UNIVERSIDADE FEDERAL DE SÃO CARLOS

DEPARTAMENTO DE FÍSICA, QUÍMICA E MATEMÁTICA

KLEBER DE SANTANA SOUZA

O TEOREMA DO PONTO FIXO DE BANACH E APLICAÇÕES NO ESTUDO DA CONVERGÊNCIA DOS MÉTODOS ITERATIVOS

Sorocaba

2015

UNIVERSIDADE FEDERAL DE SÃO CARLOS

DEPARTAMENTO DE FÍSICA, QUÍMICA E MATEMÁTICA

O TEOREMA DO PONTO FIXO DE BANACH E APLICAÇÕES NO ESTUDO DA CONVERGÊNCIA DOS MÉTODOS ITERATIVOS

Kleber de Santana Souza

Orientador: Prof. Dr. Wladimir Seixas

Sorocaba

2015

O TEOREMA DO PONTO FIXO DE BANACH E APLICAÇÕES NO ESTUDO DA CONVERGÊNCIA DOS MÉTODOS ITERATIVOS

Monografia apresentada ao Curso de Licenciatura em Matemática da Universidade Federal de São Carlos, como exigência parcial para a obtenção do título de Licenciado em Matemática sob orientação do Professor Doutor Wladimir Seixas.

Sorocaba

2015

<u>**Folha de aprovação**</u>

Kleber de Santana Souza

O Teorema do Ponto Fixo de Banach e Aplicações no Estudo da Convergência dos Métodos Iterativos

Trabalho de Conclusão de Curso

Universidade Federal de São Carlos – *Campus* Sorocaba

Sorocaba, 11/12/2015.

Orientador ______________________
Prof. Dr. Wladimir Seixas

Membro 2 ______________________
Prof.ª Dr.ª Marta Cilene Gadotti

Membro 3 ______________________
Prof. Dr. Adilson José Vieira Brandão

Dedico este trabalho a minha mãe Lucilda de Santana Souza e ao meu pai Wilson Batista dos Santos Souza.

AGRADECIMENTOS

Gostaria de agradecer a todos os professores com os quais tive aula durante a graduação, mas principalmente ao Prof. Dr. Paulo Cesar Oliveira, a Profª. Dra. Silvia Maria Simões de Carvalho e ao Prof. Dr. Adilson Brandão que foram em grande parte responsáveis pelo novo mundo intelectual que pude vislumbrar. Tive o prazer e a oportunidade de aprender com estas pessoas que dedicaram parte do seu tempo para minha formação. E em especial ao meu orientador Prof. Dr. Wladimir Seixas, por ter aceitado me orientar neste trabalho e nas iniciações científicas que fizemos. Gostaria de agradecer pelo incentivo, paciência, confiança e pelas orientações prestadas.

A minha esposa Lidia Aparecida França Souza e a minha filha Sofia França Souza pela compreensão e apoio durante as horas de estudo que dediquei a este trabalho.

Aprender é a única coisa de que a mente nunca se cansa, nunca tem medo e nunca se arrepende.

Leonardo Da Vinci

RESUMO

Este trabalho tem como objetivo realizar o estudo do Teorema do Ponto Fixo de Banach. São apresentadas várias definições e teoremas sobre espaços métricos, espaços normados e espaços de Banach. Todos eles pré-requisitos importantes para o estudo e entendimento deste teorema. São apresentadas também algumas aplicações à teoria da Análise Numérica onde o Teorema do Ponto Fixo de Banach se faz presente.

Palavras-chaves: Análise Funcional. Espaços de Banach. Teorema do ponto fixo de Banach. Análise Numérica.

ABSTRACT

This work aims to conduct the study of Banach Fixed Point Theorem. Several definitions and theorems on metric spaces, normed spaces and Banach spaces are presented. They are important prerequisites for the study and understanding of the Banach Fixed Point Theorem. Also we discuss some applications of this important theorem to Numerical Analysis.

Key-words: Functional Analysis. Banach Spaces. Banach Fixed Point Theorem. Numerical Analysis.

SUMÁRIO

INTRODUÇÃO

Análise Funcional é um ramo abstrato da Matemática que se originou da Análise Clássica. Seu desenvolvimento teve início aproximadamente no começo do século XX, e atualmente os métodos analíticos funcionais e seus resultados são importantes em vários campos da Matemática e suas aplicações. O ímpeto veio da Álgebra Linear, das Equações Diferenciais lineares ordinárias e parciais, do Cálculo de Variações, da Teoria da Aproximação e em particular das equações integrais lineares, cuja teoria teve grande impacto no desenvolvimento e promoção das ideias modernas.

Este trabalho tem por objetivo principal estudar o Teorema do Ponto Fixo de Banach, um importante teorema da Análise Funcional. Para tanto foi necessário a realização de um estudo introdutório de alguns conceitos da Análise Funcional. Neste estudo nos deparamos com conteúdos que não estão presentes na grade de nosso curso de Licenciatura em Matemática. Realizamos neste trabalho o estudo de várias definições e teoremas de espaços métricos, de espaços normados e espaços de Banach. E verificamos algumas aplicações onde os resultados são consequência direta da aplicação destes conteúdos. Como os de alguns métodos iterativos presentes no Cálculo Numérico.

Este trabalho está dividido nos seguintes capítulos:

Capítulo 1: Neste capítulo enunciamos e provamos as desigualdades de Cauchy-Schwarz, de Hölder e de Minkowski. Apresentamos algumas definições sobre espaços métricos, assim como exemplos de espaços métricos. Definimos conjunto aberto e conjunto fechado, topologia e continuidade. Definimos também alguns espaços métricos completos e teoremas que serão necessários aos estudos que pretendemos realizar.

Capítulo 2: São definidos neste capítulo os espaços normados e os espaços de Banach. Retomamos os exemplos do capítulo anterior para mostrar que são espaços normados. Definimos o que são operadores lineares e funcionais lineares e apresentamos os teoremas e exemplos que foram estudados.

Capítulo 3: Neste capítulo é apresentado o Teorema do Ponto Fixo de Banach e sua aplicação na demonstração de teoremas de existência e unicidade de equações algébricas, da equação integral não linear de Volterra, de equações diferenciais ordinárias em espaços de Banach e do método de Newton em espaços de Banach.

1 ESPAÇOS MÉTRICOS

Os espaços métricos são fundamentais em Análise Funcional porque desempenham um papel similar ao da reta real $\mathbb{R}$ no Cálculo. De fato, eles generalizam $\mathbb{R}$ e foram criados de modo a proporcionar as bases para um tratamento unificado de importantes problemas de vários ramos da Análise. Um espaço métrico é um conjunto X com uma métrica. A métrica associa a qualquer par de pontos (elementos) de X uma distância. Uma propriedade adicional que um espaço métrico pode ter é a completude. Outro conceito teórico e de interesse prático é o de separabilidade.

Como algumas desigualdades são fundamentais nas demonstrações de teoremas que iremos apresentar, começaremos enunciando e provando estas desigualdades.

Os teoremas e definições assim como as demonstrações que se seguem foram retirados principalmente das referências [3], [8] e [2].

Definição 1.1. *Seja I um intervalo não vazio de $\mathbb{R}$. A função $f : I \longrightarrow \mathbb{R}$ é dita convexa em I quando*

$$f(\alpha x_1 + (1-\alpha)x_0) \leqslant \alpha f(x_1) + (1-\alpha)f(x_0) \tag{1.1}$$

para todos os pares de pontos (x_0, x_1) em I e todo $\alpha \in (0,1)$.

Teorema 1.1. *(Desigualdade de Jensen) Sejam I um intervalo não vazio de $\mathbb{R}$ e f uma função convexa em I. Então, para qualquer conjunto $\{x_1, \dots, x_k\}$ de pontos em I e qualquer conjunto de números $\{\alpha_1, \dots, \alpha_k\}$ satisfazendo*

$$\alpha_i \geqslant 0 \textit{ para } i = 1, \dots, k \textit{ e } \sum_{i=1}^{k} \alpha_i = 1,$$

vale a desigualdade

$$f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leqslant \sum_{i=1}^{k} \alpha_i f(x_i).$$

Demonstração. Para $k = 2$ a relação é trivial se ou α_1 é zero ou α_2 é zero; caso contrário, temos a desigualdade (1.1) da definição anterior.

Supondo por indução que a relação é verdadeira para $k-1$; tomemos os conjuntos $\{x_i\}$ e $\{\alpha_i\}$. Se α_k é 0 ou 1, não há o que provar. Caso contrário, definindo

$$\overline{\alpha} := \sum_{i=1}^{k-1} \alpha_i \in (0,1) \text{ temos } \alpha_k = 1 - \overline{\alpha} \in (0,1)$$

$$\overline{\alpha}_i = \frac{\alpha_i}{\overline{\alpha}} \text{ para } i = 1, \dots, k-1 \text{ temos } \overline{\alpha}_i \geqslant 0 \text{ e } \sum_{i=1}^{k-1} \overline{\alpha}_i = 1.$$

Logo

$$\sum_{i=1}^{k} \alpha_i x_i = \sum_{i=1}^{k-1} \alpha_i x_i + \alpha_k x_k = \overline{\alpha} \sum_{i=1}^{k-1} \overline{\alpha}_i x_i + (1-\overline{\alpha}) x_k.$$

Nesta última relação, o ponto $\overline{x} := \sum_{i=1}^{k-1} \alpha_i x_i$ está em I, entre $\min_i x_i$ e $\max_i x_i$. Aplicando (1.1)

$$f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leqslant \overline{\alpha} f(\overline{x}) + (1-\overline{\alpha}) f(x_k) = \overline{\alpha} f(\overline{x}) + \alpha_k f(x_k).$$

O resultado segue da hipótese de indução aplicada a $\overline{x}$:

$$\overline{\alpha} f(\overline{x}) \leqslant \overline{\alpha} \sum_{i=1}^{k-1} \overline{\alpha}_i f(x_i) = \sum_{i=1}^{k-1} \alpha_i f(x_i).$$

$\square$

Teorema 1.2. *(Desigualdade de Young) Sejam $a, b > 0$ e $p, q > 1$ números reais tais que $\frac{1}{p} + \frac{1}{q} = 1$. Então*

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}.$$

A igualdade ocorre se, e somente se, $a^p = b^q$.

Demonstração. Para $f(x) = e^x$ temos $f'(x) = f''(x) = e^x > 0$, para todo $x \in \mathbb{R}$. Logo $f(x)$ é convexa em $(0, \infty)$. Tomando $x = p \ln a$ e $y = q \ln b$, e usando a desigualdade de Jensen obtemos

$$\begin{aligned}
f\left(\frac{1}{p}x + \frac{1}{q}y\right) &\leqslant \frac{1}{p} f(x) + \frac{1}{q} f(y) \\
e^{\frac{1}{p}x + \frac{1}{q}y} &\leqslant \frac{e^x}{p} + \frac{e^y}{q} \\
e^{\ln a + \ln b} &\leqslant \frac{e^{p \ln a}}{p} + \frac{e^{q \ln b}}{q} \\
e^{\ln ab} &\leqslant \frac{e^{\ln a^p}}{p} + \frac{e^{\ln b^q}}{q} \\
ab &\leqslant \frac{a^p}{p} + \frac{b^q}{q}.
\end{aligned}$$

$\square$

Teorema 1.3. *(Desigualdade de Cauchy-Schwarz) Sejam $a_1, a_2, \ldots, a_n$ e $b_1, b_2, \ldots, b_n$ números reais. Então temos*

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leqslant \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right).$$

Demonstração. Considere o seguinte trinômio quadrático

$$\sum_{i=1}^{n}(a_i x - b_i)^2 = \sum_{i=1}^{n}(a_i^2 x^2 - 2a_i b_i x + b_i^2) = x^2 \sum_{i=1}^{n} a_i^2 - 2x \sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} b_i^2.$$

Este trinômio é não negativo para todo $x \in \mathbb{R}$, portanto seu discriminante é não positivo, isto é

$$4\left(\sum_{i=1}^{n} a_i b_i\right)^2 - 4\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) \leqslant 0 \Leftrightarrow \left(\sum_{i=1}^{n} a_i b_i\right)^2 \leqslant \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right).$$

como requerido.

A igualdade vale se, e somente se, $a_i x = b_i, \forall i.$ □

Teorema 1.4. *(Desigualdade de Hölder) Sejam $a_1, a_2, \ldots, a_n$; $b_1, b_2, \ldots, b_n$ números reais e $p, q > 1$ tais que $\frac{1}{p} + \frac{1}{q} = 1$. Então,*

$$\sum_{i=1}^{n} |a_i b_i| \leqslant \left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |b_i^q|\right)^{\frac{1}{q}}.$$

Demonstração. Considerando

$$a = \frac{|a_i|}{\left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}}},\ b = \frac{|b_i|}{\left(\sum_{i=1}^{n} |b_i^q|\right)^{\frac{1}{q}}},\ i = 1, 2, \ldots, n.$$

Da desigualdade de Young obtemos

$$\frac{|a_i b_i|}{\left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |b_i^q|\right)^{\frac{1}{q}}} \leqslant \frac{1}{p}\frac{|a_i^p|}{\sum_{i=1}^{n} |a_i^p|} + \frac{1}{q}\frac{|b_i^q|}{\sum_{i=1}^{n} |b_i^q|}.$$

Somando as desigualdades para $i = 1, 2, \ldots, n$ obtemos

$$\frac{\sum_{i=1}^{n} |a_i b_i|}{\left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |b_i^q|\right)^{\frac{1}{q}}} \leqslant \frac{1}{p}\frac{\sum_{i=1}^{n} |a_i^p|}{\sum_{i=1}^{n} |a_i^p|} + \frac{1}{q}\frac{\sum_{i=1}^{n} |b_i^q|}{\sum_{i=1}^{n} |b_i^q|} = \frac{1}{p} + \frac{1}{q} = 1.$$

□

Observação: Tomando $p = q = 2$ na desigualdade de Hölder, obtemos a desigualdade de Cauchy-Schwarz.

Teorema 1.5. *(Desigualdade de Minkowski) Sejam $a_1, a_2, \ldots, a_n$; $b_1, b_2, \ldots, b_n$ números reais e $p > 1$. Então tem-se a seguinte desigualdade*

$$\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |b_i^p|\right)^{\frac{1}{p}}.$$

Demonstração. Para $p > 1$ escolhemos $q > 1$ tal que $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Aplicando a desigualdade de Hölder temos,

$$\begin{aligned}
\sum_{i=1}^{n} |a_i + b_i|^p &= \sum_{i=1}^{n} |a_i + b_i||a_i + b_i|^{p-1} \\
&\leqslant \sum_{i=1}^{n} |a_i||a_i + b_i|^{p-1} + \sum_{i=1}^{n} |b_i||a_i + b_i|^{p-1} \\
&\leqslant \left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} \left(|a_i + b_i|^{p-1}\right)^q\right)^{\frac{1}{q}} + \left(\sum_{i=1}^{n} |b_i^p|\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} \left(|a_i + b_i|^{p-1}\right)^q\right)^{\frac{1}{q}} \\
&= \left(\sum_{i=1}^{n} \left(|a_i + b_i|^{p-1}\right)^q\right)^{\frac{1}{q}} \left(\left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |b_i^p|\right)^{\frac{1}{p}}\right) \\
&= \left(\sum_{i=1}^{n} \left(|a_i + b_i|^{p-1}\right)^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \left(\left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |b_i^p|\right)^{\frac{1}{p}}\right) \\
&= \left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{\frac{p-1}{p}} \left(\left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |b_i^p|\right)^{\frac{1}{p}}\right).
\end{aligned}$$

Logo,

$$\begin{aligned}
&\sum_{i=1}^{n} |a_i + b_i|^p \leqslant \left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{\frac{p-1}{p}} \left(\left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |b_i^p|\right)^{\frac{1}{p}}\right) \\
\Leftrightarrow &\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{1-\frac{p-1}{p}} \leqslant \left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |b_i^p|\right)^{\frac{1}{p}} \\
\Leftrightarrow &\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} |a_i^p|\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |b_i^p|\right)^{\frac{1}{p}}.
\end{aligned}$$

□

Observação: As demonstrações das desigualdades aplicam-se igualmente a séries infinitas tais que $\sum |x_j|^p < \infty$.

1.1 ESPAÇOS MÉTRICOS

Definição 1.2. *(Espaço métrico, métrica) Um espaço métrico é um par* (X, d)*, onde* X *é um conjunto e* d *é uma métrica em* X *(ou função distância em* X*), que é, uma função definida em* $X \times X$ *tal que para todo* $x, y, z \in X$ *tem se:*
(M1) $d : X \times X \longrightarrow [0, \infty)$.
(M2) $d(x, y) = 0$ *se, e somente se,* $x = y$.
(M3) $d(x, y) = d(y, x)$ *(simetria).*

(M4) $d(x,y) \leqslant d(x,z) + d(z,y)$ *(desigualdade triangular).*

Seguem alguns exemplos de espaços métricos.

1.1.1 ESPAÇO EUCLIDIANO $\mathbb{R}^n$

O conjunto $\mathbb{R}^n$ é formado por todas as n-uplas $(x_1, x_2, \cdots, x_n)$, onde cada $x_i \in \mathbb{R}$. Existem três métricas importantes sobre $\mathbb{R}^n$ e que de uma certa forma são equivalentes.

Sendo $x = (x_1, x_2, \cdots, x_n)$ e $y = (y_1, y_2, \cdots, y_n)$ pontos arbitrários do $\mathbb{R}^n$ temos

$$\begin{aligned}
d(x,y) =& \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2} = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}. \\
d_1(x,y) =& |x_1 - y_1| + \cdots + |x_n - y_n| = \sum_{i=1}^{n} |x_i - y_i|. \\
d_2(x,y) =& max\{|x_1 - y_1|, \cdots, |x_n - y_n|\} = \max_{1 \leqslant i \leqslant n} |x_i - y_i|.
\end{aligned}$$

A verificação de que são métricas sobre $\mathbb{R}^n$ para os axiomas (M1),(M2) e (M3) é imediata. Mostraremos então que se verifica o axioma (M4).

De fato, sejam $x = (x_1, x_2, \cdots, x_n), y = (y_1, y_2, \cdots, y_n)$ e $z = (z_1, z_2, \cdots, z_n)$ pontos no $\mathbb{R}^n$.

Então para a métrica d temos,

$$\begin{aligned}
(d(x,y))^2 =& \sum_{i=1}^{n} (x_i - y_i)^2 \\
=& \sum_{i=1}^{n} (x_i - z_i + z_i - y_i)^2 = \sum_{i=1}^{n} (x_i - y_i)^2 + 2\sum_{i=1}^{n} (x_i - z_i)(z_i - y_i) + \sum_{i=1}^{n} (z_i - y_i)^2.
\end{aligned}$$

Aplicando a desigualdade de *Cauchy-Schwarz*,

$$\begin{aligned}
&\sum_{i=1}^{n} (x_i - y_i)^2 + 2\sum_{i=1}^{n} (x_i - z_i)(z_i - y_i) + \sum_{i=1}^{n} (z_i - y_i)^2 \leqslant \\
&\leqslant \sum_{i=1}^{n} (x_i - y_i)^2 + 2\left(\sum_{i=1}^{n} (x_i - z_i)^2\right)^{1/2} \left(\sum_{i=1}^{n} (z_i - y_i)^2\right)^{1/2} + \sum_{i=1}^{n} (z_i - y_i)^2 = \\
&= \left(\sqrt{\sum_{i=1}^{n} (x_i - z_i)^2} + \sqrt{\sum_{i=1}^{n} (z_i - y_i)^2}\right)^2 = (d(x,z) + d(z,y))^2.
\end{aligned}$$

Para a métrica d_1 temos,

$$d_1(x,y) = \sum_{i=1}^{n} |x_i - y_i| = \sum_{i=1}^{n} |x_i - z_i + z_i - y_i| \leqslant \sum_{i=1}^{n} |x_i - z_i| + \sum_{i=1}^{n} |z_i - y_i| = d_1(x,z) + d_1(z,y).$$

Para a métrica d_2 temos,

$$\begin{aligned} d_2(x,y) =& \underset{1\leqslant i\leqslant n}{max}|x_i - y_i| \\ =& \underset{1\leqslant i\leqslant n}{max}|x_i - z_i + z_i - y_i| \leqslant \underset{1\leqslant i\leqslant n}{max}|x_i - z_i| + \underset{1\leqslant i\leqslant n}{max}|z_i - y_i| = d_2(x,z) + d_2(z,y). \end{aligned}$$

Sejam d, d_1 e d_2 as métricas definidas em $\mathbb{R}^n$. Quaisquer que sejam $x, y \in \mathbb{R}^n$, tem-se:

$$d_2(x,y) \leqslant d(x,y) \leqslant d_1(x,y) \leqslant nd_2(x,y).$$

1.1.2 ESPAÇO DE FUNÇÕES $C[a,b]$

Consideremos X o conjunto de todas as funções de valores reais $x, y, \ldots$ que são funções de uma variável real t definidas e contínuas em $J = [a,b]$. Seja a métrica definida por

$$d(x,y) = \max_{t\in J}|x(t) - y(t)|$$

ou

$$d(x,y) = \int_a^b |x(t) - y(t)|dt,$$

(X, d) é espaço métrico e será denotado por $C[a,b]$.

Verifiquemos que as métricas assim definidas satisfazem os axiomas de espaço métrico.

$$d(x,y) = \max_{t\in J}|x(t) - y(t)| = 0 \Leftrightarrow |x(t) - y(t)| = 0 \Leftrightarrow x(t) = y(t),\ \forall\ t \in J \Leftrightarrow x = y.$$

$$d(x,y) = \int_a^b |x(t) - y(t)|dt = 0 \Leftrightarrow |x(t) - y(t)| = 0 \Leftrightarrow x(t) = y(t) \Leftrightarrow x = y.$$

$$d(x,y) = \max_{t\in J}|x(t) - y(t)| = \max_{t\in J}|y(t) - x(t)| = d(y,x).$$

$$d(x,y) = \int_a^b |x(t) - y(t)|dt = \int_a^b |y(t) - x(t)|dt = d(y,x).$$

$$\begin{aligned} d(x,y) = \max_{t\in J}|x(t) - y(t)| &= \max_{t\in J}|x(t) - z(t) + z(t) - y(t)| \\ &\leqslant \max_{t\in J}|x(t) - z(t)| + \max_{t\in J}|z(t) - y(t)| \\ &= d(x,z) + d(z,y). \end{aligned}$$

$$\begin{aligned} d(x,y) = \int_a^b |x(t) - y(t)|dt &= \int_a^b |x(t) - z(t) + z(t) - y(t)|dt \\ &\leqslant \int_a^b |x(t) - z(t)|dt + \int_a^b |z(t) - y(t)|dt \\ &= d(x,z) + d(z,y). \end{aligned}$$

1.1.3 ESPAÇO ℓ^p E ESPAÇO DE SEQUÊNCIAS DE HILBERT ℓ^2

Seja $p \geqslant 1$ um número real fixo. Por definição, cada elemento no espaço ℓ^p é uma sequência $x = (\xi_j) = (\xi_1, \xi_2, \dots)$ de números tais que $|\xi_1|^p + |\xi_2|^p + \cdots$ converge; assim

$$\sum_{j=1}^{\infty} |\xi_j|^p < \infty$$

e a métrica é definida por

$$d(x,y) = \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p \right)^{1/p}$$

onde $y = (\eta_j)$ e $\sum |\eta_j|^p < \infty$.

No caso $p = 2$ temos o espaço de sequências de Hilbert ℓ^2 com métrica definida por

$$d(x,y) = \sqrt{\sum_{j=1}^{\infty} |\xi_j - \eta_j|^2}.$$

Mostremos agora que ℓ^p é um espaço métrico. Como anteriormente os axiomas $(M1), (M2)$ e $(M3)$ são de verificação imediata. Mostraremos a validade do axioma $(M4)$. De fato, sejam $x = (\xi_j) = (\xi_1, \xi_2, \dots) \in \ell^p$, $y = (\eta_j) = (\eta_1, \eta_2, \dots) \in \ell^p$, $z = (\zeta_j) = (\zeta_1, \zeta_2, \dots) \in \ell^p$ e aplicando a desigualdade triangular para números e em seguida a desigualdade de Minkowski temos

$$\begin{aligned} d(x,y) = \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p \right)^{1/p} &\leqslant \left(\sum_{j=1}^{\infty} [|\xi_j - \zeta_j| + |\zeta_j - \eta_j|]^p \right)^{1/p} \\ &\leqslant \left(\sum_{j=1}^{\infty} |\xi_j - \zeta_j|^p \right)^{1/p} + \left(\sum_{j=1}^{\infty} |\zeta_j - \eta_j|^p \right)^{1/p} \\ &= d(x,z) + d(z,y). \end{aligned}$$

Apresentaremos a seguir definições e teoremas que serão importantes nos estudos que pretendemos realizar.

1.2 CONJUNTO ABERTO E CONJUNTO FECHADO

Definição 1.3. *(Bola e Esfera) Seja a um ponto no espaço métrico M. Dado um número real $r > 0$, define-se:*

*1) A **bola aberta** de centro a e raio r é o conjunto $B(a;r)$ dos pontos de M cuja distância ao ponto a é menor do que r, ou seja, $B(a;r) = \{x \in M; d(x,a) < r\}$.*

*2) A **bola fechada** de centro a e raio r é o conjunto $B[a;r]$ dos pontos de M cuja distância ao ponto a é menor do que ou igual a r, ou seja, $B[a;r] = \{x \in M; d(x,a) \leqslant r\}$.*

*3) A **esfera** de centro a e raio r é o conjunto $S(a;r)$ dos pontos de M cuja distância ao ponto a é igual a r, ou seja, $S(a;r) = \{x \in M; d(x,a) = r\}$.*

Definição 1.4. *(Distância de um ponto a um conjunto) Sejam a um ponto e X um subconjunto não-vazio de um espaço métrico M. Define-se a distância do ponto a ao conjunto X como o número real*

$$d(a, X) = \inf_{x \in X} d(a, x).$$

Definição 1.5. *(Distância entre dois conjuntos) Sejam X e Y subconjuntos não vazios do espaço métrico M. Define-se a distância entre os subconjuntos X e Y como sendo*

$$d(X, Y) = \inf\{d(x, y); x \in X, y \in Y\}.$$

Teorema 1.6. *Se x, y e z são pontos genéricos de um espaço métrico (M, d), então*

$$|d(x, y) - d(x, z)| \leqslant d(y, z).$$

Demonstração. Da desigualdade triangular $(M4)$ decorre que

$$d(x, y) - d(x, z) \leqslant d(z, y).$$

Por outro lado podemos expressar a mesma desigualdade $(M4)$ por

$$d(x, z) \leqslant d(x, y) + d(y, z)$$

$$d(x, z) - d(x, y) \leqslant d(y, z).$$

Portanto

$$|d(x, y) - d(x, z)| \leqslant d(y, z).$$

□

Definição 1.6. *(Ponto interior) Seja X um subconjunto de um espaço métrico M. Um ponto $a \in X$ diz-se um ponto interior a X quando é centro de uma bola aberta contida em X, ou seja, quando existe $r > 0$ tal que $d(x, a) < r \Rightarrow x \in X$. Chama-se o interior de X em M ao conjunto $int(X)$ formado pelos pontos interiores a X.*

Definição 1.7. *(Vizinhança) Seja M um espaço métrico. Diz-se que o conjunto V é uma vizinhança do ponto $a \in M$ quando $a \in int(V)$.*

Assim, V é uma vizinhança de a se, e somente se, V contém um aberto que contém a. Notação: V_a.

Definição 1.8. *(Ponto isolado) Seja M um espaço métrico. Um ponto $a \in M$ chama-se um ponto isolado de M quando ele é uma bola aberta em M, ou seja, quando existe $r > 0$ tal que $B(a; r) = \{a\}$.*

Isto significa, que além do próprio a, não existem outros pontos de M a uma distância de a inferior a r. Dizer que um ponto $a \in M$ não é isolado significa, portanto, afirmar que para todo $r > 0$ pode-se encontrar um ponto $x \in A$ tal que $0 < d(a, x) < r$.

