operador $T : \mathbb{R}^2 \to \mathbb{R}$ com $T(\mathbf{u}) = (u_1 + u_2)$. A dimensão da imagem de T, $Im(T)$ é $n = 1$, tal que $\mathscr{R}(T) = 1$.

■

Definição 6.1.5 *Nullity* de um operador

A dimensão do espaço nulo $\aleph(T)$ é chamada de *nullity* de T. Em outras palavras, é a cardinalidade do espaço nulo, $|\aleph(T)|$.

Definição 6.1.6 $|\aleph(T)| + |\mathscr{R}(T)| = |U|$

Seja um operador $T : U \to V$. A soma do *rank* com o *nullity* é igual a dimensão (cardinalidade) de U.

■ **Exemplo 6.9** Seja $T : U \to V$ tal que $T(\mathbf{u}) = u_1 + u_2$. Temos que o *rank* da transformação é igual a 1 e que a dimensão do espaço nulo é 1 (um parâmetro a). U corresponde ao $\mathbb{R}^2$, o que está de acordo com a definição. ■

■ **Exemplo 6.10** Seja o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ com $T(\mathbf{u}) = 7\mathbf{u}$.

A dimensão da imagem de T, $Im(T)$ é $n = 3$ e o espaço nulo é trivial, tal que a *nullity* é zero. Assim, $3 + 0 = 3$. ■

6.2 Transformação Linear de Espaços Finito Dimensionais

Todas as definições anteriores são gerais para espaços com dimensão infinita (embora os exemplos tenham sido realizados em espaços com dimensão finita, para simplificar o entendimento).

Sejam U e V dois espaços finito dimensionais e $T : U \to V$ uma transformação linear. Se $\{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, ..., \boldsymbol{\phi}_n\}$ é base de U e $\{\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, ..., \boldsymbol{\varphi}_m\}$ é base de V, podemos escrever qualquer membro de U e V como

$$\mathbf{u} = \sum_{i=1}^{n} \alpha_i \boldsymbol{\phi}_i$$

e

$$\mathbf{v} = \sum_{j=1}^{m} \beta_j \boldsymbol{\varphi}_j.$$

Para cada $\mathbf{u}$ e $\mathbf{v}$, temos que $T(\mathbf{u}) = \mathbf{v}$ pode ser escrito na forma

$$T(\mathbf{u}) = T\left(\sum_{i=1}^{n} \alpha_i \boldsymbol{\phi}_i\right) = \sum_{j=1}^{m} \beta_j \boldsymbol{\varphi}_j,$$

e, utilizando as propriedades de um operador linear, Definição 6.1.2,

$$T(\mathbf{u}) = \sum_{i=1}^{n} T(\alpha_i \boldsymbol{\phi}_i) = \sum_{j=1}^{m} \beta_j \boldsymbol{\varphi}_j,$$

e

$$T(\mathbf{u}) = \sum_{i=1}^{n} \alpha_i T(\boldsymbol{\phi}_i) = \sum_{j=1}^{m} \beta_j \boldsymbol{\varphi}_j,$$

e, como $T(\boldsymbol{\phi}_i) \in V$, podemos a aplicação do operador sobre um termo de base do domínio como uma combinação linear na base do contra-domínio

$$T(\boldsymbol{\phi}_i) = \sum_{j=1}^{m} t_{ji} \boldsymbol{\varphi}_j$$

tal que

$$\sum_{j=1}^{m} \beta_j \boldsymbol{\varphi}_j - \sum_{i=1}^{n} \alpha_i \left(\sum_{j=1}^{m} t_{ji} \boldsymbol{\varphi}_j\right) = \mathbf{0}$$

ou

$$\sum_{j=1}^{m} \left(\beta_j - \sum_{i=1}^{n} t_{ji} \alpha_i\right) \boldsymbol{\varphi}_j = \mathbf{0}$$

que pode ser escrita na forma matricial $\boldsymbol{\beta} = \mathbf{T}\boldsymbol{\alpha}$

$$\begin{bmatrix} t_{11} & \dots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{m1} & \dots & t_{mn} \end{bmatrix} \begin{Bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{Bmatrix} = \begin{Bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{Bmatrix}$$

onde $\mathbf{T}$ é uma matriz de dimensões $m \times n$, contendo os termos t_{ji}. Baseados nesta dedução, podemos verificar que:

- A transformação entre dois espaços vetoriais finitos pode ser representada pela aplicação do operador T nas bases de cada espaço;
- A matriz $\mathbf{T}$ é a forma finito dimensional (discreta) do operador T.

Estes resultados são utilizados a todo instante na mecânica computacional, pois trabalhamos com espaços finito dimensionais para representar modelos discretos de problemas contínuos (infinito dimensionais).

Vamos apresentar alguns exemplos.

■ **Exemplo 6.11** Considere a operação $T : \mathbb{R}^3 \to \mathbb{R}^3$

$$T(\mathbf{u}) = (v_1, v_2, v_3) = (10u_1 - u_2 + 3u_3, -u_1 + 15u_2 - 3u_3, 3u_1 - 3u_2 + 9u_3).$$

Podemos identificar que as bases do domínio U são $\boldsymbol{\phi}_1 = (1,0,0)$, $\boldsymbol{\phi}_2 = (0,1,0)$ e $\boldsymbol{\phi}_3 = (0,0,1)$ e as bases do contra-domínio são $\boldsymbol{\varphi}_1 = (1,0,0)$, $\boldsymbol{\varphi}_2 = (0,1,0)$ e $\boldsymbol{\varphi}_3 = (0,0,1)$. Essas são as bases canônicas no $\mathbb{R}^3$. Se aplicarmos o operador nas bases $\boldsymbol{\phi}$ do domínio

$$\begin{aligned}
T(\boldsymbol{\phi}_1) &= T(1,0,0) = (10,-1,3) \\
T(\boldsymbol{\phi}_2) &= T(0,1,0) = (-1,15,-3) \\
T(\boldsymbol{\phi}_3) &= T(0,0,1) = (3,-3,9)
\end{aligned}$$

e como queremos escrever na base de V

$$\begin{aligned}
T(\boldsymbol{\phi}_1) &= t_{11}(1,0,0) + t_{21}(0,1,0) + t_{31}(0,0,1) = (10,-1,3) \\
T(\boldsymbol{\phi}_2) &= t_{12}(1,0,0) + t_{22}(0,1,0) + t_{32}(0,0,1) = (-1,15,-3) \\
T(\boldsymbol{\phi}_3) &= t_{13}(1,0,0) + t_{23}(0,1,0) + t_{33}(0,0,1) = (3,-3,9)
\end{aligned}$$

tal que

$$\alpha_1(10,-1,3) + \alpha_2(-1,15,-3) + \alpha_3(3,-3,9) = \beta_1(1,0,0) + \beta_2(0,1,0) + \beta)3(0,0,1).$$

Se colocarmos os termos de base $(1,0,0)$, $(0,1,0)$ e $(0,0,1)$ em evidência

$$\begin{aligned}
(1,0,0) &\implies \alpha_1 \cdot 10 + \alpha_2 \cdot -1 + \alpha_3 \cdot 3 = \beta_1 \\
(0,1,0) &\implies \alpha_1 \cdot -1 + \alpha_2 \cdot 15 + \alpha_3 \cdot -3 = \beta_2 \\
(0,0,1) &\implies \alpha_1 \cdot 3 + \alpha_2 \cdot -3 + \alpha_3 \cdot 9 = \beta_3
\end{aligned}$$

observamos que o operador discreto é

$$\mathbf{T} = \begin{bmatrix} 10 & -1 & 3 \\ -1 & 15 & -3 \\ 3 & -3 & 9 \end{bmatrix}.$$

Pode-se observar que quando usamos as bases canônicas no $\mathbb{R}^n$ os coeficientes de $\mathbf{T}$ são obtidos diretamente da aplicação do operador na base.

Como exemplo,

$$\begin{bmatrix} 10 & -1 & 3 \\ -1 & 15 & -3 \\ 3 & -3 & 9 \end{bmatrix} \begin{Bmatrix} 10 \\ 20 \\ 30 \end{Bmatrix} = \begin{Bmatrix} 10 \cdot 10 - 1 \cdot 20 + 3 \cdot 30 \\ -1 \cdot 10 + 15 \cdot 20 - 3 \cdot 30 \\ 3 \cdot 10 - 3 \cdot 20 + 9 \cdot 30 \end{Bmatrix} = \begin{Bmatrix} 170 \\ 200 \\ 240 \end{Bmatrix}.$$

O operador tem espaço nulo trivial pois suas colunas são linearmente independentes. De fato

$$\begin{bmatrix} 10 & -1 & 3 \\ -1 & 15 & -3 \\ 3 & -3 & 9 \end{bmatrix} \begin{Bmatrix} a \\ b \\ c \end{Bmatrix} = \begin{Bmatrix} 10\cdot a - 1\cdot b + 3\cdot c \\ -1\cdot a + 15\cdot b - 3\cdot c \\ 3\cdot a - 3\cdot b + 9\cdot c \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

só é satisfeito se $a = b = c = 0$. Assim $\aleph(T) = 0$, $\mathscr{R}(T) = 3$ e $|U| = 3$ e verificamos que $\mathscr{R}(T) + \aleph(T) = |U|$.

Como a matriz é quadrada e a nulidade é zero, temos que o operador é um monomorfismo, tal que a operação inversa é possível. De fato, se invertermos as operações entre **u** e **v**

$$u_1 = \frac{3v_1 - v_3}{27}$$
$$u_2 = \frac{3v_2 + v_3}{42}$$
$$u_3 = \frac{-42v_1 + 27v_2 + 149v_3}{1134},$$

tal que agora temos um operador $P : V \to U$. Aplicando esse operador nas bases $\boldsymbol{\varphi}$ de V geramos elementos em U

$$P(1,0,0) = (3/27, 0, -42/1134)$$
$$P(0,1,0) = (0, 3/42, 27/1134)$$
$$P(0,0,1) = (-1/27, 1/42, 149/1134)$$

ou, na forma matricial (discreta)

$$\mathbf{P} = \begin{bmatrix} 3/27 & 0 & -1/27 \\ 0 & 3/42 & 1/42 \\ -1/27 & 1/42 & 149/1134 \end{bmatrix}$$

que é exatamente a inversa da matrix **T**

$$\begin{bmatrix} 10 & -1 & 3 \\ -1 & 15 & -3 \\ 3 & -3 & 9 \end{bmatrix}^{-1} = \begin{bmatrix} 3/27 & 0 & -1/27 \\ 0 & 3/42 & 1/42 \\ -1/27 & 1/42 & 149/1134 \end{bmatrix}$$

Testando

$$\begin{bmatrix} 3/27 & 0 & -1/27 \\ 0 & 3/42 & 1/42 \\ -1/27 & 1/42 & 149/1134 \end{bmatrix} \begin{Bmatrix} 170 \\ 200 \\ 240 \end{Bmatrix} = \begin{Bmatrix} 10 \\ 20 \\ 30 \end{Bmatrix}.$$

■

■ **Exemplo 6.12** Seja U o espaço dos polinômios de terceiro grau e V o espaço dos polinômios de primeiro grau. Embora existam infinitos polinômios de primeiro e terceiro graus, as suas bases são finitas, tendo a forma

$$\phi = \{1, x, x^2, x^3\}$$
$$\varphi = \{1, x\}$$

tal que um elemento de U tem a forma genérica $\mathbf{u} = \sum_{i=1}^{4} \alpha_i \boldsymbol{\phi}_i = \alpha_1 + \alpha_2 x + \alpha_3 x^2 + \alpha_4 x^3$ $(n = 4)$ e um elemento de V tem a forma genérica $\mathbf{v} = \sum_{j=1}^{2} \beta_j \boldsymbol{\varphi}_j = \beta_1 + \beta_2 x$ $(m = 2)$.

Um operador linear que mapeia (transforma) entre U e V é o operador diferencial

$$D = \frac{d^2}{dx^2}$$

e, portanto,

$$D : U \to V$$

significa que

$$D\left(\sum_{i=1}^{4} \alpha_i \phi_i\right) = \sum_{j=1}^{2} \beta_j \boldsymbol{\varphi}_j$$

e, como o operador D é linear, podemos passar para dentro do somatório em i, tal que

$$\sum_{i=1}^{4} \alpha_i D(\boldsymbol{\phi}_i) = \sum_{j=1}^{2} \beta_j \boldsymbol{\varphi}_j$$

e como $D(\boldsymbol{\phi}_i) \in V$, podemos escrever

$$D(\boldsymbol{\phi}_i) = \sum_{j=1}^{2} d_{ji} \boldsymbol{\varphi}_j$$

significando que estamos escrevendo este resultado na base de V. Vamos avaliar $D(\boldsymbol{\phi}_i)$ para $i = 1..4$, ou seja, $\{1, x, x^2, x^3\}$:

$$\begin{aligned}
D(\phi_1) &= D(1) = 0 \\
D(\phi_2) &= D(x) = 0 \\
D(\phi_3) &= D(x^2) = 2 \\
D(\phi_4) &= D(x^3) = 6x
\end{aligned}$$

mas como queremos escrever estes termos na base de V, observamos que

$$\begin{aligned}
D(\phi_1) &= d_{11}1 + d_{21}x = 0 \cdot 1 + 0 \cdot x \\
D(\phi_2) &= d_{12}1 + d_{22}x = 0 \cdot 1 + 0 \cdot x \\
D(\phi_3) &= d_{13}1 + d_{23}x = 2 \cdot 1 + 0 \cdot x \\
D(\phi_4) &= d_{14}1 + d_{24}x = 0 \cdot 1 + 6 \cdot x
\end{aligned}$$

tal que

$$\sum_{i=1}^{4} \alpha_i D(\boldsymbol{\phi}_i) = \sum_{j=1}^{2} \beta_j \boldsymbol{\varphi}_j$$

pode ser escrito como

$$\alpha_1 (0 \cdot 1 + 0 \cdot x) + \alpha_2 (0 \cdot 1 + 0 \cdot x) + \alpha_3 (2 \cdot 1 + 0 \cdot x) + \alpha_4 (0 \cdot 1 + 6 \cdot x) = \beta_1 \cdot 1 + \beta_2 \cdot x$$

de onde observamos que cada termo da base de V pode ser colocado em evidência:

$$\begin{aligned}
\varphi_1 &\to \alpha_1 0 + \alpha_2 0 + \alpha_3 2 + \alpha_4 0 = \beta_1 \\
\varphi_2 &\to \alpha_1 0 + \alpha_2 0 + \alpha_3 0 + \alpha_4 6 = \beta_2
\end{aligned}$$

ou, na forma matricial

$$\begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix} \left\{ \begin{array}{c} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{array} \right\} = \left\{ \begin{array}{c} \beta_1 \\ \beta_2 \end{array} \right\}.$$

tal que

$$D = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

é a forma discreta do operador diferencial que mapeia U em V. É interessante notar que a matriz nada mais é do que a coleção os coeficientes d que multiplicam a aplicação do operador nas bases $\boldsymbol{\phi}$.

Para testar, considerarmos o polinômio

$$p(x) = 7 + 2x + 15x^2 - 0.5x^3$$

e realizarmos a operação $D : U \to V$, obteremos

$$\begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix} \begin{Bmatrix} 7 \\ 2 \\ 15 \\ -0.5 \end{Bmatrix} = \begin{Bmatrix} 30 \\ -3 \end{Bmatrix}$$

ou seja, o polinômio $\mathbf{v} = 30 - 3x$.

■

■ **Exemplo 6.13** Dado o operador $A(p(x)) = (x^2 - 1)p'(x)$ e as bases $U = \{1, x, x^2\}$, $V = \{1, x, x^2, x^3\}$, obtenha a forma discreta do operador.

$$\begin{aligned} A(1) &= (x^2 - 1)0 = 0 = a_{11}1 + a_{21}x + a_{31}x^2 + a_{41}x^3 \\ A(x) &= (x^2 - 1)1 = x^2 - 1 = a_{12}1 + a_{22}x + a_{32}x^2 + a_{42}x^3 \\ A(x^2) &= (x^2 - 1)2x = 2x^3 - 2x = a_{13}1 + a_{23}x + a_{33}x^2 + a_{43}x^3. \end{aligned}$$

A primeira linha é satisfeita se e somente se $a_{11} = a_{21} = a_{31} = a_{41} = 0$

A segunda equação resulta em $a_{12} = -1$, $a_{32} = 1$, $a_{22} = a_{42} = 0$ e a terceira linha em $a_{43} = 2$, $a_{23} = -2$, $a_{13} = d_{33} = 0$. Assim, o operador discreto é

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

Testando o operador discreto $\mathbf{A}$ com $\mathbf{u} = 1 + 2x + 3x^2$

$$\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{Bmatrix} 1 \\ 2 \\ 3 \end{Bmatrix} = \begin{Bmatrix} -1 \cdot 2 \\ -2 \cdot 3 \\ 1 \cdot 2 \\ 2 \cdot 3 \end{Bmatrix} = \begin{Bmatrix} -2 \\ -6 \\ 2 \\ 6 \end{Bmatrix}.$$

Isso quer dizer que $A(\mathbf{u})$ resulta em

$$-2 - 6x + 2x^2 + 6x^3,$$

que é igual a

$$(x^2 - 1)\frac{d(1 + 2x + 3x^2)}{dx} = (x^2 - 1)(2 + 6x) = -2 - 6x + 2x^2 + 6x^3.$$

■

■ **Exemplo 6.14** Dado o operador $A(p(x)) = (x^2-1)p'(x)$ e as bases $U = \{1, x, x^2\}$, $V = \{1, (x-1), (x-1)^2, (x-1)^3\}$, obtenha a forma discreta do operador.

