Universidade Estadual Paulista “Júlio de Mesquita Filho”
Instituto de Geociências e Ciências Exatas
Campus de Rio Claro

Análise Funcional e Aplicações

Luciana Bertholdi Machado

Dissertação apresentada ao Programa de Pós-Graduação – Mestrado Profissional em Matemática Universitária como requisito parcial para a obtenção do grau de Mestre

Orientadora
Profa. Dra. Marta Cilene Gadotti

2012

517.5 Machado, Luciana Bertholdi
M149a Análise Funcional e Aplicações/ Luciana Bertholdi Machado- Rio Claro: [s.n.], 2012.
204 f.:fig.

Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Geociências e Ciências Exatas.
Orientadora: Marta Cilene Gadotti

1. Espaços Normados. 2. Espaços de Banach. 3. Espaços de Hilbert. 4. Espaço Dual. I. Título

Ficha Catalográfica elaborada pela STATI - Biblioteca da UNESP
Campus de Rio Claro/SP

TERMO DE APROVAÇÃO

Luciana Bertholdi Machado

ANÁLISE FUNCIONAL E APLICAÇÕES

Dissertação APROVADA como requisito parcial para a obtenção do grau de Mestre no Curso de Pós-Graduação Mestrado Profissional em Matemática Universitária do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", pela seguinte banca examinadora:

Profa. Dra. Marta Cilene Gadotti
Orientadora

Prof. Dr. Wladimir Seixas
Departamento de Física, Química e Matemática - UFSCar

Prof. Dr. Miguel Vinicios Santini Frasson
Departamento de Matemática Aplicada - ICMC - USP

Rio Claro, 30 de Novembro de 2012

Aos meus pais Tarcísio e Alice,
ao meu esposo Junior e ao meu filho Matheus, dedico.

Agradecimentos

Agradeço primeiramente a Deus por todas as bênçãos recebidas e por me fortalecer nos momentos de dificuldade, me possibilitando chegar até aqui.

Aos meus pais, irmãs e familiares por acreditarem em mim. Em especial ao meu esposo Junior e ao meu filho Matheus pela paciência e por suportarem a distância para que meu objetivo pudesse ser atingido.

A minha orientadora e amiga Profa. Dra. Marta C. Gadotti pelo conhecimento a mim transmitido, pela dedicação e paciência.

Aos queridos professores do programa de pós-graduação por contribuírem na minha formação.

As secretárias do departamento de Matemática Ana e Eliza pela atenção e amizade. A secretária da seção de pós-graduação e também amiga Inajara pela disposição e simpatia.

Aos grandes amigos que conquistei em Rio Claro, que durante todo este período foram minha família.

Aos professores da UNEMAT de Barra do Bugres que me incentivaram a prosseguir na carreira acadêmica.

A todas as pessoas que direta ou indiretamente contribuíram pela realização deste trabalho.

Toda a educação científica
que não se inicia com a Matemática
é, naturalmente, imperfeita na sua base.
Auguste Conte

Resumo

O presente trabalho, intitulado Análise Funcional e Aplicações, tem por objetivo realizar um estudo sobre espaços de funções, principalmente, os espaços de dimensão infinita. Em particular, apresentar resultados sobre a teoria de funcionais lineares e espaço dual, conceitos de ortogonalidade e teoremas fundamentais em Análise Funcional como, por exemplo, o Teorema da Representação de Riesz e os Teoremas de Hahn-Banach.

Palavras-chave: Espaços Normados, Espaços de Banach, Espaços de Hilbert, Espaço Dual.

Abstract

This work, entitled Functional Analysis and Applications, has by objective to carry an study on function spaces, mainly, spaces of infinite dimension. In particular, to present results on the theory of linear functionals and dual space, concepts of orthogonality and fundamental theorems in Functional Analysis as, for example, the Riesz Representation Theorem and the Hahn-Banach Theorems.

Keywords: Normed Spaces, Banach Spaces, Hilbert Spaces, Dual Spaces.

Lista de Figuras

Sumário

1 Introdução

A Análise Funcional é o ramo da Matemática, mais especificamente da Análise, que trata do estudo de espaços de funções e faz uso de muitos conceitos de Álgebra Linear, com ênfase para espaços vetoriais de dimensão infinita. Entre os resultados importantes da Análise Funcional está o Teorema de Hahn-Banach, o qual permite que funcionais lineares definidos em um subespaço de um espaço vetorial sejam estendidos a todo o espaço.

Nos objetivos traçados para este trabalho, além de abordar alguns tópicos importantes da Análise Funcional, pretende-se fornecer um material didático para fins acadêmicos, cujo intuito é facilitar a compreensão dos conceitos aqui apresentados.

No capítulo preliminar serão abordados alguns conceitos da Análise Matemática tais como convergência e continuidade. O capítulo sobre espaços métricos apresentará alguns resultados como: métrica, conjuntos abertos, conjuntos fechados e compactos, convergência, continuidade, completamento, etc..

No capítulo sobre espaços normados e de Banach seguem alguns resultados de espaços vetoriais, entre os quais se destacam Lema de Zorn e base. Também serão apresentados conceitos de espaços normados e espaços normados que são completos, chamados espaços de Banach. Além disso, importantes conceitos como funcionais lineares contínuos e espaço dual serão abordados.

O quinto capítulo tem por objetivo definir e apresentar alguns exemplos sobre espaços de Hilbert, que são espaços de Banach cuja norma provém de um produto interno. Além disso, seguem algumas propriedades referentes ao complemento ortogonal, resultados importantes envolvendo conjuntos ortonormais e, por fim, o famoso Teorema da Representação de Riesz.

Finalmente, no sexto capítulo, serão apresentadas algumas aplicações. Uma das aplicações envolve a integral de Riemann-Stieltjes e os Teoremas de Hahn-Banach, também será apresentado o Teorema do Ponto fixo de Banach com aplicação em equações diferenciais ordinárias e, por fim, alguns resultados sobre aproximação no espaço de funções.

2 Preliminares

Este capítulo traz resultados e conceitos que serão utilizados nos capítulos posteriores. Serão apresentados resultados básicos da Análise Matemática que, em sua maioria, serão abordados no capítulo seguinte para espaços de dimensão maior.

2.1 Resultados de Análise Real

Os resultados apresentados nesta seção têm por referências [10] e [11].

Definição 2.1. *Seja $A \subset \mathbb{R}$ não vazio e limitado superiormente, ou seja, existe $b \in \mathbb{R}$ tal que $x \leq b$ para qualquer $x \in A$ e, neste caso, diz-se que b é cota superior de A. O* ***supremo*** *de um conjunto A é a menor das cotas superiores de A, isto é, o número $b \in \mathbb{R}$ é supremo do conjunto A quando:*

i. Para todo $x \in A$, tem-se $x \leq b$;

ii. Para qualquer $\epsilon > 0$ existe $x \in A$ tal que $b - \epsilon < x$.

Notação: $b = \sup A$.

Definição 2.2. *Seja $A \subset \mathbb{R}$ não vazio e limitado inferiormente, ou seja, existe $a \in \mathbb{R}$ tal que $a \leq x$ para qualquer $x \in A$ e, neste caso diz-se que o número a é cota inferior de A. O* ***ínfimo*** *de um conjunto A é a maior das cotas inferiores de A, isto é, o número $a \in \mathbb{R}$ é ínfimo do conjunto A quando:*

i. Para todo $x \in A$, tem-se $a \leq x$;

ii. Para qualquer $\epsilon > 0$ existe $x \in A$ tal que $x < a + \epsilon$.

Notação: $a = \inf A$.

Observação 2.1. Vale observar a seguinte propriedade envolvendo as definições anteriores: sejam A e B subconjuntos não vazios de números reais, tais que $x \leq y$ para todo $x \in A$ e todo $y \in B$. Então $\sup A = \inf B$ se, e somente se, para todo $\epsilon > 0$ dado, podem-se obter $x \in A$ e $y \in B$ tais que $y - x < \epsilon$.

De fato, se $\sup A = \inf B$, então dado $\epsilon > 0$, existe $x \in A$ tal que $x > \sup A - \epsilon/2$ e $y < \inf B + \epsilon/2$. Logo,

$$y - x < \inf B + \frac{\epsilon}{2} - \sup A + \frac{\epsilon}{2} = \inf B - \sup A + \epsilon = \epsilon,$$

ou seja, $y - x < \epsilon$.

Reciprocamente, se $\sup A \neq \inf B$ e, portanto, $\sup A < \inf B$, considere

$$\epsilon = \inf B - \sup A > 0.$$

Por hípótese, existem $x \in A$ e $y \in B$ tais que $y - x < \epsilon$. Logo,

$$x > y - \epsilon = y - \inf B + \sup A > \sup A,$$

o que é uma contradição. Portanto, $\sup A = \inf B$.

Axioma 2.1. *Todo subconjunto $A \subset \mathbb{R}$ não vazio e limitado superiormente possui **supremo** em $\mathbb{R}$.*

Utilizando o axioma acima é possível provar que todo subconjunto $B \subset \mathbb{R}$ não vazio e limitado inferiormente possui **ínfimo** em $\mathbb{R}$.

Observação 2.2. O axioma afirma que o corpo ordenado $\mathbb{R}$ é completo.

Definição 2.3. *Uma **sequência** de números reais é uma função definida no conjunto $\mathbb{N} = \{1, 2, \ldots\}$ dos números naturais com valores no conjunto $\mathbb{R}$ dos números reais. Ou seja,*

$$\begin{aligned} x : \mathbb{N} &\to \mathbb{R} \\ n &\mapsto x(n) = x_n \end{aligned}$$

Escreve-se $(x_1, \ldots, x_n, \ldots)$, $(x_n)_{n\in\mathbb{N}}$, ou ainda (x_n) para indicar a sequência cujo n-ésimo termo é x_n.

Definição 2.4. *Uma **subsequência** da sequência $x = (x_n)_{n\in\mathbb{N}}$ é uma restrição da função $x : \mathbb{N} \to \mathbb{R}$ a um subconjunto infinito $\mathbb{N}' = \{n_1 < n_2 < \ldots < n_k < \ldots\}$ de $\mathbb{N}$. Escreve-se $(x_{n_1}, \ldots, x_{n_k}, \ldots)$, $x' = (x_n)_{n\in\mathbb{N}'}$ ou ainda $(x_{n_k})_{k\in\mathbb{N}}$ para indicar a subsequência de x.*

Definição 2.5. *Diz-se que o número real a é **limite** da sequência $(x_n)_{n\in\mathbb{N}}$ quando, para qualquer número real $\epsilon > 0$, pode-se obter um $n_0 \in \mathbb{N}$ tal que todos os termos x_n com índice $n > n_0$ cumprem a condição $|x_n - a| < \epsilon$. Escreve-se então $a = \lim\limits_{n\to\infty} x_n$. Simbolicamente,*

$$a = \lim_{n\to\infty} x_n \Leftrightarrow \forall\, \epsilon > 0, \exists\, n_0 \in \mathbb{N};\ |x_n - a| < \epsilon, \forall\, n > n_0.$$

ou seja,

$$a = \lim_{n\to\infty} x_n \Leftrightarrow \forall\, \epsilon > 0, \exists\, n_0 \in \mathbb{N};\ x_n \in (a-\epsilon, a+\epsilon), \forall\, n > n_0.$$

*Diz-se que a sequência $(x_n)_{n\in\mathbb{N}}$ **converge** (ou tende) para o número a e escreve-se também $x_n \to a$. Diz-se que (x_n) é **divergente** se **não existe** $\lim\limits_{n\to\infty} x_n$.*

Teorema 2.1. *Se o limite de uma sequência existir, então ele é único.*

Demonstração. Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência convergente e suponha que, para $a \neq b$, $a = \lim\limits_{n\to\infty} x_n$ e $b = \lim\limits_{n\to\infty} x_n$. Assim, pode-se escolher $\epsilon > 0$ tal que

$$(a-\epsilon, a+\epsilon) \cap (b-\epsilon, b+\epsilon) = \emptyset.$$

Então existe $n_0 \in \mathbb{N}$ com $n > n_0$, tal que $x_n \in (a-\epsilon, a+\epsilon)$. Por outro lado existe $n_1 \in \mathbb{N}$, tal que $x_n \in (b-\epsilon, b+\epsilon)$.

Seja $n = \max\{n_0, n_1\}$ então $x_n \in (a-\epsilon, a+\epsilon) \cap (b-\epsilon, b+\epsilon)$, o que é contradição. □

Teorema 2.2. *Se $\lim\limits_{n\to\infty} x_n = a$ então toda subsequência de (x_n) converge para o limite a.*

Demonstração. Se $x_n \to a$, então dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $n > n_0$ implica $x_n \in (a-\epsilon, a+\epsilon)$. Seja $(x_{n_k})_{k\in\mathbb{N}}$ uma subsequência de (x_n). Então, para todo $n_k > n_0$ tem-se $x_{n_k} \in (a-\epsilon, a+\epsilon)$, logo $x_{n_k} \to a$. □

Proposição 2.1. *Toda sequência convergente é limitada.*

Demonstração. Seja $\lim\limits_{n\to\infty} x_n = a$, então para $\epsilon = 1$ existe $n_0 \in \mathbb{N}$ tal que $n > n_0$ implica $x_n \in (a-1, a+1)$. Considere o conjunto finito $F = \{x_1, \ldots, x_{n_0}, a-1, a+1\}$ e sejam m o menor e M o maior elemento de F. Desse modo, todos os termos x_n da sequência estão contidos no intervalo $[m, M]$, ou seja, $m \leq x_n \leq M$, portanto a sequência (x_n) é limitada. □

Observação 2.3. A recíproca do teorema não é verdadeira. Note, por exemplo, que a sequência $(1, 0, 1, 0, \ldots)$ é limitada mas não converge porque possui duas subsequências constantes $x_{2n-1} = 1$ e $x_{2n} = 0$ com limites distintos, pois $x_{2n-1} \to 1$ e $x_{2n} \to 0$.

Observação 2.4. A negação da proposição 2.1 afirma que se uma sequência não é limitada, ela não é congervente. Por exemplo, $(x_n)_{n\in\mathbb{N}} = (1, 2, 3, \ldots)$.

Definição 2.6. *Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência. Então*

*i. Se $x_n < x_{n+1}$ para todo $n \in \mathbb{N}$, diz-se que (x_n) é **crescente**. Se $x_n \leq x_{n+1}$ para todo $n \in \mathbb{N}$, diz-se que (x_n) é **não-decrescente**.*

*ii. Se $x_n > x_{n+1}$ para todo $n \in \mathbb{N}$, diz-se que (x_n) é **decrescente**. Se $x_n \geq x_{n+1}$ para todo $n \in \mathbb{N}$, diz-se que (x_n) é **não-crescente**.*

iii. As sequências definidas em i. e ii. são chamadas ***sequências monótonas****.*

Teorema 2.3. *Toda sequência monótona limitada é convergente.*

Demonstração. Considere $(x_n)_{n\in\mathbb{N}}$ uma sequência não-descrecente e limitada. Pelo aximoma 2.1 existe $a = \sup\{x_n\}$, então $\lim_{n\to\infty} x_n = a$. Com efeito, como $a = \sup\{x_n\}$, pela definição 2.1, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $a - \epsilon < x_{n_0} < a$. Assim, para todo $n > n_0$

$$a - \epsilon < x_{n_0} \leq x_n < a + \epsilon,$$

pois (x_n) é não-decrescente. Portanto, $x_n \to a$. De maneira semelhante, prova-se que se a sequência (x_n) é não-crescente e limitada então $\lim_{n\to\infty} x_n = \inf\{x_n\}_{n\in\mathbb{N}}$. □

Teorema 2.4. ***(Teorema de Bolzano-Weierstrass)*** *Toda sequência limitada de números reais possui uma subsequência convergente.*

Demonstração. Seja (x_n) uma sequência limitada. Pelo teorema anterior, basta provar que (x_n) possui subsequência monótona. Diz-se que um termo x_n da sequência dada é *destacado* quando, para todo $m > n$ tem-se

$$x_n \geq x_m, \tag{2.1}$$

ou seja, x_n é maior ou igual aos termos que o sucedem. Seja $D \subset \mathbb{N}$ o conjunto dos índices n tais que x_n é um termo destacado, ou seja, o conjunto dos índices n que satisfazem (2.1).

i. Se D for um conjunto infinito, $D = \{n_1 < n_2 < \ldots < n_k < \ldots\}$, pela definição de termo destacado a subsequência (x_{n_k}), $n \in D$ e $k \in \mathbb{N}$, será monótona não-crescente e limitada, logo convergente.

ii. Se o conjunto D for finito, $D = \{n_1 < n_2 < \ldots < n_k\}$, seja $n_{k_1} \in \mathbb{N}$ maior do que todos os $n_k \in D$. Neste caso, $x_{n_{k_1}}$ não é destacado, ou seja, não satisfaz (2.1). Logo existe $n_{k_2} > n_{k_1}$ de forma que $x_{n_{k_1}} < x_{n_{k_2}}$. Como $x_{n_{k_2}}$ não é destacado, existe $n_{k_3} > n_{k_2}$ com $x_{n_{k_2}} < x_{n_{k_3}}$. Prosseguindo desta forma, obtem-se uma subsequência crescente e limitada

$$x_{n_{k_1}} < x_{n_{k_2}} < x_{n_{k_3}} < \ldots < x_{n_{k_k}} < \ldots,$$

portanto, convergente. □

Definição 2.7. *Diz-se que (x_n) é uma* ***sequência de Cauchy*** *quando, para qualquer $\epsilon > 0$ dado, existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ implica $|x_m - x_n| < \epsilon$.*

Lema 2.1. *Toda sequência de Cauchy é limitada.*

Demonstração. Seja (x_n) uma sequência de Cauchy. Em particular, considerando $\epsilon = 1$ e $n \geq n_0$ tem-se $|x_{n_0} - x_n| < 1$, ou seja, se $n \geq n_0$ então $x_n \in (x_{n_0} - 1, x_{n_0} + 1)$. Escrevendo $X = \{x_1, x_2, \ldots, x_{n_0} - 1, x_{n_0} + 1\}$, sejam α e β o menor e o maior elemento do conjunto X, repectivamente. Então $x_n \in [\alpha, \beta]$ para todo $n \in \mathbb{N}$, portanto (x_n) é limitada. $\square$

