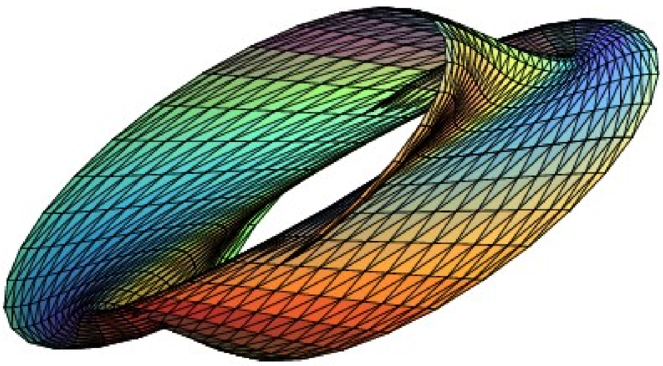
ISSN 1980–024X

Boletim de Iniciação Científica em Matemática · BICMat



Volume XVII
Outubro de 2020
Departamento de Matemática
IGCE · Rio Claro

unesp

ISSN 1980–024X

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat

Comissão editorial

Elíris Cristina Rizziolli
Renata Zotin Gomes de Oliveira
Nativi Viana Pereira Bertolo
Thiago de Melo

Editoração gráfica

Thiago de Melo

Realização

Conselho de Curso de Graduação em Matemática
Departamento de Matemática
IGCE · Unesp Rio Claro

Editorial

O Boletim de Iniciação Científica em Matemática · BICMat é uma publicação que se destina a difundir prioritariamente trabalhos de Iniciação Científica em Matemática que fazem parte de projetos desenvolvidos por alunos do Curso de Graduação em Matemática do IGCE, Unesp Rio Claro. Eventualmente trabalhos de Iniciação Científica realizados em outras instituições poderão também ser publicados neste Boletim.

O BICMat foi criado em 1998 e nessa época foram publicados dois volumes; o primeiro no ano de criação e o segundo em 2000.

Considerando a importância da Iniciação Científica para o graduando, e o sempre crescente número de projetos desta natureza desenvolvidos em nossa instituição, resolvemos reativar a publicação do BICMat em 2006, com ISSN 1980–024X.

Destacamos que a autoria dos trabalhos apresentados no BICMat é dos alunos. O orientador figura apenas como responsável científico.

Este Boletim também está aberto à divulgação de trabalhos que não sejam frutos de projetos de Iniciação Científica, mas que sejam de interesse dos alunos do curso de graduação em Matemática. Estes trabalhos serão selecionados pelos Editores.

Este volume está disponibilizado eletronicamente na página do Departamento de Matemática no endereço

`http://igce.rc.unesp.br/departamentos/matematica/`

ou através do código QR

Sumário

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 6–15

Diferenciabilidade e Otimização de Funções Definidas em Espaços Normados

Carolina Kakazu Paula[†]
Orientador(a): Profa. Dra. Marta Cilene Gadotti

Resumo: Este trabalho faz parte de uma Iniciação Científica no qual estudamos a teoria dos operadores e as aplicações. Nesta publicação em específico, ilustraremos uma parte da teoria envolvendo diferenciabilidade e otimização de funções definidas em espaços normados.

Palavras-chave: Derivada; Otimização; Ponto Crítico; Espaço de Dimensão Infinita.

1 Introdução

Os problemas de otimização de uma função dada são bastante comuns para aumento de produção, aumento de lucro, diminuição de tempo para uma fábrica e/ou empresa. Para analisar tais problemas é necessário o estudo de alguns conceitos como a diferenciabilidade e as propriedades que nos possibilitem a chegar em alguns resultados esperados.

Em algumas situações as funções a serem otimizadas estão definidas em conjuntos mais gerais que $\mathbb{R}$. Neste trabalho, consideraremos $F : X \to \mathbb{R}$, onde X é um espaço normado, que pode ter dimensão infinita.

2 A derivada

Antes da definição de diferenciabilidade de uma função definida no espaço normado, precisaremos dos seguintes conceitos:

Definição 2.1. Sejam X, Y dois espaços vetoriais sobre o mesmo corpo $\mathbb{K}$. Uma aplicação $T : D(T) \subset X \to Y$ diz-se um operador linear se

- O domínio $D(T)$ é um espaço vetorial e a imagem $R(T) \subset Y$;
- Para quaisquer $x, y \in D(T)$ e escalar $\alpha \in K$ temos

$$T(x+y) = Tx + Ty$$
$$T(\alpha x) = \alpha Tx.$$

Definição 2.2. Sejam $X.Y$ espaços normados e $T : D(T) \subset X \to Y$ um operador linear. Então T diz-se limitado se existe uma constante c tal que para todo $x \in D(T)$ ter

$$\|Tx\| \leqslant c\|x\|$$

[†]FNDE - PET

Observação: Note que, a presente definição de operador limitado difere da função limitada da análise, pois, uma função limitada é aquela cujo conjunto imagem é limitado; no caso dos operadores lineares só o operador nulo tem o conjunto imagem limitado.

Assim, quando o operador é um operador limitado definimos

$$\|T\| = \sup_{x \in D(T)-\{0\}} \frac{\|Tx\|}{\|x\|}.$$

Teorema 2.3. *Seja $T : D(T) \subset X \to Y$ um operador linear entre espaços normados. Então*

1. *T é contínuo se, e somente se, T é limitado;*
2. *Se T é contínuo num ponto, então é contínuo.*

Prova: 1. ($\Leftarrow$) Suponhamos que T é limitado. Queremos mostrar que T é contínuo. Se T é nula o resultado já é válido. Caso contrário, dado $\epsilon > 0$ e $x_0 \in D(T)$ um elemento arbitrário. Escolhendo $\delta = \frac{\epsilon}{\|T\|}$ e para qualquer $x \in D(T)$ tal que

$$\|x - x_0\| \leqslant \delta$$

temos

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leqslant \|T\|.\|x - x_0\| < \|T\|\delta = \epsilon.$$

Logo, T é contínuo em x_0; da arbitrariedade de $x_0 \in D(T)$ resulta que T é contínuo.

($\Rightarrow$) Suponhamos agora que T é contínuo e queremos mostra que T é limitado. Assim, para qualquer $x_0 \in D(T)$ e qualquer $\epsilon > 0$ existe $\delta > 0$ e para todo $x \in D(T)$ com $\|x-x_0\| < \delta$ temos

$$\|Tx - Tx_0\| < \epsilon.$$

Para $y \neq 0$ em $D(T)$ definimos

$$x := x_0 + \frac{\delta}{\|y\|}y \Longleftrightarrow x - x_0 = \frac{\delta}{\|y\|}y \Longrightarrow \|x - x_0\| = \delta.$$

Logo,

$$\|Tx - Tx_0\| = \|T(x - x_0)\| = \left|T\left(\frac{\delta}{\|y\|}y\right)\right| = \frac{\delta}{\|y\|}|Ty| < \epsilon$$

ou ainda

$$\frac{|Ty|}{\|y\|} < \frac{\epsilon}{\delta}.$$

Passando ao supremo em ambos os lados sobre todos os $y \in D(T)$ tal que $\|y\| = 1$ obtemos $\|T\| < \frac{\epsilon}{\delta}$, isto é, T é limitado.

2. Se T é contínuo num ponto, então pela segunda parte da prova de 1. T é limitado, logo por 1. T é contínuo. ■

Denotaremos o conjunto $\mathcal{L}(X, Y)$, onde

$$\mathcal{L}(X, Y) = \{F : X \to Y \mid F \text{ é um operador linear}\}.$$

Definição 2.4. Sejam X e Y espaços normados, $F : X \to Y$ uma aplicação e $x_0 \in X$. Dizemos que F é diferenciável em x_0 se existe um operador linear limitado $L \in \mathcal{L}(X, Y)$ tal que dado $\epsilon > 0$ existe $\delta > 0$ tal que se $x \in X - x_0$

$$\|x - x_0\| < \delta \Longrightarrow \frac{\|F(x) - F(x_0) - L(x - x_0)\|}{\|x - x_0\|} < \epsilon. \tag{2.1}$$

Denotamos por $DFx_0 = L$

O operador L é chamado derivada de F em x_0. Se F é derivável em todo ponto $x \in X$, então dizemos simplesmente que é derivável.

Provemos agora que se F é derivável no ponto $x_0 \in X$, a derivada é única.

Teorema 2.5. *Sejam X e Y espaços normados. Se $F : X \to Y$ é derivável em $x_0 \in X$, então esta derivada é única.*

Prova: Suponha que L_1 e L_2 são derivadas de F em x_0. Dado $\epsilon > 0$ escolha δ como (2.1) com L_1 e L_2 no lugar de L. Logo, para todo $x \in X - x_0$ temos

$$\|x - x_0\| < \delta \Longrightarrow \frac{\|L_2(x - x_0) - L_1(x - x_0)\|}{\|x - x_0\|} < 2\epsilon. \tag{2.2}$$

Dado $h \in X$ qualquer tal que $h \neq 0$, defina

$$x = x_0 + \frac{\delta}{2\|h\|}h$$

Então $\|x - x_0\| = \frac{\delta}{2} < \delta$. Disso, junto com (2.2) temos

$$\|(L_2 - L_1)h\| \leqslant 2\epsilon\|h\|.$$

Daí, $\|(L_2 - L_1)\| \leqslant 2\epsilon$ e como $\epsilon > 0$ é arbitrário, temos que $\|L_2 - L_1\| = 0$. Logo, $L_2 = L_1$, provando a unicidade. ∎

Exemplo 2.6. Considere $C[a, b] = \{f : [a, b] \to \mathbb{R} \text{ contínuas}\}$ com a norma $\|f\| = \max_{t\in[a,b]} |f(t)|$.

Defina

$$\begin{aligned} F : C[a, b] &\longrightarrow C[a, b] \\ u &\longmapsto (F(u))(t) = (u(t))^2. \end{aligned}$$

Mostremos que ela é derivável.

De fato, observe primeiro que

$$(Fu - Fu_0)(t) = u(t)^2 - u_0(t)^2 = (u(t) + u_0(t))(u(t) - u_0(t)). \tag{2.3}$$

Como u se aproxima de u_0 em $C[a, b]$, o termo $u(t)+u_0(t)$ se aproxima de $2u_0(t)$. Mostremos que operador multiplicação M dado por:

$$(Mu)(t) = 2u_0(t)u(t), \quad t \in [a, b]$$

é a $DF(u_0)$.

Dado $\epsilon > 0$, temos

$$\begin{aligned} |(Fu - Fu_0 - M(u - u_0))(t)| &= |u(t)^2 - u_0(t)^2 - 2u_0(t)(u(t) - u_0(t))| \\ &= |u(t)^2 - u_0(t)^2 + 2u_0(t)^2 - 2u_0(t)u(t)| \\ &= |u(t)^2 + u_0(t)^2 - 2u_0(t)u(t)| \\ &= |u(t) - u_0(t)|^2 \\ &\leqslant \|u - u_0\|^2 \end{aligned}$$

Consequentemente $\delta := \epsilon > 0$, para todo $u \in C[a, b]$ satisfazendo $\|u - u_0\| < \delta$, temos

$$\|Fu - Fu_0 - M(u - u_0)\| \leqslant \|u - u_0\|^2,$$

e então para todo $u \in C[a, b] - \{u_0\}$ satisfazendo $\|u - u_0\| < \delta$, obtemos

$$\frac{\|Fu - Fu_0 - M(u - u_0)\|}{\|u - u_0\|} \leqslant \|u - u_0\| < \delta = \epsilon.$$

Portanto, $DF(u_0) = M$

3 Otimização: A necessidade da derivada se tornar nula

Consideremos aqui $Y = \mathbb{R}$, e assim $I : X \to \mathbb{R}$. Queremos encontrar pontos $x_0 \in X$ que maximiza (ou minimiza) I.

Na Análise, a condição necessária para uma função derivável $f : \mathbb{R} \to \mathbb{R}$ possuir um extremo (máximo local ou mínimo local) no ponto $x_0 \in \mathbb{R}$ é que $DF(x_0) = 0$. Vamos provar uma condição necessária similar para uma função derivável $I : X \to \mathbb{R}$.

Primeiro, vamos especificar o que significa ser um máximo/mínimo local. A grosso modo, um ponto $x_0 \in X$ é um máximo local (ou mínimo local) de I se para todo ponto x numa vizinhança de x_0, todos os valores de $I(x)$ são menores (respectivamente maiores) que $I(x_0)$.

Definição 3.1. Seja X um espaço normado. Uma função $I : X \to \mathbb{R}$ possui um extremo local em $x_0 (\in X)$ se existe $\delta > 0$ tal que

$$\forall x \in X - \{x_0\} \quad \text{satisfazem} \quad \|x - x_0\| < \delta,\ I(x) \geqslant I(x_0) \quad \text{(mínimo local)}$$

ou

$$\forall x \in X - \{x_0\} \quad \text{satisfazem} \quad \|x - x_0\| < \delta,\ I(x) \leqslant I(x_0) \quad \text{(máximo local).}$$

Teorema 3.2. *Sejam X um espaço normado e $I : X \to \mathbb{R}$ uma função derivável no ponto $x_0 \in X$. Se o ponto x_0 é extremo local, então $(DI)(x_0) = 0$.*

Prova: Provaremos esta afirmação no caso em que x_0 é mínimo local de I (se x_0 é máximo local então a função $(-I)$ tem mínimo local em x_0 e vale $(DI)(x_0) = (D(-I))(x_0) = 0$).

Para simplificar a notação escreveremos $(DI)(x_0)$ por L. Suponha que $Lh \neq 0$ para algum $h \in X$. dado $\epsilon > 0$, escolha δ tal que $\forall x \in X$ com $\|x - x_0\| < \delta$ tem-se $I(x) \geqslant I(x_0)$, e ainda, se $x \neq x_0$,

$$\frac{|I(x) - I(x_0) - L(x - x_0)|}{\|x - x_0\|} < \epsilon.$$

Defina a sequência

$$x_n = x_0 - \frac{1}{n} . \frac{Lh}{|Lh|} h, \quad n \in \mathbb{N}.$$

Notamos que $\|x_n - x_0\| = \frac{\|h\|}{n}$, e escolhendo um n_0 suficientemente grande, temos $\|x_n - x_0\| < \delta$ para todo $n > n_0$.

Segue então que para todo $n > n_0$,

$$0 \leqslant \frac{I(x_n) - I(x_0)}{\|x_n - x_0\|} < \frac{L(x_n - x_0)}{\|x_n - x_0\|} + \epsilon = -\frac{|Lh|}{\|h\|} + \epsilon.$$

Como $\epsilon > 0$ é arbitrário, temos que $|Lh| \leqslant 0$, e então $Lh = 0$, o que é uma contradição. ∎

Observação: Note que isto é uma condição necessária para existência do extremo local. Mas o fato da derivada zerar não implica que x_0 será extremo local de I. Um bom exemplo para mostrarmos isso é a $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^3$, cujo $f'(0) = 0$ mas a função não tem máximo nem mínimo em 0.

4 Otimização: no caso convexo.

Nesta seção, mostraremos que se $I : X \to \mathbb{R}$ é uma função convexa, então concluiremos que o ponto crítico desta função é um ponto de mínimo global. Começamos pela definição de função convexa:

Definição 4.1. Seja X um espaço normado. Uma função $F : X \to \mathbb{R}$ é convexa se para todos $x_1, x_2 \in X$ e todo $\alpha \in [0,1]$

$$F(\alpha x_1 + (1-\alpha)x_2) \leqslant \alpha F(x_1) + (1-\alpha)F(x_2).$$

Exemplo 4.2. Se $X = \mathbb{R}$, então a função $f(x) = x^2, \quad x \in \mathbb{R}$ é convexa.
Podemos provar isto do seguinte modo:

$$\begin{aligned}
f(\alpha x_1 + (1-\alpha)x_2) &= (\alpha x_1 + (1-\alpha)x_2)^2 \\
&= \alpha^2 x_1^2 + 2\alpha(1-\alpha)x_1x_2 + (1-\alpha)^2 x_2^2 \\
&= \alpha^2 x_1^2 + 2\alpha x_1 x_2 - 2\alpha^2 x_1 x_2 + (1 - 2\alpha + \alpha^2)x_2^2 \\
&= \alpha x_1^2 + (1-\alpha)x_2^2 + (\alpha^2-\alpha)x_1^2 + (\alpha^2-\alpha)x_2^2 + 2\alpha(1-\alpha)x_1x_2 \\
&= \alpha x_1^2 + (1-\alpha)x_2^2 - \alpha(1-\alpha)(x_1^2 + x_2^2 - 2x_1x_2) \\
&= \alpha x_1^2 + (1-\alpha)x_2^2 - \alpha(1-\alpha)(x_1 - x_2)^2 \\
&\leqslant \alpha x_1^2 + (1-\alpha)x_2^2 \\
&= \alpha f(x_1) + (1-\alpha)f(x_2)
\end{aligned}$$

Segue da definição, que a função é convexa.

Afim de provar o teorema sobre a condição suficiente do ponto crítico, no caso da função convexa, precisamos do seguinte resultado que nos diz que se uma função derivável f é convexa então a derivada f' é uma função crescente, isto é, se $x \leqslant y$, então $f'(x) \leqslant f'(y)$.

Lema 4.3. *Se $f : \mathbb{R} \to \mathbb{R}$ é convexa e diferenciável, então f' é uma função crescente*

Prova: Se f é convexa, então

$$f((1-t)a + tb) \leqslant (1-t)f(a) + tf(b), \quad t \in [0,1]. \tag{4.1}$$

Fazendo $x = (1-t)a + tb$, temos

$$t = \frac{x-a}{b-a} \quad \text{e} \quad 1 - t = \frac{b-x}{b-a}.$$

Então (4.1) se torna:

$$f(x) \leqslant \left(\frac{b-x}{b-a}\right)f(a) + \left(\frac{x-a}{b-a}\right)f(b)$$

Subtraindo $f(a)$ em ambos os lados e multiplicando por $\frac{1}{x-a}$, obtemos

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{1}{x-a}\left[\left(\frac{b-x}{b-a}\right)f(a) + \left(\frac{x-a}{b-a}\right)f(b)\right]$$

$$\Rightarrow \frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a} \tag{4.2}$$

Analogamente, temos

$$\frac{f(b)-f(x)}{b-x} \geqslant \frac{f(b)-f(a)}{b-a} \tag{4.3}$$

Como f é convexa em $\mathbb{R}$, de (4.2) e(4.3) temos

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a} \leqslant \frac{f(b)-f(x)}{b-x}, \quad \forall x \in (a,b)$$

Como f é derivável, fazendo $x \to a$ na primeira desigualdade e $x \to b$ na segunda desigualdade, obtemos

$$f'(a) \leqslant \frac{f(b) - f(a)}{b - a} \leqslant f'(b)$$

Portanto, f' é crescente ■

Estamos prontos para provar o resultado sobre a existência de mínimo global. Primeiro de tudo mencionamos que se I é uma função definida em um espaço normado X com imagem em $\mathbb{R}$, então I possui um mínimo global no ponto $x_0 \in X$ se $I(x) \geqslant I(x_0)$, $\forall x \in X$. Similarmente, se $I(x) \leqslant I(x_0)$ para todo x, então I possui um máximo global em x_0.

Também notamos que o problema de encontrar um máximo para aplicação I pode sempre ser convertido em um problema de minimização considerando a função $(-I)$ em vez de I. Provemos o que segue:

Teorema 4.4. *Sejam X um espaço normado e $I : X \to \mathbb{R}$ diferenciável. Suponha que I é convexa. Se $x_0 \in X$ é tal que $(DI)(x_0) = 0$, então I possui mínimo global em x_0.*

Prova: Suponha que $x_1 \in X$ e $I(x_1) < I(x_0)$. Defina $f : \mathbb{R} \to \mathbb{R}$ por

$$f(\alpha) = I(\alpha x_1 + (1 - \alpha)x_0), \quad \alpha \in \mathbb{R}.$$

A função f é convexa, pois se $r \in [0, 1]$ e $\alpha, \beta \in \mathbb{R}$, então temos

$$\begin{aligned} f(r\alpha + (1 - r)\beta) &= I((r\alpha + (1 - r)\beta)x_1 + (1 - r\alpha - (1 - r)\beta)x_0) \\ &= I(r(\alpha x_1 + (1 - \alpha)x_0) + (1 - r)(\beta x_1 + (1 - \beta)x_0)) \\ &\leqslant rI(\alpha x_1 + (1 - \alpha)x_0) + (1 - r)I(\beta x_1 + (1 - \beta)x_0) \\ &= rf(\alpha) + (1 - r)f(\beta). \end{aligned}$$

Também, f é diferenciável em $[0, 1]$ e vale

$$f'(0) = ((DI)(x_0))(x_1 - x_0) = 0.$$

Como $f(1) = I(x_1) < I(x_0) = f(0)$, de acordo com Teorema do Valor Médio, existe um $c \in (0, 1)$ tal que

$$f'(c) = \frac{f(1) - f(0)}{1 - 0} < 0 = f'(0).$$

isto contradiz a convexidade de f, e então $I(x_1) \geqslant I(x_0)$. Consequentemente, I tem mínimo global em x_0, como queríamos demonstrar. ■

5 Um exemplo de otimização em um espaço funcional

Considere a seguinte situação baseada na referência [1]:

Uma empresa mineradora de cobre pretende extrair todo o cobre de uma região que contém Q toneladas, durante um período de T anos. Assim que extraídos, irão vender para o processamento a um preço líquido por tonelada dado por

$$p(x(t), x'(t)) = P - ax(t) - bx'(t),$$

onde P, a, b são constantes e $x(t)$ denota a tonelagem total vendida no tempo t. (Este modelo de precificação permite que o custo da mineração aumente com a extensão da região mineradora e velocidade de produção).

Se a companhia quer maximizar seu lucro total dado por

$$I(x) = \int_0^T p(x(t), x'(t))x'(t)dt$$

onde $x(0) = 0$ e $x(T) = Q$, como isto deve ser realizado?

Passo 1: Em primeiro lugar notamos que o conjunto de curvas em $C^1[0, T]$ satisfazendo $x(0) = 0$ e $x(T) = Q$ não forma um espaço linear e, portanto, não é possível aplicar diretamente o Teorema 3.2. Por isso introduziremos um espaço linear X e consideraremos uma nova função $\tilde{I} : X \to \mathbb{R}$ que é definida em termos da função I anterior.

Inserimos o espaço linear

$$X = \{x \in C^1[0, T] \quad \text{tal que} \quad x(0) = x(T) = 0\},$$

com norma de $C^1[0, T]$:

$$\|x\| = \sup_{t\in[0,T]} |x(t)| + \sup_{t\in[0,T]} |x'(t)|.$$

Observe que para todo $h \in X$ e se x_0 satisfaz $x_0(0) = 0, x_0(T) = Q$, então $x_0 + h$ satisfaz $(x_0 + h)(0) = 0$ e $(x_0 + h)(T) = Q$. Definindo $\tilde{I} = I(x_0 + h)$, pode-se notar que $\tilde{I} : X \to \mathbb{R}$ possui um extremo no ponto 0. Segue do teorema 3.2 que $(D\tilde{I})(0) = 0$. Observe que o zero da direita refere-se ao zero funcional, ou seja, a aplicação linear de X para $\mathbb{R}$ que é definido por $h \mapsto 0$ para todo $h \in X$.

Passo 2: Calculamos agora o $\tilde{I}'(0)$. Temos

$$\begin{aligned}
\tilde{I}(h) - \tilde{I}(0) &= I(x_0 + h) - I(x_0) \\
&= \int_0^T [P - a(x_0(t) + h(t)) - b(x_0'(t) + h'(t))](x_0'(t) + h'(t))dt \\
&\quad - \int_0^T [P - ax_0(t) - bx_0(t)]x_0'(t)dt \\
&= \int_0^T [P - ax_0(t) - ah(t) - bx_0'(t) - bh'(t)]x_0'(t) \\
&\quad + \{[P - ax_0(t) - ah(t) - bx_0'(t) - bh'(t)]h'(t)\}dt \\
&\quad - \int_0^T [P - ax_0(t) - bx_0'(t)]x_0'(t)dt \\
&= \int_0^T [-ah(t)x_0'(t) - bh'(t)x_0'(t) + Ph'(t) - ax_0(t)h'(t) \\
&\quad - ah(t)h'(t) - bx_0'(t)h'(t) - bh'(t)h'(t)]dt \\
&= \int_0^T [P - ax_0(t) - 2bx_0'(t)]h'(t) - ax_0'(t)h(t)]dt \\
&\quad + \int_0^T [-ah(t)h'(t) - bh'(t)h'(t)]dt.
\end{aligned}$$

Como a aplicação

$$h \mapsto \int_0^T [(P - ax_0(t) - 2bx_0'(t))h'(t) - ax_0'(t)h(t)]dt$$

é funcional de X para $\mathbb{R}$ e como

$$\begin{aligned}\left|\int_0^T -ah(t)h'(t) - bh'(t)h'(t)dt\right| &\leqslant \int_0^T |ah(t) + bh'(t)|.|h'(t)|dt \\ &\leqslant \int_0^T (a+b)\|h\|.\|h\|dt \\ &= (a+b)(T-0)\|h\|^2 \\ &= T(a+b)\|h\|^2,\end{aligned}$$

segue que

$$[(D\tilde{I})(0)](h) = \int_0^T (P - ax_0(t) - 2bx_0'(t))h'(t) - ax_0'(t)h(t)dt = \int_0^T (P - 2bx_0'(t))h'(t)dt,$$

onde a ultima igualdade segue utilizando integral por partes:

$$\int_0^T ax_0'(t)h(t)dt = -\int_0^T ax_0(t)h'(t) + ax_0(t)h(t)|_{t=0}^T = -\int_0^T ax_0(t)h'(t)dt$$

Passo 3: Tendo $(D\tilde{I})(0) = 0$, segue que

$$\int_0^T (P - ax_0(t) - 2bx_0'(t) - a\int_0^T x_0'(\tau)d\tau)h'(t)dt = 0$$

para todo $h \in C^1[0,T]$ com $h(0) = h(T) = 0$. Provamos agora a seguinte afirmação:

Se $k \in C[a,b]$ e

$$\int_a^b k(t)h'(t)dt = 0$$

para todo $h \in C^1[a,b]$ com $h(a) = h(b) = 0$ então existe uma constante c tal que $k(t) = c$ para todo $t \in [a,b]$.

Para provar isto, defina constante c e a função h por

$$\int_a^b (k(t) - c)dt = 0 \quad \text{e} \quad h(t) = \int_a^t (k(\tau) - c)d\tau.$$

Então $h \in C^1[a,b]$ e isto satisfaz $h(a) = h(b) = 0$. Além disso,

$$\int_a^b (k(t) - c)^2 dt = \int_a^b (k(t) - c)h'(t)dt = \int_a^b k(t)h'(t)dt - c(h(b) - h(a)) = 0$$

Portanto, $k(t) - c = 0$ para todo $t \in [a,b]$.

Passo 4: O resultado anterior implica, no nosso caso

$$\forall t \in [0,T], \quad P - 2bx_0'(t) = c.$$

Integrando, obtemos $x_0(t) = At + B$, $t \in [0,T]$ para algumas constantes A e B. Usando $x_0(0) = 0$ e $x_0(T) = Q$, temos

$$x_0(t) = \frac{t}{T}Q, \quad t \in [0,T].$$

Passo 5: Finalmente, mostraremos que $I(x_0) \geqslant I(x)$ para todo x tal que $x(0) = 0$ e $x(T) = Q$. Provamos isto mostrando que $-I$ é convexa, e do Teorema 4.4, $-\tilde{I}$ na verdade possui mínimo global em 0.

Sejam $h_1, h_2 \in X$, $\alpha \in [0,1]$ e defina $x_1 = x_0 + h_1$, $x_2 = x_0 + h_2$. Então temos

$$\int_0^T (\alpha x_1'(t) + (1-\alpha)x_2'(t))^2 dt \leqslant \int_0^T \alpha(x_1'(t))^2 + (1-\alpha)(x_2'(t))^2 dt,$$

usando a convexidade de $y \mapsto y^2$. Além disso, $x_1(0) = 0 = x_2(0)$ e $x_1(T) = Q = x_2(T)$, e então

$$\begin{aligned}
&\int_0^T (\alpha x_1'(t) + (1-\alpha)x_2'(t))(\alpha x_1(t) + (1-\alpha)x_2(t))dt \\
&\qquad = \frac{1}{2}\int_0^T \frac{d}{dt}(\alpha x_1'(t) + (1-\alpha)x_2'(t))^2 dt = \frac{1}{2}Q^2 = \alpha\frac{1}{2}Q^2 + (1-\alpha)\frac{1}{2}Q^2 \\
&\qquad\qquad = \alpha\int_0^T x_1'(t)x_1(t)dt + (1-\alpha)\int_0^T x_2'(t)x_2(t)dt.
\end{aligned}$$

Consequentemente,

$$\begin{aligned}
-\tilde{I}(\alpha h_1 + (1-\alpha)h_2) &= -I(x_0 + \alpha h_1 + (1-\alpha)h_2) \\
&= -I(\alpha x_0 + (1-\alpha)x_0 + \alpha h_1 + (1-\alpha)h_2) \\
&= -I(\alpha x_1 + (1-\alpha)x_2) \\
&= b\int_0^T (\alpha x_1'(t) + (1-\alpha)x_2'(t))^2 dt \\
&\quad + a\int_0^T (\alpha x_1'(t) + (1-\alpha)x_2'(t))(\alpha x_1(t) + (1-\alpha)x_2(t))dt \\
&\quad - P\int_0^T (\alpha x_1'(t) + (1-\alpha)x_2'(t))dt \\
&\leqslant \alpha\int_0^T (x_1'(t))^2 dt + (1-\alpha)\int_0^T (x_2'(t))^2 dt \\
&\quad + \alpha\int_0^T x_1'(t)x_1(t)dt + (1-\alpha)\int_0^T x_2'(t)x_2(t)dt \\
&\quad - \alpha P\int_0^T x_1'(t)dt - (1-\alpha)P\int_0^T x_2'(t)dt \\
&= \alpha(\int_0^T x_1'(t)(bx_1'(t) + ax_1(t) - P)dt) \\
&\quad + (1-\alpha)(\int_0^T x_2'(t)(bx_2'(t) + ax_2(t) - P)dt) \\
&= \alpha(-I(x_1)) + (1-\alpha)(-I(x_2)) \\
&= \alpha(-\tilde{I}(h_1)) + (1-\alpha)(-\tilde{I}(h_2))
\end{aligned}$$

Consequentemente, $-\tilde{I}$ é convexa. Assim, obtemos o resultado esperado.

Agradecimentos: Agradeço aos meus pais, Milton e Amélia, pois sem eles não teria chegado onde estou agora. Agradeço à minha orientadora Profa. Dra. Marta Cilene Gadotti pelos incentivos, paciência e realização do projeto e ao grupo PET Matemática da UNESP de Rio Claro pelos projetos desenvolvidos e o incentivo à pesquisa. Aos meus amigos, dentro e fora da UNESP, que sempre me apoiaram nos momentos difíceis e me impediram de desistir.

Abstract: This work is part of a scientific initiation in which we study the theory of perators and the applications. In this specific publication, we illustrate a part of the theory involving differentiability and optimization of functions defined in normed space.

Keywords: Derivatives; Optimization; Critical Point; Infinite Dimensional Space.

Referências Bibliográficas

[1] SASANE, Amol. **Functional analysis and its applications**. [S. l.]: London School of Economics, 2013. 92 p. Disponível em: http://personal.lse.ac.uk/sasane/ma412.pdf. Acesso em: 1 set. 2020.

[2] KREYSZIG, Erwin. Introductory functional analysis with applications. 15. ed. [S. l.]: New York, N.Y. : Wiley, 1978., 1978. 688 p. ISBN 0-471-50731-8.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 16–22

Um Breve Estudo Sobre a Derivada Fracionária de Riemann-Liouville

Daniel de Lima Pazim†
Orientador(a): Profa. Dra. Marta Cilene Gadotti

Resumo: Neste trabalho introduzimos o conceito de derivada fracionária de Riemann-Liouville e apresentamos algumas propriedades. Além disso, calculamos a derivada de ordem fracionária p da função $f(t) = (t-a)^{\nu}$.

Palavras-chave: função gama; integração; diferenciabilidade; derivada fracionária.

1 Derivada Fracionária de Riemann-Liouville

Em Matemática é muito comum a investigação de certos conceitos no sentido de generalizá-los e muitas vezes, para fazer isso, é preciso um novo contexto ou uma nova abordagem. Por exemplo a extensão do conceito de função fatorial, sabemos que se n é um inteiro positivo então $f(n) = (n-1)! = 1 \cdot 2 \cdot 3 \cdots (n-1)$, mas se x é um número real positivo como definir $f(x)$? Neste caso, a resposta é dada pela função gama:

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1}\, dt,$$

pois é fácil verificar que $\Gamma(z+1) = z\Gamma(z)$ de onde concluímos que $\Gamma(n+1) = n!$, $n \in \mathbb{N}$.

Neste trabalho iremos estender o conceito de derivada, com a possibilidade de calcularmos derivadas fracionárias de uma função e iremos também mostrar algumas propriedades. Não existe uma única abordagem para fazer isso, aqui nós optamos pela definição da derivada fracionária de Riemann-Liouville.

Definição 1.1. Seja f uma função contínua e m vezes diferenciável, com $m \in \mathbb{N}$. Para $m \leq p < m+1$, definimos a derivada fracionária de Riemann-Liouville de ordem p por

$$_aD_t^p f(t) = \left(\frac{d}{dt}\right)^{m+1} \int_a^t (t-\tau)^{m-p} f(\tau) d\tau. \tag{1.1}$$

Esta é a mais abrangente definição de derivada fracionária conhecida.

Mostraremos que quando a ordem p da derivada é um inteiro, retornamos ao caso da derivada clássica de ordem inteira p.

Observação 1.2. Para $0 < p < 1$, a derivada fracionária de Riemann-Liouville pode ser aplicada em funções não diferenciáveis.

Um fato a notar é que assumindo a função $m+1$ vezes diferenciável, podemos obter a definição de derivada fracionária segundo Grünwald-Letnikov após a realização de repetidas

†bolsista FAPESP

integrações por partes. Assim, podemos escrever

$$
\begin{aligned}
{}_aD_t^p f(t) &= \left(\frac{d}{dt}\right)^{m+1} \int_a^t (t-\tau)^{m-p} f(\tau)d\tau \\
&= \sum_{k=0}^{m} \frac{f^{(k)}(a)(t-a)^{-p+k}}{\Gamma(-p+k+1)} + \frac{1}{\Gamma(-p+m+1)} \int_a^t (t-\tau)^{m-p} f^{(m+1)}(\tau)d\tau \\
&\quad (m \le p < m+1),
\end{aligned}
$$

onde esta última expressão é a definição da derivada fracionária de Grünwald-Letnikov, veja [1].

Portanto, para as classes de funções $f(t)$ que são contínuas e $m+1$ vezes deriváveis, as definições de derivada fracionária segundo Grünwald-Letnikov e Riemann-Liouville, coincidem.

Do ponto de vista da matemática pura, tal classe de funções é limitada. No entanto, essa é uma classe de funções muito importante para as aplicações, uma vez que a maioria dos processos dinâmicos são representados por funções suaves o suficiente. Este fato é muito importante para o uso adequado da derivada fracionária, pois a definição segundo Riemann-Liouville é uma excelente oportunidade de enfraquecer as condições sob a função f. Em outras palavras, é suficiente exigir a integrabilidade de f; com isso a integral (1.1) existe para $t > a$ e pode ser $m+1$ vezes diferenciável.

1.1 Unificação da Derivada e Integral de ordem inteira

Vamos analisar agora como a definição (1.1) é vista como a unificação das noções de integração e diferenciação de ordem inteira.

Seja f uma função contínua e integrável em todo intervalo limitado (a, t). Assim, a função f pode ter uma singularidade de ordem $r < 1$ em $\tau = a$, isto é,

$$\lim_{\tau \to a} (\tau - a)^r f(\tau) = cte \neq 0.$$

Dessa forma, a integral

$$\int_a^t f(\tau)d\tau = f^{(-1)}(t) \tag{1.2}$$

existe e tem um valor finito, igual a zero quando $t \to a$. De fato, fazendo a substituição $\tau = a + y(t-a)$, com isso $d\tau = (t-a)dy$, e denotando $\epsilon = t - a$, o que faz $\epsilon \to 0$ quando $t \to a$, então

$$
\begin{aligned}
\lim_{t \to a} f^{(-1)}(t) &= \lim_{t \to a} \int_a^t f(\tau)d\tau = \lim_{t \to a} (t-a) \int_0^1 f(a + y(t-a))dy \\
&= \lim_{\epsilon \to 0} \epsilon^{1-r} \int_0^1 (\epsilon y)^r f(a + y\epsilon) y^{-r} dy = 0,
\end{aligned}
$$

pois $r < 1$. Logo, podemos integrar (1.2) e obter $f^{(-2)}(t)$, a saber

$$f^{(-2)}(t) = \int_a^t d\tau_1 \int_a^{\tau_1} f(\tau)d\tau = \int_a^t f(\tau)d\tau \int_\tau^t d\tau_1 = \int_a^t (t-\tau) f(\tau)d\tau. \tag{1.3}$$

Novamente, vemos que integrando (1.3), obtemos

$$
\begin{aligned}
f^{(-3)}(t) &= \int_a^t d\tau_1 \int_a^{\tau_1} d\tau_2 \int_a^{\tau_2} f(\tau_3)d\tau_3 = \int_a^t d\tau_1 \int_a^{\tau_1} (\tau_1 - \tau) f(\tau)d\tau \\
&= \frac{1}{2} \int_a^t (t-\tau)^2 f(\tau)d\tau.
\end{aligned}
$$

Por indução, obtemos o caso geral, chamado de fórmula de Cauchy

$$f^{(-n)}(t) = \frac{1}{\Gamma(n)} \int_a^t (t-\tau)^{n-1} f(\tau) d\tau.$$

Sendo D^{-k} a notação para k integrações iteradas, supondo $n \geq 1$ fixo e $k \in \mathbb{N}^*$, então

$$f^{(-k-n)}(t) = \frac{1}{\Gamma(n)} D^{-k} \int_a^t (t-\tau)^{n-1} f(\tau) d\tau. \tag{1.4}$$

Por outro lado, para $n \geq 1$ e $k \geq n$, $k \in \mathbb{Z}$, a $(k-n)$-ésima derivada de $f(t)$ pode ser escrita como

$$f^{(k-n)}(t) = \frac{1}{\Gamma(n)} D^{k} \int_a^t (t-\tau)^{n-1} f(\tau) d\tau, \tag{1.5}$$

em que D^k representa k diferenciações iteradas.

Deste modo, as fórmulas (1.4) e (1.5) podem ser consideradas como caso particular uma da outra. Por conveniência, escolhemos (1.4) como caso particular de (1.5). O símbolo D^k significa k integrações se $k < 0$ ou k derivações se $k > 0$.

Vale a seguinte regra de composição para integrais:

$${}_aD_t^{-p}\left({}_aD_t^{-q} f(t)\right) = {}_a D_t^{-q}\left({}_aD_t^{-p} f(t)\right) = {}_a D_t^{-p-q} f(t), \qquad p, q > 0. \tag{1.6}$$

A prova deste fato pode ser encontrada em [1].

1.2 Derivadas de Ordem Arbitrária

Uma opção para pensar em uma derivada de ordem arbitrária é a partir da expressão (1.5), onde temos a derivada de ordem $(k-n)$ inteira. Para isso, vamos substituir o valor inteiro n por um valor real α de modo que $k - \alpha > 0$. Assim,

$${}_aD_t^{k-\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \frac{d^k}{dt^k} \int_a^t (t-\tau)^{\alpha-1} f(\tau) d\tau, \qquad 0 < \alpha \leq 1. \tag{1.7}$$

A única restrição para α é $\alpha > 0$. Porém, é possível, sem perda de generalidade, estreitar essa condição para $0 < \alpha \leq 1$, como foi considerado.

Note que a expressão (1.7) pode ser reescrita pondo $p = k - \alpha$, assim:

$${}_aD_t^{p} f(t) = \frac{1}{\Gamma(k-p)} \frac{d^k}{dt^k} \int_a^t (t-\tau)^{k-p-1} f(\tau) d\tau, \qquad k-1 < p \leq k, \tag{1.8}$$

equivalentemente

$${}_aD_t^{p} f(t) = \frac{d^k}{dt^k}\left({}_aD_t^{-(k-p)} f(t)\right), \qquad k-1 < p \leq k.$$

Para $p = k-1$, voltamos ao caso da derivada convencional de ordem inteira $k-1$, isto é,

$$\begin{aligned} {}_aD_t^{k-1} f(t) &= \frac{d^k}{dt^k}\left({}_aD_t^{-(k-(k-1))} f(t)\right) \\ &= \frac{d^k}{dt^k}\left({}_aD_t^{-1} f(t)\right) = f^{(k-1)}(t). \end{aligned}$$

Mais ainda, usando $D^0 f = f$, temos para $p = k \geq 1$ e $t > a$,

$${}_aD_t^{p} f(t) = \frac{d^k}{dt^k}\left({}_aD_t^{0} f(t)\right) = \frac{d^k f(t)}{dt^k} = f^{(k)}(t).$$

O que significa que para $t > a$ a derivada fracionária de Riemann-Liouville para $p = k$ coincide com a derivada convencional de ordem k.