Definição 1.9. *(Fronteira) A fronteira de X em M é o conjunto ∂X, formado pelos pontos $b \in M$ tais que toda bola aberta de centro b contém pelo menos um ponto de X e um ponto do complementar $M - X$.*

Definição 1.10. *(Ponto aderente) Um ponto a diz-se aderente a um subconjunto X de um espaço métrico M quando $d(a, X) = 0$.*

Isto significa que existem pontos de X arbitrariamente próximos de a, ou seja, para cada $\varepsilon > 0$, podemos encontrar $x \in X$ tal que $d(a, x) < \varepsilon$.

Outras maneiras equivalentes de dizer que a é aderente a X são:

1) para todo $\varepsilon > 0$, tem-se $B(a; \varepsilon) \cap X \neq \emptyset$;

2) para todo aberto A contendo a, tem-se $A \cap X \neq \emptyset$;

3) toda vizinhança de a tem pontos em comum com X.

Definição 1.11. *(Fecho) O fecho de um conjunto X num espaço métrico M é o conjunto $\overline{X}$ dos pontos de M que são aderentes a X.*

Portanto, escrever $a \in \overline{X}$ é o mesmo que afirmar que o ponto a é aderente a X em M. Tem-se $\overline{\emptyset} = \emptyset$, $\overline{M} = M$ e $X \subset \overline{X}$ para todo $X \subset M$. É também claro que $X \subset Y \Rightarrow \overline{X} \subset \overline{Y}$.

Definição 1.12. *(Conjunto denso) Um subconjunto $X \subset M$ diz-se denso em M quando $\overline{X} = M$.*

Isto é, quando toda bola aberta em M contém algum ponto de X, ou ainda, para cada aberto não-vazio A em M, tem-se $A \cap X \neq \emptyset$

Definição 1.13. *(Ponto de acumulação) Seja X um subconjunto do espaço métrico M. Um ponto $a \in M$ chama-se ponto de acumulação de X quando toda bola de centro a contém algum ponto de X, diferente do ponto a. Indicaremos com a notação X' o conjunto dos pontos de acumulação de X em M.*

Definição 1.14. *(Conjunto aberto) Um subconjunto A de um espaço métrico M diz-se aberto em M quando todos os seus pontos são interiores, isto é, $intA = A$.*

Assim, $A \subset M$ é aberto se, e somente se, $A \cap \partial A = \emptyset$. Dito de forma equivalente, um conjunto $A \subset M$ é aberto em M se for possível obter, para cada $x \in A$, um raio $r > 0$ tal que $B(x; r) \subset A$.

Definição 1.15. *(Conjunto fechado) Seja M um espaço métrico. Um conjunto $F \subset M$ diz-se fechado em M se seu complementar $F^c = M - F$ é aberto.*

Teorema 1.7. *Seja M um espaço métrico. Então*

1) $\emptyset$ e M são abertos.

2) Se $A_1, \dots, A_n$ são abertos em M, então $A_1 \cap \cdots \cap A_n$ é aberto em M. (A interseção de um número finito de conjuntos abertos é um conjunto aberto).

3) Se $\{A_\lambda\}_{\lambda \in J}$ é uma família qualquer de abertos em M, onde J é um conjunto de índices, então $\bigcup_{\lambda \in J} A_\lambda$ é aberto em M. (A reunião de uma família qualquer de conjuntos abertos é um conjunto aberto).

Demonstração. 1) $\emptyset$ é aberto pois, como não existe $x \in \emptyset$, o conjunto $\emptyset$ não viola a condição que define abertos. Agora, M é aberto pois, considere o subconjunto $M \subset M$, para todo $x \in M$, existe $r > 0$, tal que $B(x; r) \subset M$.

2) Seja $x \in A_1 \cap \cdots \cap A_n$, então $x \in A_i$, $i = 1, 2, \dots, n$. Como A_i é aberto, para cada $i = 1, 2, \dots, n$ existe uma bola aberta $B(x; r_i) \subset A_i$. Considere $r = \min\{r_1, \dots, r_n\}$, $r > 0$. Então $B(x; r) \subset B(x; r_i) \subset A_i$ para cada i. Logo, $B(x; r) \subset (A_1 \cap \cdots \cap A_n)$ e, portanto, $A_1 \cap \cdots \cap A_n$ é aberto.

3) Seja $A = \bigcup_{\lambda \in J} A_\lambda$. Dado $x \in A$, existe um índice $\lambda \in J$ tal que $x \in A_\lambda$. Como A_λ é aberto, existe $B(x; r)$, $r > 0$, tal que $B(x; r) \subset A_\lambda$. Logo, $B(x; r) \subset A$ e, portanto, A é aberto. □

1.3 TOPOLOGIA E CONTINUIDADE

Um espaço topológico é um par $(X, \mathcal{T})$ onde X é um conjunto e $\mathcal{T}$ é uma topologia em X. Todo espaço métrico M pode ser considerado, de modo natural, como um espaço topológico, no qual a coleção $\mathcal{T}$ é formada pelos subconjuntos abertos de M. Uma topologia $\mathcal{T}$ em X se diz metrizável quando existe uma métrica em X em relação à qual os abertos são os elementos de $\mathcal{T}$. Duas métricas d_1 e d_2 sobre um conjunto M são equivalentes se, e somente se, elas determinam a mesma topologia em M.

Definição 1.16. *(Topologia) Uma topologia num conjunto X é uma coleção $\mathcal{T}$ de subconjuntos de X, chamados os abertos da topologia, com as seguintes propriedades:*

1) $\emptyset$ e M pertencem a $\mathcal{T}$;

2) Se $A_1, \dots, A_n \in \mathcal{T}$ então $A_1 \cap \cdots \cap A_n \in \mathcal{T}$;

3) Dada uma família arbitrária $(A_\lambda)_{\lambda \in L}$ com $A_\lambda \in \mathcal{T}$ para cada $\lambda \in L$, tem-se $\bigcup_{\lambda \in L} A_\lambda \in \mathcal{T}$.

Definição 1.17. *(Continuidade) Sejam M, N espaços métricos. Diz-se que a aplicação $f : M \longrightarrow N$ é contínua no ponto $a \in M$ quando, para todo $\varepsilon > 0$ dado, é possível obter $\delta > 0$ tal que $d(x, a) < \delta$ implica $d(f(x), f(a)) < \varepsilon$. Diz-se que $f : M \longrightarrow N$ é contínua quando ela é contínua em todos os pontos $a \in M$. Equivalentemente, $f : M \longrightarrow N$ é contínua no ponto $a \in M$ quando, dada qualquer bola $B' = B(f(a); \varepsilon)$ de centro $f(a)$, pode se encontrar uma bola $B = B(a; \delta)$, de centro a, tal que $f(B) \subset B'$.*

Definição 1.18. *(Continuidade uniforme) Sejam M, N espaços métricos. Uma aplicação $f : M \longrightarrow N$ diz-se uniformemente contínua quando, para todo $\varepsilon > 0$ dado, existe $\delta > 0$ tal que, sejam quais forem $x, y \in M$, $d(x, y) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon$.*

É importante e interessante notar que a continuidade pode ser caracterizada em termos de conjuntos abertos. O que reforça a importância da noção de conjunto aberto na topologia dos espaços métricos.

Teorema 1.8. *Sejam M, N espaços métricos. A fim de que uma aplicação $f : M \longrightarrow N$ seja contínua, é necessário e suficiente que a imagem inversa $f^{-1}(A')$ de todo subconjunto aberto $A' \subset N$ seja um subconjunto aberto de M.*

Demonstração. Suponhamos primeiramente que f seja contínua, tomemos $A' \subset N$ aberto e mostremos que $f^{-1}(A')$ é aberto em M. De fato, para cada $a \in f^{-1}(A')$, temos $f(a) \in A'$. Pela definição de conjunto aberto, existe $\varepsilon > 0$ tal que $B(f(a); \varepsilon) \subset A'$. Sendo f contínua no ponto a, ao ε corresponde um $\delta > 0$ tal que $f(B(a; \delta)) \subset B(f(a); \varepsilon) \subset A'$. Isto quer dizer que $B(a; \delta) \subset f^{-1}(A')$. Logo $f^{-1}(A')$ é aberto.

Reciprocamente, suponhamos que a imagem inversa por f de cada aberto em N seja um aberto em M. Seja $a \in M$. Mostraremos que f é contínua no ponto a. Com efeito, dado $\varepsilon > 0$, a bola $A' = B(f(a); \varepsilon)$ é um aberto em N, contendo $f(a)$. Logo sua imagem inversa $A = f^{-1}(A')$ é um aberto em M, contendo a. Assim existe $\delta > 0$ tal que $B(a; \delta) \subset A$, ou seja, $f(B(a; \delta)) \subset B(f(a); \varepsilon)$. □

Definição 1.19. *Uma aplicação $f : M \longrightarrow N$ é chamada de lipschitziana quando*

$$\exists\ c > 0 \text{ para } \forall x, y \in M \text{ tal que } d(f(x), f(y)) \leqslant c d(x, y).$$

Além disso, a aplicação $f : M \to N$ é chamada

1) uma ***equivalência*** *quando f é bijetora e ambas f e f^{-1} são lipschitzianas.*

2) uma ***contração*** *quando é lipschitziana com constante $c < 1$.*

3) uma ***isometria****, e M, N são ditas serem isométricas, quando f preserva distâncias, isto é,*

$$\forall x, y \in M,\ d(f(x), f(y)) = d(x, y).$$

1.4 ESPAÇOS COMPLETOS

Ser completo é uma propriedade adicional que um espaço métrico pode ou não ter. Existem várias consequências que fazem dos espaços métricos completos serem espaços reais mais interessantes que os não completos.

Definição 1.20. *(Sequência) Uma sequência num conjunto M é uma aplicação $x : \mathbb{N} \longrightarrow M$, definida no conjunto $\mathbb{N} = \{1, 2, \ldots, n, \ldots\}$. O valor que a sequência x assume no número $n \in \mathbb{N}$ será indicado por x_n. As seguintes notações são usadas para representar uma sequência, $(x_1, x_2, \ldots, x_n, \ldots), (x_n)_{n\in\mathbb{N}}$ ou (x_n). O conjunto dos valores são escritos como $\{x_1, x_2, \ldots, x_n, \ldots\}$ ou $\{x_n; n \in \mathbb{N}\}$.*

Definição 1.21. *(Sequência limitada) Uma sequência (x_n) no espaço métrico M chama-se limitada quando o conjunto dos seus termos é limitado, isto é, quando existe $c > 0$ tal que $d(x_m, x_n) \leqslant c$ para quaisquer $m, n \in \mathbb{N}$.*

Definição 1.22. *(Limite de sequência) Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência num espaço métrico M. Diz-se que o ponto $a \in M$ é limite da sequência (x_n) quando, para todo número $\varepsilon > 0$ dado arbitrariamente, pode-se obter $n_0 \in \mathbb{N}$ tal que $n > n_0 \Rightarrow d(x_n, a) < \varepsilon$. Ou ainda que $\lim\limits_{n\to\infty} d(x_n, a) = 0$. Escreve-se então $a = \lim x_n$, $a = \lim\limits_{n\to\infty} x_n$ ou $a = \lim\limits_{n\in\mathbb{N}} x_n$. Dizemos que a sequência $(x_n)_{n\in\mathbb{N}}$ é convergente em M e converge para a. Ou que x_n tende para a, e indicamos $x_n \longrightarrow a$.*

Observação: Devemos notar que o limite de uma sequência convergente deve ser um ponto do espaço M. Por exemplo, seja M o intervalo aberto $(0, 1)$ em $\mathbb{R}$ com a métrica dada por $d(x, y) = |x - y|$. Então a sequência $\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right)$ não é convergente uma vez que 0, o ponto para o qual a sequência converge não está em M.

Teorema 1.9. *Um ponto a, num espaço métrico M, é limite de uma subsequência de (x_n) se, e somente se, toda bola aberta de centro a contém termos x_n com índices n arbitrariamente grandes.*

Demonstração. Se uma subsequência $(x_{n_1}, x_{n_2}, \ldots, x_{n_k}, \ldots)$ converge para a então, dado $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow x_{n_k} \in B(a; \varepsilon)$. Logo toda bola $B(a; \varepsilon)$ de centro a contém termos com índices arbitrariamente grandes, a saber, todos os índices n_k com $k > k_0$. Reciprocamente, supondo cumprida esta condição, a bola $B(a; 1)$ contém um termo x_{n_1}, a bola $B(a; 1/2)$ contém um termo x_{n_2} com índice $n_2 > n_1$, e assim por diante: para todo $k \in \mathbb{N}$ podemos achar $x_{n_k} \in B(a; 1/k)$ com $n_k > n_{k-1} > \cdots > n_2 > n_1$. Isto define um subconjunto infinito $\mathbb{N}' = \{n_1 < n_2 < \cdots < n_k < \cdots\}$ e uma subsequência x_{n_k} tal que $d(x_{n_k}, a) < 1/k$. Segue-se que $\lim\limits_{k\to\infty} x_{n_k} = a$. □

Observação: No enunciado do teorema anterior, podemos substituir "bola aberta de centro a" por "conjunto aberto contendo a" ou "vizinhança de a".

Teorema 1.10. *Toda sequência monótona limitada de números reais é convergente.*

Demonstração. Para fixar as ideias, seja $(x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n \leqslant \dots)$ a sequência limitada em questão. Tomemos $a = \sup_{n \in \mathbb{N}} x_n$. Afirmamos que $a = \lim x_n$. Com efeito, dado arbitrariamente $\varepsilon > 0$, o número $a - \varepsilon$, sendo menor do que a, não pode ser cota superior do conjunto dos valores x_n. Logo existe $n_0 \in \mathbb{N}$ tal que $a - \varepsilon < x_{n_0} \leqslant a$. Então $n > n_0 \Rightarrow a - \varepsilon < x_{n_0} \leqslant x_n \leqslant a < a + \varepsilon \Rightarrow a - \varepsilon < x_n < a + \varepsilon$. Isto conclui a demonstração. □

Lema 1.1. *Seja $M = (M, d)$ um espaço métrico. Então:*

1) Uma sequência convergente em M é limitada e seu limite é único.

2) Se $x_n \longrightarrow x$ e $y_n \longrightarrow y$ em M, então $d(x_n, y_n) \longrightarrow d(x, y)$.

Demonstração. 1) Suponha que $x_n \longrightarrow x$. Então, seja $\epsilon = 1$, podemos determinar n_0 tal que $d(x_n, x) < 1$ para todo $n > n_0$. Consequentemente pela condição $(M4)$, para todo n temos $d(x_n, x) < 1 + a$ onde $a = \max\{d(x_1, x), \dots, d(x_{n_0}, x)\}$. Isto mostra que (x_n) é limitada.

Assumindo que $x_n \longrightarrow x$ e $x_n \longrightarrow z$, obtemos de $(M4)$

$$0 \leqslant d(x, z) \leqslant d(x, x_n) + d(x_n, z) \longrightarrow 0 + 0$$

e a unicidade $x = z$ do limite segue do postulado $(M2)$.

2)Da generalização da desigualdade triangular obtemos

$$d(x_n, y_n) \leqslant d(x_n, x) + d(x, y) + d(y, y_n)$$

$$d(x_n, y_n) - d(x, y) \leqslant d(x_n, x) + d(y, y_n).$$

E podemos escrever também

$$d(x, y) \leqslant d(x, x_n) + d(x_n, y_n) + d(y_n, y)$$

$$d(x_n, y_n) - d(x, y) \geqslant -d(x, x_n) - d(y_n, y).$$

Logo

$$-(d(x_n, x) + d(y, y_n)) \leqslant d(x_n, y_n) - d(x, y) \leqslant d(x_n, x) + d(y, y_n).$$

Ou seja

$$|d(x_n, y_n) - d(x, y)| \leqslant d(x_n, x) + d(y, y_n).$$

Pela hipótese $x_n \longrightarrow x$ e $y_n \longrightarrow y$, então

$$|d(x_n, y_n) - d(x, y)| \leqslant d(x_n, x) + d(y, y_n) \longrightarrow 0$$

quando $n \longrightarrow \infty$. Portanto $d(x_n, y_n) \longrightarrow d(x, y)$.

□

Definição 1.23. *(Sequência de Cauchy) Uma sequência* $(x_n)_{n\in\mathbb{N}}$ *num espaço métrico* M *chama-se uma sequência de Cauchy quando, para todo* $\varepsilon > 0$ *dado, existe* $n_0 \in \mathbb{N}$ *tal que para todo* $m, n \in \mathbb{N}$, $m, n > n_0 \Rightarrow d(x_m, x_n) < \varepsilon$.

Teorema 1.11. *Toda sequência convergente é de Cauchy.*

Demonstração. Se $\lim x_n = a$ no espaço métrico M então, dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $n > n_0 \Rightarrow d(x_n, a) < \varepsilon/2$. Se tomarmos $m, n > n_0$ teremos

$$d(x_m, x_n) \leqslant d(x_m, a) + d(x_n, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Logo, (x_n) é de Cauchy. □

Observação: A recíproca deste teorema não é válida. Ou seja, uma sequência de Cauchy de um espaço métrico M pode não convergir em M. Tomemos uma sequência de números racionais x_n convergindo para um número irracional a. Por exemplo, $x_1 = 1, x_2 = 1,4, x_3 = 1,41, \ldots$, com $\lim x_n = \sqrt{2}$. Sendo convergente em $\mathbb{R}$, segue do teorema anterior que $(x_n)_{n\in\mathbb{N}}$ é de Cauchy no espaço métrico $\mathbb{Q}$ dos números racionais, mas não é convergente em $\mathbb{Q}$, pois $\sqrt{2} \notin \mathbb{Q}$. Este fato também ilustra o conceito de que convergência não é uma propriedade intrínseca da sequência, mas que depende também do espaço no qual a sequência é considerada.

Teorema 1.12. *Toda sequência de Cauchy é limitada.*

Demonstração. Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de Cauchy no espaço métrico M. Dado $\varepsilon = 1$, existe $n_0 \in \mathbb{N}$ tal que $m, n > n_0 \Rightarrow d(x_m, x_n) < 1$. Logo o conjunto $\{x_{n_0+1}, x_{n_0+2}, \ldots\}$ é limitado e tem diâmetro $\leqslant 1$. Segue-se que

$$\{x_1, x_2, \ldots, x_n, \ldots\} = \{x_1, \ldots, x_{n_0}\} \cup \{x_{n_0+1}, x_{n_0+2}, \ldots\}$$

é limitado.

□

Observação: Não vale a recíproca deste teorema. Tomemos a sequência dada por $(1, 0, 1, 0, \ldots)$ na reta. Embora limitada, esta sequência não é de Cauchy pois $d(x_n, x_{n+1}) = 1$ para todo n.

Teorema 1.13. *Uma sequência de Cauchy que possua uma subsequência convergente é convergente (e tem o mesmo limite que a subsequência).*

Demonstração. Sejam (x_n) uma sequência de Cauchy no espaço métrico M e (x_{n_k}) uma subsequência que converge para o ponto $a \in M$. Afirmamos que $\lim x_n = a$. Com efeito, dado $\varepsilon > 0$, existe $p \in \mathbb{N}$ tal que $n_k > p \Rightarrow d(x_{n_k}, a) < \varepsilon/2$. Existe também $q \in \mathbb{N}$ tal que

$m, n > q \Rightarrow d(x_m, x_n) < \varepsilon/2$. Seja $n_0 = \max\{p, q\}$. Para todo $n > n_0$ e fixado $n_k > n_0$ temos

$$d(x_n, a) \leqslant d(x_n, x_{n_k}) + d(x_{n_k}, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Logo $\lim x_n = a$. □

*Observação:*A propriedade enunciada no teorema anterior é evidentemente falsa para sequências arbitrárias. Ela indica que uma sequência de Cauchy só não converge num espaço M se "faltarem pontos no espaço".

Definição 1.24. *Diz-se que o espaço métrico M é completo quando toda sequência de Cauchy em M é convergente.*

Teorema 1.14. *A reta é um espaço métrico completo.*

Demonstração. Seja (x_n) uma sequência de Cauchy em $\mathbb{R}$. Pondo, para cada $n \in \mathbb{N}$, $X_n = \{x_n, x_{n+1}, \dots\}$, temos $X_1 \supset X_2 \supset \cdots \supset X_n \supset \cdots$ e os conjuntos X_n são limitados. Seja $a_n = \inf X_n$ $(n = 1, 2, 3, \dots)$. Então $a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n \leqslant \cdots \leqslant b = \sup X_1$. Pelo teorema 1.10, existe o número $a = \lim a_n$. Afirmamos que $a = \lim x_n$. Para provar isto, basta mostrar que a é limite de uma subsequência de (x_n), ou seja, que dados arbitrariamente $\varepsilon > 0$ e $n_1 \in \mathbb{N}$, podemos obter $n > n_1$ tal que $x_n \in (a - \varepsilon, a + \varepsilon)$. (Vide teoremas 1.9 e 1.13). Ora, sendo $a = \lim a_n$, existe $m > n_1$ tal que $a - \varepsilon < a_m < a + \varepsilon$. Como $a_m = \inf X_m$ existe $n \geqslant m$ (e portanto $n > n_1$) tal que $a_m \leqslant x_n < a + \varepsilon$, isto é, $x_n \in (a - \varepsilon, a + \varepsilon)$. □

Teorema 1.15. *Seja F um subconjunto não vazio de um espaço métrico (M, d) e $\overline{F}$ o seu fecho. Então:*

(a) $x \in \overline{F}$ se, e somente se, existir uma sequência (x_n) em F tal que $x_n \longrightarrow x$.

(b) F é fechado se, e somente se, a condição $x_n \in F$, $x_n \longrightarrow x$ implica que $x \in F$.

Demonstração.

(a) Seja $x \in \overline{F}$. Se $x \in F$, uma sequência do tipo é $(x, x, \dots)$. Se $x \notin F$, ele é um ponto de acumulação de F. Consequentemente para cada $n = 1, 2, \dots$ a bola $B(x; 1/n)$ contém um $x_n \in F$, e $x_n \longrightarrow x$ porque $1/n \longrightarrow 0$ quando $n \longrightarrow \infty$.

Reciprocamente, se x_n está em F e $x_n \longrightarrow x$, então $x \in F$ ou toda vizinhança de x contém pontos $x_n \neq x$, tal que x é um ponto de acumulação de F. Consequentemente $x \in \overline{F}$, pela definição de fecho.

(b) F é fechado se, e somente se, $F = \overline{F}$, de modo que (b) segue imediatamente de (a). □

Teorema 1.16. *Um subespaço F de um espaço métrico completo M é completo se, e somente se, o conjunto F é fechado em M.*

Demonstração. Seja F completo. Pelo teorema 1.15(a), para todo $x \in \overline{F}$ existe uma sequência (x_n) em F que converge para x. Uma vez que (x_n) é Cauchy pelo teorema 1.11 e F é completo, (x_n) converge em F, o limite é único pelo lema 1.1. Consequentemente $x \in F$. Isto prova que F é fechado porque $x \in \overline{F}$ é arbitrário.

Reciprocamente, seja F fechado e (x_n) Cauchy em F. Então $x_n \longrightarrow x \in M$, o que implica $x \in \overline{F}$ pelo teorema 1.15(a), e $x \in F$ uma vez que $F = \overline{F}$ por hipótese. Consequentemente uma sequência de Cauchy (x_n) arbitrária converge em F, o que prova que F é completo. □

Com base nas definições e teoremas apresentados mostraremos que os espaços $\mathbb{R}^n, C[a,b]$ e ℓ^p que utilizamos como exemplos de espaços métricos são espaços métricos completos.

1.4.1 ESPAÇO EUCLIDIANO $\mathbb{R}^n$

Teorema 1.17. *O espaço euclidiano $\mathbb{R}^n$ é completo.*

Demonstração. Seja a métrica em $\mathbb{R}^n$ definida por

$$d(x,y) = \left(\sum_{j=1}^{n}(\xi_j - \eta_j)^2\right)^{1/2}$$

onde $x = (\xi_j)$ e $y = (\eta_j)$. Consideremos uma sequência de Cauchy $x_m \in \mathbb{R}^n$ qualquer, escrevendo $x_m = (\xi_1^{(m)}, \dots, \xi_n^{(m)})$. Uma vez que $(x_m)_{n\in\mathbb{N}}$ é Cauchy, para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $m, r > n_0$ implica em

$$d(x_m, x_r) = \left(\sum_{j=1}^{n}(\xi_j^{(m)} - \xi_j^{(r)})^2\right)^{1/2} < \varepsilon \tag{1.2}$$

Elevando ao quadrado, temos para $m, r > n_0$ e $j = 1, \dots, n$

$$(\xi_j^{(m)} - \xi_j^{(r)})^2 < \varepsilon^2 \qquad \text{e} \qquad |\xi_j^{(m)} - \xi_j^{(r)}| < \varepsilon.$$

Então para cada j fixado, $(1 \leqslant j \leqslant n)$, a sequência $(\xi_j^{(1)}, \xi_j^{(2)}, \dots)$ é uma sequência de Cauchy de números reais. Pelo teorema 1.14 esta sequência converge, ou seja, $\xi_j^{(m)} \longrightarrow \xi_j$ quando $m \longrightarrow \infty$. Usando estes n limites, definimos $x = (\xi_1, \dots, \xi_n)$. Claramente, $x \in \mathbb{R}^n$. De (3.2), com $r \longrightarrow \infty$, para $m > n_0$ temos

$$d(x_m, x) \leqslant \varepsilon$$

Isto mostra que x é o limite de $(x_m)_{n\in\mathbb{N}}$ e prova que $\mathbb{R}^n$ é completo uma vez que $(x_m)_{n\in\mathbb{N}}$ é uma sequência de Cauchy arbitrária. □

1.4.2 ESPAÇO DE FUNÇÕES $C[a,b]$

Teorema 1.18. *O espaço de funções $C[a,b]$ com a métrica $d(x,y) = \max |x(t) - y(t)|$ é completo sendo $[a,b]$ qualquer intervalo fechado em $\mathbb{R}$.*

Demonstração. Seja $(x_m)_{n\in\mathbb{N}}$ uma sequência de Cauchy qualquer em $C[a,b]$. Então, para qualquer $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ temos

$$d(x_m, x_n) = \max_{t\in J} |x_m(t) - x_n(t)| < \varepsilon \tag{1.3}$$

onde $J = [a,b]$. Consequentemente para qualquer $t = t_0 \in J$ fixo, com $m, n > n_0$

$$|x_m(t_0) - x_n(t_0)| < \varepsilon.$$

Isto mostra que $(x_1(t_0), x_2(t_0), \dots)$ é uma sequência de Cauchy de números reais. Uma vez que $\mathbb{R}$ é completo (conforme teorema 1.14), a sequência converge, isto é, $x_m(t_0) \longrightarrow x(t_0)$ quando $m \longrightarrow \infty$. Desta forma, podemos associar a cada $t \in J$ um único número real $x(t)$. Isto define (ponto a ponto) uma função x em J, e mostra que $x \in C[a,b]$ e $x_m \longrightarrow x$.