O exemplo é praticamente igual ao anterior, sendo que as bases agora não são do $\mathbb{R}^2$ e sim de espaços de polinômios. Novamente, o que queremos é descrever a aplicação do operador em uma base, em função de outra.

$$\begin{aligned}
A(1) &= (x^2-1)0 = a_{11}1 + a_{21}(x-1) + a_{31}(x-1)^2 + a_{41}(x-1)^3 \\
A(x) &= (x^2-1)1 = a_{12}1 + a_{22}(x-1) + a_{32}(x-1)^2 + a_{42}(x-1)^3 \\
A(x^2) &= (x^2-1)2x = a_{13}1 + a_{23}(x-1) + a_{33}(x-1)^2 + a_{43}(x-1)^3.
\end{aligned}$$

A primeira linha é satisfeita se e somente se $a_{11} = a_{21} = a_{31} = a_{41} = 0$

As outras duas equações resultam em

$$\begin{aligned}
x^2 - 1 &= a_{42}x^3 + (a_{32} - 3a_{42})\,x^2 + (3a_{42} - 2a_{32} + a_{22})\,x - a_{42} + a_{32} - a_{22} + a_{12} \\
2x^3 - 2x &= a_{43}x^3 + (a_{33} - 3a_{43})\,x^2 + (3a_{43} - 2a_{33} + a_{23})\,x - a_{43} + a_{33} - a_{23} + a_{13}
\end{aligned}$$

tal que, da primeira equação

$$\begin{aligned}
&a_{42} = 0 \\
&a_{32} - 3a_{42} = 1 \\
&3a_{42} - 2a_{32} + a_{22} = 0 \\
&-a_{42} + a_{32} - a_{22} + a_{12} = -1
\end{aligned}$$

que é satisfeito se $a_{42} = 0$ (da primeira igualdade), $a_{32} = 1$ (da segunda igualdade), $a_{22} = 2$ (da terceira igualdade) e $a_{12} = 0$, da última igualdade.

Da segunda equação observamos que

$$\begin{aligned}
&a_{43} = 2 \\
&a_{33} - 3a_{43} = 0 \\
&3a_{43} - 2a_{33} + a_{23} = -2 \\
&-a_{43} + a_{33} - a_{23} + a_{13} = 0
\end{aligned}$$

que é satisfeito se $a_{43} = 2$, $a_{33} = 6$ nas duas primeiras igualdades, $a_{23} = 4$ (na terceira igualdade) e $a_{13} = 0$ na última igualdade. Assim, o operador discreto é

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 1 & 6 \\ 0 & 0 & 2 \end{bmatrix}.$$

Testando o operador discreto $\mathbf{A}$ com $\mathbf{u} = 1 + 2x + 3x^2$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 1 & 6 \\ 0 & 0 & 2 \end{bmatrix} \begin{Bmatrix} 1 \\ 2 \\ 3 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 2\cdot 2 + 4\cdot 3 \\ 1\cdot 2 + 6\cdot 3 \\ 2\cdot 3 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 16 \\ 20 \\ 6 \end{Bmatrix}.$$

Isso quer dizer que $A(u)$ resulta em

$$0\cdot 1 + 16(x-1) + 20(x-1)^2 + 6(x-1)^3$$

Observem que a mudança de base provoca a alteração na forma discreta do operador A**. Isso será abordado em detalhes no próximo tópico**.

■

■ **Exemplo 6.15** Dado um vetor $\mathbf{u} = (x_1, x_2, x_3)$ no $\mathbb{R}^3$, podemos aplicar uma rotação em torno do eixo 3 para obter um novo vetor $\mathbf{v} = (x_1^{'}, x_2^{'}, x_3^{'})$. Esta transformação, também conhecida como rotação, é do tipo $T : \mathbb{R}^3 \to \mathbb{R}^3$ e tem a forma

$$T(x_1, x_2, x_3) = (x_1 \cos(\theta) - x_2 \sin(\theta),\, x_1 \sin(\theta) + x_2 \cos(\theta),\, x_3).$$

Como sabemos que as bases do $\mathbb{R}^3$ formam o conjunto $\{(1,0,0),(0,1,0),(0,0,1)\}$, se seguirmos o raciocínio apresentado no exemplo anterior, verificamos que o operador discreto associado a esta transformação pode ser obtido com

$$\begin{aligned}
T(1,0,0) &= (\cos(\theta), \sin(\theta), 0)\\
T(0,1,0) &= (-\sin(\theta), \cos(\theta), 0)\\
T(0,0,1) &= (0,0,1)
\end{aligned}$$

tal que

$$\mathbf{T} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

De fato,

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \begin{Bmatrix} x_1 \cos(\theta) - x_2 \sin(\theta) \\ x_1 \sin(\theta) + x_2 \cos(\theta) \\ x_3 \end{Bmatrix}.$$

■

Exercício 6.3 Verifique se as transformações dos exemplos acima tem inversa e, se tiverem, deduza a forma discreta do operador. ■

Exercício 6.4 Seja o operador, já na forma discreta,

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

tal que $T : \mathbb{R}^3 \to \mathbb{R}^3$. Qual é o *kernel* deste operador ? ■

Exemplo: Um polinômio cúbico tem a forma $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ e é membro de um espaço de dimensão 4, pois esta é a cardinalidade de sua base $U = \{1, x, x^2, x^3\}$. Vamos chamar este espaço de P. Na forma matricial, temos

$$\mathbf{p} = [1, x, x^2, x^3] \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{Bmatrix} = \mathbf{Ua}$$

onde diferentes valores de **a** caracterizam elementos distintos do espaço (pensem nas componentes de um vetor) e **U** é formado por colunas que possuem vetores linearmente independentes. Observem que uma coisa é o vetor (elemento do espaço dos polinômios cúbicos) $\mathbf{p} \in P$ e outra coisa é um dado valor que $p(x)$ assume para um valor de x, que é um valor real. Isto fica claro se escrevermos a relação

$$p(s) = \mathbf{C}(s)\mathbf{p}$$

onde $\mathbf{C}(s)$ é chamado de operador de avaliação em s e é um funcional linear em P, pois mapeia $\mathbf{p}$ para um número real. Assim,

$$p(s) = \mathbf{C}(s)\mathbf{U}\mathbf{a} = [1, s, s^2, s^3] \left\{ \begin{array}{c} a_0 \\ a_1 \\ a_2 \\ a_3 \end{array} \right\}.$$

É interessante notar que existem problemas práticos onde temos os valores $p(s)$ e queremos determinar quais são os polinômios $\mathbf{p}$ que estão associados a estes valores. Assim, vamos assumir que temos 4 valores distintos de s e seus correspondentes $p(s)$ conhecidos. Portanto,

$$[1, s_i, s_i^2, s_i^3] \left\{ \begin{array}{c} a_0 \\ a_1 \\ a_2 \\ a_3 \end{array} \right\} = p\left(s_i\right), i = 1..4$$

forma um sistema de equações na forma

$$\mathbf{E}(S)\mathbf{a} = \mathbf{v}$$

onde $S = \{s_1, s_2, s_3, s_4\}$ e $\mathbf{v} = \{p\left(s_1\right), p\left(s_2\right), p\left(s_3\right), p\left(s_4\right)\}$ são vetores coluna e $\mathbf{E}$ é a matriz onde cada linha contém os valores das bases avaliadas em um dos pontos conhecidos. Portanto, ao realizarmos a operação

$$\mathbf{a} = \mathbf{E}^{-1}(S)\mathbf{v}$$

estamos realizando um procedimento de interpolação (recuperação de uma função a partir de um conjunto de valores) e o operador $\mathbf{E}^{-1}$ é chamado de **Interpolação de Lagrange**.

6.3 Mudança de Base

Se um espaço pode ter mais de uma base e se espaços finito dimensionais admitem uma representação matricial, então é interessante investigar como podemos relacionar as bases de **um determinado espaço**. De forma geral, podemos definir um operador linear

$$T_{\Delta\Phi} : \mathbf{u}_\Delta \rightarrow \mathbf{u}_\Phi, \mathbf{u}_\Delta, \mathbf{u}_\Phi \in V,$$

em que $\mathbf{u}_\Delta$ é um elemento de V escrito na base Δ e $\mathbf{u}_\Phi$ o mesmo elemento descrito na base Φ. **Enfatiza-se que o elemento é um só, estando somente descrito em bases diferentes**.

Seja V um espaço $n-$dimensional, com $\mathbf{u}_\Delta = \sum_{i=1}^{n} \alpha_i \boldsymbol{\delta}_i$ é descrito na base Δ e $\mathbf{u}_\Phi = \sum_{j=1}^{n} \beta_j \boldsymbol{\phi}_j$ é descrito na base Φ. Definimos o operador **discreto** de mudança da base Δ para a base Φ como uma matriz $n \times n$

$$\mathbf{T}_{\Delta\Phi} = [\boldsymbol{\delta}_1|_\Phi \quad \dots \quad \boldsymbol{\delta}_n|_\Phi]$$

onde $\boldsymbol{\delta}_i|_\Phi$ significa que escrevemos a o termo de base $\boldsymbol{\delta}_i$ de Δ usando a base de Φ. É bom lembrar que tanto os termos de base de Δ quanto de Φ são escritos em função de uma base em comum, que chamaremos de canônica.

Demonstração. Seja um vetor $\mathbf{u}_\Delta$ escrito na base Δ

$$\mathbf{u}_\Delta = (\alpha_1, ..., \alpha_n) = \alpha_1 \boldsymbol{\delta}_1 + ... + \alpha_n \boldsymbol{\delta}_n.$$

Como o elemento é o mesmo, só escrito em bases diferentes,

$$T\left(\sum_{i=1}^{n} \alpha_i \boldsymbol{\delta}_i\right) = \sum_{j=1}^{n} \beta_j \boldsymbol{\phi}_j$$

implica em

$$\sum_{i=1}^{n} \alpha_i T(\boldsymbol{\delta}_i) = \sum_{j=1}^{n} \beta_j \boldsymbol{\phi}_j$$

e, como $T(\boldsymbol{\delta}_i)$ mapeia para a base Φ, podemos escrever

$$T(\boldsymbol{\delta}_i) = \sum_{k=1}^{n} t_{ki} \boldsymbol{\phi}_k,$$

ou seja, obtemos os valores de t_{ki} estudando combinações lineares na base destino que permitam descrever o termo da base original

$$\sum_{i=1}^{n} \alpha_i \left(\sum_{k=1}^{n} t_{ki} \boldsymbol{\phi}_k\right) = \sum_{j=1}^{n} \beta_j \boldsymbol{\phi}_j.$$

Com isso, podemos relacionar

$$\mathbf{T}_{\Delta\Phi} \mathbf{u}_\Delta = \mathbf{u}_\Phi.$$

■

A operação inversa é sempre definida (afinal, estamos no mesmo espaço), tal que

$$\mathbf{T}_{\Phi\Delta} \mathbf{u}_\Phi = \mathbf{u}_\Delta$$

com

$$\mathbf{T}_{\Phi\Delta} \mathbf{T}_{\Delta\Phi} = \mathbf{I}.$$

■ **Exemplo 6.16** Sejam duas bases no $\mathbb{R}^2$: $\Delta = [(1,2),(3,4)]$ e $\Phi = [(7,3),(4,2)]$. Para obtermos $\mathbf{T}_{\Delta\Phi}$, que mapeia da base Δ para a base Φ, iniciamos escrevendo $\boldsymbol{\delta}_1$ na base Φ

$$t_{11}(7,3) + t_{21}(4,2) = (1,2)|_\Phi$$

tal que

$$\begin{cases} 1 &= 7t_{11} + 4t_{21} \\ 2 &= 3t_{11} + 2t_{21} \end{cases}$$

que leva a $t_{11} = -3$ e $t_{21} = 5.5$. Escrevendo agora $\boldsymbol{\delta}_2$ na base Φ

$$t_{12}(7,3) + t_{22}(4,2) = (3,4)|_\Phi$$

obtemos $t_{12} = -5$ e $t_{22} = 9.5$. Com isso, obtemos

$$\mathbf{T}_{\Delta\Phi} = \begin{bmatrix} -3 & -5 \\ 5.5 & 9.5 \end{bmatrix}.$$

Para entendermos este operador, vamos considerar um vetor $\mathbf{u}$ na base canônica do $\mathbb{R}^2$, $[(1,0),(0,1)]$

$$\mathbf{u} = (-1,1).$$

Esse vetor pode ser escrito na base Δ como

$$\mathbf{u}_\Delta = (-1,1)_\Delta = \alpha_1(1,2) + \alpha_2(3,4)$$

tal que $\alpha_1 = 3.5$ e $\alpha_2 = -1.5$, com $\mathbf{u}_\Delta = (3.5,-1.5)$. Da mesma forma, podemos escrever o vetor $\mathbf{u}$ na base Φ

$$\mathbf{u}_\Phi = (-1,1)_\Phi = \beta_1(7,3) + \beta_2(4,2)$$

tal que $\beta_1 = -3$ e $\beta_2 = 5$, com $\mathbf{u}_\Phi = (-3,5)$.

O operador $\mathbf{T}_{\Delta\Phi}$ permite converter $\mathbf{u}_\Delta$ diretamente para $\mathbf{u}_\Phi$, pois

$$\begin{bmatrix} -3 & -5 \\ 5.5 & 9.5 \end{bmatrix} \begin{Bmatrix} 3.5 \\ -1.5 \end{Bmatrix} = \begin{Bmatrix} -3 \\ 5 \end{Bmatrix}.$$

Aqui cabe um comentário importante: qualquer vetor que multiplique o operador $\mathbf{T}_{\Delta\Phi}$ deve estar escrito na base Δ e não na canônica. Isso pode ser visto claramente se quisermos converter os vetores de base de Δ para a base de Φ. Começamos escrevendo $\boldsymbol{\delta}_1$ na base de Δ (está escrito na base canônica)

$$(1,2) = \alpha_1(1,2) + \alpha(3,4) = (1,0)$$

que é um resultado "óbvio" depois de calculado. Assim,

$$\begin{bmatrix} -3 & -5 \\ 5.5 & 9.5 \end{bmatrix} \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} = \begin{Bmatrix} -3 \\ 5.5 \end{Bmatrix} = \boldsymbol{\delta}_1|_\Phi,$$

que já foi calculado no início deste exemplo.

■

■ **Exemplo 6.17** Vamos considerar o mesmo exemplo anterior, mas agora vamos obter $\mathbf{T}_{\Phi\Delta}$. Iniciamos escrevendo $\boldsymbol{\phi}_1$ na base Δ

$$t_{11}(1,2) + t_{21}(3,4) = (7,3)|_\Delta$$

tal que

$$\begin{cases} 7 = 1t_{11} + 3t_{21} \\ 3 = 2t_{11} + 4t_{21} \end{cases}$$

que leva a $t_{11} = -9.5$ e $t_{21} = 5.5$. Escrevendo agora $\boldsymbol{\phi}_2$ na base Δ

$$t_{12}(1,2) + t_{22}(3,4) = (4,2)|_\Delta$$

obtemos $t_{12} = -5$ e $t_{22} = 3$. Com isso, obtemos

$$\mathbf{T}_{\Delta\Phi} = \begin{bmatrix} -9.5 & -5 \\ 5.5 & 3 \end{bmatrix}.$$

O operador $\mathbf{T}_{\Phi\Delta}$ permite converter $\mathbf{u}_\Phi$ diretamente para $\mathbf{u}_\Delta$, pois

$$\begin{bmatrix} -9.5 & -5 \\ 5.5 & 3 \end{bmatrix} \begin{Bmatrix} -3 \\ 5 \end{Bmatrix} = \begin{Bmatrix} 3.5 \\ -1.5 \end{Bmatrix}.$$

Novamente, podemos calcular $\boldsymbol{\phi}_1$ na base Φ

$$(7,3) = \beta_1(7,3) + \beta_2(4,2) = (1,0)$$

tal que

$$\begin{bmatrix} -9.5 & -5 \\ 5.5 & 3 \end{bmatrix} \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} = \begin{Bmatrix} -9.5 \\ 5.5 \end{Bmatrix} = \boldsymbol{\phi}_1|_\Delta.$$

Como esperado, os operadores $\mathbf{T}_{\Delta\Phi}$ e $\mathbf{T}_{\Phi\Delta}$ são relacionados por

$$\begin{bmatrix} -9.5 & -5 \\ 5.5 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} -3 & -5 \\ 5.5 & 9.5 \end{bmatrix}$$

.

■

■ **Exemplo 6.18** Seja a base canônica do $\mathbb{R}^2$

$$C = \{(1,0),(0,1)\}$$

e outra base

$$B = \{(1,0),(2,2)\}.$$

O operador de mudança de base de C para B é obtido com

$$B_1 = (1,0) = t_{11}(1,0) + t_{21}(0,1) \tag{6.1}$$
$$B_2 = (2,2) = t_{12}(1,0) + t_{22}(0,1) \tag{6.2}$$

que leva a

$$\mathbf{T}_{CB} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}.$$

Aqui, como o operador transforma da canônica para B, podemos usar os coeficientes de $\mathbf{u}_C$ diretamente. Assim

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

e

$$\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 2 \\ 2 \end{Bmatrix}.$$

Para obtermos o operador $\mathbf{T}_{BC}$

$$C_1 = (1,0) = t_{11}(1,0) + t_{21}(2,2) \tag{6.3}$$
$$C_2 = (0,1) = t_{12}(1,0) + t_{22}(2,2) \tag{6.4}$$

que leva a

$$\mathbf{T}_{BC} = \begin{bmatrix} 1 & -1 \\ 0 & \frac{1}{2} \end{bmatrix}, \tag{6.5}$$

de onde podemos verificar, novamente, que cada coluna de $\mathbf{T}_{BC}$ nada mais é do que o elemento da base canônica descrita na base B

$$1(1,0)+0(2,2)=(1,0)$$

e

$$-1(1,0)+0.5(2,2)=(0,1).$$

De fato,

$$\mathbf{T}_{BC}\left\{\begin{array}{c}1\\0\end{array}\right\}=\left\{\begin{array}{c}1\\0\end{array}\right\}$$

e

$$\mathbf{T}_{BC}\left\{\begin{array}{c}2\\2\end{array}\right\}=\left\{\begin{array}{c}0\\1\end{array}\right\}.$$

■

■ **Exemplo 6.19** Considere a mudança entre as bases polinomiais $\Delta=\{1,x,x^2,x^3\}$ e $\Phi=\{1,(x-1),(x-1)^2,(x-1)^3\}$.