Teorema 1.3. *Sob as condições da f dadas anteriormente, tem-se*

i) Para $p > 0$ e $t > a$,

$$ {}_aD_t^p \left({}_aD_t^{-p} f(t) \right) = f(t). \tag{1.9} $$

ii) Para uma função cuja derivada fracionária é integrável,

$$ {}_aD_t^{-p} \left({}_aD_t^p f(t) \right) = f(t) - \sum_{j=1}^{k} \left[{}_aD_t^{p-j} f(t) \right]_{t=a} \frac{(t-a)^{p-j}}{\Gamma(p-j+1)}. \tag{1.10} $$

Ou seja, a operação de derivação fracionária é uma inversa à esquerda da operação de integração fracionária de mesma ordem p e assim como para ordem inteira, a derivação e integração de ordem fracionárias não comutam.

Prova: (i) Vamos considerar o caso de $p = n \geq 1$ inteiro:

$$ \begin{aligned} {}_aD_t^n \left({}_aD_t^{-n} f(t) \right) &= \frac{1}{\Gamma(n)} \frac{d^n}{dt^n} \int_a^t (t-\tau)^{n-1} f(\tau) d\tau \\ &= \frac{d}{dt} \int_a^t f(\tau) d\tau = f(t). \end{aligned} $$

Agora, para o caso $k - 1 \leq p < k$, com a ajuda da expressão (1.6), temos

$$ {}_aD_t^{-k} f(t) ={}_a D_t^{-(k-p)} \left({}_aD_t^{-p} f(t) \right), $$

e, portanto,

$$ \begin{aligned} {}_aD_t^p \left({}_aD_t^{-p} f(t) \right) &= \frac{d^k}{dt^k} \left\{ {}_aD_t^{-(k-p)} \left({}_aD_t^{-p} f(t) \right) \right\} \\ &= \frac{d^k}{dt^k} \left\{ {}_aD_t^{-k} f(t) \right\} = f(t). \end{aligned} $$

(ii) De fato, por um lado temos

$$ \begin{aligned} {}_aD_t^{-p} \left({}_aD_t^p f(t) \right) &= \frac{1}{\Gamma(p)} \int_a^t (t-\tau)^{p-1}_a D_\tau^p f(\tau) d\tau \\ &= \frac{d}{dt} \left\{ \frac{1}{\Gamma(p+1)} \int_a^t (t-\tau)^p_a D_\tau^p f(\tau) d\tau \right\}. \end{aligned} $$

Por outro lado, como k é inteiro, podemos reescrever o termo que está sendo derivado da seguinte forma:

$$ \frac{1}{\Gamma(p+1)} \int_a^t (t-\tau)^p_a D_\tau^p f(\tau) d\tau = \frac{1}{\Gamma(p+1)} \int_a^t (t-\tau)^p \frac{d}{d^k} \left\{ {}_aD_t^{-k} \left({}_aD_\tau^p f(\tau) \right) \right\} d\tau, $$

usando (1.6) e fazendo repetidas vezes o processo de integração por partes, segue que

(espaço em branco intencional)

$$
\begin{aligned}
&\frac{1}{\Gamma(p+1)}\int_a^t (t-\tau)^p \frac{d}{d^k}\left\{{}_aD_\tau^{-(k-p)}f(\tau)\right\}d\tau\\
&= \frac{1}{\Gamma(p-k+1)}\int_a^t (t-\tau)^{p-k}\left\{{}_aD_\tau^{-(k-p)}f(\tau)\right\}d\tau\\
&\quad -\sum_{j=1}^{k}\left[\frac{d^{k-j}}{dt^{k-j}}\left({}_aD_t^{-(k-p)}f(t)\right)\right]_{t=a}\frac{(t-a)^{p-j+1}}{\Gamma(2+p-j)}\\
&= \frac{1}{\Gamma(p-k+1)}\int_a^t (t-\tau)^{p-k}\left\{{}_aD_\tau^{-(k-p)}f(\tau)\right\}d\tau-\sum_{j=1}^{k}\left[{}_aD_t^{p-j}f(t)\right]_{t=a}\frac{(t-a)^{p-j+1}}{\Gamma(2+p-j)} \quad (1.11)\\
&={}_a D_t^{-(p-k+1)}\left({}_aD_t^{-(k-p)}f(t)\right)-\sum_{j=1}^{k}\left[{}_aD_t^{p-j}f(t)\right]_{t=a}\frac{(t-a)^{p-j+1}}{\Gamma(2+p-j)}\\
&={}_a D_t^{-1}f(t)-\sum_{j=1}^{k}\left[{}_aD_t^{p-j}f(t)\right]_{t=a}\frac{(t-a)^{p-j+1}}{\Gamma(2+p-j)}. \quad (1.12)
\end{aligned}
$$

Garantimos a existência de todos os termos em (1.11) pelo fato de ${}_aD_t^pf(t)$ ser integrável, pois assim as derivadas fracionárias ${}_aD_t^{p-j}f(t)$, $(j=1,\ 2,\ 3,\ \dots,\ k)$, são todas limitadas em $t=a$.

Derivando (1.12) em relação a t, segue o resultado de (1.10). ■

Um caso particular importante é para $0<p<1$, isto é $j=1$, e assim

$$
{}_aD_t^{-p}\left({}_aD_t^pf(t)\right)=f(t)-\left[{}_aD_t^{p-1}f(t)\right]_{t=a}\frac{(t-a)^{p-1}}{\Gamma(p)}.
$$

A propriedade (1.9) é um caso particular de uma mais abrangente, a saber:

Corolário 1.4. *Sob as condições do item* $i)$ *do Teorema* (1.3)*, temos*

$$
{}_aD_t^p\left({}_aD_t^{-q}f(t)\right)={}_a D_t^{p-q}f(t), \quad (1.13)
$$

onde f *é contínua e, se* $p\geq q\geq 0$ *a derivada* ${}_aD_t^{p-q}f(t)$ *existe.*

Prova: Dois casos devem ser considerados: $q\geq p\geq 0$ e $p>q\geq 0$.

Se $q\geq p\geq 0$, usando (1.6) e (1.9), obtemos

$$
\begin{aligned}
{}_aD_t^p\left({}_aD_t^{-q}f(t)\right)&={}_a D_t^p\left({}_aD_t^{-p}\ {}_aD_t^{-(q-p)}f(t)\right)\\
&={}_a D_t^{-(q-p)}f(t)={}_a D_t^{p-q}f(t).
\end{aligned}
$$

Agora para $p>q\geq 0$, sejam m,n naturais tais que $0\leq m-1\leq p<m$ e $0\leq n\leq p-q<m$. Usando (1.8) e (1.6) obtemos

$$
\begin{aligned}
{}_aD_t^p\left({}_aD_t^{-q}f(t)\right)&=\frac{d^m}{dt^m}\left\{{}_aD_t^{-(m-p)}\left({}_aD_t^{-q}f(t)\right)\right\}\\
&=\frac{d^m}{dt^m}\left\{{}_aD_t^{p-q-m}f(t)\right\}\\
&=\frac{d^n}{dt^n}\left\{{}_aD_t^{p-q-n}f(t)\right\}={}_a D_t^{p-q}f(t).
\end{aligned}
$$
■

A propriedade (1.10) também é um caso particular de uma mais geral, a saber:

Corolário 1.5. *Sob as condições do item ii) do Teorema* (1.3)*, temos*

$$ {}_aD_t^{-p}\left({}_aD_t^{q}f(t)\right) = {}_aD^{p-q}f(t) - \sum_{j=1}^{k}\left[{}_aD_t^{q-j}f(t)\right]_{t=a}\frac{(t-a)^{p-j}}{\Gamma(1+p-j)}, \qquad (0 \le k-1 \le q < k). $$

Prova: Usemos (1.6), se $q \le p$ ou (1.13), se $q \ge p$ e também a propriedade (1.10). Temos:

$$ \begin{aligned} {}_aD_t^{-p}\left({}_aD_t^{q}f(t)\right) &= {}_aD_t^{q-p}\left\{{}_aD_t^{-q}\left({}_aD_t^{q}f(t)\right)\right\} \\ &= {}_aD_t^{q-p}\left\{f(t) - \sum_{j=1}^{k}\left[{}_aD_t^{q-j}f(t)\right]_{t=a}\frac{(t-a)^{q-j}}{\Gamma(p-j+1)}\right\} \\ &= {}_aD_t^{q-p}f(t) - \sum_{j=1}^{k}\left[{}_aD_t^{q-j}f(t)\right]_{t=a}\frac{(t-a)^{p-j}}{\Gamma(1+p-j)}, \end{aligned} $$

aqui usamos a derivada da função potência:

$$ {}_aD_t^{q-p}\left\{\frac{(t-a)^{q-j}}{\Gamma(1+q-j)}\right\} = \frac{(t-a)^{p-j}}{\Gamma(1+p-j)}, $$

a qual veremos na seção seguinte. ■

2 Alguns Exemplos da Derivada Fracionária de Riemann-Liouville

Vejamos agora alguns cálculos de derivadas fracionárias.

Exemplo 2.1. Calculemos ${}_aD_t^{p}f(t)$ para a expressão que tem a forma

$$ f(t) = (t-a)^{\nu}, $$

onde ν é um número real.

De (1.8), temos

$$ {}_aD_t^{p}(t-a)^{\beta} = \frac{1}{\Gamma(k-p)}\frac{d^k}{dt^k}\int_a^t (t-\tau)^{k-p-1}(\tau-a)^{\beta}d\tau. $$

Fazendo a mudança $s = (\tau-a)(t-a)^{-1}$, obtemos

$$ \begin{aligned} {}_aD_t^{p}(t-a)^{\beta} &= \frac{1}{\Gamma(k-p)}\frac{d^k}{dt^k}(t-a)^{\beta-p+k}\int_0^1 s^{\beta}(1-s)^{k-p-1}ds \\ &= \frac{\Gamma(\beta+1)}{\Gamma(k-p+\beta+1)}\frac{d^k}{dt^k}(t-a)^{\beta-p+k}, \end{aligned} \tag{2.1} $$

onde identificamos $\int_0^1 s^{\beta}(1-s)^{k-p-1}ds = B(k-p,\beta+1)$ com a função Beta de Euler, definida por

$$ B(z,w) = \int_0^1 \tau^{z-1}(1-\tau)^{w-1}d\tau, \qquad Re(z) > 0 \quad \text{e} \quad Re(w) > 0. $$

Dois casos são possíveis para a derivada ${}_aD_t^{p}(t-a)^{\beta}$. O primeiro é se $p-\beta \in \mathbb{N}$, pois assim estaríamos fazendo a derivada de ordem inteira n de um polinômio de grau $n-(p-\beta) < n$. O segundo, é quando $p-\beta \notin \mathbb{N}$ assim, de (2.1), sabendo que $\Gamma(z+1) = z\Gamma(z)$, temos

$$ {}_aD_t^{p}(t-a)^{\beta} = \frac{\Gamma(\beta+1)}{\Gamma(\beta-p+1)}(t-a)^{\beta-p}. \tag{2.2} $$

Exemplo 2.2. A derivada fracionária com ordem $p \notin \mathbb{N}$ de uma constante não é zero. Por exemplo, associando $1 = (t-a)^0$ e usando (2.2) temos

$$ {}_aD_t^p 1 = {}_a D_t^p (t-a)^0 = \frac{\Gamma(1)}{\Gamma(-p+1)}(t-a)^{-p} = \frac{(t-a)^{-p}}{\Gamma(1-p)}. $$

Agradecimentos: Agradeço a Deus por minha vida, família e amigos, a minha orientadora por sempre me motivar e ajudar e a FAPESP pela ajuda financeira em minha pesquisa.

Abstract: In this work we introduce the Riemann-Liouville fractional derivative concept and present some properties. In addition, we calculate the fractional derivative of order p of the function $f(t) = (t-a)^\nu$.

Keywords: gamma function; integration; differentiability; fractional derivative.

Referências Bibliográficas

[1] PODLUBNY, I. *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications.* [S.I.]: Elsevier, 1998.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 23–27

Espaço Afim e Conjunto Algébrico

Gabriela Prampolim
Orientador(a): Profa. Dra. Elíris Cristina Rizziolli

Resumo: Este trabalho tem como finalidade introduzir noções clássicas da geometria algébrica. Nesse sentido, abordaremos o espaço afim, conjunto algébrico e resultados relevantes a estes tópicos.

Palavras-chave: geometria algébrica; espaço afim; conjunto algébrico.

1 Espaço Afim

Definição 1.1. Seja k um corpo. Definimos $A^n(k)$ (ou A^n, quando k é subentendido), como o produto cartesiano de k com ele mesmo n vezes. Ou seja, $A^n(k)$ é o conjunto das n-uplas dos elementos de k. Chamamos $A^n(k)$ de **n-espaço afim sobre** k e seus elementos são chamados de *pontos*. Quando temos $A^1(k)$ chamamos de reta afim e $A^2(k)$ de plano afim.

Definição 1.2. Sejam $F \in k[X_1, \ldots, X_n]$ um polinômio e $P = (a_1, \ldots, a_n) \in A^n(k)$ um ponto. Dizemos que P é **zero de** F se $F(P) = F(a_1, \ldots, a_n) = 0$.

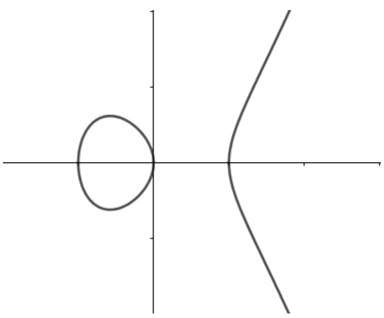
Definição 1.3. Se F não é um polinômio constante, o conjunto de zeros de F é chamado de **hipersuperfície** definido por F e é denotado por $V(F)$. A hipersuperfície de $A^2(k)$ é chamada *curva do plano afim*. Se F é um polinômio de grau 1, $V(F)$ é chamado de *hiperplano* de $A^n(k)$. Se $n = 2$, então é uma reta.

Genericamente, se tivermos S um conjunto qualquer de polinômios em $k[X_1, \ldots, X_n]$, então podemos definir $V(S) = \{P \in A^n(k);\ F(P) = 0,\ \forall F \in S\}$. Deste modo,

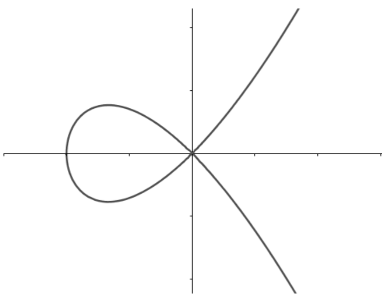
$$V(S) = \bigcap_{F \in S} V(F).$$

Vale ressaltar que, geralmente, utilizamos $V(F_1, \ldots, F_r)$ ao invés de $V(\{F_1, \ldots, F_r\})$, como um abuso de notação.

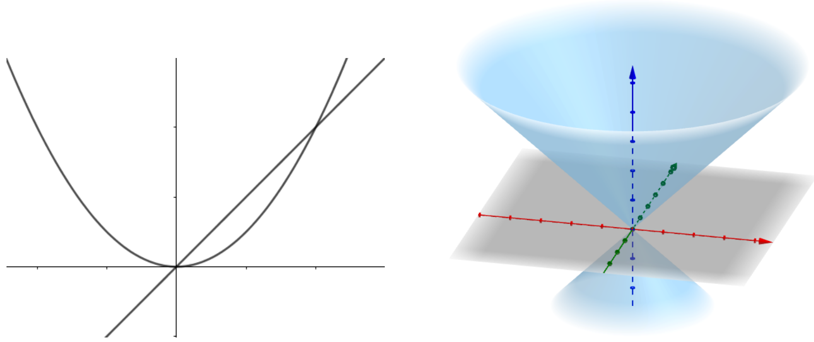
Exemplo 1.4. Seja $k = \mathbb{R}$.



(a) $V(Y^2 - X(X^2 - 1)) \subset A^2$



(b) $V(Y^2 - X^2(X + 1)) \subset A^2$



(c) $V(Y^2 - XY - X^2Y + X^3) \subset A^2$ (d) $V(Z^2 - (X^2 + Y^2) \subset A^3$

Figura 3.1: Figuras do exemplo 1.4

Definição 1.5. Um subconjunto $X \subset A^n(k)$ é um **conjunto algébrico** (ou, de maneira mais completa, conjunto algébrico afim) se $X = V(S)$, para algum $S \subset k[X_1, \ldots, X_n]$.

Proposição 1.6. *Se I é um ideal em $k[X_1, \ldots, X_n]$ gerado por S, então $V(S) = V(I)$, ou seja, todo conjunto algébrico é igual a $V(I)$, para algum ideal I.*

Prova: Queremos mostrar que $V(S) = V(I)$ e para tal precisamos mostrar que: (i) $V(S) \subset V(I)$; (ii) $V(I) \subset V(S)$.

Sejam $S = \{F_1, \ldots, F_r\}$, $F_i \in k[X_1, \ldots, X_n]$ e $I = \langle S \rangle = \langle F_1, \ldots, F_r \rangle$. Lembremos que se $F \in I$, então $F = \sum_{i=1}^r c_i F_i$.

(i) Seja $P \in V(I)$ qualquer, então $F(P) = 0, \forall F \in I = \langle F_1, \ldots, F_r \rangle$. Em particular, $F_1, \ldots, F_r \in I$. Desta forma, $F_i(P) = 0, \forall i \in \{1, \ldots, r\} \Rightarrow P \in \bigcap_{i=1}^r V(F_1) = V(S)$, tal igualdade se dá pela definição de $V(S)$.

(ii) Seja $P \in V(S)$ qualquer. Isso significa que $F_1(P) = 0, \forall i \in \{1, \ldots, r\}$. Para mostrar que $P \in V(I)$ tomemos $F \in I$ qualquer. Assim, como $F \in I$, temos que $F = \sum_{i=1}^r c_i F_i$, logo: $F(P) = \sum_{i=1}^r c_i F_i(P)$. Observe que $F_i(P) = 0, \forall i \in \{1, \ldots, r\}$, já que $P \in V(S)$, logo $\sum_{i=1}^r c_i F_i(P) = 0$. Ou seja, $F(P) = 0$, desta forma $P \in V(I)$. ■

Proposição 1.7. *Se I_α é uma coleção de ideais, então $V\left(\bigcup_\alpha I_\alpha\right) = \bigcap_\alpha V(I_\alpha)$, ou seja, a intersecção de qualquer coleção de conjuntos algébricos é um conjunto algébrico.*

Prova: Queremos mostrar que $V\left(\bigcup_\alpha I_\alpha\right) = \bigcap_\alpha V(I_\alpha)$, para tal mostremos que: (i) $V\left(\bigcup_\alpha I_\alpha\right) \subset \bigcap_\alpha V(I_\alpha)$; (ii) $\bigcap_\alpha V(I_\alpha) \subset V\left(\bigcup_\alpha I_\alpha\right)$.

(i) Seja $P \in V\left(\bigcup_\alpha I_\alpha\right)$ qualquer. Veja que como $P \in V\left(\bigcup_\alpha I_\alpha\right)$, segue que $F(P) = 0, \forall F \in \bigcup_\alpha I_\alpha$. Em particular, $I_\alpha \in V\left(\bigcup_\alpha I_\alpha\right), \forall \alpha$. Daí, se $F \in I_\alpha, \forall \alpha$, temos que $F(P) = 0$, logo $P \in V(I_\alpha), \forall \alpha$, ou seja, $P \in \bigcap_\alpha V(I_\alpha)$.

(ii) Seja $P \in \bigcap_\alpha V(I_\alpha)$ qualquer. Assim, se $P \in \bigcap_\alpha V(I_\alpha)$, então $P \in V(I_\alpha), \forall \alpha$, logo

$$F(P) = 0, \forall F \in I_\alpha, \forall \alpha. \tag{1.1}$$

Agora, seja $F \in \bigcup_\alpha I_\alpha$ qualquer. Logo, $F \in I_\alpha$, para algum $\alpha \in I$. Por (1.1) temos que $F(P) = 0$ e, assim, $P \in V\left(\bigcup_\alpha I_\alpha\right)$. ■

Proposição 1.8. *Se $I \subset J$, onde I e J são ideais, então $V(I) \supset V(J)$.*

Prova: Seja $P \in V(J)$ qualquer, logo $F(P) = 0, \forall F \in J$. Como $I \subset J$, segue que $F(P) = 0$, $\forall F \in I$. Logo, $P \in V(I)$. ■

Proposição 1.9. *As seguintes propriedades são válidas:*

(a) $V(F \cdot G) = V(F) \cup V(G)$*, para quaisquer* F, G *polinômios.*

(b) $V(I) \cup V(J) = V(\{F \cdot G;\ F \in I, G \in J\})$.

Ou seja, a união finita de conjuntos algébricos é um conjunto algébrico.

Prova: **(a)** Queremos mostrar que $V(F \cdot G) = V(F) \cup V(G)$, para isso, precisamos que: (i) $V(F \cdot G) \subset V(F) \cup V(G)$; (ii) $V(F) \cup V(G) \subset V(F \cdot G)$.

(i) Seja $P \in V(F \cdot G)$, qualquer. Desta forma, temos que:

$$(F \cdot G)(P) = 0 \Rightarrow F(P) \cdot G(P) = 0 \Rightarrow F(P) = 0 \text{ ou } G(P) = 0.$$

Lembremos que essa última passagem pode ser realizada pois k é corpo e, portanto, domínio de integridade. Como $F(P) = 0$ ou $G(P) = 0 \Rightarrow P \in V(F)$ ou $P \in V(G)$. Portanto, $P \in V(F) \cup V(G)$.

(ii) Seja $P \in V(F) \cup V(G)$, qualquer. Assim, temos que:

$$P \in V(F) \text{ ou } P \in V(G) \Rightarrow F(P) = 0 \text{ ou } G(P) = 0 \Rightarrow$$
$$F(P) \cdot G(P) = 0 \Rightarrow (F \cdot G)(P) = 0.$$

Logo, $P \in V(F \cdot G)$.

(b) Mostremos que $V(I) \cup V(J) = V(\{F \cdot G;\ F \in I, G \in J\}$, para isso, precisamos mostrar que: (i) $V(I) \cup V(J) \subset V(\{F \cdot G;\ F \in I, G \in J\}$; (ii) $V(\{F \cdot G;\ F \in I, G \in J\}) \subset V(I) \cup V(J)$.

(i) Seja $P \in V(I) \cup V(J)$, qualquer. Disto, temos que:

$$P \in V(I) \text{ ou } P \in V(J) \Rightarrow F(P) = 0,\ \forall F \in I, \text{ ou } G(P) = 0,\ \forall G \in J \Rightarrow$$
$$F(P) \cdot F(G) = 0,\ \forall F \in I,\ \forall G \in J \Rightarrow (F \cdot G)(P) = 0,\ \forall F \in I,\ \forall G \in J.$$

Desta forma, $P \in V(\{F \cdot G;\ F \in I, G \in J\})$.

(ii) Seja $P \in V(\{F \cdot G;\ F \in I, G \in J\})$, qualquer. Assim, temos que:

$$(F \cdot G)(P) = 0,\ \forall F \in I,\ \forall G \in J \Rightarrow F(P) \cdot F(G) = 0,\ \forall F \in I,\ \forall G \in J \Rightarrow$$
$$F(P) = 0,\ \forall F \in I, \text{ ou } G(P) = 0,\ \forall G \in J.$$

Esse último passo pode ser realizado pois k é corpo e, portanto, domínio de integridade. Como $F(P) = 0,\ \forall F \in I$, ou $G(P) = 0,\ \forall G \in J$, então $P \in V(I)$ ou $P \in V(J)$. Com isso, $P \in V(I) \cup V(J)$. ■

Proposição 1.10. *São válidas as seguintes propriedades:*

(a) $V(0) = A^n(k)$*, onde* 0 *é o polinômio nulo.*

(b) $V(1) = \emptyset$*, em que* 1 *é o polinômio constante* 1.

(c) *$V(X_1 - a_1, \ldots, X_n - a_n) = \{(a_1, \ldots, a_n)\}$, para qualquer $a_i \in k$. Consequentemente, um subconjunto finito de $A^n(k)$ é um conjunto algébrico.*

Prova: (a) Queremos mostrar que $V(0) = A^n(k)$, onde 0 é o polinômio nulo. De fato, $\forall P \in A^n(k)$,

$$0(P) = 0_k \Rightarrow P \in V(0).$$

Assim, $A^n(k) \subset V(0)$. Além disso, $V(0) \subset A^n(K)$ por definição.

(b) Mostremos que $V(1) = \emptyset$. De fato: $\forall P \in A^n(k)$,

$$1(P) = 1_k \neq 0 \Rightarrow V(1) = \emptyset.$$

(c) Agora, mostremos que $V(X_1 - a_1, \ldots, X_n - a_n) = \{(a_1, \ldots, a_n)\}, \forall a_i \in k$. Para tal, precisamos mostrar que:

(i) $V(X_1 - a_1, \ldots, X_n - a_n) \subset \{(a_1, \ldots, a_n)\}$;

(ii) $\{(a_1, \ldots, a_n)\} \subset V(X_1 - a_1, \ldots, X_n - a_n)$.

Com efeito:

(i) Seja $P = (b_1, \ldots, b_n) \in V(X_1 - a_1, \ldots, X_n - a_n)$, então sabemos que $F_i(P) = 0$, em que $F_i = X_i - a_i$, $\forall i \in \{1, \ldots, n\}$. Porém, isso apenas ocorre se $b_i - a_i = 0, \forall i \in \{1, \ldots, n\} \Rightarrow b_i = a_i, \forall i \in \{1, \ldots, n\}$. Assim, $P = (a_1, \ldots, a_n) \in \{(a_1, \ldots, a_n)\}$.

(ii) Se $P = (a_1, \ldots, a_n) \in \{(a_1, \ldots, a_n)\}$, então $F_i(P) = 0$, em que $F_i = X_i - a_i$, $\forall i \in \{1, \ldots, n\}$. Logo, $P \in V(X_1 - a_1, \ldots, X_n - a_n)$.

Veja que esta última parte pode ser generalizada para um conjunto $X = \{P_1, \ldots, P_s\} \subset A^n(k)$, qualquer, onde cada $P_i = (a_{i_1}, \ldots, a_{i_n}), 1 \leq i \leq s$. Já que:

$$X = \{P_1, \ldots, P_s\} = \bigcup_{i=1}^{s} \{P_i\} = \bigcup_{i=1}^{s} V(\{X_1 - a_{i_1}, \ldots, X_n - a_{i_n}\}) = \bigcup_{i=1}^{s} V(I_i) = V(\{F_1 \cdot \ \ldots \ \cdot F_s;\ F_i \in I_i\}),$$

em que $I_i = \langle X_1 - a_{i_1}, \ldots, X_n - a_{i_n} \rangle$, com $1 \leq i \leq s$. ■

Exemplo 1.11. Mostremos que $Y = \{P = (t, t^2, t^3) \in A^3(k);\ t \in k\}$ é um conjunto algébrico. Para tal, precisamos achar algum conjunto de polinômios tais que os elementos de X sejam raiz desse polinômio.

Suponha $F = X_1^3 - 2X_2X_1 + X_3$. Veja que qualquer elemento de Y utilizado irá zerar o polinômio F, de fato:

$$F(P) = (t)^3 - 2(t^2)(t) + (t^3) = t^3 - 2t^3 + t^3 = 0 \Leftrightarrow F(P) = 0, \forall P \in Y.$$

Assim, $Y = V(F)$.

Exemplo 1.12. Mostremos que $X = \{P = (\cos(t), \operatorname{sen}(t)) \in A^2(\mathbb{R});\ t \in \mathbb{R}\}$ é um conjunto algébrico. Para tal, precisamos achar algum conjunto de polinômios tais que os elementos de X sejam raiz desse polinômio.

Suponha $G = X_1^2 + X_2^2 - 1$. Veja que qualquer elemento de X que for utilizado irá zerar o polinômio G. De fato:

$$G(P) = (\cos t)^2 + (\operatorname{sen} t)^2 - 1 = 0 \Leftrightarrow G(P) = 0, \forall P \in X.$$

Logo, $X = V(G)$.

Agradecimentos: Agradeço primeiramente a minha orientadora, Elíris, pela paciência, disposição e pelos conselhos, tanto pessoais quanto profissionais. Agradeço também a minha família que tem me apoiado imensamente durante minha trajetória, mesmo que eles não compreendam o que eu faço. Agradeço ao PET e aos seus integrantes pelas oportunidades únicas que me ofereceram. Por último, agradeço meus amigos pelas dicas, pela ajuda e por todo o apoio que recebi.

Abstract: This paper presents an introduction to basic notions on algebric geometry. In this purport, we will approach affine space, algebric set and relevant results to those topics.

Keywords: algebric geometry; affine space; algebraic set.

Referências Bibliográficas

[1] Fulton, W., *Algebraic Curves: an introdution to Algebraic Geometry*, Addison-Wesley, 1989.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 28–37

Conceitos da Teoria de Conjuntos Fuzzy e Aplicações

Gilberto Murilo Lopes da Silva†
Orientador(a): Profa. Dra. Renata Zotin Gomes de Oliveira

Resumo: Neste artigo, apresentaremos conceitos básicos da Teoria de Conjuntos Fuzzy e algumas aplicações

Palavras-chave: Conjuntos fuzzy; funções de pertinência; lógica fuzzy; operações com conjuntos fuzzy

1 Conjuntos clássicos e conjuntos fuzzy

A Teoria dos Conjuntos Fuzzy, recente do ponto de vista da história da matemática, vem se desenvolvendo e, cada vez mais, está sendo usada como ferramenta para formulação de modelos nos vários campos das ciências. A palavra "fuzzy", de origem inglesa, significa "incerto", "vago", "impreciso", "subjetivo", "nebuloso", "difuso" e a primeira publicação sobre conjuntos fuzzy é de Lotfi A. Zadeh, com o artigo *Fuzzy Sets*, em 1965. O desenvolvimento da teoria e suas aplicações vêm apresentando uma evolução muito rápida.

A principal intenção de Zadeh com a lógica fuzzy é dar um tratamento matemático a certos termos linguísticos subjetivos, como "aproximadamente", "em torno de", dentre outros. Esse seria um primeiro passo no sentido de se programar e armazenar conceitos vagos em computadores, tornando possível a produção de cálculos com informações imprecisas, a exemplo do que faz o ser humano.

Vamos então, de início, apresentar o conceito de conjunto fuzzy, modelador da incerteza, relacionando-o com conjuntos clássicos, chamados de conjuntos crisp. Para formalizar matematicamente um conjunto fuzzy, Zadeh baseou-se no fato de que qualquer conjunto crisp pode ser caracterizado por sua função característica, cuja definição é dada a seguir.

Definição 1.1. Seja U um conjunto e A um subconjunto de U. A função característica de A é dada por

$$\chi_A(x) = \begin{cases} 1, & \text{se } x \in A, \\ 0, & \text{se } x \notin A. \end{cases}$$

Perceba que $\chi_A(x) = 1$ indica que o elemento x está em A, enquanto que $\chi_A(x) = 0$ indica que x não é elemento de A, ou seja, a função característica descreve completamente o conjunto A.

Entretanto, existem casos em que não sabemos dizer se um elemento pertence efetivamente a um conjunto ou não. São os casos em que os termos linguísticos utilizados para descrever o conjunto não nos dão uma informação exata. Para esclarecimento, veja o seguinte exemplo.

Exemplo 1.2. Consideremos o conjunto A dos números reais "próximos de 2":

$$A = \{x \in \mathbb{R} : x \text{ é próximo de } 2\}$$

†Bolsista PET

Se tomarmos os números 7 e $2,001$, não conseguimos dizer certamente que um deles pertence ou não ao conjunto, mas uma informação razoável que pode ser tirada é que $2,001$ é mais próximo de 2 do que 7.

1.1 Subconjuntos fuzzy

Definição 1.3. Seja U um conjunto (clássico); um subconjunto fuzzy F de U é caracterizado por uma função

$$\varphi_F : U \to [0,1]$$

pré-fixada, chamada função de pertinência do subconjunto fuzzy F.

O índice F na função de pertinência é usado em analogia à função característica de um subconjunto clássico, de acordo com a Definição 1.1.

O valor $\varphi_F(x) \in [0,1]$ indica o grau de pertinência do elemento $x \in U$ ao conjunto fuzzy F. Quando $\varphi_F(x) = 0$, dizemos que x não pertence ao conjunto (não pertinência) e quando $\varphi_F(x) = 1$, dizemos que x pertence completamente ao conjunto (pertinência completa).

Perceba que a definição de subconjunto fuzzy foi obtida simplesmente ampliando-se o contradomínio da função característica, ou seja, que o conjunto clássico (crisp) é um caso particular de um dado conjunto fuzzy.

1.2 Exemplos de conjuntos fuzzy

Em seguida serão apresentados alguns exemplos de conjuntos fuzzy.

Exemplo 1.4 (Números pares)**.** Considere o conjunto dos números naturais pares:

$$P = \{n \in \mathbb{N}^* : n \text{ é par}\}$$

O conjunto P tem a seguinte função característica:

$$\chi_P(n) = \begin{cases} 1, & \text{se } n \text{ é par,} \\ 0, & \text{se } n \text{ é ímpar.} \end{cases}$$

Assim, o conjunto dos números pares é um particular conjunto fuzzy, já que $\chi_P(n) \in [0,1]$.

Exemplo 1.5 (Números próximos de 2)**.** Considere o conjunto A dado no Exemplo 1.2.

Podemos definir a função $\varphi_A : \mathbb{R} \to [0,1]$, dada pela expressão:

$$\varphi_A(x) = \begin{cases} 1 - |x-2|, & \text{se } 1 \le x \le 3, \\ 0, & \text{se } x \notin [1,3]\,. \end{cases}$$

Desse modo, encontramos uma maneira de categorizar os números $2,001$ e 7. O subconjunto fuzzy A dos pontos próximos de 2, caracterizado por φ_A é tal que $\varphi_A(2,001) = 0,999$ e $\varphi_A(7) = 0$. Dizemos, nesse caso, que $x = 2,001$ é um ponto próximo de 2 com grau de pertinência $0,999$ ao conjunto A e $x = 7$ não é próximo de 2.

O gráfico fica da seguinte forma:

(espaço em branco intencional)

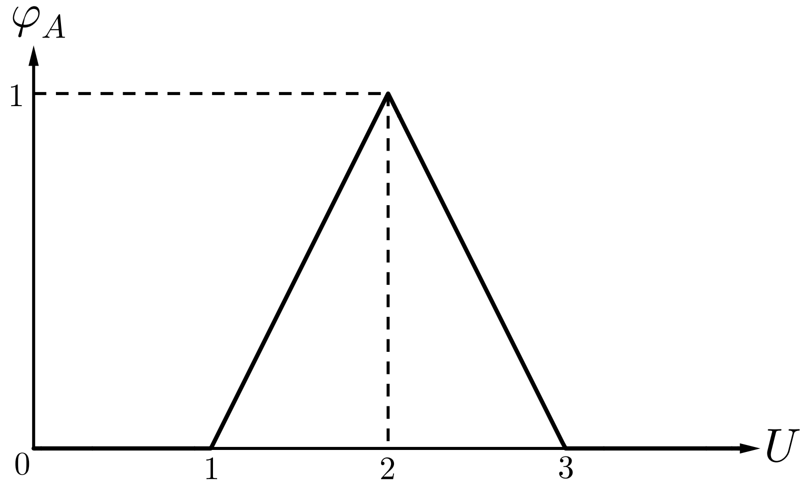


Figura 4.1: Números próximos de 2 segundo φ_A.

Perceba que a escolha da função pode variar, dependendo do problema a ser resolvido. Por exemplo, alguém poderia sugerir uma função de proximidade de 2 que fosse definida por:

$$\nu_F(x) = e^{\left[-(x-2)^2\right]}$$

com $x \in \mathbb{R}$, o que faria os elementos do conjunto fuzzy F, caracterizado pela função ν_F, terem outros graus de pertinência: $\nu_F(2,001) = 0,9999990000004999\ldots$ e $\nu_F(7) = 1,38879438\cdots 10^{-11}$. Note que, no gráfico a seguir, por se tratar de um função exponencial, nunca teremos $\nu_F = 0$.

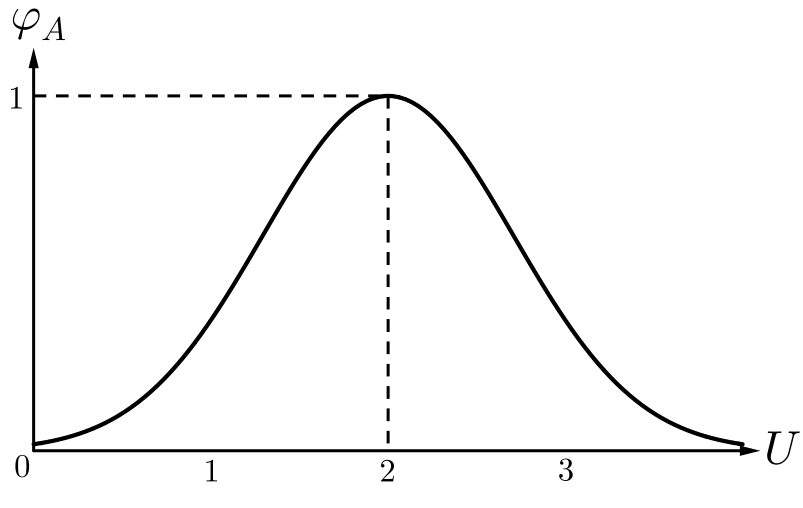


Figura 4.2: Números próximos de 2 segundo ν_F.

Isso quer dizer que a escolha da caracterização de proximidade é subjetiva e depende da função de pertinência, que pode ser dada de várias maneiras diferentes, e que depende de como se deseja avaliar o termo "próximo".

Mais além, poderíamos definir "próximo de 2" por um conjunto clássico com função de pertinência $\varphi_{\varepsilon F}$, sendo ε um valor suficientemente pequeno e a função característica dada conforme a seguinte expressão:

$$\varphi_{\varepsilon F}(x) = \begin{cases} 1, & \text{se } |x-2| < \varepsilon, \\ 0, & \text{se } |x-2| \geq \varepsilon. \end{cases}$$

Com essa função, o ser "próximo de 2" se resume em estar numa vizinhança pré determinada de 2, a escolha do raio da vizinha é totalmente subjetiva. Note que, especificamente nesse caso, todos os valores dessa vizinhança estão próximos de 2 com o mesmo grau de pertinência, que é 1.

2 Operações com conjuntos fuzzy

Agora que já vimos alguns exemplos de conjuntos fuzzy, serão apresentados algumas definições e resultados sobre operações entre conjuntos fuzzy, como união, intersecção e complementação.

Sejam A e B dois subconjuntos fuzzy de U e φ_A e φ_B suas respectivas funções de pertinência.

Dizemos que A é um subconjunto fuzzy de B, e denotamos por $A \subset B$, se $\varphi_A \leq \varphi_B$ para todo $x \in U$.

Além disso, a função de pertinência do conjunto vazio ($\emptyset$) é dada por $\varphi_{\emptyset}(x) = 0$ e a do conjunto universo (U) tem função de pertinência $\varphi_U(x) = 1$ para todo $x \in U$. Portanto, para todo conjunto A, temos que $\emptyset \subset A \subset U$.

Definição 2.1. Os subconjuntos fuzzy A e B de U são iguais se suas funções de pertinência forem iguais, ou seja:

$$\varphi_A(x) = \varphi_B(x), \forall\ x \in U.$$

Agora, sejam A e B dois subconjuntos fuzzy, com suas funções de pertinência dadas na figura a seguir. Vamos definir alguns conjuntos fuzzy que podem ser obtidos a partir deles.

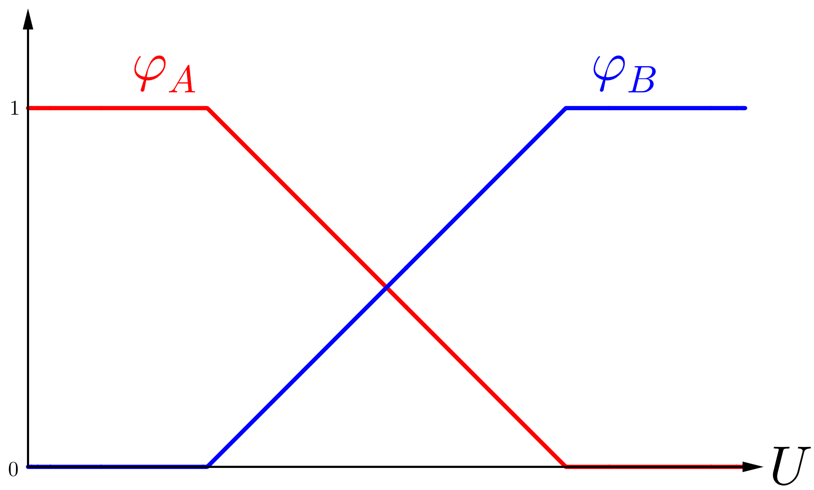


Figura 4.3: Gráfico das funções de pertinência φ_A e φ_B.