De (1.3) com $n \longrightarrow \infty$ e $m > n_0$ temos

$$\max_{t\in J} |x_m(t) - x(t)| \leqslant \varepsilon.$$

Consequentemente para todo $t \in J$ com $m > n_0$

$$|x_m(t) - x(t)| \leqslant \varepsilon.$$

Isto mostra que $(x_m(t))_{n\in\mathbb{N}}$ converge para $x(t)$ uniformemente em J. Uma vez que $x'_m s$ são funções contínuas em J e a convergência é uniforme, a função limite x é contínua em J. Consequentemente $x \in C[a,b]$. Além disso, $x_m \longrightarrow x$. O que prova que $C[a,b]$ é completo. □

1.4.3 ESPAÇO ℓ^p

Teorema 1.19. *O espaço ℓ^p é completo; aqui p é fixo e $1 \leqslant p < +\infty$.*

Demonstração. Seja (x_n) uma sequência de Cauchy qualquer no espaço ℓ^p, onde $x_m = (\xi_1^{(m)}, \xi_2^{(m)}, \dots)$. Então para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$,

$$d(x_m, x_n) = \left(\sum_{j=1}^{\infty} |\xi_j^{(m)} - \xi_j^{(n)}|^p \right)^{1/p} < \varepsilon. \tag{1.4}$$

Daí resulta que para cada $j = 1, 2, \dots$ com $m, n > n_0$ temos

$$|\xi_j^{(m)} - \xi_j^{(n)}| < \varepsilon. \tag{1.5}$$

Escolhendo um j fixo. De (1.5) vemos que $(\xi_j^{(1)}, \xi_j^{(2)}, \dots)$ é uma sequência de Cauchy de números reais convergente, pois $\mathbb{R}$ é completo (conforme teorema 1.14), isto é, $\xi_j^{(m)} \longrightarrow \xi_j$ quando $m \longrightarrow \infty$. Usando estes limites, definimos $x = (\xi_1, \xi_2, \dots)$ e mostramos que $x \in \ell^p$ e $x_m \longrightarrow x$.

De (1.4) temos para todo $m, n > n_0$ e $k = 1, 2, \dots$

$$\sum_{j=1}^{k} |\xi_j^{(m)} - \xi_j^{(n)}|^p < \varepsilon^p.$$

Quando $n \longrightarrow \infty$, obtemos para $m > n_0$ e $k = 1, 2, \dots$

$$\sum_{j=1}^{k} |\xi_j^{(m)} - \xi_j|^p \leqslant \varepsilon^p.$$

Quando $k \longrightarrow \infty$; então para $m > n_0$

$$\sum_{j=1}^{\infty} |\xi_j^{(m)} - \xi_j|^p \leqslant \varepsilon^p. \tag{1.6}$$

Isto mostra que $x_m - x = (\xi_j^{(m)} - \xi_j) \in \ell^p$. Uma vez que $x_m \in \ell^p$, segue por meio da desigualdade de Minkowski que

$$x = x_m + (x - x_m) \in \ell^p.$$

Além disso, a série em (1.6) representa $[d(x_m, x)]^p$, de modo que (1.6) implica que $x_m \longrightarrow x$. Uma vez que (x_m) representa uma sequência de Cauchy arbitrária em ℓ^p, isto prova que ℓ^p é completo, onde $1 \leqslant p < \infty$. □

1.4.4 COMPLETAMENTO DE UM ESPAÇO MÉTRICO

Todo espaço métrico M pode ser ampliado, acrescentando-lhe novos pontos, de modo a obter um espaço completo $\widehat{M}$, chamado o completamento de M.

Definição 1.25. *Um completamento de um espaço métrico M é um par $(\widehat{M}, \varphi)$, onde $\widehat{M}$ é completo e $\varphi : M \longrightarrow \widehat{M}$ é uma imersão isométrica cuja imagem $\varphi(M)$ é densa em $\widehat{M}$.*

Exemplo: $\mathbb{R}$ é um completamento de $\mathbb{Q}$. Se M é um espaço métrico completo então, para todo subconjunto $X \subset M$, seu fecho $\overline{X}$ em M é um completamento de X. Em particular, o intervalo $[0, 1]$ é um completamento de $(0, 1)$.

1.5 ESPAÇOS SEPARÁVEIS

Além da completude um conceito de interesse teórico e prático é o da separabilidade de um espaço métrico. Espaços métricos separáveis são mais simples que os espaços métricos não-separáveis.

Definição 1.26. *Um espaço métrico M chama-se separável se ele contém um subconjunto enumerável que é denso.*

$$\exists\ X \subseteq M,\ \ X \ \textit{enumerável e}\ \overline{X} = M.$$

Exemplo: $\mathbb{R}$ é separável pois o subconjunto enumerável $\mathbb{Q}$ é denso em $\mathbb{R}$.

Começando do mais simples, e falando de uma maneira informal, temos:

1) *Espaços métricos finitos.* Existe um número finito de distâncias possíveis para computar.

2) *Espaços métricos enumeráveis.* Com um número infinito de pontos, um algoritmo pode ainda calcular distâncias precisamente, mas pode levar mais e mais tempo para fazê-lo.

3) *Espaços métricos separáveis.* Pontos podem ser aproximados por um dos vários pontos enumeráveis; qualquer distância pode ser avaliada, não exatamente, mas com qualquer precisão.

4) *Espaços métricos não-separáveis.* Pode não existir nenhum algoritmo que encontre a distância entre dois pontos genéricos, mesmo que aproximadamente.

2 ESPAÇOS NORMADOS E ESPAÇOS DE BANACH

Um espaço vetorial onde a métrica é definida por meio de uma norma é chamado espaço normado. Se ele for um espaço métrico normado completo é chamado de espaço de Banach. Uma aplicação de um espaço normado X em um espaço normado Y é chamado operador. Uma aplicação de X em um conjunto escalar $\mathbb{R}$ ou $\mathbb{C}$ é chamado de funcional. De particular importância são os chamados operadores lineares limitados e os funcionais lineares limitados uma vez que são contínuos e tiram vantagem da estrutura do espaço vetorial. O conjunto de todos operadores lineares limitados de um dado espaço normado X em um dado espaço normado Y forma um espaço normado denotado por $B(X, Y)$. De forma similar o conjunto de todos os funcionais lineares limitados em X formam um espaço normado, que é chamado de espaço dual X' de X.

As definições, teoremas e demonstrações deste capítulo foram retirados principalmente das referências [5], [9] e [10].

2.1 ESPAÇO NORMADO. ESPAÇO DE BANACH

Definição 2.1. *(Espaço normado. Espaço de Banach) Um espaço normado X é um espaço vetorial com uma norma definida nele. Um espaço de Banach é um espaço normado completo com a métrica induzida pela norma.*

Aqui uma norma em um espaço vetorial X, real ou complexo, é uma função de valor real em X cujo valor em $x \in X$ é denotado por

$$\|x\|$$

e que tem as seguintes propriedades

(N1) $\|x\| \geqslant 0$.

(N2) $\|x\| = 0 \Leftrightarrow x = 0$.

(N3) $\|\alpha x\| = |\alpha|\|x\|$.

(N4) $\|x + y\| \leqslant \|x\| + \|y\|$.

sendo x e y vetores arbitrários em X e α um escalar qualquer.

Uma norma em X define uma métrica d em X que é dada por

$$d(x, y) = \|x - y\|, \qquad x, y \in X.$$

e é chamada a métrica induzida pela norma. O espaço normado definido é denotado por $(X, \|\cdot\|)$ ou simplesmente X.

As propriedades $(N1)$ a $(N4)$ de uma norma são sugeridas e motivadas pelo comprimento $|x|$ de um vetor x na Álgebra Vetorial. Neste caso podemos escrever $\|x\| = |x|$.

Teorema 2.1. *Seja $\|\cdot\|$ uma norma qualquer em X. Então, para todo $x, y \in X$ tem-se $|\|x\| - \|y\|| \leqslant \|x - y\|$.*

Demonstração. $\|x\| = \|(x-y)+y\| \leqslant \|x-y\| + \|y\| \Rightarrow \|x\| - \|y\| \leqslant \|x-y\|$

$\|y\| = \|(y-x)+x\| \leqslant \|y-x\|+\|x\| \Rightarrow \|y\|-\|x\| \leqslant \|y-x\| \Rightarrow \|x\|-\|y\| \geqslant -\|x-y\|$.

Logo

$$-\|x-y\| \leqslant \|x\| - \|y\| \leqslant \|x-y\|.$$

Portanto

$$|\|x\| - \|y\|| \leqslant \|x-y\|.$$

□

Lema 2.1. *Toda norma é uma função uniformemente contínua.*

Demonstração. Para todo $\varepsilon > 0$, existe $\delta = \varepsilon > 0$ tal que para todos $x, y \in E$ com $\|x-y\| < \delta$ tem-se, pela proposição (2.1),

$$|\|x\| - \|y\|| \leqslant \|x-y\| < \delta = \varepsilon.$$

□

Teorema 2.2. *Todo espaço normado $(X, \|\cdot\|)$ é um espaço métrico (X, d) quando a métrica é definida por $d(x,y) = \|x-y\|$, $\forall x, y \in X$.*

Demonstração. De fato para todo $x, y, z \in X$ tem-se:
(M1) d é real, finito e não-negativo.

Segue da definição de espaço normado

$$\begin{aligned} \|\cdot\| : X &\longrightarrow \mathbb{R} \\ x &\longmapsto \|x\| \end{aligned}$$

$$(N4)\ \|x\| > 0.$$

(M2) $d(x,y) = 0$ se, e somente se, $x = y$.

$$d(x,y) = \|x-y\| = 0 \Longleftrightarrow x - y = 0 \Longleftrightarrow x = y.$$

(M3) $d(x,y) = d(y,x)$ (simetria).

$$d(x,y) = \|x-y\| = \|-(y-x)\| = |-1|\|y-x\| = \|y-x\| = d(y,x).$$

(M4) $d(x,y) \leqslant d(x,z) + d(z,y)$ (desigualdade triangular).

$$d(x,y) = \|x-y\| = \|x-z+z-y\| \leqslant \|x-z\| + \|z-y\| = d(x,z) + d(z,y).$$

□

Lema 2.2. *(invariância translacional) Uma métrica d induzida por uma norma em um espaço normado X satisfaz*

a) $d(x+a, y+a) = d(x,y)$.

b) $d(\alpha x, \alpha y) = |\alpha| d(x,y)$.

Demonstração. Temos

$$d(x+a, y+a) = \|x+a-(y+a)\| = \|x-y\| = d(x,y)$$

e

$$d(\alpha x, \alpha y) = \|\alpha x - \alpha y\| = |\alpha|\|x-y\| = |\alpha| d(x,y).$$

□

Lema 2.3. *(Combinação linear) Seja $\{x_1, \ldots, x_n\}$ um conjunto de vetores linearmente independentes em um espaço normado X (de qualquer dimensão). Então existe um número $c > 0$ tal que para toda escolha de escalares $\alpha_1, \ldots, \alpha_n$ temos*

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geqslant c(|\alpha_1| + \cdots + |\alpha_n|). \tag{2.1}$$

Demonstração. Seja $s = |\alpha_1| + \cdots + |\alpha_n| = \sum_{j=1}^{n} |\alpha_j|$. Note que se $s = 0$, então todos os α_j são nulos e (2.1) é satisfeita para qualquer c. Suponha agora $s > 0$. Dividindo (2.1) por s segue que

$$\|\beta_1 x_1 + \cdots + \beta_n x_n\| \geqslant c \tag{2.2}$$

com $\sum_{j=1}^{n} |\beta_j| = 1$, onde $\beta_1 = \dfrac{\alpha_1}{|\alpha_1| + \cdots + |\alpha_n|}, \ldots, \beta_n = \dfrac{\alpha_n}{|\alpha_1| + \cdots + |\alpha_n|}$.

Então provar (2.1) equivale a provar que existe $c > 0$ tal que (2.2) seja satisfeita para quaisquer escalares $\beta_1, \ldots, \beta_n \in \mathbb{K}$ com $\sum_{j=1}^{n} |\beta_j| = 1$.

Suponha que (2.2) seja falsa. Ou seja, para todo $c > 0$ existem escalares $\beta_1^c, \dots, \beta_n^c$ com $\sum_{j=1}^{n} \beta_j^{(c)} = 1$ tais que $\|y_c\| < c$ onde $y_c = \beta_1^c x_1, \dots, \beta_n^c x_n$. Conforme c varia obtêm-se uma sequência $(y_m)_{m\in\mathbb{N}}(c = \frac{1}{m}, m \in \mathbb{N})$ satisfazendo

$$y_m = \beta_1^{(m)} x_1 + \cdots + \beta_n^{(m)} x_n \text{ e } \sum_{j=1}^{n} |\beta_j^{(m)}| = 1$$

ou seja,

$$\begin{aligned} y_1 &= \beta_1^{(1)} x_1 + \cdots + \beta_n^{(1)} x_n \\ y_2 &= \beta_1^{(2)} x_1 + \cdots + \beta_n^{(2)} x_n \\ &\vdots \\ y_k &= \beta_1^{(k)} x_1 + \cdots + \beta_n^{(k)} x_n \\ &\vdots \\ y_m &= \beta_1^{(m)} x_1 + \cdots + \beta_n^{(m)} x_n \end{aligned}$$

com $\sum_{j=1}^{n} |\beta_j^{(m)}| = 1$, de forma que $\|y_m\| \to 0$ quando $m \to \infty$.

Como $\sum_{j=1}^{n} |\beta_j^{(m)}| = 1$ então $|\beta_j^{(m)}| \leqslant 1$. Assim para cada j fixo a sequência $(\beta_j^{(m)}) = (\beta_j^{(1)}, \beta_j^{(2)}, \dots)$ é limitada. Pelo teorema de Bolzano-Weierstrass, cada uma das sequências

$$\begin{aligned} (\beta_1^{(m)}) &= (\beta_1^1, \beta_1^2, \dots) \\ (\beta_2^{(m)}) &= (\beta_2^1, \beta_2^2, \dots) \\ &\vdots \\ (\beta_n^{(m)}) &= (\beta_n^1, \beta_n^2, \dots) \end{aligned}$$

possui subsequência convergente que será denotada por $(\lambda_j^{(m)})_{n\in\mathbb{N}}$. Considere β_1 o limite da sequência $(\beta_1^{(m)})_{m\in\mathbb{N}}$ e $(y_1^{(m)})_{m\in\mathbb{N}}$ a correspondente subsequência de $(y_m)_{m\in\mathbb{N}}$.

Da mesma forma, sejam β_2 o limite da sequência $(\beta_2^{(m)})_{m\in\mathbb{N}}$ e $(y_2^{(m)})_{m\in\mathbb{N}}$ a correspondente subsequência de $(y_1^{(m)})_{m\in\mathbb{N}}$.

Prosseguindo desta maneira para $j = 3, \dots, n$ obtêm-se uma subsequência $(y_n^{(m)})$ de (y_m), onde $(y_n^{(m)}) = (y_n^{(1)}, y_n^{(2)}, \dots)$ cujos os termos são da forma

$$\begin{aligned} y_n^{(1)} &= \lambda_1^{(1)} x_1 + \lambda_2^{(1)} x_2 + \cdots + \lambda_n^{(1)} x_n \\ y_n^{(2)} &= \lambda_1^{(2)} x_1 + \lambda_2^{(2)} x_2 + \cdots + \lambda_n^{(2)} x_n \\ &\vdots \\ y_n^{(m)} &= \lambda_1^{(m)} x_1 + \lambda_2^{(m)} x_2 + \cdots + \lambda_n^{(m)} x_n \end{aligned}$$

ou seja,

$$y_n^{(m)} = \sum_{j=1}^{n} \lambda_j^{(m)} x_j \quad \text{e} \quad \sum_{j=1}^{n} |\lambda_j^{(m)}| = 1$$

com os escalares $\lambda_j^{(m)}$ satisfazendo $\lambda_1^{(m)} \to \beta_1, \lambda_2^{(m)} \to \beta_2, \ldots, \lambda_n^{(m)} \to \beta_n$, ou seja, $\lambda_j^{(m)} \to \beta_j$ quando $m \to \infty$.

Assim quando $m \to \infty, y_n^{(m)} \to y = \sum_{j=1}^{n} \beta_j x_j$, onde $\sum_{j=1}^{n} |\beta_j| = 1$. Assim, nem todos os β_j podem ser nulos e como $\{x_1, \ldots, x_n\}$ é um conjunto linearmente independente segue que $y \neq 0$.

Por outro lado, pela continuidade da norma (2.1), segue que

$$y_n^{(m)} \to y \Rightarrow \|y_n^{(m)}\| \to \|y\|, \text{ quando } m \to \infty.$$

Como $\|y_m\| \to 0$ e $y_n^{(m)}$ é uma subsequência de $(y_m)_{m\in\mathbb{N}}$ segue que $y_n^{(m)} \to 0$ quando $m \to \infty$. Pela definição de norma, $\|y\| = 0 \Leftrightarrow y = 0$. Contradição, pois $y \neq 0$. □

Teorema 2.3. *Todo subespaço dimensionalmente finito Y de um espaço normado X é completo. Em particular, todo espaço normado finito é completo.*

Demonstração. Consideremos uma sequência de Cauchy arbitrária $(y_m)_{m\in\mathbb{N}}$ em Y e mostremos que ela é convergente em Y; o limite será denotado por y. Sejam $\dim Y = n$ e $\{e_1, \ldots, e_n\}$ uma base qualquer de Y. Então cada y_m tem uma representação única dada por

$$y_m = \alpha_1^{(m)} e_1 + \cdots + \alpha_n^{(m)} e_n.$$

Uma vez que (y_m) é uma sequência de Cauchy, para cada $\varepsilon > 0$ existe um N tal que $\|y_m - y_r\| < \varepsilon$ quando $m, r > N$, $m \in \mathbb{N}$. Então em conjunto com o lema anterior para algum $c > 0$ temos

$$\varepsilon > \|y_m - y_r\| = \left\| \sum_{j=1}^{n} (\alpha_j^{(m)} - \alpha_j^{(r)}) e_j \right\| \geqslant c \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j^{(r)}|,$$

onde $m, r > N$. Dividindo por $c > 0$ temos

$$|\alpha_j^{(m)} - \alpha_j^{(r)}| \leqslant \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j^{(r)}| < \frac{\varepsilon}{c} \qquad (m, r > N).$$

Isto mostra que cada uma das sequências

$$(\alpha_j^{(m)}) = (\alpha_j^{(1)}, \alpha_j^{(2)}, \ldots), \qquad j = 1, 2, \ldots, n.$$

é Cauchy em $\mathbb{R}$ ou $\mathbb{C}$ e portanto converge. Seja α_j o limite. Usando estes n limites $\alpha_1, \ldots, \alpha_n$, definimos

$$y = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

Claramente, $y \in Y$. Além disso,

$$\|y_m - y\| = \left\| \sum_{j=1}^{n} (\alpha_j^{(m)} - \alpha_j) e_j \right\| \leqslant \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j| \|e_j\|.$$

Do lado direito, $\alpha_j^{(m)} \longrightarrow \alpha_j$. Conseguentemente $\|y_m - y\| \longrightarrow 0$, isto é, $y_m \longrightarrow y$. Isto mostra que (y_m) é convergente em Y. Uma vez que $(y_m)_{m\in\mathbb{N}}$ é uma sequência de Cauchy arbitrária em Y, isto prova que Y é completo. $\square$

Definição 2.2. *(Normas equivalentes) Uma norma $\|.\|$ em um espaço vetorial X é equivalente a uma norma $\|.\|_0$ em X se existirem números positivos a e b tais que para todo $x \in X$ temos*

$$a\|x\|_0 \leqslant \|x\| \leqslant b\|x\|_0. \tag{2.3}$$

Teorema 2.4. *Em um espaço vetorial de dimensão finita X, qualquer norma $\|.\|$ é equivalente a qualquer outra norma $\|.\|_0$.*

Demonstração. Sejam $\dim X = n$ e $e_1, \ldots, e_n$ uma base qualquer para X. Então todo $x \in X$ tem uma representação única

$$x = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

Pelo lema (2.3) existe uma constante positiva c tal que

$$\|x\| \geqslant c(|\alpha_1| + \cdots + |\alpha_n|).$$

Por outro lado, a desigualdade triangular fornece

$$\|x\|_0 \leqslant \sum_{j=1}^{n} |\alpha_j| \|e_j\|_0 \leqslant k \sum_{j=1}^{n} |\alpha_j|, \qquad k = \max_j \|e_j\|_0.$$

O que nos leva a

$$a\|x\|_0 \leqslant \|x\|, \qquad \text{onde } a = \frac{c}{k} > 0.$$

A outra desigualdade em (2.3) é obtida pela troca entre $\|\cdot\|$ e $\|\cdot\|_0$ no argumento anterior. $\square$

Este teorema tem uma importância prática considerável. Por exemplo, ele implica que a convergência ou divergência de uma sequência em um espaço vetorial finito não depende de uma escolha particular da norma neste espaço. É importante ressaltar que este teorema não vale para espaços infinitos.

Seguem alguns exemplos de espaços normados.

2.1.1 ESPAÇO EUCLIDIANO $\mathbb{R}^n$

O espaço $\mathbb{R}^n$ com a norma definida por

$$\|x\| = \left(\sum_{j=1}^{n} x_j^2\right)^{1/2}$$

é um espaço normado.

De fato,

(N1) $\|x\| \geqslant 0$.

$$\|x\| = \left(\sum_{j=1}^{n} x_j^2\right)^{1/2} = \sqrt{x_1^2 + \cdots + x_n^2}, \text{ como } x_1^2 + \cdots + x_n^2 \geqslant 0 \Rightarrow \|x\| \geqslant 0.$$

(N2) $\|x\| = 0 \Leftrightarrow x = 0$.

$$\|x\| = \left(\sum_{j=1}^{n} x_j^2\right)^{1/2} = \sqrt{x_1^2 + \cdots + x_n^2} = 0 \Leftrightarrow x_1^2 + \cdots + x_n^2 = 0 \Leftrightarrow x_1 = \cdots = x_n = 0 \Leftrightarrow x = 0.$$

(N3) $\|\alpha x\| = |\alpha| \|x\|$.

$$\|\alpha x\| = \left(\sum_{j=1}^{n} (\alpha x_j)^2\right)^{1/2} = \left(\alpha^2 \sum_{j=1}^{n} x_j^2\right)^{1/2} = |\alpha| \left(\sum_{j=1}^{n} x_j^2\right)^{1/2} = |\alpha| \|x\|.$$

(N4) $\|x + y\| \leqslant \|x\| + \|y\|$.

Da desigualdade de Minkowski temos,

$$\|x + y\| = \left(\sum_{j=1}^{n} (x_j + y_j)^2\right)^{1/2} \leqslant \left(\sum_{j=1}^{n} x_j^2\right)^{1/2} + \left(\sum_{j=1}^{n} y_j^2\right)^{1/2} = \|x\| + \|y\|.$$

2.1.2 ESPAÇO DE FUNÇÕES $C[a, b]$

O espaço de funções $C[a, b]$ com a norma definida por

$$\|x\| = \max_{t \in J} |x(t)|, \qquad J = [a, b]$$

é um espaço normado.

Verifiquemos que,

(N1) $\|x\| \geqslant 0$

$$\|x\| = \max_{t \in J} |x(t)|, \text{ como } |x(t)| \geqslant 0 \Rightarrow \|x\| \geqslant 0.$$

(N2) $\|x\| = 0 \Leftrightarrow x = 0$

$$\|x\| = \max_{t \in J} |x(t)| = 0 \Leftrightarrow |x(t)| = 0 \Leftrightarrow x(t) = 0,\ \forall\ t \in J \Leftrightarrow x = 0.$$

(N3) $\|\alpha x\| = |\alpha| \|x\|$

$$\|\alpha x\| = \max_{t \in J} |\alpha x(t)| = |\alpha| \max_{t \in J} |x(t)| = |\alpha| \|x\|.$$

(N4) $\|x + y\| \leqslant \|x\| + \|y\|$

$$\|x + y\| = \max_{t \in J} |x(t) + y(t)| \leqslant \max_{t \in J} |x(t)| + \max_{t \in J} |y(t)| = \|x\| + \|y\|.$$

2.1.3 ESPAÇO DE SEQUÊNCIAS ℓ^p

O espaço de sequências ℓ^p com a norma definida por

$$\|x\| = \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p}$$

é um espaço normado. Essa norma induz a métrica

$$d(x, y) = \|x - y\| = \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p \right)^{1/p}.$$

Verifiquemos que,

(N1) $\|x\| \geqslant 0$

$$\|x\| = \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p}, \text{ como } \ |\xi_1|^p + \cdots + |\xi_n|^p + \cdots \geqslant 0 \Rightarrow \|x\| \geqslant 0.$$

(N2) $\|x\| = 0 \Leftrightarrow x = 0$

$$\|x\| = \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} = 0 \Leftrightarrow\ |\xi_1|^p + \cdots + |\xi_n|^p + \cdots = 0 \Leftrightarrow\ \xi_1 = \cdots = \xi_n = \cdots = 0 \Leftrightarrow\ x = 0.$$

(N3) $\|\alpha x\| = |\alpha| \|x\|$

$$\|\alpha x\| = \left(\sum_{j=1}^{\infty} |\alpha \xi_j|^p \right)^{1/p} = \left(|\alpha|^p \sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} = |\alpha| \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} = |\alpha| \|x\|.$$

(N4) $\|x + y\| \leqslant \|x\| + \|y\|$

Utilizando a desigualdade de Minkowski

$$\|x + y\| = \left(\sum_{j=1}^{\infty} |\xi_j + \eta_j|^p \right)^{1/p} \leqslant \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} + \left(\sum_{j=1}^{\infty} |\eta_j|^p \right)^{1/p} = \|x\| + \|y\|.$$

2.2 OPERADOR LINEAR

Em Análise Funcional são considerados espaços mais gerais, tais como espaços métricos e espaços normados, e mapeamentos destes espaços.

Definição 2.3. *(Operador linear) Um operador linear T é uma transformação linear tal que*

(i) o domínio $\mathcal{D}(T)$ é um espaço vetorial e o conjunto imagem $\mathcal{R}(T)$ encontra-se em um espaço vetorial sobre o mesmo corpo,

(ii) para todo $x, y \in \mathcal{D}(T)$ e escalar α,

$$\begin{aligned} T(x+y) &= Tx + Ty. \\ T(\alpha x) &= \alpha T x. \end{aligned} \tag{2.4}$$

Claramente (2.4) é equivalente a

$$T(\alpha x + \beta y) = \alpha T x + \beta T y. \tag{2.5}$$

Tomando $\alpha = 0$ em (2.4) obtemos a seguinte fórmula

$$T0 = 0. \tag{2.6}$$

A equação (2.4) expressa o fato que um operador linear T é um homomorfismo de um espaço vetorial (do seu domínio) em outro espaço vetorial, isto é, T preserva as duas operações do espaço vetorial, no seguinte sentido. Em (2.4) a esquerda primeiro aplicamos uma operação do espaço vetorial (adição ou multiplicação por escalar) e então mapeamos o vetor resultante em Y, ao passo que a direita primeiro mapeamos x e y em Y e então realizamos a operação do espaço vetorial em Y, cujo resultado é o mesmo. Esta propriedade torna os operadores lineares importantes. Por sua vez, espaços vetoriais são importantes em Análise Funcional principalmente pelos operadores lineares definidos neles.

São exemplos de operadores lineares:

Operador identidade. O operador identidade $I : X \longrightarrow X$ é definido por $Ix = x$ para todo $x \in X$.

Operador nulo. O operador nulo $0 : X \longrightarrow Y$ é definido por $0x = 0$ para todo $x \in X$.

Diferenciação. Seja X o espaço vetorial de todos os polinômios em $[a, b]$. Podemos definir um operador linear T em X por

$$Tx(t) = x'(t)$$

para todo $x \in X$. Este operador T mapeia X nele mesmo.