Definição 2.8. *Uma função $f : X \to \mathbb{R}$ diz-se* ***contínua no ponto*** *$a \in X$ quando, para todo $\epsilon > 0$ dado, pode-se obter $\delta > 0$ (este δ pode depender de ϵ e de a) tal que $x \in X$, $|x - a| < \delta$ implica $|f(x) - f(a)| < \epsilon$. Simbolicamente*

$$\forall\, \epsilon > 0 \;\exists\; \delta > 0; \forall\, x \in X,\; |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Diz-se que $f : X \to \mathbb{R}$ é uma ***função contínua*** *quando f é contínua em todos os pontos $a \in X$.*

Chama-se ***descontínua no ponto*** *$a \in X$ uma função $f : X \to \mathbb{R}$ que não é contínua neste ponto. Em outras palavras,*

$$\exists\, \epsilon > 0;\; \forall\; \delta > 0; \exists\, x_\delta \in X,\; |x_\delta - a| < \delta \Rightarrow |f(x_\delta) - f(a)| \geq \epsilon.$$

Exemplo 2.1. A função $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = x^2$ para todo $x \in \mathbb{R}$, é contínua no ponto $a \in \mathbb{R}$. Com efeito, inicialmente note que $f(x) = x^2$ é contínua no ponto $a = 0$, ou seja, dado $\epsilon > 0$ basta tomar $\delta = \sqrt{\epsilon} > 0$ tal que

$$|x - 0| < \delta \Rightarrow |f(x) - f(0)| = |x^2 - 0^2| = |x|.|x| < \delta.\delta = \sqrt{\epsilon}.\sqrt{\epsilon} = \epsilon.$$

Agora, para todo ponto $a \neq 0$, dado $\epsilon > 0$ note que

$$\begin{aligned} |f(x) - f(a)| < \epsilon \Leftrightarrow &\;\; f(a) - \epsilon < f(x) < f(a) + \epsilon \\ \Leftrightarrow &\;\; a^2 - \epsilon < x^2 < a^2 + \epsilon. \end{aligned}$$

Considerando $\epsilon' < \min\{a^2, \epsilon\}$ tem-se

$$\sqrt{a^2 - \epsilon'} < |x| < \sqrt{a^2 + \epsilon'}.$$

i. Se $a > 0$, considere o intervalo aberto $I_1 = (\sqrt{a^2 - \epsilon}, \sqrt{a^2 + \epsilon})$ centrado em a. Como f é crescente neste intervalo, segue que

$$x \in I_1 \Rightarrow a^2 - \epsilon' < x^2 < a^2 + \epsilon' \Leftrightarrow |x^2 - a^2| = |f(x) - f(a)| < \epsilon',$$

e neste caso $\delta = a - \sqrt{a^2 - \epsilon'} > 0$.

ii. Se $a < 0$, considere o intervalo aberto $I_2 = (-\sqrt{a^2 + \epsilon}, -\sqrt{a^2 - \epsilon})$ centrado em a. Como f é decrescente neste intervalo, segue que

$$x \in I_2 \Rightarrow a^2 - \epsilon' < x^2 < a^2 + \epsilon' \Leftrightarrow |x^2 - a^2| = |f(x) - f(a)| < \epsilon',$$

onde $\delta = a + \sqrt{a^2 + \epsilon'} > 0$.

Ou seja, $f(x) = x^2$ é contínua em todo ponto $a \in \mathbb{R}$.

Definição 2.9. *Uma função $f : X \to \mathbb{R}$ diz-se **uniformemente contínua** no conjunto X quando, para qualquer $\epsilon > 0$ dado, pode-se obter $\delta > 0$ (este δ depende apenas de ϵ) tal que $x, y \in X$, $|x - y| < \delta$ implica $|f(x) - f(y)| < \epsilon$.*

Observação 2.5. Toda função uniformemente contínua é contínua. Com efeito, dado $\epsilon > 0$ existe $\delta > 0$ tal que $x, y \in X$,

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Agora, se $a \in X$ então para todo $x \in X$,

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon,$$

onde δ não depende do ponto $a \in X$, apenas de ϵ.

Exemplo 2.2. A função identidade $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = x$ para todo $x \in \mathbb{R}$, é uniformemente contínua. De fato, para qualquer $\epsilon > 0$ dado existe $\delta = \epsilon > 0$ tal que para todo $x, y \in \mathbb{R}$ tem-se

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| = |x - y| < \delta = \epsilon \Rightarrow |f(x) - f(y)| < \epsilon, \forall\, x, y \in \mathbb{R}.$$

Observação 2.6. Uma função ser contínua não implica que ela seja uniformemente contínua. Por exemplo, a função $f(x) = x^2$, apresentada no exemplo 2.1, é contínua mas não é uniformemente contínua. Com efeito, para $\epsilon = 1$ e qualquer que seja $\delta > 0$ considere $x > \dfrac{1}{\delta}$ e $y = x + \dfrac{\delta}{2}$. Note que

$$|x - y| = \left| x - \left(x + \frac{\delta}{2} \right) \right| = \frac{\delta}{2} < \delta.$$

Entretanto,

$$|f(x) - f(y)| = \left| x^2 - \left(x + \frac{\delta}{2} \right)^2 \right| = \left| - \left(x\delta + \frac{\delta^2}{4} \right) \right| = x\delta + \frac{\delta^2}{4} > x\delta > 1 = \epsilon.$$

Ou seja,

$$\exists\, \epsilon = 1; \forall\, \delta > 0,\ \exists\, x_\delta, y_\delta \in \mathbb{R}, |x_\delta - y_\delta| < \delta \Rightarrow |f(x_\delta) - f(y_\delta)| \geq \epsilon = 1.$$

Teorema 2.5. *Seja $f : X \to \mathbb{R}$ uma função uniformemente contínua. Se (x_n) é uma sequência de Cauchy em X, então $(f(x_n))$ é uma sequência de Cauchy em $\mathbb{R}$.*

Demonstração. Como f é uniformemente contínua, dado $\epsilon > 0$ existe $\delta > 0$ tal que

$$x, y \in X,\ |y - x| < \delta \Rightarrow |f(y) - f(x)| < \epsilon. \tag{2.2}$$

Como (x_n) é de Cauchy, para $\delta > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ tem-se $|x_m - x_n| < \delta$. Logo, por (2.2), $|f(x_m) - f(x_n)| < \epsilon$ para $m, n > n_0$. Portanto, $(f(x_n))$ é de Cauchy. $\square$

Definição 2.10. *Diz-se que uma sequência de funções $f_n : X \to \mathbb{R}$, $n \in \mathbb{N}$,* ***converge pontualmente*** *para a função $f : X \to \mathbb{R}$, quando para todo $x \in X$, a sequência de números $(f_1(x), \ldots, f_n(x), \ldots)$ converge para um número $f(x)$. Em outras palavras, $f_n \to f$ pontualmente em X quando dados $\epsilon > 0$ e $x \in X$, existe $n_0 = n_0(\epsilon, x) \in \mathbb{N}$ tal que $n > n_0$ implica $|f_n(x) - f(x)| < \epsilon$.*

Exemplo 2.3. A sequência de funções $f_n : \mathbb{R} \to \mathbb{R}$ dada por $f_n(x) = \dfrac{x}{n}$ converge pontualmente em $\mathbb{R}$ para a função nula. Com efeito, para cada $x \in \mathbb{R}$ fixo, tem-se que $\lim\limits_{n\to\infty} f_n(x) = \lim\limits_{n\to\infty} \dfrac{x}{n} = 0$. Em outras palavras, dados $\epsilon > 0$ e $x \in \mathbb{R}$, existe $n_0 \in \mathbb{N}$ tal que $n_0 > \dfrac{|x|}{\epsilon}$, então para todo $n > n_0$ segue

$$|f_n(x) - f(x)| = \left|\frac{x}{n} - 0\right| = \left|\frac{x}{n}\right| = \frac{|x|}{n} < \epsilon.$$

Ou seja, para cada $x \in \mathbb{R}$ fixo, a sequência de números $f_n(x) = \dfrac{x}{n}$ converge pontualmente para zero. Veja a figura 2.1.

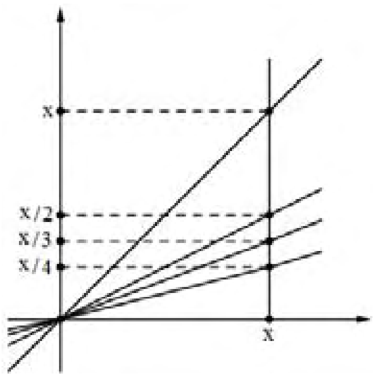

Figura 2.1: *Convergência pontual*

Exemplo 2.4. A sequência de funções (f_n) tal que $f_n : [0, 2\pi] \to \mathbb{R}$ definida por $f_n(x) = \cos(nx)$ para todo $x \in [0, 2\pi]$, **não converge pontualmente** para função alguma. De fato, considerando $x = \pi$, note que $f_n(x) = (-1)^n$, ou seja, não existe limite da função $f_n(x)$ quando $n \to \infty$.

Definição 2.11. *Diz-se que a sequência de funções $f_n : X \to \mathbb{R}$, $n \in \mathbb{N}$,* ***converge uniformente*** *para a função $f : X \to \mathbb{R}$, quando para qualquer $\epsilon > 0$ dado, existe $n_0 \in \mathbb{N}$ (n_0 só depende de ϵ) tal que $n > n_0$ implica $|f_n(x) - f(x)| < \epsilon$, seja qual for $x \in X$.*

Geometricamente, se $X = [a, b]$, a faixa de amplitude 2ϵ em torno do gráfico de $f : [a, b] \to \mathbb{R}$ é o conjunto $\{(x, y) \in \mathbb{R}^2; a \leq x \leq b, f(x) - \epsilon < y < f(x) + \epsilon\}$, como ilustra a figura 2.2 abaixo.

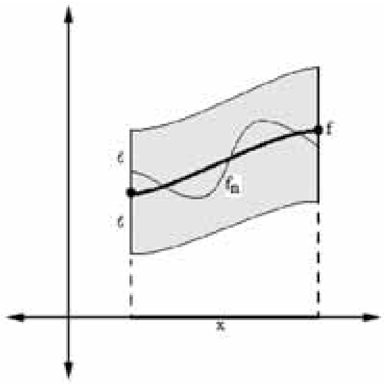

Figura 2.2: *Convergência uniforme*

Observe que para $n_0 \in \mathbb{N}$ suficientemente grande, o gráfico de $f_n(x)$ está contido na referida faixa.

Observação 2.7. Note que, para provar a convergência pontual, dado qualquer $\epsilon > 0$, pode-se obter, para cada $x \in X$, um $n_0 = n_0(\epsilon, x)$. Quando a convergência é uniforme, dado qualquer $\epsilon > 0$, pode-se obter um n_0 que satisfaz a definição para todo $x \in X$.

Exemplo 2.5. Seja $D = [0, 1]$ e para cada $n \in \mathbb{N}$, seja $f_n : D \to \mathbb{R}$ definida por $f_n(x) = x^n$. Note que esta função converge pontualmente para a função $f : D \to \mathbb{R}$ definida por $f(x) = 0$ se $0 \leq x < 1$ e $f(1) = 1$. Por outro lado, dado $0 < \delta < 1$ seja $D' = [0, 1 - \delta]$, observe que $D' \subset D$. Assim, para cada $n \in \mathbb{N}$ seja $g_n : D' \to \mathbb{R}$ a restrição da função f_n ao intervalo D', ou seja, $g_n(x) = x^n$ com $0 \leq x \leq 1 - \delta$. A sequência (g_n) converge uniformemente em D' para a função nula $g : D' \to \mathbb{R}$. De fato, dado $\epsilon > 0$, como $0 < 1 - \delta < 1$, existe $n_0 \in \mathbb{N}$ tal que $n > n_0$ implica $(1 - \delta)^n < \epsilon$. Então, para todo $x \in D'$, tem-se que $0 \leq x^n \leq (1 - \delta)^n < \epsilon$ sempre que $n > n_0$. Portanto, $x^n \to 0$ uniformemente em D'.

Observação 2.8. Convergência pontual não implica em convergência uniforme. No exemplo 2.3 a sequência $f_n(x) = \dfrac{x}{n}$ converge pontualmente para a função nula $\mathbb{R}$, entretanto esta sequência não converge uniformemente para a função nula em $\mathbb{R}$. Com efeito, para $\epsilon = 1$ e qualquer que seja $n \in \mathbb{N}$ pode-se obter $x \in \mathbb{R}$ com $x > n$, tais que

$$|f_n(x) - f(x)| = \left|\frac{x}{n} - 0\right| = \frac{|x|}{n} > \frac{n}{n} = 1 = \epsilon.$$

Teorema 2.6. *Se uma sequência de funções $f_n : X \to \mathbb{R}$ converge uniformemente para $f : X \to \mathbb{R}$ e cada f_n é contínua no ponto $a \in X$ então f é contínua no ponto a.*

Demonstração. Como f_n converge uniformemente para a função f, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que se $n > n_0$ então $|f_n(x) - f(x)| < \epsilon/3$ para todo $x \in X$. Além disso, como f_n é contínua no ponto a, fixando $n \in \mathbb{N}$, $n > n_0$, dado $\epsilon > 0$ existe $\delta > 0$ tal que se $x \in X$, então

$$|x - a| < \delta \Rightarrow |f_n(x) - f(a)| < \frac{\epsilon}{3}.$$

Assim,

$$\begin{aligned} |f(x) - f(a)| = & \ |f(x) - f_n(x) + f_n(x) - f_n(a) + f_n(a) - f(a)| \\ \leq & \ |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| \\ < & \ \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ = & \ \epsilon, \end{aligned}$$

ou seja, dado $\epsilon > 0$ existe $\delta > 0$ tal que $x \in X$,

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Portanto, f é contínua no ponto $a \in X$. □

Exemplo 2.6. A sequência de funções $f_n(x) = x^n$ não converge uniformemente em $[0, 1]$, pois converge pontualmente para a função descontínua $f : [0, 1] \to \mathbb{R}$, onde $f(x) = 0$ se $0 \leq x < 1$ e $f(1) = 1$ (ver exemplo 2.5).

Definição 2.12. *Um conjunto $X \subset \mathbb{R}$ é chamado **compacto** quando é fechado e limitado.*

Exemplo 2.7. Todo intervalo do tipo $[a, b]$ é fechado e limitado, logo compacto. Por outro lado, o intervalo (a, b) é limitado mas não é fechado, logo não é compacto.

Teorema 2.7. *Seja $X \subset \mathbb{R}$ um conjunto compacto. Então toda função contínua $f : X \to \mathbb{R}$ é uniformemente contínua.*

Demonstração. Suponha que f não seja uniformemente contínua em X. Logo existe $\epsilon_0 > 0$ tal que, para todo $n \in \mathbb{N}$ existem $x_n, y_n \in X$ com

$$|x_n - y_n| < \frac{1}{n} \quad \text{e} \quad |f(x_n) - f(y_n)| \geq \epsilon_0.$$

Como (x_n) é limitada, pelo teorema 2.4 (página 22), (x_n) possui uma subsequência (x_{n_k}) convergente. Sendo X compacto, a sequência (x_{n_k}) converge para um ponto $x \in X$. Logo, $y_{n_k} \to x$, pois $\lim\limits_{n\to\infty}(x_n - y_n) = 0$. Da continuidade de f em x, tem-se

$$\lim_{k\to\infty}[f(x_{n_k}) - f(y_{n_k})] = \lim_{k\to\infty} f(x_{n_k}) - \lim_{k\to\infty} f(y_{n_k}) = f(x) - f(x) = 0,$$

contradizendo o fato de $|f(x_{n_k}) - f(y_{n_k})| \geq \epsilon_0$ para todo $k \in \mathbb{N}$. □

Teorema 2.8. ***(Teste de Weierstrass)*** *Dada a sequência de funções $f_n : X \to \mathbb{R}$, seja $\sum a_n$ uma série convergente de números reais $a_n \geq 0$ tais que $|f_n(x)| \leq a_n$ para todo $n \in \mathbb{N}$ e $x \in X$. Nestas condições as séries $\sum |f_n|$ e $\sum f_n$ são uniformemente convergentes.*

Demonstração. Como $\sum a_n$ é uma série convergente, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $n > n_0$ então $\displaystyle\sum_{j=n+1}^{\infty} a_j < \epsilon$. Fazendo, $R(x) = \displaystyle\sum_{j=n+1}^{\infty} |f_j(x)|$ e $r(x) = \displaystyle\sum_{j=n+1}^{\infty} f_j(x)$ e observando que

$$|r(x)| \leq R(x) \leq \sum_{j=n+1}^{\infty} a_j < \epsilon,$$

para todo $n > n_0$, segue que $\sum |f_n|$ e $\sum f_n$ são uniformemente convergentes. $\square$

O capítulo seguinte traz resultados sobre Espaços Métricos. Estes conceitos são de grande importância dentro da Análise Funcional, já que em sua maioria se aplicam a espaços maiores como, por exemplo, espaços normados (que serão apresentados no capítulo 4), quando a métrica é induzida pela norma.

3 Espaços métricos

Este capítulo apresentará resultados de grande relevância para o desenvolvimento deste trabalho, dentre os quais: alguns conceitos topológicos, convergência de sequências, espaços completos, completamento, etc.. A ideia de métrica está associada à noção intuitiva de distância, a qual precisa satisfazer algumas propriedades, conforme definição abaixo.

Definição 3.1. *Seja M um conjunto não vazio. Uma **métrica** sobre o conjunto M é uma função $d : M \times M \to \mathbb{R}$, que a cada par de elementos $x, y \in M$ associa um número real positivo $d(x, y)$, que se chama distância de x a y, de modo que, para quaisquer elementos $x,\ y,\ z \in M$ tem-se:*

M1) $d(x, y) \geq 0$ *e* $d(x, y) = 0 \Leftrightarrow x = y$;

M2) $d(x, y) = d(y, x)$ *(simetria);*

M3) $d(x, z) \leq d(x, y) + d(y, z)$ *(desigualdade triangular).*

*O par (M, d) é chamado **espaço métrico** formado por um conjunto não vazio M e uma métrica d em M.*

Segue alguns exemplos de espaços métricos.