Neste caso, podemos escrever a base Δ em função da base Φ

$$\begin{aligned}
\delta_1 &= 1 = t_{11}\phi_1+t_{21}\phi_2+t_{31}\phi_3+t_{41}\phi_4\\
\delta_2 &= x = t_{12}\phi_1+t_{22}\phi_2+t_{32}\phi_3+t_{42}\phi_4\\
\delta_3 &= x^2 = t_{13}\phi_1+t_{23}\phi_2+t_{33}\phi_3+t_{43}\phi_4\\
\delta_4 &= x^3 = t_{14}\phi_1+t_{24}\phi_2+t_{34}\phi_3+t_{44}\phi_4.
\end{aligned}$$

Substituindo a base Φ e organizando por temos em comum com a base Δ obtemos

$$\begin{aligned}
1 &= t_{41}x^3+(t_{31}-3t_{41})x^2+(t_{21}-2t_{31}+3t_{41})x+(t_{11}-t_{21}+t_{31}-t_{41})\\
x &= t_{42}x^3+(t_{32}-3t_{42})x^2+(t_{22}-2t_{32}+3t_{42})x+(t_{12}-t_{22}+t_{32}-t_{42})\\
x^2 &= t_{43}x^3+(t_{33}-3t_{43})x^2+(t_{23}-2t_{33}+3t_{43})x+(t_{13}-t_{23}+t_{33}-t_{43})\\
x^3 &= t_{44}x^3+(t_{34}-3t_{44})x^2+(t_{24}-2t_{34}+3t_{44})x+(t_{14}-t_{24}+t_{34}-t_{44})
\end{aligned}$$

com solução

$$\mathbf{T}_{\Phi\Delta}=\begin{bmatrix}1&1&1&1\\0&1&2&3\\0&0&1&3\\0&0&0&1\end{bmatrix}.$$

De fato, se utilizarmos esse operador de mudança de base para relacionar os resultados dos exemplos 2 e 3 da seção anterior

$$\mathbf{T}_{\Phi\Delta}\left\{\begin{array}{c}-2\\-6\\2\\6\end{array}\right\}_{\Phi}=\left\{\begin{array}{c}0\\16\\20\\6\end{array}\right\}_{\Delta}.$$

A operação inversa resulta em

$$\mathbf{T}_{\Delta\Phi} = \mathbf{T}_{\Phi\Delta}^{-1} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

tal que a operação inversa ao exemplo é

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} 0 \\ 16 \\ 20 \\ 6 \end{Bmatrix}_{\Delta} = \begin{Bmatrix} -2 \\ -6 \\ 2 \\ 6 \end{Bmatrix}_{\Phi} .$$

Pode-se avaliar que cada coluna de $\mathbf{T}_{\Delta\Phi}$ é um dos termos de base de Φ escrito na base de Δ. Por exemplo, se considerarmos a terceira coluna

$$(x-1)^2 = x^2 - 2x + 1 = (1, -2, 1, 0)$$

■

Exercício 6.5 Proponha dois conjuntos de bases para o $\mathbb{R}^2$ e deduza a forma matricial de $\mathbf{R}$ e de $\mathbf{R}^{-1}$. Após, valide a solução e, utilizando os conceitos de isomorfismo, verifique porque o operador de mudança de base é isomórfico. ■

6.4 Mudança de Base Aplicada a Operadores Lineares

Já vimos que a mudança de base faz com que um vetor $\mathbf{x}$, observado em uma base Δ e descrito por uma combinação linear na forma

$$\boldsymbol{x}_{\Delta} = \sum_{i=1}^{n} \alpha_i \boldsymbol{\delta}_i$$

passe a ser descrito por um novo conjunto de coeficientes β_j em uma nova base Φ, tal que

$$\boldsymbol{x}_{\Phi} = \sum_{j=1}^{n} \beta_j \boldsymbol{\phi}_j .$$

Esta operação é realizada, em um espaço finito dimensional, por um produto

$$\boldsymbol{x}_{\Phi} = \boldsymbol{T}_{\Delta\Phi} \boldsymbol{x}_{\Delta}$$

onde $\boldsymbol{T}_{\Delta\Phi}$ é o operador de mudança da base Δ para a base Φ na forma matricial. A operação inversa é descrita por

$$\boldsymbol{x}_{\Delta} = \boldsymbol{T}_{\Delta\Phi}^{-1} \boldsymbol{x}_{\Phi} = \boldsymbol{T}_{\Phi\Delta} \boldsymbol{x}_{\Phi} .$$

Se o observador que utiliza a base Δ aplicar uma operação linear $\boldsymbol{A}_{\Delta}$ sobre $\boldsymbol{x}_{\Delta}$, obtendo um vetor $\boldsymbol{s}_{\Delta}$ por meio da operação

$$\boldsymbol{s}_{\Delta} = \boldsymbol{A}_{\Delta} \boldsymbol{x}_{\Delta}$$

o observador que utiliza a base Φ irá realizar um operação equivalente, tal que

$$\boldsymbol{s}_{\Phi} = \boldsymbol{A}_{\Phi} \boldsymbol{x}_{\Phi} .$$

É interessante notar que podemos aplicar a mudança de base no operador $\boldsymbol{A}_\Delta$, obtendo diretamente o operador $\boldsymbol{A}_\Phi$, pois

$$\boldsymbol{s}_\Delta = \boldsymbol{A}_\Delta \boldsymbol{x}_\Delta \rightarrow \underbrace{\boldsymbol{T}_{\Phi\Delta}\boldsymbol{s}_\Phi}_{\boldsymbol{s}_\Delta} = \boldsymbol{A}_\Delta \underbrace{\boldsymbol{T}_{\Phi\Delta}\boldsymbol{x}_\Phi}_{\boldsymbol{x}_\Delta}$$

e, multiplicando a equação por $\boldsymbol{T}_{\Delta\Phi}$, obtemos

$$\boldsymbol{s}_\Phi = \underbrace{\boldsymbol{T}_{\Delta\Phi}\boldsymbol{A}_\Delta \boldsymbol{T}_{\boldsymbol{\Phi\Delta}}}_{\boldsymbol{A}_\Phi}\boldsymbol{x}_\Phi \rightarrow \boldsymbol{s}_\Phi = \boldsymbol{A}_\Phi \boldsymbol{x}_\Phi$$

tal que

$$\boxed{\boldsymbol{A}_\Phi = \boldsymbol{T}_{\Delta\Phi}\boldsymbol{A}_\Delta \boldsymbol{T}_{\Phi\Delta}}.$$

■ **Exemplo 6.20** Um operador linear $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ é definido por $A(\mathbf{u}) = (4u_1 + u_3, -2u_1 + u_2, -2u_1 + u_3)$ e é descrito na base canônica do $\mathbb{R}^3$ pela matriz

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

Supondo que o mesmo observador descreva um valor $\mathbf{x}$ na base canônica, como sendo

$$\mathbf{x}_C = \begin{Bmatrix} 1 \\ 2 \\ 3 \end{Bmatrix}$$

então a aplicação do operador no vetor fornece como resultado

$$\mathbf{s}_C = \mathbf{A}_C \mathbf{x}_C = \begin{Bmatrix} 7 \\ 0 \\ 1 \end{Bmatrix}.$$

Um outro observador, utilizando outra base $B = [(s, -s, 0), (s, s, 0), (0, 0, 1)]$ com $s = \frac{1}{\sqrt{2}}$, irá observar $\mathbf{x}$ com valores diferentes. O operador de mudança de base da canônica para a nova base é

$$\boldsymbol{T}_{CB} = \begin{bmatrix} s & s & 0 \\ -s & s & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

pois

$$\boldsymbol{T}_{CB} \begin{Bmatrix} 1 & | & 0 & | & 0 \\ 0 & | & 1 & | & 0 \\ 0 & | & 0 & | & 1 \end{Bmatrix} = \begin{Bmatrix} s & | & s & | & 0 \\ -s & | & s & | & 0 \\ 0 & | & 0 & | & 1 \end{Bmatrix}$$

Assim, o vetor observado na nova base será

$$\mathbf{x}_B = \boldsymbol{T}_{CB}\mathbf{x}_C = \begin{Bmatrix} 3\sqrt{2} \\ 1/\sqrt{2} \\ 3.0 \end{Bmatrix}.$$

O operador nesta nova base é obtido por

$$\boldsymbol{A}_B = \boldsymbol{T}_{CB}\boldsymbol{A}_C\boldsymbol{T}_{BC} = \begin{bmatrix} 1.5 & -0.5 & 1/\sqrt{2} \\ -2.5 & 3.5 & -1/\sqrt{2} \\ -\sqrt{2} & \sqrt{2} & 1.0 \end{bmatrix}$$

e este observador irá obter, como resultado da aplicação deste operador em $\mathbf{x}_B$

$$\mathbf{s}_B = \mathbf{A}_B\mathbf{x}_B = \begin{Bmatrix} \frac{7}{\sqrt{2}} \\ -\frac{7}{\sqrt{2}} \\ 1.0 \end{Bmatrix}.$$

Finalmente, podemos observar que cada observador obteve um resultado distinto em valores numéricos, mas que tem o mesmo significado físico. De fato, se desfizermos a operação de mudança de base iremos observar que

$$\mathbf{s}_C = \boldsymbol{T}_{BC}\mathbf{s}_B = \begin{Bmatrix} 7.0 \\ 0.0 \\ 1.0 \end{Bmatrix}.$$

É interessante notar que os vetores $\mathbf{s}_B$ e $\mathbf{s}_C$ tem a mesma norma $5\sqrt{2}$, indicando que são fisicamente iguais (pensando em um vetor força, por exemplo).

■

■ **Exemplo 6.21** Seja o operador linear $A : \mathbb{R}^2 \to \mathbb{R}^2$ com $A(\mathbf{u}) = (u_1 + u_2, 5u_1 - 3u_2)$. Considerando a base canônica $C = \{(1,0),(0,1)\}$ para o domínio e para o contra-domínio.

$$\mathbf{A}_C = \begin{bmatrix} 1 & 1 \\ 5 & -3 \end{bmatrix}$$

Agora vamos considerar outra base no $\mathbb{R}^2$: $B = \{(1,0),(2,2)\}$. O operador de mudança de base de B para C é

$$\mathbf{T}_{BC} = \begin{bmatrix} 1 & -1 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Assim, o mesmo operador linear, descrito na base B, será

$$\mathbf{A}_B = \mathbf{T}_{BC}\mathbf{A}_B\mathbf{T}_{CB} = \begin{bmatrix} 1 & -1 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 5 & -3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & \frac{1}{2} \end{bmatrix}^{-1} = \begin{bmatrix} -4 & 0 \\ \frac{5}{2} & 2 \end{bmatrix}.$$

É interessante notar que o traço (soma dos termos da diagonal) do operador A não se altera, pois $Tr(\mathbf{A}_C) = Tr(\mathbf{A}_B) = -2$. Essa medida é chamada de primeiro invariante do operador.

■

6.5 Formas lineares, bilineares e quadráticas

Existem diversos casos na área de mecânica em que necessitamos de relações lineares, bilineares e quadráticas. Um exemplo de relação quadrática se dá no cálculo da energia de deformação, que é função quadrática das deformações. De forma esquemática: espaço das deformações $\to$ forma quadrática $\to$ energia de deformação. Desta forma, é importante estudarmos os conceitos envolvidos e suas propriedades.

6.5.1 Funcional Linear

Vamos considerar um espaço vetorial U sob o campo dos reais e um operador linear $l: U \to \mathbb{R}$. O operador l nada mais é do que uma transformação que mapeia para um espaço em particular, no caso $\mathbb{R}$ (ou seja, um escalar). Neste contexto, l é chamado de um **funcional linear** ou uma **forma linear**. Na literatura, os funcionais são muitas vezes chamados de **funções de funções**.

■ **Exemplo 6.22** Funcionais Lineares em $\mathbb{R}^n$

Neste caso mais simples, podemos definir um funcional na forma

$$l: \mathbb{R}^n \to \mathbb{R},\, l(\mathbf{u}) = \sum_{i=1}^{n} a_i u_i$$

onde a_i são coeficientes. ■

■ **Exemplo 6.23** Integração

O operador de integração pode ser visto como um funcional linear, pois

$$l: U[a,b] \to \mathbb{R},\, l(f) = \int_a^b f(x)\,dx,$$

e sabemos que o operador de integração é homogêneo e aditivo (portanto linear). ■

6.5.2 Forma Bilinear

Outro tipo de operador muito utilizado é a forma bilinear:

Definição 6.5.1 Forma Bilinear

Sejam U e V espaços vetoriais com $\mathbf{u}$, $\mathbf{u}_1$, $\mathbf{u}_2 \in U$, $\mathbf{v}$, $\mathbf{v}_1$, $\mathbf{v}_2 \in V$ e $\alpha \in \mathbb{R}$. Definimos uma forma bilinear B como

$$B: U \times V \to \mathbb{R}$$

que satisfaz:

- $B(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) = B(\mathbf{u}_1, \mathbf{v}) + B(\mathbf{u}_2, \mathbf{v})$
- $B(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) = B(\mathbf{u}, \mathbf{v}_1) + B(\mathbf{u}, \mathbf{v}_2)$
- $B(\lambda \mathbf{u}, \mathbf{v}) = B(\mathbf{u}, \lambda \mathbf{v}) = \lambda B(\mathbf{u}, \mathbf{v})$

É interessante notar que se U e V forem finitos, tais que

$$\mathbf{u} = \sum_{i}^{|U|} \alpha_i \boldsymbol{\phi}_i$$

e

$$\mathbf{v} = \sum_{j}^{|V|} \beta_j \boldsymbol{\varphi}_j$$

então

$$B(\mathbf{u}, \mathbf{v}) = \sum_i \sum_j \alpha_i \beta_j B(\boldsymbol{\phi}_i, \boldsymbol{\varphi}_j) = \sum_i \sum_j \alpha_i \beta_j b_{ij}$$

que permite a representação matricial

$$B(\mathbf{u}, \mathbf{v}) = \boldsymbol{\alpha}^{\mathbf{T}} \mathbf{B} \boldsymbol{\beta}$$

onde a forma discreta do operador B tem dimensão $|U| \times |V|$, com os seguintes resultados:

Definição 6.5.2 Uma forma bilinear $B(\mathbf{u},\mathbf{v}) : U \times V \to \mathbb{R}$ é dita:
- simétrica se $B(\mathbf{u},\mathbf{v}) = B(\mathbf{v},\mathbf{u})$
- anti-simétrica (skew) $B(\mathbf{u},\mathbf{u}) = 0$, $\forall \mathbf{u} \in U$

de tal forma que
- Se B é simétrica, então $b_{ij} = b_{ji}$
- Se B é skew, então $b_{ij} = -b_{ji}$

Nada melhor do que alguns exemplos:

■ **Exemplo 6.24** Exemplos de formas bilineares
1. $B : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $B(x,y) = xy$
2. $B : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $B((x,y),(z,t)) = xz - 2yt$

■

Exercício 6.6 Encontre a forma matricial para os operadores bilineares do exemplo anterior. ■

■ **Exemplo 6.25** Exemplo de forma bilinear

Seja $B : U \times V \to \mathbb{R}$, com U espaço dos polinômios de terceiro grau, de base $\{1, x, x^2, x^3\}$, e V espaço dos polinômios de segundo grau, de base $\{1, x, x^2\}$. Assim, se definirmos, por exemplo, a forma bilinear como

$$B(\mathbf{u},\mathbf{v}) = \int_0^1 (\mathbf{uv})\,dx = \int_0^1 \left[\left(a_1 + a_2 x + a_3 x^2 + a_4 x^3\right)\left(b_1 + b_2 x + b_3 x^2\right)\right] dx.$$

Se solucionarmos esta integral veremos que, ao colocar os temos com b_j em evidência, iremos obter

$$b_1)\ \frac{1}{60}(60a_1 + 30a_2 + 20a_3 + 15a_4) \tag{6.6}$$

$$b_2)\ \frac{1}{60}(30a_1 + 20a_2 + 15a_3 + 12a_4) \tag{6.7}$$

$$b_3)\ \frac{1}{60}(20a_1 + 15a_2 + 12a_3 + 10a_4)$$

tal que $B(\mathbf{u},\mathbf{v})$ pode ser colocada na forma matricial

$$\mathbf{B} = \frac{1}{60}\begin{bmatrix} 60 & 30 & 20 \\ 30 & 20 & 15 \\ 20 & 15 & 12 \\ 15 & 12 & 10 \end{bmatrix}$$

tal que

$$B(\mathbf{u},\mathbf{v}) = \left\{\ a_1 \quad a_2 \quad a_3 \quad a_4\ \right\} \frac{1}{60}\begin{bmatrix} 60 & 30 & 20 \\ 30 & 20 & 15 \\ 20 & 15 & 12 \\ 15 & 12 & 10 \end{bmatrix} \left\{ \begin{matrix} b_1 \\ b_2 \\ b_3 \end{matrix} \right\}.$$

■

6.5.3 Formas Quadráticas

Uma forma quadrática é um funcional $Q(\mathbf{u})$ tal que

$$Q(\alpha\mathbf{u}) = \alpha^2 Q(\mathbf{u})$$

e, por esta definição, observamos que se B é uma forma bilinear simétrica

$$Q(\mathbf{u}) = B(\mathbf{u},\mathbf{u}) = B_s(\mathbf{u},\mathbf{u})$$

tal que se **duas formas bilineares tem a mesma parte simétrica então geram a mesma forma quadrática**.