Integração. Um operador linear T de $C[a, b]$ nele mesmo pode ser definido por

$$Tx(t) = \int_a^t x(\tau)\, d\tau, \qquad t \in [a, b].$$

Teorema 2.5. *(Imagem e espaço nulo) Seja T um operador linear. Então:*

a) *O conjunto imagem $\mathcal{R}(T)$ é um espaço vetorial.*

b) *Se $\dim \mathcal{D}(T) = n < \infty$, então $\dim \mathcal{R}(T) \leqslant n$.*

c) *O espaço nulo $\mathcal{N}(T)$ é um espaço vetorial.*

Demonstração. **(a)** Tomamos qualquer $y_1, y_2 \in \mathcal{R}(T)$ e mostramos que $\alpha y_1 + \beta y_2 \in \mathcal{R}(T)$ para quaisquer escalares α, β. Uma vez que $y_1, y_2 \in \mathcal{R}(T)$, temos que $y_1 = Tx_1, y_2 = Tx_2$ para algum $x_1, x_2 \in \mathcal{D}(T)$, e $\alpha x_1 + \beta x_2 \in \mathcal{D}(T)$ porque $\mathcal{D}(T)$ é um espaço vetorial. A linearidade de T produz

$$T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 = \alpha y_1 + \beta y_2.$$

Consequentemente $\alpha y_1 + \beta y_2 \in \mathcal{R}(T)$. Uma vez que $y_1, y_2 \in \mathcal{R}(T)$ são arbitrários assim como os escalares, isto prova que $\mathcal{R}(T)$ é um espaço vetorial.
(b) Escolhemos $n+1$ elementos $y_1, \ldots, y_{n+1}$ de $\mathcal{R}(T)$ de uma forma arbitrária. Então temos $y_1 = Tx_1, \cdots, y_{n+1} = Tx_{n+1}$ para algum $x_1, \ldots, x_{n+1}$ em $\mathcal{D}(T)$. Uma vez que $\dim \mathcal{D}(T) = n$, este conjunto $\{x_1, \ldots x_{n+1}\}$ deve ser linearmente dependente. Consequentemente

$$\alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1} = 0$$

para alguns escalares $\alpha_1, \ldots, \alpha_{n+1}$, nem todos iguais a zero. Uma vez que T é linear e $T0 = 0$, a aplicação de T em ambos os lados fornece

$$T(\alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1}) = \alpha_1 y_1 + \cdots + \alpha_{n+1} y_{n+1} = 0.$$

Isto mostra que $\{y_1, \ldots y_{n+1}\}$ é um conjunto linearmente dependente porque os $\alpha_j' s$ não são todos iguais a zero. Lembrando que este subconjunto de $\mathcal{R}(T)$ foi escolhido de maneira arbitrária, concluímos que $\mathcal{R}(T)$ não tem subconjuntos linearmente independentes de $n+1$ ou mais elementos. Por definição isto significa que $\dim \mathcal{R}(T) \leqslant n$.
(c) Tomamos quaisquer $x_1, x_2 \in \mathcal{N}(T)$. Então $Tx_1 = Tx_2 = 0$. Uma vez que T é linear, para quaisquer escalares α, β temos

$$T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 = 0.$$

Isto mostra que $\alpha x_1 + \beta x_2 \in \mathcal{N}(T)$. Consequentemente $\mathcal{N}(T)$ é um espaço vetorial. □

É interessante notar que uma consequência imediata da parte (b) é a de que operadores lineares preservam a dependência linear.

Definição 2.4. *Uma aplicação $T : \mathcal{D}(T) \longrightarrow Y$ é dito ser injetivo se diferentes pontos no domínio tem diferentes imagens, isto é, se para qualquer $x_1, x_2 \in \mathcal{D}(T)$,*

$$x_1 \neq x_2 \Rightarrow Tx_1 \neq Tx_2. \tag{2.7}$$

Equivalentemente

$$Tx_1 = Tx_2 \Rightarrow x_1 = x_2. \tag{2.8}$$

Neste caso existe o mapeamento

$$\begin{aligned} T^{-1} :&\mathcal{R}(T) \longrightarrow \mathcal{D}(T) \\ &y_0 \longmapsto x_0 \qquad\qquad (y_0 = Tx_0) \end{aligned} \tag{2.9}$$

que mapeia todo $y_0 \in \mathcal{R}(T)$ *em* $x_0 \in \mathcal{D}(T)$ *para os quais* $Tx_0 = y_0$. *O mapeamento* T^{-1} *é chamado de o inverso de* T.

De (2.9) claramente temos

$$\begin{aligned} T^{-1}Tx =&x \quad \text{para todo} \quad x \in \mathcal{D}(T) \\ TT^{-1}y =&y \quad \text{para todo} \quad y \in \mathcal{R}(T). \end{aligned}$$

O inverso de um operador linear existe se, e somente se, o espaço nulo do operador consiste apenas do vetor zero.

Teorema 2.6. *(Operador inverso) Sejam* X, Y *espaços vetoriais, ambos reais ou complexos. Seja* $T : \mathcal{D}(T) \longrightarrow Y$ *um operador linear com domínio* $\mathcal{D}(T) \subset X$ *e imagem* $\mathcal{R}(T) \subset Y$. *Então:*

a) *O inverso* $T^{-1} : \mathcal{R}(T) \longrightarrow \mathcal{D}(T)$ *existe se, e somente se,*

$$Tx = 0 \Rightarrow x = 0.$$

b) *Se* T^{-1} *existe, ele é um operador linear.*

c) *Se* $\dim \mathcal{D}(T) = n < \infty$ *e* T^{-1} *existe, então* $\dim \mathcal{R}(T) = \dim \mathcal{D}(T)$.

Demonstração.

(a) suponha que $Tx = 0$ implique $x = 0$. Seja $Tx_1 = Tx_2$. Uma vez que T é linear,

$$T(x_1 - x_2) = Tx_1 - Tx_2 = 0$$

de modo que pela hipótese $x_1 - x_2 = 0$. Consequentemente $Tx_1 - Tx_2$ implica $x_1 = x_2$ e T^{-1} existe por (2.8). Reciprocamente, se T^{-1} existe, então (2.8) vale. De (2.8) com $x_2 = 0$ e (2.6) obtemos

$$Tx_1 = T0 = 0 \Rightarrow x_1 = 0.$$

(b) Suponhamos que T^{-1} existe e vamos mostrar que T^{-1} é linear. O domínio de T^{-1} é $\mathcal{R}(T)$ e pelo teorema 2.5(a) é um espaço vetorial. Considerando qualquer $x_1, x_2 \in \mathcal{D}(T)$ e suas imagens

$$y_1 = Tx_1 \qquad \text{e} \qquad y_2 = Tx_2.$$

Então

$$x_1 = T^{-1}y_1 \qquad \text{e} \qquad x_2 = T^{-1}y_2.$$

T é linear, de tal modo que para qualquer escalar α e β temos

$$\alpha y_1 + \beta y_2 = \alpha T x_1 + \beta T x_2 = T(\alpha x_1 + \beta x_2).$$

Uma vez que $x_j = T^{-1}y_j$, isto implica que

$$T^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha T^{-1}y_1 + \beta T^{-1}y_2$$

o que prova que T^{-1} é linear.

(c) Temos $\dim \mathcal{R}(T) \leqslant \dim \mathcal{D}(T)$ pelo teorema 2.3(b), e $\dim \mathcal{D}(T) \leqslant \dim \mathcal{R}(T)$ pelo mesmo teorema aplicado a T^{-1}. $\square$

Lema 2.4. *(Inversa do produto) Sejam $T : X \longrightarrow Y$ e $S : Y \longrightarrow Z$ operadores lineares bijetores, onde X, Y, Z são espaços vetoriais. Então o operador inverso $(ST)^{-1} : Z \longrightarrow X$ do produto ST existe, e*

$$(ST)^{-1} = T^{-1}S^{-1}. \tag{2.10}$$

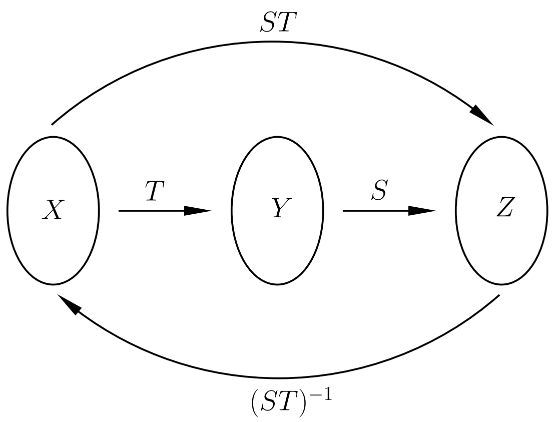

Figura 1 – Notações no lema 2.4.

Definição 2.5. *(Operador linear limitado) Sejam X e Y espaços normados e $T : \mathcal{D}(T) \longrightarrow Y$ um operador linear, onde $\mathcal{D}(T) \subset X$. O operador T é limitado se existe um número real c tal que para todo $x \in \mathcal{D}(T)$,*

$$\|Tx\| \leqslant c\|x\| \tag{2.11}$$

Em (2.11) a norma à esquerda é a norma em Y, e a norma à direita é a norma em X. A fórmula (2.11) mostra que operadores lineares limitados mapeiam conjuntos limitados em $\mathcal{D}(T)$ em conjuntos limitados em Y. Devemos notar que o uso da palavra "limitado" aqui é diferente do utilizado no Cálculo, onde a função limitada é aquela cuja imagem é um conjunto limitado.

Vamos definir agora qual o menor c possível tal que (2.11) ainda seja válido para todo $x \in \mathcal{D}(T)$ diferente de zero. A divisão

$$\frac{\|Tx\|}{\|x\|} \leqslant c, \qquad x \neq 0$$

mostra que c deve ser pelo menos tão grande quanto o supremo da expressão a esquerda tomado em $\mathcal{D}(T) - \{0\}$. Esta quantidade é denotada por $\|T\|$. Assim

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}. \tag{2.12}$$

$\|T\|$ é chamado de norma do operador T. Se $\mathcal{D}(T) = \{0\}$, definimos $\|T\| = 0$. Neste caso, $T = 0$ uma vez que $T0 = 0$.

Temos que (2.11) com $c = \|T\|$ é

$$\|Tx\| \leqslant \|T\|\|x\|. \tag{2.13}$$

Lema 2.5. *(Norma) Seja T um operador linear limitado, como definido anteriormente. Então*

a) *Uma fórmula alternativa para a norma de T é*

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ \|x\|=1}} \|Tx\|. \tag{2.14}$$

b) *A norma definida por (2.12) satisfaz (N1) a (N4).*

Demonstração.

(a) Escrevemos $\|x\| = a$ e tomamos $y = (1/a)x$, onde $x \neq 0$. Então $\|y\| = \|x\|/a = 1$, uma vez que T é linear, (2.12) será dada por

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{1}{a}\|Tx\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \left\|T\left(\frac{1}{a}x\right)\right\| = \sup_{\substack{y \in \mathcal{D}(T) \\ \|y\|=1}} \|Ty\|.$$

Escrevendo x para y a direita, temos (2.14).

(b) $(N1)$ é óbvio, então $\|0\| = 0$. De $\|T\| = 0$ temos $Tx = 0$ para todo $x \in \mathcal{D}(T)$. Logo $T = 0$. Consequentemente vale $(N2)$. Além disso, $(N3)$ é obtida de

$$\sup_{\|x\|=1} \|\alpha Tx\| = \sup_{\|x\|=1} |\alpha|\|Tx\| = |\alpha| \sup_{\|x\|=1} \|Tx\|$$

onde $x \in \mathcal{D}(T)$. Finalmente, $(N4)$ segue de

$$\sup_{\|x\|=1} \|(T_1 + T_2)x\| = \sup_{\|x\|=1} \|T_1x + T_2x\| \leqslant \sup_{\|x\|=1} \|T_1x\| + \sup_{\|x\|=1} \|T_2x\|$$

aqui, $x \in \mathcal{D}(T)$. □

Vejamos dentre os operadores lineares dados como exemplo anteriormente quais são limitados.

Operador identidade. O operador identidade $I : X \longrightarrow X$ em um espaço normado $X \neq \{0\}$ é limitado e tem norma $\|I\| = 1$.

Operador nulo. O operador nulo $0 : X \longrightarrow Y$ em um espaço normado X é limitado e tem norma $\|0\| = 0$.

Operador diferencial. Seja X o espaço normado de todos os polinômios em $J = [0, 1]$ com norma dada por $\|x\| = \max |x(t)|, t \in J$. Um operador diferencial T é definido em X por

$$Tx(t) = x'(t)$$

Este operador é linear mas não é limitado. De fato, seja $x_n(t) = t^n$, onde $n \in \mathbb{N}$. Então $\|x_n\| = 1$ e

$$Tx_n(t) = x'_n(t) = nt^{n-1}$$

de modo que $\|Tx_n\| = n$ e $\|Tx_n\|/\|x_n\| = n$. Uma vez que $n \in \mathbb{N}$ é arbitrário, isto mostra que não existe um número c tal que $\|Tx_n\|/\|x_n\| \leqslant c$.

Operador integral. Podemos definir um operador integral $T : C[0, 1] \longrightarrow C[0, 1]$ por

$$y = Tx \quad \text{onde} \quad y(t) = \int_0^1 k(t, \tau)x(\tau)d\tau.$$

Aqui k é uma função dada, que é chamada de *núcleo* de T e é assumida como sendo contínua no quadrado fechado $G = J \times J$ no plano $t\tau$, onde $J = [0, 1]$. Este operador é linear.

Para provar que T é limitado, primeiro notamos que a continuidade de k no quadrado fechado implica que k é limitado, $|k(t, \tau)| \leqslant k_0$ para todo $(t, \tau) \in G$, onde k_0 é um número real. Além disso,

$$|x(t)| \leqslant \max_{t \in J} |x(t)| = \|x\|$$

Consequentemente

$$\begin{aligned} \|y\| = \|Tx\| = &\max_{t \in J} \left| \int_0^1 k(t, \tau)x(\tau)d\tau \right| \\ \leqslant &\max_{t \in J} \int_0^1 |k(t, \tau)||x(\tau)|d\tau \\ \leqslant &k_0\|x\|. \end{aligned}$$

O resultado é $\|Tx\| = k_0\|x\|$. Isto é (2.11) com $c = k_0$. Consequentemente T é limitado.

Teorema 2.7. *(Dimensão finita) Se um espaço normado X é dimensionalmente finito, então todo operador em X é limitado.*

Demonstração. Sejam $\dim X = n$ e $\{e_1, \ldots, e_n\}$ uma base de X. Consideremos $x = \sum_{j=1}^{n} \xi_j e_j$ e consideramos qualquer operador linear T em X. Uma vez que T é linear,

$$\|Tx\| = \left\| \sum_{j=1}^{n} \xi_j T e_j \right\| \leqslant \sum_{j=1}^{n} |\xi_j| \|Te_j\| \leqslant \max_k \|Te_j\| \sum_{j=1}^{n} |\xi_j|.$$

Para a última soma aplicamos o lema 2.3 com $\alpha_j = \xi_j$ e $x_j = e_j$. Então,

$$\sum |\xi_j| \leqslant \frac{1}{c} \left\| \sum \xi_j e_j \right\| = \frac{1}{c} \|x\|.$$

Juntamente com

$$\|Tx\| \leqslant \gamma \|x\| \quad \text{onde} \quad \gamma = \frac{1}{c} \max_k \|Te_j\|.$$

Desta equação e de (2.11) podemos verificar que T é limitado. □

Operadores são mapeamentos, logo a definição de continuidade se aplica a eles. É um fato fundamental que para operadores lineares, continuidade e limitação tornam-se conceitos equivalentes.

Seja $T : \mathcal{D}(T) \longrightarrow Y$ qualquer operador, não necessariamente linear, onde $\mathcal{D}(T) \subset X$ e X e Y são espaços normados. Por definição o operador T é contínuo em um $x_0 \in \mathcal{D}(T)$ se para todo $\varepsilon > 0$ existir um $\delta > 0$ tal que

$$\|Tx - Tx_0\| < \varepsilon \text{ para todo } x \in \mathcal{D}(T) \text{ satisfazendo } \|x - x_0\| < \delta.$$

T é contínuo se é contínuo em todo $x \in \mathcal{D}(T)$.

Teorema 2.8. *(Continuidade e limitação) Seja $T : \mathcal{D}(T) \longrightarrow Y$ um operador linear, onde $\mathcal{D}(T) \subset X$ e X, Y são espaços normados. Então*

(a) T é contínuo se, e somente se é limitado.

(b) Se T é contínuo em um único ponto, ele é contínuo.

Demonstração.

(a) Para $T = 0$ a afirmação é trivial. Seja $T \neq 0$. Então $\|T\| \neq 0$. Supomos que T seja limitado e consideramos qualquer $x_0 \in \mathcal{D}(T)$. Dado qualquer $\varepsilon > 0$. Então como T é linear, para todo $x \in \mathcal{D}(T)$ tal que

$$\|x - x_0\| < \delta \quad \text{onde} \quad \delta = \frac{\varepsilon}{\|T\|}$$

obtemos

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

Uma vez que $x_0 \in \mathcal{D}(T)$ é arbitrário, isto mostra que T é contínuo.

Reciprocamente, assumindo que T é contínuo em um $x_0 \in \mathcal{D}(T)$ arbitrário. Então, dado qualquer $\varepsilon > 0$, existe um $\delta > 0$ tal que

$$\|Tx - Tx_0\| < \varepsilon \text{ pra todo } x \in \mathcal{D}(T) \text{ satisfazendo } \|x - x_0\| \leqslant \delta. \tag{2.15}$$

Tomando qualquer $y \neq 0$ em $\mathcal{D}(T)$ e definindo

$$x = x_0 + \frac{\delta}{\|y\|}y. \text{ Então } x - x_0 = \frac{\delta}{\|y\|}y.$$

Consequentemente $\|x - x_0\| = \delta$, logo podemos usar (2.15). Uma vez que T é linear, temos

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

e (2.15) implica

$$\frac{\delta}{\|y\|}\|Ty\| \leqslant \varepsilon. \text{ Assim } \|Ty\| \leqslant \frac{\varepsilon}{\delta}\|y\|.$$

Então podemos escrever $\|Ty\| \leqslant c\|y\|$, onde $c = \varepsilon/\delta$, o que mostra que T é limitado.

(b) Continuidade de T no ponto a implica limitação de T pela segunda parte da demonstração de (a), que por sua vez implica na continuidade de T por (a). □

Teorema 2.9. *$B(X, Y)$ é um espaço vetorial com uma norma definida por*

$$\|T\| := \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|=1} \|Tx\|_Y.$$

$B(X, Y)$ é completo quando Y é completo. Em particular, X' é um espaço de Banach com norma

$$\|\phi\| = \sup_{x \neq 0} \frac{|\phi x|}{\|x\|}.$$

2.3 FUNCIONAL LINEAR

Um funcional é um operador cuja imagem está na reta real $\mathbb{R}$ ou no campo complexo $\mathbb{C}$. Denotaremos funcionais por letras minúsculas $f, g, h, \ldots$, o domínio de f por $\mathcal{D}(f)$, a imagem por $\mathcal{R}(f)$ e o valor de f em um $x \in \mathcal{D}(f)$ por $f(x)$. Como funcionais são operadores as definições prévias se aplicam.

Definição 2.6. *(Funcional linear) Um funcional linear f é um operador linear com domínio em um espaço vetorial X e imagem no corpo dos escalares $\mathbb{K}$ de X, assim*

$$f = \mathcal{D}(f) \longrightarrow \mathbb{K},$$

onde $\mathbb{K} = \mathbb{R}$ se X é real e $\mathbb{K} = \mathbb{C}$ se X é complexo.

Definição 2.7. *(Funcional linear limitado) Um funcional linear limitado f é um operador linear limitado com imagem no corpo dos escalares do espaço normado X em que o domínio $\mathcal{D}(f)$ se encontra. Assim existe um número real c tal que para todo $x \in \mathcal{D}(f)$,*

$$|f(x)| \leqslant c\|x\|.$$

Além disso, a norma de f é

$$\|f\| = \sup_{\substack{x \in \mathcal{D}(f) \\ x \neq 0}} \frac{|f(x)|}{\|x\|} \tag{2.16}$$

ou

$$\|f\| = \sup_{\substack{x \in \mathcal{D}(f) \\ \|x\|=1}} |f(x)|. \tag{2.17}$$

A fórmula (2.11) agora implica em

$$|f(x)| = \|f\|\|x\|. \tag{2.18}$$

Um caso especial do teorema (2.8) é

Teorema 2.10. *(Continuidade e limitação) Um funcional linear f com domínio $\mathcal{D}(f)$ em um espaço normado é contínuo se, e somente se, f é limitado.*

Vejamos alguns exemplos de funcionais.

Norma. A norma $\|\cdot\| : X \longrightarrow \mathbb{R}$ em um espaço normado $(X, \|\cdot\|)$ é um funcional em X que não é linear.

Produto escalar. O produto escalar com um ponto fixo define um funcional $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ por meio de

$$f(x) = x \cdot a = \xi_1\alpha_1 + \xi_2\alpha_2 + \xi_3\alpha_3,$$

onde $a = (\alpha_j) \in \mathbb{R}^3$ é fixo.

f é linear e limitado. De fato,

$$|f(x)| = |x \cdot a| \leqslant \|x\|\|a\|,$$

de modo que $\|f\| \leqslant \|a\|$ segue de (2.17) se nós tomarmos o supremo sobre todos x de norma um. Por outro lado, tomando $x = a$ e usando (2.18) obtemos

$$\|f\| \geqslant \frac{|f(a)|}{\|a\|} = \frac{\|a\|^2}{\|a\|} = \|a\|.$$

Consequentemente a norma de f é $\|f\| = \|a\|$.

Integral definida. A integral definida é um número se consideramos uma única função, como fazemos em cálculo a maior parte do tempo. Contudo, a situação muda completamente se considerarmos a integral para todas as funções em um certo espaço de função. Então a integral torna-se um funcional naquele espaço. Como um espaço vamos escolher $C[a, b]$. Então f é definido por

$$f(x) = \int_a^b x(t)dt, \qquad x \in C[a, b]$$

f é linear. Provaremos que f é limitado e tem norma $\|f\| = b - a$.

De fato, escrevendo $J = [a, b]$ e lembrando a norma de $C[a, b]$, obtemos

$$|f(x)| = \left|\int_a^b x(t)dt\right| \leqslant (b - a)\max_{t \in J}|x(t)| = (b - a)\|x\|.$$

Tomando o supremo sobre todos os x de norma 1, obtemos $\|f\| \leqslant b - a$. Para obter $\|f\| \geqslant b - a$, escolhemos em particular $x = x_0 = 1$, notemos que $\|x_0\| = 1$ e usando (2.18)

$$\|f\| \geqslant \frac{|f(x_0)|}{\|x_0\|} = |f(x_0)| = \int_a^b dt = b - a.$$

Espaço $C[a, b]$. Outro funcional importante em $C[a, b]$ é obtido se escolhemos um $t_0 \in J = [a, b]$ fixo e estabelecemos

$$f_1(x) = x(t_0), \qquad x \in C[a, b].$$

f_1 é linear. f_1 é limitado e tem norma $\|f_1\| = 1$. De fato, temos

$$|f_1(x)| = |x(t_0)| \leqslant \|x\|,$$

e isto implica $\|f_1\| \leqslant 1$ por (2.16). Por outro lado, para $x_0 = 1$ temos $\|x_0\| = 1$ e obtemos de (2.18)

$$\|f_1\| \geqslant |f_1(x_0)| = 1.$$

Espaço ℓ^2. Podemos obter um funcional linear f no espaço de Hilbert ℓ^2 escolhendo um $a = (\alpha_j) \in \ell^2$ fixo e estabelecendo

$$f(x) = \sum_{j=1}^{\infty} \xi_j \alpha_j$$

onde $x = (\xi_j) \in \ell^2$. Esta série converge absolutamente e f é limitado, uma vez que a desigualdade de Cauchy-Schwarz fornece (somatória sobre j de 1 a ∞)

$$|f(x)| = \left|\sum \xi_j \alpha_j\right| \leqslant \sum |\xi_j \alpha_j| \leqslant \sqrt{\sum |\xi_j|^2}\sqrt{\sum |\alpha_j|^2} = \|x\|\|a\|.$$

3 APLICAÇÕES

Análise funcional não-linear é o estudo de operadores sem a propriedade da linearidade.

Nós começaremos considerando operadores da equação que tem a forma

$$u = T(u), \qquad u \in K. \tag{3.1}$$

Onde, K é um subconjunto de um espaço de Banach V, e $T : K \longrightarrow V$. As soluções desta equação são chamadas de pontos fixos do operador T, a medida que não são alterados por T. O método mais importante para analisar a solução para tais equações é o do teorema do ponto fixo de Banach.

As definições, teoremas e demonstrações deste capítulo foram retirados principalmente das referências [1], [4] e [5].

3.1 TEOREMA DO PONTO FIXO DE BANACH

Seja V um espaço de Banach com norma $\|\cdot\|_V$, e seja K um subconjunto de V. Considere um operador $T : K \longrightarrow V$ definido em K. Nós estamos interessados na existência de uma solução do problema de ponto fixo (3.1) e na possibilidade de aproximar a solução u pelo seguinte método iterativo. Escolha uma aproximação inicial $u_0 \in K$, e defina uma sequência $\{u_n\}$ pela fórmula de iteração

$$u_{n+1} = T(u_n), \qquad n = 0, 1, \ldots \tag{3.2}$$

Para que isto tenha sentido, nós identificamos outra condição que deve ser imposta sobre T:

$$T(v) \in K, \qquad \forall v \in K. \tag{3.3}$$

O problema de resolver uma equação

$$f(u) = 0 \tag{3.4}$$

para algum operador $f : K \subset V \longrightarrow V$ pode ser reduzido a um problema de ponto fixo equivalente na forma de (3.1) definindo $T(v) = v - c_0 f(v)$ para algum escalar constante $c_0 \neq 0$, ou de forma mais geral, $T(v) = v - F(f(v))$ com um operador $F : V \longrightarrow V$ satisfazendo

$$F(w) = 0 \iff w = 0.$$

Dessa forma qualquer resultado em um problema de ponto fixo (3.1) pode ser reformulado como um resultado para uma equação (3.4). Além disso, o método iterativo (3.2) então provê um possível procedimento de aproximação para resolver a equação (3.4)

Definição 3.1. *Dizemos que um operador $T : K \subset V \longrightarrow V$ é uma contração com constante de contração $\alpha \in [0, 1)$ se*

$$\|T(u) - T(v)\|_V \leqslant \alpha\|u - v\|_V \qquad \forall u, v \in K.$$

O operador T é chamado não-expansivo se

$$\|T(u) - T(v)\|_V \leqslant \|u - v\|_V \qquad \forall u, v \in K,$$

e Lipschitz contínua se existe uma constante $L \geqslant 0$ tal que

$$\|T(u) - T(v)\|_V \leqslant L\|u - v\|_V \qquad \forall u, v \in K.$$

Temos as seguintes implicações:

$$\begin{aligned} \text{contração} &\Longrightarrow \text{não-expansão} \\ &\Longrightarrow \text{Lipschitz contínua} \\ &\Longrightarrow \text{continuidade.} \end{aligned}$$

Teorema 3.1. *(Teorema do ponto fixo de Banach) Assumindo que K é um conjunto fechado não-vazio em um espaço de Banach V, e mais, que $T : K \longrightarrow K$ é um mapeamento de contração com constante de contração α, $0 \leqslant \alpha < 1$. Então vale o seguinte resultado.*

(1) Existência e unicidade: Existe um único $u \in K$ tal que

$$u = T(u).$$

(2) Convergência e estimativa de erro de iteração: Para qualquer $u_0 \in K$, a sequência $\{u_n\} \subset K$ definida por $u_{n+1} = T(u_n)$, $n = 0, 1, \ldots,$ converge para u:

$$\|u_n - u\|_V \longrightarrow 0 \text{ quando } n \longrightarrow \infty.$$

Para o erro, os seguintes limites são válidos:

$$\|u_n - u\|_V \leqslant \frac{\alpha^n}{1-\alpha}\|u_0 - u_1\|_V, \tag{3.5}$$

$$\|u_n - u\|_V \leqslant \frac{\alpha}{1-\alpha}\|u_{n-1} - u_n\|_V, \tag{3.6}$$

$$\|u_n - u\|_V \leqslant \alpha\|u_{n-1} - u\|_V. \tag{3.7}$$

Demonstração. Posto que $T : K \longrightarrow K$, a sequência $\{u_n\}$ é bem definida. Primeiro provaremos que $\{u_n\}$ é uma sequência de Cauchy. Usando o fato do mapeamento T ser uma contração, temos

$$\|u_{n+1} - u_n\|_V \leqslant \alpha\|u_n - u_{n-1}\|_V \leqslant \cdots \leqslant \alpha^n\|u_1 - u_0\|_V.$$

Então para qualquer $m \geqslant n \geqslant 1$,

$$\begin{aligned} \|u_m - u_n\|_V \leqslant & \sum_{j=0}^{m-n-1} \|u_{n+j+1} - u_{n+j}\|_V \\ \leqslant & \sum_{j=0}^{m-n-1} \alpha^{n+j} \|u_1 - u_0\|_V \\ \leqslant & \frac{\alpha^n}{1-\alpha} \|u_1 - u_0\|_V. \end{aligned} \tag{3.8}$$

Uma vez que $\alpha \in [0,1)$, $\|u_m - u_n\|_V \longrightarrow 0$ quando $m, n \longrightarrow \infty$. Deste modo $\{u_n\}$ é uma sequência de Cauchy; e uma vez que K é um conjunto fechado no espaço de Banach V, $\{u_n\}$ tem um limite $u \in K$. Fazendo $n \longrightarrow \infty$ em $u_{n+1} = T(u_n)$ para ver que $u = T(u)$ pela continuidade de T, isto é, u é um ponto fixo de T.