Exemplo 3.1. Espaço métrico discreto. Seja M um conjunto qualquer não vazio. A função $d : M \times M \to \mathbb{R}$, definida por

$$d(x, y) = \begin{cases} 1, & \text{se } x \neq y \\ 0, & \text{se } x = y \end{cases}$$

é uma métrica em M. Logo (M, d) é um espaço métrico.

Observação 3.1. Todo conjunto torna-se um espaço métrico com a métrica definida no exemplo acima.

Exemplo 3.2. O conjunto dos números reais $\mathbb{R}$ com a métrica usual definida por

$$d(x, y) = |x - y|, \forall\, x, y \in \mathbb{R}$$

é um espaço métrico onde $|.|$ denota o valor absoluto.

Exemplo 3.3. Espaço métrico $\mathbb{R}^n$ é o conjunto de todas as n-uplas reais. Considera-se a métrica definida por

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + ... + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}.$$

onde $x = (x_1, ..., x_n)$ e $y = (y_1, ..., y_n)$ são elementos de $\mathbb{R}^n$. Claramente as condições $M1$ e $M2$ da definição 3.1 estão satisfeitas. Resta provar a desigualdade triangular, isto é, dados $x, y, z \in \mathbb{R}^n$

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

ou, equivalentemente,

$$\sqrt{\sum_{i=1}^{n}(x_i - z_i)^2} \leq \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} + \sqrt{\sum_{i=1}^{n}(y_i - z_i)^2}.$$

Faça $x_i - y_i = a_i$ e $y_i - z_i = b_i$ então, a desigualdade acima pode ser reescrita como

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

e, elevando ambos os membros ao quadrado, obtem-se

$$\sum_{i=1}^{n}(a_i)^2 + 2\sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n}(b_i)^2 \leq \sum_{i=1}^{n}(a_i)^2 + 2\sqrt{\sum_{i=1}^{n}(a_i)^2}\sqrt{\sum_{i=1}^{n}(b_i)^2} + \sum_{i=1}^{n}(b_i)^2$$

ou seja,

$$\sum_{i=1}^{n} a_i b_i \leq \sqrt{\sum_{i=1}^{n}(a_i)^2}\sqrt{\sum_{i=1}^{n}(b_i)^2}. \tag{3.1}$$

Esta desigualdade é uma consequência da *Desigualdade de Cauchy-Schwarz*:

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

A fim de demonstrar a desigualdade considere o trinômio do segundo grau em t:

$$f(t) = (a_1 t - b_1)^2 + ... + (a_n t - b_n)^2 = \sum_{i=1}^{n}(a_i t - b_i)^2$$

Desenvolvendo o quadrado, obtem-se

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

Note que $f(t) \geq 0$ para qualquer t real, pois é soma de quadrados e, sendo assim, seu discriminante é menor do que ou igual a zero, ou seja,

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

Deste modo

$$\sum_{i=1}^{n} a_ib_i \leq \sqrt{\sum_{i=1}^{n}(a_i)^2}\sqrt{\sum_{i=1}^{n}(b_i)^2},$$

satisfazendo a condição $M3$ da definição 3.1.

Pode-se definir outras métricas em $\mathbb{R}^n$. Por exemplo, a *métrica da soma* e a *métrica do máximo* definidas, repectivamente, por:

$$d(x,y) = \sum_{i=1}^{n} |x_i - y_i|, \, \forall \, x, \, y \in \mathbb{R}^n,$$

$$d(x,y) = \max\{|x_i - y_i|, \, \forall \, x, \, y \in \mathbb{R}^n\}.$$

Note que é possível em um dado conjunto M ter mais que uma métrica.

Exemplo 3.4. Espaço funcional $\mathcal{C}([a,b],\mathbb{R})$ é o conjunto de todas as funções reais definidas e contínuas no intervalo fechado $[a,b]$ com imagem em $\mathbb{R}$. São métricas neste conjunto:

$$d(f,g) = \max\{|f(t) - g(t)|; t \in [a,b]\}$$

$$d(f,g) = \int_a^b |f(t) - g(t)| \, dt$$

para toda função $f, \, g \in C[a,b]$.

Exemplo 3.5. Métrica do supremo ou métrica da convergência uniforme.

Seja M um conjunto qualquer não vazio. Uma função $f : M \to \mathbb{R}$ diz-se limitada quando existe uma constante $c_f > 0$ tal que $\forall \, x \in M$, $|f(x)| \leq c_f$. Considere $\mathcal{B}(M,\mathbb{R})$ o conjuntos de todas as funções limitadas $f : M \to \mathbb{R}$. Diz-se que $\mathcal{B}(M,\mathbb{R})$ é um espaço métrico com a métrica d definida por

$$d(f,g) = \sup_{x \in M} |f(x) - g(x)| \tag{3.2}$$

para toda função f, $g \in \mathcal{B}(M,\mathbb{R})$. Note que esta métrica está bem definida, pois se $f, g \in \mathcal{B}(M,\mathbb{R})$, então existem constantes $c_f > 0$ e $c_g > 0$ tais que $|f(x)| \leq c_f$ e $|g(x)| \leq c_g$, respectivamente. Logo, por desigualdade triangular

$$|f(x) - g(x)| \leq |f(x)| + |-g(x)| = |f(x)| + |g(x)| \leq c_f + c_g = C.$$

Logo, $(f - g) \in \mathcal{B}(M,\mathbb{R})$ e, portanto, (3.2) está bem definida. Com a métrica definida em (3.2), o espaço $\mathcal{B}(M,\mathbb{R})$ é métrico. Com efeito,

$M1)$ $\forall\, f \in \mathcal{B}(M,\mathbb{R})$ verifica-se que

$$d(f,f) = \sup_{x\in M} |f(x) - f(x)| = \sup 0 = 0.$$

Se $f \neq g$ então existe pelo menos um ponto $x_0 \in M$ tal que $f(x_0) \neq g(x_0)$ e, assim, $|f(x_0) - g(x_0)| > 0$. Como

$$\sup_{x\in M} |f(x) - g(x)| \geq |f(x_0) - g(x_0)| > 0$$

então

$$d(f,g) > 0.$$

$M2)$ De $|f(x) - g(x)| = |g(x) - f(x)|$ segue que

$$d(f,g) = \sup_{x\in M} |f(x) - g(x)| = \sup_{x\in M} |g(x) - f(x)| = d(g,f).$$

$M3)$ $\forall\, f,\ g,\ h \in \mathcal{B}(M,\mathbb{R})$ tem-se

$$d(f,g) = \sup_{x\in M} |f(x) - g(x)| \text{ e } d(g,h) = \sup_{x\in M} |g(x) - h(x)|.$$

Como $|f(x) - g(x)| \geq 0$ e $|g(x) - h(x)| \geq 0$ vale que

$$\sup_{x\in M} |f(x) - g(x)| + \sup_{x\in M} |g(x) - h(x)| = \sup_{x\in M}\{|f(x) - g(x)| + |g(x) - h(x)|\}.$$

Por desigualdade triangular, para todo $x \in M$ ocorre

$$|f(x) - h(x)| \leq |f(x) - g(x)| + |g(x) - h(x)|.$$

Assim,

$$\sup_{x\in M} |f(x) - h(x)| \leq \sup_{x\in M}\{|f(x) - g(x)| + |g(x) - h(x)|\}$$

ou seja,

$$\sup_{x\in M} |f(x) - h(x)| \leq \sup_{x\in M} |f(x) - g(x)| + \sup_{x\in M} |g(x) - h(x)|.$$

Portanto,

$$d(f,h) \leq d(f,g) + d(g,h).$$

Exemplo 3.6. Espaço de sequência ℓ^∞ é o conjunto de todas as sequência limitadas de números reais ou complexos. Se $x \in \ell^\infty$, $x = (\eta_k)_{k\in\mathbb{N}}$, então

$$|\eta_k| \leq c_k,$$

onde c_k é um número real que pode depender de x. A métrica em ℓ^∞ é definida por

$$d(x,y) = \sup_{k\in N} |\eta_k - \mu_k|,$$

onde $y = (\mu_k)_{k\in\mathbb{N}}$. Note que a métrica está bem definida, pois se $x = (\eta_k)$ e $y = (\mu_k)$ pertencem ao espaço ℓ^∞, então existem constantes α_k e β_k, respectivamente, tais que $|\eta_k| \leq \alpha_k$ e $|\mu_k| \leq \beta_k$. Logo,

$$|x - y| = |\eta_k - \mu_k| \leq |\eta_k| + |\mu_k| \leq \alpha_k + \beta_k = c,$$

onde c é um constante real.

Exemplo 3.7. Espaço de sequência $\mathbf{s}$ é o conjunto de todas as sequências (limitadas ou ilimitadas) de números complexos ou reais cuja métrica é definida por

$$d(x,y) = \sum_{i=i}^{\infty} \frac{1}{2^i} \frac{|\eta_i - \mu_i|}{1 + |\eta_i - \mu_i|},$$

onde $x = (\eta_i)$ e $y = (\mu_i)$. Note que a métrica está bem definida, pois

$$\frac{1}{2^i} \frac{|\eta_i - \mu_i|}{1 + |\eta_i - \mu_i|} \leq \frac{1}{2^i}$$

e como $\displaystyle\sum_{i=i}^{\infty} \frac{1}{2^i}$ converge então, pelo critério da comparação, $\displaystyle\sum_{i=i}^{\infty} \frac{1}{2^i} \frac{|\eta_i - \mu_i|}{1 + |\eta_i - \mu_i|}$ também converge.

Facilmente verifica-se que as condições $M1$ e $M2$ estão satisfeitas. Para provar $M3$ considere a função definida por

$$f(t) = \frac{t}{1+t},\ t \in \mathbb{R}.$$

Derivando f na variável t obtem-se

$$f'(t) = \frac{1}{(1+t)^2} > 0,$$

para todo t. Assim f é monótona crescente, ou seja, se $t_1 \leq t_2$, então $f(t_1) \leq f(t_2)$. Por consequência,

$$|a + b| \leq |a| + |b| \Rightarrow f(|a+b|) \leq f(|a| + |b|),$$

ou seja,

$$
\begin{aligned}
\frac{|a+b|}{1+|a+b|} &\leq \frac{|a|+|b|}{1+|a|+|b|} \\
&= \frac{|a|}{1+|a|+|b|} + \frac{|b|}{1+|a|+|b|} \\
&\leq \frac{|a|}{1+|a|} + \frac{|b|}{1+|b|}
\end{aligned}
$$

Admitindo que $a = \eta_i - \mu_i$ e $b = \mu_i - \xi_i$, segue que $a + b = \eta_i - \xi_i$, onde $x = (\eta_i)$, $y = (\mu_i)$ e $z = (\xi_i)$. Substituindo na desigualdade anterior

$$\frac{|\eta_i - \xi_i|}{1+|\eta_i - \xi_i|} \leq \frac{|\eta_i - \mu_i|}{1+|\eta_i - \mu_i|} + \frac{|\mu_i - \xi_i|}{1+|\mu_i - \xi_i|}.$$

Multiplicando ambos os lados da desigualdade por $1/2^i$ e somando os infintos termos, resulta que

$$
\begin{aligned}
\sum_{i=i}^{\infty} \frac{1}{2^i} \frac{|\eta_i - \xi_i|}{1+|\eta_i - \xi_i|} &\leq \sum_{i=i}^{\infty} \frac{1}{2^i} \left(\frac{|\eta_i - \mu_i|}{1+|\eta_i - \mu_i|} + \frac{|\mu_i - \xi_i|}{1+|\mu_i - \xi_i|} \right) \\
&= \sum_{i=i}^{\infty} \frac{1}{2^i} \frac{|\eta_i - \mu_i|}{1+|\eta_i - \mu_i|} + \sum_{i=i}^{\infty} \frac{1}{2^i} \frac{|\mu_i - \xi_i|}{1+|\mu_i - \xi_i|}
\end{aligned}
$$

ou seja,

$$d(x, z) \leq d(x, y) + d(y, z).$$

Define-se para $1 \leq p < \infty$ o conjunto $\ell^p = \{x = (\eta_i)_{i \in \mathbb{N}}; \sum |\eta_i|^p < \infty\}$. Para provar que ℓ^p é um espaço métrico são necessárias as duas proposições abaixo.

Proposição 3.1. ***Desigualdade de Hölder***: *Sejam $1 < p, q < \infty$ tais que $\dfrac{1}{p} + \dfrac{1}{q} = 1$ (diz-se que p e q são expoentes conjugados). Então, para todo $x = (\eta_i) \in \ell^p$, $y = (\mu_i) \in \ell^p$ a série $\displaystyle\sum_{i=1}^{\infty} |\eta_i \, \mu_i|$ é absolutamente convergente e tem-se a seguinte desigualdade:*

$$\sum_{i=1}^{\infty} |\eta_i \, \mu_i| \leq \left(\sum_{i=1}^{\infty} |\eta_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^{\infty} |\mu_i|^q \right)^{\frac{1}{q}}. \tag{3.3}$$

Demonstração. Sejam $\alpha > 0$ e $\beta > 0$ tais que $\alpha + \beta = 1$, e $\rho \geq 0$, $\nu \geq 0$. Então vale a desigualdade

$$\rho^\alpha \nu^\beta \leq \alpha\rho + \beta\nu. \tag{3.4}$$

Com efeito, se $\nu = 0$ (ou $\rho = 0$) a desigualdade é óbvia. Seja $\nu \neq 0$. Divida a desigualdade por ν e tome $t = \rho/\nu \geq 0$,

$$t^{\alpha} \leq \alpha t + \beta.$$

Como $\beta = 1 - \alpha$ então,

$$t^{\alpha} - \alpha t \leq 1 - \alpha$$

Considere $f(t) = t^{\alpha} - \alpha t$. Provar a desigualdade (3.4) equivale a mostrar que $f(t) = t^{\alpha} - \alpha t \leq 1 - \alpha$, $\forall\, t \geq 0$.

De fato, analise a derivada de f em relação a t. Como $f'(t) = \alpha(t^{\alpha-1} - 1)$, segue que, para $0 \leq t < 1$, f é estritamente crescente, para $t > 1$ f é estritamente decrescente e, para $t = 1$, f atinge seu ponto máximo, ou seja, $1 - \alpha$. Portanto, vale a desigualdade (3.4).

Para provar a Desigualdade de Hölder note que se $x = 0$ ou $y = 0$ a desigualdade é óbvia. Suponha então que $x \neq 0$ e $y \neq 0$. Seja $j \in \mathbb{N}$, fixo. Considere

$$\rho = \frac{|\eta_j|^p}{\displaystyle\sum_{i=1}^{\infty} |\eta_i|^p},\ \nu = \frac{|\mu_j|^q}{\displaystyle\sum_{i=1}^{\infty} |\mu_i|^q},\ \alpha = \frac{1}{p} \text{ e } \beta = \frac{1}{q}$$

Da desigualdade (3.4), segue que

$$\frac{|\eta_j\, \mu_j|}{\left(\displaystyle\sum_{i=1}^{\infty} |\eta_i|^p\right)^{\frac{1}{p}} \left(\displaystyle\sum_{i=1}^{\infty} |\mu_i|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p}\, \frac{|\eta_j|^p}{\displaystyle\sum_{i=1}^{\infty} |\eta_i|^p} + \frac{1}{q}\, \frac{|\mu_j|^q}{\displaystyle\sum_{i=1}^{\infty} |\mu_i|^q}.$$

Agora, somando para cada índice $j \in \mathbb{N}$, tem-se

$$\frac{\displaystyle\sum_{j=1}^{\infty} |\eta_j\, \mu_j|}{\left(\displaystyle\sum_{i=1}^{\infty} |\eta_i|^p\right)^{\frac{1}{p}} \left(\displaystyle\sum_{i=1}^{\infty} |\mu_i|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p}\, \frac{\displaystyle\sum_{j=1}^{\infty} |\eta_j|^p}{\displaystyle\sum_{i=1}^{\infty} |\eta_i|^p} + \frac{1}{q}\, \frac{\displaystyle\sum_{j=1}^{\infty} |\mu_j|^q}{\displaystyle\sum_{i=1}^{\infty} |\mu_i|^q}$$

ou seja,

$$\frac{\displaystyle\sum_{j=1}^{\infty} |\eta_j\, \mu_j|}{\left(\displaystyle\sum_{i=1}^{\infty} |\eta_i|^p\right)^{\frac{1}{p}} \left(\displaystyle\sum_{i=1}^{\infty} |\mu_i|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

Assim,

$$\sum_{i=1}^{\infty} |\eta_i\, \mu_i| \leq \left(\sum_{i=1}^{\infty} |\eta_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{\infty} |\mu_i|^q\right)^{\frac{1}{q}}.$$

□

Note que se $p = 2$ então $q = 2$ e, esta desigualdade, torna-se a *Desigualdade de Cauchy-Schwarz* para soma.