Definição 2.2 (União)**.** A união entre os subconjuntos fuzzy A e B é o subconjunto fuzzy de U cuja função de pertinência é dada por:

$$\varphi_{A \cup B}(x) = \max\{\varphi_A(x), \varphi_B(x)\}, x \in U.$$

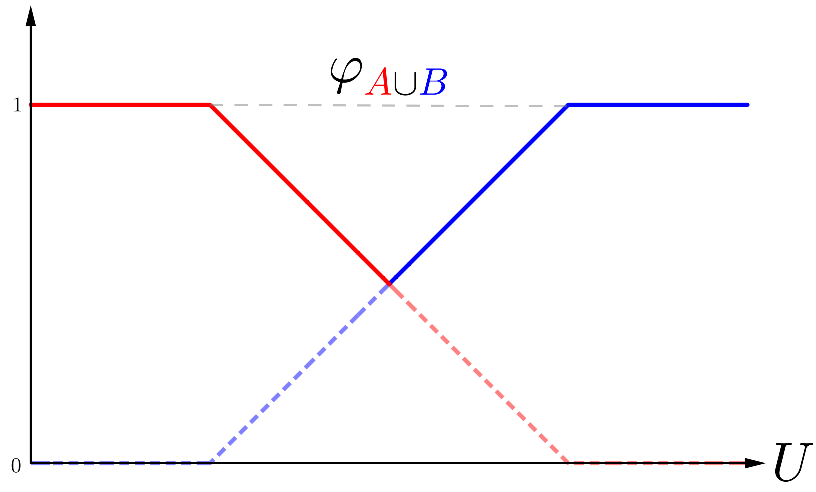


Figura 4.4: Gráfico da função de pertinência $\varphi_{A\cup B}$.

Perceba que esta definição é apenas uma extensão do caso clássico de união de conjuntos, já que para quaisquer conjuntos crisp A e B, temos:

$$\max\{\chi_A(x), \chi_B(x)\} = \begin{cases} 1, & \text{se } x \in A \vee x \in B \\ 0, & \text{se } x \notin A \wedge x \notin B \end{cases} = \begin{cases} 1, & \text{se } x \in A \cup B \\ 0, & \text{se } x \notin A \cup B \end{cases} = \chi_{A\cup B}(x),\ x \in U.$$

Definição 2.3 (Intersecção). A intersecção entre os subconjuntos fuzzy A e B é o subconjunto fuzzy de U cuja função de pertinência é dada por:

$$\varphi_{A\cap B}(x) = \min\{\varphi_A(x), \varphi_B(x)\}, x \in U.$$

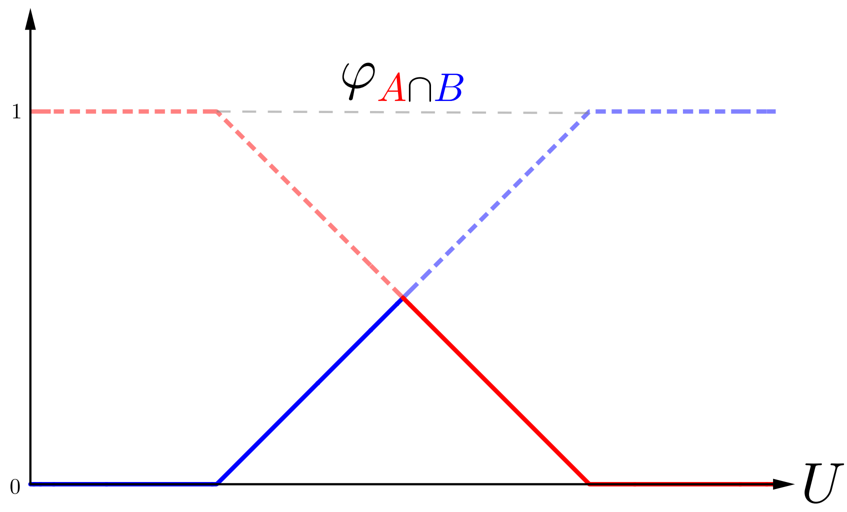


Figura 4.5: Gráfico da função de pertinência $\varphi_{A\cap B}$.

Analogamente, a definição de intersecção fuzzy é uma extensão do caso clássico de intersecção de conjuntos, uma vez que para todo conjunto crisp A e B, temos:

$$\min\{\chi_A(x), \chi_B(x)\} = 1 \iff \chi_A(x) = 1 \text{ e } \chi_B(x) = 1 \iff x \in A \text{ e } x \in B \iff x \in A \cap B.$$
$$\min\{\chi_A(x), \chi_B(x)\} = 0 \iff \chi_A(x) = 0 \text{ ou } \chi_B(x) = 0 \iff x \notin A \text{ ou } x \notin B \iff x \notin A \cap B.$$

Logo, $\min\{\chi_A(x), \chi_B(x)\} = \begin{cases} 1, & \text{se } x \in A \cup B, \\ 0, & \text{se } x \notin A \cup B. \end{cases}$

Definição 2.4 (Complementar de subconjuntos fuzzy). O complementar de A é o subconjunto fuzzy A' de U, cuja função de pertinência é dada por:

$$\varphi_{A'}(x) = 1 - \varphi_A(x), x \in U.$$

Veja o gráfico a seguir:

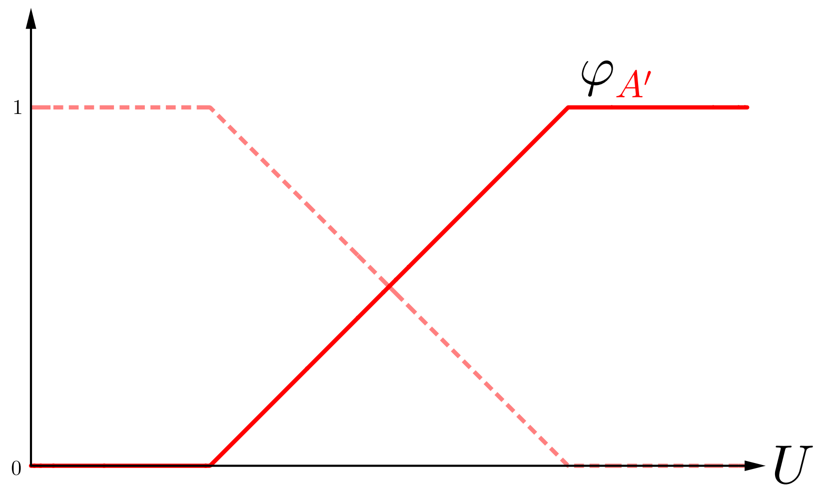


Figura 4.6: Gráfico da função de pertinência $\varphi_{A'}$.

Ao analisarmos o caso clássico, conseguimos alguns resultados mais diretos de intersecção e complementação de conjuntos.

Para a intersecção, observe que, para todo $x \in U$:

$$\chi_A(x) \cdot \chi_B(x) = 1 \iff \chi_A(x) = 1 \text{ e } \chi_B(x) = 1 \iff x \in A \text{ e } x \in B \iff x \in A \cap B. \quad (2.1)$$
$$\chi_A(x) \cdot \chi_B(x) = 0 \iff \chi_A(x) = 0 \text{ ou } \chi_B(x) = 0 \iff x \notin A \text{ ou } x \notin B \iff x \notin A \cap B. \quad (2.2)$$

De (2.1) e (2.2), temos que a intersecção também pode ser representada da seguinte maneira:

$$\chi_{A\cap B}(x) = \chi_A(x) \cdot \chi_B(x), \forall\, x \in U.$$

E para a complementação, basta notarmos que para todo conjunto crisp A e seu complementar A':

$$A \cap A' = \emptyset \text{ e } A \cup A' = U.$$

Logo, $\chi_{A\cap A'}(x) = \chi_\emptyset(x)$ e $\chi_{A\cup A'}(x) = \chi_U(x) = 1, \forall\, x \in U$.

Porém, quando trabalhamos com conjuntos fuzzy, essas propriedades descritas acima não são necessariamente verdadeiras, como veremos no exemplo a seguir.

Exemplo 2.5. Sejam A e B os seguintes conjuntos fuzzy e suas respectivas funções de pertinência:

$$A = \{x \in \mathbb{R} : x \text{ é próximo de } 2\} \text{ e } \varphi_A(x) = \begin{cases} 1 - |x-2|, & \text{se } x \in [1,3], \\ 0, & \text{se } x \notin [1,3]. \end{cases}$$

$$B = \{x \in \mathbb{R} : x \text{ é próximo de } 2,5\} \text{ e } \varphi_B(x) = \begin{cases} 1 - |x - 2,5|, & \text{se } x \in [1,5; 2,5]\,, \\ 0, & \text{se } x \notin [1,5; 2,5]\,. \end{cases}$$

Com isso, se tomarmos, por exemplo $x = 2,6$, temos que:

$$\varphi_{A\cap B}(2,6) = \min\{\varphi_A(2,6), \varphi_B(2,6)\} = \min\{0,4; 0,9\} = 0,4.$$

Por outro lado, $\varphi_A(2,6) \cdot \varphi_B(2,6) = 0,4 \cdot 0,9 = 0,36$.
E também, no contexto fuzzy, podemos ter

$$\varphi_{A\cap A'}(x) \neq 0 = \varphi_\emptyset(x), \text{ ou seja}, A \cap A' \neq \emptyset;$$
$$\varphi_{A\cup A'}(x) \neq 1 = \varphi_U(x), \text{ ou seja}, A \cup A' \neq U.$$

2.1 Os conjuntos fuzzy e a jovialidade

Sabendo as operações básicas, vamos refletir um pouco sobre a capacidade desses conjuntos fuzzy por meio de um exemplo.

Exemplo 2.6 (Conjunto fuzzy dos jovens). O IBGE considera jovem um indivíduo de até 24 anos de idade, considerando desde a infância até a adolescência. Então, vamos supor que um indivíduo tem 24 anos e está na véspera de seu aniversário, ou seja, no limite de sua jovialidade. Ele dorme tranquilamente em sua cama e, no dia seguinte, está totalmente idoso. O gráfico a seguir descreve a situação apresentada:

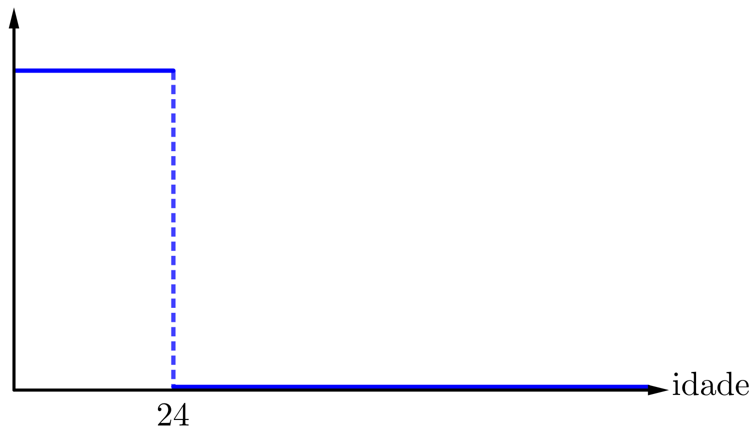


Figura 4.7: Gráfico do limite da jovialidade.

Sabemos que não é exatamente isso que acontece na vida real e, com a lógica fuzzy, conseguimos tratar essa situação de uma maneira diferente.

Como estamos tratando de idade de população, tomemos $U = [0, 122]$, sendo 0 a idade inicial e 122 a idade da pessoa mais velha que já existiu no planeta Terra. Considere a seguinte função de pertinência:

$$\varphi_{J_1}(x) = \begin{cases} 1, & \text{se } x \leq 24, \\ \dfrac{65 - x}{41}, & \text{se } 24 < x \leq 65, \\ 0, & \text{se } x > 65. \end{cases}$$

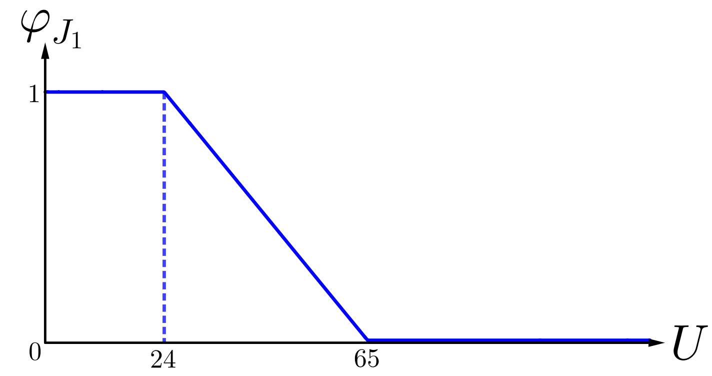


Figura 4.8: Gráfico da função de pertinência φ_{J_1}.

Note que iniciamos no valor 1, porém, quando chegamos na idade 24, não temos mais uma descontinuidade, e sim uma reta que decai até a idade 65. A partir dessa idade, o grau de pertinência é 0, o que consideramos uma pessoa totalmente idosa.

Portanto, um indivíduo de 25 anos não seria completamento idoso, uma vez que $\varphi_{J_1}(25) = 0,975609756097\ldots$. Assim, mesmo o indivíduo não sendo mais totalmente jovem, também não é totalmente idoso, o que retrata melhor a situação real.

As situações são modeladas por funções de pertinência de acordo com o problema a ser estudado. Por exemplo, poderíamos escolher a função φ_{J_2} e ao invés de um decaimento por meio de uma reta, teríamos uma parábola, decaindo até 40 anos:

$$\varphi_{J_2}(x) = \begin{cases} \left(\dfrac{40-x}{40}\right)^2, & \text{se } 0 \leq x \leq 40, \\ 0, & \text{se } 40 < x \leq 122. \end{cases}$$

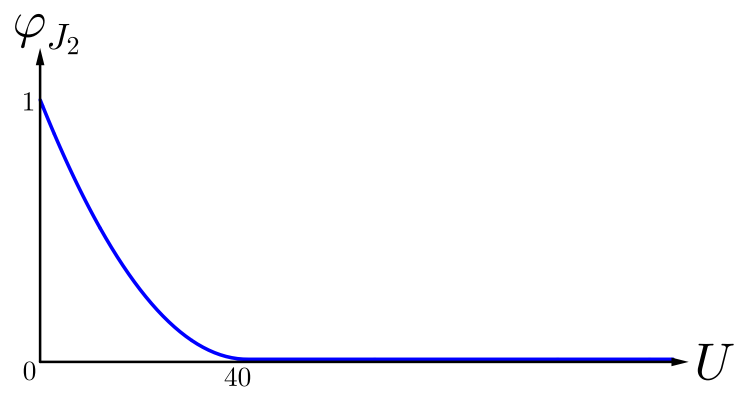


Figura 4.9: Gráfico da função de pertinência φ_{J_2}.

Perceba que, diferentemente da função anterior, o grau de jovialidade aos 25 anos é bem menor, já que $\varphi_{J_2}(25) = 0,140625$.

A escolha da função de pertinência adequada depende do problema estudado. A jovialidade pode variar dependendo do ponto de vista que está sendo analisado. Podemos ter a função de jovialidade para o mercado de trabalho, para a conclusão dos estudos, para planos de saúde, aposentadoria de atletas, entre outros. Por isso, os dados são coletados por especialistas nas determinadas áreas de conhecimento.

Vamos considerar uma última função φ_J relacionada à jovialidade. Note que o gráfico agora decai muito mais lentamente:

$$\varphi_J(x) = \begin{cases} \left[1 - \left(\dfrac{x-10}{122}\right)^2\right]^4, & \text{se } x \in [10,122]\,, \\ 1, & \text{se } x \notin [10,122]\,. \end{cases}$$

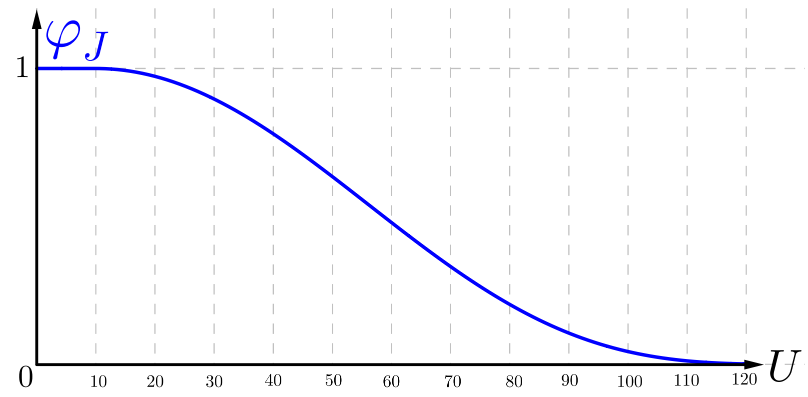


Figura 4.10: Gráfico da função de pertinência φ_J.

Agora, podemos refletir um pouco sobre seu complementar. A situação oposta relacionada ao conjunto dos jovens, quando consideramos a idade dos seus elementos, pode ser facilmente um conjunto fuzzy I dos idosos. Assim, uma possibilidade para a função de pertinência de I é:

$$\varphi_I = 1 - \varphi_J.$$

Ou seja, nesse caso, o complementar do conjunto fuzzy dos jovens é o conjunto fuzzy dos idosos. Os gráficos sobrepostos ficam da seguinte maneira:

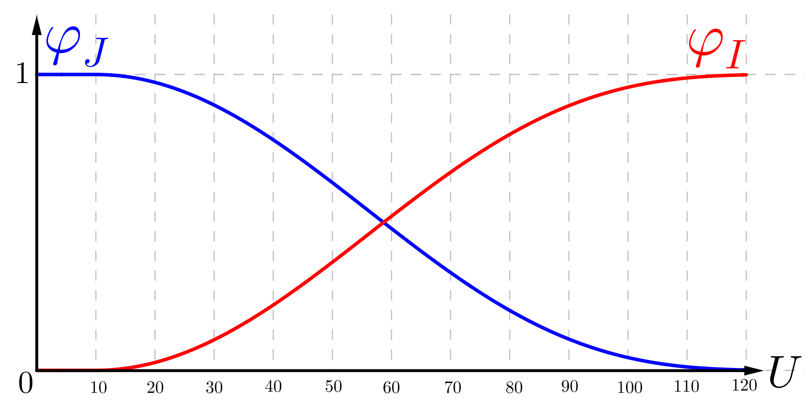


Figura 4.11: Gráfico do complementar fuzzy dos jovens.

Note que, segundo nossas funções de pertinência definidas, um indivíduo de "meia idade", isto é, grau $0,5$ tanto de jovialidade como de velhice, considerando nosso conjunto universo

$U = [0, 122]$, estaria por volta dos seus 58 anos de idade, pois:

$$\varphi_J(x) = \varphi_I(x) = 0,5$$

$$\left[1 - \left(\frac{x-10}{122}\right)^2\right]^4 = 0,5 \Longrightarrow x = 58,6631$$

Neste caso, a idade de 58 anos tem mesmo grau de pertinência tanto no conjunto dos jovens quanto no seu complementar, isso é uma consequência da imprecisão dos conjuntos fuzzy. Perceba que há uma sobreposição do conjunto fuzzy e seu complemento fuzzy. Um indivíduo que pertence ao conjunto fuzzy J com grau $0,7$, pertence também ao seu complementar I com grau $0,3$. Esta é a grande diferença da teoria clássica de conjuntos; na qual, um elemento, excludentemente, ou pertence a um conjunto ou ao seu complementar.

Novamente, enfatizamos que o conjunto dos idosos não necessariamente precisa ser o complementar de uma função dos jovens, isso é apenas uma possibilidade que faz sentido; porém, poderíamos trabalhar com a φ_J e com uma função φ_{I_1} que não é complementar da φ_J, conforme a figura a seguir:

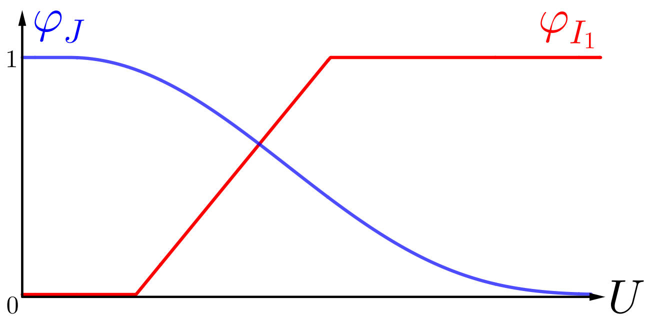


Figura 4.12: Gráfico das funções de pertinência não-complementares φ_J e φ_{I_1}.

A lógica fuzzy já tem um lugar de destaque, com aplicações práticas visíveis. Conseguimos concluir que o grande ganho quando modelamos com conjuntos fuzzy é um tratamento mais adequado a conjuntos que não definidos por termos linguísticos, cujas fronteiras são imprecisas.

Agradecimentos: Agradeço à professora Renata, por me orientar, sempre me ajudando, auxiliando e esclarecendo minhas dúvidas.

Abstract: In this article, we will show basic concepts of the Fuzzy Set Theory and some applications.

Keywords: Fuzzy sets; membership functions; fuzzy logic; operations with fuzzy sets

Referências Bibliográficas

[1] Barros, L. C. de; Bassanesi, R. C. *Tópicos de Lógica Fuzzy e Biomatemática*. 3. ed. Campinas: IMECC - UNICAMP, 2010.

[2] Wasques, V. F. *Lógica Fuzzy Aplicada à Geologia*. Rio Claro: Dissertação (Mestrado) - Universidade Estadual Paulista, Instituto de Geociências e Ciências Exatas, 2015.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 38–43

Determinante Mínimo Normalizado de Códigos de Bloco Espaço-Tempo Diagonais sobre $\mathbb{Q}(i)$

João Gabriel Oliveira de Jesus[†] e Carina Alves
Orientador(a): Profa. Dra. Carina Alves

Resumo: Neste trabalho apresentamos como obter extensões quadráticas ótimas sobre $\mathbb{Q}(i)$ a fim de maximizar o determinante mínimo normalizado de códigos de bloco espaço-tempo diagonais com duas antenas transmissoras e duas antenas receptoras.

Palavras-chave: extensões quadráticas; matriz geradora; determinante mínimo normalizado; códigos de bloco espaço-tempo

1 Introdução

Os sistemas de comunicação MIMO - *Multiple Input Multiple Output* - são formados por estruturas que utilizam várias antenas, tanto no transmissor como no receptor. Diversas adversidades interferem no desenvolvimento desse tipo de sistema de comunicação, como problemas de ruídos e de multipercursos. Nesse sentido, muitos estudos discutem certas propriedades dos sinais enviados com o intuito de minimizar os efeitos sofridos na informação durante a transmissão.

Diversidade máxima (maior confiabilidade do sinal) e maior determinante mínimo normalizado (menor probabilidade de erro ponto a ponto) são os dois mais importantes critérios para construir bons códigos de bloco espaço-tempo. Existem vários tipos de códigos STBC para duas antenas transmissoras. Neste trabalho consideramos códigos com diversidade máxima e o objetivo é que o determinante mínimo normalizado de tais códigos seja o maior possível. Os códigos considerados aqui são os códigos de bloco espaço-tempo diagonais.

2 Extensões quadráticas sobre $\mathbb{Q}(i)$

Um código de bloco espaço-tempo diagonal 2×2, $\mathcal{C} := \mathcal{C}(\mathbb{Q}(i), \alpha_1, \alpha_2, 0)$ baseado em um polinômio quadrático irredutível $x^2 + ax + b$ sobre $\mathbb{Z}[i]$, $a, b \in \mathbb{Z}$ e com raízes θ_1, θ_2 é definido por

$$D_2 = \left\{ X = \begin{pmatrix} z_1 + \theta_1 z_2 & 0 \\ 0 & z_1 + \theta_2 z_2 \end{pmatrix} \mid z_1, z_2 \in \mathbb{Z}[i] \right\} \tag{2.1}$$

cuja matriz geradora G do reticulado complexo é dada por

$$G = \begin{pmatrix} 1 & \theta_1 \\ 1 & \theta_2 \end{pmatrix}$$

e a matriz do reticulado real obtido via a extensão $\mathbb{Q}(i)$ sobre $\mathbb{Q}$ é dada por

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

[†]Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP

Podemos descrever explicitamente a relação entre o discriminante Δ do polinômio quadrático $x^2 + ax + b$ e o determinante da matriz geradora G desenvolvendo a expressão

$$\Delta = a^2 - 4b = (\theta_1 + \theta_2)^2 - 4\theta_1\theta_2 = \theta_1^2 + \theta_2^2 - 2\theta_1\theta_2 = (\theta_1 - \theta_2)^2 = \det(G)^2.$$

Usando a relação entre o determinante do reticulado complexo e o determinante do reticulado real introduzida em [3], podemos expressar o determinante mínimo normalizado do código D_2, denotado por d_N, como

$$d_N = \frac{1}{|\det(G)||\det(M)|} = \frac{1}{\sqrt{|\Delta|}}. \tag{2.2}$$

Como o código D_2 é determinado por um polinômio minimal quadrático sobre $\mathbb{Z}[i]$, segue que estamos interessados em determinar esse polinômio de modo que o determinante mínimo normalizado do código D_2 seja o maior possível, isto é, obter uma extensão quadrática ótima sobre $\mathbb{Q}(i)$.

De (2.2) notamos que maximizar d_N equivale à minimizar o valor absoluto do discriminante. Nesse sentido, é de grande interesse investigar a existência de uma cota inferior para o valor absoluto do discriminante de um polinômio quadrático minimal sobre $\mathbb{Z}[i]$ qualquer.

Exemplo 2.1. Considere o polinômio x^2-ix-1 irredutível sobre $\mathbb{Z}[i]$. Através desse polinômio construímos o código D_2, dado por

$$D_2 = \left\{ X = \begin{pmatrix} z_1 + \theta_1 z_2 & 0 \\ 0 & z_1 + \theta_2 z_2 \end{pmatrix}; z_1, z_2 \in \mathbb{Z}[i] \right\}$$

em que $\theta_1 = e^{\frac{i\pi}{6}}$ e $\theta_2 = e^{\frac{i5\pi}{6}}$ são as raízes do polinômio $x^2 - ix - 1$. Da expressão (2.2) segue que D_2 possui determinante mínimo normalizado d_N dado por

$$d_N = \frac{1}{\sqrt{|\Delta|}} = \frac{1}{\sqrt{i^2 + 4}} = \frac{1}{\sqrt{3}}.$$

Estamos interessados em saber se 3 é menor valor absoluto do discriminante de polinômios quadráticos irredutíveis sobre $\mathbb{Z}[i]$.

3 Resultados

Utilizando as referências [1, 2, 5], apresentamos resultados que solucionam as questões propostas neste trabalho. Mais especificamente, provamos que para qualquer polinômio x^2+ax+b sobre $\mathbb{Z}[i]$, o valor absoluto do seu discriminante é limitado inferiormente por 3 e, consequentemente, exibimos um limitante para o determinante mínimo normalizado do código.

Lema 3.1 ([2]). *Para qualquer polinômio $x^2 + ax + b$ irredutível sobre $\mathbb{Z}[i]$ o valor absoluto do seu discriminante $|\Delta|$ é limitado inferiormente por 3, isto é*

$$|\Delta| \geq 3.$$

Prova: Suponha θ_1 e θ_2 raízes de um polinômio irredutível $p(x) = x^2 + ax + b$ sobre $\mathbb{Z}[i]$, com $a, b \in \mathbb{Z}[i]$. Se considerarmos $a = l + mi$ e $b = s + ti$, em que $l, s, m, t \in \mathbb{Z}$, temos que $p(x) = x^2 + (l + mi)x + (s + ti)$. Logo, o discriminante desse polinômio é dado por

$$\Delta = a^2 - 4b = (l + mi)^2 - 4(s + ti) = l^2 - m^2 - 4s + 2(lm - 2t)i. \tag{3.1}$$

Assumimos que $|\Delta| < 3$ e vamos mostrar que isso é um absurdo considerando os seguintes casos:

(i) l par e m ímpar, isto é, $l = 2u$ e $m = 2v+1$, $u, v \in \mathbb{Z}$. Segue da expressão (3.1) que

$$\begin{aligned}\Delta &= (2u)^2 - (2v+1)^2 - 4s + 2[(2u)(2v+1) - 2t]i \\ &= 4(u^2 - v^2 + v - s) - 1 + 4(2uv + u - t)i. \end{aligned} \tag{3.2}$$

Como $|\Delta| < 3$, obtemos

$$\sqrt{[4(u^2 - v^2 + v - s)]^2 + [4(2uv + u - t)]^2} < 3.$$

Tomando $k_1 = u^2 - v^2 + v - s$ e $k_2 = 2uv + u - t$, segue da expressão acima que

$$\sqrt{(4k_1 - 1)^2 + (4k_2)^2} < 3 \Leftrightarrow (4k_1 - 1)^2 + (4k_2)^2 < 9. \tag{3.3}$$

Desenvolvendo (3.3), obtemos

$$16(k_1^2 + k_2^2) - 8k_1 + 1 < 9 \Leftrightarrow 2(k_1^2 + k_2^2) - k_1 < 1.$$

A desigualdade acima só é satisfeita quando $k_1 = k_2 = 0$, ou seja,

$$k_1 = 0 \Leftrightarrow u^2 - v^2 + v - s = 0 \Leftrightarrow s = u^2 - v^2 + v. \tag{3.4}$$

$$k_2 = 0 \Leftrightarrow 2uv + u - t = 0 \Leftrightarrow t = 2uv + u. \tag{3.5}$$

Com as expressões em (3.4) e (3.5) podemos reescrever $p(x)$ como

$$\begin{aligned}p(x) &= x^2 + (l + mi)x + (s + ti) \\ &= x^2 + [2u + (2v+1)i]x + (u^2 - v^2 - v) + (2uv + u)i \\ &= [x + u + (1 + v)i](x + u + vi). \end{aligned} \tag{3.6}$$

Claramente, a expressão (3.6) contradiz a irredutibilidade de $p(x)$ sobre $\mathbb{Z}[i]$. Portanto, para esse caso, devemos ter $|\Delta| \geq 3$.

(ii) l ímpar e m par, isto é, $l = 2u+1$ e $m = 2v$, com $u, v \in \mathbb{Z}$. De (3.1) podemos escrever o discriminante de $p(x)$ como

$$\begin{aligned}\Delta &= (2u+1)^2 - (2v)^2 - 4s + 2[(2u+1)(2v)]i \\ &= 4(u^2 + u - v^2 - s) + 1 + 4(2u + v - t)i. \end{aligned}$$

Sendo $|\Delta| < 3$, temos

$$\sqrt{[4(u^2 + u - v^2 - s) + 1]^2 + [4(2u + v - t)]^2} < 3. \tag{3.7}$$

Tomando $k_1 = u^2 - u - v^2 - s$ e $k_2 = 2uv + v - t$, temos de (3.7) que

$$\begin{aligned}(4k_1 + 1)^2 + (4k_2)^2 < 9 &\Leftrightarrow 16(k_1^2 + k_2^2) + 8k_1 + 1 < 9 \\ &\Leftrightarrow 2(k_1^2 + k_2^2) + (k_1 - 1) < 0. \end{aligned}$$

A desigualdade acima só é satisfeita quando $k_1 = k_2 = 0$, ou seja,

$$\begin{aligned}k_1 = 0 &\Leftrightarrow u^2 - u - v^2 - s = 0 \Leftrightarrow s = u^2 - u - v^2 \\ k_2 = 0 &\Leftrightarrow 2uv + v - t = 0 \Leftrightarrow t = 2uv + v. \end{aligned}$$

Reescrevendo $p(x)$ obtemos

$$\begin{aligned}p(x) &= x^2 + (l + mi)x + (s + ti) \\ &= x^2 + (2u + 1 + 2vi)x + u^2 + u - v^2 + (2uv + v)i \\ &= (x + u + 1 + vi)(x + u + vi), \end{aligned}$$

contradizendo a irredutibilidade de $p(x)$ sobre $\mathbb{Z}[i]$. Logo, para esse caso devemos ter $|\Delta| \geq 3$.

(iii) l e m pares, ou seja, $l = 2u$ e $m = 2v$, com $u, v \in \mathbb{Z}$. Da expressão (3.1) podemos escrever o discriminante de $p(x)$ como

$$\begin{aligned}\Delta &= (2u)^2 - (2v)^2 - 4s + 2[(2u)(2v) - 2t]i \\ &= 4(u^2 - v^2 - s) + 4(2uv - t)i.\end{aligned}$$

Tomando $k_1 = u^2 - v^2 - s$ e $k_2 = 2uv - t$, obtemos, como $|\Delta| < 3$,

$$|\Delta| = \sqrt{(4k_1)^2 + (4k_2)^2} < 3 \Leftrightarrow 16(k_1^2 + k_2^2) < 9.$$

Claramente, a desigualdade acima só é satisfeita quando $k_1 = k_2 = 0$, ou seja,

$$\begin{aligned}k_1 = 0 &\Leftrightarrow u^2 - v^2 - s = 0 \Leftrightarrow s = u^2 - v^2 \\ k_2 = 0 &\Leftrightarrow 2uv - t = 0 \Leftrightarrow t = 2uv.\end{aligned}$$

Com os valores de s e t obtidos, podemos reescrever $p(x)$ como

$$\begin{aligned}p(x) &= x^2 + (l + mi)x + (s + ti) \\ &= x^2 + (2u + 2vi)x + (u^2 - v^2 + 2uvi) \\ &= (x + u + vi)^2.\end{aligned}$$

A expressão acima contradiz a irredutibilidade de $p(x)$ sobre $\mathbb{Z}[i]$. Logo, para esse caso, $|\Delta| \geq 3$.

(iv) l e m ímpares, isto é, $l = 2u + 1$ e $m = 2v + 1$, com $u, v \in \mathbb{Z}$. Da expressão (3.1) podemos escrever o discriminante de $p(x)$ como

$$\begin{aligned}\Delta &= (2u + 1)^2 - (2v + 1)^2 - 4s + 2[(2u + 1)(2v + 1) - 2t]i \\ &= 4(u^2 + u - v^2 - v - s) + 4(2uv + v + u - t + 2)i.\end{aligned}$$

Como $|\Delta| < 3$, temos

$$|\Delta| = \sqrt{[4(u^2 + u - v^2 - v - s)]^2 + [4(2uv + v + u - t + 2)]^2} < 3. \tag{3.8}$$

Tomando $k_1 = u^2 + u - v^2 - v - s$ e $k_2 = 2uv + v + u - t$ temos de (3.8) que

$$16k_1^2 + (4k_2 + 2)^2 < 9.$$

Note que a desigualdade acima só é satisfeita quando $k_1 = k_2 = 0$, ou $k_1 = 0$ e $k_2 = -1$, isto é, quando $s = u^2 + u - v^2 - v$ e $t = 2uv + u + v$, ou $s = u^2 + u - v^2 - v$ e $t = 2uv + u + v + 1$. Consideremos estes dois casos, respectivamente.

(a) Para este caso temos,

$$\begin{aligned}x^2 + ax + b &= x^2 + (l + mi)x + s + ti \\ &= x^2 + [2u + 1 + (2v + 1)i]x + u^2 + u - v^2 - v + (2uv + u + v)i \\ &= [x + u + 1 + (v + 1)i](x + u + vi).\end{aligned}$$

contradizendo a irredutibilidade de $p(x)$ sobre $\mathbb{Z}[i]$.

(b) Neste caso temos,

$$\begin{aligned} x^2+ax+b &= x^2+(l+mi)x+s+ti \\ &= x^2+[2u+1+(2v+1)i]x+u^2+u-v^2-v+(2uv+u+v+1)i \\ &= (x+u+1+vi)[x+u+(v+1)i]. \end{aligned}$$

contradizendo a irredutibilidade de $p(x)$ sobre $\mathbb{Z}[i]$.

Segue dos itens **(i)**, **(ii)**, **(iii)** e **(iv)** a validade do lema. ■

A partir do Lema 3.1 temos condições de estabelecer o seguinte teorema.

Teorema 3.2 ([2]). *O determinante mínimo normalizado de qualquer código de bloco espaço-tempo diagonal D_2, como definido em* (2.1)*, é limitado superiormente por $\frac{1}{\sqrt{3}}$, isto é, $d_N \leq \frac{1}{\sqrt{3}}$.*

Prova: Um polinômio minimal x^2+ax+b sobre $\mathbb{Q}(i)$ é sempre irredutível sobre $\mathbb{Q}(i)$ e portanto, irredutível sobre $\mathbb{Z}[i]$. Logo, segue do Lema 3.1 que

$$|\Delta| \geq 3 \Leftrightarrow d_N = \frac{1}{\sqrt{|\Delta|}} \leq \frac{1}{\sqrt{3}}.$$ ■

Agradecimentos: Agradeço à minha orientadora, Profa. Dra. Carina Alves, pelo apoio e confiança, e à Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP, processo nº 2019/20800-8, pelo suporte financeiro.

Abstract: In this work we present how to obtain optimal quadratic extensions on $\mathbb{Q}(i)$ in order to maximize the minimum normalized determinant of diagonal space-time block codes with two transmitting and two receiving antennas.

Keywords: quadratic extensions; generator matrix; normalized minimum determinant; space-time block codes

Referências Bibliográficas

[1] Damen, M.O., Tewfik, A. and Belfiore, J.-C. A construction of a space-time code based on number theory, *IEEE Transactions on Information Theory*, 48:753-760, 2002. DOI: 10.1109/18.986032.

[2] Liao, H., Wang, H. and Xia, X.-G. Some designs and normalized diversity product upper bound for lattice-based diagonal and full-rate space-time block codes, *IEEE Transactions on Information Theory*, 55:569-583, 2009. DOI: 10.1109/TIT.2008.2009851.

[3] Wang, G., Liao, H., Wang, H. and Xia, X.-G., Systematic and optimal cyclotomic lattice and diagonal space-time block code designs, *IEEE Transactions on Information Theory*, 50:3348-3360, 2004. DOI: 10.1109/TIT.2004.838096.

[4] Wang, G., Zhang, J.K. and Amin, M. Space-time block code designs based on quadratic field extension for two-transmitter antennas, *IEEE Transactions on Information Theory*, 58:4005-4013, 2012. DOI: 10.1109/TIT.2012.2184633.

[5] Wang, G. and Xia, X.-G. On optimal multilayer cyclotomic spacetime code designs, *IEEE Transactions on Information Theory*, 51:1102–1135, 2005. DOI: 10.1109/TIT.2004.842630.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 44–46

Um Estudo Sobre as Proporções dos Sorteios da Mega-Sena

Luana Aparecida Pasetto†
Orientador(a): José Silvio Govone

Resumo: Neste artigo para exemplificar como é realizada uma amostragem probabilística, mais especificamente a amostragem casual simples, utilizamos os dados da Mega-Sena, um jogo de azar realizado pelas Caixas Lotéricas valido para todo o Brasil. Um estudo da aleatoriedade dos sorteios realizados, através do teste quiquadrado de igualdade de Proporções.

Palavras-chave: amostragem; probabilística

1 Introdução

A Amostragem probabilística nos permite a obtenção de informações desconhecidas, a partir da observação de uma "pequena" parte de uma população. Neste estudo foi utilizado a amostragem casual simples, um processo de amostragem probabilística. Dada uma população de N elementos é realizado um sorteio, sem reposição e, em cada etapa do sorteio todos os elementos remanescentes têm igual chance de seleção. O número de possíveis amostras é calculado pela combinação de N elementos n a n, $C_{(N;n)}$, onde n é o tamanho amostral. A probabilidade de cada amostra ser sorteada é definida por $\frac{1}{C_{(N;n)}}$ [1].

Assim obtemos uma série de n números aleatórios entre 1 e N (sem repetição) sendo o sorteio feito com o apoio de uma tabela de números aleatórios ou de rotinas em computadores.

Neste estudo a partir de uma amostra aleatória dos resultados dos sorteios realizados entre 11 de Março de 1996 a 26 de junho de 2019, um total de 2163 sorteios, esses dados foram obtidos no site da CAIXA [2]. Foi analisado se o jogo de fato é um sorteio, isto é se os resultados são aleatórios, ou se há algum número com maior probabilidade de ser sorteado; para esta análise foi utilizado o Teste Qui-Quadrado de igualdade de proporções.

2 Obtenção da Amostra

Retirou-se uma amostra aleatória simples com 1081 jogos, correspondendo a metade dos jogos disponíveis, com 6486 números sorteados (6 números por jogo). Para realizar a amostragem foi utilizado o comando "RANDOM" no LibreOficce Calc, este comando nomeou cada jogo com número entre 0 e 1 de forma aleatória, estes números de 0 a 1 foram reordenados em ordem crescente, e foi tomado como amostra os 1081 primeiros números.

A tabela 6.1 apresenta as frequências absolutas dos dígitos final dos números sorteados

Podemos notar na tabela 6.1 que os números terminados em 4 tem maior prevalência na amostra, enquanto os números terminados 9 tem menor.

Assim, nos perguntamos, será que a partir da análise feita na tabela podemos afirmar que em todos os jogos existe maior probabilidade de ser sorteado um número terminado em 4? Ou que existe menor chance de sair números terminados em 9?

†FNDE-PET

Tabela 6.1: Frequências observadas dos dígitos finais dos números amostrados dos sorteios da Mega-Sena

Números terminados	Frequência Observada
0	647
1	637
2	650
3	653
4	704
5	634
6	661
7	652
8	651
9	597
TOTAL	6486

Para responder estas questões temos que verificar se essas diferenças tem significância estatística, para isso utilizamos o teste quiquadrado de igualdade proporções, em que as hipóteses são:

$$H_0\text{: Proporções Iguais} \qquad H_i\text{: Proporções diferentes}$$

Portanto, se os jogos são de fato aleatórios, a hipótese nula H_0 é satisfeita. Dessa forma aplicamos o teste quiquadrado com 9 Graus de liberdade e com 5% de significância, isso nos garante que $\chi^2_{\text{crítico}} = 16,92$ (valor encontrado na tabela quiquadrado). Se acontecer de $\chi^2 \leq \chi^2_{\text{crítico}}$; sendo χ^2 o valor calculado do teste, podemos concluir que H_0 é satisfeita.