Suponha que $u_1, u_2 \in K$ são ambos pontos fixos de T. Então de $u_1 = T(u_1)$ e $u_2 = T(u_2)$, obtemos

$$u_1 - u_2 = T(u_1) - T(u_2)$$

Consequentemente

$$\|u_1 - u_2\|_V = \|T(u_1) - T(u_2)\|_V \leqslant \alpha \|u_1 - u_2\|_V$$

o que implica $\|u_1 - u_2\|_V = 0$ uma vez que $\alpha \in [0,1)]$. Logo o ponto fixo de uma contração é único.

Agora provemos a estimativa de erro. Fazendo $m \longrightarrow \infty$ em (3.8), nó temos a estimativa (3.5). De

$$\|u_n - u\|_V = \|T(u_{n-1}) - T(u)\|_V \leqslant \alpha \|u_{n-1} - u\|_V$$

nós obtemos a estimativa (3.7). Esta estimativa em conjunto com

$$\|u_{n-1} - u\|_V \leqslant \|u_{n-1} - u_n\|_V + \|u_n - u\|_V$$

implica na estimativa (3.6). □

3.2 APLICAÇÕES DO MÉTODO ITERATIVO

O teorema do ponto fixo de Banach contém a maioria das propriedades desejáveis em um método numérico. Sob as condições indicadas, a sequência de aproximação é bem definida e converge para a solução única do problema. Além disso em (3.7) podemos ver que a taxa de convergência é linear. Temos um erro estimado *a priori* (3.5) que pode ser usado para determinar o número de iterações necessárias para alcançar a acurácia da solução prescrita antes de realizarmos os cálculos e temos uma estimativa de erro *a posteriori* (3.6) que dá o limite do erro uma vez que alguma solução numérica é calculada.

3.2.1 EQUAÇÕES ALGÉBRICAS NÃO LINEARES

Dada uma função real de uma variável real, $f : \mathbb{R} \longrightarrow \mathbb{R}$, estamos interessados em computar suas raízes reais, isto é, estamos interessados em resolver a equação

$$f(x) = 0, \qquad x \in \mathbb{R}. \tag{3.9}$$

Existe uma variedade de maneiras para reformular esta equação como um problema equivalente de ponto fixo da forma

$$x = T(x), \qquad x \in \mathbb{R}. \tag{3.10}$$

Um exemplo é $T(x) \equiv x - c_0 f(x)$ para alguma constante $c_0 \neq 0$. Um exemplo mais sofisticado é $T(x) = x - f(x)/f'(x)$, nesse caso o método iterativo torna-se o método de Newton. Para este último exemplo, geralmente usamos o método de Newton para achar raízes simples de $f(x)$, o que significa que precisamos assumir $f'(x) \neq 0$ quando $f(x) \neq 0$. Restringindo o teorema do ponto fixo de Banach ao problema (3.10), temos o seguinte resultado.

Teorema 3.2. *Seja $-\infty < a < b < \infty$ e $T : [a,b] \longrightarrow [a,b]$ uma função contrátil com constante de contração $\alpha \in [0,1)$. Então valem os seguintes resultados*

(1) Existência e unicidade: Existe uma única solução $\alpha \in [a,b]$ para a equação $x = T(x)$.

(2) Convergência e estimativa de erro de iteração: Para qualquer $x_0 \in [a,b]$, a sequência $\{x_n\} \subset [a,b]$ definida por $x_{n+1} = T(x_n)$, $n = 0, 1, \ldots,$ converge para x:

$$x_n \longrightarrow x \text{ quando } n \longrightarrow \infty.$$

Para o erro, valem os limites

$$\begin{aligned} |x_n - x| &\leqslant \frac{\alpha^n}{1-\alpha}|x_0 - x_1|, \\ |x_n - x| &\leqslant \frac{\alpha}{1-\alpha}|x_{n-1} - x_n|, \\ |x_n - x| &\leqslant \alpha|x_{n-1} - x|. \end{aligned}$$

A contratividade da função T é garantida a partir do pressuposto que

$$\sup_{a \leqslant x \leqslant b} |T'(x)| < 1.$$

De fato, usando o Teorema do valor Médio, podemos ver que T é uma contração com constante de contração $\alpha = \sup\limits_{a \leqslant x \leqslant b} |T'(x)|$.

3.2.2 SISTEMAS ALGÉBRICOS LINEARES

Seja $A \in \mathbb{R}^{m\times m}$ uma matriz, e consideremos o sistema linear

$$Ax = b, \qquad x \in \mathbb{R}^m \tag{3.11}$$

onde $b \in \mathbb{R}^m$ é dado. Sabemos que (3.11) tem solução única x para qualquer b dado se, e somente se, A é não singular, ou seja, $\det(A) \neq 0$.

Uma prática comum para desenvolver o método iterativo pra resolução de (3.11) é através da divisão da matriz

$$A = N - M$$

com N escolhido de tal maneira que o sistema $Nx = k$ seja fácil e resolvível de forma única para qualquer k. Então o sistema linear (3.11) é reescrito como

$$Nx = Mx + b.$$

O que leva naturalmente ao método iterativo para resolução de (3.11):

$$Nx_n = Mx_{n-1} + b, \qquad n = 1, 2, \ldots \tag{3.12}$$

sendo x_0 uma aproximação inicial da solução x.

Para analisar mais facilmente a iteração, reescrevemos estas duas últimas equações como

$$\begin{aligned} x =& N^{-1}Mx + N^{-1}b, \\ x_n =& N^{-1}Mx_{n-1} + N^{-1}b. \end{aligned}$$

A matriz $N^{-1}M$ é chamada de matriz de iteração. Subtraindo as duas equações, nós obtemos a equação do erro

$$x - x_n = N^{-1}M(x - x_{n-1}).$$

Por indução

$$x - x_n = (N^{-1}M)^n(x - x_0), \qquad n = 0, 1, 2, \ldots \tag{3.13}$$

Podemos ver que o método iterativo converge se $\|N_{-1}M\| < 1$, onde $\|\cdot\|$ é a norma de algum operador de matriz, isto é, é a norma induzida por alguma norma vetorial $\|\cdot\|$:

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}. \tag{3.14}$$

Recordando que para uma matriz quadrada A, uma condição necessária e suficiente para $A^n \longrightarrow 0$ quando $n \longrightarrow \infty$ é $r_\sigma(A) < 1$. Isto segue da forma canônica de Jordan para matrizes quadradas. Aqui $r_\sigma(A)$ é o raio espectral de A:

$$r_\sigma(A) = \max_i |\lambda_i(A)|,$$

com $\{\lambda_i(A)\}$ o conjunto de todos os autovalores de A. Da relação de erro (3.13), temos convergência $x_n \longrightarrow x$ quando $n \longrightarrow \infty$ para qualquer valor inicial x_0 se, e somente se, $(N^{-1}M)^n \longrightarrow 0$ quando $n \longrightarrow \infty$. Portanto, para o método iterativo (3.12), uma condição necessária e suficiente para convergência é $r_\sigma(N^{-1}M) < 1$.

O raio espectral de uma matriz é uma quantidade intrínseca da matriz, enquanto que a norma matricial não é. Portanto é natural que a condição necessária e suficiente para convergência do método iterativo seja descrita em termos do raio espectral da matriz de iteração. Seria também de esperar algo deste tipo baseado na propriedade que em espaços dimensionalmente finitos, a convergência de $\{x_n\}$ em uma norma seja equivalente a convergência em qualquer outra norma.

Temos as seguintes relações entre o raio espectral e as normas de uma matriz $A \in \mathbb{R}^{m\times m}$.

1. 1. $r_\sigma(A) \leqslant \|A\|$ para qualquer norma de operador de matriz $\|\cdot\|$. Este resultado segue imediatamente da definição de $r_\sigma(A)$, da relação que define autovalor, e do fato que a norma da matriz é gerada por uma norma vetorial.

2. 2. Para qualquer $\varepsilon > 0$, existe uma norma de operador de matriz $\|\cdot\|_{A,\varepsilon}$ tal que

   $$r_\sigma(A) \leqslant \|A\|_{A,\varepsilon} \leqslant r_\sigma(A) + \varepsilon.$$

   Então

   $$r_\sigma(A) = \inf\{\|A\| \mid \|\cdot\| \text{ é uma norma de operador de matriz}\}.$$

3. 3. $r_\sigma(A) = \lim_{n\to\infty} \|A^n\|^{1/n}$ para qualquer norma de matriz. Aqui a norma pode ser qualquer norma de matriz, não necessariamente a gerada pela norma vetorial como em (3.14).

3.2.3 EQUAÇÃO INTEGRAL NÃO LINEAR DE VOLTERRA

Uma equação da forma

$$u(t) = f(t) + \int_a^t k(t, s, u(s))ds \qquad t \in [a, b] \tag{3.15}$$

é chamada de equação integral não linear de Volterra do segundo tipo.

A forma da equação (3.15) leva naturalmente ao método iterativo

$$u_n(t) = f(t) + \int_a^t k(t, s, u_{n-1}(s))ds \qquad t \in [a, b],\ n \geqslant 1. \tag{3.16}$$

Lema 3.1. *(Ponto fixo) Seja $T : X \longrightarrow X$ um mapeamento contínuo em um espaço de Banach X, e suponha que T^m é uma contração em X para algum inteiro positivo m. Então T tem um único ponto fixo.*

Demonstração. Por hipótese, $B = T^m$ é uma contração em X, isto é, $\|Bx - By\| \leqslant \alpha\|x - y\|$ para todo $x, y \in X$; aqui $\alpha < 1$.

Assim, para cada $x_0 \in X$,

$$\begin{aligned} \|B^n T x_0 - B^n x_0\| \leqslant& \alpha\|B^{n-1}Tx_0 - B^{n-1}x_0\| \\ \cdots \leqslant& \alpha^n\|Tx_0 - x_0\| \quad \longrightarrow 0 \quad (n \longrightarrow \infty). \end{aligned} \tag{3.17}$$

O teorema do ponto fixo de Banach 3.1 implica que B tem um único ponto fixo, chamemos ele de x, e $B^n x_0 \longrightarrow x$. Uma vez que o mapeamento T é contínuo, isto implica em $B^n T x_0 = T B^n x_0 \longrightarrow Tx$. Assim pelo lema 1.1 2)

$$\|B^n T x_0 - B^n x_0\| \longrightarrow \|Tx - x\|,$$

tal que $\|Tx - x\| = 0$ por (3.17). Isto mostra que x é um ponto fixo de T. Uma vez que cada ponto fixo de T é também ponto fixo de B, vemos que T não pode ter mais que um ponto fixo. □

Teorema 3.3. *Assuma que $k(t, s, u)$ é contínuo para $a \leqslant s \leqslant t \leqslant b$ e $u \in \mathbb{R}$; e seja $f \in C[a, b]$. Além disso, assuma*

$$|k(t, s, u_1) - k(t, s, u_2)| \leqslant M|u_1 - u_2|, \qquad a \leqslant s \leqslant t \leqslant b,\ u_1, u_2 \in \mathbb{R}$$

para alguma constante M. Então a equação integral (3.15) tem uma única solução $u \in C[a, b]$. Além disso, o método iterativo (3.16) converge, qualquer que seja a função inicial $u_0 \in C[a, b]$.

Demonstração. Há pelo menos duas abordagens para aplicar o teorema do ponto fixo de Banach para provar a existência de uma solução única de (3.15). Assumindo as condições estabelecidas no teorema 3.3 definimos o operador integral não linear

$$T : C[a, b] \to C[a, b], \qquad Tu(t) \equiv \int_a^t k(t, s, u(s))ds + f(t).$$

Abordagem 1: Vamos mostrar que para um m suficientemente grande, o operador T^m é uma contração em $C[a, b]$. Para $u, v \in C[a, b]$,

$$Tu(t) - Tv(t) = \int_a^t [k(t, s, u(s)) - k(t, s, v(s))]ds.$$

Então

$$|Tu(t) - Tv(t)| \leqslant M \int_a^t |u(s) - v(s)|ds \tag{3.18}$$

e

$$|Tu(t) - Tv(t)| \leqslant M\|u - v\|_\infty (t - a).$$

Uma vez que

$$T^2u(t) - T^2v(t) = \int_a^t [k(t, s, Tu(s)) - k(t, s, Tv(s))]ds.$$

temos

$$\begin{aligned} |T^2u(t) - T^2v(t)| \leqslant & M \int_a^t |Tu(s) - Tv(s)|ds \\ \leqslant & \frac{[M(t-a)]^2}{2!}\|u - v\|_\infty. \end{aligned}$$

Por indução obtemos

$$|T^m u(t) - T^m v(t)| \leqslant \frac{[M(t-a)]^m}{m!}\|u - v\|_\infty.$$

Assim

$$\|T^m u - T^m v\|_\infty \leqslant \frac{[M(b-a)]^m}{m!}\|u - v\|_\infty.$$

Uma vez que

$$\frac{[M(b-a)]^m}{m!} \to 0 \quad \text{quando } m \to \infty,$$

o operador T^m é uma contração em $C[a, b]$ quando m é suficientemente grande. Pelo lema 3.1 o operador T tem um único ponto fixo em $C[a, b]$ e a sequência de iteração converge para a solução.

Abordagem 2: Sobre o espaço $C[a, b]$, definimos a norma

$$|||v||| = \max_{a \leqslant t \leqslant b} e^{-\beta t}|v(t)|$$

que é equivalente a norma padrão $\|v\|_\infty$ em $C[a, b]$. O parâmetro β é escolhido para satisfazer $\beta > M$. Então modificamos a relação (3.18) como segue:

$$e^{-\beta t}|Tu(t) - Tv(t)| \leqslant Me^{-\beta t} \int_a^t e^{\beta s}e^{-\beta s}|u(s) - v(s)|ds.$$

Consequentemente,

$$\begin{aligned} e^{-\beta t}|Tu(t) - Tv(t)| \leqslant & Me^{-\beta t}|||u - v||| \int_a^t e^{\beta s}ds. \\ = & \frac{M}{\beta}e^{-\beta t}(e^{\beta t} - e^{\beta a})|||u - v|||. \end{aligned}$$

Portanto,

$$|||Tu - Tv||| \leqslant \frac{M}{\beta}|||u - v|||.$$

Uma vez que $\beta > M$, o operador T é uma contração no espaço de Banach $(V, |||\cdot|||)$. Então T tem um único ponto fixo que é a única solução da equação integral (3.15) e a sequência de iteração converge. $\square$

Exemplo 3.1. Use o método das aproximações sucessivas para resolver a equação integral não linear de Volterra

$$u(x) = e^x + \frac{1}{3}x(1 - e^{3x}) + \int_0^x xu^3(t)dt$$

O método das aproximações sucessivas admite o uso da fórmula de iteração. Logo temos

$$u_n(x) = e^x + \frac{1}{3}x(1 - e^{3x}) + \int_0^x xu_{n-1}^3(t)dt, \qquad n \geqslant 0$$

Para realização dos cálculos utilizaremos a linguagem de programação Python e a biblioteca SymPy. SymPy é uma biblioteca Python para matemática simbólica. Nosso objetivo aqui é ilustrar a teoria. Por isso utilizamos matemática simbólica e não pacotes ou softwares para cálculo numérico ou programação utilizando algoritmos computacionalmente mais eficientes.

Segue a implementação do algoritmo.

classoffset

```
# -*- coding: utf-8 -*-
"""
Created on Mon Sep 29 22:21:14 2015

@author: Kleber
"""

from sympy import *
init_printing(use_unicode=True)

x, y, z = symbols("x y z") # Define simbolos para variaveis genericas
f, g, h = map(Function, 'fgh') # define simbolos para funcoes

A=[]
S=[]

def IE(u0):
    A.append(u0)
    for i in range (0,3):
        u1=exp(x) + Rational(1,3)*x*(1-exp(3*x))+integrate(x*u0**3, (t, 0, x))
        u0=u1
        A.append(u1)
        S.append(series(u1,x,0,7))
        u0=u0.subs(x,t)
    return S[i]
```

Escolhendo $u_0 = 1$ para a aproximação inicial. Obtivemos os seguintes resultados.

$$
\begin{aligned}
u_0(x) =& 1\\
u_1(x) =& e^x + \frac{1}{3}x(1-e^{3x}) + \int_0^x xu_0^3(t)dt = 1 + x + \frac{x^2}{2} - \frac{4x^3}{3} - \frac{35x^4}{24} - \frac{67x^5}{60} - \frac{97x^6}{144} + O(x^7)\\
u_2(x) =& e^x + \frac{1}{3}x(1-e^{3x}) + \int_0^x xu_0^3(t)dt = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} - \frac{67x^5}{60} - \frac{1943x^6}{720} + O(x^7)\\
u_3(x) =& e^x + \frac{1}{3}x(1-e^{3x}) + \int_0^x xu_0^3(t)dt = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \frac{x^6}{720} + O(x^7)
\end{aligned}
$$

Escolhendo $u_0 = x$ para a aproximação inicial temos os seguintes resultados.

$$
\begin{aligned}
u_0(x) =& x\\
u_1(x) =& e^x + \frac{1}{3}x(1-e^{3x}) + \int_0^x xu_0^3(t)dt = 1 + x - \frac{x^2}{2} - \frac{4x^3}{3} - \frac{35x^4}{24} - \frac{7x^5}{60} - \frac{97x^6}{144} + O(x^7)\\
u_2(x) =& e^x + \frac{1}{3}x(1-e^{3x}) + \int_0^x xu_0^3(t)dt = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} - \frac{23x^4}{24} - \frac{157x^5}{60} - \frac{475x^6}{144} + O(x^7)\\
u_3(x) =& e^x + \frac{1}{3}x(1-e^{3x}) + \int_0^x xu_0^3(t)dt = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} - \frac{431x^6}{720} + O(x^7)
\end{aligned}
$$

Como esperado qualquer que seja a função inicial $u_0(x)$ o método iterativo converge para a mesma solução.

Consequentemente, a solução $u(x)$ é dada por

$$u(x) = \lim_{n\to\infty} u_n(x) = e^x.$$

3.2.4 EQUAÇÕES DIFERENCIAIS ORDINÁRIAS EM ESPAÇOS DE BANACH

Seja V um espaço de Banach e considere o problema de valor inicial

$$
\begin{cases}
u'(t) = f(t, u(t)), \qquad |t - t_0| < a,\\
u(t_0) = z
\end{cases}
\tag{3.19}
$$

Aqui $z \in V$ e $f : [t_0 - a, t_0 + a] \times V \longrightarrow V$ é contínuo. Por exemplo, f pode ser um operador integral; e então (3.19) pode ser uma equação integro diferencial.

O problema da equação diferencial (3.19) é equivalente a equação integral

$$u(t) = z + \int_{t_0}^{t} f(s, u(s))ds, \qquad |t - t_0| < a, \tag{3.20}$$

que é da forma $u = T(u)$. O que leva naturalmente ao método de iteração do ponto fixo

$$u_n(t) = z + \int_{t_0}^{t} f(s, u_{n-1}(s))ds, \qquad |t - t_0| < a,\ n \geqslant 1. \tag{3.21}$$

Para $b > 0$,

$$Q_b \equiv \{(t,u) \in \mathbb{R} \times V \mid |t - t_0| \leqslant a, \|u - z\| \leqslant b\}.$$

Temos o seguinte teorema de existência e unicidade para (3.19).

Teorema 3.4. *(Teorema da existência e unicidade de Picard) Seja f contínua em um retângulo*

$$R = \{(t,x)|\ |t - t_0| \leqslant a, |x - x_0| \leqslant b\}$$

conforme mostrado na figura 2, e portanto limitado em R. E seja

$$|f(t,x)| \leqslant c \tag{3.22}$$

para todo $(t,x) \in R$.

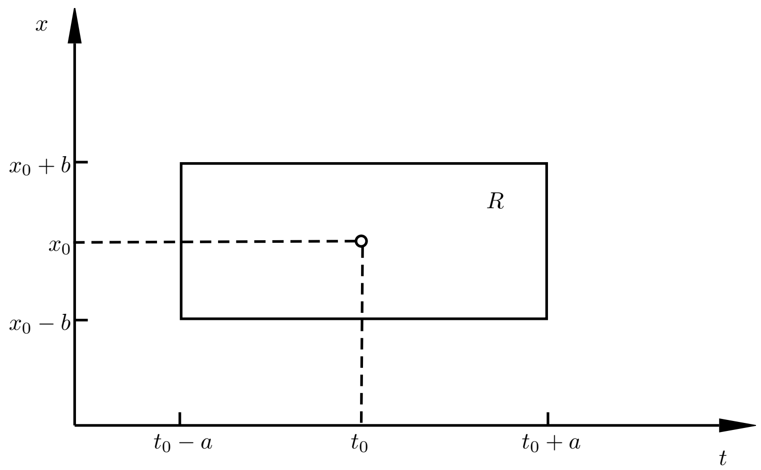

Figura 2 – O retângulo R.

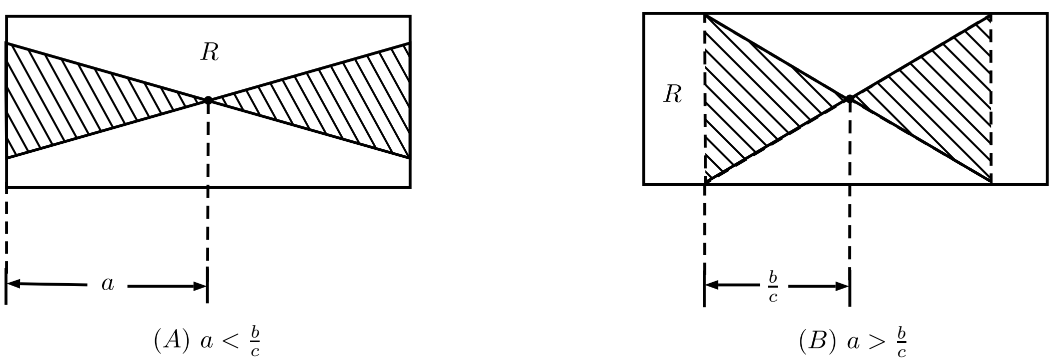

Figura 3 – Ilustração geométrica da desigualdade (3.22) para c relativamente pequeno (Figura A), c relativamente grande (Figura B). A curva solução deve permanecer na região hachurada delimitada pelos segmentos de reta com inclinação $\pm c$.

Suponha que f satisfaça a condição de Lipschitz em R com respeito ao seu segundo argumento, isto é, existe uma constante k (constante de Lipschitz) tal que para $(t,x), (t,v) \in$

R

$$|f(t,x) - f(t,v)| \leqslant k|x - v|. \tag{3.23}$$

Então o problema de valor inicial tem uma única solução. Esta solução existe em um intervalo $[t_0 - \beta, t_0 + \beta]$*,onde*

$$\beta < \min\left\{a, \frac{b}{c}, \frac{1}{k}\right\}. \tag{3.24}$$

Demonstração. Seja $C(J)$ o espaço métrico de todas as funções contínuas de valor real no intervalo $J = [t_0 - \beta, t_0 + \beta]$ com a métrica definida por

$$d(x,y) = \max_{t \in J} |x(t) - y(t)|$$

$C(J)$ é completo. Seja $\tilde{C}$ o subespaço de $C(J)$ consistindo de todas as funções $x \in C(J)$ que satisfazem

$$|x(t) - x_0| \leqslant c\beta. \tag{3.25}$$

Mostremos que $\tilde{C}$ é fechado em $C(J)$. Seja (x_n) uma sequência convergente em $\tilde{C}$, ou seja, existe $\tilde{x} \in C(J)$ tal que $x_n \longrightarrow \tilde{x}$, isto é, dado $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que

$$d(x_n, \tilde{x}) < \varepsilon,$$

para todo $n > n_0$. Como $(x_n) \subset \tilde{C}$, tem-se

$$|x_n(t) - x_0| \leqslant c\beta, \quad \forall t \in J.$$

Como $c\beta$ não depende de t, segue que

$$d(x_n, x_0) = \max_{t \in J} |x_n(t) - x_0| \leqslant c\beta.$$

Da desigualdade triangular temos

$$d(\tilde{x}, x_0) \leqslant d(\tilde{x}, x_n) + d(x_n, x_0) \leqslant \varepsilon + c\beta, \quad n > n_0.$$

Fazendo $\varepsilon \longrightarrow 0$, obtêm-se

$$|\tilde{x}(t) - x_0| \leqslant \max_{t \in J} |\tilde{x}(t) - x_0| = d(\tilde{x}, x_0) \leqslant c\beta$$

Segue que $\tilde{x} \in \tilde{C}$ e $\tilde{C}$ é um espaço fechado em $C(J)$. Segue do teorema (1.16) que $\tilde{C}$ é completo.

Por integração vemos que (3.19) pode ser escrito como $x = Tx$, onde $T : \tilde{C} \longrightarrow \tilde{C}$ é definido por

$$Tx(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau))\, d\tau. \tag{3.26}$$

De fato, T é definido para todo $x \in \tilde{C}$, porque $c\beta < b$ por (3.24), de modo que se $x \in \tilde{C}$, então $\tau \in J$ e $(\tau, x(\tau)) \in R$, e a integral (3.26) existe uma vez que f é contínua em R. Para ver que T mapeia $\tilde{C}$ em si mesmo, podemos usar (3.26) e (3.22), obtendo

$$|Tx(t) - x_0| = \left| \int_{t_0}^{t} f(\tau, x(\tau))\, d\tau \right| \leqslant c|t - t_0| \leqslant c\beta.$$

Vamos mostrar que T é uma contração em $\tilde{C}$. Pela condição de Lipschitz (3.23)

$$\begin{aligned}|Tx(t) - Tv(t)| &= \left| \int_{t_0}^{t} [f(\tau, x(\tau)) - f(\tau, v(\tau))]\, d\tau \right| \\ &\leqslant |t - t_0| \max_{\tau \in J} k|x(\tau) - v(\tau)| \\ &\leqslant k\beta\ d(x, v).\end{aligned}$$

Uma vez que a última expressão não depende de t, podemos considerar o máximo na esquerda e ter

$$d(Tx, Tv) \leqslant \alpha d(x, v), \qquad \alpha = k\beta.$$

De (3.24) verificamos que $\alpha = k\beta < 1$, de modo que T é de fato uma contração em $\tilde{C}$. O teorema do ponto fixo de Banach portanto implica que T tem um único ponto fixo $x \in \tilde{C}$, isto é, uma função contínua x em J satisfazendo $x = Tx$. Escrevendo $x = Tx$, temos de (3.26)

$$x(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau))\, d\tau. \tag{3.27}$$

Uma vez que $(\tau, x(\tau)) \in R$ onde f é contínua, (3.27) pode ser diferenciável. Consequentemente x é ainda diferenciável e satisfaz (3.19). Reciprocamente, toda solução de (3.19) deve satisfazer (3.27). Isto completa a prova. □

Exemplo 3.2. Utilize o método previsor-corretor para aproximar valores da solução do seguinte problema de valor inicial

$$\begin{cases} y' = y - x^2 + 1, \qquad 0 \leqslant x \leqslant 4, \qquad N = 20 \\ y(0) = 0.5 \end{cases}$$

Para obter os valores iniciais utilizaremos o método de Runge-Kutta de ordem 4. Em seguida utilizaremos o método de Adams-Bashforth de 4-passos como previsor e melhoraremos este resultado utilizando o método de Adams-Moulton de 3-passos como corretor. Faremos uma única iteração.