Proposição 3.2. ***Desigualdade de Minkowski****: Seja $1 \leq p < \infty$. Se $x = (\eta_i) \in \ell^p$, $y = (\mu_i) \in \ell^p$, então $x + y = (x_i + y_i) \in \ell^p$ e tem-se a seguinte desigualdade:*

$$\left(\sum_{i=1}^{\infty} |\eta_i + \mu_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{\infty} |\eta_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |\mu_i|^p\right)^{\frac{1}{p}}. \tag{3.5}$$

Demonstração. Para $p = 1$ a Desigualdade de Minkowski é imediata. Suponha $1 < p, q < \infty$ tal que $\dfrac{1}{p} + \dfrac{1}{q} = 1$ e, pela Desigualdade de Hölder, segue que

$$\begin{aligned}
\sum_{i=1}^{m} |\eta_i + \mu_i|^p = & \sum_{i=1}^{m} |\eta_i + \mu_i|^{p-1} |\eta_i + \mu_i| \\
\leq & \sum_{i=1}^{m} |\eta_i + \mu_i|^{p-1} (|\eta_i| + |\mu_i|) \\
\leq & \sum_{i=1}^{m} |\eta_i||\eta_i + \mu_i|^{p-1} + \sum_{i=1}^{m} |\mu_i||\eta_i + \mu_i|^{p-1} \\
\leq & \left(\sum_{i=1}^{m} |\eta_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{m} |\eta_i + \mu_i|^{(p-1)q}\right)^{\frac{1}{q}} \\
+ & \left(\sum_{i=1}^{m} |\mu_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{m} |\eta_i + \mu_i|^{(p-1)q}\right)^{\frac{1}{q}} \\
= & \left(\sum_{i=1}^{m} |\eta_i + \mu_i|^p\right)^{\frac{1}{q}} \left[\left(\sum_{i=1}^{m} |\eta_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{m} |\mu_i|^p\right)^{\frac{1}{p}}\right].
\end{aligned}$$

Suponha $\displaystyle\sum_{i=1}^{m} |\eta_i + \mu_i|^p \neq 0$ e como $\dfrac{1}{p} = 1 - \dfrac{1}{q}$ segue

$$\left(\sum_{i=1}^{m} |\eta_i + \mu_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{m} |\eta_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{m} |\mu_i|^p\right)^{\frac{1}{p}}.$$

Fazendo $m \to \infty$, a direita da desigualdade obtem-se duas séries convergentes, pois $x, y \in \ell^p$. Assim, a série a esquerda também converge, ou seja, vale a desigualdade (3.5), de Minkowski. □

Exemplo 3.8. Espaço ℓ^p.

Seja $p \geq 1$ um número real fixo. Por definição cada elemento de ℓ^p é uma sequência $x = (\eta_i)_{i\in\mathbb{N}} = (\eta_1, \eta_2, \eta_3, ...)$ de números reais ou complexos cuja soma converge, ou seja,

$$\sum_{i=1}^{\infty} |\eta_i|^p < \infty. \tag{3.6}$$

A métrica é definida por

$$d(x,y) = \left(\sum_{i=1}^{\infty} |\eta_i - \mu_i|^p\right)^{\frac{1}{p}}. \tag{3.7}$$

onde $x = (\eta_i)_{i\in\mathbb{N}}$, $y = (\mu_i)_{i\in\mathbb{N}} \in \ell^p$. Note que (3.7) satisfaz as condições $M1$ e $M2$ da definição de espaços métricos, desde que (3.6) seja satisfeita. Resta provar $M3$. Com efeito, sejam $x = (\eta_i)$, $y = (\mu_i)$, $z = (\xi_i) \in \ell^p$ e, pela proposição 3.2, tem-se

$$\begin{aligned}
d(x,z) = & \left(\sum_{i=1}^{m} |\eta_i - \xi_i|^p\right)^{\frac{1}{p}} \\
\leq & \left(\sum_{i=1}^{m} [|\eta_i - \mu_i| + |\mu_i - \xi_i|]^p\right)^{\frac{1}{p}} \\
\leq & \left(\sum_{i=1}^{m} |\eta_i - \mu_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{m} |\mu_i - \xi_i|^p\right)^{\frac{1}{p}} \\
= & \; d(x,y) + d(y,z).
\end{aligned}$$

Portanto, ℓ^p é um espaço métrico com a métrica definida por (3.7).

Definição 3.2. *Seja (M,d) um espaço métrico. Um subespaço (Q,d_Q) de (M,d) é um subconjunto $Q \subset M$ com a métrica d_Q sobre Q definida por*

$$d_Q = d\,|_{Q\times Q}\colon Q \times Q \to \mathbb{R}.$$

*d_Q é chamada **a métrica induzida** em Q por d.*

3.1 Conjuntos abertos, fechados e compactos

Dentre os inúmeros subconjuntos em um espaço métrico, há alguns que naturalmente se destacam, como por exemplo, os conjuntos abertos, fechados e compactos. Em geral, as propriedades que envolvem estes conjuntos possuem várias aplicações, e por isso algumas dessas propriedades serão demonstradas nesta seção. Será convencionada a notação M para indicar o espaço métrico (M, d) nesta e nas demais seções.

Definição 3.3. *Seja a um ponto qualquer no espaço métrico M. Dado um número real $r > 0$ define-se:*

*i. 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,*

$$B(a, r) = \{x \in M;\, d(x, a) < r\}.$$

*ii. A **bola fechada** de centro a e raio r é o conjunto $B[a, r]$ dos pontos de M que estão a uma distância menor do que ou igual a r do ponto a, ou seja,*

$$B[a, r] = \{x \in M;\, d(x, a) \leq r\}.$$

*iii. A **esfera** de centro a e raio r é o conjunto $S(a, r)$ dos pontos $x \in M$ tais que $d(x, a) = r$, ou seja,*

$$S(a, r) = \{x \in M;\, d(x, a) = r\}.$$

Definição 3.4. *Sejam X um subconjunto não vazio do espaço métrico M e a um ponto de M. Define-se a **distância do ponto** a **ao subconjunto** X como o número real*

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

Exemplo 3.9. Sejam $S_1 = \{(x, y) \in \mathbb{R}^2;\, x^2 + y^2 = 1\}$ o círculo unitário do plano e $0 \in \mathbb{R}^2$ a origem. Então, para todo $z \in S_1$ tem-se $d(0, z) = 1$, logo $d(0, S_1) = 1$.

Definição 3.5. *Sejam X e Y subconjuntos não vazios do espaço métrico M. Define-se a **distância entre os subconjuntos** X e Y como sendo*

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

Observação 3.2. Note que quando $X \cap Y \neq \emptyset$, tem-se $d(X, Y) = 0$. Por outro lado, $d(X, Y) = 0$ não implica em $X \cap Y \neq \emptyset$. Com efeito, considere $X = (-\infty, 0)$ e $Y = (0, +\infty)$ subconjuntos da reta $\mathbb{R}$. Neste caso, $d(X, Y) = 0$ mas $X \cap Y = \emptyset$.

Exemplo 3.10. Considere o espaço $\mathbb{R}^2$ com a métrica definida por

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

para todo $x = (x_1, x_2)$ e $y = (y_1, y_2)$ em $\mathbb{R}^2$. Sejam $A = \{(a, b) \in \mathbb{R}^2;\, b = 0\}$ e $B = \{(a, b) \in \mathbb{R}^2;\, ab = 1\}$ subconjuntos de $\mathbb{R}^2$. Note que $d(A, B) = 0$. Para ver isto, basta verificar que dado $\epsilon > 0$ existe $x \in A$ e $y \in B$ tal que $d(x, y) < \epsilon$. De fato, dado $\epsilon > 0$ existe $n \in \mathbb{N}$ tal que $\dfrac{1}{n} < \epsilon$. Considere $x = (n, 0) \in A$ e $y = (n, 1/n) \in B$, logo

$$d(x, y) = \sqrt{(n-n)^2 + \left(0 - \frac{1}{n}\right)^2} = \sqrt{\frac{1}{n^2}} = \frac{1}{n} < \epsilon.$$

Portanto, $d(A, B) = 0$.

Definição 3.6. *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 $d(x, a) < r$ implica $x \in X$. Chama-se o interior de X em M ao conjunto int(X) formado pelos pontos interiores a X.*

Definição 3.7. *Chama-se* ***fronteira*** *de X em M 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 $M - X$.*

Exemplo 3.11. Seja $\mathbb{Q}$ o conjunto dos números racionais. Note que, no conjunto $\mathbb{R}$, o $\text{int}(\mathbb{Q}) = \emptyset$, pois nenhum intervalo aberto pode ser formado apenas por números racionais. Por outro lado, a fronteira $\partial\,\mathbb{Q} = \mathbb{R}$ pois qualquer intervalo aberto contém números racionais e irracionais.

Definição 3.8. *Um subconjunto A de um espaço métrico M diz-se* ***aberto*** *em M quando todos os seus pontos são interiores, ou seja, int$(A) = A$. Assim, $A \subset M$ é aberto se, e somente se, $A \cap \partial A = \emptyset$. Ou seja, $A \subset M$ é um aberto em M se para todo $x \in A$, existe $r > 0$ tal que $B(x, r) \subset A$.*

Proposição 3.3. *Em qualquer espaço métrico M, uma bola aberta $B(a, r)$ é um conjunto aberto.*

Demonstração. Seja $x \in B(a, r)$. Então $d(x, a) < r$ e escolhendo $\epsilon = r - d(x, a) > 0$ tem-se $B(x, \epsilon) \subset B(a, r)$. Com efeito, se $y \in B(x, \epsilon)$ então $d(x, y) < \epsilon$ e portanto $d(a, y) \leq d(a, x) + d(x, y) < d(a, x) + \epsilon = r$. Logo $y \in B(a, r)$. $\square$

Proposição 3.4. *Seja M um espaço métrico. Então*

i. *$\emptyset$ e M são abertos.*

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

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

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

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

iii. 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. □

Observação 3.3. A interseção de uma família infinita de abertos pode não ser um conjunto aberto. Note que se $A = \{a\}$, $a \in M$, A não é aberto. Mas todo ponto $a \in M$ é interseção de uma família enumerável de abertos, ou seja, $\{a\} = \bigcap_{n \in \mathbb{N}} B\left(a, \frac{1}{n}\right)$. De fato, se $x \neq a$ então $d(x,a) > 0$, logo existe $n_x \in \mathbb{N}$ tal que $d(x,a) > \frac{1}{n_x}$, isto mostra que $x \notin B\left(a, \frac{1}{n_x}\right)$, ou seja, apenas o ponto a pertence a todas as bolas $B\left(a, \frac{1}{n}\right)$, $n \in \mathbb{N}$.

Proposição 3.5. *Seja $X \subset M$. Um subconjunto U é aberto em X se, e somente se, $U = V \cap X$, onde V é um aberto em M.*

Demonstração. ($\Rightarrow$) Seja U um aberto em X. Então, para cada $x \in U$, existe $r_x > 0$ tal que $B_X(x, r_x) \subset U$, onde $B_X(x, r_x)$ denota a bola aberta em X, de centro em x e raio r_x. Agora, para todo $x \in X$, $B_X(x, r_x) = B_M(x, r_x) \cap X$, assim

$$\{x\} \subset B_M(x, r_x) \cap X = B_X(x, r_x) \subset U \Rightarrow$$

$$\bigcup_{x \in U} \{x\} \subset \bigcup_{x \in U} B_M(x, r_x) \cap X \subset U \Rightarrow$$

$$U \subset \bigcup_{x \in U} B_M(x, r_x) \cap X \subset U \Rightarrow$$

$$U = \bigcup_{x \in U} B_M(x, r_x) \cap X.$$

Basta tomar $V = \bigcup_{x \in U} B_M(x, r_x)$ que, pela proposição 3.4, é aberto.

($\Leftarrow$) Seja $U = V \cap X$, onde V é um aberto em M. Então, para cada $x \in U$, existe $r_x > 0$ tal que $B_M(x, r_x) \subset V$. Como $B_X(x, r_x) = B_M(x, r_x) \cap X$, para todo $x \in M$, tem-se $B_X(x, r_x) \subset V \cap X = U$. Portanto, U é aberto em X. □

Definição 3.9. *Sejam M e N espaços métricos. Diz-se que a aplicação $f : M \to N$ é* ***contínua*** *no ponto $a \in M$ quando para qualquer $\epsilon > 0$ dado, é possível obter $\delta > 0$ tal que $d(x, a) < \delta$ implica $d(f(x), f(a)) < \epsilon$. Diz-se que $f : M \to N$ é contínua quando ela é contínua em todos os pontos $a \in M$.*

Dizer que f é contínua no ponto $a \in M$ equivale a dizer que para todo ponto $x \in M$ que pertence a bola $B(a, \delta)$, f os transforma em pontos da bola $B(f(a), \epsilon)$.

Exemplo 3.12. O *conjunto das aplicações limitadas descontínuas é aberto* em $\mathcal{B}(M, N)$, onde $\mathcal{B}(M, N)$ é o conjunto das funções $f : M \to N$ limitadas. Com efeito, considere D_a o conjunto das aplicações $f : M \to N$ que são descontínuas no ponto $a \in M$. Assim, para $f \in D_a$, existe $\epsilon > 0$ tal que para todo $\delta > 0$ pode-se obter $x_\delta \in M$ com $d(x_\delta, a) < \delta$ e $d(f(x_\delta), f(a)) \geq 3\epsilon$.

Se $g \in \mathcal{B}(M, N)$ e $d(f, g) < \epsilon$, então $g \in D_a$. De fato, nestas condições, para todo $\delta > 0$ temos

$$3\epsilon \leq d(f(x_\delta), f(a)) \leq d(f(x_\delta), g(x_\delta)) + d(g(x_\delta), g(a)) + d(g(a), f(a)).$$

Na soma acima, $d(f(x_\delta), g(x_\delta)) < \epsilon$ e $d(g(a), f(a)) < \epsilon$, assim a desigualdade torna-se

$$3\epsilon \leq d(f(x_\delta), g(x_\delta)) + d(g(x_\delta), g(a)) + d(g(a), f(a)) < 2\epsilon + d(g(x_\delta), g(a))$$

ou seja,

$$d(g(x_\delta), g(a)) > \epsilon,$$

logo, $g \in D_a$. Agora, seja D o conjunto de todas as aplicações limitadas descontínuas $f : M \to N$. Então $D = \bigcup_{a \in M} D_a$, que é uma união de abertos, é aberto em $\mathcal{B}(M, N)$.

Definição 3.10. *Sejam M e N espaços métricos. Um aplicação $f : M \to N$ diz-se* ***uniformemente contínua*** *quando, para qualquer $\epsilon > 0$ dado, pode-se obter $\delta > 0$ tal que, para quaisquer $x, y \in M$, $d(x, y) < \delta$ então $d(f(x), f(y)) < \epsilon$.*

Definição 3.11. *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.*

Da definição 3.9, para que a aplicação $f : M \to N$ seja contínua no ponto $a \in M$, basta mostrar que para cada vizinhança $U_{f(a)} \subset N$ de $f(a)$ existe uma vizinhança $V_a \subset M$ de a tal que $f(V_a) \subset U_{f(a)}$.

Definição 3.12. *Seja M um espaço métrico. Um conjunto $F \subset M$ diz-se* ***fechado*** *em M se seu complementar $F^c = M - F$ é aberto.*

Exemplo 3.13. A bola $B[a, r]$, definida anteriormente, é um conjunto fechado. Note que $M - B[a, r]$ é um conjunto aberto. De fato, seja $y \in M - B[a, r]$ então $d(a, y) > r$.

Considere $s = d(a,y) - r > 0$, então $B(y,s) \subset M - B[a,r]$. Assim, se $x \in B(y,s)$ então $d(x,y) < s$. Agora, por desigualdade triangular

$$d(a,y) \leq d(a,x) + d(x,y),$$

ou seja,

$$d(a,x) \geq d(a,y) - d(x,y) > d(a,y) - s = d(a,y) - d(a,y) + r = r.$$

Logo, $x \in M - B[a,r]$, portanto o conjunto $B[a,r]$ é fechado.

Definição 3.13. *Um ponto x diz-se **aderente** a um subconjunto X do espaço métrico M se para todo $r > 0$, $B(x,r) \cap X \neq \emptyset$, ou seja, quando toda vizinhança de x contiver pelo menos um ponto de X.*

Observação 3.4. Todo ponto que pertence a X é um ponto aderente a X. Além disso, os pontos da fronteira ∂X também são aderentes a X.

Definição 3.14. *O conjunto de todos os pontos aderentes a X chama-se **fecho** e denota-se por $\overline{X}$. Portanto, escrever $a \in \overline{X}$ é o mesmo que afirmar que o ponto a é aderente a X em M.*

Definição 3.15. *Um subconjunto $X \subset M$ diz-se **denso** em M quando $\overline{X} = M$, ou seja, quando toda bola aberta em M contém algum ponto de X, ou ainda, para cada aberto $A \neq \emptyset$ em M, tem-se $A \cap X \neq \emptyset$.*

Exemplo 3.14. O conjunto $\mathbb{Q}$ dos números racionais é denso em $\mathbb{R}$. Note que $\overline{\mathbb{Q}} = \mathbb{R}$, pois toda bola aberta em $\mathbb{R}$ contém números racionais.

Definição 3.16. *Um espaço métrico M chama-se **separável** se ele contém um subconjunto enumerável que é denso.*

Exemplo 3.15. O conjunto $\mathbb{R}$ é separável, pois o subconjunto $\mathbb{Q}$ é enumerável e denso em $\mathbb{R}$.

Proposição 3.6. *Seja $F \subset M$, tem-se $\overline{F} = F$ se, e somente se, $M - F$ é aberto. Em outras palavras, um conjunto é fechado se, e somente se, contém todos os seus pontos aderentes.*

Demonstração. ($\Rightarrow$) Como $\overline{F} = F$, por hipótese, segue que F é fechado. Seja $a \notin F$, logo a não é aderente a F. Segue disto, que para todo ponto $a \in M - F$ existe uma bola aberta $B(a,r)$ que não contém os pontos de F. Assim, existe $r > 0$ tal que $B(a,r) \subset M - F$, ou seja, $M - F$ é aberto.

($\Leftarrow$) Se $M - F$ é aberto, então para todo ponto $a \in M - F$, existe $r > 0$ tal que $B(a,r) \subset M - F$, ou seja, existe uma bola aberta $B(a,r)$ que não contém pontos de F. Neste caso, os pontos que não pertencem a F não são aderentes a F. Logo, $\overline{F} = F$ e, portanto, F é fechado. $\square$

Observação 3.5. Quando um conjunto não é fechado, não se pode concluir que ele seja aberto. Por exemplo, um intervalo do tipo $(a, b] \subset \mathbb{R}$ não é aberto e nem fechado em $\mathbb{R}$. No espaço $\mathbb{Q}$, dos números racionais com a métrica $|x - y|$ (induzida de $\mathbb{R}$), o intervalo $(\sqrt{2}, \pi) = \{x \in \mathbb{Q};\ \sqrt{2} < r < \pi\}$ além de ser um subconjunto aberto é também fechado, pois seu complementar $\mathbb{Q} - (\sqrt{2}, \pi)$ é o conjunto aberto $(-\infty, \sqrt{2}) \cup (\pi, +\infty) \cap \mathbb{Q}$.