Para aplicar o teste é necessário termos conhecimento da Frequência Observada (o_i) em cada umas das classes da Tabela 1, da Frequência Esperada (e_i) que é dada pelo total de dados observados dividido pelo número de classes, este valor é igual para todas as classes.

A tabela 6.2 mostra estes valores a serem calculados.

Tabela 6.2: Tabela de frequências

Números terminados	o_i	e_i
0	647	648,6
1	637	648,6
2	650	648,6
3	653	648,6
4	704	648,6
5	634	648,6
6	661	648,6
7	652	648,6
8	651	648,6
9	597	648,6
TOTAL	6486	

O teste Qui-Quadrado é dado pela expressão:

$$\chi^2 = \sum_{i=1}^{10} \frac{(o_i - e_i)^2}{e_i}.$$

Realizando este calculo obtemos $\chi^2 = 9,67$.

Como $\chi^2 = 9,67 \leq 16,92 = \chi^2_{\text{crítico}}$ podemos concluir que há igualdade de proporções, ou seja, nos sorteios realizados pela Mega-Sena o último dígito dos números tem iguais probabilidades de serem sorteados.

Abstract: In this work, to exemplify how a probabilistic sampling is performed, more specifically simple casual sampling. We used data from Mega-Sena, a game of chance performes by lottery, valid for all Brazil. A study of the randomness of the draws is carried out through the chi square test of equality of proportions.

Keywords: randon sampling; proportions

Referências Bibliográficas

[1] Silva, N.N., *Amostragem probabilística: um curso introdutório*, [S.l: s.n] 1998.

[2] CAIXA. Loterias. Resultado da Mega-sena. [S.l.: s.n.], 2020. Disponivel em: `http://www1.caixa.gov.br/loterias/_arquivos/loterias/D_megase.zip`. Acesso em: 12 ago. 2019

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 47–51

Plano de Isometrias

Luana Guimarães Nogueira†
Orientador(a): Profa. Dra. Elíris Cristina Rizziolli

Resumo: Isometrias são transformações geométricas que mantém as distâncias entre os pontos. Assim, os ângulos mantém sua amplitude e também, os segmentos e pontos dessa nova figura são geometricamente iguais ao da figura original. Entretanto, a direção pode mudar. Nesse trabalho estudaremos o plano de isometrias e 4 tipos de isometrias, a saber: rotação, translação, reflexão e reflexão com deslizamento.

Palavras-chave: isometria; reflexão; translação; rotação; reflexão com deslizamento; transformação geométrica; plano de isometrias.

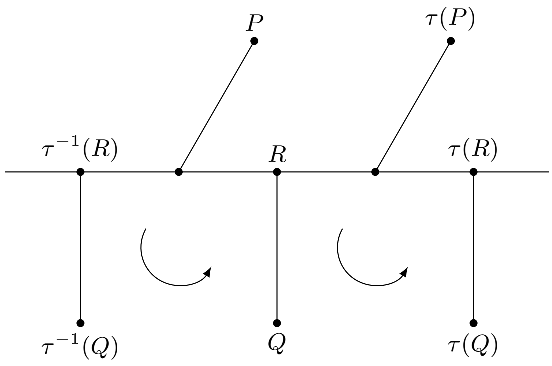
1 Plano de Isometrias

Considere o plano Euclidiano $\mathbb{R}^2$. Uma isometria de $\mathbb{R}^2$ é uma permutação em $\mathbb{R}^2$ (isto é, uma bijeção) $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ que preserva distância, ou seja, tal que a distância entre dois pontos P e Q é a mesma que a distância entre $\phi(P)$ e $\phi(Q)$ para todos os pontos P e Q em $\mathbb{R}^2$. Se ψ é também uma isometria de $\mathbb{R}^2$, então a distância entre $\psi(\phi(P))$ e $\psi(\phi(Q))$ deve ser a mesma que a distância entre $\phi(P)$ e $\phi(Q)$, que por sua vez é a distância entre P e Q, o que se mostra com isso é que a composição de duas isometrias é também uma isometria. Uma vez que a função identidade é uma isometria e a inversa de uma isometria é uma isometria, segue que as isometrias de $\mathbb{R}^2$ formam um subgrupo do grupo de todas as permutações de $\mathbb{R}^2$.

Dado qualquer subconjunto S de $\mathbb{R}^2$, as isometrias de $\mathbb{R}^2$ que aplicam S sobre ele mesmo formam um subgrupo do grupo de isometrias. Esse subgrupo é o **grupo de simetrias de S em $\mathbb{R}^2$**.

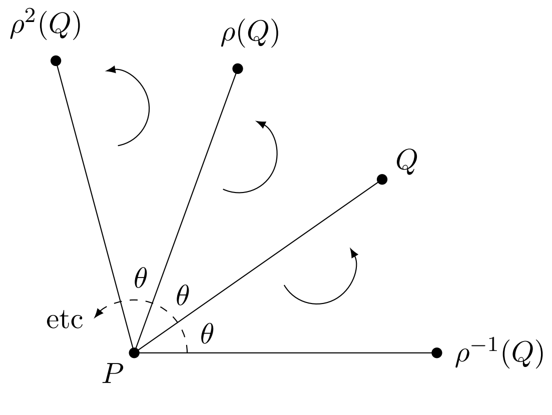
Em cada uma das figuras a seguir, apresentamos as 4 isometrias que abordaremos nesse trabalho:

- **Translação τ:** Deslize todo ponto a mesma distância na mesma direção. (Exemplo: $\tau(x, y) = (x, y) + (2, -3) = (x + 2, y - 3)$);

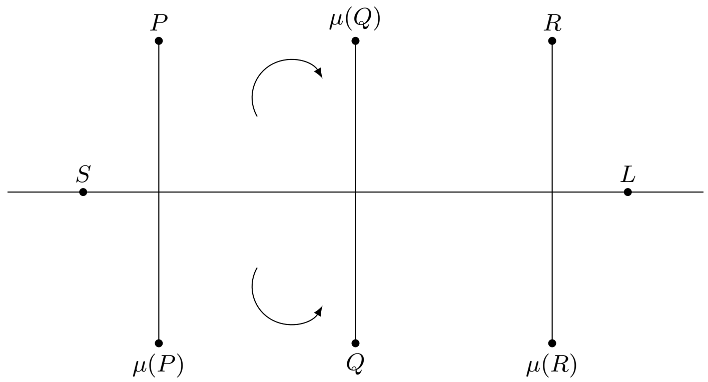


†PET/MEC-Sesu

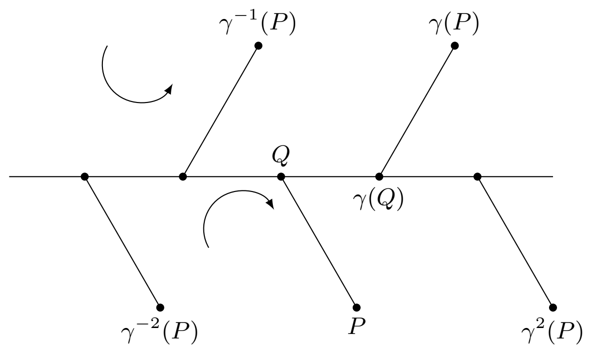
- **Rotação ρ:** Rotacione o plano sobre um ponto P através de um ângulo θ. (Exemplo: $\rho(x, y) = (-y, x)$ é uma rotação de 90° em sentido anti-horário sobre a origem $(0, 0)$);



- **Reflexão** μ: Aplique cada ponto dentro de sua imagem espelhada (μ para espelho) através de uma reta L, cada ponto do qual é deixado fixo por μ. A reta L é o *eixo de reflexão*. (Exemplo: $\mu(x, y) = (y, x)$ é uma reflexão através da reta $y = x$.);



- **Reflexão com deslizamento γ:** O produto de uma translação e uma reflexão através da reta aplicada nela mesma por uma translação. (Exemplo: $\gamma(x, y) = (x + 4, -y)$ é uma reflexão com deslizamento ao longo do eixo x).



Veja que em cada uma das figuras existe uma sequência de setas curvadas, cujas direções têm os seguintes significados: observe que para translação e rotação, as direções das setas curvadas no sentido anti-horário permanecem as mesmas, mas para a reflexão e para a reflexão com deslizamento, uma seta no sentido anti-horário é levada em uma seta no sentido horário. Neste contexto, dizemos que translações e rotações preservam a orientação, enquanto a reflexão e a reflexão com deslizamento invertem a orientação.

Observe ainda que não classificamos a isometria identidade como nenhum dos quatro tipos listados, entretanto tal isometria poderia ser igualmente olhada como a translação pelo vetor nulo ou uma rotação sobre qualquer ponto através de um ângulo de 0°. Sempre consideramos que uma reflexão com deslizamento é o produto de uma reflexão e uma translação, quando esta é diferente da isometria identidade.

A seguir definimos os grupos D_n e $\mathbb{Z}_n$ que serão usados no Teorema principal desse trabalho.

Definição 1.1. O grupo $\mathbb{Z}_n$ é o grupo das classes de equivalência $\overline{a}$, $a \in \mathbb{Z}$, sob a seguinte relação de equivalência: $\forall a, b \in \mathbb{Z}, a \sim b \Leftrightarrow a = b + nk$, para algum $k \in \mathbb{Z}$, munido da operação $\overline{a} + \overline{b} := \overline{a+b}$. Mais especificamente, $\mathbb{Z}_n$ é o grupo de ordem n dado por:

$$\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}.$$

Definição 1.2. O grupo diedral D_n é o grupo das simetrias de um polígono regular de n lados, munido da operação composição. Mais especificamente, D_n é o grupo de ordem $2n$ dado por:

$$D_n = \{R, R^2, \ldots, R^n, X, X \circ R, X \circ R^2, \ldots, X \circ R^{n-1}\},$$

em que R é a rotação $\dfrac{2\pi}{n}$ e X é a reflexão com relação a um vértice fixado e o centro O.

O teorema a seguir descreve as possíveis estruturas de um grupo finito constituído por simetrias do plano.

Teorema 1.3. *Todo grupo finito de isometrias do plano é isomorfo a $\mathbb{Z}_n$ ou a algum grupo diedral D_n.*

Prova: Seja G um grupo finito de isometrias do plano. Primeiro mostramos que existe um ponto P no plano que é fixado por cada isometria em G. Isso pode ser feito usando coordenadas no plano, como segue: suponha que $G = \{\phi_1, \phi_2, \ldots, \phi_n\}$ e seja

$$(x_i, y_i) = \phi_i(0,0).$$

Então o ponto

$$P = (\overline{x}, \overline{y}) = \left(\frac{x_1 + x_2 + \cdots + x_n}{n}, \frac{y_1 + y_2 + \cdots + y_n}{n}\right)$$

é o centroide do conjunto $S = \{(x_i, y_i); i = 1, 2, \ldots, n\}$. As isometrias em G permutam os pontos em S entre si, pois se $\phi_i \phi_j = \phi_k$ então $\phi_i(x_j, y_j) = \phi_i[\phi_j(0,0)] = \phi_k(0,0) = (x_k, y_k)$. Pode ser demonstrado que o centroide de um conjunto de pontos é determinado unicamente pelas suas distâncias entre os pontos, e como cada isometria em G apenas permuta o conjunto S, esta deve deixar o centroide $(\overline{x}, \overline{y})$ fixo. Assim, G consiste na identidade (que denotaremos por e), rotações em torno de P e reflexões através de uma reta que passa por P e, como sabemos, a identidade e as rotações preservam orientação, enquanto que as reflexões através de uma reta que passa por P, não preservam orientação.

As isometrias em G que preservam orientação formam um subgrupo H de G que, como veremos, é todo o G ou da ordem $\frac{n}{2}$ (onde $n = o(G)$), pois:

- Se $o(H) = o(G)$, então $H = G$;
- Se $o(H) \neq o(G)$, então, pelo Teorema de Lagrange, $o(G) = (G : H)o(H)$. A seguir calculamos o índice $(G : H)$:

  Sabemos que $(G : H)$ é o número de classes laterais distintas ϕH, com $\phi \in G$. Nesse caso, temos duas possibilidades:

  1. ϕ preserva orientação, logo ϕH é o próprio H;
  2. ϕ não preserva orientação e portanto $\phi H = G - H$ (complementar do H em G). Assim temos só duas classes laterais distintas e, pelo Teorema de Lagrange:

$$o(H)2 = o(G) \Rightarrow o(H) = \frac{o(G)}{2} \Rightarrow o(H) = \frac{n}{2}.$$

Veja que H consiste na identidade e nas rotações em G, já que estas são as isometrias que preservam orientação. Veremos a seguir que, se $H \neq \{e\}$, então existe uma rotação em G que rotaciona o plano com o menor ângulo $\theta > 0$ e que, tal rotação gera o subgrupo H. Mostremos isso:

- Se $H = \{e\} \Rightarrow e$ (a rotação nula, de ângulo 0°) gera H, ou seja, H é cíclico.
- Se $H \neq \{e\}$:

  Afirmação: Existe $\theta \in H$ tal que H é gerado por I_θ (onde I_θ é a rotação de ângulo θ). Para isso, usamos o Princípio da Boa Ordem para o conjunto:

$$A = \{\beta \in \mathbb{Z}_+^*; I_\beta \in H\}.$$

  O conjunto A é limitado inferiormente pelo ângulo 0°, já que a rotação de 0° é a identidade e está em H. Portanto, pelo Princípio da Boa Ordem, A tem um menor elemento.

  Seja θ o menor elemento de A. Agora, mostramos que H é gerado por I_θ, para isso, seja $I_\alpha \in H$, $I_\alpha \neq e$, uma rotação qualquer de ângulo $\alpha \in \mathbb{Z}_+^*$. Temos, pelo Algoritmo da Divisão para números inteiros, que:

$$\alpha = q\theta + r,$$

  em que $0 \leq r < \theta$. Então:

$$r = \alpha - q\theta,$$

  mas $I_\theta \in H$, logo $I_{q\theta} \in H$ e como I_α também pertence a H, segue que $I_{\alpha - q\theta} \in H$, ou seja, $I_r \in H$. Logo, teremos necessariamente $r = 0$, pois se $r \neq 0$, chegaríamos em um absurdo, já que $r < \theta$ e θ é o menor ângulo, tal que $I_\theta \in H$. Sendo assim,

$$r = 0 \Rightarrow \alpha = q\theta,$$

  ou seja, $I_\alpha = I_{q\theta} \in \langle I_\theta \rangle$. Portanto, $H = \langle I_\theta \rangle$, ou seja H é cíclico.

Com isso, vimos que se $H = \{e\}$ ou se $H \neq \{e\}$, nos dois casos ele é cíclico. Consequentemente, se $G = H$, ou seja, se G só tem rotações, então G é cíclico (já que, como vimos, H é cíclico) e como $o(G) = n$, G é isomorfo Z_n. Agora suponha que $H \neq G$ e, portanto, G não tem apenas rotações, logo G contém necessariamente algumas reflexões. Seja $H = \{e, \rho_1, \ldots, \rho_{m-1}\}$ com $m = \frac{n}{2}$. Se μ é uma reflexão em G, então a classe lateral

$H\mu$ consiste de todas as m reflexões em G, assim como $G = H \cup H_\mu$ e $H \cap H_\mu = \emptyset$, com $o(G) = m + m = \dfrac{n}{2} + \dfrac{n}{2} = n$.

Considere agora um polígono regular de m lados no plano tendo P como seu centro e com um vértice pertencente à reta que passa por P, fixado pela reflexão μ. Os elementos de H giram esse polígono por todas as m posições (em função das m rotações de H) e, os elementos de $H\mu$ primeiro refletem em um eixo através de um vértice, efetivamente virando o polígono e, em seguida, giram em todas as posições. Assim, a ação de G nesse polígono de m lados é a ação de $D_{2m} = D_{2\frac{n}{2}} = D_n$, então G é isomorfo a D_n. ■

Mesmo sem perceber, as isometrias estão presentes em nosso cotidiano. Como a roda gigante (rotação), borboleta (simetria axial) e também a arte, a pintura, a cerâmica e a tecelagem, que são exemplos de isometrias devido aos padrões geométricos que seguem.

Agradecimentos: Agradeço à minha orientadora, Elíris, pelo tempo investido nesse trabalho comigo, pela paciência, pelos ensinamentos e correções feitas. Agradeço também ao PET, por ser um programa que enriqueceu e enriquece minha formação com a disciplina, comprometimento, trabalho em equipe, dedicação e responsabilidade que venho adquirindo, ao fazer parte desse grupo. Por fim, agradeço aos familiares e amigos que tanto me apoiaram e apoiam durante a graduação.

Abstract: Isometries are geometric transformations that maintain the distances between points. Thus, the angles maintain their amplitude and also, the segments and points of this new figure are geometrically the same as the original figure, however the direction can change. In this work we will study the plan of isometries and 4 types of isometries, namely: rotation, translation, reflection and glide reflection.

Keywords: isometry; reflection; translation; rotation; glide reflection; geometric transformation; isometries plan.

Referências Bibliográficas

[1] Fraleigh, J.B., *A First Course in Abstract Algebra*, Addison Wesley, Massachusetts, 7th edition, 2003.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM
MATEMÁTICA · BICMat
Vol. 17 (2020) 52–59

O Teorema de Heine Borel, Compacidade e Funções

Lucas Gattera Begiato
Orientador(a): Prof. Dr. João Peres Vieira

Resumo: Em Matemática, muitos são os conceitos que varrem os seus diferentes ramos. Em particular, na Topologia, estuda-se os espaços topológicos, que em muitas vezes, podem ser vistos como uma extensão da geometria, abrangendo-a de forma mais ampla e abstrata. Segundo o trabalho intitulado *"Sobre a Noção de Compacidade"* realizado por Cecília Fernandez e Luiz Silva [2], o conceito de compacidade foi introduzido por Maurice Fréchet (1878-1973) em sua tese de doutorado que fora defendida em 1906, conferindo ao conceito de compacidade uma extensão topológica das ideias de finitude e limitação. A partir disso, diversos matemáticos começaram a desenvolver este assunto. Entre os muitos que se destacaram, *Heinrich Eduard Heine* e *Félix Édouard Justin Émile Borel* foram responsáveis por demonstrar um teorema bastante conhecido na área da Topologia, chamado **Teorema de Heine-Borel**. Neste trabalho, será feita demonstração e uma aplicação deste teorema, bem como obteremos resultados sobre compacidade e funções.

Palavras-chave: Compacidade; Heine-Borel; Topologia; Continuidade; Homeomorfismo; Conjunto-Aberto; Conjunto-Fechado

1 Preliminares

Nesta seção daremos alguns conceitos e resultados que serão de suma importância para a próxima seção. As demonstrações dos resultados poderão ser encontradas em [1], que é a referência principal deste trabalho.

Definição 1.1. Se (X, d) é um espaço métrico e $x \in X$, então a δ-vizinhança de x em X é definida como: $U_X(x; \delta) = \{y \in X : d(x, y) < \delta\}$.

Definição 1.2. Qualquer subconjunto de X que contém $U(x, \delta)$ para algum $\delta > 0$ é chamado de vizinhança de x.

Definição 1.3. Um subconjunto de X é dito aberto em X se tem uma vizinhança para cada um de seus pontos.

Definição 1.4. Se $B \subset Y$ e $f : X \to Y$ é uma função, a imagem inversa de B por f é o subconjunto de X dado por: $f^{\dashv}(B) = \{x \in X | f(x) \in B\}$.

Teorema 1.5. *Sejam X, Y espaços métricos e $f : X \longrightarrow Y$ uma função. Então as seguintes afirmações são equivalentes:*

1. *f é contínua em $x \in X$.*
2. *Para qualquer vizinhança N de $f(x)$, em Y, $f^{\dashv}(N)$ é aberta em X.*

As seguintes afirmações também são equivalentes:

3. *f é contínua em $x \in X$.*

4. *Para algum V aberto em Y, $f^{-1}(V)$ é aberta em X.*

Definição 1.6. Dado qualquer subconjunto Y de X há um maior subconjunto aberto de X contido em Y. Esse conjunto é chamado de **interior** de Y em X, escrito como $Int_X(Y)$, que é a união de todos os subconjuntos abertos de X contidos em Y.

Definição 1.7. Um subconjunto Y de X é fechado em X se $X - Y$ é aberto.

Definição 1.8. O conjunto de pontos de X que não são nem interiores de Y nem interiores de $X - Y$ é chamado de **fronteira** de Y em X, denotada por $Fr_X(Y)$.

Definição 1.9. Dado qualquer subconjunto Y de X, há um menor subconjunto fechado de X contendo Y. Esse conjunto é chamado de **fecho** de Y em X, escrito como $Cl_X(Y)$, que é a união do interior e da fronteira de Y em X.

Proposição 1.10. *Y é um subconjunto fechado de X se, e somente se, $Y = Cl_X(Y)$.*

Definição 1.11. Para A e B subconjuntos de um espaço métrico (X, d) definimos $d(A, B) = \inf\{d(a, b) : a \in A, b \in B\}$.

Definição 1.12. Para todo subconjunto não-vazio Y de um espaço métrico (X, d) e $x \in X$, escrevemos $d(x, Y) = \inf\{d(x, y) : y \in Y\}$.

Lema 1.13. *Seja Y um subespaço não-vazio do espaço métrico X e $x \in X$. As seguintes afirmações são equivalentes:*

(i) *$x \in Cl_X(Y)$;*

(ii) *Qualquer conjunto aberto contendo x intersecta Y;*

(iii) *$d(x, Y) = 0$.*

Definição 1.14. Sejam X e Y espaços métricos e $f : X \longrightarrow Y$ uma função. Dizemos que f é um homeomorfismo se:

1. f é contínua;
2. f é bijetora;
3. f^{-1} é contínua.

Quando existe um homeomorfismo entre X e Y, dizemos que X e Y são homeomorfos.

Definição 1.15. Se X e Y são espaços métricos, usaremos a notação $X \underset{ab}{\subset} Y$ quando X for um subconjunto aberto de Y. Ainda, se X for um subconjunto fechado de Y, denotaremos por $X \underset{f}{\subset} Y$.

Definição 1.16. Sejam X e Y espaços métricos e $f : X \to Y$ uma função. Dizemos que f é aberta se $U \underset{ab}{\subset} X \Rightarrow f(U) \underset{ab}{\subset} Y$ e dizemos que f é fechada se $F \underset{f}{\subset} X \Rightarrow f(F) \underset{f}{\subset} Y$.

2 Compacidade

Neste trabalho, veremos alguns resultados importantes de compacidade, partindo de uma definição intrínseca que descreve a propriedade mais importante de espaços compactos. Após isso, demonstraremos o conhecido Teorema de Heine-Borel, o qual nos diz que um cubo fechado C, em $\mathbb{R}^n$, é compacto. Feito isso, mostraremos algumas aplicações deste Teorema.

Definição 2.1. Um espaço métrico X é chamado compacto se, para toda coleção $\{U_\alpha\}_{\alpha\in\Lambda}$ de conjuntos abertos em X, com $X = \bigcup\limits_{\alpha\in\Lambda} U_\alpha$, for possível encontrar uma subcoleção finita $U_{\alpha_1}, \ldots, U_{\alpha_n}$ com $X = U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$.

Exemplo 2.2. O conjunto $\mathbb{R}$, com a topologia usual, não é compacto.

Demonstração. Considere $\mathbb{R}$ com a topologia usual. Então dada a coleção de abertos $\{]-n,n[\}_{n\in\mathbb{N}}$ tal que $\mathbb{R} = \bigcup\limits_{n\in\mathbb{N}}]-n,n[$, note que para qualquer subcoleção finita

$$\{]-n_1,n_1]\,,]-n_2,n_2]\,,\ldots,]-n_{k_0},n_{k_0}]\,,\ldots,]-n_k,n_k]\}$$

em que $\max\{n_1,n_2,\ldots,n_{k_0},\ldots,n_k\} = n_{k_0}$, temos:

$$]-n_1,n_1[\ \cup\ ]-n_2,n_2[\ \cup\cdots\cup\ ]-n_{k_0},n_{k_0}[\ \cup\cdots\cup\ ]-n_k,n_k[\ =\ ]-n_{k_0},n_{k_0}[\ \neq \mathbb{R}.$$

Portanto, $\mathbb{R}$ não é compacto com a topologia usual. ∎

Lema 2.3. *Se X é um subconjunto compacto de Y, então X é um subconjunto fechado de Y.*

Demonstração. Para mostrarmos que X é um subconjunto fechado de Y, basta provarmos que $Cl_Y(X) = X$. Primeiro mostremos que se $y \in Y - X$ então $y \notin Cl_Y(X)$. De fato, suponha que $y \in Y - X$. Então $y \in Y$ e $y \notin X$. Logo, para cada $x \in X$, $d(x,y) > 0$. Agora, os conjuntos $U_X(x, 1/2\, d(x,y))$, $\forall x \in X$ são abertos em X e $\bigcup\limits_{x\in X} U_X(x, 1/2\, d(x,y)) = X$. Como X é compacto, $X = U_X(x_1, 1/2\, d(x_1,y)) \cup \cdots \cup U_X(x_n, 1/2\, d(x_n,y))$. Então, $\forall x \in X$, $x \in U_X(x_{i_0}, 1/2\, d(x_{i_0},y))$ para algum $1 \le i_0 \le n$. Então,

$$d(x,x_{i_0}) < 1/2\, d(x_{i_0},y),$$

ou equivalentemente,

$$-d(x,x_{i_0}) > -1/2\, d(x_{i_0},y).$$

Pela desigualdade triangular,

$$d(x_{i_0},y) \le d(x,y) + d(x,x_{i_0}),$$

ou equivalentemente,

$$d(x,y) \ge d(x_{i_0},y) - d(x,x_{i_0}) > d(x_{i_0},y) - 1/2\, d(x_{i_0},y)$$

Portanto, temos que

$$d(x,y) > 1/2\, d(x_{i_0},y).$$

Logo,

$$1/2 \min_{1\le i\le n}\{d(x_i,y)\} \le 1/2\, d(x_{i_0},y) < d(x,y),\ \forall x \in X$$

e portanto

$$\begin{aligned} d(y,X) &= \inf\{d(x,y) : x \in X\} \\ &= \text{maior dos limitantes inferiores de } \{d(x,y) : x \in X\} \\ &\ge 1/2 \min_{1\le i\le n}\{d(x_i,y)\} > 0. \end{aligned}$$

Então, pelo Lema 1.13, $y \notin Cl_Y(X)$, $\forall y \in Y - X$. Portanto $Cl_Y(X) \subset X$. Mas $X \subset Cl_Y(X)$ e portanto $Cl_Y(X) = X$ e X é fechado em Y. ∎

Corolário 2.4. *Um subconjunto A de um espaço compacto X é fechado em X se, e somente se, ele é compacto.*

Demonstração. ($\Leftarrow$) Se A é um subconjunto compacto de X, então, pelo Lema anterior, A é um subconjunto fechado de X.
($\Rightarrow$) Se A é um subconjunto fechado em X e $\{U_\alpha\}_{\alpha\in\Lambda}$ é uma coleção de subconjuntos abertos de A, com união A, então para cada $\alpha \in \Lambda$, $(A - U_\alpha)$ é fechado em A, pois U_α é aberto em A. Consequentemente, $(A - U_\alpha)$ é fechado em X, pois $(A - U_\alpha)$ é um subconjunto fechado de A e A é um subconjunto fechado de X. Então

$$X - (A - U_\alpha) = (X - A) \cup U_\alpha \text{ é aberto em } X.$$

Assim,

$$\bigcup_{\alpha\in\Lambda} \{(X - A) \cup U_\alpha\} = (X - A) \cup \bigcup_{\alpha\in\Lambda} U_\alpha = (X - A) \cup A = X.$$

Como X é compacto, podemos escolher $\alpha_1, \dots, \alpha_n$ em Λ tais que

$$X = \bigcup_{1\leq i\leq n} \{(X - A) \cup U_{\alpha_i}\}.$$

Logo,

$$(X - A) \cup A = X = \bigcup_{1\leq i\leq n} \{(X - A) \cup U_{\alpha_i}\} = (X - A) \cup \bigcup_{1\leq i\leq n} U_{\alpha_i}.$$

Portanto,

$$A = \bigcup_{1\leq i\leq n} U_{\alpha_i}$$

e A é compacto. ∎

Teorema 2.5. *Suponha que A e B são subconjuntos fechados e disjuntos de $\mathbb{R}^n$, com A compacto. Então:*

$$d(A, B) = \inf\{d(a, b) : a \in A, b \in B\} > 0.$$

Demonstração. Considere uma coleção $\left\{U_A(a, \frac{1}{2}d(a, B)) : a \in A\right\}$ de subconjuntos abertos de A, cuja reunião é o próprio A. Ou seja,

$$A = \bigcup_{a\in A} U_A(a, \frac{1}{2}d(a, B)).$$

Como A é compacto, então conseguimos extrair uma subcoleção finita

$$\left\{U_A(a_1, \frac{1}{2}d(a_1, B)), \dots, U_A(a_i, \frac{1}{2}d(a_i, B)), \dots, U_A(a_k, \frac{1}{2}d(a_k, B))\right\}$$

tal que

$$A = \bigcup_{1\leq i\leq k} U_A(a_i, \frac{1}{2}d(a_i, B)).$$

Assim, para cada $a \in A$, temos que existe $i \in \{1, \dots, k\}$ tal que $d(a, a_i) < \frac{1}{2}d(a_i, B)$. Desde que $B \underset{f}{\subset} \mathbb{R}^n$, então $B = Cl_{\mathbb{R}^n}(B)$. Como $A \cap B = \emptyset$, $a \notin B$, seque que $a \notin Cl_{\mathbb{R}^n}(B)$ e portanto, pelo Lema 1.13, $d(a, B) > 0, \forall a \in A$. Como $d(a, B) = \inf\{d(a, b) : b \in B\}$, segue que $d(a, B) \leq d(a, b), \forall b \in B$. Daí, como $a \in A$ é qualquer, temos que $d(a, B) \leq d(a, b), \forall a \in A$ e $\forall b \in B$.

Logo, $d(a,B)$ é um limitante inferior para o conjunto $\{d(a,b) : a \in A, b \in B\}$. Como $d(A,B) = \inf\{d(a,b) : a \in A, b \in B\}$, então

$$d(A,B) > d(a,B) > 0.$$

Portanto, temos então que $d(A,B) = \inf\{d(a,b) : a \in A, b \in B\} > 0$. ∎

2.1 Teorema de Heine-Borel e Aplicação

Teorema 2.6 (Teorema de Heine-Borel). *Um cubo fechado C, em $\mathbb{R}^n$, é compacto.*

Demonstração. Vamos supor $\bigcup_{\alpha\in\Lambda} U_\alpha = C$, com cada U_α aberto em C, e que nenhuma subcoleção finita tem união C.

Divida C por hiperplanos paralelos as suas faces em 2^n cubos iguais. Alguns destes cubos menores, C', podem ter a propriedade de encontrarmos um conjunto finito $\alpha_1, \dots, \alpha_k$, com $C' \subset \bigcup_{1\leq i\leq k} U_{\alpha_i}$, mas não todos. Caso contrário, a união do conjunto finito de todos estes U_{α_i} daria C, o que nos diria que existiria uma subcoleção finita com reunião C, o que contraria a nossa hipótese. Portanto, escolhamos um C' sem a propriedade $C' \subset \bigcup_{1\leq i\leq k} U_{\alpha_i}$. Chamemos-no de C_1.

Divida C_1 em 2^n cubos iguais por hiperplanos paralelos as suas faces e, repita o argumento, escolhendo assim, um C_2 sem essa propriedade. Então, por indução, obteremos uma sequência de cubos

$$C \supset C_1 \supset C_2 \supset \cdots \supset C_r \supset C_{r+1} \supset \cdots \tag{2.1}$$

tal que, para cada r, o comprimento de uma aresta de C_r seja 2^{-n} vezes a aresta de C, e não existe nenhum conjunto finito $\alpha_1, \dots, \alpha_k$ tal que $C_r \subset \bigcup_{1\leq i\leq k} U_{\alpha_i}$. A seguir, note que os cubos em (2.1) se intersectam em um único ponto. Chamemos-no de P. Porém, $P \in U_\alpha$, para algum $\alpha \in \Lambda$. Agora, desde que U_α é aberto, nós podemos encontrar um $\varepsilon > 0$, de modo que $U(P,\varepsilon) \subset U_\alpha$. Então, se C tem diâmetro δ, C_n tem diâmetro $2^{-n}\delta$. Consequentemente, se n for tão frande que $2^{-n}\delta < \varepsilon$, então nós encontramos que, $\forall Q \in C_n$,

$$d(P,Q) \leq 2^{-n}\delta < \varepsilon,$$

de modo que $Q \in U(P,\epsilon) \subset U_\alpha$. Logo $C_n \subset U_\alpha$. Porém, isso contradiz a nossa hipótese de que C_n não tenha essa propriedade. Logo, segue que um cubo C fechado em $\mathbb{R}^n$ é compacto. ∎

Teorema 2.7. *Um espaço $X \subset \mathbb{R}^n$ é compacto se, e somente se, ele é fechado e limitado.*

Demonstração. ($\Rightarrow$) Suponhamos que X é compacto. Logo, pelo Lema 2.3, se $X \subset \mathbb{R}^n$ e X é compacto, então X é um subconjunto fechado de $\mathbb{R}^n$. Provemos agora que X é limitado. Sejam,

$$X \cap U(0,m), m > 0$$

subconjuntos abertos de X. Então

$$\bigcup_{m>0} (X \cap U(0,m)) = X \cap \bigcup_{m>0} U(0,m) = X \cap \mathbb{R}^n = X.$$

Consequentemente, conseguimos encontrar uma subcoleção finita $X \cap U(0,m_i)$, $i \in \{1,\dots,k\}$ com união X, isto é,

$$X = (X \cap U(0,m_1)) \ \cup \ \cdots \ \cup \ (X \cap U(0,m_k)) = X \cap \bigcup_{1\leq i\leq k} U(0,m_i) \subset \bigcup_{1\leq i\leq k} U(0,m_i).$$

Então, tomando $N = \max\{m_1, \ldots, m_k\}$, segue que $X \subset U(0, N)$ e portanto X é limitado.

($\Leftarrow$) Agora, suponhamos que X é fechado e limitado. Queremos mostrar que X é compacto. Para tanto, suponha que $X \subset U(0, N)$. Então, X está contido no cubo C dado por $|x_i| \leq N$, $\forall i$. Mas, X é um subconjunto fechado de C. Pelo Teorema 2.6 (Teorema de Heine Borel), temos que C é compacto. Então, pelo Corolário 2.4, como $X \underset{f}{\subset} C$ e C é compacto, então X é compacto. ∎

2.2 Compacidade e Funções

Teorema 2.8. *Se X é um espaço compacto e $f : X \longrightarrow Y$ é uma função contínua e sobrejetora, então Y é compacto.*

Demonstração. Seja U_α uma coleção de subconjuntos abertos de Y tal que $\bigcup\limits_{\alpha\in\Lambda} U_\alpha = Y$. Como f é contínua, pelo Teorema 1.5, $f^{\dashv}(U_\alpha)$ são subconjuntos abertos de X, $\forall \alpha \in \Lambda$.

Além disso, $\forall x \in X$, $f(x) \in Y = \bigcup\limits_{\alpha\in\Lambda} U_\alpha$. Então, por definição de imagem inversa, $\forall x \in X$, $x \in f^{\dashv}(U_\alpha)$ para algum $\alpha \in \Lambda$. Logo, a união desses conjuntos abertos $f^{\dashv}(U_\alpha)$, $\alpha \in \Lambda$ é todo o X. Porém, X é compacto e então $X = \bigcup\limits_{1\leq i\leq n} f^{\dashv}(U_{\alpha_i})$. Assim, para todo $x \in X$ temos que $x \in f^{\dashv}(U_{\alpha_j})$, para algum $j \in \{1, \ldots, n\}$. Como f é sobrejetora então para cada $y \in Y$ existe $x \in X$ tal que $y = f(x)$. Logo, para qualquer $y \in Y$, $y \in U_{\alpha_j}$, para algum $j \in \{1, \ldots, n\}$, o que implica que $y \in \bigcup\limits_{1\leq i\leq n} U_{\alpha_i}$.

Como y é qualquer, então $Y \subset \bigcup\limits_{1\leq i\leq n} U_{\alpha_i}$ Por outro lado, $Y \supset \bigcup\limits_{1\leq i\leq n} U_{\alpha_i}$ desde que $Y = \bigcup\limits_{\alpha\in\Lambda} U_\alpha$. Logo, $Y = \bigcup\limits_{1\leq i\leq n} U_{\alpha_i}$ e portanto Y é compacto. ∎

Lema 2.9. *Se $f : X \longrightarrow Y$ é uma função contínua e bijetora, então $(f^{-1})^{\dashv}(U) = f(U)$.*

Demonstração. Primeiramente, provemos que $(f^{-1})^{\dashv}(U) \subset f(U)$. Para isso, tomemos $z \in (f^{-1})^{\dashv}(U)$. Então, por definição de imagem inversa, $f^{-1}(z) \in U$. Logo, $f(f^{-1}(z)) \in f(U)$, ou seja, $Id_Y(z) \in f(U)$. Consequentemente, $z \in f(U)$ e portanto, $(f^{-1})^{\dashv}(U) \subset f(U)$. Agora, vejamos a outra inclusão. Se $x \in f(U), x = f(u)$, com $u \in U$. Assim, $f^{-1}(x) = f^{-1}(f(u)) = Id_X(u) = u \in U$. Então, $f^{-1}(x) \in U$. Por definição de imagem inversa, segue que $x \in (f^{-1})^{\dashv}(U)$. Portanto, $f(U) \subset (f^{-1})^{\dashv}(U)$. Provamos então que $(f^{-1})^{\dashv}(U) = f(U)$. ∎

Lema 2.10. *Se $f : X \longrightarrow Y$ é uma função sobrejetora então $f(X - F) = Y - f(F)$.*

Demonstração. Mostremos inicialmente que $f(X - F) \subset Y - f(F)$. Se $z \in f(X - F)$, então $z = f(v), v \in X - F$. Agora, se $v \in X - F$, então $z = f(v) \in f(X)$ e $z = f(v) \notin f(F)$. Como f é sobrejetora, então $f(X) = Y$. Logo, $z \in f(X) = Y$ e $z \notin f(F)$. Portanto, $z \in Y - f(F)$ o que implica $f(X - F) \subset Y - f(F)$. Por outro lado, se $x \in Y - f(F)$, $x \in Y$ e $x \notin f(F)$. Como f é sobrejetora, $Y = f(X)$. Assim, $x \in f(X)$ e $x \notin f(F)$. Logo, $x = f(h)$ tal que $h \in X$ e $h \notin F$. Portanto, $h \in X - F$. Consequentemente, $x = f(h) \in f(X - F)$. Logo, $Y - f(F) \subset f(X - F)$. Segue daí que $f(X - F) = Y - f(F)$. ∎

Lema 2.11. *Se $f : X \to Y$ é uma função sobrejetora então f é aberta se, e somente se, f é fechada.*

Demonstração. De fato, se $F \underset{f}{\subset} X$, então $X - F \underset{ab}{\subset} X$. Logo, $f(X - F) \underset{ab}{\subset} Y$. Mas, como f é uma função sobrejetora, pelo Lema 2.10, $f(X - F) = Y - f(F)$. Portanto, $Y - f(F) \underset{ab}{\subset} Y$

e então $f(F) = Y - (Y - f(F)) \underset{f}{\subset} Y$. Reciprocamente, se $U \underset{ab}{\subset} X$, então $X - U \underset{f}{\subset} X$. Logo, $f(X - U) \underset{f}{\subset} Y$. Mas, de modo análogo ao que fizemos no Lema 2.10, sendo f sobrejetora, $f(X - U) = Y - f(U)$ (basta substituir F por U na demonstração deste lema). Assim, $Y - f(U) \underset{f}{\subset} Y$. Portanto, $f(U) = Y - (Y - f(U)) \underset{ab}{\subset} Y$. ■

Corolário 2.12. *Se X é um espaço compacto e $f : X \longrightarrow Y$ é uma função contínua e bijetora então f é um homeomorfismo.*

Demonstração. Para f ser um homeomorfismo, $f^{-1} : Y \longrightarrow X$ deve ser contínua. Para isso, devemos mostrar que para qualquer U aberto de X, $(f^{-1})^{\dashv}(U) \underset{ab}{\subset} Y$. Mas, como f é bijetora, pelo Lema 2.9, temos que $(f^{-1})^{\dashv}(U) = f(U)$. Assim, devemos mostrar que para qualquer U aberto de X, $f(U)$ é aberto de Y, isto é , que f é aberta. Mas, pelo Lema 2.11, isso equivale a provar que f é uma aplicação fechada, isto é, que se F é um subconjunto fechado de X então $f(F)$ é um subconjunto fechado de Y.