■ **Exemplo 6.26** Formas quadráticas

1. $B : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, $B(x,y) = xy$ gera a forma quadrática $Q(x) = x^2$;
2. $B : \mathbb{R}^2\times\mathbb{R}^2 \to \mathbb{R}$, $B((x,y),(z,t)) = xz - 2yt$ gera a forma quadrática $Q((x,y)) = x^2 - 2y^2$.

■

Exercício 6.7 Encontre a forma matricial para as formas quadráticos do exemplo anterior. ■

■ **Exemplo 6.27** Forma quadrática

$B : U \times V \to \mathbb{R}$, com U e V espaço dos polinômios de terceiro grau. Vamos definir a seguinte forma bilinear

$$B(\mathbf{u},\mathbf{v}) = \int_0^1 (3\mathbf{u}\cdot\mathbf{v})\,dx$$

que gera a forma quadrática

$$Q(\mathbf{u}) = \int_0^1 (3\mathbf{u}\cdot\mathbf{u})\,dx$$

ou, na forma discreta

$$\frac{1}{140}\begin{bmatrix} 420 & 210 & 140 & 105 \\ 210 & 140 & 105 & 84 \\ 140 & 105 & 84 & 70 \\ 105 & 84 & 70 & 60 \end{bmatrix}$$

■

6.6 Continuidade de Operadores

No capítulo sobre funções discutimos sobre funções contínuas. Essa definição pode ser estendida para operadores:

Definição 6.6.1 Continuidade de um Operador

Seja $T : U \to V$ um operador entre os espaços normados U e V. O operador é dito contínuo em U se existe um número $M > 0$ tal que

$$\|T\mathbf{u}_1 - T\mathbf{u}_2\|_V \le M\,\|\mathbf{u}_1 - \mathbf{u}_2\|_U\,,\ \forall \mathbf{u}_1,\mathbf{u}_2 \in U$$

e deve ficar claro que a continuidade de um operador depende das normas $||\cdot||_U$ e $||\cdot||_V$ empregadas.

Definição 6.6.2 Operador Limitado

Um operador T é dito limitado (*bounded*) se existe $M > 0$ tal que

$$\|T\mathbf{u}\|_V \le M\,\|\mathbf{u}\|_U\,,\ \forall \mathbf{u} \in U$$

e, se o operador for linear, então ser contínuo e ser limitado significam a mesma coisa.

Demonstração. Da definição de operador limitado temos que

$$\|T\mathbf{u}\|_V \leq M\,\|\mathbf{u}\|_U\,,\ \forall \mathbf{u} \in U.$$

Se assumirmos que $\mathbf{u} = \mathbf{u}_1 - \mathbf{u}_2$

$$\|T(\mathbf{u}_1 - \mathbf{u}_2)\|_V \leq M\,\|\mathbf{u}_1 - \mathbf{u}_2\|_U$$

e, se o operador for linear

$$\|T(\mathbf{u}_1) - T(\mathbf{u}_2)\|_V \leq M\,\|\mathbf{u}_1 - \mathbf{u}_2\|_U$$

que implica em ser contínuo.

■

■ **Exemplo 6.28** Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ uma transformação linear definida por $T(\mathbf{u}) = (u_1 + u_2, u_2 - u_3, u_3)$. Considerando a norma 2, temos que

$$\|T\mathbf{u}\|_2 = \left((u_1+u_2)^2 + (u_2-u_3)^2 + u_3^2\right)^{\frac{1}{2}}$$

e

$$\|\mathbf{u}\|_2 = \left(u_1^2 + u_2^2 + u_3^2\right)^{\frac{1}{2}}.$$

Assim

$$\left((u_1+u_2)^2 + (u_2-u_3)^2 + u_3^2\right)^{\frac{1}{2}} \leq M\left(u_1^2 + u_2^2 + u_3^2\right)^{\frac{1}{2}}$$

Elevando ambos os lados ao quadrado, temos

$$\left((u_1+u_2)^2 + (u_2-u_3)^2 + u_3^2\right) \leq M^2\left(u_1^2 + u_2^2 + u_3^2\right)$$

que pode ser re-escrito como

$$u_1^2 + 2u_2^2 + 2u_3^2 + 2u_1u_2 - 2u_2u_3 \leq M^2\left(u_1^2 + u_2^2 + u_3^2\right).$$

É interessante notar que para $M \geq 2$ a desigualdade é satisfeita (caso limite quando $u_2 \to \infty$ ou quando $u_3 \to \infty$), de tal forma que o operador é limitado.

■

■ **Exemplo 6.29** Seja o espaço dos polinômios $p(x)$ em $x \in [0,1]$, com a base canônica $B = \{x^n\}$, $n \in \mathbb{N}$. Considere o operador

$$T(p) = \frac{dp}{dx}.$$

Assim, considerando um monômio qualquer x^n, temos que

$$||T(p)||_2^2 = ||nx^{n-1}||_2^2 = \int_0^1 \left(nx^{n-1}\right)^2 dx = \frac{n^2}{2n-1}$$

e

$$||x^n||_2^2 = \int_0^1 x^{2n} dx = \frac{1}{2n-1}.$$

Com isso, vemos que esse operador não é limitado, pois

$$n^2 \leqslant M^2$$

não permite estabelecer um valor finito para M.

■

7. Autovalores e Autovetores

Se A é um operador escrito em termos de uma base qualquer, então existe uma base no qual a representação discreta de A, $\mathbf{A}$, é diagonal. No $\mathbb{R}^2$, por exemplo, isto é facilmente visualizado na operação $T : U \to V$ representada pela matriz

$$\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

tal que

$$\mathbf{Au} = \begin{Bmatrix} au_1 + bu_2 \\ bu_1 + cu_2 \end{Bmatrix}.$$

Uma outra forma de apresentar a MESMA operação, porém em outra base, seria realizada com o operador $\mathbf{A}^{'}$ diagonal, tal que

$$\mathbf{A}^{'} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

tal que

$$\mathbf{A}^{'}\mathbf{u}^{'} = \begin{Bmatrix} \lambda_1 u_1^{'} \\ \lambda_2 u_2^{'} \end{Bmatrix}.$$

A forma diagonal de $\mathbf{A}$, $\mathbf{A}^{'}$ é conhecida como *forma canônica* do operador (neste caso, no $\mathbb{R}^2$). Essa nomenclatura não deve ser confundida com a base canônica, como veremos mais adiante.

Como desenvolvido no Capítulo 6.4, sabemos que a mudança de base de um operador é realizada pela operação

$$\mathbf{A}^{'} = \mathbf{TAT}^{-1},$$

onde $\mathbf{T}$ é o operador de mudança de base original para a que diagonaliza o operador. O problema aqui é que não conhecemos o operador de mudança de base que permita obter a forma diagonal do operador e nem os valores de λ. A determinação desta mudança de base específica e do operador diagonal associado, $\mathbf{A}^{'}$ é o objetivo de uma classe de problemas muito útil, chamada de problema de autovalores e autovetores.

Definição 7.0.1 Problema de autovalores e autovetores
Considere a transformação $T : U \to V$, na forma

$$T(\mathbf{u}) = \lambda \mathbf{u}.$$

Assumindo que U e V são finito dimensionais, com a mesma cardinalidade n. Como toda a transformação linear finito dimensional pode ser escrita pela multiplicação de uma matriz por um vetor, então

$$T(\mathbf{u}) : \mathbf{Tu} = \lambda \mathbf{u}$$

e, neste caso estamos dizendo que a aplicação de $\mathbf{T}$ sobre $\mathbf{u}$ é equivalente a multiplicação de um escalar λ sobre o mesmo $\mathbf{u}$. De forma equivalente

$$(\mathbf{T} - \lambda \mathbf{I})\,\mathbf{u} = \mathbf{0}.$$

Assim, descartando a solução trivial $\mathbf{u} = \mathbf{0}$, o problema homogêneo definido pela equação acima tem solução se e somente se

$$\det(\mathbf{T} - \lambda \mathbf{I}) = 0$$

resultando em um polinômio $p(\lambda)$ de grau n. As raízes λ_i são chamadas de *autovalores* (ou valores característicos) do operador e os vetores $\mathbf{u}_i$ que satisfazem a igualdade

$$\mathbf{Tu_i} = \lambda_i \mathbf{u}_i$$

são chamados de autovetores do operador.

■ **Exemplo 7.1** Seja o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$

$$T(x,y,z) = (3x - y + z, -x + 5y - z, x - y + 3z)$$

que, na base canônica do $\mathbb{R}^3$, tem a forma matricial

$$\mathbf{T} = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix},$$

pois

$$T(x,y,z) = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{Bmatrix} x \\ y \\ z \end{Bmatrix}.$$

A equação característica associada a este operador é obtida por meio da operação

$$\det\left(\begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) = 0$$

resultado em

$$-\lambda^3 + 11\lambda^2 - 36\lambda + 36 = 0$$

com raízes $\lambda_1 = 2$, $\lambda_2 = 3$ e $\lambda_3 = 6$. Os autovetores associados a cada um dos autovalores são obtidos por

$$\lambda_1 : \begin{bmatrix} 3-2 & -1 & 1 \\ -1 & 5-2 & -1 \\ 1 & -1 & 3-2 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix}_1 = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

tal que

$$\begin{aligned} u_1 \; -u_2 \; +u_3 &= 0 \quad (7.1)\\ -u_1 \; +3u_2 \; -u_3 &= 0 \quad (7.2)\\ u_1 \; -u_2 \; +u_3 &= 0 \end{aligned}$$

tem infinitas soluções, na forma

$$\begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix}_1 = \begin{Bmatrix} -a \\ 0 \\ a \end{Bmatrix}, \forall a \in \mathbb{R}$$

e

$$\lambda_2 : \begin{bmatrix} 3-3 & -1 & 1 \\ -1 & 5-3 & -1 \\ 1 & -1 & 3-3 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix}_2 = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

resulta em

$$\begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix}_2 = \begin{Bmatrix} b \\ b \\ b \end{Bmatrix}, \forall b \in \mathbb{R}$$

e finalmente,

$$\lambda_{3:} \begin{bmatrix} 3-6 & -1 & 1 \\ -1 & 5-6 & -1 \\ 1 & -1 & 3-6 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix}_3 = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

resulta em

$$\begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix}_3 = \begin{Bmatrix} c \\ -2c \\ c \end{Bmatrix}, \forall c \in \mathbb{R}.$$

Por fim, podemos agrupar os três vetores em uma matriz $\mathbf{\Phi}$

$$\mathbf{\Phi} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

tem colunas que satisfazem as soluções obtidas e que são $L.I$ (verifique). Os valores de a, b e c foram escolhidos para que a norma 2 de cada um dos vetores base fosse unitária.

■

Definição 7.0.2 Mudança de base provocada pelos autovetores
Da definição do problema de autovalores e autovetores, vimos que

$$\mathbf{Tu_i} = \lambda_i \mathbf{u}_i\,, \mathbf{u}_i \neq \mathbf{0}\,, i = 1..n$$

Definimos a matriz $\mathbf{\Phi} = [\mathbf{u}_1 \mathbf{u}_2 ... \mathbf{u}_n]$, onde cada coluna é um autovetor e assumindo, por hora, que todos os autovetores são linearmente independentes. Neste caso, observamos que $\mathbf{\Phi}$ *spans* o espaço e, portanto, serve como conjunto de base para o espaço. O operador $\mathbf{T}$ pode ser então reescrito na base $\mathbf{\Phi}$, gerando $\mathbf{T}'$. Os coeficientes t'_{ij} podem ser obtidos com

$$\mathbf{Tu}_j = \sum_{i=1}^{n} t'_{ij} \mathbf{u}_i.$$

No entanto, $\mathbf{Tu}_j = \lambda_j \mathbf{u}_j$, tal que

$$\lambda_j \mathbf{u}_j = \sum_{i=1}^{n} t'_{ij} \mathbf{u}_i.$$

implica em $t'_{ij} = \lambda_j \delta_{ij}$, tal que

$$\mathbf{T}_{\Phi ii} = \lambda_i\,, i = 1..n.$$

Com isso, verificamos que o operador $\mathbf{\Phi}$ mapeia da base que diagonaliza para a base canônica, de tal forma que

$$\mathbf{u} = \mathbf{\Phi u}'$$

o que leva a

$$\mathbf{A}' = \mathbf{\Phi}^{-1} \mathbf{A \Phi}.$$

■ **Exemplo 7.2** Retornando ao problema do exemplo anterior. Conforme a teoria de mudança de bases associada aos autovetores, observamos que

$$\mathbf{\Phi}^{-1}\mathbf{T\Phi} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

é a forma diagonal de $\mathbf{T}$.

■

■ **Exemplo 7.3** Um operador linear é descrito, utilizando a base canônica do $\mathbb{R}^3$, na forma

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

tem autovalores $\lambda_1 = 1$, $\lambda_2 = 3$ e $\lambda_3 = 2$ e autovetores

$$\mathbf{\Phi} = \begin{bmatrix} 0 & \frac{\sqrt{3}}{3} & -\frac{1}{3} \\ 1 & -\frac{\sqrt{3}}{3} & \frac{2}{3} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{2}{3} \end{bmatrix}.$$

Supondo que o mesmo observador descreva um valor $\mathbf{x}$ na base canônica, como sendo

$$\mathbf{x} = \begin{Bmatrix} 1 \\ 2 \\ 3 \end{Bmatrix}$$

então a aplicação do operador no vetor fornece como resultado

$$\mathbf{s} = \mathbf{A}\mathbf{x} = \begin{Bmatrix} 7 \\ 0 \\ 1 \end{Bmatrix}.$$

Um outro observador, utilizando a base que diagonaliza o operador (autovetores) irá observar $\mathbf{x}$ como sendo (lembre-se que $\mathbf{\Phi}$ mapeia de $\mathbf{x}'$ para $\mathbf{x}$)

$$\mathbf{x}' = \mathbf{\Phi}^{-1}\mathbf{x} = \begin{Bmatrix} -1 \\ 8,66 \\ 12 \end{Bmatrix}$$

e este observador irá obter, como resultado da aplicação do operador diagonalizado em $\mathbf{x}'$

$$\mathbf{s}' = \mathbf{A}'\mathbf{x}' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{Bmatrix} -1 \\ 8,66 \\ 12 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 25,98 \\ 24 \end{Bmatrix}.$$

Finalmente, podemos observar que cada observador obteve um resultado descrito em sua base, mas como sabemos que o operador de mudança de base $\mathbf{\Phi}$ permite descrever um vetor da base diagonalizada na base canônica, verificamos que

$$\mathbf{s} = \mathbf{\Phi}\mathbf{s}' = \begin{Bmatrix} 7 \\ 0 \\ 1 \end{Bmatrix}.$$

■

Exercício 7.1 Comente a frase: Um autovetor $\mathbf{u}$ pertence ao *kernel* do operador $(\mathbf{T} - \lambda \mathbf{I})$ ■

Exercício 7.2 Determine os autovalores e autovetores dos operadores:
- $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (3x, 8x - y)$
- $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (3x, 2y, y + 2z)$

■

7.1 Autovalores e Autovetores Reais

Como os autovalores são raízes do polinômio característico, fica claro que, sob determinadas condições, podemos ter raízes (autovalores) complexos. Da mesma, forma, ao solucionarmos o sistema linear homogêneo com λ complexo, podemos ter autovetores complexos. No entanto, se a matriz for Hermitiana, podemos provar que os autovalores serão reais.

Definição 7.1.1 Matriz Hermitiana
Uma matriz $\mathbf{A}$ é dita Hermitiana se for auto-adjunta, ou seja, se

$$\mathbf{A} = \mathbf{A}^{\mathrm{H}}$$

onde $\mathbf{A}^{\mathrm{H}}$ é o conjugado transposto da matriz $(a_{ij} = a^*_{ji})$, ou seja, as posições diagonais são reais e as posições fora da diagonal são tais que $a_{ij} + b_{ij}i = a_{ji} - b_{ji}i$. Se a matriz é real, o complexo conjugado se reduz ao fato de a matriz ser simétrica, isto é, $\mathbf{A} = \mathbf{A}^T$.

■ **Exemplo 7.4** A matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2+i & -1 \\ 2-i & 3 & -3i \\ -1 & 3i & 2 \end{bmatrix}$$

é Hermitiana pois os termos da diagonal são reais e $a_{ij} = a_{ji}^*$. ■

Com isso, podemos provar que os autovalores de uma matriz Hermitiana são reais.