Proposição 3.7. *Os subconjuntos fechados de um espaço métrico M satifazem as seguintes propriedades,*

i. *$\emptyset$ e M são fechados.*

ii. *A reunião $F_1 \cup \ldots \cup F_n$ de um número finito de subconjuntos fechados $F_1, \ldots, F_n \subset M$ é um subconjunto fechado em M.*

iii. *A interseção $\bigcap_{\lambda \in J} F_\lambda$ de uma família qualquer $(F_\lambda)_{\lambda \in J}$, onde J é um conjunto de índices, de subconjuntos fechados $F_\lambda \subset M$ é um conjunto fechado em M.*

Demonstração. Para mostrar cada um dos itens basta considerar o complementar de cada conjunto fechado e usar a proposição 3.4 (página 39) para abertos.

i. Imediato.

ii. Seja $A_1 = F_1^c, \ldots, A_n = F_n^c$ abertos em M. Desse modo, $A_1 \cap \ldots \cap A_n = F_1^c \cap \ldots \cap F_n^c = (F_1 \cup \ldots \cup F_n)^c$ é aberto e portanto $F_1 \cup \ldots \cup F_n$ é fechado em M.

iii. Seja $A_\lambda = F_\lambda^c$ para cada $\lambda \in J$. Então cada A_λ é aberto e portanto sua união $\bigcup_{\lambda \in J} A_\lambda = \bigcup_{\lambda \in J} F_\lambda^c = \left[\bigcap_{\lambda \in J} F_\lambda\right]^c$ é aberto em M. Logo, $\bigcap_{\lambda \in J} F_\lambda$ é fechado. □

Proposição 3.8. *Seja $X \subset M$. Um conjunto U é fechado em X se, e somente se, $U = V \cap X$, onde V é um fechado em M.*

A demonstração será omitida pois é feita de modo semelhante ao da proposição 3.5(página 40).

Observação 3.6. A reunião de uma família infinita de fechados pode não ser um conjunto fechado. Por exemplo, considere $M = \mathbb{R}$, cada conjunto unitário $\{x\} \subset \mathbb{R}$ é fechado em $\mathbb{R}$, pois $\mathbb{R} - \{x\}$ é aberto. Seja $(a, b) = \{a < x < b;\ x \in \mathbb{R}\}$ um intervalo aberto. Se a união infinita de fechados fosse um conjunto fechado, então (a, b) seria um conjunto fechado, pois (a, b) é reunião de infinitos conjuntos $\{x\}$, onde $x \in \mathbb{R}$.

Definição 3.17. *Seja X um subconjunto do espaço métrico M. Um ponto $a \in M$ chama-se* ***ponto de acumulação*** *de X quando toda bola aberta de centro a contém algum ponto de X distinto de a. Em outras palavras, quando toda vizinhança V_a contém algum ponto diferente de a, ou seja, $V_a \cap (X - \{a\}) \neq \emptyset$. Denota-se por X' o conjunto de todos os pontos de acumulação de X em M. O conjunto X' é chamado o derivado do conjunto X.*

Exemplo 3.16. Seja $X = \left\{\frac{1}{n},\, n \in \mathbb{N}\right\} \subset \mathbb{R}$. Note que o único ponto de acumulação de X é o zero.

O estudo de convergência de sequências de um dado espaço é importante e tem várias aplicações.

Definição 3.18. *Uma* ***sequência*** *em um espaço métrico M é uma função $\phi : \mathbb{N} \to M$ que a cada $n \in \mathbb{N}$ associa um único elemento $x_n \in M$. Denota-se qualquer sequência por (x_n). Diz-se que uma sequência (x_n) é* ***convergente*** *para um ponto $x \in M$ se, para todo $\epsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $n > n_0$ implica $d(x_n, x) < \epsilon$. x é chamado* ***o limite*** *de (x_n) e escreve-se $\lim\limits_{n\to\infty} x_n = x$, ou ainda, $x_n \to x$ quando $n \to \infty$.*

Definição 3.19. *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 $d(x_m, x_n) \leq c$ para quaisquer $m, n \in \mathbb{N}$.*

Lema 3.1. *Seja M um espaço métrico. Então,*

i. Toda sequência convergente em M é limitada e seu limite é único.

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

Demonstração. i. Seja (x_n) uma sequência em M e suponha que $x_n \to x$, onde $x \in M$. Para $\epsilon = 1$, existe $n_0 \in \mathbb{N}$ tal que $n > n_0$ implica $d(x_n, x) < 1$. Considere $a = \max\{d(x_1, x), \ldots, d(x_{n_0}, x)\}$ então

$$d(x_n, x) < 1 < 1 + a, \text{ se } n > n_0$$

e

$$d(x_n, x) \leq a < 1 + a, \text{ se } n \leq n_0,$$

ou seja, $d(x_n, x) < 1 + a$ para todo $n \in \mathbb{N}$.

Logo, pela desigualdade triangular

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < 2(1 + a).$$

Considerando $c = 2(1 + a)$, então $d(x_n, x_m) < c$. Logo $(x_n)_{n\in\mathbb{N}}$ é limitada.

Agora, sejam $x, z \in M$ tais que $x = \lim\limits_{n\to\infty} x_n$ e $z = \lim\limits_{n\to\infty} x_n$. Então

$$0 \leq d(x, z) \leq d(x, x_n) + d(x_n, z) \to 0 + 0,$$

e segue disto que $d(x, z) = 0$ e, portanto, $x = z$.

ii. Por desigualdade triangular tem-se

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

ou seja,

$$d(x_n, y_n) - d(x, y) \leq d(x_n, x) + d(y_n, y). \tag{3.8}$$

Por outro lado,

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

ou seja,

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

Com isto tem-se

$$-d(x_n, x) - d(y_n, y) \leq d(x_n, y_n) - d(x, y). \tag{3.9}$$

Das equações (3.8) e (3.9) segue,

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

ou seja,

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

Como por hipótese $x_n \to x$ e $y_n \to y$, da desigualdade acima resulta que

$$|d(x_n, y_n) - d(x, y)| \to 0,$$

quando $n \to \infty$. Portanto, $d(x_n, y_n) \to d(x, y)$ em M. □

Proposição 3.9. *Se $x_n \to a$ então toda subsequência de (x_n) converge para a.*

Demonstração. Como $x_n \to a$, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $n > n_0$ então $d(x_n, a) < \epsilon$. Seja $\mathbb{N}' = \{n_1 < n_2 < \ldots < n_k < \ldots\}$ um subconjunto infinito de $\mathbb{N}$. Neste caso, existe $k_0 \in \mathbb{N}$ tal que $n_{k_0} > n_0$. Logo, para todo $k \in \mathbb{N}$, com $k > k_0$ tem-se $n_k > n_{k_0} > n_0$, então $d(x_{n_k}, a) < \epsilon$. □

Definição 3.20. *Uma sequência $(x_n)_{n\in\mathbb{N}}$ em um espaço métrico M chama-se uma **sequência de Cauchy** quando para qualquer $\epsilon > 0$ dado, existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ implica $d(x_m, x_n) < \epsilon$.*

Note que se a sequência é de Cauchy, seus termos vão se tornando cada vez mais próximos à medida que cresce o índice n.

Teorema 3.1. *Toda sequência convergente em um espaço métrico M é uma sequência de Cauchy em M.*

Demonstração. Seja (x_n) uma sequência convergente em M. Se $x_n \to x$, então para todo $\epsilon > 0$, existe n_0 tal que $n > n_0$ implica que

$$d(x_n, x) < \frac{\epsilon}{2}.$$

Assim, se $m, n > n_0$, por desigualdade triangular, segue que

$$d(x_m, x_n) \leq d(x_m, x) + d(x_n, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Portanto, (x_n) é uma sequência de Cauchy. □

Observação 3.7. Nem toda sequência de Cauchy em M é convergente em M. De fato, considere a sequência de números racionais $(x_n) = (1; 1,4; 1,41; 1,414; ...)$ que converge para $\sqrt{2}$. Pela proposição anterior, (x_n) é de Cauchy em $\mathbb{R}$ e, portanto, é de Cauchy em $\mathbb{Q}$. Mas (x_n) não é convergente em $\mathbb{Q}$, pois $\sqrt{2} \notin \mathbb{Q}$.

Proposição 3.10. *Toda sequência de Cauchy é limitada.*

Demonstração. Seja (x_n) uma sequência de Cauchy no espaço métrico M, então dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ implica $d(x_m, x_n) < \epsilon$. Em particular, considerando $\epsilon = 1$, existe $n_0 \in \mathbb{N}$ tal que $m, n > n_0$ então $d(x_m, x_n) < 1$. Logo o conjunto $X = \{x_{n_0+1}, x_{n_0+2}, \ldots\}$ dos termos da sequência é limitado, ou seja, $d(x_m, x_n) \leq 1$ para $m, n > n_0$. Por outro lado,

$$\{x_1, x_2, \ldots, x_n, \ldots\} = \{x_1, x_2, \ldots, x_{n_0}\} \cup X.$$

Logo, como cada conjunto a direita da igualdade é limitado, segue que a sequência (x_n) é limitada. □

Observação 3.8. Nem toda sequência limitada é de Cauchy. Basta observar que uma sequência do tipo $(1, 0, 1, 0, \ldots)$ embora limitada, tem-se $d(x_n, x_{n+1}) = 1$ para todo $n \in \mathbb{N}$, logo não é de Cauchy.

Teorema 3.2. *Seja F um subconjunto não vazio do espaço métrico M. Então:*

i. $x \in \overline{F}$ se, e somente se, existe uma sequência $(x_n)_{n\in\mathbb{N}} \subset F$ tal que $x_n \to x$.

ii. F é fechado se, e somente se, para uma sequência (x_n) em F e $x_n \to x$ implicar $x \in F$.

Demonstração. (i) Seja x um elemento de $\overline{F}$. Se $x \in F$, considere a sequência $(x, x, x, \ldots)$ em F. Esta sequência tende a x. Se $x \notin F$, então x é um ponto de acumulação de elementos de F. Assim, para cada $n \in \mathbb{N}$, a bola aberta $B\left(x, \dfrac{1}{n}\right)$ contém um ponto $x_n \in F$. Note que quando $n \to \infty$, $\dfrac{1}{n} \to 0$ e, portanto, $x_n \to x$.

Reciprocamente, se (x_n) é uma sequência em F e $x_n \to x$ então ou $x \in F$ ou x é um ponto de acumulação de F, com $x_n \neq x$, ou seja, x é aderente a F em M. Assim, $x \in \overline{F}$, pela definição de fecho.

(ii) Pela proposição 3.6 (página 42), F é fechado se, e somente se, $F = \overline{F}$, e o resultado segue pelo item (i) deste teorema. □

Definição 3.21. *Um espaço métrico M é dito ser* **compacto** *se toda sequência em M possui uma subsequência convergente. Um subconjunto $F \subset M$ é compacto se toda sequência em F possui uma subsequência convergente cujo limite é um elemento de F.*

Lema 3.2. *Um subconjunto compacto F de um espaço métrico M é fechado e limitado.*

Demonstração. Para todo $x \in \overline{F}$ existe uma sequência (x_n) em F tal que $x_n \to x$. Como F é compacto, pela definição 3.21, segue que $x \in F$. Ainda, como x é um elemento arbitrário tem-se que $\overline{F} \subset F$ e como $F \subset \overline{F}$ segue que $F = \overline{F}$, ou seja, F é fechado.

Agora, suponha que F não seja limitado. Então para cada $n \in \mathbb{N}$ existe $y_n \in F$ tal que $d(y_n, a) > n$, onde $a \in F$ é qualquer elemento fixo. Neste caso, note que (y_n) não pode ter uma subsequência convergente, mas isto implica que F não é compacto, contrariando a hipótese. Portanto F é limitado. $\square$

Como uma consequência deste lema e outros resultados da Análise pode-se demonstrar a proposição abaixo, cuja a prova pode ser encontrada na referência [1].

Proposição 3.11. *$X \subset \mathbb{R}^n$ é compacto se, e somente se, é fechado e limitado.*

Proposição 3.12. *Uma sequência de Cauchy que possui uma subsequência convergente é convergente e tem o mesmo limite que a subsequência.*

Demonstração. Seja (x_n) uma sequência convergente em um espaço métrico M. Considere (x_{n_k}) uma subsequência convergente para um ponto $x \in M$. Assim, dado $\epsilon > 0$, existe $p \in \mathbb{N}$ tal que para $n_k > p$ tem-se $d(x_{n_k}, x) < \dfrac{\epsilon}{2}$.

Como (x_n) é de Cauchy, para todo $\epsilon > 0$ existe $q \in \mathbb{N}$ tal que $m, n > q$ implica $d(x_m, x_n) < \dfrac{\epsilon}{2}$. Seja $n_0 = \max\{p, q\}$. Assim, para todo $n > n_0$ existe $n_k > n_0$ tal que

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Portanto, $x_n \to x$. $\square$

Teorema 3.3. *Sejam M e N espaços métricos. Uma aplicação $f : M \to N$ é contínua no ponto $x_0 \in M$ se, e somente se, para qualquer sequência (x_n) em M com $x_n \to x_0$ tem-se $f(x_n) \to f(x_0)$.*

Demonstração. Seja f uma aplicação contínua em x_0, então dado qualquer $\epsilon > 0$ existe $\delta > 0$ tal que

$$d(x, x_0) < \delta \Rightarrow d(f(x), f(x_0)) < \epsilon.$$

Agora, se $x_n \to x_0$ então existe $n_0 \in \mathbb{N}$ tal que para $n > n_0$ tem-se $d(x_n, x_0) < \delta$. Assim, para todo $n > n_0$, $d(f(x_n), f(x_0)) < \epsilon$. Portanto, $f(x_n) \to f(x_0)$.

Por outro lado, basta mostrar que se $x_n \to x_0$ implica $f(x_n) \to (x_0)$, então f será uma aplicação contínua. Suponha que isto não ocorra, ou seja, existe $\delta > 0$ tal que para qualquer $\epsilon > 0$ existe $x \neq x_0$ tal que

$$d(x, x_0) < \delta \Rightarrow d(f(x), f(x_0)) \geq \epsilon.$$

Em particular, para $\delta = \dfrac{1}{n}$ existe x_n satisfazendo

$$d(x_n, x_0) < \frac{1}{n} \Rightarrow d(f(x_n), f(x_0)) \geq \epsilon.$$

Note que $x_n \to x_0$ mas $f(x_n) \nrightarrow f(x_0)$. Isto contradiz $f(x_n) \to (x_0)$. □

Definição 3.22. *Diz-se que uma sequência de funções $f_n : X \to M$, onde X é um conjunto qualquer,* ***converge pontualmente*** *em X para a função $f : X \to M$ quando para cada $x \in X$, a sequência $(f_1(x), \ldots, f_n(x), \ldots)$ tem limite $f(x)$ em M, ou seja, para cada $x \in X$, tem-se $\lim_{n\to\infty} f_n(x) = f(x)$. Em outras palavras, $f_n \to f$ pontualmente em X quando dados $\epsilon > 0$ e $x \in X$ existe $n_0 = n_0(\epsilon, x) \in \mathbb{N}$ tal que $n > n_0$ implica $d(f_n(x), f(x)) < \epsilon$.*

Definição 3.23. *Diz-se que uma sequência de funções $f_n : X \to M$, onde X é um conjunto qualquer,* ***converge uniformemente*** *em X para a função $f : X \to M$ quando para qualquer $\epsilon > 0$ dado, pode-se obter $n_0 \in \mathbb{N}$ (dependendo apenas de ϵ) tal que $n > n_0$ então $d(f_n(x), f(x)) < \epsilon$ para qualquer $x \in X$.*

Proposição 3.13. *Toda aplicação uniformemente contínua transforma sequências de Cauchy em sequências de Cauchy.*

Demonstração. Considere $f : M \to N$ uniformemente contínua e (x_n) uma sequência de Cauchy em M. Assim, da continuidade de f, para todo $\epsilon > 0$ existe $\delta > 0$ tal que se $x, y \in M$ e $d(x, y) < \delta$ então $d(f(x), f(y)) < \epsilon$. Por outro lado, sendo (x_n) de Cauchy, dado $\delta > 0$, existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ tem-se $d(x_m, x_n) < \delta$ o que implica, $d(f(x_m), f(y_n)) < \epsilon$, ou seja, $(f(x_n))$ é uma sequência de Cauchy em N. □

3.2 Espaços métricos completos

Definição 3.24. *Diz-se que um espaço métrico M é* ***completo*** *quando toda sequência de Cauchy em M é convergente em M.*

Todo espaço métrico M que não é completo pode ser estendido por adição de novos elementos até que se obtenha um espaço métrico completo, este processo é chamado de completamento do espaço métrico. Para obter este completamento basta adicionar os elementos aderentes a M. Para tanto, é preciso encontrar uma aplicação $f : M \to N$, onde N é um espaço métrico completo. A aplicação f é chamada de imersão isométrica e será definida mais abaixo.

Exemplo 3.17. A conjunto $\mathbb{R}$ dos números reais é um espaço métrico completo. Com efeito, seja (x_n) uma sequência de Cauchy em $\mathbb{R}$, então para qualquer $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ tem-se $d(x_m, x_n) < \epsilon$. Fazendo para cada $n \in \mathbb{N}$, $X_n = \{x_n, x_{n+1}, x_{n+2}, \ldots\}$, de forma que $X_1 = \{x_1, x_2, x_3, \ldots\}$, $X_2 = \{x_2, x_3, x_4, \ldots\}$, e assim sucessivamente. Note que $X_1 \supset X_2 \supset \ldots \supset X_n \supset \ldots$ e os conjuntos X_n são limitados, uma vez que (x_n) é uma sequência limitada, pela proposição 3.10 (página 46).