Como X é compacto, se $F \underset{f}{\subset} X$, pelo Corolário 2.4, F é compacto. Pelo Teorema 2.8, como F é compacto e $f|_F : F \longrightarrow f(F)$ é uma função contínua e sobrejetora, então $f(F) \subset Y$ é compacto. Pelo Lema 2.3, $f(F) \underset{f}{\subset} Y$. Portanto, f é um homeomorfismo. ■

2.3 Conclusão

É de fácil percepção que ao tratarmos de compacidade, esta noção nos leva a trabalhar conceitos que não nos são tão naturais quanto achávamos que seria ao dizermos que espaços topológicos podem ser vistos como uma expansão da geometria. Um exemplo disso é tratarmos de um cubo, que a princípio, o conhecemos em $\mathbb{R}^3$. Vimos, de maneira abstrata, que um cubo fechado não apenas pode ser definido com n dimensões, mas como também estabelecemos uma regra para ele, ou seja, que ele deve ser compacto em $\mathbb{R}^n$. Ademais, expandimos a ideia de compacidade para funções, no entanto, não paramos por aí. Ao final, conseguimos mostrar que a compacidade de um certo espaço X juntamente com a bijetividade de uma função contínua $f : X \longrightarrow Y$, nos resultam em um homeomorfismo f. Ou seja, os conceitos trabalhados neste artigo são alguns de muitos que existem por aí. Os resultados, aqui apresentados, nos mostram que a Matemática pode ser bem mais ampla e abstrata do que imaginamos.

Agradecimentos: Agradeço primeiramente à minha família e amigos, aos quais sempre posso recorrer. Agradeço também ao Prof. Dr. João Peres Vieira, o qual me orientou durante dois anos e me proporcionou muito conhecimento ao longo desta jornada.

*Abstract: In Mathematics, many are the concepts that sweep their different branches. In particular, in Topology, topological spaces are studied, which can often be seen as an extension of geometry, covering it in a broader and abstract way. According to the work entitled "About the Notion of Compactness" done by Cecília Fernandez and Luiz Silva [2], the concept of compactness was introduced by Maurice Fréchet (1878-1973) in his doctoral thesis, which had been defended in 1906, giving the concept of compactness a topological extension of the ideas of finitude and limitation. From that, several mathematicians started to develop this subject. Among the many who stood out, Heinrich Eduard Heine and Félix Édouard Justin Émile Borel were responsible for demonstrating a well-known theorem in the area of Topology, called **Heine-Borel Theorem**. In this work, we will demonstrate and apply this theorem, as well as obtain results on compactness and functions.*

Keywords: Closed Set; Compactness; Continuity; Heine-Borel; Homeomorphism; Open Set; Topology;

Referências Bibliográficas

[1] Wall, C.T.C., *A geometric Introduction to Topology*, Addison-Wesley Publishing Company, 1972.

[2] Fernandez, C.S., Silva, L.A.V., *Sobre a Noção de Compacidade*, Universidade Federal Fluminense, Setembro de 2014.

[3] Lima, E.L., *Espaços Métricos*, IMPA, 1993.

[4] Lima, E.L., *Elementos de Topologia Geral*, IMPA, Ao Livro Técnico S.A., 1970.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 60–67

Sistemas de Amortização: elaborando as tabelas SAC e Price

Lucas Lopes Gueiros de Souza[†]
Orientador(a): Prof. Dr. Lucas Carato Mazzi

Resumo: Neste trabalho desenvolveremos a ideia de amortização em Matemática Financeira e a aplicaremos em alguns problemas do cotidiano. Amortização, de maneira geral, é o pagamento de uma dívida por meio de prestações em um prazo pré-estipulado. Apresentaremos, ao longo deste texto, dois sistemas de amortização comuns, a saber: o Sistema de Amortização Constante (SAC) e um variante do Sistema de Amortização Francês, o chamado Sistema Price. De modo geral, temos que no SAC o valor da prestação é decrescente e o valor da amortização se mantém constante. Já no Price, a prestação é fixa, enquanto a amortização é crescente.

Palavras-chave: Matemática Financeira; Sistema de Amortização Constante; Sistema de Amortização Francês; Price

1 Introdução

O sistema de amortização é desenvolvido basicamente para os negócios de empréstimos e financiamentos de longo prazo, envolvendo pagamentos regulares de principal e despesas financeiras.

Existem vários métodos de reembolso de dívidas, e os termos de cada transação devem ser determinados no contrato celebrado entre o credor (a quem ou a que se deve dinheiro) e o devedor.

A característica básica do sistema de amortização a ser estudado neste trabalho é o uso exclusivo de padrões de juros compostos, e os juros são dedicados ao saldo devedor apurado no período anterior.

Neste trabalho, trataremos de dois sistemas de amortização:

1. Sistema de Amortização Constante (SAC)

2. Sistema de Amortização Francês (Price)

2 Definições

Amortização (A) é o pagamento de uma dívida, geralmente, por meio de prestações periódicas em um prazo pré-estipulado.

A prestação (R) é composta pelo valor da amortização mais os encargos financeiros devidos (juros, impostos e outras possíveis taxas). Temos que essas prestações podem ser fixas ou variáveis, de acordo com o sistema utilizado.

O saldo devedor (SD) representa o valor do capital emprestado, em um dado momento, após a dedução do valor já pago ao credor a título de amortização.

Na sequência, elencamos algumas notações necessárias para a compreensão do trabalho.

- R - Valor da prestação;

[†]Bolsista PET-SESu/MEC

- P - Capital a ser amortizado ou valor presente da dívida;
- i - Taxa unitária de juro;
- J - Juros do período;
- A - Amortização do período;
- n - Número de períodos do financiamento/empréstimo.

3 Sistemas de Amortização

3.1 Sistema de Amortização Constante (SAC)

Também conhecido como Sistema de Amortização Hamburguês, esse sistema diz respeito às amortizações constantes em prestações periódicas, sucessivas e decrescentes em progressão aritmética, em que o valor da prestação é composto por uma parte devida aos juros e a outra à amortização (VIEIRA SOBRINHO, 2000).

Dessa forma, o valor da prestação é decrescente, porque os juros incidem somente sobre o saldo devedor restante após a dedução do pagamento da parcela mensal referente ao capital. Então o que muda é o valor decorrente dos juros aplicados sobre o Saldo Devedor (SD) do período antecedente (ZANNA, 2011).

Temos assim que a equação que corresponde ao SAC equivale a:

$$Rn_n = A + P \cdot i - (n - 1) \cdot A \cdot i. \tag{3.1}$$

Sendo:

- Rn – O valor da prestação em cada período;
- A – Amortização do período;
- P – Capital a ser amortizado ou Valor presente da dívida;
- i – Taxa unitária de juro;
- n – Número de períodos de financiamento.

Nesse caso, o valor das prestações iniciais são maiores, necessitando que o devedor tenha um valor maior para desembolsar inicialmente, o qual com o passar do tempo temos que as prestações vão ficando menores.

Exemplo 3.1. Iremos montar uma planilha financeira do Sistema de Amortização Constante para um empréstimo de R\$10.000 a 3% ao mês, com prazo de 12 meses sendo amortizável em 12 prestações mensais.

A Tabela 9.1 abaixo representa todos os dados que obtemos.

Para elaborar essa tabela seguimos três passos, sendo eles:

1. Calcular a amortização, que será igual ao Saldo Devedor (SD) do período zero que em outras palavras, o valor que foi financiado dividido pelo número de parcelas:

$$A = \frac{SD_0}{n},\ n \geq 0.$$

Tabela 9.1: Elaboração da tabela SAC

Mês	Prestação (R$)	Juros (R$)	Amortização (R$)	Saldo Devedor (R$)
0				10.000,00
1	1133,33	300,00	833,33	9.167,67
2	1108,33	275,00	833,33	8.333,34
3	1083,33	250,00	833,33	7.500,01
4	1058,33	225,00	833,33	6.667,68
5	1033,33	200,00	833,33	5.833,35
6	1008,33	175,00	833,33	5.000,02
7	983,33	150,00	833,33	4.167,69
8	958,33	125,00	833,33	3.333,36
9	933,33	100,00	833,33	2.500,03
10	908,33	75,00	833,33	1.667,70
11	883,33	50,00	833,33	833,37
12	858,33	25,00	833,33	0,00

2. Calcular os juros em termos monetários, que será a taxa de juros multiplicada pelo Saldo Devedor no período anterior:

$$J_n = SD_{n-1} \cdot i, \ n \geq 1.$$

3. Calcular a parcela (ou prestação), que será simplesmente a soma dos juros com a amortização:

$$Rn = J_n + A.$$

Inicialmente, devemos encontrar o valor da amortização. Como estamos tratando com tabela SAC, utilizaremos o seguinte cálculo:

$$A = \frac{SD_0}{n} \Rightarrow A = \frac{10.000}{12} \Rightarrow A = 833,33.$$

Esse valor, R$ 833,33, representa a amortização e, como o SAC propõe, será fixo (constante) em todos os 12 meses de pagamentos. Após esse cálculo, calcularemos na sequência o valor dos juros para o primeiro período.

$$J_n = SD_{n-1} \cdot i \Rightarrow J_1 = SD_0 \cdot i \Rightarrow J_1 = 10.000 \cdot 0,03 \Rightarrow J_1 = 300,00.$$

Com o valor dos Juros, agora podemos obter o valor da prestação da seguinte forma

$$Rn_n = J_n + A \Rightarrow Rn_1 = J_1 + A \Rightarrow Rn = 300,00 + 833,33 \Rightarrow Rn = 1133,33.$$

Com esse valor temos todos os dados para colocar no primeiro mês, agora, indo para o segundo mês temos que a amortização será de R$ 833,33 e que o novo Saldo Devedor será feito a partir da diferença entre o Saldo Devedor anterior e a amortização ficando

$$SD_n = SD_{n-1} - A \Rightarrow SD_1 = SD_0 - A \Rightarrow SD_n = 10.000 - 833,33 \Rightarrow SD_n = 9.167,00.$$

Agora com o Saldo Devedor no segundo mês, seguimos para encontrar o valor dos Juros no segundo mês, ficando

$$J_n = SD_{n-1} \cdot i \Rightarrow J_2 = SD_1 \cdot i \Rightarrow J_2 = 9.167 \cdot 0,03 \Rightarrow J_2 = 275,00.$$

Com o valor de Juros, podemos encontrar o valor da parcela no segundo mês

$$Rn_n = J_n + A \Rightarrow Rn_2 = J_2 + A \Rightarrow Rn_2 = 275,00 + 833,33 \Rightarrow Rn_2 = 1108,33.$$

Agora por último, encontraremos o novo valor do Saldo Devedor

$$SD_n = SD_{n-1} - A \Rightarrow SD_2 = SD_1 - A \Rightarrow SD_2 = 9.167 - 833,33 \Rightarrow SD_2 = 8.333,00.$$

O cálculo dos outros meses é feita de maneira análoga ao mês 1 e ao mês 2.

3.2 Sistema Price

Também conhecido como Sistema de Amortização Francês que consiste em um plano de amortização de uma dívida em prestações periódicas, iguais e sucessivas, dentro do conceito de termos vencidos, em que o valor de cada prestação, é composto por duas parcelas distintas: uma de juros e outra de capital (VIEIRA SOBRINHO, 2000).

O que diferencia o Sistema Price do SAC é que no SAC a Amortização é constante e o valor das parcelas diminuem com o passar do tempo, enquanto no Sistema Price, a Amortização aumenta com o passar do tempo e o valor das prestações são constantes.

Esse método de amortização implica na capitalização mensal dos juros. Os valores fixos da prestação são obtidos pela divisão do fator de amortização. Portanto, é considerado uma série de pagamentos uniformes, para ratear o pagamento de uma dívida em prestações mensais, iguais e consecutivas, agregando-se juros ao capital acordado (ZANNA, 2011).

A expressão que representa o Tabela Price (TP), é chamada de “Fator de Capitalização”, sendo dada a seguir.

$$R = P \cdot \left[\frac{i \cdot (1+i)^n}{(1+i)^n - 1}\right]. \tag{3.2}$$

Sendo:

- R – Valor da Prestação;
- P – Capital a ser Amortizado ou Valor Presente da Dívida;
- i – Taxa unitária de juro;
- n – Número de períodos de financiamento.

É importante observar durante as contas qual é a taxa de Juros utilizada, se é a Taxa Nominal ou Taxa Efetiva. A Taxa Efetiva tem incidência somente uma vez em cada prestação enquanto a Taxa Nominal tem incidência mais de uma vez a cada prestação.

É importante ressaltar que há autores que consideram dois Sistemas de Amortização, o Sistema de Amortização Francês e o TP. Segundo eles, no primeiro a taxa utilizada é Efetiva enquanto no segundo a taxa utilizada é Nominal.

Exemplo 3.2. Iremos construir a planilha financeira da Tabela Price para um empréstimo de R$10.000 a 3% ao mês, com prazo de 12 meses sendo amortizável em 12 prestações mensais.

A tabela a seguir representa todos os dados que obtemos:

(espaço em branco intencional)

Tabela 9.2: Elaboração da tabela Price

Mês	Prestação (R$)	Juros (R$)	Amortização (R$)	Saldo Devedor (R$)
0				10.000,00
1	1004,62	300,00	704,62	9.295,38
2	1004,62	278,86	725,76	8.569,62
3	1004,62	257,09	747,53	7.822,09
4	1004,62	234,66	769,96	7.052,13
5	1004,62	211,56	793,06	6.259,07
6	1004,62	187,77	816,85	5.442,22
7	1004,62	163,27	841,35	4.600,87
8	1004,62	138,03	866,59	3.734,27
9	1004,62	112,03	892,59	2.841,68
10	1004,62	85,25	919,37	1.922,31
11	1004,62	57,67	946,95	975,36
12	1004,62	29,26	975,36	0,00

Para a construção dessa tabela seguimos 3 passos, sendo eles:

1. Calcular a parcela (ou prestação), que será simplesmente a soma dos juros com a amortização:
$$R = P \cdot \left[\frac{i \cdot (1+i)^n}{(1+i)^n - 1}\right].$$

2. Calcular os juros em termos monetários, que será a taxa de juros multiplicada pelo saldo devedor no período anterior:
$$J_n = SD_{n-1} \cdot i.$$

3. Por fim, calcular a amortização, que será igual ao valor da prestação menos os juros do período:
$$A_n = R - J_n.$$

Abrindo as contas por trás dos dados da tabela temos que, no mês zero, ou seja, o momento em que é feito o empréstimo temos que o Saldo Devedor é de R$ 10.000,00 indo para o mês 1, temos já o primeiro pagamento da dívida no valor de R$ 1.004,62 e que agora a dívida está no valor de R$ 9.295,00, o passo a passo deste primeiro mês é:

Primeiro vamos encontrar o valor das parcelas para pagar, como estamos tratando com o Sistema Price, o valor dessas parcelas serão as mesmas para todos meses, tendo assim

$$R = P \cdot \left[\frac{i \cdot (1+i)^n}{(1+i)^n - 1}\right] \Rightarrow R = 10.000,00 \cdot \left[\frac{0,03 \cdot (1,03)^{12}}{(1,03)^{12} - 1}\right]$$
$$\Rightarrow R = 1.004,62.$$

Agora que temos o valor das parcelas, vamos calcular os Juros neste mês:

$$J_n = SD_{n-1} \cdot i \Rightarrow J_1 = SD_0 \cdot i \Rightarrow J_1 = 10.000,00 \cdot 0,03 \Rightarrow J_1 = 300,00.$$

Agora, usando o valor dos juros, podemos calcular o valor da amortização e obter

$$A_n = R - J_n \Rightarrow A_1 = R - J_1 \Rightarrow A_1 = 1004,62 - 300,00 \Rightarrow A_1 = 704,62.$$

Com a amortização podemos fazer a diferença entre o Saldo Devedor e ela, encontrando assim o novo valor da dívida

$$SD_n = SD_{n-1} - A_n \Rightarrow SD_1 = SD_0 - A_1 \Rightarrow SD_1 = 10.000 - 704,62 \Rightarrow SD_1 = 9.295,00$$

com isso conseguimos encontrar os dados do segundo mês mas agora não precisamos encontrar o valor da parcela já que como falado anteriormente, as prestações utilizadas no Sistema Price são constantes para todos os meses, ou seja,

$$R = 1004,62.$$

Com o valor da prestação, podemos encontrar o valor dos Juros, sendo ele

$$J_n = SD_{n-1} \cdot i \Rightarrow J_2 = SD_1 \cdot i \Rightarrow J_2 = 9.295 \cdot 0,03 \Rightarrow J_2 = 278,86.$$

Por meio do valor dos juros, podemos obter o valor da amortização

$$A_n = R - J_n \Rightarrow A_2 = R - J_2 \Rightarrow A_2 = 1004,62 - 278,86 \Rightarrow A_2 = 725,76.$$

Agora com essa amortização conseguimos obter o Saldo Devedor do terceiro mês a partir de

$$SD_n = SDn - 1 - A_n \Rightarrow SD_2 = SD1 - A_2 \Rightarrow SD_2 = 9.295 - 725,76 \Rightarrow SD_2 = 8.570.$$

Com esses dados para os próximos meses a maneira de se fazer é análoga às anteriores o qual devemos fazer até o Saldo Devedor ser igual ao valor da Amortização que significará que acabou a dívida.

4 Outros Exemplos

Exemplo 4.1. Para o próximo caso, vamos supor que Janete fez um empréstimo de R$ 8.000,00 para ser pago em 5 parcelas mensais e consecutivas, à taxa de 5% a.m., pelo sistema de amortização constante (SAC). O valor da última parcela a ser paga por Janete é igual a:

a) R$ 1.760,00.

b) R$ 1.680,00.

c) R$ 1.720,00.

d) R$ 1.700,00.

e) R$ 1.600,00.

Antes temos que saber o valor de cada parcela sem o acréscimo do juros, para isso devemos somente dividir o valor total da dívida pelo número de parcelas, ficando:

$$A = \frac{8.000,00}{5} \Rightarrow A = 1.600,00.$$

Com esse dado agora podemos utilizar a relação (5.1), tendo assim:

$$Rn = A + P \cdot i - (n-1) \cdot A \cdot i \Rightarrow Rn = 1.600,00 + 8.000,00 \cdot 0,05 - (5-1) \cdot 1.600,00 \cdot 0,05$$

$$\Rightarrow Rn = 2.000,00 - 320 \Rightarrow Rn = 1.680,00.$$

Exemplo 4.2. Nesse caso, temos que uma geladeira é vendida, pelo pagamento à vista, por R$ 1.200,00. Caso o consumidor queira parcelar o eletrodoméstico, será cobrada uma taxa de juros no valor de 2% ao mês. Vamos considerar que uma pessoa comprou a geladeira em 12 prestações iguais e queremos o valor das parcelas utilizando o sistema Price. Para isso iremos utilizar a relação (5.2), tendo:

$$R = P \cdot \left[\frac{i \cdot (1+i)^n}{(1+i)^n - 1}\right] \Rightarrow R = 1.200 \cdot \left[\frac{0,02 \cdot (1+0,02)^{12}}{(1+0,02)^{12} - 1}\right]$$

$$\Rightarrow R = 1.200 \cdot \left[\frac{1,268 \cdot 0,02}{1,268 - 1}\right] \Rightarrow R = 1.200 \cdot \frac{0,025}{0,26}$$

$$\Rightarrow R = 1.200 \cdot 0,0945 \Rightarrow R = 113,47.$$

Exemplo 4.3. Vamos supor o financiamento de um imóvel no valor de R$120.000,00 e entrada de $R$$ 20.000,00, com um prazo de financiamento de 24 parcelas e juros de 10,6% ao ano, o que dá uma taxa de 0,883% ao mês. Para isso utilizaremos a equação (5.3), tendo então:

$$R = \left[\frac{P}{n}\right] + P \cdot i \Rightarrow R = \left[\frac{100.000,00}{24}\right] + 100.000,00 \cdot 0,00883$$

$$\Rightarrow R = 4.166,66 + 883 \Rightarrow R = \$5.049,66.$$

As 12 primeiras parcelas pelo SACRE terão o valor de R$5.049,67. Para as próximas 12 parcelas, o cálculo é repetido, mas sobre o saldo devedor de R$51.535,22 ficando da seguinte forma:

$$R = \left[\frac{P}{n}\right] + P \cdot i \Rightarrow R = \left[\frac{51.535,22}{12}\right] + 51.535,22 \cdot 0,00883$$

$$\Rightarrow R = 4.294,60 + 455,05 \Rightarrow R = 4.749,65.$$

5 Algumas considerações

O SAC é o sistema mais utilizado nos financiamentos e empréstimos para imóveis, tem um valor inicial de prestações maior do que o sistema Price, mas que, com o passar do tempo, acaba sendo a opção com menor valor de prestação, sendo preferível assim para as pessoas que querem parcelas mais baratas ao final do financiamento. Nesse primeiro caso, por exemplo, deve-se ressaltar que como os valores das parcelas iniciais são maiores, tenha em mente que a fonte de sua renda deve ser inicialmente estável ou com uma perspectiva salarial boa para não ter problemas com atrasos e até mesmo com inadimplência,

No segundo caso discutido, temos que a Tabela Price é mais comum de ser vista em amortização para veículos, no entanto pode ser utilizada no lugar do SAC em empréstimos para imóveis também. Como o valor das prestações nesse sistema são constantes, não há surpresas ao longo do pagamento, no entanto, também não há redução de valores. Em relação às prestações, temos que em relação ao SAC, as prestações do Sistema Price começa sendo menor, porém com o passar do tempo ela acaba ficando maior.

Vale destacar, por fim, que a escolha pelo sistema de amortização utilizado deve ser feito a partir da realidade de cada pessoa.

Agradecimentos: Gostaria de agradecer a minha família e amigos pelo apoio e incentivo durante a realização desse trabalho, ao meu orientador Dr. Lucas Carato Mazzi por ter me apoiado academicamente com todas as dificuldades encontrada durante o projeto, também gostaria de agradecer o grupo PET-Matemática pelo apoio financeiro e profissional.

Abstract: In this work we will develop the idea of amortization in Financial Mathematics and apply it in some problems of everyday life. Amortization, in general, is the payment of a debt through installments in a pre-stipulated period. Throughout this text, we will present two common amortization systems, i.e. the Constant Amortization System (SAC) and a variant of the French Amortization System, the so-called Price System. In general, we have that in the SAC the value of the installment is decreasing and the amortization value remains constant. In Price, the installment is fixed, while the amortization is increasing.

Keywords: Financial Mathematics; Constant Amortization System; French Amortization System; Price

Referências Bibliográficas

[1] ASSAF NETO, A. *Matemática Financeira e suas aplicações*, - 8. ed. - São Paulo : Atlas, 2003.

[2] JÚNIOR, M. J. G. *Comparação dos Sistemas Bancários de Financiamento na Aquisição de Automóveis e Imóveis*, 2018. 87 p. Dissertação (Mestrado em Engenharia Automotiva) - Universidade Federal de Santa Catarina, Centro Tecnológico de Joinville., Joinville.

[3] PUCCINI, A. L., *Matemática Financeira - Objetiva e Aplicada*, Editora Saraiva, São Paulo, 7ª Edição, 2006.

[4] SOUZA, F. H. A. *Matemática Financeira: Uma Importante Ferramenta no Cotidiano*, 2017. 51 p. Dissertação (Mestrado em Matemática) - Universidade Federal de Goiás Instituto de Matemática e Estatística, Goiânia.

[5] VIEIRA, J. D. S., DUTRA, J. *Matemática financeira*, Atlas, Sao Paulo, 1997.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 68–83

Teorema do Ponto Fixo de Brouwer e Suas Equivalências

Lucas Ozaki Mizuguti†
Orientador(a): Profa. Dra. Suzete Maria Silva Afonso

Resumo: Neste artigo apresentamos uma prova do Teorema do Ponto Fixo de Brouwer utilizando a teoria do grau topológico de Brouwer. Além disso, mostramos algumas equivalências entre o Teorema do Ponto Fixo de Brouwer e outros resultados conhecidos da teoria de Análise Funcional, tais como o Teorema de Knaster–Kuratowski–Mazurkiewicz, o Teorema de Hartman–Stampachia e a Desigualdade do min–max de Ky–Fan.

Palavras-chave: Teorema do Ponto Fixo de Brouwer; teoria do grau topológico de Brouwer; Análise Funcional; Teorema de Knaster–Kuratowski–Mazurkiewicz; Teorema de Hartman–Stampachia; Desigualdade do min–max de Ky–Fan.

Introdução

Os teoremas de ponto fixo são ferramentas fundamentais para a obtenção de resultados de existência de soluções para problemas matemáticos. Dentre os teoremas de ponto fixo, destaca-se o Teorema do Ponto Fixo de Brouwer pela sua relevante contribuição em áreas da matemática, como análise [2, 5, 10], topologia [5, 7], álgebra [7], lógica [3], teoria dos jogos [9], e em outras áreas do conhecimento, como biologia e economia, veja [6] e [12], por exemplo.

Um ponto fixo de uma função é um ponto de seu domínio que não se altera pela sua aplicação, ou seja, $x \in X$ é ponto fixo de uma aplicação $f\colon X \to X$ se $f(x) = x$. O Teorema do Ponto Fixo de Brouwer garante que toda aplicação contínua de $\overline{B_1(0)}$ em $\overline{B_1(0)}$ admite pelo menos um ponto fixo, onde $\overline{B_1(0)} = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ e $\| \cdot \|$ é uma norma em $\mathbb{R}^n$, e pode ser generalizado de forma a assegurar que toda aplicação contínua de um subconjunto convexo e compacto de $\mathbb{R}^n$ em si mesmo admite pelo menos um ponto fixo.

Este trabalho destina-se a apresentar uma prova do Teorema do Ponto Fixo de Brouwer utilizando a teoria do grau de Brouwer e a demonstrar alguns resultados equivalentes a ele, como os conhecidos Teorema de Knaster–Kuratowski–Mazurkiewicz, Teorema de Hartman–Stampachia e Desigualdade do $\min - \max$ de Ky–Fan.

Dividimos este artigo em três seções. A primeira seção será dedicada a uma breve introdução da teoria do grau de Brouwer, ferramenta topológica a ser utilizada para demonstrar o Teorema do Ponto Fixo de Brouwer. Na segunda seção, enunciaremos e demonstraremos o Teorema do Ponto Fixo de Brouwer, bem como a versão generalizada dele. Por fim, a terceira seção será dedicada às equivalências entre o Teorema do Ponto Fixo de Brouwer e outros resultados da teoria de Análise Funcional.

1 Sobre o grau de Brouwer

Nesta seção apresentaremos de forma breve a teoria básica do grau topológico de Brouwer, ferramenta topológica fundamental para a prova do Teorema do Ponto Fixo de Brouwer que exibiremos. As principais referências para a teoria exibida aqui são [1, 2, 11].

†Bolsista PET-SESu/MEC

Consideraremos $\Omega \subset \mathbb{R}^n$ um subconjunto aberto e limitado e $C^k(\overline{\Omega}, \mathbb{R}^n)$, $k \geqslant 1$, o espaço das aplicações k-vezes diferenciáveis em $\overline{\Omega}$, ou seja, o espaço das aplicações diferenciáveis em $\overline{\Omega}$ que possuem todas as suas derivadas até ordem k, sendo as derivadas restrições de aplicações contínuas definidas em $\overline{\Omega}$. O símbolo $C(\overline{\Omega}, \mathbb{R}^n)$ denotará o espaço das aplicações contínuas $\varphi\colon \overline{\Omega} \to \mathbb{R}^n$.

Consideraremos o espaço $C^k(\overline{\Omega}, \mathbb{R}^n)$ munido com a norma $\| \ \|_k\colon C^k(\overline{\Omega}, \mathbb{R}^n) \to \mathbb{R}_+$ definida por

$$\|\varphi\|_k = \max_{0 \leq j \leq k} \sup_{x \in \Omega} \|D^j\varphi(x)\|, \quad \text{para } \varphi \in C^k(\overline{\Omega}, \mathbb{R}^n),$$

onde $D^j\varphi$ representa a derivada de ordem j de φ, com $j \in \{1, \ldots, k\}$, e $\| \cdot \|$ é uma norma em $\mathbb{R}^n$.

Sejam $S = \{x \in \Omega : J_\varphi(x) = 0\}$, onde J_φ representa o determinante da matriz jacobiana de φ, e $b \in \mathbb{R}^n$ tal que

$$b \notin \varphi(\partial\Omega) \cup \varphi(S), \tag{1.1}$$

onde $\partial\Omega$ denota a fronteira de Ω.

Note que se $x \in \varphi^{-1}(\{b\})$, então $J_\varphi(x) \neq 0$ por (1.1). Assim, $\varphi'(x)$ é invertível. Então, pelo Teorema da Aplicação Inversa, φ é um difeomorfismo de uma vizinhança U_x de x sobre uma vizinhança V_b de b, ou seja, $\varphi|_{U_x}\colon U_x \to \varphi(U_x) = V_b$ é um difeomorfismo.

Afirmação: O conjunto $\varphi^{-1}(\{b\})$ é compacto e finito.

De fato, como $\varphi^{-1}(\{b\}) \subset \overline{\Omega} \subset \mathbb{R}^n$ e Ω é limitado (e portanto, $\overline{\Omega}$ também é limitado), temos que $\varphi^{-1}(\{b\})$ é limitado. Além disso, $\varphi^{-1}(\{b\})$ é fechado, já que φ é contínua e $\{b\}$ é um conjunto fechado em $\mathbb{R}^n$. Portanto, $\varphi^{-1}(\{b\})$ é compacto.

Agora, para cada $x \in \varphi^{-1}(\{b\})$, considere a bola aberta de centro x e raio $r_x > 0$, $B_{r_x}(x) \subset U_x$. Assim,

$$\varphi^{-1}(\{b\}) \subset \bigcup_{x_j \in \varphi^{-1}(\{b\})} B_{r_j}(x_j).$$

Dessa forma, $\cup_{x_j \in \varphi^{-1}(\{b\})} B_{r_j}(x_j)$ é uma cobertura para o compacto $\varphi^{-1}(\{b\})$, constituída por conjuntos abertos. O Teorema de Borel–Lebesgue garante que $\varphi^{-1}(\{b\})$ admite um subcobertura finita, digamos

$$\varphi^{-1}(\{b\}) \subset \bigcup_{j=1}^{k} B_{r_j}(x_j).$$

Como $\varphi|_{U_{x_j}}: U_{x_j} \to \varphi(U_{x_j})$ é um difeomorfismo para cada j, concluímos que

$$\varphi^{-1}(\{b\}) = \{x_1, x_2, \ldots, x_k\},$$

de onde segue que $\varphi^{-1}(\{b\})$ é um conjunto finito.

A partir dessas considerações iniciais, podemos definir o grau topológico de Brouwer.

Definição 1.1. Sejam $\varphi \in C^1(\overline{\Omega}, \mathbb{R}^n)$ e $b \notin \varphi(\partial\Omega) \cup \varphi(S)$. Definimos o grau topológico de Brouwer da aplicação φ em relação a Ω no ponto b como sendo o número inteiro

$$d(\varphi, \Omega, b) = \sum_{\xi_i \in \varphi^{-1}(\{b\})} \operatorname{sgn}(J_\varphi(\xi_i)), \tag{1.2}$$

onde sgn é a função sinal, definida por:

$$\begin{aligned} \operatorname{sgn}\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto \operatorname{sgn}(t) = \begin{cases} 1, & \text{se } t > 0, \\ 0, & \text{se } t = 0, \\ -1, & \text{se } t < 0. \end{cases} \end{aligned}$$

Se $b \notin \varphi(\overline{\Omega})$, definimos $d(\varphi, \Omega, b) = 0$.

Observação 1.2. Segue da definição do grau topológico de Brouwer a validade da igualdade:

$$d(\varphi, \Omega, b) = d(\varphi - b, \Omega, 0).$$

A fim de elucidar a Definição 1.1, apresentamos um exemplo a seguir.

Exemplo 1.3. Seja $\varphi \colon \Omega \to \mathbb{R}$ a função diferenciável dada por $\varphi(x) = \text{sen}(x)$, onde $\Omega = \left(0, \frac{5\pi}{2}\right)$ e $b = \frac{\sqrt{2}}{2}$. Calculemos o grau topológico de Brouwer de φ em relação a Ω em b.

Sabemos que $\varphi \in C^1(\overline{\Omega}, \mathbb{R})$. Precisamos mostrar que $d(\varphi, \Omega, b)$ está bem definido, ou seja, que $b \notin \varphi(\partial\Omega) \cup \varphi(S)$. De fato:

- como $\Omega = \left(0, \frac{5\pi}{2}\right)$, temos $\partial\Omega = \left\{0, \frac{5\pi}{2}\right\}$ e, portanto, $\varphi(\partial\Omega) = \{0, 1\}$;
- como $S = \{x \in \Omega : J_\varphi(x) = 0\} = \left\{x \in \left(0, \frac{5\pi}{2}\right) : \cos(x) = 0\right\} = \left\{\frac{\pi}{2}, \frac{3\pi}{2}\right\}$, segue que $\varphi(S) = \{-1, 1\}$.

Consequentemente, $\varphi(\partial\Omega) \cup \varphi(S) = \{-1, 0, 1\}$ e $b = \frac{\sqrt{2}}{2} \notin \varphi(\partial\Omega) \cup \varphi(S)$, provando que $d(\varphi, \Omega, b)$ está bem definido. Ademais,

$$\varphi^{-1}\left(\left\{\frac{\sqrt{2}}{2}\right\}\right) = \left\{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{9\pi}{4}\right\}.$$

Por definição, temos

$$d\left(\text{sen}(x), \left(0, \frac{5\pi}{2}\right), \frac{\sqrt{2}}{2}\right) = \sum_{\xi_i \in \varphi^{-1}\left(\left\{\frac{\sqrt{2}}{2}\right\}\right)} \text{sgn}(J_\varphi(\xi_i)),$$

ou seja,

$$\begin{aligned} d\left(\text{sen}(x), \left(0, \frac{5\pi}{2}\right), \frac{\sqrt{2}}{2}\right) &= \text{sgn}(\varphi'(\xi_1)) + \text{sgn}(\varphi'(\xi_2)) + \text{sgn}(\varphi'(\xi_3)) \\ &= \text{sgn}(\cos(\xi_1)) + \text{sgn}(\cos(\xi_2)) + \text{sgn}(\cos(\xi_3)), \end{aligned}$$

onde $\xi_1 = \frac{\pi}{4}$, $\xi_2 = \frac{3\pi}{4}$ e $\xi_3 = \frac{9\pi}{4}$.

Portanto,

$$\begin{aligned} d\left(\text{sen}(x), \left(0, \frac{5\pi}{2}\right), \frac{\sqrt{2}}{2}\right) &= \text{sgn}\left(\cos\left(\frac{\pi}{4}\right)\right) + \text{sgn}\left(\cos\left(\frac{3\pi}{4}\right)\right) + \text{sgn}\left(\cos\left(\frac{9\pi}{4}\right)\right) \\ &= 1 + (-1) + 1 = 1 \end{aligned}$$

ou seja,

$$d\left(\text{sen}(x), \left(0, \frac{5\pi}{2}\right), \frac{\sqrt{2}}{2}\right) = 1.$$

Na Definição 1.1 apresentamos a definição do grau topológico de Brouwer para uma aplicação φ de classe C^1 com relação a Ω em b, em que $b \notin \varphi(\partial\Omega) \cup \varphi(S)$. Entretanto, a definição do grau pode ser estendida para casos mais gerais, mais especificamente, para aplicações $\varphi \in C(\overline{\Omega}, \mathbb{R}^n)$, em que $b \notin \varphi(\partial\Omega)$, conforme veremos adiante.

A seguir, enunciaremos uma série de resultados auxiliares para a definição do grau para aplicações contínuas e suas propriedades. Convém mencionarmos que o leitor interessado pode consultar [2] para verificar as provas destes resultados.

Proposição 1.4. *Se $\varphi \in C^1(\overline{\Omega}, \mathbb{R}^n)$ e $b \notin \varphi(\partial\Omega) \cup \varphi(S)$, então existe uma vizinhança U de φ na topologia de $C^1(\overline{\Omega}, \mathbb{R}^n)$ tal que para qualquer $\psi \in U$ tem-se:*

(i) $b \notin \psi(\partial\Omega)$;

(ii) $x \in \psi^{-1}(\{b\}) \Rightarrow J_\psi(x) \neq 0$;

(iii) $d(\psi, \Omega, b) = d(\varphi, \Omega, b)$.

Atentamos para o fato de que o item (iii) da Proposição 1.4 fornece a informação de que o grau topológico de Brouwer é localmente constante na topologia de $C^1(\overline{\Omega}, \mathbb{R}^n)$.

O próximo resultado diz que o grau de Brouwer é constante em componentes conexas de $\mathbb{R}^n \setminus \varphi(\partial\Omega)$, quando consideramos $\varphi \in C^2(\overline{\Omega}, \mathbb{R}^n)$.

Proposição 1.5. *Sejam $\varphi \in C^2(\overline{\Omega}, \mathbb{R}^n)$ e b_1, b_2 pontos de $\mathbb{R}^n$ que não pertencem a $\varphi(\partial\Omega) \cup \varphi(S)$. Se b_1 e b_2 estão na mesma componente conexa de $\mathbb{R}^n \setminus \varphi(\partial\Omega)$, tem-se*

$$d(\varphi, \Omega, b_1) = d(\varphi, \Omega, b_2).$$

Seja C_b a componente conexa de $\mathbb{R}^n \setminus \varphi(\partial\Omega)$ que contém b. Como $\mathbb{R}^n \setminus \varphi(\partial\Omega)$ é aberto, então C_b também é. Por outro lado, pelo Teorema de Sard,[††] $\varphi(S)$ tem medida nula em $\mathbb{R}^n$. Dessa forma, podemos concluir que $\varphi(S)$ não contém C_b e podemos inferir que $C_b \setminus \varphi(S) \neq \emptyset$. Este fato, juntamente com a Proposição 1.5, nos permite definir o grau topológico $d(\varphi, \Omega, b)$ quando φ pertence a $C^2(\overline{\Omega}, \mathbb{R}^n)$ e b é imagem de um ponto crítico de φ.

Definição 1.6. Sejam $\varphi \in C^2(\overline{\Omega}, \mathbb{R}^n)$ e $b \in \varphi(S)$ tal que $b \notin \varphi(\partial\Omega)$. Considere C_b a componente conexa de $\mathbb{R}^n \setminus \varphi(\partial\Omega)$ que contém b. Definimos

$$d(\varphi, \Omega, b) = d(\varphi, \Omega, x), \quad x \in C_b \setminus \varphi(S),$$

sendo o membro direito da igualdade acima dado pela Definição 1.1.

O próximo resultado implica que o grau estabelecido na Definição 1.6 é localmente constante na topologia de $C^1(\Omega, \mathbb{R})$.

Proposição 1.7. *Se $\varphi \in C^2(\overline{\Omega}, \mathbb{R}^n)$ e $b \notin \varphi(\partial\Omega)$, então existe uma vizinhança U de φ, na topologia de $C^1(\overline{\Omega}, \mathbb{R}^n)$ tal que, para cada $\psi \in U$, vale*

(i) $b \notin \psi(\partial\Omega)$;

(ii) $d(\psi, \Omega, b) = d(\varphi, \Omega, b)$.

O resultado a seguir é conhecido como propriedade da invariância do grau de Brouwer por homotopia de classe C^2.

Proposição 1.8 (Invariância por Homotopia de Classe C^2)**.** *Se $H \in C^2(\overline{\Omega} \times [0,1], \mathbb{R}^n)$ e $b \notin H(\partial(\Omega) \times [0,1])$, então*

$$d(H(\cdot, t_1), \Omega, b) = d(H(\cdot, t_2), \Omega, b),$$

para quaisquer $t_1, t_2 \in [0,1]$.

[††] **Teorema de Sard**: Sejam Λ um aberto de $\mathbb{R}^n$ e $f\colon \Lambda \to \mathbb{R}^n$ uma aplicação de classe C^1 ($f \in C^1(\Lambda, \mathbb{R}^n)$). Se $S = \{x \in \Lambda : J_f(x) = 0\}$, então $f(S)$ é um conjunto de medida nula.

A próxima proposição permitirá definir o grau de Brouwer para aplicações contínuas.

Proposição 1.9. *Se $\varphi \in C(\overline{\Omega}, \mathbb{R}^n)$ e $b \notin \varphi(\partial\Omega)$, então existe uma vizinhança U de φ, na topologia de $C^2(\overline{\Omega}, \mathbb{R}^n)$, tal que*

$$d(\psi_1, \Omega, b) = d(\psi_2, \Omega, b),$$

para quaisquer $\psi_1, \psi_2 \in U$.

Demonstração. Sejam $\varphi \in C(\overline{\Omega}, \mathbb{R}^n)$ e $b \notin \varphi(\partial\Omega)$. Como $b \notin \varphi(\partial\Omega)$, a distância de b a $\varphi(\partial\Omega)$ é positiva. Chamemos $r = \operatorname{dist}(b, \varphi(\partial\Omega)) > 0$. Pelo Teorema de Aproximação de Weierstrass, $C^2(\overline{\Omega}, \mathbb{R}^n)$ é denso em $C(\overline{\Omega}, \mathbb{R}^n)$. Portanto,

$$U = \left\{\psi \in C^2(\overline{\Omega}, \mathbb{R}^n) : \ \|\varphi - \psi\|_\infty < \frac{r}{2}\right\} \neq \emptyset,$$

onde $\|\varphi\|_\infty = \sup\limits_{x \in \overline{\Omega}} \|\varphi(x)\|$ e $\|\cdot\|$ é uma norma em $\mathbb{R}^n$.

Sejam $\psi_1, \psi_2 \in U$ e defina $H\colon \overline{\Omega} \times [0,1] \longrightarrow \mathbb{R}^n$ por

$$H(x,t) = t\psi_1(x) + (1-t)\psi_2(x), \text{ para } (x,t) \in \overline{\Omega} \times [0,1].$$

Afirmação: $b \notin H(\partial\Omega \times [0,1])$.