Demonstração. Os autovalores de uma matriz Hermitiana são reais

Partindo da definição

$$\mathbf{Au} = \lambda \mathbf{u},$$

podemos multiplicar ambos os lados por $\mathbf{u}^H$ tal que

$$\mathbf{u}^H\mathbf{Au} = \lambda \mathbf{u}^H\mathbf{u} = \lambda ||\mathbf{u}||^2.$$

Se utilizarmos a relação $\mathbf{u}^H\mathbf{Au} = \mathbf{u}^T\mathbf{A}^T\mathbf{u}^*$, obtemos

$$\mathbf{u}^T\mathbf{A}^T\mathbf{u}^* = \lambda ||\mathbf{u}||^2.$$

Podemos aplicar o complexo conjugado a ambos os lados, obtendo

$$\mathbf{u}^H\mathbf{A}^H\mathbf{u} = \lambda^* ||\mathbf{u}||^2.$$

e como assumimos que a matriz $\mathbf{A}$ é Hermitiana, $\mathbf{A}^H = \mathbf{A}$ tal que

$$\mathbf{u}^H\mathbf{A}^H\mathbf{u} = \mathbf{u}^H \underbrace{\mathbf{Au}}_{\lambda \mathbf{u}} = \lambda^* ||\mathbf{u}||^2.$$

Por fim, verificamos que a ultima igualdade leva a

$$\lambda ||\mathbf{u}||^2 = \lambda^* ||\mathbf{u}||^2$$

que implica em $\lambda = \lambda^*$, ou seja, λ é real. ■

■ **Exemplo 7.5** A matriz Hermitiana

$$\mathbf{A} = \begin{bmatrix} 1 & 2+i & -1 \\ 2-i & 3 & -3i \\ -1 & 3i & 2 \end{bmatrix}$$

tem autovalores $\lambda_1 = -1,9359708$, $\lambda_2 = 1,8748776$ e $\lambda_3 = 6,0610932$ com autovetores

$$\mathbf{\Phi} = \begin{bmatrix} 0,4666037 - 0,3527762i & -0,4221890 + 0,6140749i & -0,0032597 - 0,3201372i \\ -0,1175921 + 0,5766664i & 0,2046916 + 0,1662744i & -0,1067124 - 0,7567659i \\ 0,5580841 & 0,6124726 & 0,5598387 \end{bmatrix},$$

que, como esperado, diagonaliza a matriz. Observe que os autovalores são reais, mas os autovetores não são, necessariamente, reais. ■

7.2 Operador Diagonalizável e Multiplicidade de Autovalores

Nos exercícios propostos na seção anterior podemos verificar que é possível obter autovalores idênticos (múltiplos autovalores). Por exemplo, $T : \mathbb{R}^3 \to \mathbb{R}^3, T(x,y,z) = (3x, 2y, y+2z)$ tem autovalores $\lambda_1 = \lambda_2 = 2$ e $\lambda_3 = 3$, de tal forma que o autovalor 2 aparece duas vezes. Neste caso, dizemos que existe uma Multiplicidade Algébrica do autovalor.

Definição 7.2.1 Multiplicidade Algébrica de um autovalor
Seja um operador $\mathbf{A}$ e um autovalor λ. A Multiplicidade Algébrica do autovalor é o número de vezes em que este autovalor é repetido.

Outra definição importante é:

Definição 7.2.2 Multiplicidade Geométrica de um autovalor
Seja um operador $\mathbf{A}$ e um autovalor λ. A Multiplicidade Geométrica de λ é a *nullity* do operador $(\mathbf{A} - \lambda\mathbf{I})$.
A multiplicidade Geométrica é sempre menor ou igual a multiplicidade algébrica.

Com essas definições, podemos estabelevcer as condições para um operador ser diagonalizavel

Definição 7.2.3 Operador Diagonalizável
Seja um operador $\mathbf{A}$. O operador é diagonalizável se, para cada autovalor, a Multiplicidade Algébrica for igual a Multiplicidade Geométrica.

■ **Exemplo 7.6** O operador

$$\mathbf{A} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

tem como autovalores $\lambda_1 = \lambda_2 = 2$ e $\lambda_3 = 8$ e seus autovetores são

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -2 & 0 & 1 \end{bmatrix},$$

que são linearmente independentes. Substituindo o autovalor $\lambda_{1,2}$ na expressão $\mathbf{A} - \lambda\mathbf{I}$, obtemos a matriz

$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

que tem *Rank* igual a 1 e *nullity* igual a 2. Assim, verificamos que a **Multiplicidade Geométrica** do autovalor $\lambda_{1,2} = 2$ é dada pela dimensão do espaço nulo de $\mathbf{A} - \lambda\mathbf{I}$, ou seja, 2 (também chamado de número de parâmetros livres do sistema de equações). Observe que a multiplicidade geométrica do autovalor é, neste caso, igual ao número de vezes em que o autovalor aparece (se repete). Da mesma forma, substituindo o autovalor $\lambda_3 = 8$ na expressão $\mathbf{A} - \lambda\mathbf{I}$, obtemos a matriz

$$\mathbf{A} - 8\mathbf{I} = \begin{bmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{bmatrix}$$

que tem *Rank* igual a 2 e *nullity* igual a 1. Assim, a multiplicidade geométrica deste autovalor é igual a 1 que é, neste caso, igual ao número de vezes em que o autovalor se repete (**Multiplicidade Algébrica**). Devido ao fato de a multiplicidade geométrica de **todos** os autovalores desta matriz serem iguais ao número de vezes em que o autovalor se repete, verificamos que é possível utilizar os autovetores como um operador de mudança de base, ou seja, diagonalizar a matriz **A** por meio da operação $\mathbf{\Phi}^{-1}\mathbf{A}\mathbf{\Phi}$.

■

■ **Exemplo 7.7** O operador

$$\mathbf{B} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

tem como autovalores $\lambda_1 = \lambda_2 = 2$ e $\lambda_3 = 3$. Sua matriz de autovetores é dada por

$$\mathbf{\Phi} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix},$$

que possui duas colunas linearmente dependentes. substituindo o autovalor $\lambda_1 = \lambda_2 = 2$ na expressão $\mathbf{B} - \lambda\mathbf{I}$, obtemos a matriz

$$\mathbf{B} - 2\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

que tem *Rank* igual a 2 e *nullity* igual a 1. Assim, a multiplicidade geométrica deste autovalor é 1 que, neste caso, não é igual ao número de vezes no qual o autovalor aparece, 2. Por sua vez, a substituição do autovalor $\lambda_3 = 3$ na expressão $\mathbf{B} - \lambda\mathbf{I}$ resulta na matriz

$$\mathbf{B} - 3\mathbf{I} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

que também tem *Rank* igual a 2 e *nullity* igual a 1, tal que sua multiplicidade geométrica é 1 que, neste caso, é igual ao número de vezes em que o autovalor é repetido. Desta forma, devido ao fato de não termos a igualdade entre a multiplicidade geométrica e o número de vezes em que o autovalor é repetido para **todos** os autovalores, verificamos que o operador não é diagonalizável pela operação $\mathbf{\Phi}^{-1}\mathbf{B}\mathbf{\Phi}$.

■

A explicação para a possibilidade ou não de podermos diagonalizar o operador fica clara com a definição

Definição 7.2.4 Subespaços Próprios
O conjunto U_λ dos autovetores associados a um autovalor λ é definido por

$$U_\lambda = \{\mathbf{X} : \mathbf{A}\mathbf{X} = \lambda\mathbf{X}\}.$$

Estes autovetores, se linearmente independentes, com a dimensão e número correto, podem formar (*span*) uma base para um subespaço, chamado de Subespaço Próprio do autovalor λ.

■ **Exemplo 7.8** O operador

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 0 & 0 \end{bmatrix}$$

tem autovalores $\lambda_1 = \lambda_2 = 0$ e $\lambda_3 = 1$. Os autovetores associados a estes autovalores são

$$\mathbf{\Phi} = \begin{bmatrix} 0 & 0 & 0,2672 \\ 0 & 1 & 0,5345 \\ 1 & 0 & 0,8018 \end{bmatrix}.$$

Avaliando-se a dimensão do espaço nulo de $\mathbf{A} - 0\mathbf{I}$, obtém-se 2, que é igual a multiplicidade algébrica deste autovalor. Com isto, pode-se verificar que os autovetores $(0,0,1)$ e $(0,1,0)$ formam uma base (os vetores são L.I.) para um espaço de dimensão 2, tal que o conjunto de vetores que formam U_0 é dado por

$$U_0 = \left\{ a \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix} + b \begin{Bmatrix} 0 \\ 1 \\ 0 \end{Bmatrix}, a, b \in \mathbb{R} \right\}$$

assim como o autovetor associado ao autovalor $\lambda_3 = 1$, que forma o conjunto

$$U_1 = \left\{ a \begin{Bmatrix} 0,2672 \\ 0,5345 \\ 0,8018 \end{Bmatrix}, a \in \mathbb{R} \right\}.$$

■

■ **Exemplo 7.9** O operador

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

tem 3 autovalores iguais a 5 com autovetores

$$\mathbf{\Phi} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

A multiplicidade geométrica deste autovalor é 1 (verifique), tal que o mesmo não é diagonalizável. Outra consequência importante é o fato de os autovetores associados ao autovalor não formarem uma base adequada para o subespaço próprio de $\lambda = 5$, pois os mesmos são linearmente dependentes. De fato, podemos verificar que o subespaço próprio tem apenas uma dimensão, pois

$$U_5 = \left\{ a \begin{Bmatrix} 1 \\ 0 \\ 0 \end{Bmatrix}, a \in \mathbb{R} \right\},$$

o que justifica a falta de informações para realizar a diagonalização do operador.

■

7.3 Problema Generalizado de Autovalores e Autovetores

Definição 7.3.1 Problema Generalizado de Autovalores e Autovetores
Um problema generalizado de autovalores e autovetores tem a forma

$$\mathbf{Ax} = \lambda \mathbf{Bx} \tag{7.3}$$

tal que a forma padrão é obtida quando $\mathbf{B} = \mathbf{I}$. Os valores de λ e $\mathbf{x}$ que satisfazem esta equação são obtidos pelo mesmo procedimento desenvolvido no começo deste capítulo, pois

$$(\mathbf{A} - \lambda \mathbf{B})\,\mathbf{x} = \mathbf{0}$$

leva a condição

$$\det(\mathbf{A} - \lambda\mathbf{B}) = 0.$$

Se as matrizes $\mathbf{A}$ e $\mathbf{B}$ forem auto-adjuntas (que no caso de matrizes reais implica em simetria) e se $\mathbf{B}$ for positivo definida, então as colunas de $\mathbf{\Phi}$ (autovetores) podem ser escritas como um conjunto ortogonal.

Um caso especial ocorre quando a matriz $\mathbf{B}$ é inversível, pois podemos multiplicar a Eq. 7.3 por $\mathbf{B}^{-1}$, obtendo

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{x} = \lambda\mathbf{B}^{-1}\mathbf{B}\mathbf{x}$$

ou

$$\mathbf{C}\mathbf{x} = \lambda\mathbf{I}\mathbf{x} = \lambda\mathbf{x},$$

com $\mathbf{C} = \mathbf{B}^{-1}\mathbf{A}$. A expressão obtida tem a mesma forma que o problema de autovalores e autovetores tradicional estudado ao longo deste capítulo, tal que os autovalores e autovetores de $\mathbf{C}$ satisfazem a equação generalizada.

■ **Exemplo 7.10** Seja o problema generalizado de autovalores e autovetores definido por

$$\mathbf{A} = \begin{bmatrix} 10 & 5 \\ 6 & 15 \end{bmatrix}$$

e

$$\mathbf{B} = \begin{bmatrix} 4 & 1 \\ 2 & 6 \end{bmatrix}.$$

Os autovalores deste problema podem ser obtidos por

$$\det\left(\begin{bmatrix} 10 & 5 \\ 6 & 15 \end{bmatrix} - \lambda\begin{bmatrix} 4 & 1 \\ 2 & 6 \end{bmatrix}\right) = 0$$

resultando no polinômio característico $p(\lambda) = 22\lambda^2 - 104\lambda + 120$ com raízes (autovalores) $\lambda_1 = 2,7272$ e $\lambda_2 = 2$. Os autovetores correspondentes podem ser obtidos pela solução dos problemas homogêneos

$$(\mathbf{A} - \lambda_1\mathbf{B})\,\mathbf{u} = \mathbf{0}$$

e

$$(\mathbf{A} - \lambda_2\mathbf{B})\,\mathbf{v} = \mathbf{0},$$

resultando em

$$\mathbf{\Phi} = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2' \end{bmatrix} = \begin{bmatrix} 2,5 & -1,5 \\ 1 & 1 \end{bmatrix}.$$

Como, neste exemplo, a matriz $\mathbf{B}$ é inversível

$$\mathbf{B}^{-1} = \begin{bmatrix} \frac{3}{11} & -\frac{1}{22} \\ -\frac{1}{11} & \frac{2}{11} \end{bmatrix}$$

podemos obter a matriz $\mathbf{C} = \mathbf{B}^{-1}\mathbf{A}$,

$$\mathbf{C} = \begin{bmatrix} \frac{27}{11} & \frac{15}{22} \\ \frac{2}{11} & \frac{25}{11} \end{bmatrix}$$

tal que o problema de autovalores e autovetores na forma tradicional,

$$\det(\mathbf{C} - \lambda \mathbf{I})\,\mathbf{x} = \mathbf{0}$$

resulta no polinômio característico

$$p(\lambda) = (2.2727 - 1.0\lambda)\,(2.4545 - 1.0\lambda) - 0.1239$$

com raízes $\lambda_1 = 2,7272$ e $\lambda_2 = 2$, com os mesmos autovetores obtidos anteriormente. Neste caso, como esperado, observamos que

$$\mathbf{\Phi}^{-1}\mathbf{C}\mathbf{\Phi} = \begin{bmatrix} 2,72 & 0 \\ 0 & 2 \end{bmatrix}.$$

■

Exercício 7.3 Verificar se o vetor $(-2,1)$ é autovetor do o operador

$$\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$$

resp. Sim ■

Exercício 7.4 Verificar se o vetor $(-2,1,3)$ é autovetor do o operador

$$\begin{bmatrix} 1 & -1 & 0 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

resp. Não ■

Exercício 7.5 Os vetores $(1,1)$ e $(2,-1)$ são autovetores de um operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ associados aos autovalores 5 e -1, respectivamente. Determinar $T(4,1)$.

Resp. $T(x,y) = (x+4y, 2x+3y)$ e $T(4,1) = (8,11)$.

■

Exercício 7.6 Determinar o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ cujos autovalores são 1 e 3, associados aos autoespaços $V_1 = \{(-y,y), y \in \mathbb{R}\}$ e $V_3 = \{(0,y), y \in \mathbb{R}\}$.

Resp. $T(x,y) = (x, 2x+3y)$

■

Exercício 7.7 Determinar os autovalores e autovetores dos seguintes operadores lineares

$$\begin{aligned} T(x,y) &= (x+2y, -x+4y) \qquad (7.4) \\ T(x,y) &= (y, -x) \end{aligned}$$

Resp. a)2 e 3 , b) não possui autovalores reais. Investigue o motivo e observe a forma matricial deste operador. ■

Exercício 7.8 Determinar um conjunto de vetores base que diagonalize o operador

$$T(x,y) = (-3x-5y, 2y)$$

Resp. $\{(1,-1),(1,0)\}$

Exercício 7.9 Verificar se existe uma base de autovetores para

$$T(x,y,z) = (x+y+z, 2y+z, 2y+3z) \tag{7.5}$$
$$T(x,y,z) = (x, -2x+3y-z, -4y+3z)$$

Resp. a) Sim e b) Não.

Exercício 7.10 Determine os autovalores e autovetores do operador $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x,y,z) = (4x+z, -2x+y, -2x+z)$. O operador é diagonalizável ?

Resp: É diagonalizável.

8. Produto Interno

O conceito de norma permite estabelecer o "tamanho" de um elemento do espaço e, ainda, definir a distância entre dois elementos quaisquer. O conceito de produto interno permite, por sua vez, obter uma medida de "ângulo" entre 2 elementos do espaço.

Embora os axiomas que definem o produto interno já tenham sido apresentados no capítulo sobre o $\mathbb{R}^n$, iremos apresentar novamente para facilitar a discussão

Definição 8.0.1 Produto interno

Seja U um espaço vetorial linear. A forma bilinear $U \times U \to \mathbb{R}$ que associa a cada dois elementos do espaço um escalar é chamada de produto interno e é denotada por $\mathbf{u}_1 \cdot \mathbf{u}_2$, $(\mathbf{u}_1, \mathbf{u}_2)$ ou $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle$. O produto interno deve seguir os seguintes axiomas:

- Simétrico: $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}_2, \mathbf{u}_1 \rangle \ \forall \mathbf{u}_1, \mathbf{u}_2 \in U$
- Homogêneo: $\langle \alpha \mathbf{u}_1, \mathbf{u}_2 \rangle = \alpha \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \ \forall \mathbf{u}_1, \mathbf{u}_2 \in U$ e $\alpha \in \mathbb{R}$
- Aditivo: $\langle \mathbf{u}_1 + \mathbf{u}_2, \mathbf{u}_3 \rangle = \langle \mathbf{u}_1, \mathbf{u}_3 \rangle + \langle \mathbf{u}_2, \mathbf{u}_3 \rangle \ \forall \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \in U$
- Positivo-definido: $\langle \mathbf{u}, \mathbf{u} \rangle > 0$ e $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ sse $\mathbf{u} = \mathbf{0}$

Iremos utilizar a notação $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle$ para denotar o produto interno neste capítulo.

Quando um espaço vetorial tem produto interno definido, chamamos este espaço de Espaço Pré-Hilbert e, portanto, podemos interpretar estes espaços de funções como generalizações do $\mathbb{R}^n$.