Seja $a_n = \inf X_n$, $n \in \mathbb{N}$. Então, $a_1 \leq a_2 \leq \ldots \leq a_n \leq \ldots \leq b = \sup X_1$, ou seja, (a_n) é uma sequência monótona não-decrescente. Pelo teorema 2.3 (página 22), existe o número $a = \lim\limits_{n\to\infty} a_n$. Pela proposição 3.12 (página 47), basta mostrar que o número a é limite de uma subsequência de (x_n) para poder concluir que $a = \lim\limits_{n\to\infty} x_n$.

Como $a_n \to a$ então dado $\epsilon > 0$ existe $n_1 \in \mathbb{N}$ tal que $m > n_1$ tem-se $a_m \in (a-\epsilon, a+\epsilon)$, ou seja, $a-\epsilon < a_m < a+\epsilon$. Sendo $a_m = \inf X_m$, existe $n \geq m$ (e portanto $n > n_1$) tal que $a_m \leq x_n < a + \epsilon$, isto é, $a - \epsilon < x_n < a + \epsilon$ e, portanto, $x_n \to a$.

Definição 3.25. *Se $E \subset M$ e $\{V_a\}_{a\in A}$ é uma família de conjuntos tal que $E \subset \bigcup\limits_{a\in A} V_a$ diz-se que $\{V_a\}_{a\in A}$ é uma* ***cobertura*** *de E. Se M é um espaço métrico, diz-se que $E \subset M$ é* ***totalmente limitado*** *se para cada $\epsilon > 0$, E pode ser coberto por um número finito de bolas de raio ϵ.*

A partir desta definição tem-se o seguinte teorema.

Teorema 3.4. *Se E é um subconjunto de um espaço métrico M, as seguintes afirmações são equivalentes:*

i. E é completo e totalmente limitado.

ii. Toda sequência em E tem uma subsequência que converge para um ponto de E.

Demonstração. $(i) \Rightarrow (ii)$ Suponha que E é completo e totalmente limitado. Seja (x_n) uma sequência em E. Pela definição 3.25, E pode ser coberto por um número finito de bolas de raio $\epsilon = \dfrac{1}{2}$. Note que ao menos uma dessas bolas deve conter x_n para um número infinito de índices. Considere $x_n \in B_1$ para $n \in \mathbb{N}_1$.

Agora, $E \cap B_1$ pode ser coberto por um número finito de bolas de raio $\epsilon = \dfrac{1}{2^2}$ e portanto uma dessas bolas contém x_n para um número infinito de índices e seja $x_n \in B_2$ para $n \in \mathbb{N}_2$.

Procedendo desta forma, obtem-se uma sequência de bolas B_j de raio $\epsilon = \dfrac{1}{2^j}$ e uma sequência decrescente de subconjuntos infinitos $\mathbb{N}_j$ de $\mathbb{N}$ tal que $x_n \in B_j$ para todo $n \in \mathbb{N}_j$. Escolhendo $n_1 \in \mathbb{N}_1$, $n_2 \in \mathbb{N}_2$, $\ldots$ tal que $n_1 < n_2 < \ldots$ então (x_{n_k}) é uma sequência de Cauchy pois para $k > j$, $d(x_{n_j}, x_{n_k}) < \dfrac{2}{2^j}$ e como E é completo segue que esta subsequência converge em E.

$(ii) \Rightarrow (i)$ Esta implicação será demonstrada pela contra positiva. Suponha que E não é completo, ou seja, existe uma sequência de Cauchy (x_n) em E que não converge para um ponto de E. Neste caso, nenhuma subsequência de (x_n) pode convergir em E pois caso contrário (x_n) seria convergente e ambas teriam o mesmo limite. Suponha então que E não seja totalmente limitado, ou seja, existe $\epsilon > 0$ tal que E não pode ser coberto por um número finito de bolas de raio ϵ. Escolha $x_n \in E$ da seguinte forma: comece com qualquer $x_1 \in E$ e tendo escolhido $x_1, \ldots, x_n$ escolha $x_{n+1} \in E - \bigcup_{i=1}^{n} B_\epsilon(x_i)$. Então $d(x_n, x_m) \geq \epsilon$ para todo m, n e, portanto (x_n) não possui subsequência convergente. □

Proposição 3.14. *Todo subespaço fechado $F \subset M$ em um espaço métrico M completo é completo. Reciprocamente, todo subespaço completo de qualquer espaço métrico é fechado.*

Demonstração. Seja $F \subset M$ fechado, com M completo. Considere (x_n) uma sequência de Cauchy em F, então existe um ponto $x \in M$ tal que $x_n \to x$ em M. Como F é fechado em M, tem-se que $x \in F$ e, portanto, F é completo.

Por outro lado, seja $F \subset M$ um subespaço completo e M um espaço métrico. Então, considere (x_n) uma sequência em F convergente para algum ponto $x \in M$. Pela proposição 3.1 (página 45), (x_n) é de Cauchy. Logo existe $x_1 \in F$ tal que $x_n \to x_1$ em F. Pela unicidade do limite, tem-se $x = x_1$ e portanto, F é fechado em M. □

Proposição 3.15. *Se o espaço métrico M é completo, então o conjunto $\mathcal{B}_\alpha(X, M)$, formado pelas funções que estão a uma distância finita de α, é completo, sejam quais forem X e $\alpha : X \to M$.*

Demonstração. Dizer que o conjunto $\mathcal{B}_\alpha(X, M)$ é formado pelas funções que estão a uma distância finita de α equivale a dizer que $\mathcal{B}_\alpha(X, M)$ representa o conjunto das aplicações $f : X \to M$ tais que $d(f, \alpha) = \sup d(f(x), \alpha(x)) < \infty$ (métrica da convergência uniforme ou métrica do supremo).

Seja (f_n) uma sequência de Cauchy em $\mathcal{B}_\alpha(X, M)$. Esta sequência é limitada, logo existe uma constante $c > 0$ tal que $d(f_n(x), \alpha(x)) \leq d(f_n, \alpha) \leq c$ para todo $n \in \mathbb{N}$ e todo $x \in X$. Fixando $x \in X$, a sequência $(f_n(x))_{n\in\mathbb{N}}$ é de Cauchy em M. Como M é completo, existe $f \in M$ tal que para cada $x \in X$, $\lim_{n\to\infty} f_n(x) = f(x)$. Isto define a aplicação $f : X \to M$ como sendo o limite pontual da sequência (f_n). Como $d(f_n(x), \alpha(x)) \leq c$ para todo $n \in \mathbb{N}$ e $x \in X$, fazendo $n \to \infty$ nesta desigualdade, segue que $d(f(x), \alpha(x)) \leq c$ para todo $x \in X$ e, portanto, $f \in \mathcal{B}_\alpha(X, M)$.

Agora, como (f_n) é de Cauchy, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ tem-se $d(f_m(x), f_n(x)) < \epsilon$ para qualquer $x \in X$. Fazendo $m \to \infty$ nesta desigualdade segue que $d(f(x), f_n(x)) \leq \epsilon$ para qualquer $x \in X$ e $n > n_0$, ou seja, $f_n \to f$ uniformemente em X. □

Segue abaixo o critério de Cauchy para convergência uniforme.

Corolário 3.1. *Seja M um espaço métrico completo. A fim de que uma sequência de aplicações $f_n : X \to M$ convirja uniformemente em X, é necessário e suficiente que, para qualquer $\epsilon > 0$ dado, existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ implique $d(f_m(x), f_n(x)) < \epsilon$ para todo $x \in X$.*

Demonstração. Se $f_n \to f$ uniformemente em X então dado qualquer $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $n > n_0$ obtem-se $d(f_n(x), f(x)) < \epsilon$ para todo $x \in X$. Assim, $f_n \in \mathcal{B}_f(X, M)$ para todo n suficientemente grande e $\lim\limits_{n\to\infty} f_n = f$ neste espaço, pois $\mathcal{B}_f(X, M)$ é um espaço métrico completo e então a sequência (f_n) é de Cauchy em $\mathcal{B}_f(X, M)$, ou seja, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ tem-se $d(f_m(x), f_n(x)) < \epsilon$.

Suponha agora que (f_n) seja de Cauchy, isto é, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ tem-se $d(f_m(x), f_n(x)) < \epsilon$. Deste fato segue que $d(f_n(x), f_{n_0+1}(x)) < \epsilon$. Considere então $\alpha = f_{n_0+1}$ e $\epsilon = 1$, assim $d(f_n(x), f_{n_0+1}(x)) < 1$ para todo $x \in X$, logo $d(f_n, f_{n_0+1}) \leq 1$, ou seja, $f_n \in \mathcal{B}_\alpha(X, M)$ se $n > n_0$.

Resta provar que f_n converge uniformemente em X. Note que a sequência (f_n) é de Cauchy no espaço métrico completo $\mathcal{B}_\alpha(X, M)$, pois $d(f_n(x), f_{n_0+1}(x)) < \epsilon$ para todo $x \in X$. Logo, existe $f \in \mathcal{B}_\alpha(X, M)$ tal que $f_n \to f$. Portanto, pela proposição 3.15 (página 50), $f_n \to f$ uniformemente em X. □

Corolário 3.2. *Sejam M e N espaços métricos, onde N é completo. Se uma sequência de aplicações contínuas $f_n : M \to N$ converge uniformemente em um subconjunto $X \subset M$ então (f_n) converge uniformemente em $\overline{X}$.*

Demonstração. Antes de demonstrar o corolário, note que se $\varphi : M \to \mathbb{R}$ é uma função contínua tal que $\varphi(x) < \epsilon$ para todo $x \in X$, então $\varphi(x) \leq \epsilon$ para todo $x \in \overline{X}$. Com efeito, o conjunto de todos os pontos $x \in M$ tais que $\varphi(x) \leq \epsilon$ é fechado em M e contém X, logo contém $\overline{X}$, o que vale dizer que $\varphi(x) \leq \epsilon$ para todo $x \in \overline{X}$.

Se $f_n : M \to N$ converge uniformemente em X então, pelo corolário 3.1, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, n < n_0$ tem-se $d(f_m(x), f_n(x)) < \epsilon$ para todo $x \in X$. Fixando m, n e escrevendo $\varphi(x) = (f_m(x), f_n(x))$ segue, pela observação acima, que para todo $m, n < n_0$ implica $d(f_m(x), f_n(x)) \leq \epsilon$ para todo $x \in \overline{X}$. Pela recíproca do corolário 3.1, tem-se que (f_n) converge uniformemente em $\overline{X}$. □

Observação 3.9. Como foi visto na observação 3.7 (página 46) que o conjunto $\mathbb{Q}$ não é um espaço métrico completo, mas $\mathbb{Q}$ admite um completamento que, neste caso, é o conjunto $\mathbb{R}$.

As definições seguintes darão uma ideia de como um espaço métrico pode ser completado.

Definição 3.26. *Sejam M e N espaços métricos. Uma aplicação $f : M \to N$ chama-se uma* ***imersão isométrica*** *quando $d(f(x), f(y)) = d(x, y)$ para quaisquer $x, y \in M$. Neste caso, diz-se também que f preserva distâncias.*

Definição 3.27. *Um **completamento** de um espaço métrico M é um par $(\hat{M}, \varphi)$, onde $\hat{M}$ é completo e $\varphi : M \to \hat{M}$ é uma imersão isométrica cuja imagem $\varphi(M)$ é densa em $\hat{M}$, ou seja, $\hat{M} = \overline{\varphi(M)}$.*

Na referência [12] é demonstrado que todo espaço métrico possui um completamento e que este é único.

Exemplo 3.18. $\mathbb{R}$ é o completamento de $\mathbb{Q}$. Com efeito, considere a imersão isométrica $\varphi : \mathbb{Q} \to \mathbb{R}$ tal que $\varphi(x) = x$, para todo $x \in \mathbb{Q}$. Considere em $\mathbb{Q}$ a métrica definida por $d(x, y) = |x - y|$, para quaisquer $x, y \in \mathbb{Q}$. Note que φ preserva distâncias, pois

$$d(\varphi(x), \varphi(y)) = |\varphi(x) - \varphi(y)| = |x - y| = d(x, y),\ \forall\, x, y \in \mathbb{Q}.$$

Ainda, como $\varphi(\mathbb{Q}) = \mathbb{Q}$ e $\overline{\mathbb{Q}} = \mathbb{R}$ segue que $\overline{\varphi(\mathbb{Q})} = \mathbb{R}$, ou seja, $\varphi(\mathbb{Q})$ é denso em $\mathbb{R}$. Portanto, o completamento para $\mathbb{Q}$ é o par $(\mathbb{R}, \varphi)$, em que $\mathbb{R}$ é um espaço métrico completo e $\varphi : \mathbb{Q} \to \mathbb{R}$ é uma imersão isométrica.

O próximo capítulo é sobre Espaços Normados e Espaços de Banach. Um espaço normado é um espaço vetorial com uma norma definida e um espaço de Banach é um espaço vetorial normado que é completo. Será demonstrado que todo espaço normado é um espaço métrico quando $d(x, y)$ é igual a norma de $(x - y)$. Pode-se concluir disto que um espaço de Banach é um espaço completo na métrica induzida pela norma.

4 Espaços normados e de Banach

Para falar de espaço normado é preciso, inicialmente, introduzir alguns conceitos de espaços vetoriais, uma vez que um espaço normado é qualquer espaço vetorial que possui uma norma definida. Dentre os conceitos que serão apresentados destacam-se: independência linear, base e dimensão de um espaço vetorial, além disso, o famoso Lema de Zorn. Importantes exemplos de espaços vetoriais são os normados e os de Banach (espaço normado completo), que serão abordados neste capítulo.

4.1 Espaços vetoriais

As operações definidas são as de adição de vetores e multiplicação de vetor por escalar, como mostra a definição abaixo.

Definição 4.1. *Um* ***espaço vetorial*** *sobre um corpo* $\mathbb{K}$ *(*$\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{K} = \mathbb{C}$*) é um conjunto* V *não vazio cujos elementos são chamados de vetores.* V *é munido de duas operações algébricas.*

Define-se ***adição de vetores*** *da seguinte forma:*

$$\begin{array}{rccc} (+): & V \times V & \to & V \\ & (x, y) & \mapsto & x + y \end{array}$$

satisfazendo as seguintes propriedades:

A1) $x + y = y + x,\ \forall\, x,\ y \in V$ *(comutativa);*

A2) $x + (y + z) = (x + y) + z,\ \forall\, x,\ y,\ z \in V$ *(associativa);*

A3) Existe um vetor 0 *em* V*, chamado vetor nulo tal que*

$$x + 0 = x,\ \forall\, x \in V;$$

A4) Para cada vetor x *existe um elemento* $-x$ *em* V*, chamado elemento oposto, tal que*

$$x + (-x) = 0,\ \forall\, x \in V.$$

Define-se ***multiplicação por escalar*** *como segue:*

$$\begin{array}{rccc} (\,.\,): & V \times \mathbb{K} & \to & V \\ & (x, \alpha) & \mapsto & \alpha x \end{array}$$

satisfazendo:

$M1)$ $\alpha(\beta x) = (\alpha\beta)x,\ \forall\, x \in V$ *e* $\forall\, \alpha,\ \beta \in \mathbb{K}$;
$M2)$ $\alpha(x + y) = \alpha x + \alpha y,\ \forall\, x,\ y \in V$ *e* $\forall\, \alpha \in \mathbb{K}$;
$M3)$ $(\alpha + \beta)x = \alpha x + \beta x,\ \forall\, x \in V$ *e* $\forall\, \alpha,\ \beta \in \mathbb{K}$;
$M4)$ $1x = x,\ \forall\, x \in V$.

Uma consequência da definição é a unicidade do elemento neutro e do elemento oposto. De fato,

i. Sejam 0 e $0'$ em V satisfazendo $A3)$ então, pelas propriedades $A1)$ e $A3)$ tem-se

$$0' = 0 + 0' = 0' + 0 = 0.$$

ii. Seja $x \in V$, pela propriedade $A4)$ $-x \in V$. Considere $y \in V$ tal que $x + y = 0$. Utilizando as propriedades $A1)$, $A2)$ e $A3)$ tem-se

$$-x = -x + 0 = -x + (x + y) = (-x + x) + y = 0 + y = y.$$

Note que se V é um espaço vetorial e $\mathbb{K} = \mathbb{R}$ diz-se que V é um *espaço vetorial real.* Se os elementos de $\mathbb{K}$ forem números complexos, ou seja, $\mathbb{K} = \mathbb{C}$, diz-se que V é um *espaço vetorial complexo.*

Segue abaixo alguns exemplos de espaços vetoriais.

Exemplo 4.1. O espaço vetorial $\mathbb{R}^n$ é definido como o conjunto de todas as n-uplas $x = (x_1, x_2, ..., x_n)$ em $\mathbb{R}^n$ considerando as seguintes operações

$$x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n),\ \forall\, x,\ y \in \mathbb{R}^n,$$

$$\alpha x = (\alpha x_1, \alpha x_2, ..., \alpha x_n),\ \forall\, x \in \mathbb{R}^n,\ \forall\, \alpha \in \mathbb{R}.$$

Além disso, o elemento nulo para a adição é o vetor $(0, 0, ..., 0)$ e o elemento oposto de x é o vetor $-x = (-x_1, -x_2, ..., -x_n)$.