Com efeito, suponha que exista $(x_0, t_0) \in \partial\Omega \times [0,1]$ tal que $H(x_0, t_0) = b$. Neste caso, como a distância entre b e $\varphi(\partial\Omega)$ é r, ou seja, $\inf\{\|b - \varphi(x)\| : x \in \partial\Omega\} = r$, segue que

$$\|H(x_0,t_0) - \varphi(x_0)\| = \|b - \varphi(x_0)\| \geq r. \tag{1.3}$$

Em contrapartida, utilizando a expressão que define H, obtemos

$$\begin{aligned}\|H(x_0,t_0) - \varphi(x_0)\| &= \|t_0\psi_1(x_0) + (1-t_0)\psi_2(x_0) - t_0\varphi(x_0) - (1-t_0)\varphi(x_0)\| \\ &\leq t_0\|\psi_1(x_0) - \varphi(x_0)\| + (1-t_0)\|\psi_2(x_0) - \varphi(x_0)\|,\end{aligned}$$

donde

$$\|H(x_0,t_0) - \varphi(x_0)\| \leq t_0\|\psi_1 - \varphi\|_\infty + (1-t_0)\|\psi_2 - \varphi\|_\infty.$$

Como $\psi_1, \psi_2 \in U$, temos

$$\|H(x_0,t_0) - \varphi(x_0)\| < t_0\frac{r}{2} + (1-t_0)\frac{r}{2} = \frac{r}{2},$$

que contradiz (1.3) e prova a afirmação.

Agora, como $H \in C^2(\overline{\Omega}, \mathbb{R}^n)$, a Proposição 1.8 assegura que

$$d(H(\cdot,1), \Omega, b) = d(H(\cdot,0), \Omega, b),$$

ou seja,

$$d(\psi_1, \Omega, b) = d(\psi_2, \Omega, b),$$

como queríamos demonstrar. ■

Definição 1.10. Sejam $\varphi \in C(\overline{\Omega}, \mathbb{R}^n)$ e $b \notin \varphi(\partial\Omega)$. Definimos o grau topológico de Brouwer de φ em relação a Ω no ponto b por

$$d(\varphi, \Omega, b) = d(\psi, \Omega, b), \quad \psi \in U,$$

sendo U a vizinhança de φ na topologia de $C^2(\overline{\Omega}, \mathbb{R}^n)$, cuja existência é assegurada pela Proposição 1.9.

Finalizaremos esta seção apresentando as principais propriedades do grau topológico de Brouwer para aplicações contínuas, cujas demonstrações podem ser encontradas em [1] e [2].

Proposição 1.11 (Continuidade). *Se $\varphi \in C(\overline{\Omega}, \mathbb{R}^n)$ e $b \notin \varphi(\partial\Omega)$, então existe uma vizinhança V de φ em $C(\overline{\Omega}, \mathbb{R}^n)$ tal que para qualquer $\psi \in V$ tem-se*

(i) $b \notin \psi(\partial\Omega)$;

(ii) $d(\psi, \Omega, b) = d(\varphi, \Omega, b)$.

Proposição 1.12 (Invariância do Grau por Homotopia). *Seja $H \in C(\overline{\Omega} \times [0,1], \mathbb{R}^n)$ uma homotopia, com $b \notin H(\partial\Omega \times [0,1])$. Então, $d(H(\cdot, t), \Omega, b)$ é constante para todo $t \in [0,1]$.*

Proposição 1.13. *O grau de Brouwer é constante em componentes conexas de $\mathbb{R}^n \setminus \varphi(\partial\Omega)$. Ou seja, se b_1 e b_2 estão na mesma componente conexa de $\mathbb{R}^n \setminus \varphi(\partial\Omega)$, então*

$$d(\varphi, \Omega, b_1) = d(\varphi, \Omega, b_2).$$

Proposição 1.14 (Aditividade). *Se $\Omega = \Omega_1 \cup \Omega_2$, onde Ω_1 e Ω_2 são subconjuntos abertos, disjuntos de $\mathbb{R}^n$, e $b \notin \varphi(\partial\Omega_1) \cup \varphi(\partial\Omega_2)$, então*

$$d(\varphi, \Omega, b) = d(\varphi, \Omega_1, b) + d(\varphi, \Omega_2, b).$$

Proposição 1.15. *Se I é a aplicação identidade de $\overline{\Omega}$ em $\mathbb{R}^n$, ou seja, $I(x) = x$ para $x \in \overline{\Omega}$, então*

$$d(I, \Omega, b) = \begin{cases} 1, & se\ b \in \Omega, \\ 0, & se\ b \notin \overline{\Omega}. \end{cases}$$

Proposição 1.16 (Existência). *Se $d(\varphi, \Omega, b) \neq 0$, então existe $x_0 \in \Omega$ tal que $\varphi(x_0) = b$.*

Proposição 1.17. *Se $d(\varphi, \Omega, b) \neq 0$, então $\varphi(\Omega)$ é uma vizinhança de b. Ou seja, existe $\delta > 0$ tal que*

$$B_\delta(b) \subset \varphi(\Omega).$$

2 Teorema do Ponto Fixo de Brouwer

Esta seção é dedicada ao Teorema do Ponto Fixo de Brouwer. Este teorema possui diferentes demonstrações. Apresentaremos aqui uma demonstração dele utilizando a teoria do grau topológico de Brouwer.

No que segue, $\overline{B_1(0)}$ denota a bola fechada de centro 0 e raio 1 em $\mathbb{R}^n$, $B_1(0)$ denota a bola aberta de centro 0 e raio 1 em $\mathbb{R}^n$ e $\partial B_1(0)$ denota a fronteira de $B_1(0)$.

Teorema 2.1 (Teorema do Ponto Fixo de Brouwer). *Toda aplicação contínua de $\overline{B_1(0)}$ em $\overline{B_1(0)}$ admite pelo menos um ponto fixo. Ou seja, se $f\colon \overline{B_1(0)} \to \overline{B_1(0)}$ é contínua, então existe $x \in \overline{B_1(0)}$ tal que $f(x) = x$.*

Demonstração. Defina $\varphi\colon \overline{B_1(0)} \to \mathbb{R}^n$ por $\varphi(x) = x - f(x)$ para $x \in \overline{B_1(0)}$. Claramente, φ é contínua, pois f e a aplicação identidade são contínuas.

Se existir $x \in \partial B_1(0)$ tal que $f(x) = x$, então não há o que provar. Suponhamos, então, que $f(x) \neq x$ para $x \in \partial B_1(0)$. Portanto, $\varphi(x) \neq 0$ para $x \in \partial B_1(0)$ e consequentemente $0 \notin \varphi(\partial B_1(0))$.

A partir de agora, por simplicidade, denotaremos $B_1(0) = B_1$.

Consideremos a homotopia entre a aplicação identidade e φ, definida por

$$\begin{aligned} H\colon \overline{B_1} \times [0,1] &\longrightarrow \mathbb{R}^n \\ (x,t) &\longmapsto H(x,t) = x \ - \ t\ f(x). \end{aligned}$$

Afirmamos que $0 \notin H(\partial B_1 \times [0,1])$. Com efeito, para $t = 1$, temos

$$H(x,1) = x - 1 \cdot f(x) = \varphi(x) \neq 0, \ \forall x \in \partial B_1.$$

Se $t_0 \in [0,1)$, então para $x \in \partial B_1$ temos

$$\|t_0 f(x)\| = t_0 \|f(x)\| \leq t_0 \cdot 1 < 1 = \|x\|,$$

donde $t_0 f(x) \neq x$ para todo $x \in \partial B_1$. Logo, $H(x,t_0) \neq 0$ para todo $x \in \partial B_1$ e, portanto, $0 \notin H(\partial B_1 \times [0,1])$.

Pela invariância do grau por homotopia (Proposição 1.12), sabemos que $d(H(\cdot,t), B_1, 0)$ é constante para todo $t \in [0,1]$. Sendo assim, temos

$$d(H(\cdot,0), B_1, 0) = d(H(\cdot,1), B_1, 0). \tag{2.1}$$

Por outro lado, como

$$H(x,0) = x \quad \text{e} \quad H(x,1) = \varphi(x), \qquad \forall x \in \overline{B_1},$$

segue de (2.1) que

$$d(\varphi, B_1, 0) = d(Id, B_1, 0) \overset{(*)}{=} 1.$$

A igualdade $(*)$ se deve à Proposição 1.15. Então, como $d(\varphi, B_1, 0) \neq 0$, existe $x_0 \in B_1$ tal que $\varphi(x_0) = 0$ pela Proposição 1.16. Logo, existe $x_0 \in B_1$ tal que

$$f(x_0) = x_0,$$

o que completa a prova. ■

Observação 2.2. Cabe mencionar que o Teorema 2.1 vale para aplicações contínuas $f\colon \overline{B_\mu(a)} \to \overline{B_\mu(a)}$, onde $\overline{B_\mu(a)} = \{x \in \mathbb{R}^n : \|x - a\| \leq \mu\}$, $a \in \mathbb{R}^n$ é arbitrário e μ é qualquer número real positivo. De fato, sabemos que existe um homeomorfismo $\varphi\colon \overline{B_1(0)} \to \overline{B_\mu(a)}$, a saber, $\varphi(x) = \mu x + a$, com $\varphi^{-1}\colon \overline{B_\mu(a)} \to \overline{B_1(0)}$ dada por $\varphi^{-1}(x) = \frac{x-a}{\mu}$. Sendo f uma aplicação contínua de $\overline{B_\mu(a)}$ em $\overline{B_\mu(a)}$, a composta $\psi = \varphi^{-1} \circ f \circ \varphi\colon \overline{B_1(0)} \to \overline{B_1(0)}$ é contínua e admite um ponto fixo pelo Teorema 2.1, ou seja, existe $x_0 \in \overline{B_1(0)}$ tal que $\psi(x_0) = x_0$. Daí, $f(\varphi(x_0)) = \varphi(x_0)$ e, por conseguinte, $\varphi(x_0) \in \overline{B_\mu(a)}$ é um ponto fixo de f.

Na sequência apresentamos uma versão mais geral do Teorema do Ponto Fixo de Brouwer.

Teorema 2.3 (Teorema do Ponto Fixo de Brouwer Generalizado)**.** *Toda aplicação contínua φ de um subconjunto convexo compacto K de $\mathbb{R}^n$ em si mesmo admite um ponto fixo.*

Demonstração. Sejam K um subconjunto convexo e compacto de $\mathbb{R}^n$ e $\varphi\colon K \to K$ uma aplicação contínua.

Vamos supor que $\operatorname{int} K \neq \emptyset$ e $0 \in \operatorname{int} K$. Se $0 \notin K$, poderíamos considerar o conjunto compacto convexo $\widetilde{K} = K - \widetilde{x} = \{x - \widetilde{x} : x \in K\}$, com $\widetilde{x} \in \operatorname{int} K$, e usar os mesmos argumentos que serão utilizados abaixo, uma vez que $0 \in \operatorname{int} \widetilde{K}$.

Seja $p_K : \mathbb{R}^n \to \mathbb{R}^+$ a função dada por

$$p_K(x) = \inf\{\lambda \geq 0 : \ x \in \lambda K\}, \quad x \in \mathbb{R}^n.$$

Usando a definição de p_k e propriedades de ínfimo, é fácil verificar que p_K é um funcional sublinear, ou seja,

$$p_K(\lambda x) = \lambda p_K(x), \quad \forall \lambda > 0, \forall x \in \mathbb{R}^n, \tag{2.2}$$

e

$$p_K(x+y) \leq p_K(x) + p_K(y), \quad \forall x, y \in \mathbb{R}^n. \tag{2.3}$$

Ademais, é importante destacar as seguintes propriedades de p_K:

(P1) para todo $x \in \mathbb{R}^n$, $x \in p_K(x) \cdot K$;

(P2) $K = \{x \in \mathbb{R}^n : p_K(x) \leq 1\}$;

(P3) p_K é uma função contínua.

Para provarmos (P1), tomamos $x \in \mathbb{R}^n$ arbitrariamente. Como $p_K(x) = \inf\{\lambda \geq 0 : x \in \lambda K\}$, existe uma sequência $(\lambda_n)_{n\in\mathbb{N}} \subset \{\lambda \geq 0 : \ x \in \lambda K\}$ tal que $\lambda_n \to p_K(x)$ quando $n \to \infty$. Agora, temos dois casos a considerar: quando $p_K(x) = 0$ e quando $p_K(x) > 0$.

Suponhamos, inicialmente, que $p_K(x) = 0$. Note que $\|x\| \leq |\lambda_n| \operatorname{diam} K$, uma vez que $\lambda_n x \in K$. Sendo K compacto, temos $\operatorname{diam} K < \infty$. E como $\lambda_n \to 0$ quando $n \to \infty$, concluímos que $x = 0 \in 0 \cdot K$. Note ainda que $p_K(0) = 0$ pela definição de p_K. Portanto, $p_K(x) = 0$ se, e somente se, $x = 0$.

Agora, se $p_K(x) > 0$, existe $n_0 \in \mathbb{N}$ tal que $\lambda_n > 0$ para todo $n > n_0$. Dessa forma, temos

$$\frac{x}{\lambda_n} \in K \quad \text{sempre que } n > n_0. \tag{2.4}$$

Como $\frac{x}{\lambda_n} \to \frac{x}{p_K(x)}$ quando $n \to \infty$ e K é fechado, segue de (2.4) que $\frac{x}{p_K(x)} \in K$, de onde concluímos que $x \in p_K(x) \cdot K$.

Vamos provar (P2). Se $x \in K$, então $x \in 1 \cdot K$ e claramente $p_K(x) \leq 1$. Se $p_K(x) = 0$, então $x = 0$ (provamos acima) e $0 \in K$. Suponhamos, pois, que $0 < p_K(x) \leq 1$. Como $\frac{x}{p_K(x)} \in K$ e $0 \in K$, segue pela convexidade de K que

$$p_K(x) \cdot \frac{x}{p_K(x)} + (1 - p_K(x)) \cdot 0 \in K \quad \Longrightarrow \quad x \in K.$$

Agora, prosseguiremos para provar (P3). Usando a subaditividade de p_K (propriedade (2.3)), podemos inferir que

$$|p_K(x) \ - \ p_K(y)| \leq \max\{p_K(x-y), p_K(y-x)\}. \tag{2.5}$$

Afirmação 1: Existe $M > 0$ tal que

$$p_K(x) \leq M\|x\|, \quad \forall x \in \mathbb{R}^n. \tag{2.6}$$

De fato, como $0 \in \operatorname{int} K$, existe $\varepsilon > 0$ tal que $B(0, \varepsilon) \subset K$. Dado $x \in \mathbb{R}^n \setminus \{0\}$, temos

$$y = \frac{x}{\|x\|} \in \overline{S}(0,1) = \{x \in \mathbb{R}^n : \|x\| = 1\}.$$

Portanto,

$$\frac{\varepsilon}{2} \cdot y \in B(0, \varepsilon),$$

pois

$$\left\|\frac{\varepsilon}{2}\cdot y\right\| = \frac{\varepsilon}{2} < \varepsilon.$$

Então, $\dfrac{\varepsilon}{2}\dfrac{x}{\|x\|} \in K$ e, consequentemente, $x \in \dfrac{2}{\varepsilon}\|x\|\cdot K$, donde

$$\frac{2}{\varepsilon}\|x\| \in \{\lambda \geq 0 : x \in \lambda K\}.$$

Logo,

$$p_K(x) \leq M\|x\|,$$

onde $M = \frac{2}{\varepsilon}$. Se $x = 0$, temos $p_K(x) = 0$ e, portanto, $p_K(x) = 0 = M \cdot \|0\|$, provando a Afirmação 1.

Por (2.5) e (2.6), obtemos

$$|p_K(x) \ - \ p_K(y)| \leq M\|x-y\|. \tag{2.7}$$

Finalmente, a relação (2.7) mostra que p_K é uma aplicação lipschitziana e, portanto, contínua.

Agora, vamos definir $h\colon K \to \overline{B(0,1)}$ por

$$h(x) = \begin{cases} 0, & x = 0, \\ p_K(x)\dfrac{x}{\|x\|}, & x \neq 0. \end{cases}$$

A aplicação h é contínua. Realmente, h é contínua em $x \neq 0$, pois é produto de funções contínuas, e é contínua em $x = 0$, pois

$$\lim_{x\to 0} h(x) = \lim_{x \to 0} p_K(x)\frac{x}{\|x\|} \overset{(*)}{=} 0 = h(0),$$

onde em $(*)$ usamos o fato de p_K ser contínua em 0, $p_K(0) = 0$ e o fato da função $\tau(x) = \frac{x}{\|x\|}$ ser limitada.

Agora, mostremos que $h\colon K \to \overline{B(0,1)}$ é uma bijeção. Com efeito, seja $g\colon \overline{B(0,1)} \to K$ a aplicação definida por

$$g(y) = \begin{cases} 0, & y = 0, \\ \dfrac{\|y\|}{p_K(y)}y, & y \neq 0. \end{cases}$$

Inicialmente, note que se $\|y\| \leq 1$ e $y \neq 0$, então $g(y) \in K$, conforme podemos constatar abaixo:

$$p_K\left(\frac{\|y\|}{p_K(y)}y\right) = \frac{\|y\|}{p_K(y)}p_K(y) = \|y\| \ \leq \ 1 \quad \overset{(**)}{\Longrightarrow} \quad \frac{\|y\|}{p_K(y)}y \in K,$$

onde em $(**)$ usamos a propriedade (P2).

<u>Afirmação 2</u>: g é a inversa de h.

De fato, para $y \in \overline{B(0,1)}$, $y \neq 0$, vale:

$$\begin{aligned} h(g(y)) &= h\left(\frac{\|y\|}{p_K(y)}y\right) = p_K\left(\frac{\|y\|}{p_K(y)}y\right)\cdot \frac{\frac{\|y\|}{p_K(y)}y}{\left\|\frac{\|y\|}{p_K(y)}y\right\|} \\ &= \frac{\frac{\|y\|}{p_K(y)}\cdot p_K(y)\cdot\frac{\|y\|}{p_K(y)}\cdot y}{\frac{\|y\|\cdot\|y\|}{p_K(y)}} \\ &= y. \end{aligned}$$

Além disso, para $x \in K$, $x \neq 0$, tem-se

$$\begin{aligned} g(h(x)) &= g\Big(p_K(x)\frac{x}{\|x\|}\Big) = \frac{\Big\|p_K(x)\cdot\frac{x}{\|x\|}\Big\|\cdot\frac{p_K(x)x}{\|x\|}}{p_K\big(p_K(x)\cdot\frac{x}{\|x\|}\big)} \\ &= \frac{p_K(x)\cdot\frac{p_K(x)}{\|x\|}\cdot x}{\frac{p_K(x)}{\|x\|}\cdot p_K(x)} \\ &= x. \end{aligned}$$

Logo, $g\circ h = id_K$ e $h\circ g = id_{\overline{B(0,1)}}$, de onde segue a validade da Afirmação 2, a qual nos permite concluir que h é uma bijeção.

Sendo h uma bijeção contínua definida sobre um compacto, sua inversa g também é contínua. Portanto, h é um homeomorfismo. Com isso, podemos afirmar que a aplicação $f = h\circ\varphi\circ h^{-1}\colon \overline{B_1(0)} \to \overline{B_1(0)}$ é contínua, posto que $\varphi\colon K \to K$ é contínua. Sendo assim, o Teorema 2.1 garante que existe $x_0 \in \overline{B_1(0)}$ tal que $f(x_0) = (x_0)$, ou seja,

$$h\circ\varphi\circ h^{-1}(x_0) = x_0 \quad\Longrightarrow\quad \varphi(h^{-1}(x_0)) = h^{-1}(x_0),$$

donde concluímos que $h^{-1}(x_0)\in K$ é ponto fixo de φ e completamos a prova. ∎

3 Resultados equivalentes ao Teorema do Ponto Fixo de Brouwer

Nesta seção, apresentaremos algumas equivalências entre o Teorema do Ponto Fixo de Brouwer e outros resultados da teoria de Análise Funcional.

Cabe mencionar que alguns conceitos e resultados da teoria de Análise Funcional serão fundamentais para a compreensão dos resultados apresentados aqui nesta seção. Recomendamos ao leitor interessado que consulte a referência [4] para ter acesso aos pré-requisitos necessários.

A seguir, enunciamos quatro resultados equivalentes ao Teo. do Ponto Fixo de Brouwer. As demonstrações serão feitas na sequência.

Teorema 3.1 (Teorema de Knaster–Kuratowski–Mazurkiewicz). *Seja E um espaço vetorial topológico separável. Considere m pontos de E, $x_1, \ldots, x_m$, e m fechados de E, $X_1,\ldots,X_m$ para os quais a seguinte propriedade é válida:* $\operatorname{conv}\{x_{i_1},\ldots,x_{i_k}\} \subset X_{i_1}\cup\cdots\cup X_{i_k}$*, para toda família de índices $\{i_1,\ldots,i_k\}\subseteq\{1,2,\ldots,m\}$, onde* $\operatorname{conv}\{x_{i_1},\ldots,x_{i_k}\}$ *é a envoltória convexa do conjunto $\{x_{i_1},\ldots,x_{i_k}\}$, ou seja, a interseção de todos os conjuntos convexos que contêm $\{x_{i_1},\ldots,x_{i_k}\}$. Então,*

$$\bigcap_{i=1}^{m} X_i \neq \emptyset.$$

Teorema 3.2 (Generalização do Teorema 3.1). *Sejam E um espaço vetorial topológico separável e $X\subset E$ um subconjunto qualquer de E. Cada $x\in X$ está associado a um conjunto fechado $F(x)$ de E, e existe pelo menos um $x\in X$ tal que $F(x)$ é compacto.*

Se a envoltória convexa de toda família finita $\{x_1,\ldots,x_m\}\subset X$ estiver contida em $\bigcup_{i=1}^m F(x_i)$, então $\bigcap_{x\in X} F(x)\neq\emptyset$.

Teorema 3.3 (Desigualdade do $\min-\max$ de Ky–Fan). *Sejam E um espaço vetorial topológico separável e X um subconjunto compacto e convexo de E. Seja $f\colon X\times X\to\mathbb{R}$ tal que:*

(i) *$\forall x\in X$ fixado, $f(x,y)$ é* semi-contínua inferiormente *(s.c.i) em $y\in X$;*

(ii) $\forall y \in X$ *fixado,* $f(x, y)$ *é* quase-côncava *em* $x \in X$.

Nessas condições, tem-se

$$\min_{y\in X} \max_{x\in X} f(x, y) \leq \max_{x\in X} f(x, x).$$

Teorema 3.4 (Teorema de Hartman–Stampacchia)**.** *Sejam C um subconjunto compacto e convexo de $\mathbb{R}^n$ e A uma aplicação contínua de C em $\mathbb{R}^n$, $A\colon C \to \mathbb{R}^n$. Então, existe $u \in C$ tal que*

$$\langle Au,\, v - u\rangle \geq 0,\ \forall v \in C,$$

onde $\langle\ ,\ \rangle$ denota um produto interno em $\mathbb{R}^n$.

Provaremos as equivalências entre os resultados enunciados acima e o Teorema do Ponto Fixo de Brouwer (Teorema 2.1) seguindo o roteiro abaixo.

$$\begin{aligned} &\text{Teorema 2.1} \Rightarrow \text{Teorema 2.3} \Rightarrow \text{Teorema 3.1}\\ &\text{Teorema 3.1} \Rightarrow \text{Teorema 3.2} \Rightarrow \text{Teorema 3.3}\\ &\text{Teorema 3.3} \Rightarrow \text{Teorema 3.4} \Rightarrow \text{Teorema 2.1.}\end{aligned} \tag{3.1}$$

A prova da implicação <u>Teorema 2.1 ⇒ Teorema 2.3</u> foi feita na Seção 2.

<u>Teorema 2.3 ⇒ Teorema 3.1</u>: Sejam $x_1, \ldots, x_m$ pontos de E e $X_1, \ldots, X_m$ conjuntos fechados para os quais

$$\operatorname{conv}\{x_{i_1}, \ldots, x_{i_k}\} \subset X_{i_1} \cup \cdots \cup X_{i_k}, \tag{3.2}$$

para toda família de índices $\{i_1, \ldots, i_k\} \subseteq \{1, 2, \ldots, m\}$.

Queremos mostrar que $\cap_{i=1}^m X_i \neq \emptyset$. Prosseguiremos por absurdo. Suponhamos que $\cap_{i=1}^m X_i = \emptyset$ e consideremos $K = \operatorname{conv}\{x_1, \ldots, x_m\}$.

Seja $\varphi_i(x) = d(x, X_i)$. Para todo $x \in K$, existe $i_0 \in \{1, 2, \ldots, m\}$ tal que $x \notin X_{i_0}$, pois caso contrário concluiríamos que $\cap_{i=1}^m X_i \neq \emptyset$ por meio da relação (3.2). Neste caso, $\varphi_{i_0}(x) = d(x, X_{i_0}) > 0$ e

$$\sum_{i=1}^{m} \varphi_i(x) > 0.$$

Agora, seja $\Phi\colon K \to K$ a aplicação dada por

$$\Phi(x) = \frac{\sum_{i=1}^m \varphi_i(x) x_i}{\sum_{i=1}^m \varphi_i(x)} \quad \text{para } x \in K.$$

Como φ_i é contínua para todo $i \in \{1, \ldots, m\}$, Φ também é contínua. Como K é compacto e convexo, existe $x_0 \in K$ tal que $\Phi(x_0) = x_0$, pelo Teorema 2.3.

Por outro lado, pela relação (3.2), existe i_1 tal que $x_0 \in X_{i_1}$. Então, $\varphi_{i_1}(x_0) = d(x_0, X_{i_1}) = 0$. Logo,

$$x_0 = \Phi(x_0) = \frac{\sum_{j\neq i_1} \varphi_j(x_0) x_j}{\sum_{j\neq i_1} \varphi_j(x_0)}.$$

Repetindo esse argumento, como $x_0 \in \operatorname{conv}_{i\neq i_1}\{x_i\} \subset \bigcup_{i\neq i_1} X_i$, podemos encontrar um índice i_2 tal que $x_0 \in X_{i_2}$ e assim $\varphi_{i_2}(x_0) = 0$. Com isso, x_0 deve pertencer a $\operatorname{conv}\{x_i\}, i \neq i_1$ e $i \neq i_2$. Prosseguindo dessa maneira, inferimos que

$$x_0 \in \bigcap_{i=1}^m X_i,$$

o que é um absurdo, pois estamos supondo que $\bigcap_{i=1}^{m} X_i = \emptyset$. Logo, devemos ter $\bigcap_{i=1}^{m} X_i \neq \emptyset$.

<u>Teorema 3.1 ⇒ Teorema 3.2</u>: Temos, por hipótese, que cada $x \in X$ está associado a um fechado $F(x)$ de E e existe ao menos um $x_0 \in X$ tal que $F(x_0)$ seja compacto, além disso a envoltória convexa de toda família finita $\{x_1, \dots, x_m\} \subset X$ está contida em $\bigcup_{i=1}^{m} F(x_i)$. Devemos mostrar que $\bigcap_{x\in X} F(x) \neq \emptyset$. Pois bem, suponhamos que isso não ocorra, ou seja, $\bigcap_{x\in X} F(x) = \emptyset$.

Como $F(x_0)$ é compacto para algum $x_0 \in X$, $\bigcap_{x\in X} F(x) \subset F(x_0)$ e estamos supondo $\bigcap_{x\in X} F(x) = \emptyset$, podemos garantir que existe uma subfamília finita $\{x_{i_1}, \ \dots \ , x_{i_m}\} \subset X$ tal que $\bigcap_{i=1}^{m} F(x_i) = \emptyset$, mas isso é um absurdo, visto que pelo Teorema 3.1 devemos ter $\bigcap_{x\in X} F(x_i) \neq \emptyset$. Dessa forma, a prova dessa implicação está completa.

<u>Teorema 3.2 ⇒ Teorema 3.3</u>: Seja $\mu = \max\limits_{x\in X} f(x,x)$ e considere a família de fechados $F(x)$, $x \in X$, onde

$$F(x) = \{y \in X : f(x,y) \leq \mu\}.$$

Seja $\{x_1, \dots, x_n\}$ uma família finita de elementos de X.

<u>Afirmação</u>: $\operatorname{conv}\{x_1, \dots, x_n\} \subset \bigcup\limits_{i=1}^{n} F(x_i)$.

Com efeito, suponha que tal inclusão não ocorra. Ou seja, existe uma combinação convexa $\sum_{j=1}^{n} \alpha_j x_j$ tal que

$$\sum_{j=1}^{n} \alpha_j x_j \notin \bigcup_{i=1}^{n} F(x_i),$$

em que $\alpha_j \in [0,1]$ para todo j e $\sum_{j=1}^{n} \alpha_j = 1$. Com isso, segue pela definição de $F(x)$ que

$$f\left(x_i, \sum_{j=1}^{n} \alpha_j x_j\right) > \mu, \quad \forall i = 1, 2, \dots, n.$$

Como $f(x,y)$ é quase-concava em $x \in X$, temos

$$f\left(\sum_{j=1}^{n} \alpha_j x_j, \sum_{j=1}^{n} \alpha_j x_j\right) > \mu,$$

o que é um absurdo, pois $\mu = \max\limits_{x\in X} f(x,x)$. Isso confirma a veracidade da afirmação. Portanto, pelo Teorema 3.2, temos que $\bigcap_{x\in X} F(x) \neq \emptyset$.

Por fim, como $\bigcap_{x\in X} F(x) \neq \emptyset$ e $f(x,y)$ é s.c.i em $y \in X$, existe $y_0 \in \bigcap_{x\in X} F(x)$, de onde segue que

$$\max_{x\in X} \ f(x, y_0) \leq \mu = \max_{x\in X} f(x,x).$$

Por conseguinte,

$$\min_{y\in X} \ \max_{x\in X} \ f(x,y) \leq \max_{x\in X} \ f(x,x).$$

<u>Teorema 3.3 ⇒ Teorema 3.4</u>: Considere $f \colon C \times C \to \mathbb{R}$ a aplicação dada por $f(x,y) = \langle -Ay, \, x - y\rangle$, onde $\langle \ , \ \rangle$ denota um produto interno em $\mathbb{R}^n$.

Inicialmente, observamos que f é uma aplicação contínua com respeito à segunda variável. Com efeito, fixando $x_0 \in \mathbb{R}^n$ e tomando quaisquer $y_1, y_2 \in \mathbb{R}^n$, temos

$$\begin{aligned} f(x_0,y_1) - f(x_0,y_2) &= \langle -Ay_1, x_0 - y_1\rangle - \langle -Ay_2, x_0 - y_2\rangle \\ &= -\langle Ay_1, x_0\rangle + \langle Ay_1, y_1\rangle + \langle Ay_2, x_0\rangle - \langle Ay_2, y_2\rangle \\ &= -\langle Ay_1, x_0\rangle + \langle Ay_1, y_1\rangle - \langle Ay_2, y_1\rangle \\ &\quad + \langle Ay_2, y_1\rangle + \langle Ay_2, x_0\rangle - \langle Ay_2, y_2\rangle \\ &= \langle Ay_2 - Ay_1, x_0\rangle + \langle Ay_2 - Ay_1, y_1\rangle + \langle Ay_2, y_1 - y_2\rangle \\ &\overset{(*)}{\leq} \|Ay_2 - Ay_1\| \cdot \|x_0\| + \|Ay_2 - Ay_1\| \cdot \|y_1\| \\ &\quad + \|Ay_2\| \cdot \|y_1 - y_2\|, \end{aligned} \tag{3.3}$$

onde em $(*)$ usamos a Desigualdade de Cauchy–Schwarz.

Se $(y_n)_{n\in\mathbb{N}}$ é uma sequência qualquer em C tal que $y_n \to y_0$ quando $n \to \infty$, então $Ay_n \to Ay_0$, haja vista que A é contínua e, por conseguinte, $f(x_0, y_n) \to f(x_0, y_0)$ quando $n \to \infty$, uma vez que $(y_n)_{n\in\mathbb{N}}$ e $(Ay_n)_{n\in\mathbb{N}}$ são sequências limitadas por serem convergentes e por (3.3) temos

$$f(x_0,y_n) - f(x_0,y_0) \leq \|Ay_n - Ay_0\| \cdot \|x_0\| + \|Ay_n - Ay_0\| \cdot \|y_n\| + \|Ay_n\| \cdot \|y_n - y_0\| \to 0 \tag{3.4}$$

quando $n \to \infty$. Isso prova a continuidade de f com respeito à segunda variável.

Além disso, f é quase-côncava em com respeito à primeira variável. De fato, dado $\alpha \in \mathbb{R}$, seja

$$D_f(\alpha) := \{(x,y) \in C \times C : f(x,y) > \alpha\} = \{(x,y) \in C \times C : \langle -Ay,\, x - y\rangle > \alpha\}.$$

Fixando $y_0 \in C$, temos que se $(x_1, y_0) \in D_f(\alpha)$ e $(x_2, y_0) \in D_f(\alpha)$, então $((1-t)x_1 + tx_2, y_0) \in C \times C$ se $t \in [0,1]$ e

$$\begin{aligned} f((1-t)x_1 + tx_2, y_0) &= \langle -Ay_0, (1-t)x_1 + tx_2 - y_0\rangle \\ &= \langle -Ay_0, (1-t)x_1 + tx_2 + ty_0 - ty_0 - y_0\rangle \\ &= \langle -Ay_0, (1-t)(x_1 - y_0) + t(x_2 - y_0)\rangle \\ &= (1-t)\langle -Ay_0, x_1 - y_0\rangle + t\langle -Ay_0, x_2 - y_0\rangle \\ &> (1-t)\alpha + t\alpha = \alpha, \end{aligned}$$

donde segue que

$$[(x_1,y_0),(x_2,y_0)] \subset \{(x,y) \in C \times C : f(x,y) > \alpha\} = D_f(\alpha).$$

Daí, $D_f(\alpha)$ é convexo e, portanto, $f(x,y)$ é quase-côncava em x. Sendo assim, o Teorema 3.3 garante que

$$\min_{y\in C} \max_{x\in C} f(x,y) \leq \max_{x\in C} f(x,x) = 0,$$

uma vez que $f(x,x) = \langle -Ax,\, x - x\rangle = \langle -Ax, 0\rangle = 0$.

Portanto, existe $u \in C$ tal que $f(v,u) \leqslant 0$ para todo $v \in C$, ou seja, existe $u \in C$ tal que $f(v,u) = \langle -Au,\, v - u\rangle \leqslant 0$ para todo $v \in C$. Daí, concluímos que

$$\langle Au,\, v - u\rangle \geqslant 0, \quad \forall v \in C.$$

Teorema 3.4 $\Rightarrow$ Teorema 2.1: Sejam $f\colon \overline{B_1(0)} \to \overline{B_1(0)}$ uma aplicação contínua e $A = I - f$. Aplicando o Teorema 3.4 em A, existe $u \in \overline{B_1(0)}$ tal que para todo $v \in \overline{B_1(0)}$ tem-se

$$\langle Au,\, v - u\rangle \geq 0,$$

ou seja,

$$\langle u - f(u),\, v - u\rangle \geq 0, \quad \text{para todo } v \in \overline{B_1(0)}. \tag{3.5}$$

Vamos mostrar que $f(u) = u$. Como $f(u) \in \overline{B_1(0)}$ para tal $u \in \overline{B_1(0)}$, considerando $v = f(u)$ em (3.5), obtemos

$$\langle u - f(u), f(u) - u\rangle = \langle -(f(u) - u), f(u) - u\rangle = -\langle f(u) - u, f(u) - u\rangle = -\|f(u) - u\|^2 \;\geq\; 0,$$

donde segue que

$$\|f(u) - u\|^2 \leq 0 \quad \Longrightarrow \quad \|f(u) - u\|^2 = 0$$

e, portanto,

$$f(u) = u.$$

Concluímos, assim, as provas das implicações presentes em (3.1).

Finalizaremos este artigo provando outras implicações entre alguns dos teoremas vistos, a saber, provaremos que Teorema 2.3 ⇒ Teorema 3.4 e Teorema 3.3 ⇒ Teorema 2.3.

Teorema 2.3 ⇒ Teorema 3.4: Sejam $C \subset \mathbb{R}^n$ um conjunto compacto e convexo e $A\colon C \to \mathbb{R}^n$ uma aplicação contínua. Consideremos $f\colon C \to C$ a aplicação dada por

$$\begin{aligned} f\colon C &\longrightarrow C \\ u &\longmapsto f(u) = \mathrm{proj}_C(-Au + u) = \min_{v \in C} \| - Au + u - v\|. \end{aligned}$$

É sabido que $\mathrm{proj}_C(-Au + u)$ se caracteriza por satisfazer

$$\langle (-Au + u) - \mathrm{proj}_C(-Au + u),\, v - \mathrm{proj}_C(-Au + u)\rangle \leq 0,\ \forall v \in C. \tag{3.6}$$

Além disso, a projeção sobre um conjunto convexo fechado é contínua. Como A também é contínua, podemos deduzir que f é contínua. Então, pelo Teorema 2.3, existe $u \in C$ tal que u é ponto fixo de f, ou seja,

$$u = f(u) = \mathrm{proj}_C(-Au + u) \tag{3.7}$$

Por (3.6) e (3.7), obtemos

$$\langle (-Au + u) - u,\, v - u\rangle \leq 0,\ \forall v \in C,$$

donde concluímos que

$$\langle Au,\, v - u\rangle \geq 0, \quad \forall v \in C,$$

e completamos a prova.

Teorema 3.3 ⇒ Teorema 2.3: Sejam $\varphi\colon K \to K$ uma aplicação contínua e $f\colon K \times K \to \mathbb{R}$ a aplicação definida por

$$f(x, y) = \|\varphi(y) - y\| - \|x - \varphi(y)\|, \quad x, y \in K.$$

Afirmamos que f é côncava em x. De fato, considerando a aplicação $g(x, y) = -f(x, y) = \|x - \varphi(y)\| - \|\varphi(y) - y\|$ para $x, y \in K$, fixando $y_0 \in \mathbb{R}^n$, tomando $x_1, x_2 \in \mathbb{R}^n$ e $t \in [0, 1]$, obtemos

$$\begin{aligned} g(tx_1 + (1-t)x_2, y_0) = {} & \|tx_1 + (1-t)x_2 - \varphi(y_0)\| - \|\varphi(y_0) - y_0\| \\ & = \|t(x_1 - \varphi(y_0)) + (1-t)(x_2 - \varphi(y_0)\| - t\|\varphi(y_0) - y_0\| - (1-t)\|\varphi(y_0) - y_0\| \\ & \leq t[\|x_1 - \varphi(y_0)\| - \|\varphi(y_0) - y_0\|] + (1-t)[\|x_2 - \varphi(y_0)\| - \|\varphi(y_0) - y_0\|] \\ & = t\ g(x_1, y_0) + (1-t)\ g(x_2, y_0). \end{aligned}$$

Daí, g é convexa em x e, portanto, f é côncava em x. Ainda, pela continuidade de φ, podemos verificar facilmente que f é contínua em y.

Dessa forma, o Teorema 3.3 assegura que

$$\min_{y\in K}\ \max_{x\in K}\ f(x,y)\ \leq\ \max_{x\in K}\ f(x,x) = \max_{x\in K}\ (\|\varphi(x)-x\| - \|x-\varphi(x)\|) = 0.$$

Então, existe $y_0 \in K$ tal que $f(x,y_0) \leq 0$ para todo $x \in K$. Isso implica

$$\|\varphi(y_0)-y_0\| \leq \|x-\varphi(y_0)\|, \quad \forall x \in K.$$

Em particular, tomando $x = \varphi(y_0)$, obtemos

$$\varphi(y_0) = y_0,$$

de onde segue que y_0 é um ponto fixo para φ.

Agradecimentos: Primeiramente, agradeço a Deus pela força a cada minuto de apreensão; a Profa. Dra. Suzete Maria Silva Afonso pela amizade, incentivo, por sempre acreditar que eu poderia ir além e por me ajudar com diversas questões que perpassam o âmbito acadêmico, tornando-se um exemplo de pessoa pra mim... Por isso, sou extremamente grato. Agradeço a minha família por todo apoio e carinho. Por fim, agradeço aos amigos do grupo PET - Matemática que sempre me auxiliam em momentos de dúvida, e ao amigo Lucas Yudy pelos conselhos e conversas descontraídas.

Abstract: In this article we present a proof of the Brouwer Fixed Point Theorem using the Brouwer degree theory. In addition, we show some equivalences between the Brouwer Fixed Point Theorem and others known results from Functional Analysis theory, such as Theorem of Knaster–Kuratowiski–Mazurkiewicz, Theorem of Hartman–Stampachia and the min–max *Inequality of Ky–Fan.*

Keywords: Brouwer Fixed Point Theorem; Brouwer degree theory; Functional Analysis; Theorem of Knaster–Kuratowiski–Mazurkiewicz; Theorem of Hartman–Stampachia; min–max *Inequality of Ky–Fan.*

Referências Bibliográficas

[1] ALMEIDA, O. B. *Teoria do grau e aplicações*. Dissertação (Mestrado em Matemática)– Universidade Federal de Campina Grande, Campina Grande, 2006.

[2] BERESTYCKI, H. *Méthodes topologiques et problèmes aux limites non linéaires*. Tese (Doctorat en Mathématiques) – Université Paris VI – Université Pierre et Marie Curie, 1975.