Para realização dos cálculos utilizaremos a linguagem de programação Python e a biblioteca NumPy. NumPy é uma biblioteca fundamental para computação científica com Python. Nosso objetivo aqui foi o de obter um implementação computacionalmente eficiente.

Segue a implementação do algoritmo.

classoffset

```
# -*- coding: utf-8 -*-
"""
Created on Wed Oct 14 21:31:35 2015

@author: Kleber
e-mail:1kleber1@gmail.com
"""
from numpy import *
#intervalo
a=0.0
b=4.0
#N
N=20
#valor inicial
y=arange(N+2,dtype=float64)
y[0]=0.5
#calculo de h
h=(b-a)/N
#valores de x
x=arange(a,b+2*h,h)
# f(x,y)
def f(x,y):
    return y-x**2+1
# Metodo preditor corretor de ordem 4
def PC1(f):
    f1=f2=f3=0
# Runge-Kutta de ordem 4
    for i in range(0,3,1):
        f0=f(x[i],y[i])
        k1=h*f0
        k1=h*f(x[i],y[i])
        k2=h*f(x[i]+0.5*h,y[i]+0.5*k1)
        k3=h*f(x[i]+0.5*h,y[i]+0.5*k2)
        k4=h*f(x[i+1],y[i]+k3)
        y[i+1]=y[i]+(k1+2*k2+2*k3+k4)/6.0
        f1, f2, f3 = ( f0, f1, f2 )
# Adams-Bashfortth de 4-passos e Adams-Moulton de 3-passos
    for i in range(3,N+1,1):
        f0=f(x[i],y[i])
        y[i+1]=y[i]+h*(55*f0-59*f1+37*f2-9*f3)/24.0
        fp=f(x[i+1],y[i+1])
```

```
        y[i+1]=y[i]+h*(9*fp+19*f0-5*f1+f2)/24.0
        f1, f2, f3 = ( f0, f1, f2 )
    return y
```

Obtivemos os seguintes resultados.

Tabela 1 – Resultados obtidos usando o método previsor-corretor.

x_i	solução analítica	solução aproximada	erro
0.0	0.5000000	0.5000000	0.0000000
0.2	0.8292986	0.8292933	0.0000053
0.4	1.2140877	1.2140762	0.0000114
0.6	1.6489406	1.6489220	0.0000186
0.8	2.1272295	2.1272056	0.0000239
1.0	2.6408591	2.6408286	0.0000305
1.2	3.1799415	3.1799026	0.0000389
1.4	3.7324000	3.7323505	0.0000495
1.6	4.2834838	4.2834208	0.0000630
1.8	4.8151763	4.8150964	0.0000799
2.0	5.3054720	5.3053707	0.0001013
2.2	5.7274933	5.7273651	0.0001282
2.4	6.0484118	6.0482498	0.0001621
2.6	6.2281310	6.2279264	0.0002046
2.8	6.2176766	6.2174185	0.0002581
3.0	5.9572315	5.9569063	0.0003252
3.2	5.3737349	5.3733255	0.0004094
3.4	4.3779500	4.3774350	0.0005150
3.6	2.8608828	2.8602356	0.0006472
3.8	0.6894078	0.6885950	0.0008127
4.0	-2.2990750	-2.3000948	0.0010198

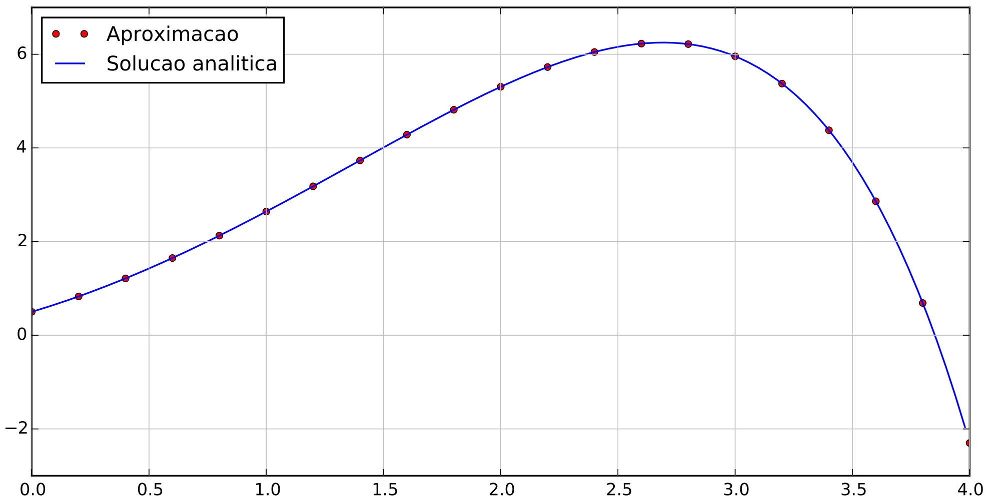

Figura 4 – Solução analítica versus solução aproximada.

3.3 DERIVADAS DE FRÉCHET E GÂTEAUX

Temos as seguintes definições de derivada para uma função real.

Seja I um intervalo em $\mathbb{R}$, e x_0 um ponto interior de I. Uma função $f : I \longrightarrow \mathbb{R}$ é derivável em x_0 se, e somente se

$$f'(x) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \text{ existe;} \tag{3.28}$$

ou equivalentemente, para algum número a,

$$f(x_0 + h) = f(x_0) + ah + o(|h|), \qquad \text{quando } h \longrightarrow 0, \tag{3.29}$$

tem-se $f'(x_0) = a$ para denotar a derivada.

Embora as duas definições sejam equivalentes. Mostraremos através de uma função vetorial de várias variáveis reais que a definição (3.29) pode ser estendida diretamente para definir a derivada de um operador geral. Enquanto a forma (3.28) é útil para definir uma derivada direcional ou parcial de um operador.

Seja K um subconjunto do espaço $\mathbb{R}^d$, com x_0 como um ponto interior. Seja $f : K \longrightarrow \mathbb{R}^m$. Seguindo (3.29), dizemos que f é diferenciável em x_0 se existir uma matriz (operador linear) $A \in \mathbb{R}^{m \times d}$ tal que

$$f(x_0 + h) = f(x_0) + Ah + o(\|h\|), \qquad \text{quando } \ h \longrightarrow 0, h \in \mathbb{R}^d. \tag{3.30}$$

Podemos mostrar que $A = \nabla f(x_0) = (A_{ij})$, o gradiente ou jacobiano de f em x_0:

$$A_{ij} = \frac{\partial f_i}{\partial x_j}, \qquad 1 \leqslant i \leqslant m, \ 1 \leqslant j \leqslant d.$$

Seja h um vetor fixo em $\mathbb{R}^d$, e consideremos a função $f(x_0 + th)$ para $t \in \mathbb{R}$ em uma vizinhança de 0. Podemos dizer que f é diferenciável em x_0 com respeito a h, se existe uma matriz A tal que

$$\lim_{t \to 0} \frac{f(x_0 + th) - f(x_0)}{t} = Ah. \tag{3.31}$$

No caso $\|h\| = 1$, chamamos a quantidade Ah de derivada direcional de f em x_0 ao longo da direção h.

Passando para o caso de um operador $f : K \subset V \longrightarrow W$ entre dois espaços normados V e W. Adotaremos a convenção de sempre que discutirmos a diferenciabilidade em um ponto u_0, implicitamente assumiremos que u_0 é um ponto interior de K; com isso queremos dizer que há um $r > 0$ tal que

$$B(u_0, r) \equiv \{u \in V | \|u - u_0\| \leqslant r\} \subset K.$$

Definição 3.2. *O operador f é Fréchet diferenciável em u_0 se, e somente se, existe $A \in \mathcal{L}(V, W)$ tal que*

$$f(u_0 + h) = f(u_0) + Ah + o(\|h\|), \qquad h \longrightarrow 0 \tag{3.32}$$

O mapa A é chamado de derivada de Fréchet de f em u_0, e escrevemos $A = f'(u_0)$. A quantidade $df(u_0; h) = f'(u_0)h$ é chamada de diferencial de Fréchet de f em u_0 ao longo de h. Se f é Fréchet diferenciável em todos os pontos em $K_0 \subset K$, dizemos que f é Fréchet diferenciável em K_0 e chamamos $f' : K_0 \subset V \longrightarrow \mathcal{L}(V, W)$ de derivada de Fréchet de f em K_0.

Se f é diferenciável em u_0, então a derivada $f'(u_0)$ é única. Isto é verificado da seguinte forma. Suponha que exista outro mapa $\tilde{A} \in \mathcal{L}(V, W)$ tal que

$$f(u_0 + h) = f(u_0) + \tilde{A}h + o(\|h\|), \qquad h \longrightarrow 0.$$

Então

$$\tilde{A}h - f'(u_0)h = o(\|h\|), \qquad h \longrightarrow 0.$$

Para qualquer $h_0 \in V$ com $\|h_0\| = 1$, seja $h = th_0$, $0 \neq t \in \mathbb{R}$. Dividindo a relação por t e tomando o limite $t \longrightarrow 0$, obtemos

$$\tilde{A}h_0 - f'(u_0)h_0 = 0.$$

Consequentemente, $\tilde{A} = f'(u_0)$.

Definição 3.3. *O operador f é Gâteaux diferenciável em u_0 se, e somente se, existe $A \in \mathcal{L}(V, W)$ tal que*

$$\lim_{t \longrightarrow 0} \frac{f(u_0 + th) - f(u_0)}{t} = Ah, \qquad \forall h \in V. \tag{3.33}$$

O mapa A é chamado de derivada de Gâteaux de f em u_0, e escrevemos $A = f'(u_0)$. A quantidade $df(u_0; h) = f'(u_0)h$ é chamada de diferencial de Gâteaux de f em u_0 ao longo de h. Se f é Gâteaux diferenciável em todos os pontos em $K_0 \subset K$, dizemos que f é Gâteaux diferenciável em K_0 e chamamos $f' : K_0 \subset V \longrightarrow \mathcal{L}(V, W)$ de derivada de Gâteaux de f em K_0.

3.4 MÉTODO DE NEWTON

Seja $f : \mathbb{R} \longrightarrow \mathbb{R}$ continuamente diferenciável e considere a equação

$$f(x) = 0.$$

Supondo conhecida uma solução aproximada x_n perto de uma raiz x^* da equação. Então pela expansão de Taylor,

$$\begin{aligned} 0 =& f(x^*) \\ =& f(x_n) + f'(x_n)(x^* - x_n) + o(|x^* - x_n|) \\ \approx& f(x_n) + f'(x_n)(x^* - x_n). \end{aligned}$$

Então,

$$x^* \approx x_n - [f'(x_n)]^{-1} f(x_n).$$

Isto leva ao método de Newton para resolução de equações $f(x) = 0$:

$$x_{n+1} \approx x_n - [f'(x_n)]^{-1} f(x_n), \qquad n = 0, 1, \ldots$$

3.4.1 MÉTODO DE NEWTON EM ESPAÇOS DE BANACH

Sejam U e V dois espaços de Banach. Assumindo que $F : U \longrightarrow V$ é Fréchet diferenciável. Estamos interessados em resolver a equação

$$F(u) = 0 \tag{3.34}$$

O método de Newton estabelece o seguinte: Escolha um vetor inicial $u_0 \in U$; para $n = 0, 1, \ldots$, compute

$$u_{n+1} = u_n - [F'(u_n)]^{-1} F(u_n) \tag{3.35}$$

Teorema 3.5. *(convergência local) Assuma que u^* é uma solução da equação (3.34) tal que $[F'(u^*)]^{-1}$ existe e é um mapa linear contínuo de V em U. Assuma adicionalmente que $F'(u)$ é localmente Lipschitz contínua em u^*,*

$$\|F'(u) - F'(v)\| \leqslant L\|u - v\| \qquad \forall u, v \in N(u^*),$$

onde $N(u^*)$ é uma vizinhança de u^*, e $L > 0$ é uma constante. Então existe um $\delta > 0$ tal que se $\|u_0 - u^*\| \leqslant \delta$, a sequência de Newton $\{u_n\}$ é bem definida e converge para u^*. Além disso, para alguma constante M com $M\delta < 1$, nós temos os limites de erro

$$\|u_{n+1} - u^*\| \leqslant M\|u_n - u^*\|^2 \tag{3.36}$$

e

$$\|u_n - u^*\| \leqslant (M\delta)^{2^n}/M. \tag{3.37}$$

Demonstração. Após a redefinição da vizinhança $N(u^*)$ se necessário, podemos assumir que $[F'(u)]^{-1}$ existe em $N(u^*)$ e

$$c_0 = \sup_{u \in N(u^*)} \|[F'(u)]^{-1}\| < \infty.$$

Definindo

$$T(u) = u - [F'(u)]^{-1}F(u), \qquad u \in N(u^*).$$

Podemos observar que $T(u^*) = u^*$. Para $u \in N(u^*)$, temos

$$\begin{aligned} T(u) - T(u^*) =& u - u^* - [F'(u)]^{-1}F(u) \\ =& [F'(u)]^{-1}[F(u^*) - F(u) - F'(u)(u^* - u)] \\ =& [F'(u)]^{-1} \int_0^1 [F'(u + t(u^* - u)) - F'(u)]dt(u^* - u) \end{aligned}$$

e tomando a norma,

$$\begin{aligned} \|T(u) - T(u^*)\| \leqslant& \|[F'(u)]^{-1}\| \int_0^1 \|F'(u + t(u^* - u)) - F'(u)\|dt\|u^* - u\| \\ \leqslant& \|[F'(u)]^{-1}\| \int_0^1 Lt\|u^* - u\|dt\|u^* - u\|. \end{aligned}$$

Consequentemente,

$$\|T(u) - T(u^*)\| \leqslant \frac{c_0 L}{2}\|u^* - u\|^2. \tag{3.38}$$

Escolha $\delta < 2/(c_0 L)$ com a propriedade $\overline{B}(u^*, \delta) \subset N(u^*)$; e note que

$$\alpha \equiv c_0 L\delta/2 < 1.$$

Então (3.38) implica em

$$\begin{aligned}\|T(u)-u^*\| =&\|T(u)-T(u^*)\| \\ \leqslant&\alpha\|u-u^*\|,\ \ u\in\overline{B}(u^*,\delta).\end{aligned} \tag{3.39}$$

Suponha uma estimativa inicial $u_0 \in \overline{B}(u^*,\delta)$. Então (3.39) implica

$$\|u_1-u^*\| = \|T(u_0)-u^*\| \leqslant \alpha\|u_0-u^*\| \leqslant \alpha\delta < \delta.$$

Assim $u_1 \in \overline{B}(u^*,\delta)$. Repetindo este argumento indutivamente, temos $u_n \in \overline{B}(u^*,\delta)$ para todo $n \geqslant 0$.

Para obter a convergência de $\{x_n\}$, começamos com

$$\|u_{n+1}-u^*\| = \|T(u_n)-u^*\| \leqslant \alpha\|u_n-u^*\|, \qquad n \geqslant 0.$$

Por indução,

$$\|u_n-u^*\| \leqslant \alpha^n\|u_0-u^*\|, \qquad n\geqslant 0$$

e $u_n \longrightarrow u^*$ quando $n \longrightarrow \infty$.

Retornando para (3.38), denotamos $M = c_0L/2$. Note que $M\delta = \alpha < 1$. Então obtemos a estimativa

$$\|u_{n+1}-u^*\| = \|T(u_n)-u^*\| \leqslant M\|u_n-u^*\|^2,$$

provando (3.36). Multiplicando os dois lados por M, obtemos

$$M\|u_{n+1}-u^*\| \leqslant (M\|u_n-u^*\|)^2.$$

Uma aplicação indutiva desta desigualdade leva a

$$M\|u_n-u^*\| \leqslant (M\|u_0-u^*\|)^{2^n},$$

provando então (3.37). □

O teorema claramente mostra que o método de Newton é localmente convergente com convergência quadrática. A principal desvantagem do resultado é a dependência das suposições sobre a raiz da equação, que é a quantidade a ser calculada. A teoria de Kantorovich supera esta dificuldade.

Teorema 3.6. *(Kantorovich) Suponha que*

(a) $F:\mathcal{D}(F)\subset U \longrightarrow V$ é diferenciável em um conjunto aberto convexo $\mathcal{D}(F)$, e a derivada é Lipschitz contínua:

$$\|F'(u)-F'(v)\| \leqslant L\|u-v\| \qquad \forall u,v\in\mathcal{D}(F).$$

(b) Para algum $u_0 \in \mathcal{D}(F)$, $[F'(u_0)]^{-1}$ *existe e é um operador contínuo de* V *em* U*, tal que* $h = abL \leqslant 1/2$ *para algum* $a \geqslant \|[F'(u_0)]^{-1}\|$ *e* $b \geqslant \|[F'(u_0)]^{-1}F(u_0)\|$*. Denote*

$$t^* = \frac{1-(1-2h)^{1/2}}{aL}, \qquad t^{**} = \frac{1+(1-2h)^{1/2}}{aL}.$$

(c) u_0 *é escolhido de modo que* $\overline{B}(u_1, r) \subset \mathcal{D}(F)$*, onde* $r = t^* - b$*.*

Então a equação (3.34) tem uma solução $u^* \in \overline{B}(u_1, r)$ *e a solução é única em* $\overline{B}(u_0, t^{**}) \cap \mathcal{D}(F)$*; a sequência* $\{u_n\}$ *converge para* u^**, e temos a estimativa de erro*

$$\|u_n - u^*\| \leqslant \frac{[1-(1-2h)^{1/2}]^{2^n}}{2^n aL}, \qquad n = 0, 1, \ldots$$

O teorema de Kantorovich provêm condições suficientes para a convergência do método de Newton. Estas condições são usualmente difíceis de verificar. No entanto, pelo menos teoricamente, o resultado é de grande importância.

3.4.2 APLICAÇÕES DO MÉTODO DE NEWTON

3.4.2.1 SISTEMAS NÃO LINEARES

Seja $F : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ uma função continuamente diferenciável. Um sistema não-linear é da forma

$$x \in \mathbb{R}^d, \qquad F(x) = 0.$$

Então o método de Newton é

$$x_{n+1} = x_n - [F'(x_n)]^{-1}F(x_n), \tag{3.40}$$

que também pode ser escrito sob a forma

$$F'(x_n)\delta_n = -F(x_n), \qquad x_{n+1} = x_n + \delta_n.$$

Então a cada passo, resolvemos um sistema linear. O método falha quando $F'(x_n)$ é singular ou aproximadamente singular.

Exemplo 3.3. Use o método de Newton para aproximar soluções para o sistema de equações

$$\begin{cases} x^2 + y^2 + z^2 = 1 \\ 2x^2 + y^2 - 4z = 0 \\ 3x^2 - 4y + z^2 = 0 \end{cases}$$

Para problemas com duas ou três variáveis podemos utilizar gráficos para estabelecer os valores da aproximação inicial.

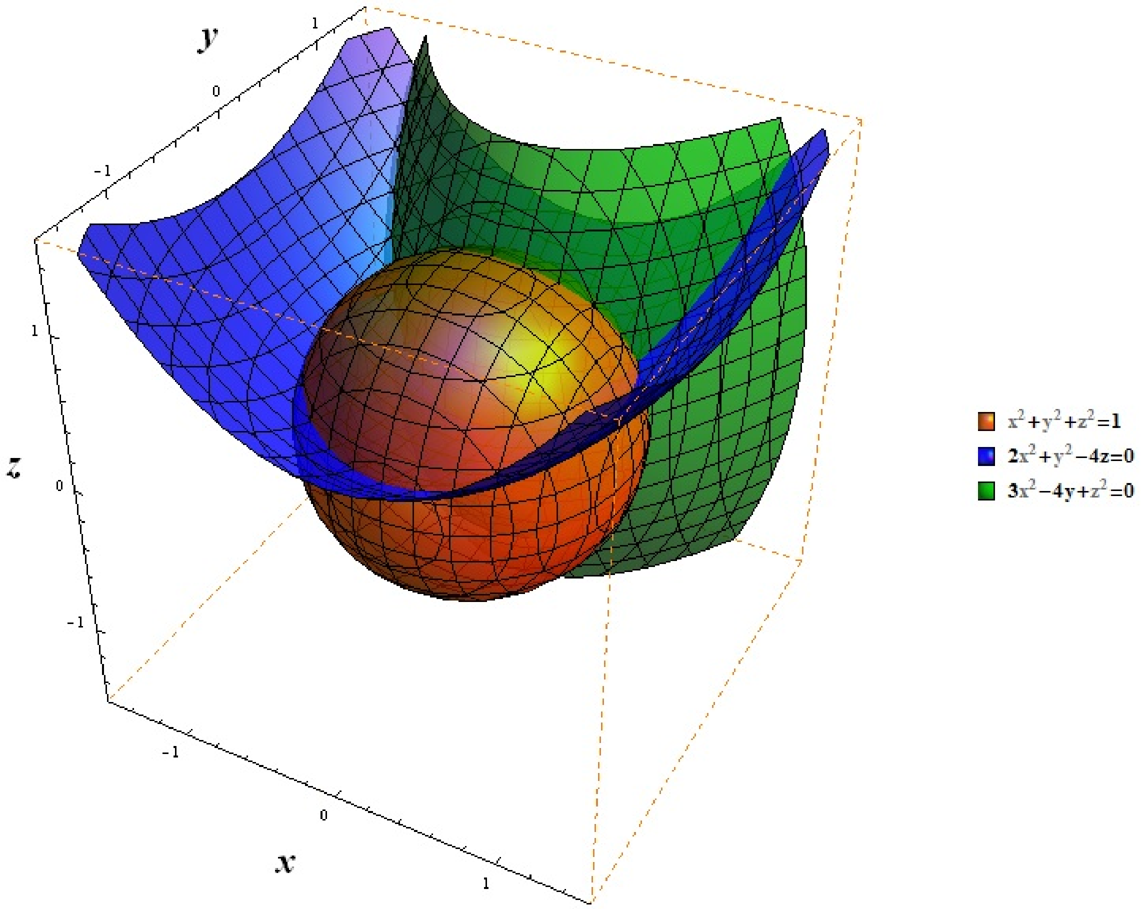

Figura 5 – Visualização gráfica do problema.

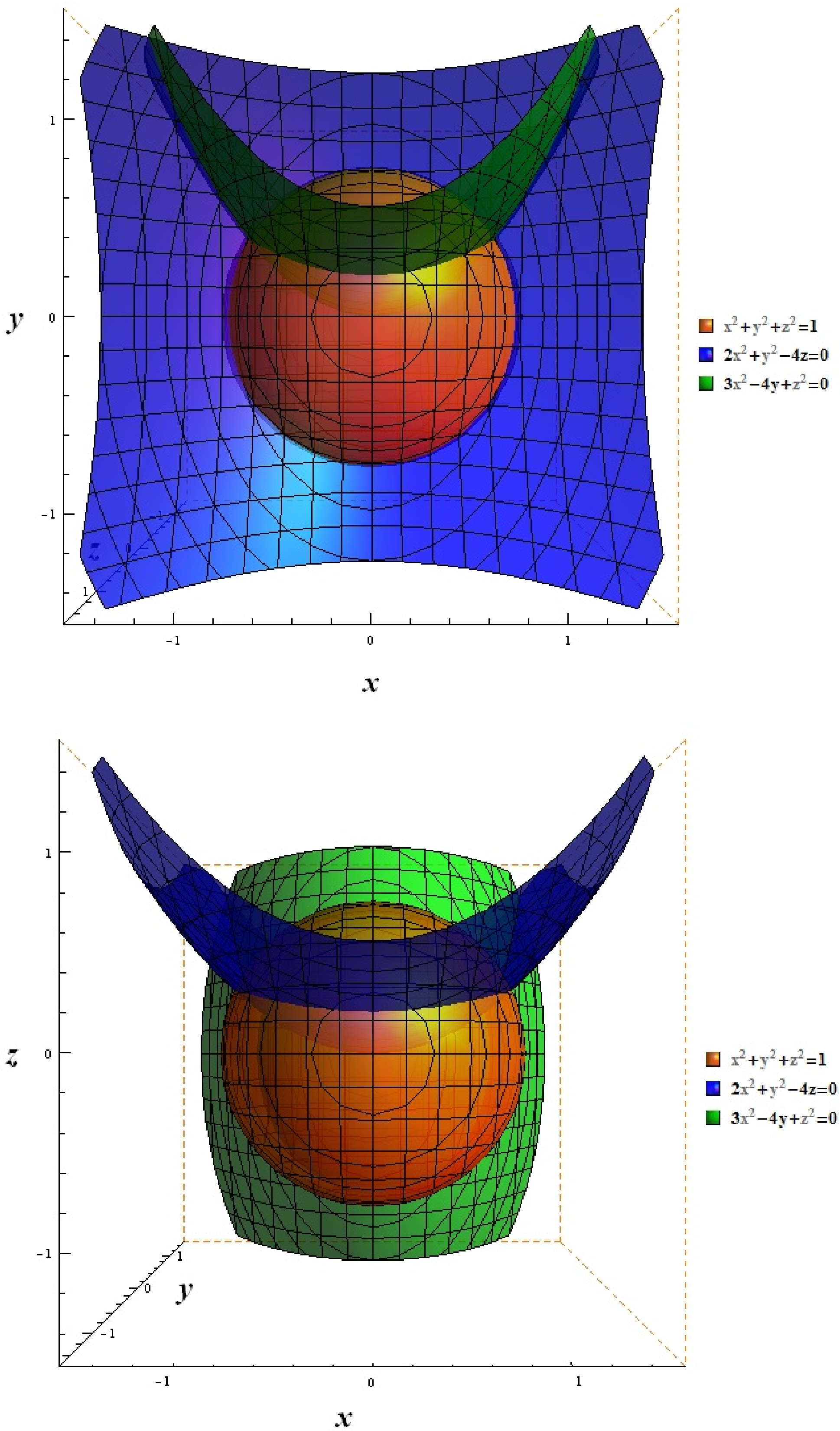

Figura 6 – Vistas auxiliares para determinação da aproximação inicial.

Pelos gráficos podemos verificar que temos raízes próximas aos pontos (0.6,0.4,0.4) e (-0.6,0.4,0.4).

Faremos a aproximação da solução para a raiz próxima ao ponto (0.6,0.4,0.4).

Segue a implementação do algoritmo utilizando Python e SymPy.

classoffset

```
# -*- coding: utf-8 -*-
"""
Created on Mon Sep 28 16:02:46 2015

@author: Kleber
"""

from sympy import *
init_printing(use_unicode=True)

x, y, z = symbols("x y z") # Define simbolos para variaveis genericas
f, g, h = map(Function, 'fgh') # define simbolos para funcoes

R=[]
R0=[]

# Funcoes
f=x**2+y**2+z**2-1
g=2*x**2+y**2-4*z
h=3*x**2-4*y+z**2

# Calculo das aproximacoes
def MN(x0,y0,z0):
    for i in range(0,4):
        X0=Matrix([x0,y0,z0])
        F=Matrix([f,g,h])
        F0=F.subs([(x,x0),(y,y0),(z,z0)])
        J=F.jacobian([x,y,z])
        J0=J.subs([(x,x0),(y,y0),(z,z0)])
        W=J0.inv()
        X1=X0-W*F0
        x0=X1[0]
        y0=X1[1]
        z0=X1[2]
        R.append(X1)
        R0.append(F0)
    return R[2],R0[3]
```

Obtivemos os seguintes resultados:

$$x_0 = \begin{pmatrix} 0.600000000000000 \\ 0.400000000000000 \\ 0.400000000000000 \end{pmatrix} \qquad F(x_0) = \begin{pmatrix} -0.320000000000000 \\ -0.720000000000000 \\ -0.360000000000000 \end{pmatrix}$$

$$x_1 = \begin{pmatrix} 0.818518518518518 \\ 0.500925925925926 \\ 0.371296296296296 \end{pmatrix} \qquad F(x_1) = \begin{pmatrix} 0.0587602880658435 \\ 0.105686728395062 \\ 0.144074931412893 \end{pmatrix}$$

$$x_2 = \begin{pmatrix} 0.785881651752030 \\ 0.496619345757194 \\ 0.369925459730971 \end{pmatrix} \qquad F(x_2) = \begin{pmatrix} 0.00108559089797203 \\ 0.00214887677741671 \\ 0.00319737440988979 \end{pmatrix}$$

$$x_3 = \begin{pmatrix} 0.785197231374475 \\ 0.496611392971886 \\ 0.369922830755001 \end{pmatrix} \qquad F(x_3) = \begin{pmatrix} 0.000000468501411 \\ 0.000000936925753 \\ 0.000001405300671 \end{pmatrix}$$

Como esperado a medida que aumenta o número de iterações $F(x_n) \longrightarrow 0$.