Exemplo 4.2. Espaço $\mathbb{C}^n$ sobre $\mathbb{R}$ ou sobre $\mathbb{C}$. Este espaço é o conjunto de todas as n-uplas $x = (x_1, x_2, ..., x_n) \in \mathbb{C}^n$ com as operações

$$x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n),\ \forall\, x,\ y \in \mathbb{C}^n,$$

$$\alpha x = (\alpha x_1, \alpha x_2, ..., \alpha x_n),\ \forall\, x \in \mathbb{C}^n,\ \forall\, \alpha \in \mathbb{K}.$$

Exemplo 4.3. O espaço $\mathcal{P}_n(\mathbb{R})$ é o conjunto formado por todos os polinônimos de grau menor ou igual a n, mais o polinômio nulo. Sejam $p(x) = a_0 + a_1 x + ... + a_n x^n$ e $q(x) = b_0 + b_1 x + ... + b_n x^n$ em V e $\lambda \in \mathbb{R}$. As operações consideradas em V são

$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + ... + (a_n + b_n)x^n,$$

$$\lambda p(x) = (\lambda a_0) + (\lambda a_1)x + ... + (\lambda a_n)x^n.$$

Exemplo 4.4. O espaço $M_{m\times n}(\mathbb{R})$ é o conjunto das matrizes $m \times n$ munido das operações:

$$A + B = (A_{ij} + B_{ij})_{m\times n},\ \forall\, A,\ B \in V,$$

onde $A = (A_{ij})$ e $B = (B_{ij})$, e

$$\alpha A = (\alpha A)_{ij},\ \forall\, A \in V \text{ e } \alpha \in \mathbb{R}.$$

Exemplo 4.5. O espaço $\mathcal{C}([a, b], \mathbb{R})$ é o conjunto de todas as funções contínuas definidas no intervalo $[a, b]$ com imagem em $\mathbb{R}$. As operações consideradas neste conjunto são as usuais:

i. A soma de dois vetores f e g em V é o vetor $f + g \in V$ tal que

$$(f + g)(x) = f(x) + g(x),\ \forall\ x \in [a, b],$$

ii. O produto do escalar $\alpha \in \mathbb{R}$ e a função $f \in V$ é a função αf tal que

$$(\alpha f)(x) = \alpha(f(x)),\ \forall\ x \in [a, b],$$

onde $f + g,\ \alpha f : [a, b] \to \mathbb{R}$ contínuas.

Exemplo 4.6. O espaço ℓ^2 é o conjunto de todas as sequências $x = (\eta_i)_{i\in\mathbb{N}} = (\eta_1, \eta_2, ...)$ reais (ou complexas) de *quadrado-somável*, isto é,

$$|\eta_1|^2 + |\eta_2|^2 + |\eta_3|^2 + ... = \sum_{i=1}^{\infty} |\eta_i|^2 < \infty.$$

Dados $x = (\eta_i)_{i\in\mathbb{N}}$ e $y = (\mu_i)_{i\in\mathbb{N}}$, as operações consideradas são:

$$x + y = (\eta_1, \eta_2, ...) + (\mu_1, \mu_2, ...) = (\eta_1 + \mu_1, \eta_2 + \mu_2, ...) = (\eta_i + \mu_i)_{i\in\mathbb{N}},$$

$$\alpha x = \alpha(\eta_1, \eta_2, ...) = (\alpha\eta_1, \alpha\eta_2, ...) = (\alpha\eta_i)_{i\in\mathbb{N}}, \forall\ \alpha \in \mathbb{K}.$$

Note que essas operações estão bem definidas. Com efeito, sejam $x, y \in \ell^2$ então

$$\sum_{i=1}^{\infty} |\eta_i|^2 < \infty \quad \text{e} \quad \sum_{i=1}^{\infty} |\mu_i|^2 < \infty.$$

Pela proposição 3.2 (página 36), considerando $p = 2$, segue que

$$\sum_{i=1}^{\infty} |\eta_i + \mu_i|^2 < \infty,$$

e, portanto, $x + y \in \ell^2$. Da mesma forma,

$$\sum_{i=1}^{\infty} |\alpha||\eta_i|^2 < \infty,$$

logo, $\alpha x \in \ell^2$. Não é difícil verificar que as oito propriedades da definição 4.1 (página 53) estão satisfeitas, pois cada η_i (i-ésima componente de x) é um elemento do corpo $\mathbb{K}$.

Definição 4.2. *Seja V um espaço vetorial sobre $\mathbb{K}$. Diz-se que um subconjunto não vazio W de V é* ***subespaço vetorial*** *de V se para quaisquer dois vetores w_1 e w_2 de W e cada escalar α em $\mathbb{K}$ o vetor $w_1 + \alpha w_2$ pertence a W.*

Exemplo 4.7. Todo espaço vetorial não nulo V admite pelo menos dois subespaços, o subespaço nulo, denotado por $\{0\}$, e o próprio espaço V que são chamados *subespaços triviais ou impróprios.* Os demais são chamados *subespaços próprios* de V.

4.1.1 Lema de Zorn, base e dimensão

Em Álgebra Linear, uma base para um espaço vetorial V é um conjunto de vetores que, satisfazendo algumas propriedades, produz todo o espaço V. Este é um dos conceitos mais importantes quando o assunto é espaço vetorial. Antes, porém, seguem outras definições.

Definição 4.3. *Sejam V um espaço vetorial sobre $\mathbb{K}$ e vetores $x_1, x_2, ..., x_n$ elementos de V. Diz-se que $x \in V$ é uma* ***combinação linear*** *de $x_1, x_2, ..., x_n$ se existem escalares $\alpha_1, \alpha_2, ..., \alpha_n$ tais que*

$$x = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n. \tag{4.1}$$

Seja Y o conjunto de todas as combinações lineares dos vetores $x_1, x_2, ..., x_n$. Assim, Y é um subespaço de V e é chamado de **subespaço gerado** pelos vetores $x_1, x_2, ..., x_n$. Denota-se o subespaço gerado por $Y = [x_1, x_2, ..., x_n]$.

Dois conceitos importantes em Álgebra Linear e também muito utilizados em outras áreas da Matemática, são os conceitos de dependência linear e independência linear, definidos como segue.

Definição 4.4. *Seja V um espaço vetorial sobre $\mathbb{K}$. Considere os vetores $x_1, x_2, ..., x_n \in V$. Diz-se que os vetores $x_1, x_2, ..., x_n$ são* ***linearmente independentes*** *(l.i.) se a combinação linear*

$$\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n = 0 \tag{4.2}$$

implicar em $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$.

Diz-se que os vetores $x_1, x_2, ..., x_n$ são ***linearmente dependentes*** *(l.d.) se não forem linearmente independentes, ou seja, se for possível encontrar ao menos um $\alpha_i \neq 0$, $i = 1, ...n$, tal que $\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n = 0$.*

No caso em que o conjunto de vetores $\{x_1, x_2, ..., x_n\}$ for linearmente dependente, pelo menos um vetor deste conjunto pode ser escrito como combinação linear dos outros, isto é, se na equação $\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n = 0$ houver $\alpha_i \neq 0$ para algum $i \in \{1, 2, ...n\}$ então,

$$x_i = \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n \text{ onde } \beta_j = -\alpha_j/\alpha_i \text{ com } i \neq j.$$

A partir dos conceitos descritos acima define-se base e dimensão de um espaço vetorial.

Definição 4.5. *Seja V um espaço vetorial sobre $\mathbb{K}$. Diz-se que V é* ***finitamente gerado*** *se existe um subconjunto finito $U \subset V$ tal que $V = [U]$.*

Exemplo 4.8. O espaço $\mathbb{R}^n$ é gerado pelos vetores $e_1 = (1, 0, ..., 0)$, $e_2 = (0, 1, 0, ..., 0)$, ..., $e_n = (0, 0, ...1)$ e escreve-se $\mathbb{R}^n = [e_1, e_2, ..., e_n]$.

Exemplo 4.9. Seja $\mathcal{P}_n(\mathbb{R})$ o espaço dos polinômios com coeficientes em $\mathbb{R}$. O conjunto $U = \{1, x, ..., x^n\}$ é um conjunto gerador de $\mathcal{P}_n(\mathbb{R})$ e escreve-se $\mathcal{P}_n(\mathbb{R}) = [1, x, ..., x^n]$.

Observação 4.1. Note que o espaço vetorial $\mathcal{P}(\mathbb{R})$ formado por todos os polinômios *não é finitamente gerado.* Com efeito, se $\mathcal{P}(\mathbb{R})$ fosse finitamente gerado existiriam polinômios $p_1(x), p_2(x), ..., p_n(x)$ tais que $\mathcal{P}(\mathbb{R}) = [p_1(x), p_2(x), ..., p_n(x)]$. Seja N o maior grau dentre todos os polinômios $p_1(x), p_2(x), ..., p_n(x)$. Note que $x^{N+1} \notin [p_1(x), p_2(x), ..., p_n(x)]$. Contradição, logo $\mathcal{P}(\mathbb{R})$ é de dimensão infinita.

Proposição 4.1. *Seja $V \neq \{0\}$ um $\mathbb{K}$-espaço vetorial finitamente gerado e considere $\{v_1, \ldots, v_m\}$ um conjunto gerador de V. Então todo conjunto linearmente independente de vetores em V tem no máximo m elementos.*

Demonstração. Será demonstrado que todo conjunto de elementos de V que contenha mais do que m vetores é linearmente dependente. Para tanto, considere o subconjunto $\{u_1, \ldots, u_n\}$ de V com $n > m$. Como $\{v_1, \ldots, v_m\}$ gera V, então para cada elemento de V, em particular os vetores de A, existem escalares $\alpha_{ij} \in \mathbb{K}$ tais que, para cada $j = 1, \ldots, n$,

$$u_j = \alpha_{1j} v_1 + \ldots + \alpha_{mj} v_m = \sum_{i=1}^{m} \alpha_{ij} v_i.$$

Como $A = \{u_1, \ldots, u_n\}$ é linearmente independente, então existem escalares $\lambda_1, \ldots, \lambda_n$ tais que

$$\begin{aligned} \lambda_1 u_1 + \ldots + \lambda_n u_n = \sum_{j=1}^{n} \lambda_j u_j = & \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{m} \alpha_{ij} v_i \right) \\ = & \sum_{j=1}^{n} \sum_{i=1}^{m} \lambda_j \alpha_{ij} v_i \qquad (4.3) \\ = & \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \lambda_j \alpha_{ij} \right) v_i. \end{aligned}$$

Analisando o caso em que $\sum_{j=1}^{n} \lambda_j \alpha_{ij} = 0$ para cada $i = 1, \ldots, m$, considere o seguinte sistema

$$\begin{cases} \alpha_{11}\lambda_1 + \ldots + \alpha_{1n}\lambda_n = 0 \\ \alpha_{21}\lambda_1 + \ldots + \alpha_{2n}\lambda_n = 0 \\ \qquad\qquad \vdots \\ \alpha_{m1}\lambda_1 + \ldots + \alpha_{mn}\lambda_n = 0 \end{cases} \tag{4.4}$$

nas incógnitas $\lambda_1, \ldots, \lambda_n$ e com coeficientes $\alpha_{ij} \in \mathbb{K}$. Note que o número de equações em (4.4) é estritamente menor do que o número de incógnitas. Segue disto que (4.4) possui uma solução não nula, isto é, existem escalares $\gamma_1, \ldots, \gamma_n \in \mathbb{K}$, nem todos nulos, tais que $\sum_{j=1}^{n} \gamma_j \alpha_{ij} = 0$ para cada $i = 1, \ldots, m$. Portanto, de (4.3) deve-se ter que

$$\gamma_1 u_1 + \ldots + \gamma_n u_n = 0$$

sem que os escalares $\gamma_1, \ldots, \gamma_n$ sejam todos nulos, resultando que o conjunto $\{u_1, \ldots, u_n\}$ é linearmente dependente. Logo, qualquer conjunto linearmente independente de vetores de V possui no máximo m elementos. □

Definição 4.6. *Seja V um espaço vetorial finitamente gerado, uma* ***base*** *para V é um conjunto $\{x_1, x_2, ..., x_n\}$ linearmente independente que gera V.*

Corolário 4.1. *Seja $V \neq \{0\}$ um $\mathbb{K}$-espaço vetorial finitamente gerado. Então quaisquer duas bases de V contém o mesmo número de elementos.*

Demonstração. Sejam B e B' duas bases de V. Como V é finitamente gerado, pela proposição anterior os conjuntos B e B' são finitos com, por exemplo, m e n elementos, respectivamente. Considerando B como conjunto gerador de V e B' linearmente independente, segue da proposição acima que $n \leq m$. Por outro lado, considerando B' como conjunto gerador de V e B linearmente independente, tem-se que $m \leq n$. Logo, $m = n$. □

Definição 4.7. *Seja V um espaço vetorial finitamente gerado. Define-se a* ***dimensão*** *de V como sendo o número de elementos de uma base qualquer de V. Neste caso V é chamado um espaço de* ***dimensão finita****. Caso V não seja finitamente gerado, diz-se que V é um espaço vetorial de* ***dimensão infinita****.*

Observação 4.2. Se V for um espaço de dimensão infinita, diz-se que um subconjunto B de V é linearmente independente se para todo subconjunto finito de B a definição 4.5 estiver satisfeita.

Observação 4.3. Sejam V um espaço vetorial sobre $\mathbb{K}$ de dimensão não necessariamente finita e B um conjunto linearmente independente em V. Se existir um elemento

$v \in V$ que não seja combinação linear de elementos de B, então $B' = B \cup \{v\}$ é linearmente independente. Com efeito, considere V um espaço vetorial n-dimensional e seja $B = \{v_1, v_2, \ldots, v_n\}$ um conjunto *l.i.* em V. Suponha que existe $v \in V$ que não é combinação linear de B e sejam $\alpha_1, \ldots, \alpha_n, \alpha \in \mathbb{K}$ escalares tais que

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n + \alpha v = 0.$$

Se $\alpha \neq 0$ então pode-se escrever

$$v = -\frac{\alpha_1}{\alpha} v_1 - \frac{\alpha_2}{\alpha} v_2 - \ldots - \frac{\alpha_n}{\alpha} v_n,$$

absurdo, pois por hipótese v não é combinação linear de B. Portanto, $\alpha = 0$ e como B é *li* segue da combinação

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_n v_n + \alpha v = 0$$

que $\alpha_1 = \alpha_2 = \ldots = \alpha = 0$, ou seja, $B' = B \cup \{v\}$ é *l.i.*. Se V for de dimensão infinita, a demonstração segue de forma análoga.

Em termos mais gerais, sendo V um espaço vetorial de dimensão finita ou não, cada elemento $x \in V$, $x \neq 0$, se escreve de modo único como combinação linear dos vetores da base. De fato, sejam $x_1, x_2, ..., x_n$ vetores linearmente independentes em um espaço vetorial V qualquer. Então

$$\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n = \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n$$

assim

$$(\alpha_1 - \beta_1)x_1 + (\alpha_2 - \beta_2)x_2 + ... + (\alpha_n - \beta_n)x_n = 0$$

Como $x_1, x_2, ..., x_n$ são *l.i.* segue que $\alpha_i - \beta_i = 0$, ou seja, $\alpha_i = \beta_i$, $i \in \{1, 2, ..., n\}$.

Um resultado importante na teoria de espaços vetoriais é o fato de que todo espaço vetorial possui uma base. Para demonstrar este resultado será usado um lema chamado *Lema de Zorn*, mas para isto é preciso apresentar algumas definições, como abaixo.

Definição 4.8. *Seja X um conjunto qualquer, então*

i. Diz-se que X é um ***conjunto parcialmente ordenado*** *se há sobre X uma relação de ordem parcial, isto é, uma relação binária denotada por $\leq$ satisfazendo as seguintes condições:*

1) $x \leq x$ para todo $x \in X$ (propriedade reflexiva),

2) Se $x \leq y$ e $y \leq x$, com $x, y \in X$, então $x = y$ (propriedade antissimétrica),

3) Se $x \leq y$ e $y \leq z$, com $x, y, z \in X$, então $x \leq z$ (propriedade transitiva).

*Parcialmente enfatiza que X pode conter elementos x e y para os quais nem $x \leq y$ e nem $y \leq x$. Então x e y são chamados **elementos imcomparáveis**. Por outro lado, se $x \leq y$ ou $y \leq x$ (ou ambos acontecem), então x e y são chamados **elementos comparáveis**.*

ii. *Um **conjunto totalmente ordenado** ou **cadeia** é um conjunto parcialmente ordenado tal que cada dois elementos do conjunto são comparáveis. Em outras palavras, uma cadeia é um conjunto parcialmente ordenado que não tem elementos incomparáveis.*

iii. *Sejam X um conjunto parcialmente ordenado e A um subcojunto de X. Um **limitante superior** para A é um elemento $x \in X$ tal que para todo $a \in A$, $a \leq x$. (Conforme forem X e A, o limitante superior pode ou não existir.)*

iv. *Seja X um conjunto parcialmente ordenado. Um elemento $x_0 \in X$ é chamado **elemento maximal** de X se para todo $x \in X$ com $x_0 \leq x$ implica $x_0 = x$. (Note que X pode ou não conter um elemento maximal e, além disso, que um elemento maximal não precisa ser um limitante superior.)*

Exemplo 4.10. O conjunto $\mathbb{R}$ dos números reais com a relação $\leq$ ("menor do que ou igual") é um conjunto totalmente ordenado e não possui elemento maximal. Notação: $(\mathbb{R}, \leq)$.

Exemplo 4.11. Seja X um conjunto qualquer. O conjunto $\wp(X)$, que é conjunto de todos os subconjuntos de X, com a relação inclusão $(\subseteq)$ é um conjunto parcialmente ordenado e o único elemento maximal de $\wp(X)$ é X. Notação: $(\wp(X), \subseteq)$.

A partir dos conceitos descritos na definição 4.8, segue o Lema de Zorn, que será utilizado na demonstração do teorema 4.1.