[3] BRATTKA, V.; LE ROUX, S.; MILLER, J. S.; PAULY, A. Connected choice and the Brouwer Fixed Point Theorem. *Journal of Mathematical Logic*, v. 19, n. 01, p. 1950004, 2019.

[4] BREZIS, H.; CIARLET, P. G.; LIONS, J. L. *Analyse fonctionnelle: théorie et applications*. [S.l.]: Dunod Paris, 1999. v. 91.

[5] BROWN R.F. Brouwer Fixed Point Theory. In: *A Topological Introduction to Nonlinear Analysis*. Birkhäuser, Boston, MA, 2004.

[6] BOROGOVAC, M. Two applications of Brouwer's fixed point theorem: in insurance and in biology models. *Journal of Difference Equations and Applications*, Taylor & Francis, v. 22, n. 6, p. 727–744, 2016.

[7] DABROWSKI, L. Towards a noncommutative Brouwer fixed-point theorem, *Journal of Geometry and Physics*, vol. 105, 2016.

[8] FOLLAND, G. B. *Real analysis: modern techniques and their applications.* [S.l.]: John Wiley & Sons, 1999. v. 40.

[9] GALE, D. The Game of Hex and the Brouwer Fixed-Point Theorem. *The American Mathematical Monthly*, 86(10), 818-827, 1979.

[10] MAWHIN, J. Variations on the Brouwer Fixed Point Theorem: A Survey. *Mathematics*, 8(4), n. 501, p. 1–14, 2020.

[11] SOUZA, C. S. *Existência de soluções periódicas e permanência de soluções de equações diferenciais funcionais com retardo.* Dissertação (Mestrado em Matemática) — Universidade Estadual Paulista (UNESP), São José do Rio Preto, 2018.

[12] URAI, K. Fixed point theorems and the existence of economic equilibria based on conditions for local directions of mappings. In: *Kusuoka S., Maruyama T. (eds) Advances in Mathematical Economics. Advances in Mathematical Economics*, vol 2. Springer, Tokyo, 2000.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 84–91

Espaços Vetoriais Quocientes e Teoremas do Isomorfismo

Rafael Froner Prando[†]
Orientador(a): Profa. Dra. Marta Cilene Gadotti

Resumo: Esse texto apresenta um recorte de um trabalho de iniciação científica que tem como objetivo estudar resultados clássicos da álgebra linear aplicados em espaços de dimensão infinita. Neste trabalho definimos espaços vetoriais quocientes, demonstramos alguns resultados e propriedades deles e de famílias de subespaços afins. Também apresentamos extensões dos três teoremas clássicos do isomorfismo para espaços vetoriais.

Palavras-chave: relações de equivalência; espaços afins; espaços vetoriais de dimensão infinita; espaços de Banach

1 Espaços Vetoriais Quocientes e Afins

Para definir espaços quocientes de um espaço vetorial V precisamos primeiro relembrar o conceito de relação de equivalência. Se A é um conjunto não vazio, $R \subset A \times A$ é uma relação sobre A e escrevemos $x \sim y \iff (x, y) \in R$.

Definição 1.1. Uma relação é chamada relação de equivalência se valem as seguintes propriedades:

a) Reflexiva: $x \sim x \ \forall x \in A$.

b) Simétrica: se $x \sim y$, então $y \sim x \ \forall x, y \in A$.

c) Transitiva: se $x \sim y$ e $y \sim z$, então $x \sim z \ \forall x, y, z \in A$.

Exemplo 1.2. Se $A = \mathbb{Z}$, a relação $x \sim y \iff x \equiv y \pmod m$ é uma relação de equivalência.

Daqui em diante usaremos o símbolo $\equiv$ para denotar uma relação de equivalência qualquer. Suponha agora uma relação $\equiv$ em um conjunto A. Para cada $x \in A$, definimos $\overline{x} = \{y \in A \mid y \equiv x\}$. $\overline{x}$ é chamada classe de equivalência de x e é um subconjunto de A.

Teorema 1.3. *As classes de equivalência satisfazem as seguintes propriedades:*

a) $x \in \overline{x}$.

b) $\overline{x} = \overline{y} \iff x \equiv y$.

c) $\forall x, y \in A$, $\overline{x} = \overline{y}$ *ou* $\overline{x} \cap \overline{y} = \emptyset$.

d) $A = \bigcup\limits_{x \in A} \overline{x}$.

Prova: a) é consequência direta da propriedade reflexiva. Para mostrar b), basta notar que $\overline{x} = \overline{y} \iff x \in \overline{y}$ e $y \in \overline{x} \iff x \equiv y$. c) Suponha $\overline{x} \neq \overline{y}$ e tome $a \in \overline{x} \cap \overline{y}$. Daí $a \equiv x$

[†]Bolsista do Programa de Educação Tutorial - PET

e $a \equiv y \Rightarrow x \equiv y \Rightarrow \overline{x} = \overline{y}$, uma contradição. Portanto $\overline{x} \cap \overline{y} = \emptyset$ ou $\overline{x} = \overline{y}$. Por fim, d) é consequência direta do item a). ∎

Vemos então que as classes de equivalência $\overline{0}, \overline{1}, \ldots, \overline{m-1}$ do exemplo 1.2 formam uma união disjunta de $\mathbb{Z}$. Pelo teorema anterior observamos que qualquer conjunto A pode ser particionado da mesma forma, isto é, por meio das classes de equivalência. O conjunto das classes de equivalência em $\mathbb{Z}$ herda as operações de adição e multiplicação. Esse é um fenômeno bastante comum na álgebra e será bastante útil no estudo de espaços vetoriais.

Sejam V espaço vetorial sobre um corpo F e W subespaço de V. W determina a seguinte relação de equivalência sobre V:

$$\alpha \equiv \beta \iff \alpha - \beta \in W. \tag{1.1}$$

Essa relação é reflexiva, pois $\alpha - \alpha = 0 \in W$. É simétrica porque se $\alpha - \beta \in W$, então $-(\alpha - \beta) = \beta - \alpha \in W$. E também é transitiva, já que se $\alpha - \beta, \beta - \gamma \in W$, então $(\alpha - \beta) + (\beta - \gamma) = \alpha - \gamma \in W$. É também evidente que essa relação é inteiramente dependente do subespaço W considerado.

Cabe mencionar que todos os resultados apresentados a seguir são válidos para espaços tanto de dimensão finita quanto de dimensão infinita, exceto quando especificado espaço de dimensão finita.

Definição 1.4. Seja W subespaço de V e $\equiv$ a relação em (1.1). Se $\alpha \in V$, o conjunto de todas as classes de equivalência $\{\overline{\alpha} \mid \alpha \in V\}$ será denotado por V/W.

Assim $\overline{\alpha} = \{\beta \in V \mid \beta \equiv \alpha\}$ e $V/W = \{\overline{\alpha} \mid \alpha \in V\}$. Veja que os elementos de V/W são subconjuntos de V. Logo, V/W é uma coleção de elementos do conjunto das partes de V: $P(V)$.

Definição 1.5. Se W é subespaço de V e $\alpha \in V$, então o conjunto $\alpha + W = \{\alpha + \gamma \mid \gamma \in W\}$ é chamado de classe lateral de W.

Nota-se que $\beta \in \alpha + W \iff \alpha - \beta \in W$, ou seja, a classe lateral $\alpha + W$ é o mesmo conjunto que a classe de equivalência $\overline{\alpha}$ sobre $\equiv$. Então V/W é o conjunto de todas as classes laterais de W. Em particular, a classe de equivalência $\overline{\alpha} = \alpha + W$ pode ser interpretada geometricamente como a translação do subespaço W pelo vetor α. A classe lateral $\alpha + W$ é também chamada de subespaço afim de V.

Definição 1.6. O conjunto de todos os subespaços afins de V é denotado por $\mathcal{A}(V)$.

Assim, $A \in \mathcal{A}(V) \iff A = \alpha + W$ para algum subespaço W de V e $\alpha \in V$. Veja que um subespaço afim $A = \alpha + W$ não é subespaço vetorial de V, exceto se $\alpha = 0$. Como $\mathcal{A}(V)$ consiste em todas as classes laterais de todos os subespaços de V, temos $V/W \subseteq \mathcal{A}(V) \subseteq P(V)$. V/W é chamado de quociente de V por W. Antes de mostrarmos que V/W tem estrutura de espaço vetorial, vejamos algumas propriedades a respeito de subespaços afins.

Teorema 1.7. *Sejam V espaço vetorial sobre F, $\mathrm{Hom}_F(V, V')$, o conjunto das transformações lineares de V em V' e $\mathcal{A}(V)$ o conjunto dos subespaços afins de V.*

a) *Se $\{A_i \mid i \in \Delta\}$ é uma coleção de subespaços afins, então $\bigcap_{i\in\Delta} A_i = \emptyset$ ou $\bigcap_{i\in\Delta} A_i \in \mathcal{A}(V)$.*

b) *Se $A, B \in \mathcal{A}(V)$, então $A + B \in \mathcal{A}(V)$.*

c) *Se* $A \in \mathcal{A}(V)$ *e* $x \in F$*, então* $xA \in \mathcal{A}(V)$*.*

d) *Se* $A \in \mathcal{A}(V)$ *e* $T \in \mathrm{Hom}(V, V')$*, então* $T(A) \in \mathcal{A}(V')$*.*

e) *Se* $A' \in \mathcal{A}(V')$ *e* $T \in \mathrm{Hom}(V, V')$*, então* $T^{-1}(A') = \emptyset$ *ou* $T^{-1}(A') \in \mathcal{A}(V)$*.*

Prova: Os itens b), c), d) são imediatos. Mostremos a) e e). Suponha $A_i = \alpha_i + W_i \ \forall i \in \Delta$ com W_i subespaço de V e $\alpha_i \in V$. Suponha $\bigcap_{i\in\Delta} A_i \neq \emptyset$ e $\beta \in \bigcap_{i\in\Delta} A_i$. Então, para cada $i \in \Delta$, $\beta = \alpha_i + \gamma_i$, $\gamma_i \in W_i$. Então $\beta + W_i = \alpha_i + W_i$ e $\bigcap_{i\in\Delta} A_i = \bigcap_{i\in\Delta}(\beta + W_i)$. Mostremos que $\bigcap_{i\in\Delta}(\beta + W_i) = \beta + (\bigcap_{i\in\Delta} W_i)$.

Imediatamente observamos que $\beta + (\bigcap_{i\in\Delta} W_i) \subseteq \bigcap_{i\in\Delta}(\beta + W_i)$. Seja $\alpha \in \bigcap_{i\in\Delta}(\beta + W_i)$. Então, para $i \neq j$, $\alpha = \beta + \delta_i = \beta + \delta_j$, $\delta_i \in W_i$, $\delta_j \in W_j$. Porém isso implica que $\delta_i = \delta_j$ e $\alpha \in \beta + (W_i \cap W_j)$. Logo $\bigcap_{i\in\Delta}(\beta + W_i) \subseteq \beta + (\bigcap_{i\in\Delta} W_i)$. Portanto $\bigcap_{i\in\Delta}(\beta + W_i) = \beta + (\bigcap_{i\in\Delta} W_i)$ e como $\beta + (\bigcap_{i\in\Delta} W_i) \in \mathcal{A}(V)$, a demonstração de *a*) está completa.

Para o item e), se $T^{-1}(A') = \emptyset$, não há o que provar. Suponha então $T^{-1}(A') \neq \emptyset$, ou seja, $\{\alpha \in V \mid T(\alpha) \in A'\} \neq \emptyset$. Como $A' \in \mathcal{A}(V')$, então $A' = \alpha' + W'$ para algum $\alpha' \in V'$ e algum W' subespaço vetorial de V'. Mostremos que $T^{-1}(A') = \alpha + W$ para algum $\alpha \in V$ e W subespaço vetorial de V.

Como W' é subespaço de V', considere $W = T^{-1}(W')$ (que é um subespaço vetorial de V). Tome também o vetor $\alpha \in T^{-1}(\{\alpha'\})$.

Se $\beta \in \alpha + W$, então $\beta = \alpha + w$ com $w \in W$ e

$$T(\beta) = T(\alpha) + T(w) = \alpha' + w' \in \alpha' + W'.$$

Portanto $\alpha + W \subset T^{-1}(A')$.

Agora, se $\beta \in T^{-1}(A')$, então $T(\beta) = \alpha' + w'$ com $w' \in W'$. Como $W = T^{-1}(W')$, $\exists w \in W$ tal que $T(w) = w'$, então $T(\beta) = \alpha' + T(w)$. Basta então escolher $\alpha \in T^{-1}(\{\alpha'\})$ tal que $\alpha = \beta - w$, o que sabemos ser possível uma vez que

$$T(\alpha) = T(\beta) - T(w) = \alpha' + T(w) - T(w) = \alpha'. \qquad \blacksquare$$

Podemos generalizar o Teorema 1.7 introduzindo o conceito de uma aplicação afim entre espaços vetoriais. Se $\alpha \in V$, a translação por α é a função $S_\alpha : V \to V$ dada por $S_\alpha(\beta) = \alpha + \beta$, então qualquer classe lateral $\alpha + W$ pode ser vista como $S_\alpha(W)$. Note que quando $\alpha \neq 0$, S_α não é transformação linear.

Definição 1.8. Sejam V e V' espaços vetoriais sobre F. A função $f : V \to V'$ é chamada transformação afim se $f = S_\alpha T$ para alguma $T \in \mathrm{Hom}_F(V, V')$ e $\alpha \in V'$. O conjunto de todas as funções afins de V em V' será denotado por $\mathrm{Aff}_F(V, V')$.

É fácil perceber que $\mathrm{Hom}_F(V, V') \subseteq \mathrm{Aff}_F(V, V') \subseteq (V')^V$, onde $(V')^V$ denota o conjunto de todas as funções de V' em V. O Teorema 1.7*d*) pode então ser reescrito da seguinte forma:

Teorema 1.9. *Se* $A \in \mathcal{A}(V)$ *e* $f \in \mathrm{Aff}_F(V, V')$*, então* $f(A) \in \mathcal{A}(V)$*.*

Voltamos nossa atenção agora para o subconjunto V/W de $\mathcal{A}(V)$. O conjunto das classes laterais de W pode ter estrutura de espaço vetorial. Definindo a soma $\dot{+}$ por $\overline{\alpha} \dot{+} \overline{\beta} = \overline{\alpha + \beta}$, verifiquemos que ela independe da escolha de representantes. Sejam $\alpha_1 \in \overline{\alpha}$ e $\beta_1 \in \overline{\beta}$, então $\alpha_1 - \alpha, \beta_1 - \beta \in W$. Logo $(\alpha_1 + \beta_1) - (\alpha + \beta) \in W$ e $\overline{\alpha_1 + \beta_1} = \overline{\alpha + \beta}$. Portanto $\dot{+} : V/W \times V/W \to V/W$ é uma função bem definida que satisfaz as propriedades da adição em um espaço vetorial. Sendo assim, a denotaremos a partir de agora apenas por $+$.

Definimos a multiplicação por escalar de forma semelhante: $x\overline{\alpha} = \overline{x\alpha}$. Novamente a operação independe dos representantes, pois: $\alpha_1 \in \overline{\alpha} \Rightarrow \alpha_1 - \alpha \in W \Rightarrow x\alpha_1 - x\alpha \in W \Rightarrow$

$\overline{x\alpha_1} = \overline{x\alpha}$. Logo $(x, \overline{\alpha}) \mapsto \overline{x\alpha}$ é uma função bem definida de $F \times V/W$ em V/W que satisfaz as propriedades da multiplicação por escalar em um espaço vetorial.

Teorema 1.10. *O conjunto V/W munido das operações $(\overline{\alpha}, \overline{\beta}) \mapsto \overline{\alpha + \beta}$ e $(x, \overline{\alpha}) \mapsto \overline{x\alpha}$ definidas acima é um espaço vetorial sobre F.*

Veja que a forma como definimos as operações de soma e multiplicação por escalar implica que a aplicação natural $\Pi : V \to V/W$ dada por $\Pi(\alpha) = \overline{\alpha}$ é uma transformação linear. Observamos também que Π é sobrejetora com $\ker \Pi = W$. Assim, se $i : W \to V$ denota a inclusão de W em V, temos a seguinte sequência:

$$W \overset{i}{\hookrightarrow} V \overset{\Pi}{\to} \frac{V}{W},$$

ou seja, podemos usar o Teorema do Núcleo e Imagem para inferir que:

Teorema 1.11. *Se V é espaço vetorial de dimensão finita sobre F e W é subespaço de V, então* $\dim V = \dim W + \dim V/W$.

2 O Espaço de Banach V/W

Nessa seção mostraremos sob quais condições o subespaço vetorial V/W é um espaço de Banach.

Definição 2.1. Seja V um espaço vetorial sobre F. Uma função $\| \cdot \| : V \to [0, +\infty)$ é chamada norma se para todo $\alpha, \beta \in V$ e $x \in F$ satisfaz as seguintes propriedades:

a) $\|\alpha\| = 0 \iff \alpha = 0$.

b) $\|x\alpha\| = |x| \cdot \|\alpha\|$.

c) $\|\alpha + \beta\| \leqslant \|\alpha\| + \|\beta\|$.

Chamamos de seminorma uma função que satisfaz apenas as propriedades *b*) e *c*).

Definição 2.2. Seja $(V, \| \cdot \|_V)$ um espaço vetorial normado e $\alpha_n \in V$. Dizemos que uma sequência (α_n) é sequência de Cauchy se dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $m, n > n_0 \Rightarrow \|\alpha_n - \alpha_m\|_V < \varepsilon$.

Definição 2.3. Um espaço vetorial normado V é chamado de completo se toda sequência se Cauchy em V converge para um elemento de V.

Definição 2.4. Um espaço vetorial V é chamado espaço de Banach se é normado e completo.

Definição 2.5. Um subconjunto K de um espaço vetorial normado V é fechado quando toda sequência convergente (α_n) de vetores de K converge para um vetor de K.

Para garantir que V/W é espaço de Banach, precisamos primeiramente definir uma norma nesse espaço. Veja que é interessante nesse caso definirmos essa norma em termos da norma de V para podermos aproveitar algumas de suas propriedades.

Teorema 2.6. *Seja W subespaço vetorial de um espaço normado $(V, \| \cdot \|_V)$ sobre um corpo F. Dado $\alpha \in V$, defina:*

$$\|\alpha + W\| = d(\alpha, W) = \inf_{w \in W} \|\alpha - w\|_V.$$

São válidas as seguintes propriedades:

a) $\|\cdot\|$ *está bem definida.*

b) $\|\cdot\|$ *é uma seminorma em* V/W*.*

c) *Se* W *é fechado, então* $\|\cdot\|$ *é uma norma em* V/W*.*

Prova: a) Mostremos que se $\alpha, \beta \in V$ e $\alpha \equiv \beta \iff \alpha - \beta \in W$, então $\|\alpha + W\| = \|\beta + W\|$. Note que se $w \in W$, então:

$$\|\alpha - w\|_V = \|\alpha - \beta + \beta - w\|_V = \|\beta + (\alpha - \beta) - w\|_V = \|\beta - w'\|_V,$$

onde $w' = -\alpha + \beta + w \in W$. Portanto

$$\|\alpha + W\| = \inf_{w \in W} \|\alpha - w\|_V = \inf_{w' \in W} \|\beta - w'\| = \|\beta + W\|.$$

Para o item b), tome primeiro $x \in F$. Se $x = 0$, temos que $x\alpha = 0$. Daí, como W é subespaço vetorial, $0 \in W$ e consequentemente

$$\|x(\alpha + W)\| = \|(x\alpha) + W\| = d(0, W) = 0 = 0 \cdot \inf_{w \in W} \|\alpha - w\|_V = |x| \cdot \|\alpha + W\|.$$

Se $x \neq 0$, existe $x^{-1} \in F$, $x^{-1}w \in W$ e podemos usar que $\|\cdot\|_V$ é norma para obter:

$$\begin{aligned} \|x(\alpha + W)\| &= \inf_{w \in W} \|x\alpha - w\|_V \\ &= \inf_{w \in W} |x| \cdot \|\alpha - x^{-1}w\|_V \\ &= |x| \cdot \inf_{w \in W} \|\alpha - x^{-1}w\|_V \\ &= |x| \cdot \inf_{w \in W} \|\alpha - w\|_V \\ &= |x| \cdot \|\alpha + W\|. \end{aligned}$$

Agora, sejam $\alpha + W, \beta + W \in V/W$. Pela forma como definimos a soma em V/W, vale que $\|(\alpha+W)+(\beta+W)\| = \|(\alpha+\beta)+W\|$. Veja que como $-w \in W$, temos que $-w-w \in W$, ou seja:

$$\begin{aligned} \|(\alpha + W) + (\beta + W)\| &= \inf_{w \in W} \|\alpha + \beta - w\|_V \\ &= \inf_{w \in W} \|\alpha + \beta - w - w\|_V \\ &\leqslant \inf_{w \in W} \|\alpha - w\|_V + \|\beta - w\|_V \\ &\leqslant \inf_{w \in W} \|\alpha - w\|_V + \inf_{w \in W} \|\beta - w\|_V \\ &= \|\alpha + W\| + \|\beta + W\|. \end{aligned}$$

Portanto $\|\cdot\|$ é seminorma em V/W.

Em c) mostremos que se W é fechado, então $\|\alpha + W\| = 0 \Leftrightarrow \alpha + W = 0 + W$.

$(\Rightarrow)$: Se $\|\alpha + W\| = 0$, então $\inf_{w \in W} \|\alpha - w\|_V = 0$. Veja que $\|\alpha - w\|_V = 0 \Leftrightarrow \alpha = w$. Por W ser fechado, sabemos que deve existir uma sequência de vetores $(\alpha_n) \subset W$ tal que $\alpha_n \to \alpha$. Como W é subespaço de V, temos que $\alpha \in W$ e portanto $\alpha + W = 0 + W$.

$(\Leftarrow)$: Se $\alpha + W = 0 + W$, então $\|\alpha + W\| = \|0 + W\| = \inf_{w \in W} \|w\|$. Como W é subespaço vetorial, $0 \in W$ e, portanto, $\inf_{w \in W} \|w\| = 0$. Ou seja, $\|\alpha + W\| = 0$. ∎

Por fim, provemos sob quais condições V/W é um espaço vetorial normado e completo.

Teorema 2.7. *Se W é subespaço vetorial fechado de um espaço de Banach V, então V/W é também espaço de Banach.*

Prova: Sendo W fechado, já provamos que V/W é espaço normado com a norma definida no Teorema 2.6. Resta apenas provar que V/W é completo. Para isso, tome $(\alpha_n + W)$ sequência de Cauchy em V/W. Mostremos que $(\alpha_n + W)$ possui subsequência convergente e consequentemente converge.

Sendo $(\alpha_n + W)$ sequência de Cauchy, sabemos que existe uma subsequência $(\alpha_{n_k} + W)$ tal que $\forall k \in \mathbb{N}$ vale que:

$$\|(\alpha_{n_{k+1}} - \alpha_{n_k}) + W\| = \|(\alpha_{n_{k+1}} + W) - (\alpha_{n_k} + W)\| < \frac{1}{2^k}.$$

Nosso objetivo será construir vetores $\beta_k \in V$ tais que a sequência $(\alpha_{n_k} - \beta_k)$ seja convergente em V. Veja também que do fato que $\beta_k \in V$, as classes laterais determinadas por α_{n_k} e $\alpha_{n_k} - \beta_k$ são idênticas. Ponha $\beta_1 = 0$. Temos:

$$\inf_{\beta \in W} \|(\alpha_{n_1} - \beta_1) - (\alpha_{n_2} - \beta)\|_V = \inf_{\beta \in W} \|(\alpha_{n_1} - \alpha_{n_2}) - \beta\|_V = \|(\alpha_{n_1} - \alpha_{n_2}) + W\| < \frac{1}{2}.$$

Isso mostra que existe $\beta_2 \in W$ tal que:

$$\|(\alpha_{n_1} - \beta_1) - (\alpha_{n_2} - \beta_2)\|_V < \frac{1}{2}.$$

E como $\beta_2 \in W$:

$$\inf_{\beta \in W} \|(\alpha_{n_2} - \beta_2) - (\alpha_{n_3} - \beta)\|_V = \inf_{\beta \in W} \|(\alpha_{n_2} - \alpha_{n_3}) - \beta\|_V = \|(\alpha_{n_2} - \alpha_{n_3}) + W\| < \frac{1}{2^2}.$$

Portanto existe $\beta_3 \in W$ tal que:

$$\|(\alpha_{n_2} - \beta_2) - (\alpha_{n_3} - \beta_3)\|_V < \frac{1}{2^2}.$$

Prosseguindo indutivamente com esse raciocínio, construímos $\delta_k = \alpha_{n_k} - \beta_k$ tais que

$$\|\delta_k - \delta_{k+1}\|_V < \frac{1}{2^k}.$$

Veja que isso é suficiente para mostrar que (δ_k) é sequência de Cauchy em V. Como V é completo, a sequência converge. Digamos $\delta_k \to \delta$.

Para concluir a demonstração, basta notar que como $\beta_k \in W$:

$$\|(\alpha_{n_k} + W) - (\delta + W)\| = \|\alpha_{n_k} - \beta_k - \delta + W\| = \|\delta_k - \delta + W\| \leqslant \|\delta_k - \delta\|_V.$$

Como $\delta_k \to \delta$, temos que $\|\delta_k - \delta\|_V \to 0$. Assim concluímos que $(\alpha_{n_k} + W)$ é subsequência convergente de $(\alpha_n + W)$. Ou seja, a sequência $(\alpha_n + W)$ é convergente, V/W é completo e portanto de Banach. ■

3 Teoremas do Isomorfismo

Finalmente, demonstraremos os três teoremas do isomorfismo em espaços vetoriais.

Teorema 3.1. *Seja $T \in \mathrm{Hom}_F(V, V')$ e W subespaço de V tal que $T(W) = 0$. Seja $\Pi : V \to V/W$ a aplicação natural. Então existe uma única $\overline{T} \in \mathrm{Hom}(V/W, V')$ tal que o diagrama a seguir é comutativo.*

$$\begin{tikzcd} V \arrow[rr, "T"] \arrow[dr, "\Pi"'] & & V' \\ & \dfrac{V}{W} \arrow[ur, "\overline{T}"'] & \end{tikzcd} \tag{3.1}$$

Prova: Defina $\overline{T}$ pondo $\overline{T}(\overline{\alpha}) = T(\alpha)$. Se $\alpha_1 \in \overline{\alpha}$: $\alpha_1 - \alpha \in W \Rightarrow T(\alpha_1 - \alpha) = 0 \Rightarrow T(\alpha_1) = T(\alpha) \Rightarrow \overline{T}(\overline{\alpha_1}) = \overline{T}(\overline{\alpha})$. Assim, $\overline{T}$ independe dos representantes. Também temos que :

$$\overline{T}(\overline{\alpha} + x\overline{\beta}) = \overline{T}(\overline{\alpha + x\beta}) = T(\alpha + x\beta) = T(\alpha) + xT(\beta) = \overline{T}(\overline{\alpha}) + xT(\overline{\beta}),$$

ou seja, $\overline{T} \in \mathrm{Hom}_F(V/W, V')$. E como $\overline{T}\Pi(\alpha) = \overline{T}(\overline{\alpha}) = T(\alpha)$, o diagrama (3.1) comuta. Resta apenas provar a unicidade. Se $T' \in \mathrm{Hom}_F(V/W, V')$ é outra aplicação tal que $T'\Pi = T$, então $T' = \overline{T}$ em $\mathrm{Im}\,\Pi$. Porém, como Π é sobrejetora, $T' = \overline{T}$. ■

Teorema 3.2 **(Primeiro Teorema do Isomorfismo).** *Se $T \in \mathrm{Hom}_F(V, V')$, então $V/\ker T \cong \mathrm{Im}\,T$.*

Prova: T pode ser interpretada como uma transformação linear sobrejetora de V em $\mathrm{Im}\,T$. Aplicando o Teorema 3.1. obtemos uma transformação linear única $\overline{T} : V/\ker T \to \mathrm{Im}\,T$ para a qual o diagrama a seguir é comutativo

$$\begin{tikzcd} V \arrow[rr, "T"] \arrow[dr, "\Pi"'] & & \mathrm{Im}\,T \\ & \dfrac{V}{\ker T} \arrow[ur, "\overline{T}"'] & \end{tikzcd}$$

Provemos que $\overline{T}$ é isomorfismo. Como $\overline{T}\Pi = T$ e $T : V \to \mathrm{Im}\,T$ é sobrejetora, então $\overline{T}$ é sobrejetora. Suponha agora $\overline{\alpha} \in \ker \overline{T}$. Então $0 = \overline{T}(\overline{\alpha}) = \overline{T}\Pi(\alpha) = T(\alpha) \Rightarrow \alpha \in \ker T$. Isso implica então que $\Pi(\alpha) = \overline{0}$ e $\overline{\alpha} = \overline{0}$, portanto $\overline{T}$ é injetora. ■

O Segundo Teorema do Isomorfismo trabalha com vários quocientes. Suponha W subespaço de V e considere $\Pi : V \to V/W$. Se W' é subespaço de V tal que $W \subseteq W'$, então $\Pi(W')$ é subespaço de V/W. Ou seja, podemos formar o quociente $(V/W)/\Pi(W')$. Pelo Teorema 3.2, $\Pi(W')$ é isomorfo a W'/W. Assim, podemos reescrever $(V/W)/\Pi(W')$ como $(V/W)/(W'/W)$.

Teorema 3.3 **(Segundo Teorema do Isomorfismo).** *Suponha $W \subseteq W'$ subespaços vetoriais de V. Então $\dfrac{V/W}{W'/W} \cong \dfrac{V}{W'}$.*

Prova: Sejam $\Pi : V \to V/W$ e $\Pi' : (V/W) \to (V/W)/\Pi(W')$ as projeções naturais. Tome $T = \Pi'\Pi : V \to (V/W)/\Pi(W')$. Como Π e Π' são sobrejetoras, T é uma transformação linear sobrejetora. Evidentemente $W' \subseteq \ker T$. Seja $\alpha \in \ker T$, então $0 = \Pi'\Pi(\alpha)$. Então

$\overline{\alpha} = \Pi(\alpha) \in \Pi(W')$. Seja $\beta \in W'$ tal que $\Pi(\beta) = \Pi(\alpha)$, então $\Pi(\beta - \alpha) = 0 \Rightarrow \beta - \alpha \in \ker \Pi = W \subseteq W'$. Em particular $\alpha \in W'$. Provamos então $\ker T = W'$. Aplicando o Teorema 3.2, temos então $(V/W)/\Pi(W') = \operatorname{Im} T \cong V/\ker T = V/W'$. ■

O Terceiro Teorema do Isomorfismo trata de somas e quocientes.

Teorema 3.4 (Terceiro Teorema do Isomorfismo)**.** *Suponha W e W' subespaços de V. Então $(W + W')/W \cong W'/(W \cap W')$.*

Prova: Seja $\Pi : W + W' \to (W + W')/W$ a projeção natural. A inclusão de W em $W + W'$ composta com Π nos dá uma transformação linear $T : W' \to (W + W')/W$. Como $\ker \Pi = W$, $\ker T = W \cap W'$. Mostremos que T é sobrejetora. Considere $\gamma \in (W + W')/W$. γ é a classe lateral de W da forma $\gamma = \delta + W$ com $\delta \in W + W'$. Logo, $\delta = \alpha + \beta$, $\alpha \in W, \beta \in W'$. Porém $\alpha + W = W$. Então $\gamma = \delta + W = (\beta + \alpha) + W = \beta + W$. Em particular $T(\beta) = \beta + W = \gamma$ e T é sobrejetora. Pelo Teorema 3.2: $(W + W')/W = \operatorname{Im} T \cong W'/\ker T = W'/(W \cap W')$. ■

Por fim, vejamos uma aplicação típica dos teoremas do isomorfismo.

Exemplo 3.5. Suponha V soma direta dos subespaços $V_1, \ldots, V_n$. Assim, $V = V_1 \oplus \cdots \oplus V_n$. Como $V_i \cap (\sum_{j \neq i} V_j) = \{0\}$, o Teorema 3.4 implica que:

$$\begin{aligned} \frac{V}{V_i} = \frac{V_i + \sum_{j \neq i} V_j}{V_i} &\cong \frac{\sum_{j \neq i} V_j}{V_i \cap \left(\sum_{j \neq i} V_j\right)} = \frac{\sum_{j \neq i} V_j}{\{0\}} \\ &= V_1 \oplus \cdots \oplus V_{i-1} \oplus V_{i+1} \oplus \cdots \oplus V_n. \end{aligned}$$

Agradecimentos: Agradeço à minha orientadora Marta pela ajuda com o desenvolvimento do projeto e paciência durante a orientação e também a meus amigos do PET pelo apoio com assuntos extra-acadêmicos.

Abstract: This text presents a fraction of a scientific initiation project which aims to study the classic results of linear algebra applied to infinite dimensional vector spaces. In this text we define quotient vector spaces and prove some theorems and properties concerning those as well as families of affine subspaces. We also present extensions of the three classic isomorphism theorems for vector spaces.

Keywords: equivalence relations; affine spaces; infinite dimensional vector spaces; Banach spaces

Referências Bibliográficas

[1] BROWN, William Clough. **A second course in Linear Algebra.** Wiley, 1988.

[2] HEIL, Christopher. **Functional Analysis lecture notes.**

[3] KREYSZIG, Erwin. **Introductory functional analysis with applications.** New York: Wiley, 1978.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 92–99

Grupos Abelianos Livres

Richard Guilherme dos Santos
Orientador(a): Prof. Dr. João Peres Vieira

Resumo: Grupos Abelianos são um dos conceitos inicias vistos em cursos de Álgebra Abstrata, muito úteis na elaboração de conceitos como anéis, espaços vetoriais e corpos. Alguns desses grupos, chamados de Grupos Abelianos Livres, possuem propriedades interessantes, além de induzirem definições em outras áreas, como a da Topologia Algébrica e Álgebra Geométrica. Neste trabalho serão vistos alguns grupos com tal propriedade e alguns resultados da definição, como a unicidade da propriedade, o grupo abeliano livre Z^m e o grupo das funções quase nulas.

Palavras-chave: Grupos Abelianos Livres; Grupos Abelianos; Topologia Algébrica; Álgebra Abstrata

1 Definições

Definição 1.1. Um grupo abeliano é um conjunto A, junto de uma função "+" definida como;

$$\begin{aligned} + : A \times A &\longrightarrow A \\ (a,b) &\longmapsto a+b \end{aligned}$$

que segue as seguintes propriedades:

1. Existe um elemento $0 \in A$ tal que $0 + a = a = a + 0, \forall a \in A$;
2. Para cada $a \in A$, existe $(-a) \in A$ tal que $(-a) + a = 0$;
3. Para quaisquer $a, b, c \in A$, $(a + b) + c = a + (b + c)$;
4. Para quaisquer $a, b \in A, a + b = b + a$.

Definição 1.2. Denotaremos por $\mathrm{Hom}(A, X)$ o conjunto de todos os homomorfismos $A \longrightarrow X$.

Definição 1.3. Se $(A, +_A), (B, +_B)$ são dois grupos, abelianos ou não, definimos sua soma direta $A \oplus B := A \times B$ como a soma $\boxplus$ dada por

$$(a_1, b_1) \boxplus (a_2, b_2) = (a_1 +_A a_2, b_1 +_B b_2).$$

Definição 1.4. Se $\phi : (A, +_A) \longrightarrow (X, +_X)$ e $\psi : (A, +_A) \longrightarrow (X, +_X)$ são homomorfismos, então definimos sua soma como:

$$\begin{aligned} \phi \mathbin{\triangle\!\!\!\!+} \psi : \quad (A, +_A) &\longrightarrow (X, +_X) \\ a &\longmapsto (\phi \mathbin{\triangle\!\!\!\!+} \psi)(a) := \phi(a) +_X \psi(a) \end{aligned}$$

Definição 1.5. Dado um conjunto X, um grupo abeliano G e uma função $i : X \longrightarrow G$, dizemos que (G, i) tem a propriedade universal para X se, para qualquer grupo abeliano A e qualquer função $j : X \longrightarrow A$ existe um único ϕ que comuta o diagrama, isto é $\phi \circ i = j$;

$$
\begin{tikzcd}
& X \arrow[dl, "i"'] \arrow[dr, "j"] & \\
G \arrow[rr, "\phi"'] & & A
\end{tikzcd}
$$

Quando isso acontece dizemos que G é um grupo abeliano livre sobre os geradores $i(X)$. Além disso, se $i(X)$ for linearmente independente dizemos que $i(X)$ é uma base para G.

2 Teoremas e Proposições

Proposição 2.1. *Suponha que (G, i) e (H, j) tenham ambos a propriedade universal para X. Então existe um único isomorfismo $\phi : G \longrightarrow H$ com $j = \phi \circ i$.*

Prova: Pela propriedade universal de (G, i) e (H, j), existem únicos homomorfismos $\phi : G \longrightarrow H$ e $\psi : H \longrightarrow G$ tais que

$$
\begin{aligned}
j &= \phi \circ i \\
i &= \psi \circ j
\end{aligned}
$$

Resta provarmos que tal ϕ é bijetor. Separemos em dois casos:

1. ϕ é injetor;
   Considere os seguintes diagramas:

$$
\begin{tikzcd}
& X \arrow[dl, "i"'] \arrow[d, "j"] \arrow[dr, "i"] \arrow[dl, phantom, shift left=3ex, "\circlearrowleft"] \arrow[dr, phantom, shift right=3ex, "\circlearrowleft"] & \\
G \arrow[r, "\phi"] \arrow[rr, bend right, "\psi \circ \phi"'] & H \arrow[r, "\psi"] & G
\end{tikzcd}
\qquad
\begin{tikzcd}
& X \arrow[dl, "i"'] \arrow[dr, "i"] & \\
G \arrow[rr, "1_G"'] & & G
\end{tikzcd}
$$

   Assim, pela unicidade da propriedade universal

$$\psi \circ \phi = 1_G \Rightarrow \phi \text{ é injetor.}$$

2. ϕ é sobrejetor;
   Considere os seguintes diagramas:

$$
\begin{tikzcd}
& X \arrow[dl, "j"'] \arrow[d, "i"] \arrow[dr, "j"] \arrow[dl, phantom, shift left=3ex, "\circlearrowleft"] \arrow[dr, phantom, shift right=3ex, "\circlearrowleft"] & \\
H \arrow[r, "\psi"] \arrow[rr, bend right, "\phi \circ \psi"'] & G \arrow[r, "\phi"] & H
\end{tikzcd}
\qquad
\begin{tikzcd}
& X \arrow[dl, "j"'] \arrow[dr, "j"] & \\
H \arrow[rr, "1_H"'] & & H
\end{tikzcd}
$$

Assim, pela unicidade da propriedade universal

$$\phi \circ \psi = 1_H \Rightarrow \phi \text{ é sobrejetor.}$$

Portanto ϕ é um isomorfismo, ou seja, grupos com a mesma propriedade universal são isomorfos. ■

Proposição 2.2. *Se X é um conjunto unitário, isto é, $X = \{a\}$, defina;*

$$\begin{aligned} i : X &\longrightarrow \mathbb{Z} \\ a &\longmapsto i(a) := 1 \end{aligned}$$

Então $(\mathbb{Z}, i)$ tem a propriedade universal para X, ou seja, $\mathbb{Z}$ é um grupo abeliano livre sobre os geradores $i(X) = \{i(a) := 1\}$.

Prova: Sejam $(A, +_A)$ um grupo abeliano qualquer, $j : X \longrightarrow A$ e $\phi : \mathbb{Z} \longrightarrow A$ funções, onde ϕ é definida como:

$$\phi(n) = \phi(n1) = \begin{cases} \underbrace{j(a) +_A \cdots +_A j(a)}_{n \text{ vezes}}, & \text{se } n > 0, \\ 0, & \text{se } n = 0, \\ \underbrace{(-_A j(a)) +_A \cdots +_A (-_A j(a))}_{-n \text{ vezes}}, & \text{se } n < 0. \end{cases}$$

Note que

$$\phi(n) = n\phi(1) = nj(a).$$

Assim, temos que ϕ é um homomorfismo, pois para $n, m \in \mathbb{Z}$:

$$\begin{aligned} \phi(n+m) = \phi(n1+m1) = \phi((n+m)1) = (n+m)\phi(1) = (n+m)j(a) = \\ nj(a) +_A mj(a) = \phi(n) +_A \phi(m). \end{aligned}$$

Além disso, ϕ é o único homomorfismo de $\mathbb{Z}$ para A, pois suponha que exista um homomorfismo $\psi : \mathbb{Z} \longrightarrow A$ tal que $\psi \circ i = j$, temos que

$$\psi(1) = \psi(i(a)) = \psi \circ i(a) = j(a) = \phi(1).$$

Assim $\forall n \in \mathbb{Z}$:

$$\psi(n) = \psi(n1) = n\psi(1) = n\phi(1) = \phi(n1) = \phi(n).$$

Portanto $\psi = \phi$. ■

Corolário 2.3. *Para qualquer grupo abeliano A a função*

$$ev : \mathrm{Hom}(\mathbb{Z}, A) \longrightarrow A$$

definida para $f : \mathbb{Z} \longrightarrow A$ como $ev(f) = f(1)$ é um isomorfismo.