Definição 8.0.2 Espaço de (pré)-Hilbert

Um espaço vetorial H é um espaço vetorial equipado com produto interno e cuja norma

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$

é induzida por este produto interno. Desta forma, a distância entre dois elementos de um espaço de pré-Hilbert é dada por

$$d(\mathbf{u}_1, \mathbf{u}_2) = \|\mathbf{u}_1 - \mathbf{u}_2\| = \sqrt{\langle \mathbf{u}_1 - \mathbf{u}_2, \mathbf{u}_1 - \mathbf{u}_2 \rangle}$$

Desta forma, podemos estender o conceito de produto interno canônico, visto no capítulo sobre o $\mathbb{R}^n$ para o espaço de funções

Definição 8.0.3 Produto interno (canônico) no espaço das funções
Sejam dois elementos $\mathbf{u}_1$ e $\mathbf{u}_2$ de um espaço L^2 de funções quadrado integráveis, isto é

$$\int_\Omega |\mathbf{u}|^2 d\Omega \leq \infty.$$

Definimos o produto interno canônico como

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \int_\Omega \mathbf{u}_1 \mathbf{u}_2 d\Omega.$$

Exercício 8.1 Verifique que a definição acima atende aos axiomas de produto interno. ■

■ **Exemplo 8.1** Sejam as funções $f = sin(\pi x) sin(\pi y)$ e $g = x$ em $L^2[0,1] \times [0,1]$. O produto interno é dado por

$$\langle f, g \rangle = \int_0^1 \int_0^1 sin(\pi x) sin(\pi y) x dx dy = \frac{2}{\pi^2}.$$

■

8.1 Ortogonalidade

De posse da definição de produto interno, podemos avaliar o ângulo entre dois elementos do espaço. Assim, da mesma forma do que no $\mathbb{R}^n$, podemos avaliar se dois elementos são ortogonais:

Definição 8.1.1 Ortogonalidade
Dois elementos $\mathbf{u}_1$ e $\mathbf{u}_2$ de um espaço vetorial com produto interno são ditos ortogonais se

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = 0$$

■ **Exemplo 8.2** Sejam os vetores $\mathbf{u}_1 = \sin(\pi x)$ e $\mathbf{u}_2 = \cos(\pi x)$ em $x = [-1, 1]$. Esses vetores (funções) são ortogonais, pois

$$\int_{-1}^1 \sin(\pi x) \cos(\pi x) dx = 0.$$

■

De posse da definição de ortogonalidade entre dois elementos do espaço, podemos então definir um conjunto ortogonal:

Definição 8.1.2 Conjunto Ortogonal
O conjunto de elementos não nulos $O = \{\mathbf{u}_i\}, i = 1..n$ é dito ortogonal se cada par i, j for ortogonal, isto é

$$< \mathbf{u}_i, \mathbf{u}_j >= 0, \forall \mathbf{u}_i, \mathbf{u}_j \in O \, e \, i \neq j.$$

■ **Exemplo 8.3** Seja o conjunto $O = \left\{1, x, x^2 - \frac{1}{3}\right\}$ em $x = [-1, 1]$. Esse conjunto é ortogonal pois

$$\int_{-1}^1 1 x dx = 0,$$

$$\int_{-1}^1 1 \left(x^2 - \frac{1}{3}\right) dx = 0,$$

e

$$\int_{-1}^1 x \left(x^2 - \frac{1}{3}\right) dx = 0,$$

■

e, com a definição de conjunto ortogonal, podemos definir um conjunto ortonormal:

Definição 8.1.3 Conjunto Ortonormal
Um conjunto de elementos não nulos $B = \{\mathbf{u}_i\}, i = 1..n$ é dito ortonormal se cada par i, j for ortogonal, isto é

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}, \forall \mathbf{u}_i, \mathbf{u}_j \in B$$

■ **Exemplo 8.4** Seja o conjunto $O = \{\sin(\pi x), \sin(2\pi x), \sin(3\pi x)\}$ em $x = [-1, 1]$. Esse conjunto é ortogonal pois

$$\int_{-1}^{1} \sin(\pi x)\sin(2\pi x)dx = 0,$$

$$\int_{-1}^{1} \sin(\pi x)\sin(3\pi x)dx = 0,$$

e

$$\int_{-1}^{1} \sin(2\pi x)\sin(3\pi x)dx = 0.$$

Além disso,

$$\int_{-1}^{1} \sin(\pi x)\sin(\pi x)dx = 1,$$

$$\int_{-1}^{1} \sin(2\pi x)\sin(2\pi x)dx = 1,$$

e

$$\int_{-1}^{1} \sin(3\pi x)\sin(3\pi x)dx = 1,$$

tal que o conjunto é ortonormal.

■

Outro resultado bem conhecido e que é naturalmente estensível para o espaço das funções é o teorema de Pitágoras:

Definição 8.1.4 Teorema de Pitágoras
Se dois elementos $\mathbf{u}_1$ e $\mathbf{u}_2$ forem ortogonais, então

$$\|\mathbf{u}_1 + \mathbf{u}_2\|^2 = \langle \mathbf{u}_1 + \mathbf{u}_2, \mathbf{u}_1 + \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, \mathbf{u}_1 \rangle + 2\underbrace{\langle \mathbf{u}_1, \mathbf{u}_2 \rangle}_{0} + \langle \mathbf{u}_2, \mathbf{u}_2 \rangle = \|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2$$

■ **Exemplo 8.5** Seja o conjunto $B = \{cos(nx)\}_{n=1}^{\infty} \in L^2[-\pi, \pi]$. Esse conjunto é ortogonal, pois

$$\int_{-\pi}^{\pi} cos(nx)cos(mx)\,dx = 0, \forall m, n \in \mathbb{N}^*, m \neq n$$

e podemos verificar o teorema de Pitágoras para dois elementos quaisquer, pois

$$\|\mathbf{v}_n + \mathbf{v}_m\|^2 = \int_{-\pi}^{\pi} [cos(nx) + cos(mx)]\,[cos(nx) + cos(mx)]\,dx = 2\pi$$

com

$$\|\mathbf{v}_n\|^2 = \int_{-\pi}^{\pi} [cos(nx)]\,[cos(nx)]\,dx = \pi$$

e

$$\|\mathbf{v}_m\|^2 = \int_{-\pi}^{\pi} [cos(mx)]\,[cos(mx)]\;dx = \pi.$$

■

Além de definirmos conjuntos contendo elementos ortogonais entre si, podemos definir conjuntos que são ortogonais entre si:

Definição 8.1.5 Complemento Ortogonal

Seja V um espaço com produto interno e M subconjunto de V. O complemento ortogonal de M, denotado por $M^{\perp}$, é definido como

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

sendo também um subconjunto de V.

■ **Exemplo 8.6** Seja $V \subseteq \mathbb{R}^3$ e $M = \{\mathbf{v}\}$ não nulo. O complemento ortogonal de M é composto por todos os vetores que são ortogonais a $\mathbf{v}$ e, portanto, forma um plano que passa pela origem e que é ortogonal a $\mathbf{v}$. ■

■ **Exemplo 8.7** Seja $V \subseteq \mathbb{R}^3$ e $M = \{\mathbf{v}_1, \mathbf{v}_2\}$ não nulos e não paralelos. O complemento ortogonal de M é composto por todos os vetores que são ortogonais ao mesmo tempo a $\mathbf{v}_1$ e a ,$\mathbf{v}_2$. Desta forma, $M^{\perp}$ é formado pela interseção dos complementos ortogonais de $\mathbf{v}_1$ e $\mathbf{v}_2$. ■

■ **Exemplo 8.8** Seja $S \subseteq \mathbb{R}^3$ um sub-espaço gerado por $\{(1,0,1),(0,2,3)\}$. Um elemento típico de S pode ser representado por

$$\mathbf{s} = \alpha(1,0,1) + \beta(0,2,3) = (s_1, s_2, s_1 + \frac{3}{2}s_2)$$

pois

$$\begin{aligned} s_1 &= \alpha \\ s_2 &= 2\beta \\ s_3 &= \alpha + 3\beta \end{aligned}$$

Assim, o complemento ortogonal de S pode ser definido por um vetor genérico (x_1, x_2, x_3) tal que

$$\langle \mathbf{x}, \mathbf{s} \rangle = 0 = x_1 s_1 + x_2 s_2 + x_3 \left(s_1 + \frac{3}{2}s_2\right)$$

ou

$$0 = (x_1 + x_3)\,s_1 + \left(x_2 + \frac{3}{2}x_3\right) s_2$$

tal que, agrupando por temos de s_i (pois o resultado tem que ser independente do $\mathbf{s}$ escolhido)

$$\left\{ \begin{array}{rcl} x_1 + x_3 & = & 0 \\ x_2 + \frac{3}{2}x_3 & = & 0 \end{array} \right.$$

permite verificar que

$$S^{\perp} = \left\{ \mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \left(-x_3, -\frac{3}{2}x_3, x_3\right) \right\}$$

■

Por fim, definimos um dos teoremas mais importantes na álgebra:

Teorema 8.1.1 Um conjunto de vetores ortogonais é linearmente independente

Seja $O = \{\mathbf{u}_i\}$ um conjunto ortogonal com $|O| = n$. Se os elementos deste conjunto forem linearmente independentes, então

$$\sum_{i=1}^{n} \alpha_i \mathbf{u}_i = \mathbf{0} \implies \alpha_i = 0\, \forall i \in [1,n].$$

Da propriedade de ortogonalidade, sabemos que

$$0 = \left\langle \sum_{i=1}^{n} \alpha_i \mathbf{u}_i, \mathbf{u}_j \right\rangle = \alpha_1 \langle \mathbf{u}_1, \mathbf{u}_j \rangle + \alpha_2 \langle \mathbf{u}_2, \mathbf{u}_j \rangle + \ldots + \alpha_n \langle \mathbf{u}_n, \mathbf{u}_j \rangle$$

com $j \in [1,n]$. Como assumimos que O é um conjunto ortogonal, verificamos que

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \beta_{ij} \delta_{ij}$$

em que β_{ij} é um escalar qualquer. Assim,

$$0 = \left\langle \sum_{i=1}^{n} \alpha_i \mathbf{u}_i, \mathbf{u}_j \right\rangle = \alpha_i \beta_{ij} \delta_{ij} \implies \alpha_j = 0$$

para $j \in [1,n]$. Podemos notar que o mesmo teorema se aplica para um conjunto ortonormal, com a diferença de que $\beta_{ij} = 1$.

■ **Exemplo 8.9** Já provamos que o conjunto $O = \left\{1, x, x^2 - \frac{1}{3}\right\}$ em $x = [-1,1]$ é ortogonal. Para verificarmos se é linearmente independente podemos construir dois elementos com esse conjunto

$$\mathbf{v}_1 = \alpha_1 + \alpha_2 x + \alpha_3 \left(x^2 - \frac{1}{3}\right)$$

e

$$\mathbf{v}_2 = \beta_1 + \beta_2 x + \beta_3 \left(x^2 - \frac{1}{3}\right).$$

Uma maneira de provar a independência linear é verificando que $\mathbf{v}_1 = \mathbf{v}_2$ se e somente se $\alpha_i = \beta_i = 0$, $\forall i$. De fato

$$\mathbf{v}_1 - \mathbf{v}_2 = (\alpha_1 - \beta_1) + (\alpha_2 - \beta_2)x + (\alpha_3 - \beta_3)\left(x^2 - \frac{1}{3}\right) = \mathbf{0}$$

implica em $\alpha_i = \beta_i = 0$, tal que o conjunto é independente. ■

Observe que o teorema acima não diz que um conjunto linearmente independente é ortogonal, mas sim o contrário. No entanto, se tivermos um conjunto linearmente independente, podemos operar sobre o conjunto para que o mesmo se torne ortogonal. Para isto, utilizamos o procedimento de ortogonalização de Gram-Schmidt, que é baseado nos conceitos que veremos a seguir.

8.2 Projeção

Definição 8.2.1 Operador de Projeção

Um operador de projeção P é uma transformação linear de um espaço em si mesmo tal que $P(P) = P$.

■ **Exemplo 8.10** Considere o operador P tal que

$$P(x,y) = (2y,y).$$

Este operador mapeia pontos do $\mathbb{R}^2$ no próprio $\mathbb{R}^2$ e é linear. Ainda, da propriedade de que $P(P) = P$, temos que

$$P(P(x,y)) = P(x,y) = (2y,y)$$

ou, na forma matricial, podemos escrever que

$$\mathbf{P} = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}$$

tal que

$$P(\mathbf{P}) = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} = \mathbf{P}.$$

■

■ **Exemplo 8.11** Considere o operador P tal que

$$P(x,y,z) = (x,y,0).$$

Este operador é uma projeção do espaço tridimensional no plano xy. Podemos verificar que mapeia pontos do $\mathbb{R}^3$ no próprio $\mathbb{R}^3$ e é linear. Ainda, da propriedade de que $P(P) = P$, temos que

$$P(P(x,y,z)) = P(x,y,0) = (x,y,0)$$

ou, na forma matricial, podemos escrever que

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

tal que

$$P(\mathbf{P}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{P}.$$

■

Um caso especial de projeção é a chamada **Projeção Ortogonal**, de fundamental importância para o desenvolvimento do conteúdo deste texto e, também, base teórica para vários métodos de aproximação que serão vistos mais adiante.

Definição 8.2.2 Projeção Ortogonal
Um operador de projeção P em um espaço U com produto interno é dito ortogonal se seu alcance (*range*) é ortogonal ao seu espaço nulo, i.e.,

$$\mathscr{R}(P) \perp \aleph(P)$$

ou seja, se $\mathbf{u} \in \mathscr{R}(P)$ e $\mathbf{v} \in \aleph(P)$, então $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. De fato, sabemos que $P(\mathbf{u}) \in \mathscr{R}(P)$ e $\mathbf{u} - P(\mathbf{u}) \in \aleph(P)$, tal que

$$\langle P(\mathbf{u}), \mathbf{u} - P(\mathbf{u}) \rangle = 0.$$

Se o espaço for finito dimensional, representamos o operador por uma matriz **P**, tal que

$$\langle \mathbf{Pu}, \mathbf{u} - \mathbf{Pu} \rangle = (\mathbf{Pu})^T (\mathbf{u} - \mathbf{Pu}) = \mathbf{u}^T \mathbf{P}^T \mathbf{u} - \mathbf{u}^T \mathbf{P}^T \mathbf{Pu}$$

que será zero quando $\mathbf{P} = \mathbf{P}^T$ (pois neste caso $\mathbf{P}^T\mathbf{P} = \mathbf{P}^T$), ou seja, quando a matriz do operador for simétrica (ou hermitiana se for complexa) o que também equivale ao fato de o operador ser auto-adjunto.

■ **Exemplo 8.12** A projeção do exemplo anterior é uma projeção ortogonal, pois

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

■

■ **Exemplo 8.13** A projeção

$$\mathbf{P} = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}$$

não é ortogonal, pois

$$\begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 \\ 2 & 1 \end{bmatrix} \neq \mathbf{P}.$$

■

Com as definições acima, podemos estabelecer o importante Teorema da Projeção:

Teorema 8.2.1 Teorema da Projeção
Seja V um espaço com produto interno e P uma projeção ortogonal em V. Então:

- cada elemento $\mathbf{w} \in V$ pode ser escrito exclusivamente como $\mathbf{w} = \mathbf{u} + \mathbf{v}$, onde $\mathbf{u} \in \mathscr{R}(P)$ e $\mathbf{v} \in \aleph(P)$;
- $\|\mathbf{w}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$;
- $\mathscr{R}(P)$ e $\aleph(P)$ são subespaços lineares fechados de V;
- $\aleph(P) = \mathscr{R}(P)^\perp$ e $\mathscr{R}(P) = \aleph(P)^\perp$

■ **Exemplo 8.14** Ainda considerando o operador P dos exemplos anteriores, sabemos que para este operador $\aleph(P) = (0,0,a)$ e $\mathscr{R}(\mathbf{P}) = (u_1, u_2, 0)$. Assim,

$$\langle (0,0,a), (u_1,u_2,0) \rangle = 0.$$

Com isso, podemos construir qualquer elemento de $V \in \mathbb{R}^3$, pois

$$\mathbf{w} = (u_1,u_2,0) + (0,0,a) = (u_1,u_2,a).$$

■

■ **Exemplo 8.15** Vamos considerar a projeção ortogonal de um vetor $\mathbf{u} \in \mathbb{R}^3$ sobre uma direção indicada por um vetor $\mathbf{a} \in \mathbb{R}^3$. Neste caso, a projeção nada mais é do que o comprimento de $\mathbf{x}$ sobre $\mathbf{a}$, que iremos chamar de c. A forma geral seria

$$\mathbf{Pu} = c\mathbf{a} \tag{8.1}$$

e, sabendo que $\mathbf{u} - \mathbf{Pu}$ deve ser perpendicular a $\mathbf{a}$, podemos afirmar que

$$0 = \langle \mathbf{a}, \mathbf{u} - \mathbf{Pu} \rangle = \mathbf{a}^T\mathbf{u} - \mathbf{a}^T\mathbf{Pu}$$

e, da Eq. (8.1), podemos escrever

$$\mathbf{a}^T\mathbf{u} - \mathbf{a}^T c\mathbf{a} = 0.$$

Isolando o comprimento da projeção, obtemos

$$c = \frac{\mathbf{a}^T\mathbf{u}}{\mathbf{a}^T\mathbf{a}}.$$

Assim, se $\mathbf{u} = (0,6)$ e se considerarmos uma linha com direção dada por $\mathbf{a} = (1,1)$, teremos

$$c = \frac{1\cdot 0 + 1\cdot 6}{1\cdot 1 + 1\cdot 1} = 3$$

tal que

$$\mathbf{Pu} = (3,3)$$

e, conforme já afirmamos,

$$\mathbf{u} - \mathbf{Pu} = (6,0) - (3,3) = (3,-3)$$

é ortogonal a $\mathbf{a}$, pois $\langle (1,1),(3-3)\rangle = 3-3 = 0$.