Infelizmente, achar um conjunto de boas condições iniciais, ou valores adequados para a aproximação inicial, não é um objetivo trivial. E não trataremos deste problema aqui.

Exemplo 3.4. Como exemplo podemos citar também um sistema de equações não lineares que tem origem em um problema clássico de geometria da antiguidade. Que é o problema da duplicação do cubo ou problema de Delos. Que consiste em obter um método para, dada a aresta de um cubo, construir com régua e compasso a aresta do cubo cujo volume é o dobro do cubo inicial. Se a é o lado de um cubo e k o lado de um cubo que estamos procurando, os respectivos volumes dos dois cubos são a^3 e k^3. Consequentemente somos confrontados com o problema de achar, quando o segmento a é dado, um segundo segmento k tal que

$$k^3 = 2a^3.$$

A partir dos trabalhos de Ruffini, Abel e Galois, no século XIX, demonstrou-se que é impossível resolver este problema usando apenas régua e compasso.

Várias soluções foram dadas a este problema, uma das mais impressionantes, principalmente quando levamos em consideração a época em que foi proposta, foi a solução tridimensional dada pelo matemático grego Arquitas (c. 400a.c.). Usando a linguagem moderna da geometria analítica a solução pode ser descrita da seguinte forma. Seja a o lado do cubo a ser duplicado, e seja o ponto $(a, 0, 0)$ o centro de três circunferências de raio a mutuamente perpendiculares e cada uma sobre um plano perpendicular a um eixo coordenado.

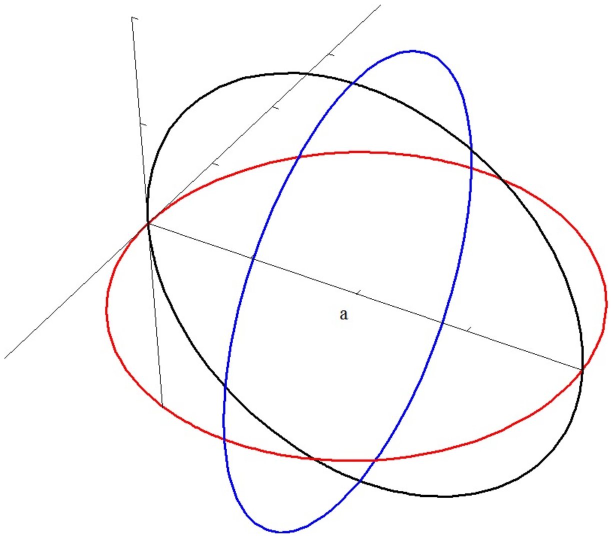

Figura 7 – Visualização das circunferências.

Sobre a circunferência perpendicular ao eixo Ox, construa um cone circular reto com vértice $(0, 0, 0)$; sobre a circunferência no plano xy, construa um cilindro circular reto; e girando a circunferência do plano xz sobre o eixo z construa um toro. As equações destas superfícies são respectivamente, $x^2 = y^2 + z^2$, $2ax = x^2 + y^2$ e $(x^2 + y^2 + z^2)^2 = 4a^2(x^2 + y^2)$.

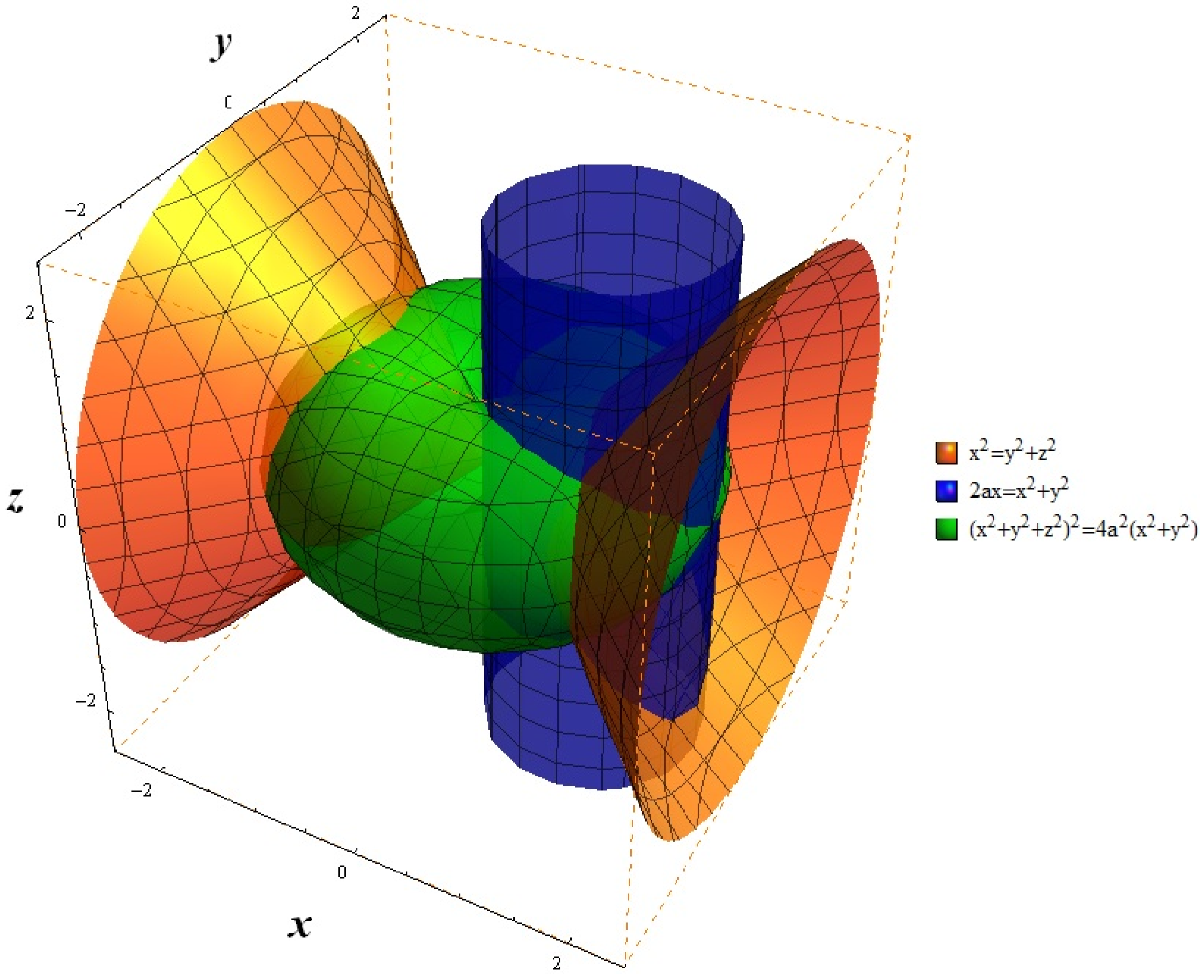

Figura 8 – Visualização das superfícies.

Essas três superfícies se intersectam em um ponto cuja coordenada x é $a\sqrt[3]{2}$; consequentemente o comprimento deste segmento de reta é o lado do cubo desejado.

Considerando $a = 1$ obtemos o seguinte sistema não linear

$$\begin{cases} x^2 = y^2 + z^2 \\ 2x = x^2 + y^2 \\ (x^2 + y^2 + z^2)^2 = 4(x^2 + y^2) \end{cases}$$

Novamente utilizaremos o gráfico das superfícies para determinarmos o valor da aproximação inicial, que neste caso será $(1, 1, 1)$. Aplicando o método de Newton obtemos a seguinte solução aproximada para uma das raízes.

$$x_0 = \begin{pmatrix} 1.00000000000000 \\ 1.00000000000000 \\ 1.00000000000000 \end{pmatrix} \qquad F(x_0) = \begin{pmatrix} 1.00000000000000 \\ 0.00000000000000 \\ 1.00000000000000 \end{pmatrix}$$

$$x_1 = \begin{pmatrix} 1.31250000000000 \\ 1.00000000000000 \\ 0.81250000000000 \end{pmatrix} \qquad F(x_1) = \begin{pmatrix} -0.06250000000000 \\ 0.09765625000000 \\ 0.55279541015625 \end{pmatrix}$$

$$x_2 = \begin{pmatrix} 1.26285264371895 \\ 0.96668667383783 \\ 0.81176297974559 \end{pmatrix} \qquad F(x_2) = \begin{pmatrix} -0.00135453908685 \\ 0.00357463768568 \\ 0.04774860582349 \end{pmatrix}$$

$$x_3 = \begin{pmatrix} 1.25993303466525 \\ 0.96563163510392 \\ 0.80931167975550 \end{pmatrix} \qquad F(x_3) = \begin{pmatrix} -0.00000140213865 \\ 0.00000963722376 \\ 0.00024018787338 \end{pmatrix}$$

$$x_4 = \begin{pmatrix} 1.25992105011590 \\ 0.96562987104449 \\ 0.80929599330363 \end{pmatrix} \qquad F(x_4) = \begin{pmatrix} 0.00000000010555 \\ 0.00000000014674 \\ 0.00000000538782 \end{pmatrix}$$

A solução aparentemente converge para a solução prevista por Arquitas quando $n \longrightarrow \infty$.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] ATKINSON, Kendall; HAN, Weimin. **Theoretical Numerical Analysis: A Functional Analysis Framework**. New York: Springer-Verlag, 2009.

[2] CVETKOVSKI, Zdravko. **Inequalities: Theorems, Techniques and Selected Problems**. Berlin: Springer-Verlag, 2012.

[3] DOMINGUES, Hygino Hugueros. **Espaços métricos e introdução a topologia**. São Paulo, SP: Atual Ed.: Editora da USP, 1982.

[4] HÄMMERLIN, Günther; HOFFMANN, Karl-Heinz. **Numerical Mathematics**. New York: Springer-Verlag, 1991.

[5] KREYSZIG, Erwin. **Introductory Functional Analysis with Applications**. New York: Wiley, 1989.

[6] LIMA, Elon L. **Análise Real - Volume 1**. 10. ed. Rio de Janeiro: IMPA, 2009.

[7] LIMA, Elon L. **Curso de Análise - Volume 1**. 11. ed. Rio de Janeiro: IMPA, 2004.

[8] LIMA, Elon L. **Espaços Métricos**. 4. ed. Rio de Janeiro: IMPA, 2011.

[9] MACHADO, Luciana B. **Análise Funcional e suas Aplicações**. Dissertação (Mestrado em Matemática Universitária). IGCE-Unesp-RioClaro, 2012.

[10] MUSCAT, Joseph. **Functional Analysis: An Introduction to Metric Spaces, Hilbert Spaces, and Banach Algebras**. Springer International Publishing, 2014.

A TEOREMAS DE ANÁLISE EM $\mathbb{R}$

Neste capítulo relacionaremos os principais teoremas de Análise em $\mathbb{R}$ que foram estudados durante nossa pesquisa e que são importantes para compreensão da teoria tratada neste trabalho. Os teoremas e definições, assim como as demonstrações que se seguem, foram retirados das referências [6] e [7].

A.1 SUPREMO E ÍNFIMO

Um conjunto $X \subset \mathbb{R}$ diz-se limitado superiormente quando existe algum $b \in \mathbb{R}$ tal que $x \leqslant b$ para todo $x \in X$. Neste caso diz-se que b é cota superior de X. O conjunto $X \subset \mathbb{R}$ é limitado inferiormente quando existe $a \in \mathbb{R}$ tal que $a \leqslant x$ para todo $x \in X$. Neste caso diz-se que a é cota inferior de X. Se o conjunto X é limitado superior e inferiormente, diz-se que X é um conjunto limitado.

Seja $X \neq \emptyset \subset \mathbb{R}$ limitado superiormente. Um número $b \in \mathbb{R}$ chama-se o supremo do conjunto X e denota-se $b = supX$, quando este é a menor das cotas superiores de X. Isto é, cumpre as seguintes condições:

S1. Para todo $x \in X$, tem-se $x \leqslant b$;

S2. Se $c \in \mathbb{R}$ é tal que $x \leqslant c$ para todo $x \subset X$ então $b \leqslant c$.

A condição S2 admite a seguinte reformulação:

S2'. Para todo $\varepsilon > 0$ existe $x \in X$ tal que $b - \varepsilon < x$.

Analogamente, seja $X \neq \emptyset \subset \mathbb{R}$ limitado inferiormente. Um número $a \in \mathbb{R}$ chama-se o ínfimo do conjunto X e denota-se $a = infX$, quando este é a maior das cotas inferiores de X. Ou seja, cumpres as seguintes condições:

I1. Para todo $x \in X$, tem-se $x \geqslant a$;

I2. Se $c \in \mathbb{R}$ é tal que $x \geqslant c$ para todo $x \subset X$ então $a \geqslant c$.

A condição I2 admite a seguinte reformulação:

I2'. Para todo $\varepsilon > 0$ existe $x \in X$ tal que $x < a + \varepsilon$.

Exemplo: Dado o conjunto $X = \{1/n; n \in \mathbb{N}\}$ temos que $0 = infX$ e $1 = supX$.

A.2 SEQUÊNCIAS DE NÚMEROS REAIS

Uma sequência de números reais é uma função $x : \mathbb{N} \longrightarrow \mathbb{R}$, que associa a cada número natural n um número real x_n, chamado o n-ésimo termo da sequência. Escrevemos

$(x_n)_{n\in\mathbb{N}}$ ou (x_n) para indicar a sequência x.

Uma sequência $(x_n)_{n\in\mathbb{N}}$ é limitada superiormente quando existe um número $b \in \mathbb{R}$ tal que $x_n \leqslant b$ para todo $n \in \mathbb{N}$. Analogamente, diz-se que $(x_n)_{n\in\mathbb{N}}$ é limitada inferiormente quando existe $a \in \mathbb{R}$ tal que $x_n \geqslant a$ para todo $n \in \mathbb{N}$. Uma sequência é limitada se,e somente se,é limitada superior e inferiormente.

Dada uma sequência $x = (x_n)_{n\in\mathbb{N}}$ de números reais. Uma subsequência de x é a restrição da função x a um subconjunto infinito $\mathbb{N}' = \{n_1 < n_2 < \cdots < n_k < \cdots\}$ de $\mathbb{N}$. Escreve-se $x' = (x_n)_{n\in\mathbb{N}'}$ ou $(x_{n_k})_{k\in\mathbb{N}}$ para indicar a subsequência $x' = x|\mathbb{N}'$.

Um número real a é limite da sequência $(x_n)_{n\in\mathbb{N}}$ quando, para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que todos os termos x_n com índice $n > n_0$ cumprem a condição $|x_n - a| < \varepsilon$. Simbolicamente:

$$a = \lim x_n \equiv \forall\ \varepsilon > 0\ \exists\ n_0 \in \mathbb{N}; n > n_0 \Rightarrow |x_n - a| < \varepsilon$$

Teorema A.1. *Se* $\lim x_n = a$ *e* $\lim x_n = b$ *então* $a = b$*.*

Demonstração. Seja $\lim x_n = a$. Dado qualquer número real $b \neq a$ podemos tomar $\varepsilon > 0$ tal que os intervalos abertos $I = (a-\varepsilon, a+\varepsilon)$ e $J = (b-\varepsilon, b+\varepsilon)$ sejam disjuntos. Existe $n_0 \in \mathbb{N}$ tal que $n > n_0$ implica $x_n \in I$. Então, para todo $n > n_0$, temos $x_n \notin J$. Logo não é $\lim x_n = b$.

□

Teorema A.2. *Se* $\lim x_n = a$ *então toda subsequência de* $(x_n)_{n\in\mathbb{N}}$ *converge para o limite* a*.*

Demonstração. Seja $(x_{n_1}, \ldots, x_{n_k}, \ldots)$ uma subsequência qualquer de $(x_n)_{n\in\mathbb{N}}$. Dado qualquer intervalo aberto I de centro a, existe $n_0 \in \mathbb{N}$ tal que todos termos x_n, com $n > n_0$, pertencem a I. Em particular, todos os termos x_{n_k}, com $n_k > n_0$ também pertencem a I. Logo, $\lim x_{n_k} = a$. □

Teorema A.3. *Toda sequência convergente é limitada.*

Demonstração. Se $a = \lim x_n$. Tomando $\varepsilon = 1$, vemos que existe $n_0 \in \mathbb{N}$ tal que $n > n_0 \Rightarrow x_n \in (a-1, a+1)$. Sejam b o menor e c o maior elemento do conjunto finito $\{x_1, \ldots, x_{n_0}, a-1, a+1\}$. Todos os termos x_n da sequência estão contidos no intervalo $[b, c]$, logo ela é limitada. □

Teorema A.4. *Toda sequência monótona limitada é convergente.*

Demonstração. Sem perda de generalidade, seja $(x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n \leqslant \ldots)$ uma sequência não-decrescente limitada. Tomemos $a = \sup\{x_n; n = 1, 2, \ldots\}$. Afirmamos que $a = \lim x_n$.

Com efeito, dado qualquer $\varepsilon > 0$, como $a - \varepsilon < a$, o número $a - \varepsilon$ não é cota superior do conjunto dos x_n. Logo existe algum $n_0 \in \mathbb{N}$ tal que $a - \varepsilon < x_n$. Como a sequência é monótona, $n > n_0 \Rightarrow x_{n_0} \leqslant x_n$ e, portanto, $a - \varepsilon < x_n$. Como $x_n < a$ para todo n, vemos que $n > n_0 \Rightarrow a - \varepsilon < x_n < a + \varepsilon \Rightarrow \lim x_n = a$. □

Corolário A.1. *(Teorema de Bolzano-Weierstrass) Toda sequência limitada de números reais possui uma subsequência convergente.*

Demonstração. Com efeito, basta mostrar que toda sequência (x_n) possui uma subsequência monótona. Diremos que um termo x_n da sequência dada é *destacado* quando $x_n \geqslant x_p$ para todo $p > n$. Seja $D \subset \mathbb{N}$ o conjunto dos índices n tais que x_n é um termo destacado. Se D for um conjunto infinito, $D = \{n_1 < n_2 < \cdots < n_k < \cdots\}$, então a subsequência $(x_n)_{n \in D}$ será monótona não-crescente. Se, entretanto, D for finito seja $n_1 \in \mathbb{N}$ maior do que todos os $n \in D$. Então x_{n_1} não é destacado, logo existe $n_2 > n_1$ com $x_{n_1} < x_{n_2}$. Por sua vez, x_{n_2} não é destacado, logo existe $n_3 > n_2$ com $x_{n_1} < x_{n_2} < x_{n_3}$. Prosseguindo, obtemos uma subsequência crescente $x_{n_1} < x_{n_2} < \cdots < x_{n_k} < \cdots$. □

Seja $(x_n)_{n \in \mathbb{N}}$ uma sequência de números reais. $(x_n)_{n \in \mathbb{N}}$ é uma *sequência de Cauchy* quando cumpre a seguinte condição:

- dado arbitrariamente um número real $\varepsilon > 0$, pode se obter $n_0 \in \mathbb{N}$ tal que $m > n_0$ e $n > n_0$ implica $|x_m - x_n| < \varepsilon$.

Teorema A.5. *Toda sequência convergente é de Cauchy.*

Demonstração. Seja $\lim x_n = a$. Dado arbitrariamente $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $m > n_0 \Rightarrow |x_m - a| < \dfrac{\varepsilon}{2}$ e $n > n_0 \Rightarrow |x_n - a| < \dfrac{\varepsilon}{2}$. Logo $m, n > n_0 \Rightarrow |x_m - x_n| \leqslant |x_m - a| + |x_n - a| < \dfrac{\varepsilon}{2} + \dfrac{\varepsilon}{2} = \varepsilon$. Portanto, (x_n) é sequência de Cauchy. □

A.3 SÉRIES DE NÚMEROS REAIS

Uma série é uma soma $s = a_1 + a_2 + \cdots + a_n + \cdots$ com um número infinito de parcelas.

Seja $(a_n)_{n \in \mathbb{N}}$ uma sequência de números reais. Seja a sequência $(s_n)_{n \in \mathbb{N}}$ cujos os elementos são as somas

$$s_1 = a_1, s_2 = a_1 + a_2, \ldots, s_n = a_1 + a_2 + \cdots + a_n$$

chamadas as reduzidas da série $\displaystyle\sum_{n=1}^{\infty} a_n$. A parcela a_n é chamada o termo geral da série.

Se existir o limite

$$s = \lim s_n = \lim_{n\to\infty} (a_1 + a_2 + \cdots + a_n),$$

diremos que a série $\sum_{n=1}^{\infty} a_n$ é convergente e o limite s será chamado a soma da série. Escreveremos então

$$s = \sum_{n=1}^{\infty} a_n = a_1 + a_2 + \cdots + a_n + \cdots$$

Se a sequência das reduzidas não convergir, diremos que a série $\sum_{n=1}^{\infty} a_n$ é divergente.

Teorema A.6. *Se* $\sum_{n=1}^{\infty} a_n$ *é uma série convergente então* $\lim a_n = 0$.

Demonstração. Seja $s_n = a_1 + \cdots + a_n$. Então, existe $s = \lim_{n\to\infty} s_n$. Evidentemente, tem-se também $s = \lim_{n\to\infty} s_{n-1}$. Logo $0 = s - s = \lim s_n - \lim s_{n-1} = \lim(s_n - s_{n-1}) = \lim a_n$. $\square$

Teorema A.7. *(Critério de comparação) Sejam* $\sum_{n=1}^{\infty} a_n$ *e* $\sum_{n=1}^{\infty} b_n$ *séries de termos não-negativos. Se existem* $c > 0$ *e* $n_0 \in \mathbb{N}$ *tais que* $a_n \leqslant cb_n$ *para todo* $n > n_0$ *então a convergência de* $\sum_{n=1}^{\infty} b_n$ *implica a de* $\sum_{n=1}^{\infty} a_n$ *enquanto a divergência de* $\sum_{n=1}^{\infty} a_n$ *implica a de* $\sum_{n=1}^{\infty} b_n$.

Demonstração. Sem perda de generalidade, podemos supor $a_n \leqslant cb_n$ para todo $n \in \mathbb{N}$. Então as reduzidas s_n e t_n, de $\sum_{n=1}^{\infty} a_n$ e $\sum_{n=1}^{\infty} b_n$ respectivamente, formam sequências não-decrescentes tais que $s_n \leqslant ct_n$ para todo $n \in \mathbb{N}$. Como $c > 0$, $(t_n)_{n\in\mathbb{N}}$ limitada implica $(s_n)_{n\in\mathbb{N}}$ limitada e $(s_n)_{n\in\mathbb{N}}$ ilimitada implica $(t_n)_{n\in\mathbb{N}}$ ilimitada, pois $t_n \geqslant s_n/c$. $\square$

A.4 TOPOLOGIA NA RETA

Definição A.1. *(ponto interior) Um ponto* a *é interior ao conjunto* $X \subset \mathbb{R}$ *quando existe um número real* $\varepsilon > 0$ *tal que o intervalo aberto* $(a - \varepsilon, a + \varepsilon)$ *está contido em* X.

Definição A.2. *(ponto aderente) Um ponto* a *é aderente ao conjunto* $X \subset \mathbb{R}$ *quando* a *é limite de alguma sequência de pontos* $x_n \in X$.

Teorema A.8. *Um ponto* $a \in \mathbb{R}$ *é aderente a um conjunto* $X \subset \mathbb{R}$ *se, e somente se, para todo* $\varepsilon > 0$ *tem se* $X \cap (a - \varepsilon, a + \varepsilon) \neq \emptyset$

Demonstração. Seja a aderente a X. Então, $a = \lim x_n$ onde $x_n \in X$ para todo $n \in \mathbb{N}$. Dada uma vizinhança qualquer V do ponto a temos $x_n \in V$ para todo n suficientemente grande (pela definição de limite). Logo, $V \cap X \neq \varnothing$. Reciprocamente, se toda vizinhança de a contém pontos de X podemos escolher, em cada intervalo $(a - 1/n, a + 1/n), n \in \mathbb{N}$, um ponto $x_n \in X$. Então $|x_n - a| < 1/n$. Logo, $\lim x_n = a$ e a é aderente a X. □

Definição A.3. *(ponto de acumulação) Seja $X \subset \mathbb{R}$. Um número $a \in \mathbb{R}$ chama-se ponto de acumulação do conjunto X quando todo intervalo aberto $(a - \varepsilon, a + \varepsilon)$, de centro a, contém algum ponto $x \in X$ diferente de a. Simbolicamente:*

$$\forall \varepsilon > 0 \; \exists \; x \in X; 0 < |x - a| < \varepsilon$$

.

Definição A.4. *(ponto isolado) Um ponto $a \in X$ que não é ponto de acumulação de X chama-se um ponto isolado de X. Se, e somente se, existir $\varepsilon > 0$ tal que $(a-\varepsilon, a+\varepsilon) \cap X = \{a\}$.*

O conjunto dos pontos interiores a X chama-se o interior do conjunto X. Notação $intX$. Quando $a \in intX$ então X é uma vizinhança do ponto a. Exemplo: $int[a, b] = (a, b)$.

O conjunto dos pontos aderentes a X chama-se o fecho do conjunto X. Notação $\overline{X}$. Exemplo: todo ponto $a \in X$ é aderente a X, basta tomar $x_n = a$ para todo $n \in \mathbb{N}$.

O conjunto dos pontos de acumulação de X é representado pela notação X'. Exemplo: seja $X = \{1, 1/2, 1/3, \ldots, 1/n, \ldots\}$ então $X' = \{0\}$.

Definição A.5. *(conjunto aberto) Um conjunto $A \subset \mathbb{R}$ chama-se aberto quando $A = intA$.*

Definição A.6. *(conjunto fechado) Um conjunto $A \subset \mathbb{R}$ chama-se fechado quando $A = \overline{A}$.*

Teorema A.9. *Um conjunto $F \subset \mathbb{R}$ é fechado se, e somente se, seu complementar $\mathbb{R} - F$ é aberto.*

Demonstração. Basta observar que cada uma das afirmações abaixo é equivalente à seguinte:

1. F é fechado;
2. todo ponto aderente a F pertence a F;
3. se $a \in \mathbb{R} - F$ então a não é aderente a F;
4. se $a \in \mathbb{R} - F$ então existe um intervalo aberto I tal que $a \in I$ e $I \cap F = \emptyset$;
5. se $a \in \mathbb{R} - F$, então existe um intervalo aberto I tal que $a \in I \subset \mathbb{R} - F$;
6. todo ponto $a \in \mathbb{R} - F$ é interior a $\mathbb{R} - F$;
7. $\mathbb{R} - F$ é aberto. □

Definição A.7. *(conjunto denso) Um conjunto* $A \subset \mathbb{R}$ *chama-se denso em* $\mathbb{R}$ *quando todo intervalo aberto* (a, b) *contém algum ponto de* A*. Ou ainda, um conjunto* A *é denso em* B *quando* $B \subset \overline{A}$*, isto é, quando todo* $b \in B$ *é aderente a* A*.*

Definição A.8. *(conjunto compacto) Um conjunto* $X \subset \mathbb{R}$ *chama-se compacto quando é limitado e fechado.*

Teorema A.10. *Um conjunto* $X \subset \mathbb{R}$ *é compacto se, e somente se, toda sequência de pontos em* X *possui uma subsequência que converge para um ponto de* X*.*

Demonstração. Se $X \subset \mathbb{R}$ é compacto, toda sequência de pontos de X é limitada. Pelo teorema de Bolzano-Wierstrass possui uma subsequência convergente, cujo limite é um ponto de X, pois X é fechado. Reciprocamente, seja $X \in \mathbb{R}$ um conjunto tal que toda sequência de pontos $x_n \in X$ possui uma subsequência que converge para um ponto de X. Então, X é limitado pois, do contrário, para cada $n \in \mathbb{N}$ poderíamos encontrar $x_n \in X$ com $|x_n| > n$. A sequência $(x_n)_{n \in \mathbb{N}}$, assim obtida, não possuiria subsequência limitada. Logo, não teria subsequência convergente. Além disso, X é fechado pois do contrário existiria um ponto $a \notin X$ com $a = \lim x_n$, onde cada $x_n \in X$. A sequência $(x_n)_{n \in \mathbb{N}}$ não possuiria então subsequência alguma convergindo para um ponto de X pois todas suas subsequências teriam limite a. Logo X é compacto. □

A.5 FUNÇÃO CONTÍNUA

Definição A.9. *Uma função* $f : X \longrightarrow \mathbb{R}$*, definida no conjunto* $X \in \mathbb{R}$*, diz-se contínua no ponto* $a \in X$ *quando para todo* $\varepsilon > 0$ *dado arbitrariamente, pode-se obter* $\delta > 0$ *tal que* $x \in X$ *satisfazendo* $|x - a| < \delta$ *temos* $|f(x) - f(a)| < \varepsilon$*.*

$$\forall \varepsilon > 0 \exists \delta > 0; x \in X, |x - a| < \delta \ temos \ |f(x) - f(a)| < \varepsilon.$$

Diz-se que $f : X \longrightarrow \mathbb{R}$ é uma função contínua quando f é contínua em todos os pontos $a \in X$.