Demonstração. Devemos mostrar que ev é um homomorfismo bijetor, ou seja, mostremos que:

1. ev é um homomorfismo;
   Sejam $f, g \in \mathrm{Hom}(\mathbb{Z}, A)$:

   $$ev(f \triangle\!\!\!\!+ g) = (f \triangle\!\!\!\!+ g)(1) = f(1) +_A g(1) = ev(f) +_A ev(g).$$

   Logo ev é um homomorfismo.

2. *ev* é bijetora;
   Seja $a \in A$ qualquer, tome

$$\begin{aligned} j: \quad X = \{x\} &\longrightarrow A \\ x &\longmapsto j(x) := a \end{aligned}$$

Como $\mathbb{Z}$ tem a propriedade universal para qualquer conjunto com um elemento ele a tem para $X = \{x\}$. Então existe um único homomorfismo $f : \mathbb{Z} \longrightarrow A$ onde o diagrama

$$\begin{tikzcd} & X = \{x\} \arrow[dl, "i"'] \arrow[dr, "j"] \arrow[d, phantom, "\circlearrowleft"] & \\ \mathbb{Z} \arrow[rr, "f"'] & & A \end{tikzcd}$$

comuta, com $i(x) = 1$.

Dessa forma temos que $\exists! f \in \mathrm{Hom}(\mathbb{Z}, A)$ tal que

$$f \circ i(x) = j(x) \Rightarrow f(i(x)) = a \Rightarrow f(1) = a \Rightarrow ev(f) = a.$$

Assim *ev* é bijetor.

Portanto *ev* é um isomorfismo, ou seja, para qualquer grupo abeliano A, temos que $\mathrm{Hom}(\mathbb{Z}, A) \cong A$. ■

Proposição 2.4. *Suponha que* (G_1, i_1) *tenha a propriedade universal para* X_1 *e* (G_2, i_2) *para* X_2*. Sejam* X_1, X_2 *disjuntos. Defina:*

$$i : X_1 \cup X_2 \longrightarrow G_1 \oplus G_2$$

por

$$\begin{aligned} i(x) &= (i_1(x), 0_2) \ \textit{se } x \in X_1, \\ i(x) &= (0_1, i_2(x)) \ \textit{se } x \in X_2. \end{aligned}$$

Então $(G_1 \oplus G_2, i)$ *tem a propriedade universal para* $X_1 \cup X_2$*.*

Demonstração. Queremos provar que existe um único homomorfismo $\phi : G_1 \oplus G_2 \longrightarrow A$ tal que

$$\phi \circ i = j,$$

isto é, ϕ comuta o diagrama

$$\begin{tikzcd} & X_1 \cup X_2 \arrow[dl, "i"'] \arrow[dr, "j"] & \\ G_1 \oplus G_2 \arrow[rr, "\phi"'] & & A \end{tikzcd}$$

Por hipótese temos que existem únicos homomorfismos ϕ_1 e ϕ_2 tais que os diagramas

$$
\begin{tikzcd}
 & X_1 \arrow[dl, "i_1"'] \arrow[dr, "j|_{X_1}"] & \\
G_1 \arrow[rr, "\phi_1"'] & & A
\end{tikzcd}
\qquad
\begin{tikzcd}
 & X_2 \arrow[dl, "i_2"'] \arrow[dr, "j|_{X_2}"] & \\
G_2 \arrow[rr, "\phi_2"'] & & A
\end{tikzcd}
$$

comutam.

Agora defina ϕ como

$$\phi(g_1, g_2) = \phi_1(g_1) +_A \phi_2(g_2).$$

Como ϕ_1 e ϕ_2 são unicamente determinados, temos que ϕ é unicamente determinado. Além disso, como ϕ_1 e ϕ_2 são homomorfismos, então ϕ é um homomorfismo.

Por fim, observe que

$$
\begin{aligned}
\phi \circ i(x) = \phi(i(x)) &= \begin{cases} \phi(i_1(x), 0_2), & \text{se } x \in X_1 \\ \phi(0_1, i_2(x)), & \text{se } x \in X_2 \end{cases} \\
&= \begin{cases} \phi_1(i_1(x)) +_A \phi_2(0_2), & \text{se } x \in X_1 \\ \phi_1(0_1) +_A \phi_2(i_2(x)), & \text{se } x \in X_2 \end{cases} \\
&= \begin{cases} j\big|_{X_1}(x), & \text{se } x \in X_1 \\ j\big|_{X_2}(x), & \text{se } x \in X_2 \end{cases} \\
&= j(x) \text{ se } x \in X_1 \cup X_2
\end{aligned}
$$

Portanto $\phi \circ i = j$. ■

Assim, se X_1 e X_2 são disjuntos, G_1 é um grupo abeliano livre sobre os geradores $i(X_1)$ e G_2 é um grupo abeliano livre sobre os geradores $i_2(X_2)$ então $G_1 \oplus G_2$ é um grupo abeliano livre sobre os geradores $i(X_1 \cup X_2)$, onde i:

$$
\begin{aligned}
i: \quad X_1 \cup X_2 &\longrightarrow G_1 \cup G_2 \\
x &\longmapsto i(x) := \begin{cases} (i_1(x), 0_2) \text{ se } x \in X_1 \\ (0_1, i_2(x)) \text{ se } x \in X_2 \end{cases}
\end{aligned}
$$

Usando o mesmo argumento provaremos por indução a proposição para somas diretas finitas. Nesse processo, usaremos a Proposição 2.2 para obter que se X é um conjunto finito, então existirá (G, i) com a propriedade universal de X; além disso, pela Proposição 2.1 já temos sua unicidade. Vamos definir agora tal grupo G.

Tome X com n elementos, $x_1, x_2, \ldots, x_n$. Nós devemos formar uma soma direta de n grupos, cada um uma cópia do grupo $\mathbb{Z}$. Um elemento dessa soma direta é uma sequência de n inteiros $(r_1, r_2, \ldots, r_n)$, os quais podem associados a uma função g tal que

$$
\begin{aligned}
\mathbb{Z}^n \quad &\xrightarrow{g} \mathbb{F}(X, \mathbb{Z}) \\
(r_1, \ldots, r_n) \quad &\longmapsto \begin{aligned} f\colon X &\longrightarrow \mathbb{Z} \\ x_i &\longmapsto r_i \end{aligned}
\end{aligned}
$$

ou ainda nós podemos pensar nesses elementos como pontos do $\mathbb{R}^n$ com coordenadas inteiras, onde a estrutura de grupo abeliano é obtida pela adição de vetores com coordenadas inteiras no $\mathbb{R}^n$, com a soma sendo dada por

$$(r_1, \ldots, r_n) + (s_1, \ldots, s_n) = (r_1 + s_1, \ldots, r_n + s_n).$$

Observação 2.5. Se X tiver de 2 a $n-1$ elementos o par $(\mathbb{Z}^m, i)$, onde $m \in \{2, \ldots, n-1\}$ e

$$\begin{aligned} i: \quad X &\longrightarrow \mathbb{Z}^m \\ x_k &\longmapsto (\delta_{k1}, \ldots, \delta_{km}) \end{aligned}$$

possui a propriedade universal para X.

Demonstração. Esse é o segundo passo do processo de indução para a indução de somas diretas finitas da proposição 2.4, pois já provamos que a proposição é válida para $m = 1$ na proposição 2.2.

Primeiramente, observe que é utilizado a função delta de Kronecker, isto é

$$\delta_{ij} = \begin{cases} 0 \text{ se } i \neq j \\ 1 \text{ se } i = j \end{cases}$$

Suponha que se X possui de 2 a $n-1$ elementos, o par $(\mathbb{Z}^m, i)$ onde $m \in \{2, \ldots n-1\}$ possui a propriedade universal para X, provemos para $X = \{x_1, \ldots, x_n\} = \{x_1, \ldots, x_{n-1}\} \cup \{x_n\}$.

Pela hipótese de indução o par $(\mathbb{Z}^{n-1}, i_1)$ onde

$$\begin{aligned} i: \quad \{x_i, \ldots, x_{n-1}\} &\longrightarrow \mathbb{Z}^{n-1} \\ x_k &\longmapsto (\delta_{k1}, \ldots, \delta_{kn-1}) \end{aligned}$$

tem a propriedade universal para $\{x_1, \ldots, x_{n-1}\}$. Além disso, pelo primeiro passo de indução $(n = 1)$ temos que $(\mathbb{Z}, i_2)$ tem a propriedade universal para $\{x_n\}$ onde

$$\begin{aligned} i_2: \quad X &\longrightarrow \mathbb{Z} \\ x_n &\longmapsto 1 \end{aligned}$$

Assim, pela proposição 2.4, temos que $(\mathbb{Z}^{n-1} \oplus \mathbb{Z} = \mathbb{Z}^n, i)$ tem a propriedade universal para $\{x_1, \ldots, x_{n-1}\} \cup \{x_n\} = X$, onde

$$\begin{aligned} i: \quad X &\longrightarrow \mathbb{Z}^{n-1} \oplus \mathbb{Z} = \mathbb{Z}^n \\ x_j &\longmapsto \begin{cases} (i_1(x_j), 0), j = 1, \ldots, n-1 \\ (0, i_2(x_j)), j = n \end{cases} \end{aligned}$$

■

Vamos agora construir um grupo abeliano livre de posto arbitrário(infinito). Tome X um conjunto qualquer. Defina

$$F(X) = \{\varepsilon : X \longrightarrow \mathbb{Z} : \varepsilon(x) = 0 \text{ para todos exceto um número finito de elementos } x \in X\}.$$

Dado $\varepsilon, \eta \in F(X)$ nós definimos:

$$\left.\begin{aligned} (-\varepsilon)(x) &= -\varepsilon(x) \\ (\varepsilon + \eta)(x) &= \varepsilon(x) + \eta(x) \end{aligned}\right\} \forall x \in X \tag{2.1}$$

Observação 2.6. Com essa operação $F(X)$ é um grupo abeliano.

Demonstração. De fato, para quaisquer $\varepsilon, \eta, \alpha \in F(X)$, temos que;

1. Vale a associativa;

$$\begin{aligned} ((\varepsilon + \eta) + \alpha)(x) = (\varepsilon + \eta)(x) + \alpha(x) = \varepsilon(x) + \eta(x) + \alpha(x) = \\ \varepsilon(x) + (\eta + \alpha)(x) = (\varepsilon + (\eta + \alpha))(x) \end{aligned}$$

2. Existe elemento oposto;
   Para qualquer $\varepsilon \in F(X)$, basta tomar $-\varepsilon$;
$$(\varepsilon + (-\varepsilon))(x) = \varepsilon(x) + (-\varepsilon(x)) = \varepsilon(x) - \varepsilon(x) = 0.$$

3. Existe elemento neutro;
   Para isso, basta ver que $0 \in F(X)$, uma vez que 0 é uma função quase nula (a função nula tem o conjunto vazio como pontos diferentes de zero, o qual é finito).

4. Vale a comutatividade;
   Basta ver que
$$(\varepsilon + \eta)(x) = \varepsilon(x) + \eta(x) = \eta + \varepsilon(x) = (\eta + \varepsilon)(x). \qquad \blacksquare$$

Por fim, defina $i_X : X \longrightarrow F(X)$ por:
$$\begin{aligned} i_X : \quad X &\longrightarrow F(X) \\ x &\longmapsto i_X(x) \end{aligned}$$

onde
$$\begin{aligned} i_X(x) : \quad X &\longrightarrow \mathbb{Z} \\ y &\longmapsto \begin{cases} 1, y = x \\ 0, y \neq x \end{cases} \end{aligned}$$

Observação 2.7. $(F(Y), i_Y)$ tem a propriedade universal para Y.

Demonstração. Queremos provar que $(F(Y), i_Y)$ tem a propriedade universal para $Y = \{y_1, \ldots, y_n\}$. Já sabemos que $\{i_Y(y_1), \ldots, i_Y(y_k)\}$ é uma base para $F(Y)$. Considere o diagrama:

$$\begin{tikzcd} & Y = \{y_1, \ldots, y_n\} \arrow[dl, "i_Y"'] \arrow[dr, "j"] \arrow[d, phantom, "\circlearrowleft"] & \\ F(Y) \arrow[rr, "\phi"'] & & A \end{tikzcd}$$

onde A é um grupo abeliano qualquer, j é uma função e
$$\begin{aligned} i_Y : \quad Y &\longrightarrow F(Y) \\ y_j &\longmapsto i_Y(y_j) \end{aligned}$$

onde
$$\begin{aligned} i_Y(y_j) : \quad Y &\longrightarrow \mathbb{Z} \\ y_k &\longmapsto i_Y(y_j)(y_k) \end{aligned}$$

Devemos definir um homomorfismo $\phi : F(Y) \longrightarrow A$ tal que $\phi \circ i_Y = j$. Basta então definirmos ϕ nos elementos da base $\{i_Y(y_1), \ldots, i_Y(y_k)\}$ de $F(Y)$ e estender por linearidade, isto é, tomar ϕ tal que
$$\begin{aligned} \phi : \quad F(Y) &\longrightarrow A \\ i_Y(y_j) &\longmapsto j(y_j) \end{aligned}$$

Agora, estendamos por linearidade da seguinte forma. Seja $\varepsilon \in F(Y)$, qualquer. Como $\{i_Y(y_1), \ldots, i_Y(y_k)\}$ é uma base de $F(Y)$,

$$\varepsilon = r_1 i_Y(y_1) + \cdots + r_k i_Y(y_K)$$

então

$$\phi(\varepsilon) = \phi(r_1 i_Y(y_1) + \cdots + r_k i_k(y_k)) := r_1\phi(i_Y(y_1)) + \cdots + r_k\phi(i_Y(y_k)) = \\ r_1 j(y_1) + \cdots + r_k j(y_k)$$

Mostremos que ϕ é um homomorfismo de grupos. De fato, se $\varepsilon = t_1 i_Y(y_1) + \cdots + t_k i_Y(y_k)$ e $\eta = s_1 i_Y(y_1) + \cdots + s_k i_Y(y_k)$, então

$$\phi(\varepsilon + \eta) = \phi((t_1 + s_1)i_Y(y_1) + \cdots + (t_k + s_k)i_Y(y_k)) = (t_1 + s_1)j(y_1) + \cdots + (t_k + s_k)j(y_k).$$

Como A é abeliano

$$\phi(\varepsilon + \eta) = t_1 j(y_1) + \cdots + t_k j(y_k) + s_1 j(y_1) + \cdots + s_k j(y_k) = \phi(\varepsilon) + \phi(\eta).$$

Logo ϕ é um homomorfismo.

Note agora que $\phi \circ i_Y = j$. De fato, pela forma como definimos ϕ temos que $\forall k \in Y$

$$(\phi \circ i_Y)(k) = \phi(i_Y(k)) = j(k), \forall k \in Y.$$

Logo $\phi \circ i_Y = j$. Por fim, note que ϕ é único. Para isso, suponha que exista $\psi : F(Y) \longrightarrow A$ homomorfismo tal que $\psi \circ i_Y = j, \forall \varepsilon \in F(Y), \varepsilon = r_1 i_Y(y_1) + \cdots + r_k i_Y(y_k)$, temos

$$\psi(\varepsilon) = \psi(r_1 i_Y(y_1) + \cdots + r_k i_Y(y_k)) = r_1 \psi \circ i_Y(y_1) + \cdots + r_k \psi \circ i_Y(y_k) = \\ r_1 j(y_1) + \cdots + r_k j(y_k) = \phi(\varepsilon)$$

Logo $\psi = \phi$. Portanto $(F(Y), i_Y)$ tem a propriedade universal para Y. ■

Agradecimentos: Inicialmente agradeço a minha família e amigos que tornam a minha jornada nesse cosmos mais alegre e significante. Agradeço também ao Prof. Dr. João Peres Vieira, por ter me apresentado e orientado em um campo tão belo quanto o da Topologia.

Abstract: Abelian groups are one of the initial concepts seen in Abstract Algebra courses, very useful in the elaboration of concepts such as rings, vector spaces and fields. Some groups, called Free Abelian Groups, have interesting properties, in addition to inducing definitions in other areas, such as Algebraic Topology and Geometric Algebra. In this work there will be some groups with such a property and some results of the definition, such as the uniqueness of the property, the free Abelian group Z^m and the group of almost null functions.

Keywords: Free Abelian Group; Abelian Group; Algebraic Topology; Abstract Algebra

Referências Bibliográficas

[1] Wall, C. T. C. "A geometric Introduction to Topology", Addison-Wesley Publishing company, 1972.

[2] IEZZI, G. DOMINGUES, Hygino H. Álgebra moderna. 5 ed. São Paulo: editora Saraiva Uni, 2017.

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA · BICMat
Vol. 17 (2020) 100–106

Teoremas de Sylow

Thiago Moraes Rizzieri†
Orientador(a): Profa. Dra. Elíris Cristina Rizziolli

Resumo: Neste projeto é trabalhado aspectos pertinentes da Teoria de Grupos. Entre esses, destacamos os Teoremas de Sylow e suas demonstrações.

Palavras-chave: Teoria de Grupos; Ação de grupo; Teorema de Burnside; Teoremas de Sylow.

Inicialmente, abordamos a definição de ação de grupo no intuito de enunciar o Teorema de Burnside que será utilizado no desenvolvimento desse trabalho. Observamos que esse estudo foi baseado no livro [1], mais especificamente, foi baseado nos capítulos 3, seção 16 e no capítulo 7, seção 36.

1 Ação de grupo em um conjunto

Definição 1.1. Sejam X um conjunto e $(G, *)$ um grupo. Uma **ação de G em X** é uma aplicação, em que denotamos $*(g, x)$ por $g * x$ ou por gx, $* : G \times X \to X$ tal que

1. Para todo $x \in X : e * x = x$;
2. Para todo $x \in X$ e todo $g_1, g_2 \in G : (g_1 * g_2) * (x) = g_1 * (g_2 * x)$.

Nesse caso dizemos que X é um G-conjunto.

Teorema 1.2. *Seja X um G-conjunto. Para cada x_1,$x_2 \in X$, a relação $x_1 \sim x_2$ se, e somente se, existe $g \in G$ tal que $gx_1 = x_2$ é uma relação de equivalência.*

Demonstração. Para mostrar que $\sim$ é uma relação de equivalência, essa relção deve ser reflexiva, simétrica e transitiva.

- Para cada $x \in X$, nós temos $ex = x$, então $x \sim x$ e $\sim$ é reflexiva.
- Suponha que $x_1 \sim x_2$, assim $gx_1 = x_2$ para algum $g \in G$. Então $g^{-1}x_2 = g^{-1}(gx_1) = (g^{-1}g)x_1 = ex_1 = x_1$, portanto $x_2 \sim x_1$ e $\sim$ é simétrica.
- Suponha que $x_1 \sim x_2$ e $x_2 \sim x_3$, assim $g_1x_1 = x_2$ e $g_2x_2 = x_3$ para algum $g_1, g_2 \in G$. Então $(g_2g_1)x_1 = g_2(g_1x_1) = g_2x_2 = x_3$, portanto $x_1 \sim x_3$ e $\sim$ é transitiva. ■

Definição 1.3. Sejam X um G-conjunto e $x \in X$ qualquer. A classe de equivalência de x segundo a relação de equivalência descrita no Teorema anterior é chamada **órbita de x em X sob G** e denotada por $G * x$. Ou seja

$$G * x = \{x_0 \mid x_0 = gx,\ g \in G\}.$$

†Bolsista PET/FNDE

Observação 1.4. No que segue, usamos as seguintes denominações:

$$X_g := \{x \in X \mid gx = x\} \subset X;$$

e

$$G_x := \{g \in G \mid gx = x\} \subset G.$$

Pode-se mostrar que G_x é subgrupo de G para todo $x \in X$.

Teorema 1.5. *Se X é um G-conjunto e $x \in X$, então a ordem da órbita $G * x$ é igual ao índice do subgrupo $(G : Gx)$, ou seja, $|G * x| = (G : G_x)$. Além disso se G é finito então $|G * x|$ divide $|G|$.*

Demonstração. Para demonstrar a primeira afirmação, vamos definir uma aplicação bijetiva ψ de $G * x$ até a coleção de classes laterais à esquerda de G_x em G, ou seja, até o conjunto $L = \{gG_x | g \in G\}$. Se $x_1 \in G * x$, então existe $g_1 \in G$ tal que $g_1 x = x_1$. Assim

$$\begin{aligned}\psi : G * x &\to L \\ x_1 &\mapsto \psi(x_1) = \psi(g_1 x) := g_1 G_x.\end{aligned}$$

Precisamos mostrar que a aplicação está bem definida e é bijetiva.

- ψ está bem definida. De fato, suponha que $x_1 = x_2$, $\forall x_1, x_2 \in G * x$. Assim, existem $g_1, g_2 \in G$ tais que $x_1 = g_1 x$ e $x_2 = g_2 x$. Como $x_1 = x_2$ então $g_1 x = g_2 x$.

  Logo, $g_1 x = g_2 x \Rightarrow g_1^{-1}(g_1 x) = g_1^{-1}(g_2 x) \Rightarrow (g_1^{-1} g_1)x = (g_1^{-1} g_2)x \Rightarrow ex = (g_1^{-1} g_2)x \Rightarrow x = (g_1^{-1} g_2)x \Rightarrow (g_1^{-1} g_2) \in G_x$.

  Veja que se $(g_1^{-1} g_2) \in G_x$ então existe $h \in G_x$ tal que $(g_1^{-1} g_2) = h$ que por consequência $(g_1^{-1} g_2) = h \Rightarrow g_2 = g_1 h \Rightarrow g_2 \in g_1 G_x \Rightarrow g_2 G_x = g_1 G_x \Rightarrow \psi(g_2 x) = \phi(g_1 x) \Rightarrow \psi(x_2) = \psi(x_1)$.

- ψ é injetiva. Pois, suponha $x_1, x_2 \in G * x$, tal que $\psi(x_1) = \psi(x_2)$. Como $x_1, x_2 \in G * x$ então existem $g_1, g_2 \in G$ de modo que $x_1 = g_1 x$ e $x_2 = g_2 x$, portanto, $\psi(g_1 x) = \psi(g_2 x) \Rightarrow g_1 G_x = g_2 G_x \Rightarrow g_2 \in g_1 G_x$, assim deve existir $h \in G_x$ tal que $g_2 = g_1 h$.

  Então $x_2 = g_2 x = (g_1 h)x = g_1(hx)$, como $h \in G_x$, então $g_1(hx) = g_1 x = x_1$.

- ψ é sobrejetiva. De fato, seja $g_1 G_x \in L$. Como X é um G-conjunto, deve existir $x_1 \in G * x$ tal que $g_1 x = x_1$. Assim, como ψ é bem definida, $g_1 G_x = \psi(x_1)$. Portanto, como $g_1 G_x$ é um elemento qualquer de L, então podemos dizer que

  $$L = \{g_1 G_x \mid g \in G\} = \{\psi(x_1) \mid x_1 \in G * x\} = \psi[G * x].$$

Por fim, temos então ψ uma aplicação bijetora em que $|G * x| = |L|$. Veja que $|L|$ representa a quantia de classes laterais de G_x em G, logo $|G * x| = |L| = (G : G_x)$.

Se G é finito então, segue do Teorema de Lagrange que $(G : G_x) = |G|/|G_x|$. Assim, $|G * x| = (G : G_x) = |G|/|G_x| \Rightarrow |G| = |G * x||G_x|$ então $|G * x|$ divide $|G|$. ■

Teorema 1.6 (Fórmula de Burnside). *Sejam G um grupo finito e X um G-conjunto finito. Se r é o número de órbitas distintas em X sobre G, então*

$$r \cdot |G| = \sum_{g \in G} |X_g|.$$

Demonstração. Inicialmente, defina o seguinte conjunto

$$N = \{\text{quantidade de pares } (g, x) \in G \times X \mid gx = x\}.$$

Assim, para cada $g_i \in G$ existem $|X_{g_i}|$ pares tendo g_i como primeiro membro. Portanto,

$$N = \sum_{g \in G} |X_g|. \tag{1.1}$$

Por outro lado, para cada $x_j \in X$ existem $|G_{x_j}|$ pares tendo x_j como segundo membro. Portanto,

$$N = \sum_{x \in X} |G_x|.$$

Mas ja vimos pelo Teorema 1.5 que $(G : G_x) = |G * x|$ e $(G : G_x) = |G|/|G_x|$. Desse modo, temos $|G_x| = |G|/|G * x|$. Assim,

$$N = \sum_{x \in X} \frac{|G|}{|G * x|} = |G| \cdot \left(\sum_{x \in X} \frac{1}{|G * x|} \right).$$

Veja que $|G * x|$ possui o mesmo valor para todos os x de uma mesma órbita. Seja $\mathcal{O}$ uma órbita qualquer de n elementos. Observe que

$$\sum_{x \in \mathcal{O}} \frac{1}{|G * x|} = \sum_{x \in \mathcal{O}} \frac{1}{|\mathcal{O}|} = \sum_{x \in \mathcal{O}} \frac{1}{n} = n \frac{1}{n} = 1$$

Assim, sendo r o número de órbitas

$$N = |G| \cdot r. \tag{1.2}$$

Juntando as equações 1.1 e 1.2 temos

$$r \cdot |G| = \sum_{g \in G} |X_g|. \qquad \blacksquare$$

Os resultados dessa seção, por definir e explorar as ações de grupo, são aplicados na compreensão dos p-grupos e nas demonstrações dos Teoremas de Sylow.

2 Teoremas de Sylow

Seja X um finito G-grupo, suponha que existam r órbitas em X sob G e que o conjunto $\{x_1, x_2, \ldots, x_r\}$ contenha um elemento para cada órbita em X. Assim, cada elemento de X está em apenas uma órbita, desta forma, a quantia de elementos do conjunto X equivale à soma da quantia de elementos de cada órbita em X, ou seja

$$|X| = \sum_{i=1}^{r} |G * x_i|.$$

Veja que podem existir órbitas unitárias em X, dessa forma, se denotamos $X_G := \{x \in X | gx = x, \text{ para todo } g \in G\}$. Segue que, X_G é exatamente a união de todas as órbitas unitárias. Digamos que que existam s órbitas unitárias, em que $0 \leq s \leq r$, dessa forma $|X_G| = s$ e

$$|X| = |X_G| + \sum_{i=s+1}^{r} |G * x_i|. \tag{2.1}$$

Essa última equação é imprescindível para a compreensão do próximo resultado.

Teorema 2.1. *Se p é um número primo, G um grupo de ordem p^n e X um G-conjunto, então $|X| \equiv |X_G| \pmod{p}$.*

Demonstração. Pelo Teorema 1.5, sabemos que, dado $x_i \in X$, $|G * x_i|$ divide $|G|$. Consequentemente, p divide $|G * x_i|$ para $s + 1 \leq i \leq r$. Veja que da equação (2.1), temos

$$|X| - |X_G| = \sum_{i=s+1}^{r} |G * x_i|.$$

Portanto, p divide $|X| - |X_G|$, ou seja $|X| \equiv |X_G| \pmod{p}$. ■

Definição 2.2. Seja p um número primo.

- Um grupo é um **p-grupo** se cada elemento de G possui ordem de uma potência de p.
- Um subgrupo de G é um **p-subgrupo de** G se o subgrupo for ele mesmo um p-grupo.

Teorema 2.3 (Teorema de Cauchy)**.** *Se p é um número primo e G é um grupo finito tal que p divide $|G|$. Então G possui um elemento de ordem p e, consequentemente, um subgrupo de ordem p.*

Demonstração. Considere X o conjunto de todas as p-uplas $(g_1, g_2, \ldots, g_p)$ dos elementos de G tal que o produto de suas coordenadas seja o elemento neutro e. Ou seja,

$$X = \{(g_1, g_2, \ldots, g_p) \mid g_i \in G \text{ e } g_1 g_2 \ldots g_p = e\}.$$

Isso é possível pois dado $g_p \in G$ deve existir seu elemento inverso $g_p^{-1} = g_1 g_2 \cdots g_{p-1}$ tal que $g_p g_p^{-1} = e$. Ao formar a p-upla $(g_1, g_2, \ldots, g_p)$, observemos que $g_1 g_2 \ldots g_p = e \Rightarrow g_p = (g_1 g_2 \cdots g_{p-1})^{-1}$. Assim, podemos ver $g_1, g_2, \ldots, g_{p-1}$ como elementos quaisquer de G, e g_p unicamente determinado por $(g_1 g_2 \cdots g_{p-1})^{-1}$. Logo, $|X| = |G|^{p-1}$ e como p divide $|G|$ então p divide $|X|$.

Seja σ o ciclo $(1, 2, 3, \ldots, p)$ do grupo de permutações S_p. Considere que σ aja em X da seguinte forma

$$\sigma(g_1, g_2, \ldots, g_p) = (g_{\sigma(1)}, g_{\sigma(2)}, \cdots, g_{\sigma(p)} = (g_2, g_3, \ldots, g_p, g_1).$$

Perceba que, como $(g_1, g_2, \ldots, g_p) \in X$ então $g_1(g_2 \ldots g_p) = e$ que implica que $g_1 = (g_2 \ldots g_p)^{-1}$, e então $(g_2 \ldots g_p)g_1 = e$.

Com isso, temos que $\sigma(g_1, g_2, \ldots, g_p) = (g_2, g_3, \ldots, g_p, g_1) \in X$. Assim σ age em X, então consideramos o subgrupo $\langle\sigma\rangle$ de S_p agindo em X por iteração, ou seja, por composições de σ.

Veja que $|\langle\sigma\rangle| = p$, então pelo Teorema 2.1 sabemos que $|X| \equiv |X_{\langle\sigma\rangle}| (\text{mod } p)$. Como p divide $|X|$, logo p divide $|X_{\langle\sigma\rangle}|$ também. Vamos examinar $|X_{\langle\sigma\rangle}|$. Lembremos que $X_{\langle\sigma\rangle} = \{(g_1, g_2, \ldots, g_p) \in X | \sigma(g_1, g_2, \ldots, g_p) = (g_1, g_2, \ldots, g_p)$, ou seja,$(g_1, g_2, \ldots, g_p)$ pertence à $|X_{\langle\sigma\rangle}|$ apenas quando $g_1 = g_2 = \ldots = g_p$ e $g_1 g_2 \ldots g_p = e$. Sabemos que existe pelo menos um elemento que satisfaz essas condições,a saber, o elemento $(e, e, \ldots, e)$. Porém, como p divide $|X_{\langle\sigma\rangle}|$, então devem existir pelo menos p elementos em $X_{\langle\sigma\rangle}$.

Assim, deve existir um elemento $a \in G$, $a \neq e$ tal que $(a, a, \ldots, a) \in X_{\langle\sigma\rangle}$ e portanto, $a^p = e$, indicando que a ordem de a é p. Consequentemente, $\langle a\rangle$ é um subgrupo de G de ordem p. ■

Corolário 2.4. *Seja G um grupo finito. Então G é um p-grupo se e somente se $|G|$ é uma potência de p.*

Demonstração. ($\Rightarrow$) Suponha que G seja um p-grupo. Seja r um número primo que divide $|G|$. Pelo Teorema de Cauchy (Teorema 2.3), existe um elemento $g \in G$ de ordem r. Porém, pela definição de p-grupo, todo elemento de G possui ordem p^k, $k \in \mathbb{N}$. Assim r deve ser igual à p^k, porém como r é primo, isso só é verdade se $k = 1$ e então $p = r$. Isso significa que $|G|$ não é divisível por nenhum outro número primo além de p. Portanto, $|G|$ é uma potência de p.

($\Leftarrow$) Suponha que $|G| = p^m$ para algum $m \in \mathbb{N}$. Seja $a \in G$ qualquer. Como G é um grupo finito, então pelo Teorema de Lagrange $|G| = (G : \langle a \rangle)|\langle a \rangle|$, então $p^m = (G : \langle a \rangle)|\langle a \rangle|$. Isso implica que $|\langle a \rangle|$ divide $|G|$, e portanto, $|\langle a \rangle| = p^l$ para algum $l \leq m$. ■

Observação 2.5. Sejam G um grupo e $\mathscr{S}$ a coleção de todos os subgrupos de G. Transformamos $\mathscr{S}$ em um G-conjunto por uma ação de G em $\mathscr{S}$ por conjugação. Isto é, se $H \in \mathscr{S}$ então H é subgrupo de G e, dado $g \in G$, a ação de g em H produz o subgrupo conjugado gHg^{-1}. Seja $N[H] = \{g \in G \mid gHg^{-1} = H\}$. Veja que NH é subgrupo de G, H é normal à $N[H]$ e mais ainda, $N[H]$ é o maior subgrupo de G que possui H como subgrupo normal. O subgrupo $N[H]$ é chamado **normalizador de H em G**.

Lema 2.6. *Seja H um p-subgrupo de um grupo finito G, então*

$$(N[H] : H) \equiv (G : H) \pmod{p}.$$

Demonstração. Páginas 323 e 324 do livro [1]. ■

Corolário 2.7. *Seja H um p-subgrupo de um grupo finito G. Se p divide $(G : H)$, então $N[H] \neq H$.*

Demonstração. Segue do Lema 2.6 que se p divide $(G : H)$, então p divide $(N[H] : H)$. Assim, $(N[H] : H)$ deve ser diferente de 1, logo, $|N[H]|$ deve ser diferente de $|H|$ e por isso $N[H] \neq H$. ■

Teorema 2.8 (Primeiro Teorema de Sylow). *Seja G um grupo finito tal que $|G| = p^n m$, em que $n \geq 1$ e p não divide m. Nessas condições,*

1. *G contém um subgrupo de ordem p^i, para cada i em que $1 \leq i \leq n$;*

2. *Todo subgrupo H de G de ordem p^i é um subgrupo normal a algum subgrupo de ordem p^{i+1} para $1 \leq i < n$.*

Demonstração. Provamos cada caso separadamente.

(1) Sabemos que G contém um subgrupo de ordem p pelo Teorema de Cauchy (Teorema 2.3), de fato pois p divide $|G|$.

Assim, existe um subgrupo de G de ordem p^i para $i = 1$.

Vamos usar o princípio de indução e mostrar que a existência de um subgrupo de ordem p^i, para $i < n$, implica que existe um subgrupo de ordem p^{i+1}.

Seja H um subgrupo de ordem p^i, com $i < n$. Como $i < n$, pelo Teorema de Lagrange $|G| = (G : H)|H|$, ou seja, $p^n m = (G : H)p^i$. Segue que p divide $(G : H)$. Pelo Lema 2.6, como p divide $(G : H)$, então p divide $(N[H] : H)$. Como H é normal à $N[H]$, então podemos formar o grupo quociente $N[H]/H$, e vemos que p divide $|N[H]/H|$. Pelo Teorema de Cauchy (Teorema 2.3), $N[H]/H$ possui um subgrupo K de ordem p.

Se $\gamma : N[H] \to N[H]/H$ é homomorfismo canônico, então $\gamma^{-1}[K] = \{x \in N[H] \mid \gamma(x) \in K\}$ é um subgrupo de $N[H]$ e de G também. Esse subgrupo contém H e possui ordem p^{i+1}.

De fato, pois dado $x+H \in K = \{H, a_1+H, a_2+H, \ldots, a_p+H\}$. Temos $x+H = a_j+H$ para algum j tal que $1 \leq j \leq p$, e por consequência, $x - a_j \in H$.

Assim, $x - a_j \in H = \{h_1, h_2, \ldots, h_{p^i}\}$. Temos $x - a_j = h_l$ e então $x = a_j + h_l$ para todo j e l tal que $1 \leq j \leq p$ e $1 \leq l \leq p^i$.

(2) Vamos repetir a construção feita no item 1. Perceba que $H < \gamma^{-1}[K] \leq N[H]$ onde $|\gamma^{-1}[K]| = p^{i+1}$ (denotamos subgrupo pelo símbolo $\leq$). Como H é normal à $N[H]$, ele é normal à $\gamma^{-1}[K]$. ■

Definição 2.9. Dizemos que um subgrupo P de um grupo G é um **p-subgrupo de Sylow** se P é o p-subgrupo maximal de G, isto é, o p-subgrupo que não está contido em nenhum outro p-subgrupo.

O Segundo Teorema de Sylow nos garante que todo p-subgrupo de Sylow pode ser obtido a partir de P, uma vez que, se P é um p-subgrupo de Sylow, então todo o conjugado gPg^{-1} de P também é um p-subgrupo de Sylow.

Teorema 2.10 (Segundo Teorema de Sylow)**.** *Se P_1 e P_2 p-subgrupos de Sylow de um grupo finito G. Então P_1 e P_2 são subgrupos conjugados de G.*

Demonstração. Aqui, iremos fazer um dos subgrupos agir nas classes laterais à esquerda do outro, e usar o Teorema 2.1. Seja $\mathcal{L} = \{xP_1 \mid x \in G\}$ a coleção de classes laterais à esquerda de P_1 em G. O subgrupo P_2 age em $\mathcal{L}$ pela ação $y(xP_1) = (yx)P_1$, para todo $y \in P_2$.

Então $\mathcal{L}$ é um P_2-conjunto. Pelo Teorema 2.1, $|\mathcal{L}_{P_2}| \equiv |\mathcal{L}| \pmod{p}$, e por outro lado $|\mathcal{L}| = (G : P_1)$ não é divisível por p, pois supondo que $|G| = p^n m$ sendo que p não divide m, temos que $|P_1| = p^n$, e assim $|\mathcal{L}| = (G : P_1) = m$. Portanto $|\mathcal{L}_{P_2}| \neq 0$.

Lembremos que $\mathcal{L}_{P_2} = \{xP_1 \in \mathcal{L} | y(xP_1) = xP_1 \text{ para todo } y \in P_2\}$. Seja $xP_1 \in \mathcal{L}_{P_2}$. Então $y(xP_1) = xP_1$ para todo $y \in P_2$ e assim, $x^{-1}yxP_1 = P_1$ para todo $y \in P_2$. Logo, $x^{-1}yx \in P_1$ para todo $y \in P_2$, e então $x^{-1}P_2x \subset P_1$. Como $|P_1| = |P_2| = p^n$, consequentemente $P_1 = x^{-1}P_2x$ e, portanto, P_1 e P_2 são subgrupos conjugados. ■

O Terceiro Teorema de Sylow nos garante a quantia dos p-subgrupos de Sylow em um grupo.

Teorema 2.11 (Terceiro Teorema de Sylow)**.** *Se G é um grupo finito e p divide $|G|$ então a quantia de p-subgrupos de Sylow é congruente a 1 módulo p e divide $|G|$.*

Demonstração. Seja P um p-subgrupo de Sylow de G. Considere $\mathscr{S}$ o conjunto de todos os p-subgrupos de Sylow e que P aja em $\mathscr{S}$ por conjugação, então $x \in P$ leva $T \in \mathscr{S}$ em $x^{-1}Tx$. Pelo Teorema 2.1, $|\mathscr{S}| \equiv |\mathscr{S}_P| \pmod{p}$.

Vamos determinar $\mathscr{S}_P$. Se $T \in \mathscr{S}_P$, então $x^{-1}Tx = T$ para todo $x \in P$. Assim, $x \in N[T]$ e por consequência, $P \leq N[T]$. Vimos anteriormente que $H \leq N[H]$ para todo subgrupo H, logo $T \leq N[T]$ também.

Como P e T são p-subgrupos de Sylow de G, então são p-subgrupos de Sylow de $N[T]$. Mas, pelo Segundo Teorema de Sylow (Teorema 2.10), P e T são conjugados em $N[T]$. Porém, como T é normal à $N[T]$, ele é seu próprio e único conjugado.

Assim, $P = T$. Consequentemente, $\mathscr{S}_P = \{P\}$, logo, $|\mathscr{S}| \equiv 1 \pmod{p}$, ou seja, a quantia de p-subgrupos de Sylow de G é congruente a 1 módulo p.

Assuma que G aja em $\mathscr{S}$ por conjugação, então $g \in G$ leva $T \in \mathscr{S}$ em $g^{-1}Tg$. Como todos os p-subgrupos de Sylow são conjugados, então existe uma única órbita de $\mathscr{S}$ sob G, pois para qualquer $T \in \mathscr{S}$, $T = g^{-1}Pg$ que pertence à órbita $G{*}P$. Assim $|G{*}P| = (G : G_P)$

pelo Teorema 1.5. Mas veja que pelo Teorema de Lagrange $(G : G_P)$ é divisor de $|G|$ e que $(G : G_P) = |\mathscr{S}|$, que é justamente a quantia de p-subgrupos de Sylow de G. Ou seja a quantia de p-subgrupos de Sylow de G divide $|G|$. ■

Agradecimentos: Gostaria de agradecer à minha orientadora Dra. Elíris Cristina Rizziolli pelo apoio, incentivo e orientação durante toda a realização desse projeto. Também agradeço à minha família e amigos pelo apoio diário.

Abstract: This project deals with relevant aspects of Group Theory. Among these, we highlight the Sylow Theorems and their proofs.

Keywords: Group theory; Group action; Burnside Theorem; Sylow Theorems.

Referências Bibliográficas

[1] Fraleigh, John B. A First Course in Abstract Algebra. 7ª edição. Pearson, 2014.

ISSN 1980–024X

Boletim de Iniciação Científica em Matemática · BICMat

Orientação aos autores

Ao redigir o material a ser divulgado o autor deve observar que o alvo principal é o aluno de graduação, devendo a redação ser clara e objetiva incentivando-o à leitura.

O trabalho deve ser enviado à Comissão Editorial, via e-mail, na linguagem LaTeX, usando a classe `bicmat`. Mais informações sobre a formatação do trabalho podem ser encontradas em `www.rc.unesp.br/igce/matematica/bicmat`, assim como o endereço para o envio do trabalho.

A responsabilidade de cada artigo é exclusiva do autor e respectivo orientador.