■

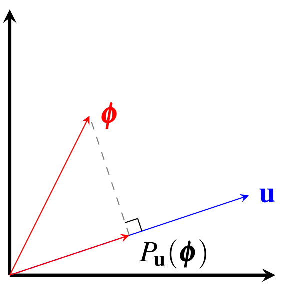
A operação realizada no exemplo anterior pode ser formalizada para qualquer espaço com produto interno. Esse resultado é muito importante e será muito utilizado no resto deste texto.

Definição 8.2.3 Operador de Projeção Ortogonal sobre um elemento $\mathbf{u}$

O operador

$$P_{\mathbf{u}}(\boldsymbol{\phi}) = \frac{\langle \boldsymbol{\phi}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}$$

projeta o vetor $\boldsymbol{\phi}$ ortogonalmente em $\mathbf{u}$. No $\mathbb{R}^2$ podemos representar essa operação graficamente como



Da álgebra sabemos que

$$\langle \mathbf{u}, \boldsymbol{\phi}\rangle = \|\mathbf{u}\|\,\|\boldsymbol{\phi}\|\cos(\theta)$$

tal que

$$\|\boldsymbol{\phi}\|\cos(\theta) = \frac{\langle \mathbf{u}, \boldsymbol{\phi}\rangle}{\|\mathbf{u}\|}$$

é (um escalar com) o tamanho da projeção de $\boldsymbol{\phi}$ sobre $\mathbf{u}$. A direção de $\mathbf{u}$ pode ser representada por um vetor unitário com a mesma direção de $\mathbf{u}$

$$\mathbf{e} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$$

tal que

$$P_{\mathbf{u}}(\boldsymbol{\phi}) = \frac{\langle \mathbf{u}, \boldsymbol{\phi} \rangle}{\|\mathbf{u}\|}\mathbf{e} = \frac{\langle \mathbf{u}, \boldsymbol{\phi} \rangle}{\|\mathbf{u}\|}\frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{\langle \mathbf{u}, \boldsymbol{\phi} \rangle}{\|\mathbf{u}\|^2}\mathbf{u} = \frac{\langle \boldsymbol{\phi}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}.$$

8.3 Ortogonalização de Gramm-Schmidt

A obtenção de um conjunto ortogonal é necessária para o desenvolvimento de diversos métodos utilizados na mecânica computacional. Desta forma, o procedimento de Gramm-Schmidt é de fundamental importância.

Definição 8.3.1 Procedimento de Ortogonalização de Gram-Schmidt

Seja um espaço com produto interno V, com dimensão n. Dado um conjunto linearmente independente $\Phi = \{\boldsymbol{\phi}_i\} \subseteq V$. Podemos aplicar uma transformação que gera um novo conjunto ortogonal $U = \{\mathbf{u}_k\} \subseteq V$ com

$$\mathbf{u}_i = \boldsymbol{\phi}_i - \sum_{k=1}^{i-1} \frac{\langle \boldsymbol{\phi}_i, \mathbf{u}_k \rangle}{\langle \mathbf{u}_k, \mathbf{u}_k \rangle}\mathbf{u}_k \quad , i = 1..n.$$

■ **Exemplo 8.16** Sejam os vetores $\mathbf{a}_1 = \begin{pmatrix}1 & -1 & -1\end{pmatrix}$, $\mathbf{a}_2 = \begin{pmatrix}-1 & 2 & 3\end{pmatrix}$ e $\mathbf{a}_3 = \begin{pmatrix}1 & 1 & 1\end{pmatrix}$, que são linearmente independentes, mas não são ortogonais.

Usando o vetor $\mathbf{a}_1$ como referência, $\mathbf{u}_1 = \mathbf{a}_1$, projetamos $\mathbf{a}_2$ sobre $\mathbf{u}_1$

$$P_{\mathbf{u}_1}(\mathbf{a}_2) = \frac{\langle \mathbf{a}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 = \begin{pmatrix}-2 & 2 & 2\end{pmatrix}$$

e geramos um novo vetor que é ortogonal a $\mathbf{a}_1$

$$\mathbf{u}_2 = \mathbf{a}_2 - P_{\mathbf{u}_1}(\mathbf{a}_2) = \begin{pmatrix}1 & 0 & 1\end{pmatrix}.$$

De fato, $\langle \mathbf{u}_2, \mathbf{u}_1 \rangle = 1 + 0 - 1 = 0$.

O vetor $\mathbf{u}_3$, ortogonal a $\mathbf{u}_2$ e a $\mathbf{u}_1$, é obtido com duas projeções

$$P_{\mathbf{u}_2}(\mathbf{a}_3) = \begin{pmatrix}1 & 0 & 1\end{pmatrix}$$

e

$$P_{\mathbf{u}_1}(\mathbf{a}_3) = \begin{pmatrix}-1/3 & 1/3 & 1/3\end{pmatrix}$$

que, somados, resultam em um vetor

$$\mathbf{P}_3 = \begin{pmatrix}2/3 & 1/3 & 4/3\end{pmatrix}$$

Com isso, geramos o terceiro vetor

$$\mathbf{u}_3 = \mathbf{a}_3 - \mathbf{P}_3 = \begin{pmatrix}1/3 & 2/3 & -1/3\end{pmatrix}$$

que é ortogonal a $\mathbf{u}_1$ e a $\mathbf{u}_2$. ■

■ **Exemplo 8.17** Seja o conjunto $\Phi = \{1, x, x^2, x^3\}$ definido no $L^2[-1,1]$. Este conjunto é sabidamente linearmente independente, mas não é ortogonal (verifique). No entanto, aplicando o procedimento de ortogonalização temos

$$\mathbf{u}_1 = 1 \tag{8.2}$$

$$\mathbf{u}_2 = x - \left[\frac{\langle x, 1\rangle}{\langle 1, 1\rangle} 1\right] = x \tag{8.3}$$

$$\mathbf{u}_3 = x^2 - \left[\frac{\langle x^2, 1\rangle}{\langle 1, 1\rangle} 1 + \frac{\langle x^2, x\rangle}{\langle x, x\rangle} x\right] = x^2 - \frac{1}{3} \tag{8.4}$$

$$\begin{aligned}\mathbf{u}_4 &= \{x^3\} - \left[\frac{\langle x^3, 1\rangle}{\langle 1, 1\rangle} 1 + \frac{\langle x^3, x\rangle}{\langle x, x\rangle} x + \frac{\langle x^3, x^2 - \frac{1}{3}\rangle}{\langle x^2 - \frac{1}{3}, x^2 - \frac{1}{3}\rangle}\left(x^2 - \frac{1}{3}\right)\right] \\ &= x^3 - \frac{3x}{5}\end{aligned} \tag{8.5}$$

tal que o conjunto ortogonal (e, portanto, continua linearmente independente) resultante é

$$U = \left\{1, x, x^2 - \frac{1}{3}, x^3 - \frac{3x}{5}\right\}.$$

Podemos verificar a ortogonalidade, por exemplo, com

$$\left\langle \left\{x^2 - \frac{1}{3}\right\}, \left\{x^3 - \frac{1}{15}x\right\} \right\rangle = \int_{-1}^{1} \left\{x^2 - \frac{1}{3}\right\} \left\{x^3 - \frac{x}{15}\right\} dx = 0.$$

Verifique os demais valores.

■

É importante salientar que embora tenhamos demostrado o conceito de projeção usando o $\mathbb{R}^n$, aplicamos a ortogonalização a uma base polinomial sem problemas. De fato, esses conceitos se aplicam a espaços de funções de maneira absolutamente geral (desde que o espaço seja dotado de produto interno).

9. Sequências

Definição 9.0.1 Sequência

Uma sequência em um espaço U nada mais é do que uma função de $\mathbb{N}$ em U, na forma $X : \mathbb{N} \to U$, com notação $(\mathbf{x}_1, \mathbf{x}_2, ...)$ ou, de forma mais compacta $(\mathbf{x}_k)$.

■ **Exemplo 9.1** Exemplos de sequências:

1. $S = (1, 1, x_{k-1} + x_{k-2}) = (1, 1, 2, 3, 5, 8, 13, ...)$;
2. $S = \left(\frac{1}{n}\right) = \left(\frac{1}{1}, \frac{1}{2}, ...\right)$;
3. $S = (\sin(n\pi x)) = (\sin(\pi x), \sin(2\pi x), ...)$

■

Sequências são utilizadas para descrever o comportamento de funções a medida que o contador varia (conceito de limite).

Definição 9.0.2 Limite de uma sequência

Seja uma sequência $S = (\mathbf{x}_k)$ em U. Dizemos que $\mathbf{x} \in U$ é limite de S se, para todo $\varepsilon > 0$, existe $K \in \mathbb{N}$ tal que $\|\mathbf{x} - \mathbf{x}_K\| < \varepsilon$ para todo $k > K$. A notação para este comportamento é $\mathbf{x}_k \to \mathbf{x}$ ou,

$$\mathbf{x} = \lim_{k \to \infty} \mathbf{x}_k$$

Definição 9.0.3 Sequências Convergentes e Sequências Divergentes

Se uma sequência $S = (\mathbf{x}_k)$ possui limite, então é dita convergente. Do contrário, é dita divergente. Uma sequência S pode ter no máximo um limite (unicidade do limite).

■ **Exemplo 9.2** Seja uma sequência definida por $S = \left(\frac{1}{n}\right)$ (ou seja, $1/1, 1/2, 1/3....$). Vamos mostrar que a sequência é convergente e tem limite em 0. Para isto,

$$\left|\frac{1}{n} - 0\right| < \varepsilon$$

e vamos selecionar um $n > M = \frac{1}{\varepsilon}$, que chamaremos de ε^+, para indicarmos que será um pouco maior do que ε. Neste caso, a sua inversa será ε^-, que será um pouco menor do que ε, tal que

$|\varepsilon^-| < \varepsilon$. Assim, a sequência é convergente com limite em 0. ■

■ **Exemplo 9.3** A série de Fibonacci, $S = (1, 1, 2, 3, 5, 8, 13, ...)$ é divergente, pois tende a infinito (que não é um limite definido).

■

Definição 9.0.4 Sequência Limitada
Seja uma sequência S. Esta é dita limitada se existe um número real M tal que

$$\|\mathbf{x}_k\| \leq M, \forall k \in \mathbb{N}.$$

É importante ressaltar que toda a sequência convergente é limitada.

■ **Exemplo 9.4** A sequência $S = \left(\frac{1}{n}\right)$ é convergente e, portanto, limitada. Podemos ver isso pois $||0|| \leqslant M$, onde M é qualquer real maior ou igual a zero. ■

■ **Exemplo 9.5** A série de Fibonacci não é convergente e nem limitada, pois seus termos tendem a infinito. ■

9.1 Sequências de Cauchy

Dada um sequência $\{x\}_k$, dizemos que a mesma é de *Cauchy* se os elementos desta sequência forem arbitrariamente próximos a medida que k aumenta (a medida que a sequência se desenvolve).

Definição 9.1.1 Sequência de Cauchy
Um sequência $\{x\}_k$ é dita de Cauchy se existem $\varepsilon > 0$ e $m, n > N$ tal que:

$$|x_n - x_m| < \varepsilon.$$

Observe que o valor do limite não é necessário para estabelecermos que uma sequência é de Cauchy.

O interessante sobre a sequência de Cauchy vem do fato de que toda a sequência convergente é de Cauchy. Assim, se provarmos que uma sequência tem ao menos uma sub sequência de Cauchy, então fica configurada a prova de que a mesma é convergente. Geralmente esta é a abordagem utilizada na maioria das provas. Outra característica interessante sobre sequências de Cauchy vem do fato de que utilizamos esta descrição para analisarmos a convergência de algoritmos.

■ **Exemplo 9.6** Dada a sequência

$$\{x_n\} = \frac{3n^2 - 2}{4n^3 + n^2 + 3}$$

é fácil verificar que o limite para $n \to \infty$ é zero. No entanto, a definição da sequência de Cauchy não necessita do cálculo do limite. Assim, utilizando a desigualdade triangular

$$\left| \frac{3m^2 - 2}{4m^3 + m^2 + 3} - \frac{3n^2 - 2}{4n^3 + n^2 + 3} \right| \leqslant \left| \frac{3m^2 - 2}{4m^3 + m^2 + 3} \right| + \left| \frac{3n^2 - 2}{4n^3 + n^2 + 3} \right| \leqslant \varepsilon$$

e, para m e n suficientemente grandes, observamos que

$$\frac{3m^2}{4m^3} + \frac{3n^2}{4n^3} = \frac{3}{4}\left(\frac{1}{m} + \frac{1}{n}\right) < \left(\frac{1}{m} + \frac{1}{n}\right) \leqslant \varepsilon$$

pois $\frac{3}{4} < 1$. Assim, se $m = n = N$, observamos que

$$\frac{1}{N} + \frac{1}{N} = \frac{2}{N} \leqslant \varepsilon.$$

■

■ **Exemplo 9.7** Dada a sequência

$$\{x_n\} = 1 + (-1)^n = 0, 2, 0, 2, 0, 2, \ldots$$

é fácil verificar que a mesma não é convergente. Utilizando a desigualdade triangular

$$|(1 + (-1)^m) - (1 + (-1)^n)| \leqslant |1 + (-1)^m| + |1 + (-1)^n| \leqslant \varepsilon$$

e, assumindo $m = n = N$,

$$2 + 2(-1)^N \leqslant \varepsilon.$$

Assim, isolando N em função de ε, obtemos

$$N = \frac{\log\left(\frac{\varepsilon}{2} - 1\right)}{\log(-1)}$$

que não é real, tal que a sequência não é de Cauchy.

■

9.2 Convergência de sequência de funções

De especial importância é o conceito de convergência de uma sequência de funções. No entanto, não existe apenas uma defināo de convergência, o que requer um cuidado especial.

Definição 9.2.1 Convergência forte de uma sequência de funções
Seja uma sequência de funções $f_n(x)$ com domínio U. Dizemos que a sequência de funções converge fortemente para uma função limite $f_0(x)$ se

$$\forall x \in U, \forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ tal que } n > N \implies |f_n(x) - f_0(x)| < \varepsilon.$$

ou, para cada x do domínio da função

$$\lim_{n\to\infty} f_n(x) = f_0(x) \quad \forall x \in U.$$

Ou seja, teremos uma convergência forte se a sequência converge ponto a ponto (x) do domínio U da função (o limite deve ser avaliado para cada x de forma independente). Em outras palavras, olhamos ponto a ponto e avaliamos se neste ponto temos convergência.

■ **Exemplo 9.8** Considere a função $f_n(x) = x^n$ para $U = [0,1]$. Essa sequência converge fortemente para

$$f(x) = \begin{cases} 0 & x < 1 \\ 1 & x = 1 \end{cases}.$$

De fato, se $x = 0$, temos que $f_n(0) = 0^n = 0$. Para $x = 1$, temos que $f_n(1) = 1^n = 1$. Para os pontos no interior do intervalo $f_n(x) = x^n$ e

$$\lim_{n\to\infty} x^n = 0, \quad 0 < x < 1.$$

■

Uma analogia muito interessante que eu li é que podemos olhar cada ponto da sequência $f_n(x)$ como um elevador. Se cada elevador chegar no seu andar $f_0(x)$, então dizemos que a convergência é forte. No entanto, não estamos preocupados com a forma como cada elevador chega (mais rápido ou mais devagar). Um conceito diferente de convergência é a convergência uniforme, onde ao invés de olhar cada elevador de forma independente, perguntamos se eles chegam "juntos"aos seus andares (aqui, então, estamos preocupados com a "forma"da convergência e não com valores pontuais). Vamos a definição:

Definição 9.2.2 Convergência uniforme de uma sequência de funções
Seja uma sequência de funções $f_n(x)$ com domínio U. A sequência de funções é dita uniformemente convergente se

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ tal que } |f_n(x) - f_0(x)| < \varepsilon \forall x \in U.$$

Ou seja, se, para cada $\varepsilon > 0$ existe $N > 0$, dependente de ε mas não de x, tal que

$$\|f_n(x) - f_0(x)\| < \varepsilon, \forall n > N \text{ e }, \forall x \in U$$

e, neste caso, escrevemos

$$\lim_{n\to\infty} f_n = f_0.$$

Uma maneira alternativa de descrevermos a mesma definição é

$$f_n \to f_0 \implies \lim_{n\to\infty} \|f_n - f_0\|_{uni} = 0$$

onde a norma uniforme é definida como

$$\|f\|_{uni} = \sup\{|f(x)| : \forall x \in U\}.$$

A principal diferença em relação a convergência forte é que na uniforme não nos preocupamos com o comportamento em cada ponto, mas sim da função como um todo no intervalo. Uma maneira de pensar nisso é que a função **como um todo** se aproxime de f_0 (o supremo indica que não estamos deixando nenhum ponto para trás).

■ **Exemplo 9.9** Considere a função $f_n(x) = x^n$ para $U = [0,1]$. Essa sequência não converge uniformemente para

$$f(x) = \begin{cases} 0 & x < 1 \\ 1 & x = 1 \end{cases}$$

pois, para $0 < x < 1$

$$|x^n - 0| < \varepsilon.$$

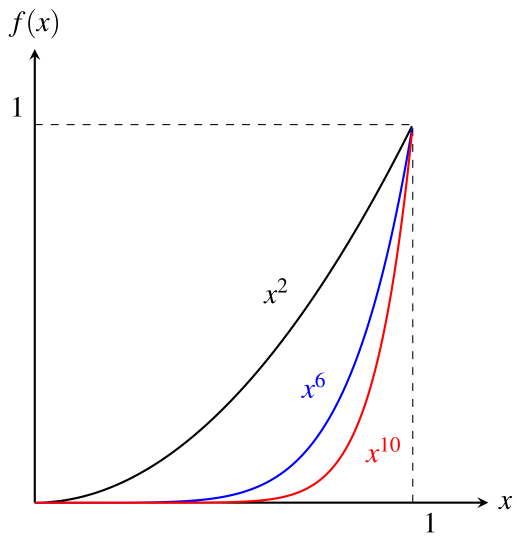
Assim, para esse exemplo, escolhendo $x = 1 - \frac{1}{n}$ (para nos aproximarmos da direita do intervalo)

$$\left| \left(1 - \frac{1}{n}\right)^n \right| < \varepsilon$$

e, para n suficientemente alto,

$$\left| \frac{1}{e} \right| < \varepsilon$$

Assim, não importa quão grande n seja, sempre teremos uma distância não nula entre a sequência e a função 0. O gráfico abaixo permite visualizar a sequência de funções x^n, de onde podemos notar que para $n \to \infty$ termos, eventualmente, convergência forte (os pontos vão ao valor zero em $0 < x < 1$, mas não uniforme (curvas não tendem uniformemente para o 0).