Observações:

1) Se a é um ponto isolado de X então toda função $f : X \longrightarrow \mathbb{R}$ é contínua no ponto a.

2) Se $a \in X$ é um ponto de acumulação de X. Então $f : X \longrightarrow \mathbb{R}$ é contínua no ponto a se, e somente se, $\lim\limits_{x \to a} f(x) = f(a)$.

Teorema A.11. *Sejam* $f, g : X \longrightarrow \mathbb{R}$ *contínua no ponto* $a \in X$*, com* $f(a) < g(a)$*. Existe* $\delta > 0$ *tal que* $f(x) < g(x)$ *para todo* $x \in X \cap (a - \delta, a + \delta)$*.*

Demonstração. Tomemos $c = [g(a) + f(a)]/2$ e $\varepsilon = g(a) - c = c - f(a)$. Então pela definição de função contínua, existem $\delta_1 > 0$ e $\delta_2 > 0$ tais que $x \in X, |x - a| < \delta_1 \Rightarrow f(a) - \varepsilon < f(x) < c$ e $x \in X, |x - a| < \delta_2 \Rightarrow c < g(x) < g(a) + \varepsilon$. Seja $\delta = min\{\delta_1, \delta_2\}$. Então $x \in X, |x - a| < \delta \Rightarrow f(x) < c < g(x)$. $\square$

Corolário A.2. *Seja $f : X \longrightarrow \mathbb{R}$ contínua no ponto $a \in X$. Se $f(a) \neq 0$, existe $\delta > 0$ tal que, para todo $x \in X \cap (a - \delta, a + \delta)$, $f(x)$ tem o mesmo sinal de $f(a)$.*

Teorema A.12. *Sejam $f : X \longrightarrow \mathbb{R}$ contínua no ponto $a \in X$, $g : Y \longrightarrow \mathbb{R}$ contínua no ponto $b = f(a) \in Y$ e $f(X) \subset Y$ de modo que a composta $g \circ f : X \longrightarrow \mathbb{R}$ está bem definida. Então $g \circ f$ é contínua no ponto a.*

Demonstração. Dado $\varepsilon > 0$ existe, pela continuidade de g no ponto b, um número $\eta > 0$ tal que $y \in Y$, $|g(y) - g(b)| < \varepsilon$. Por sua vez, a continuidade de f no ponto a assegura que existe $\delta > 0$ tal que $x \in X$, $|x - a| < \delta$ implicam $|f(x) - b| < \eta$. Consequentemente, $x \in X \cap (a - \delta, a + \delta) \Rightarrow |g(f(x)) - g(b)| = |(g \circ f)(x) - (g \circ f)(a)| < \varepsilon$. $\square$

Teorema A.13. *Seja $f : X \longrightarrow \mathbb{R}$ contínua. Se X é compacto então $f(X)$ é compacto.*

Demonstração. Seja $(y_n)_{n \in \mathbb{N}}$ uma sequência em $f(X)$. Para cada $n \in \mathbb{N}$, existe $x_n \in X$ tal que $f(x_n) = y_n$. Como X é compacto, podemos obter uma subsequência convergente $x_{n_k} \to x \in X$. Sendo f contínua, temos $\lim y_{n_k} = \lim f(x_{n_k}) = f(x) = y \in f(X)$. Logo $f(X)$ é compacto. $\square$

Corolário A.3. *(Weierstrass) Toda função contínua $f : X \longrightarrow \mathbb{R}$ definida num compacto X é limitada e atinge seus extremos, isto é, existem x_1, x_2 pertencentes à X tais que $f(x_1) \leqslant f(x) \leqslant f(x_2)$ para todo $x \in X$.*

Demonstração. Com efeito $f(X)$, sendo compacto, é limitado e fechado. Logo, $\sup f(X) \in f(X)$ e $\inf f(X) \in f(X)$. Portanto, existem x_1, x_2 pertencentes à X tais que $\inf f(X) = f(x_1)$ e $\sup f(X) = f(x_2)$. $\square$

A.6 CONTINUIDADE UNIFORME

Definição A.10. *Uma função $f : X \longrightarrow \mathbb{R}$ diz-se uniformemente contínua no conjunto X quando, para todo $\varepsilon > 0$ dado arbitrariamente, pode -se obter $\delta > 0$ tal que x, y pertencentes à $X, |y - x| < \delta \Rightarrow |f(y) - f(x)| < \varepsilon$.*

Uma função uniformemente contínua $f : X \longrightarrow \mathbb{R}$ é contínua em todos os pontos do conjunto X. A recíproca é falsa.

Teorema A.14. *A fim de que $f : X \longrightarrow \mathbb{R}$ seja uniformemente contínua é necessário e suficiente que, para todo par de sequências $(x_n), (y_n)$ em X com $\lim(y_n - x_n) = 0$, tenha-se $\lim[f(y_n) - f(x_n)] = 0$.*

Demonstração. Se f é uniformemente contínua e $\lim(y_n - x_n) = 0$ então, dado arbitrariamente $\varepsilon > 0$, existe $\delta > 0$ tal que x, y pertencentes à $X, |y - x| < \delta$ implicam $|f(y) - f(x)| < \varepsilon$. Existe também $n_0 \in \mathbb{N}$ tal que $n > n_0$ implica $|y_n - x_n| < \delta$. Logo, $n > n_0$ implica $|f(y_n) - f(x_n)| < \varepsilon$ e daí $\lim[f(y_n) - f(x_n)] = 0$. Reciprocamente, suponhamos válida a condição estipulada no enunciado do teorema. Se f não fosse uniformemente contínua, existiria um $\varepsilon > 0$ com a seguinte propriedade: para todo $n \in \mathbb{N}$ poderíamos achar pontos x_n, y_n em X tais que $|y_n - x_n| < 1/n$ e $|f(y_n) - f(x_n)| \geqslant \varepsilon$. Então teríamos $\lim(y_n - x_n) = 0$ sem que fosse $\lim[f(y_n) - f(x_n)] = 0$. Esta contradição conclui a prova do teorema. □

Teorema A.15. *Seja $X \subset \mathbb{R}$ compacto. Toda função contínua $f : X \longrightarrow \mathbb{R}$ é uniformemente contínua.*

Demonstração. Se f não fosse uniformemente contínua, existiriam $\varepsilon > 0$ e duas sequências $(x_n), (y_n)$ em X satisfazendo $\lim(y_n - x_n) = 0$ e $|f(y_n) - f(x_n)| \geqslant \varepsilon$ para todo $n \in \mathbb{N}$. Passando a uma subsequência, se necessário, podemos supor, em virtude da compacidade de X, que $\lim x_n = a \in X$. Então, como $y_n = (y_n - x_n) + x_n$, vale também $\lim y_n = a$. Sendo f contínua no ponto a, temos $\lim[f(y_n) - f(x_n)] = \lim f(y_n) - \lim f(x_n) = f(a) - f(a) = 0$, contradizendo que seja $|f(y_n) - f(x_n)| \geqslant \varepsilon$ □

Definição A.11. *Uma função $f : X \longrightarrow \mathbb{R}$ chama-se lipschitziana quando existe uma constante $k > 0$ (chamada constante de Lipschitz da função f) tal que $|f(x) - f(y)| \leqslant k|x - y|$ sejam quais forem x, y pertencentes à X.*

A fim de que $f : X \longrightarrow \mathbb{R}$ seja lipschitziana é necessário e suficiente que o quociente $[f(y) - f(x)]/(y - x)$ seja limitado, isto é, que exista uma constante $k > 0$ tal que $x, y \in X, x \neq y \Rightarrow |f(y) - f(x)|/|y - x| \leqslant k$. Toda função lipschitziana $f : X \longrightarrow \mathbb{R}$ é uniformemente contínua pois dado $\varepsilon > 0$, tome-se $\delta = \varepsilon/k$. Então $x, y \in X, |x - y| < \delta \Rightarrow |f(y) - f(x)| \leqslant k|x - y| < k \cdot \varepsilon/k = \varepsilon$.

A.7 DERIVADA

Definição A.12. *Sejam $f : X \longrightarrow \mathbb{R}$ e $a \in X \cap X'$. A derivada da função f no ponto a é o limite*

$$f'(a) = \lim_{x \longrightarrow a} \frac{f(x) - f(a)}{dx - a} = \lim_{h \longrightarrow 0} \frac{f(a + h) - f(a)}{h}.$$

Se o limite existir diz-se que f é derivável no ponto a. Quando existe a derivada $f'(x)$ em todos os pontos $x \in X \cap X'$ diz-se que a função $f : X \longrightarrow \mathbb{R}$ é derivável no conjunto X e obtém-se uma nova função $f' : X \cap X' \longrightarrow \mathbb{R},\ x \longmapsto f'(x)$, chamada a função derivada de f. Se f' é contínua, diz-se que f é de classe C^1.

Teorema A.16. *A fim de que $f : X \longrightarrow \mathbb{R}$ seja derivável no ponto $a \in X \cap X'$ é necessário e suficiente que exista $c \in \mathbb{R}$ tal que $a + h \in X \Rightarrow f(a + h) = f(a) + c \cdot h + r(h)$, onde $\lim\limits_{h \longrightarrow 0} r(h)/h = 0$. No caso afirmativo, tem-se $c = f'(a)$.*

Demonstração. Seja $Y = \{h \in \mathbb{R}; a + h \in X\}$. Então $0 \in Y \cap Y'$. Supondo que $f'(a)$ exista, definimos $r : Y \longrightarrow \mathbb{R}$ pondo $r(h) = f(a + h) - f(a) - f'(a) \cdot h$. Então

$$\frac{r(h)}{h} = \frac{f(a + h) - f(a)}{h} - f'(a),$$

logo $\lim\limits_{h \longrightarrow 0} r(h)/h = 0$. A condição é, portanto, necessária. Reciprocamente, se vale a condição, então $r(h)/h = [f(a + h) - f(a)]/h - c$, logo $\lim\limits_{h \longrightarrow 0}(f(a + h) - f(a))/h - c = \lim\limits_{h \longrightarrow 0} r(h)/h = 0$, portanto $f'(a)$ existe e é igual a c. □

A.8 INTEGRAL DE RIEMANN

Uma partição do intervalo $[a, b]$ é um subconjunto finito de pontos $P = \{t_0, t_1, \ldots, t_n\} \subset [a, b]$ tal que $a \in P$ e $b \in P$. O intervalo $[t_{i-1}, t_i]$, será chamado o i-ésimo intervalo da partição P. Sendo $\sum\limits_{i=1}^{n}(t_i - t_{i-1}) = b - a$.

Seja $f : [a, b] \longrightarrow \mathbb{R}$ uma função limitada, usaremos as seguintes notações,

$m = \inf\{f(x); x \in [a, b]\}$

$M = \sup\{f(x); x \in [a, b]\}$

$m_i = \inf\{f(x); t_{i-1} \leqslant x \leqslant t_i\}$

$M_i = \sup\{f(x); t_{i-1} \leqslant x \leqslant t_i\}$

$\omega = M_i - m_i$

onde ω indica a oscilação de f no i-ésimo intervalo de P.

Quando f é contínua, m_i e M_i são valores assumidos por f em $[t_{i-1}, t_i]$ e neste caso existem $x_i, y_i \in [t_{i-1}, t_i]$ tais que $\omega_i = |f(y_i) - f(x_i)|$.

A soma inferior de f relativamente à partição P é o número

$$s(f; P) = m_1(t_1 - t_0) + \cdots + m_n(t_n - t_{n-1}) = \sum_{i=1}^{n} m_i(t_i - t_{i-1})$$

A soma superior de f relativamente à partição P é o número

$$S(f;P) = M_1(t_1 - t_0) + \cdots + M_n(t_n - t_{n-1}) = \sum_{i=1}^{n} M_i(t_i - t_{i-1})$$

Seja qual for a partição P tem-se

$$m(b-a) \leqslant s(f;P) \leqslant S(f;P) \leqslant M(b-a)$$

$$S(f;P) - s(f;P) = \sum_{i=1}^{n} \omega_i(t_i - t_{i-1}).$$

A integral inferior e a integral superior da função limitada $f : [a,b] \longrightarrow \mathbb{R}$ são definidas, respectivamente, por

$$\underline{\int_a^b} f(x)\,\mathrm{d}x = \sup_P s(f;P)$$

$$\overline{\int_a^b} f(x)\,\mathrm{d}x = \inf_P S(f;P)$$

Definição A.13. *Uma função limitada $f : [a,b] \longrightarrow \mathbb{R}$ diz-se integrável quando sua integral inferior e sua integral superior são iguais. Esse valor comum chama-se a integral de Riemann de f e é indicado por $\int_a^b f(x)dx$.*

Lema A.1. *Sejam $f, g : [a,b] \longrightarrow \mathbb{R}$ integráveis. Então*

1) Se $f(x) \leqslant g(x)$ para todo $x \in [a,b]$, então $\int_a^b f(x)dx \leqslant \int_a^b g(x)dx$.

2) $|f|$ é integrável e $\left|\int_a^b f(x)dx\right| \leqslant \int_a^b |f(x)|dx$

Demonstração. 1) Se $f(x) \leqslant g(x)$ para todo $x \in [a,b]$ então $s(f;P) \leqslant s(g;P)$ e $S(f;P) \leqslant S(g;P)$ para toda partição P, donde $\int_a^b f(x)dx \leqslant \int_a^b g(x)dx$.

2) A desigualdade evidente $||f(y)| - |f(x)|| \leqslant |f(y) - f(x)|$ mostra que a oscilação de $|f|$ em qualquer conjunto não supera a de f. Logo, f integrável $|f|$ integrável. Além disso, como $-|f(x)| \leqslant f(x) \leqslant |f(x)|$ para todo $x \in [a,b]$, resulta de (1) que

$$-\int_a^b |f(x)|dx \leqslant \int_a^b f(x)dx \leqslant \int_a^b |f(x)|dx,$$

ou seja

$$\left|\int_a^b f(x)dx\right| \leqslant \int_a^b |f(x)|dx.$$

□

A.9 SEQUÊNCIAS E SÉRIES DE FUNÇÕES

Seja X um conjunto de números reais. De acordo com as definições gerais, uma sequência de funções $f_n : X \longrightarrow \mathbb{R}$ é uma correspondência que associa a cada número natural $n \in \mathbb{N}$ uma função f_n, definida em X e tomando valores reais.

Definição A.14. *Diz-se que a sequência de funções $f_n : X \longrightarrow \mathbb{R}$ converge simplesmente para a função $f : X \longrightarrow \mathbb{R}$ quando, para cada $x \in X$, a sequência de números $f_1(x), f_2(x), \ldots, f_n(x), \ldots$ converge para o número $f(x)$. Ou seja, para todo $x \in X$ fixado, tem-se $\lim\limits_{n \longrightarrow \infty} f_n(x) = f(x)$. De maneira mais formal, dado qualquer $\varepsilon > 0$, pode-se obter, para cada $x \in X$, um inteiro $n_0 = n_0(\varepsilon, x)$, o qual depende de ε e de x, tal que $n > n_0 \Rightarrow |f_n(x) - f(x)| < \varepsilon$.*

Definição A.15. *Uma sequência de funções $f_n : X \longrightarrow \mathbb{R}$ converge uniformemente para a função $f : X \longrightarrow \mathbb{R}$ quando, para todo $\varepsilon > 0$ dado, existe $n_0 \in \mathbb{N}$ (dependendo apenas de ε) tal que $n > n_0 \Rightarrow |f_n(x) - f(x)| < \varepsilon$ seja qual for $x \in X$.*

Teorema A.17. *Se uma sequência de funções $f_n : X \longrightarrow \mathbb{R}$ converge uniformemente para $f : X \longrightarrow \mathbb{R}$ e cada f_n é contínua no ponto $a \in X$ então f é contínua no ponto a.*

Demonstração. Dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $n > n_0 \Rightarrow |f_n(x) - f(x)| < \varepsilon/3$ para todo $x \in X$. Fixemos um número natural $n > n_0$. Como f_n é contínua no ponto a, existe $\delta > 0$ tal que $x \in X, |x - a| < \delta \Rightarrow |f_n(x) - f_n(a)| < \varepsilon/3$, donde

$$|f(x) - f(a)| \leqslant |f_n(x) - f(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

□

Teorema A.18. *(Passagem ao limite sob o sinal de integral) Se a sequência de funções integráveis $f_n : [a, b] \longrightarrow \mathbb{R}$ converge uniformemente para $f : [a, b] \longrightarrow \mathbb{R}$ então f é integrável e*

$$\int_a^b f(x)dx = \lim_{n \to \infty} \int_a^b f_n(x)dx.$$

Podemos dizer de outra forma que:

$$\int_a^b \lim_n f_n = \lim_n \int_a^b f_n$$

se a convergência é uniforme.

Demonstração. Dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $n > n_0 \Rightarrow |f(x) - f_n(x)| < \varepsilon/4(b - a)$ para todo $x \in [a, b]$. Fixemos $m > n_0$. Como f_m é integrável, existe uma partição P de $[a, b]$ tal que, indicando com ω_i e ω_i' respectivamente as oscilações de f e f_m no intervalo $[t_{i-1}, t_i]$ de P, tem-se $\sum \omega_i'(t_i - t_{i-1}) < \varepsilon/2$. Mas, para $x, y \in [t_{i-1}, t_i]$ quaisquer, vale:

$$|f(y) - f(x)| \leqslant |f(y) - f_m(y)| + |f_m(y) - f_m(x)| + |f_m(x) - f(x)| < \omega_i' + \frac{\varepsilon}{2(b - a)}.$$

Portanto $\omega_i \leqslant \omega_i' + \dfrac{\varepsilon}{2(b - a)}$. Segue-se que

$$\sum \omega_i(t_i - t_{i-1}) \leqslant \sum \omega_i'(t_i - t_{i-1}) + [\varepsilon/2(b - a)] \sum (t_i - t_{i-1}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Isto mostra que f é integrável. Além disso,

$$\left|\int_a^b f(x)dx - \int_a^b f_n(x)dx\right| = \left|\int_a^b [f(x) - f_n(x)]dx\right| \leqslant \int_a^b |f(x) - f_n(x)|dx \leqslant \frac{(b-a)\varepsilon}{4(b-a)} < \varepsilon$$

se $n > n_0$. Consequentemente, $\lim\limits_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx$. □

B ESPAÇO VETORIAL

Espaços vetoriais desempenham um papel importante em vários ramos da matemática e suas aplicações. Em vários problemas práticos e teóricos temos um conjunto X cujos elementos podem ser vetores num espaço tridimensional, ou uma sequência de números, ou funções, e estes elementos podem ser somados ou multiplicados por constantes de uma forma natural, sendo o resultado ainda um elemento de X.

Definição B.1. *(Espaço vetorial) Um espaço vetorial, ou espaço linear, sobre um corpo $\mathbb{K}$, sendo $\mathbb{K}$ igual a $\mathbb{R}$ ou $\mathbb{C}$, é um conjunto não-vazio X de elementos $x, y, \dots,$ chamados vetores, munido de duas operações algébricas. Estas operações são chamadas adição de vetores e multiplicação de vetores por escalares, ou seja, por elementos de $\mathbb{K}$.*

*A **adição vetorial** associa a cada par ordenado (x, y) de vetores um vetor $x + y$, chamado a soma de x e y, de uma forma que a seguintes propriedades são satisfeitas. A adição é comutativa e associativa, isto é, para todos os vetores nós temos*

$$x + y = y + x$$

$$x + (y + z) = (x + y) + z.$$

Além disso, existe um vetor 0, chamado de vetor nulo, e para cada vetor x existe um vetor $-x$, tal que para todos os vetores temos

$$x + 0 = x$$

$$x + (-x) = 0.$$

*A **multiplicação por escalares** associa a cada vetor x e escalar α um vetor αx, chamado de produto de α e x, de uma forma que para todos os vetores x, y e escalares α, β temos*

$$\alpha(\beta x) = (\alpha\beta)x$$

$$1x = x$$

e as leis distributivas

$$\alpha(x + y) = \alpha x + \alpha y$$

$$(\alpha + \beta)x = \alpha x + \beta x.$$

Exemplos:

(1) O espaço $\mathbb{R}^n$ que consiste nas n-uplas de números reais, $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n)$,$\cdots$, é um espaço vetorial real com duas operações algébricas definidas da

maneira usual

$$x + y = (\xi_1 + \eta_1, \cdots, \xi_n + \eta_n)$$
$$\alpha x = (\alpha\xi_1, \cdots, \alpha\xi_n), \quad \alpha \in \mathbb{R}.$$

(2) No espaço $C[a, b]$ cada ponto do espaço é uma função real contínua em $[a, b]$. O conjunto de todas estas funções formam um espaço vetorial real com as operações algébricas definidas da forma usual

$$(x + y)(t) = x(t) + y(t)$$
$$(\alpha x)(t) = \alpha x(t), \quad \alpha \in \mathbb{R}.$$

(3) O espaço ℓ^2 é um espaço vetorial com as operações algébricas definidas como é usual em conexão com sequências, isto é,

$$(\xi_1, \xi_2, \dots) + (\eta_1, \eta_2, \dots) = (\xi_1 + \eta_1, \xi_2 + \eta_2, \dots)$$
$$\alpha(\xi_1, \xi_2, \dots) = (\alpha\xi_1, \alpha\xi_2, \dots).$$

Definição B.2. *(Subespaço vetorial) Um subespaço de um espaço vetorial X é um subconjunto não vazio Y de X tal que para todo $y_1, y_2 \in Y$ e todos os escalares α, β temos $\alpha y_1 + \beta y_2 \in Y$. Consequentemente Y é um espaço vetorial e as duas operações algébricas são induzidas de X.*

Um subespaço especial de X é o subespaço impróprio $Y = X$. Qualquer outro subespaço de X $(\neq \{0\})$ é chamado próprio. Outro subespaço especial de qualquer espaço vetorial X é $Y = \{0\}$.

A combinação linear de vetores $x_1, \dots, x_n$ de um espaço vetorial X é uma expressão da forma

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m$$

onde os coeficientes $\alpha_1, \dots, \alpha_m$ são todos escalares.

Para qualquer subconjunto não-vazio $M \subset X$ o conjunto de todas as combinações lineares de vetores de M é chamado de subespaço gerado de M.

Definição B.3. *(Independência e dependência linear) A independência e dependência linear de um dado conjunto M de vetores $x_1, \dots, x_r$ $(r \geqslant 1)$ em um espaço vetorial X são definidas por meio da equação*

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_r x_r = 0 \tag{B.1}$$

onde $\alpha_1, \dots, \alpha_r$ são escalares. Claramente a equação (B.1) vale para $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$. Se esta é a única r-upla de escalares para a qual (B.1) vale, o conjunto M é linearmente *independente. M é linearmente dependente se não for linearmente independente, isto é, se (B.1) vale para alguma r-upla de escalares, onde não são todos iguais a zero.*

Um subconjunto arbitrário M de X é linearmente independente se todo conjunto finito não-vazio de M é linearmente independente. M é linearmente dependente se não for linearmente independente.

A motivação para esta terminologia resulta do fato de que se $M = \{x_1, \ldots, x_r\}$ é linearmente dependente, pelo menos um vetor de M pode ser escrito como a combinação linear dos outros. Por exemplo, se (B.1) é satisfeita com um $\alpha_r \neq 0$, então M é linearmente dependente e podemos resolver (B.1) para x_r, para obter

$$x_r = \beta_1 x_1 + \cdots + \beta_{r-1} x_{r-1} \qquad (\beta_j = -\alpha_j/\alpha_r).$$

Podemos usar os conceitos de dependência e independência linear para definir a dimensão do espaço vetorial.

Definição B.4. *(espaço vetorial finito e infinito) Um espaço vetorial X é dito ser dimensionalmente finito se existe um inteiro positivo n tal que X contém um conjunto linearmente independente de n vetores onde qualquer conjunto de $n+1$ ou mais vetores de X é linearmente dependente. n é chamado de dimensão de X, escrevemos $n = \dim X$.*

Por definição $X = \{0\}$ é dimensionalmente finito e $\dim = 0$. Se X não é dimensionalmente finito, é dimensionalmente infinito.

Em Análise, espaços vetoriais infinitos são de maior interesse que os dimensionalmente finitos. Por exemplo, $C[a,b]$ e ℓ^2 são dimensionalmente infinitos, enquanto $\mathbb{R}^n$ e $\mathbb{C}^n$ são n-dimensional.

Se $\dim X = n$, uma n-upla linearmente independente de vetores de X é chamada uma base de X. Se $\{e_1, \ldots, e_n\}$ é uma base de X, todo $x \in X$ tem uma única representação como uma combinação linear de vetores da base:

$$x = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

Por exemplo, uma base de $\mathbb{R}^n$ é

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0), \\ e_2 &= (0, 1, 0, \ldots, 0), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n &= (0, 0, 0, \ldots, 1). \end{aligned}$$

Esta base é chamada de base canônica de $\mathbb{R}^n$.

De uma forma mais geral, se X é qualquer espaço vetorial, não necessariamente dimensionalmente finito, e B é um subconjunto linearmente independente de X que gera X, então B é chamado de base de X, ou base de Hamel. Consequentemente se B é uma base de X, então todo $x \in X$ diferente de zero tem uma única representação como uma combinação linear (em um número finito) de elementos de B com escalares diferentes de zero como coeficientes.

Todo espaço vetorial $X \neq \{0\}$ tem uma base.

Em um espaço vetorial dimensionalmente finito isto é claro. Para espaços vetoriais arbitrários dimensionalmente infinitos a prova da existência é dada com o uso do lema de Zorn e pode ser encontrada na referência [4].

Todas as bases de um dado espaço vetorial X (dimensionalmente finito ou infinito) tem o mesmo número cardinal. Este número é chamado de dimensão de X.

Teorema B.1. *Seja X um espaço vetorial n-dimensional. Então qualquer subespaço próprio Y de X tem dimensão menor que n.*

Demonstração. Se $n = 0$, então $X = 0$ e não tem subespaço próprio. Se $\dim Y = 0$, então $Y = \{0\}$ e $X \neq Y$ implica $\dim X \geqslant 1$. Claramente, $\dim Y \leqslant \dim X = n$. Se $\dim Y$ for n, então Y pode ter uma base de n elementos, que também pode ser uma base para X uma vez que $\dim X = n$, logo $X = Y$. Isto mostra que qualquer conjunto de vetores linearmente independentes em Y deve ter menos de n elementos, e $\dim Y < n$. $\square$