■

Definição 9.2.3 Convergência uniforme implica em convergência forte
Para entendermos essa afirmação, lembramos que a convergência uniforme implica em

$$f_n \to f_0 \implies \lim_{n\to\infty} \|f_n - f_0\|_{uni} = 0$$

garantindo que

$$\sup(|f_n(x) - f_0(x)|) < \varepsilon, \quad \forall x \in U$$

o que também garante que

$$|f_n(x) - f_0(x)| < \varepsilon, \quad \forall x \in U,$$

que é a condição para convergência forte. Resumindo,

1. Convergência uniforme implica em convergência forte;
2. Convergência forte não implica, necessariamente, em convergência uniforme.

■ **Exemplo 9.10** Seja a sequência de funções

$$f_k(x) = \frac{1}{3} + \sum_{n=1}^{k} 2\left(\frac{(\pi^2 n^2 - 2)\sin(\pi n) + 2\pi n \cos(\pi n)}{\pi^3 n^3}\right)\cos(n\pi x).$$

Essa sequência converge uniformemente para $f_0(x) = x^2$ em $x = [-1,1]$ pois

$$\lim_{k\to\infty} \|f_k - f_0\|_{sup} = 0.$$

Embora esse não seja um cálculo muito trivial, podemos utilizar um recurso gráfico como auxiliar no entendimento desse resultado.

■

Um conceito de fundamental importância para a área de mecânica computacional é o de convergência fraca.

Definição 9.2.4 Convergência Fraca (*weak convergence*)
Seja uma sequência $\{f_n\}$ e uma função genérica w em um espaço de (pré-)Hilbert H. Esta

sequência é dita fracamente convergente para f_0 em H se

$$\langle f_n, w\rangle \to \langle f_0, w\rangle\,, \quad \forall w \in H$$

e, neste caso, empregamos a notação

$$f_n \rightharpoonup f_0.$$

Desta forma, observamos que

$$\|f\| \leq \lim_{n\to\infty} \inf \|f_n\|$$

■ **Exemplo 9.11** Seja a sequência $f_n(x) = sin\left(\frac{n}{2}x\right)\; x \in [0,\pi]$ em L^2.

Essa sequência irá apresentar diversas passagens por zero ao longo do intervalo, a medida que aumentarmos sua frequência n. No entanto, isto não significa que ela irá tender para a função $z(x) = 0$, isto é, não temos convergência uniforme da sequência. Isso pode ser avaliado pois sabemos que os valores de pico da função $f_n(x)$ são ± 1, tal que

$$\|f_n - z\|_{sup} = 1$$

e esse valor não depende de n.

No entanto,

$$\lim_{n\to\infty} \langle f_n(x), v(x)\rangle = \lim_{n\to\infty} \int_0^{\pi} \sin\left(\frac{n}{2}x\right) v(x)\, dx \to 0 = \langle z(x), v(x)\rangle$$

independente de $v(x)$, pois ficamos sempre com um n no denominador. Portanto, dizemos que $f_n \rightharpoonup z$, ou seja, temos convergência fraca. ■

A convergência fraca é muito importante, pois como o nome diz, implica em uma medida mais "tolerante"do conceito de convergência. É importante ressaltar que se uma sequência converge uniformemente então também converge fracamente.

Por fim, podemos avaliar que todos os conceitos de convergência discutidos acima também podem ser aplicados a operadores. Por exemplo:

Definição 9.2.5 Convergência Uniforme de um operador

Uma sequência de Operadores $\{T_n\}$ é dita uniformemente convergente para T_0 em um domínio Ω se, para cada $\varepsilon > 0$ existe $N > 0$, dependente de ε mas não de $\mathbf{u}$, tal que

$$|T_n(\mathbf{u}) - T_0(\mathbf{u})| < \varepsilon, \forall n > N\;, \;\forall \mathbf{u} \in \Omega$$

e, neste caso, escrevemos

$$\lim_{n\to\infty} T_n = T_0.$$

9.3 Espaço Completo

Se todas as sequências convergem para elementos dentro do espaço, então dizemos que o mesmo é completo. Este conceito é muito importante, uma vez que está associado a capacidade de obtermos efetivamente o limite de uma seqência. Por exemplo: A sequência $x_n = \frac{1}{n}$ pode ser definida no intervalo $(0,K)$, onde K é um número positivo maior do que 1. Esta sequência é claramente de Cauchy, mas seu limite converge para um elemento fora do domínio (espaço). Outro exemplo clássico será discutido no capítulo sobre Séries de Fourier, onde será demonstrado que uma sequência de funções contínuas pode convergir para uma função contínua por partes. Um outro exemplo clássico é a sequência

$$x_0 = 1, x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}$$

que é formada por números racionais, mais converge para $\sqrt{2}$, que é um número irracional. Isto comprova que o conjunto dos racionais não é completo, necessitando assim da definição dos números irracionais, que **completam** os racionais, formando o Espaço dos Números Reais.

Definição 9.3.1 Espaço de Hilbert
Um espaço de Hilbert H é um espaço com produto interno, norma induzida pelo produto interno e é completo (todas as sequências convergem para elementos do espaço quando avaliadas por esta norma).

10. Série de Fourier

Já vimos que espaços vetoriais são a extensão dos conceitos associados a um espaço que é bastante conhecido (e intuitivo), como o $\mathbb{R}^N$. Desta forma, definimos e trabalhamos com conceitos fundamentais, como o de função, operadores, norma, produto interno e sequências. De fundamental interesse para os nossos estudos em fundamentos de matemática é uma série de funções especialmente importante, conhecida como Série de Fourier.

Seja um conjunto C^0, das funções contínuas, definidas no intervalo $[-L, L]$. Se este conjunto for munido das operações binárias básicas definidas anteriormente, então constitui um espaço de dimensão infinita (cardinalidade da base é infinita), com produto interno definido por

$$\langle u(x), v(x) \rangle = \int_{-L}^{L} u(x)v(x)\,dx$$

e norma induzida pelo produto interno

$$\|u\|_2 = \langle u, u \rangle^{\frac{1}{2}}.$$

Conforme já estudamos, existem diferentes bases para um mesmo espaço. Das diversas bases das quais podemos dispor para descrever as funções contínuas C^0, uma é de interesse especial é

$$B = \left\{ \frac{1}{2} \right\} \cup \left\{ \cos\left(\frac{n\pi x}{L} \right) \right\} \cup \left\{ \sin\left(\frac{n\pi x}{L} \right) 101 \right\}, n \in \mathbb{N}^*$$

tal que, nesta base, um elemento do espaço C^0 é descrito por uma combinação linear

$$u(x) = a_0 \frac{1}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L} \right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{L} \right)$$

onde o somatório vai a infinito, pois esta é a cardinalidade da base (a combinação linear implica em usar todos os termos da base, pois essa é a dimensão do espaço).

Os coeficientes da série são obtidos por meio das projeções da função nas direções definidas pela base. Desta forma, se chamarmos os termos da base de 101

$$B = \{c_0\} \cup \{c_n\} \cup \{s_n\}, n \in \mathbb{N}^*$$

então

$$a_0 = \frac{\langle u, c_0 \rangle}{\langle c_0, c_0 \rangle} = \frac{2}{L}\int_{-L}^{L} \frac{u(x)}{2}\,dx \tag{10.1}$$

$$a_n = \frac{\langle u, c_n \rangle}{\langle c_n, c_n \rangle} = \frac{1}{L}\int_{-L}^{L} u(x) c_n(x)\,dx \tag{10.2}$$

$$b_n = \frac{\langle u, s_n \rangle}{\langle s_n, s_n \rangle} = \frac{1}{L}\int_{-L}^{L} u(x) s_n(x)\,dx$$

tal que uma sequência que aproxime $u(x)$ pode ser definida como (observe o limite do somatório)

$$u_n(x) = a_0\frac{1}{2} + \sum_{k=1}^{n} a_k \cos\left(\frac{k\pi x}{L}\right) + \sum_{k=1}^{n} b_k \sin\left(\frac{k\pi x}{L}\right).$$

Teorema 10.0.1 Teorema de Fourier

Seja $f : [-L, L] \to \mathbb{R}$ uma função derivável por partes.

Então, em um ponto $x \in (-L, L)$ a série de Fourier converge fortemente para

$$f_n(x) = \frac{f(x^+) + f(x^-)}{2}$$

.

Nos extremos $x = -L$ e $x = L$, temos as seguintes situações:

- se f for periódica com período igual a L, então $f_n(-L) = f_n(L) = f(L)$;
- se f não for periódica, então f_n converge para a média dos valores laterais, ie,

$$f_n = \frac{f(-L^+ + L^-)}{2}$$

É importante complementar a definição acima com as chamadas condições de Dirichlet. Estas 3 condições devem ser observadas para que uma função seja corretamente descrita pela série de Fourier:

Definição 10.0.1 Condições de Dirichlet

Uma função $f(x)$, para ser descrita por uma série de Fourier, deve satisfazer as seguintes condições:

- A função deve ser absolutamente integrável, ie, $\int |f(x)| dx < \infty$, implicando que os coeficientes serão limitados;
- Não podem existir infinitos pontos de máximo e/ou de mínimo. Um exemplo que não atende a este requisito é a função $f(x) = \sin\left(\frac{\pi}{x}\right)$ quando $x \to 0$;
- Não podem existir infinitas descontinuidades.

■ **Exemplo 10.1** Vamos considerar a função $f(x) = x^2$, para $x \in [-1, 1]$. A representação desta função na base que está sendo discutida será exata para $n \to \infty$, pois $f(x) \in C^0$ e $f(-1) = f(1)$. No entanto, como exercício vamos estudar o comportamento da aproximação a medida que aumentamos o número de termos na combinação linear.

1. Um termo (n=0):
   Neste caso, $f_0 = \frac{a_0}{2}$, com $a_0 = \frac{2}{(1)}\int_{-1}^{1} x^2 \frac{1}{2}\,dx = \frac{2}{3}$. Portanto, $f_0(x) = \frac{1}{3}$.

2. Termo constante mais os primeiros trigonométricos (n=1):

Neste caso, $f_1(x) = \frac{1}{3} + a_1 \cos\left(\frac{\pi x}{L}\right) + b_1 \sin\left(\frac{\pi x}{L}\right)$, com

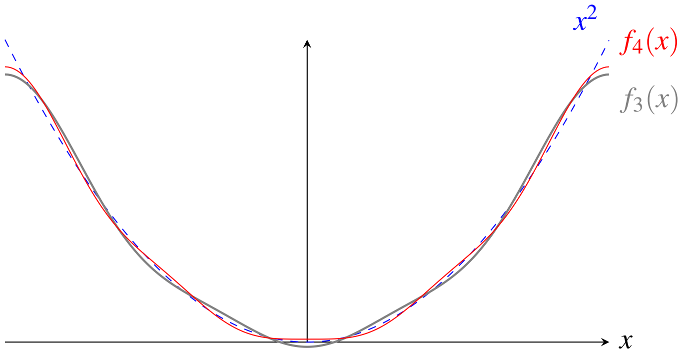
$$a_1 = \frac{1}{1}\int_{-1}^{1} x^2 \cos\left(\frac{\pi x}{L}\right)\, dx = -\frac{4}{\pi^2} \tag{10.3}$$
$$b_1 = \frac{1}{1}\int_{-1}^{1} x^2 \sin\left(\frac{\pi x}{L}\right)\, dx = 0$$

tal que

$$f_1(x) = \frac{1}{3} - \frac{4}{\pi^2}\cos\left(\frac{\pi x}{1}\right).$$

3. (n=2) $f_2(x) = \frac{\cos(2\pi x)}{\pi^2} - \frac{4\cos(\pi x)}{\pi^2} + \frac{1}{3}$
4. (n=3) $f_3(x) = -\frac{4\cos(3\pi x)}{9\pi^2} + \frac{\cos(2\pi x)}{\pi^2} - \frac{4\cos(\pi x)}{\pi^2} + \frac{1}{3}$
5. (n=4) $f_4(x) = \frac{\cos(4\pi x)}{4\pi^2} - \frac{4\cos(3\pi x)}{9\pi^2} + \frac{\cos(2\pi x)}{\pi^2} - \frac{4\cos(\pi x)}{\pi^2} + \frac{1}{3}$

Podemos visualizar a melhora na representação da função x^2 a medida que vamos acrescentando termos à série:



■

Do exemplo acima, podemos observar que os coeficientes b_n sempre nulos para a aproximação da função x^2. Isso ocorre pois a função é par.

Definição 10.0.2 Funções Pares e Funções Ímpares
Seja $f(x)$ em $x \in [-L, L]$. Essa função é dita:
- par se $f(-x) = f(x)$
- ímpar se $f(-x) = -f(x)$

$\forall x \in [-L, L]$.

Assim, pelas propriedades trigonométricas do sin e do cos, observamos que em funções pares os coeficientes b_n são zero, enquanto em funções ímpares os coeficientes a_n serão nulos.

■ **Exemplo 10.2** A função x^2 é par, pois $(-x)^2 = x^2$ em todo o intervalo $[-1, 1]$. Conforme ilustrado no exemplo anterior, todos os b_n se anulam. ■

■ **Exemplo 10.3** A função x^3 é ímpar, pois $(-x)^3 = -x^3$ em todo o intervalo $[-1, 1]$. Assim, os coeficientes a_n serão anulados. ■

Embora a definição da Série de Fourier tenha sido definida em um intervalo $[-L, L]$, não estamos restritos a aplicar o método somente nesta situação. Para isso, podemos utilizar o conceito de extensão periódica

Definição 10.0.3 Extensão periódica de uma função

Caso uma função seja definida em um intervalo $x \in [0,L]$, por exemplo, então é necessário estender o domínio da função para que seja possível aplicar as equações deduzidas acima. Neste caso, existem duas possibilidades:

- Extensão Par:

$$f_{par}(x) = \begin{cases} f(x) & x \in \quad [0,L] \\ f(-x) & x \in \quad [-L,0) \end{cases}$$

- Extensão Ímpar:

$$f_{ímpar}(x) = \begin{cases} f(x) & x \in \quad [0,L] \\ -f(x) & x \in \quad [-L,0) \end{cases}$$

■ **Exemplo 10.4** Seja $f(x) = 4x - 3x^2 - 2x^3$ e $x \in [0,1]$.

Temos as seguintes opções:

$$f_{\text{par}}(x) = \begin{cases} 4x - 3x^2 - 2x^3 & x \in \quad [0,1] \\ -4x - 3x^2 + 2x^3 & x \in \quad [-1,0) \end{cases}$$

e

$$f_{\text{ímpar}}(x) = \begin{cases} 4x - 3x^2 - 2x^3 & x \in \quad [0,1] \\ 4x + 3x^2 - 2x^3 & x \in \quad [-1,0) \end{cases}$$

Assim, estudando a série de Fourier para a extensão par para essa função, obtemos :

$$a_0 = \frac{2}{1}\left(\int_{-1}^{0} f_{\text{par}}(x)\frac{1}{2}\,dx + \int_{0}^{1} f(x)\frac{1}{2}\,dx\right) \tag{10.4}$$

$$a_n = \left(\int_{-1}^{0} f_{\text{par}}(x)\cos\left(\frac{n\pi x}{L}\right)\,dx + \int_{0}^{1} f(x)\cos\left(\frac{n\pi x}{L}\right)\,dx\right) \tag{10.5}$$

$$b_n = \left(\int_{-1}^{0} f_{\text{par}}(x)\sin\left(\frac{n\pi x}{L}\right)\,dx + \int_{0}^{1} f(x)\sin\left(\frac{n\pi x}{L}\right)\,dx\right)$$

resultando, para $n = 3$, em

$$f_3(x) = -\frac{\left(32 - 48\,\pi^2\right)\cos(3\,\pi\,x) + 324\,\pi^2\cos(2\,\pi\,x) + \left(2592 - 432\,\pi^2\right)\cos(\pi\,x) - 27\,\pi^4}{54\,\pi^4}$$

.

■

Exercício 10.1 Obtenha a série de Fourier para $f(x) = x^3$ e $x \in [0,5]$. Estude quantos termos são necessários para que obtenhamos um erro máximo de 5%. ■

Exercício 10.2 Obtenha a série de Fourier de $f(x) = e^x$, $x \in [-L,L]$. ■

Exercício 10.3 Considere a função $t(x) = \frac{x+|x|}{2}$ para $x \in [-5,5]$. Obtenha a série e investigue os valores da aproximação em $x = 0$. ■

Exercício 10.4 Obtenha a série de Fourier para $f(x) = H_0(x)$, função *Heaviside* com centro em 0, para $x \in [-1, 1]$. ■

Exercício 10.5 A relação entre o operador exponencial e as funções trigonométricas é bem conhecida (Fórmula de Euler). Desta forma, é possível escrever a base que utilizamos para obter a série de Fourier utilizando uma notação mais compacta. Estude esta base e obtenha as expressões dos coeficientes para este caso. ■