Lema 4.1. *(**Lema de Zorn**) Seja $X \neq \emptyset$ um conjunto parcialmente ordenado. Suponha que toda cadeia $C \subset X$ tenha um limitante superior. Então X tem pelo menos um elemento maximal.*

Teorema 4.1. *Sejam V um espaço vetorial sobre $\mathbb{K}$ e C um conjunto linearmente independente em V. Então existe uma base B de V contendo C.*

Demonstração. Seja $\mathcal{P}$ a classe de todos os subconjuntos linearmente independentes de V que contém C. Note que $\mathcal{P} \neq \emptyset$ uma vez que $C \subset \mathcal{P}$ e, além disso, $\mathcal{P}$ é parcialmente ordenado por inclusão. Para usar o Lema de Zorn, é necessário mostrar que todo subconjunto totalmente ordenado de $\mathcal{P}$ tem um limitante superior. Seja $\mathcal{D} = \{A_\alpha\}_{\alpha \in L}$ um subconjunto totalmente ordenado de $\mathcal{P}$, então $\mathcal{A} = \bigcup\limits_{\alpha \in L} A_\alpha$ é um limitante superior para $\mathcal{D}$. É preciso mostrar que $\mathcal{A}$ é um subconjunto linearmente indepedente. Com efeito, seja $W = \{w_1, \ldots, w_n\}$ um subconjunto finito de $\mathcal{A}$. Então, para cada $i = 1, \ldots, n$, existe $\alpha_i \in L$ tal que $w_i \in A_{\alpha_i}$. Como $\mathcal{D}$ é totalmente ordenado, reordenando os elementos de W, se necessário, $A_{\alpha_1} \subseteq \ldots \subseteq A_{\alpha_n}$, de forma que $w_i \in A_{\alpha_n}$ para cada $i = 1, \ldots, n$. Assim, W é linearmente independente como um subconjunto finito do conjunto linearmente independente A_{α_n}. Como W é qualquer, segue que $\mathcal{A}$ é linearmente independente. Logo, $\mathcal{D}$ tem $\mathcal{A}$ como limitante superior. Pelo Lema de Zorn, $\mathcal{P}$ tem um elemento maximal B. Resta mostrar que B gera todo o espaço V. De fato, se existisse $v \in V$ que não fosse gerado por B, então $B_1 = B \cup \{v\}$ seria linearmente independente (conforme observação 4.3), contrariando a maximalidade de B. Portanto, como B gera V e é *l.i.*, segue que B é uma base para V. □

Definição 4.9. *Uma **base de Hamel**, ou simplesmente base, em um espaço vetorial V é um conjunto B linearmente independente maximal, como no teorema acima.*

Segue então um teorema sobre a dimensão de um subespaço próprio de X.

Teorema 4.2. *Seja V um espaço vetorial n-dimensional. Então qualquer subespaço próprio W de V tem dimensão menor do que n, isto é,* $\dim W < \dim V$.

Demonstração. *i*) Se $n = 0$, então $V = \{0\}$ é o subespaço nulo, ou seja, V não possui subespaço próprio.

ii) Se a $\dim W = 0$, então $W = \{0\}$ e, pelo item anterior, $V \neq W$, daí $\dim V \geq 1$. Evidentemente $\dim W \leq \dim V = n$.

iii) Se a $\dim W = n$, então W teria uma base de n elementos que também seria uma base para V. Como a $\dim V = n$ ocorre que $W = V$ o que implica que W não é subespaço próprio de V, contrariando a hipótese.

Portanto, qualquer subconjunto linearmente independente de vetores de W deve ter menos de n elementos e, assim, $\dim W < n$. □

4.1.2 Soma direta de subespaços

No estudo de espaços vetoriais, muitas vezes é conveniente representar tais espaços como soma de dois ou mais subespaços. Este conceito é chamado de soma direta e além da Álgebra Linear é aplicado, por exemplo, em diversas Estruturas Algébricas como Anéis e Grupos.

Definição 4.10. *Seja V um espaço vetorial sobre $\mathbb{K}$ e sejam U e W subespaços vetoriais de V.*

i. Diz-se que a soma $U + W = \{u + w;\ u \in U,\ w \in W\}$ é direta se $U \cap W = \{0\}$ e denota-se $U \oplus W$.

*ii. Diz-se que V é **soma direta** dos subespaços U e W se $V = U \oplus W$.*

Exemplo 4.12. O espaço $\mathbb{R}^3$ é a soma direta dos subespaços U e V, onde

$$U = \{(x, 0, 0);\ x \in \mathbb{R}\} \quad \text{e} \quad V = \{(0, y, z);\ y, z \in \mathbb{R}\}.$$

Note que $U \cap V = \{(0,0,0)\}$ e ainda, para todo $(x, y, z) \in \mathbb{R}^3$ tem-se

$$(x, y, z) = (x, 0, 0) + (0, y, z) = U + V,$$

portanto, $\mathbb{R}^3 = U \oplus V$.

O resultado seguinte será utilizado no capítulo sobre espaços de Hilbert, pois este espaço, sob certas condições, pode ser escrito como soma direta.

Teorema 4.3. *Seja V um espaço vetorial sobre $\mathbb{K}$ e sejam U e W subespaços vetoriais de V. Então, $V = U \oplus W$ se, e somente se, para cada elemento $v \in V$ existem únicos $u \in U$ e $w \in W$ tais que $v = u + w$.*

Demonstração. ($\Rightarrow$) Suponha que $V = U \oplus W$, então para cada $v \in V$ existem $u \in U$ e $w \in W$ tais que $v = u + w$. Para verificar que estes elementos são únicos, considere $u, u_1 \in U$ e $w, w_1 \in W$ tais que

$$v = u + w \quad \text{e} \quad v = u_1 + w_1,$$

ou seja,

$$u + w = u_1 + w_1,$$

o que resuta

$$u - u_1 = w_1 - w,$$

onde $u - u_1 \in U$ e $w_1 - w \in W$. Logo, $u - u_1 = w_1 - w \in U \cap W$ e como $U \cap W = \{0\}$ segue que $u - u_1 = w_1 - w = 0$, resultando que $u = u_1$ e $w = w_1$. Portanto, para cada elemento $v \in V$ existem únicos $u \in U$ e $w \in W$ tais que $v = u + w$.

($\Leftarrow$) Suponha que cada elemento $v \in V$ se escreve de forma única como uma soma $v = u + w$ com $u \in U$ e $w \in W$, ou seja, $V = U + W$. Basta mostrar que $U \cap W = \{0\}$. Considere um elemento v não nulo tal que $v \in U \cap W$. Neste caso, existem únicos $u \in U$ e $w \in W$ de forma que $v = u + w$. Como $v \in U$ e $v \in W$ então, $v = u + w = (u + v) + (w - v)$, onde $(u + v) \in U$ e $(w - v) \in W$. Pela unicidade da decomposição, deve-se ter que $u = u + v$ e $w = w - v$ resultando que $v = 0$, ou seja, $U \cap W = \{0\}$ e, portanto, $V = U \oplus W$. □

Exemplo 4.13. Seja $V = M_n(\mathbb{K})$ o espaço das matrizes quadradas de ordem n sobre o corpo $\mathbb{K}$. Uma matriz $B \in M_n(\mathbb{K})$ é chamda simétrica se $B^T = B$ e é antissimétrica se $B^T = -B$. Considere $\mathcal{S}$ o conjunto das matrizes simétricas e $\mathcal{A}$ o conjunto das matrizes antissimétricas. Os conjuntos $\mathcal{S}$ e $\mathcal{A}$ são subespaços vetoriais do espaço $M_n(\mathbb{K})$ e além disso $M_n(\mathbb{K}) = \mathcal{S} \oplus \mathcal{A}$. Com efeito, sejam $A, B \in \mathcal{S}$ matrizes simétricas e $\alpha \in \mathbb{K}$ um escalar, então

$$(\alpha A + B)^T = \alpha A^T + B^T = \alpha A + B,$$

logo, $(\alpha A + B) \in \mathcal{S}$. Da mesma forma, sejam $A, B \in \mathcal{A}$ matrizes antissimétricas e $\alpha \in \mathbb{K}$ um escalar, então

$$(\alpha A + B)^T = \alpha A^T + B^T = \alpha(-A) + (-B) = -(\alpha A + B),$$

ou seja, $(\alpha A + B) \in \mathcal{A}$.

Seja $A \in M_n(\mathbb{K})$ uma matriz qualquer. Defina as matrizes $B, C \in M_n(\mathbb{K})$ como segue,

$$B = \frac{1}{2}(A + A^T) \quad \text{e} \quad C = \frac{1}{2}(A - A^T).$$

Então $A = B + C$. De fato,

$$B + C = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T) = \frac{1}{2}A + \frac{1}{2}A^T + \frac{1}{2}A - \frac{1}{2}A^T = A.$$

Note que a matriz B é simétrica, pois

$$B^T = \frac{1}{2}(A + A^T)^T = \frac{1}{2}(A^T + (A^T)^T) = \frac{1}{2}(A^T + A) = B$$

e a matriz C é antissimétrica,

$$C^T = \frac{1}{2}(A - A^T)^T = \frac{1}{2}(A^T - (A^T)^T) = \frac{1}{2}(A^T - A) = -\frac{1}{2}(A - A^T) = -C.$$

Além disso, se $A \in \mathcal{S} \cap \mathcal{A}$ então $A \in \mathcal{S}$ e $A \in \mathcal{A}$, ou seja,

$$A^T = A \quad \text{e} \quad A^T = -A, \tag{4.5}$$

resultando que $A = -A$, isto é, em (4.5) deve-se ter que $A = 0$. Portanto, $\mathcal{S} \cap \mathcal{A} = \{0\}$.

4.2 Espaços normados

O objetivo desta seção é introduzir a definição de norma e algumas propriedades como, por exemplo, normas equivalentes, também alguns exemplos de espaços normados e as conhecidas desigualdades de Hölder e Minkowski para integrais.

Definição 4.11. *Seja E um espaço vetorial qualquer. Uma **norma** em E é uma função real*

$$\begin{array}{rccc} \|.\| : & E & \to & \mathbb{R} \\ & x & \mapsto & \|x\| \end{array}$$

que a cada elemento de E associa um número real, satisfazendo as seguintes propriedades:

N1) $\|x\| \geq 0 \ \forall\, x \in E$ *e* $\|x\| = 0 \ \Leftrightarrow x = 0,$

N2) $\|\alpha x\| = |\alpha| \|x\|, \forall\, \alpha \in \mathbb{K}, \forall\, x \in E,$

N3) $\|x + y\| \leq \|x\| + \|y\|, \forall\, x,\, y \in E$ *(desigualdade triangular).*

Diz-se que o par $(E, \|.\|)$ *é um* ***espaço vetorial normado***[1].

Proposição 4.2. *Seja* $\|.\|$ *uma norma qualquer em* E*. Então, para todo* $x, y \in E$ *tem-se* $|\|x\| - \|y\|| \leq \|x - y\|$.

Demonstração. Note que $\|x\| = \|(x-y)+y\| \leq \|x-y\|+\|y\| \Rightarrow \|x\|-\|y\| \leq \|x-y\|$.

Analogamente,

$\|y\| = \|(y-x)+x\| \leq \|y-x\|+\|x\| \Rightarrow \|y\|-\|x\| \leq \|y-x\| \Rightarrow \|x\|-\|y\| \geq -\|x-y\|$.

Assim,

$$-\|x - y\| \leq \|x\| - \|y\| \leq \|x\| - \|y\|.$$

Portanto,

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

□

Proposição 4.3. *Todo espaço normado* $(E, \|.\|)$ *é um espaço métrico* (E, d) *quando a métrica é definida por* $d(x, y) = \|x - y\|, \forall\, x, y \in E$.

Demonstração. Para todo $x, y, z \in E$ tem-se:

M1) $d(x, y) > 0$ pela definição de norma e $d(x, y) = 0 \Leftrightarrow \|x - y\| = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y$.

M2) $d(x, y) = \|x - y\| = \| - (y - x)\| = | - 1|\, \|y - x\| = \|y - x\| = d(y, x)$.

M3) $d(x, z) = \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$.

Portanto, (E, d) é um espaço métrico. □

Lema 4.2. *Toda norma é uma função uniformemente contínua.*

Demonstração. Para todo $\epsilon > 0$, existe $\delta = \epsilon > 0$ tal que para todos $x, y \in E$ com $\|x - y\| < \delta$ tem-se, pela proposição 4.2,

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

□

Exemplo 4.14. Normas em $\mathbb{R}^n$:

[1]Também chamado espaço linear normado ou, simplesmente, espaço normado.

i. Norma euclidiana : $\quad \|x\| = \left(\sum_i^n x_i^2\right)^{1/2};$

ii. Norma da soma : $\quad \|x\|_1 = \sum_i^n |x_i|;$

iii. Norma do máximo : $\|x\|_\infty = \max\{|x_i|; 1 \leq i \leq n\}.$

O espaço $\mathbb{R}^n$ com a **norma euclidiana**, **norma da soma** ou **norma do máximo** é um espaço normado. Será provado para o caso com a norma euclidiana, os demais casos não são difíceis de provar.

Sejam $\alpha \in \mathbb{R}$ e vetores $x, y \in \mathbb{R}^n$ onde $\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$.

N1) $\|x\| \geq 0$ *e* $\|x\| = 0 \Leftrightarrow x = 0$;
De $\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$, como $x_1^2 + x_2^2 + \ldots + x_n^2 \geq 0$, tem-se $\|x\| \geq 0$. Ainda,
$\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = 0 \Leftrightarrow x_1^2 + x_2^2 + \ldots + x_n^2 = 0 \Leftrightarrow x_1 = x_2 = \ldots =$
$= x_n = 0$, ou seja, $\|x\| = 0 \Leftrightarrow x = 0$.

N2) $\|\alpha x\| = |\alpha|\|x\|$;
De fato,

$$\begin{aligned}
\|\alpha x\| &= \sqrt{(\alpha x_1)^2 + (\alpha x_2)^2 + \ldots + (\alpha x_n)^2} \\
&= \sqrt{\alpha^2 x_1^2 + \alpha^2 x_2^2 + \ldots + \alpha^2 x_n^2} \\
&= \sqrt{\alpha^2 (x_1^2 + x_2^2 + \ldots + x_n^2)} \\
&= |\alpha| \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \\
&= |\alpha|\|x\|.
\end{aligned}$$

N3) $\|x + y\| \leq \|x\| + \|y\|$.
Para demonstrar este item utiliza-se um resultado chamado *Desigualdade de Cauchy-Schwarz*. Pela desigualdade (3.1) (página 30) tem-se:

$$\left|\sum_{i=1}^n x_i y_i\right| \leq \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}$$

para quaisquer vetores $x, y \in \mathbb{R}^n$.

Assim, elevando ao quadrado o termo $\|x + y\|$, obtem-se

$$\begin{aligned}\|x+y\|^2 &= \sum_{i=1}^{n}(x_i+y_i)^2\\ &= \|x\|^2+2\sum_{i=1}^{n}x_iy_i+\|y\|^2\\ &\leq \|x\|^2+2\|x\|\|y\|+\|y\|^2\\ &= (\|x\|+\|y\|)^2\end{aligned}$$

de forma que $\|x+y\| \leq \|x\|+\|y\|$.

Portanto $(\mathbb{R}^n, \|.\|)$ é um espaço normado em que, nesta situação, $\|.\|$ é a norma euclidiana.

Definição 4.12. *Uma norma $\|.\|_0$ de um espaço vetorial E é equivalente a norma $\|.\|_1$ de E se existem constantes positivas α, β tais que*

$$\alpha\|.\|_1 \leq \|.\|_0 \leq \beta\|.\|_1. \tag{4.6}$$

Teorema 4.4. *Em um espaço vetorial normado de dimensão finita, todas as normas são equivalentes.*

Demonstração. A ideia é mostrar que qualquer norma é equivalente a norma da soma. Para tanto, seja E um espaço vetorial normado n-dimensional e considere $\{e_1, e_2, ..., e_n\}$ uma base de E. Assim para cada $x \in E$ existem únicos $\lambda_1, \lambda_2, ...\lambda_n \in \mathbb{K}$ tais que

$$x = \lambda_1e_1+\lambda_2e_2+...+\lambda_ne_n = \sum_{i=1}^{n}\lambda_ie_i.$$

Resta mostrar que qualquer norma $\|.\|_0$ em E é equivalente a norma $\|x\|_1 = \sum_{i=1}^{n}|\lambda_i|$.

De fato,

$$\|x\|_0 = \left\|\sum_{i=1}^{n}\lambda_ie_i\right\| \leq \sum_{i=1}^{n}|\lambda_i|\|e_i\| \leq \max_{1\leq i\leq n}\|e_i\|\sum_{i=1}^{n}|\lambda_i| = \beta\|x\|_1,$$

onde $\beta = \max_{1\leq i\leq n}\|x_i\|$. Portanto, $\|x\|_0 \leq \beta\|x\|_1$.

Para a outra desigualdade, suponha que não exista $\alpha > 0$ tal que $\alpha\|x\|_1 \leq \|x\|_0$, para todo $x \in E$. Então, para cada $n \in \mathbb{N}$ existe $x_n \in E$ tal que $\|x_n\|_1 > n\|x_n\|_0$. Defina $y_n = \dfrac{x_n}{\|x_n\|_1}$, assim obtem-se uma sequência (y_n) tal que $\|y_n\| = 1$. Como $S(0;1)$ é compacta existe subsequência (y_{n_j}) de (y_n) que converge para um ponto y

em $(E, \|.\|_1)$. Pela continuidade da norma, lema 4.2 (página 64), vem que $\|y\|_1 = 1$. Usando a desigualdade obtida acima, tem-se

$$\|y\|_0 = \|y - y_{n_j} + y_{n_j}\|_0 \leq \|y - y_{n_j}\|_0 + \|y_{n_j}\|_0 \leq \beta\|y - y_{n_j}\|_1 + \frac{1}{n_j}.$$

Quando $j \to \infty$ tem-se $\|y\|_0 = 0$ e, portanto, $y = 0$. Contradição, pois $\|y\|_1 = 1$. Logo, $\alpha\|x\|_1 \leq \|x\|_0$, $\alpha > 0$. $\square$

Exemplo 4.15. No espaço $\mathbb{R}^n$ as normas definidas no exemplo 4.14 (página 64) são equivalentes.

Observação 4.4. Se o espaço não for de dimensão finita pode ocorrer a existência de normas não equivalentes, como será mostrado mais adiante.

Exemplo 4.16. Seja X um conjunto não vazio e considere E a totalidade das funções $f : X \to \mathbb{R}$ limitadas, ou seja, para cada $f \in E$ existe um número $c_f > 0$ tal que

$$|f(x)| \leq c_f,\ \forall\, x \in X.$$

Note que E é um espaço vetorial munido das operações usuais soma e produto por escalar no espaço de funções, como no exemplo 4.5 (página 55).

Considere a função que associa a cada função $f \in E$ o número real

$$\sup\{|f(x)|;\ x \in X\} = \sup_{x \in X} |f|.$$

Para toda $f,\ g \in E$ e $\alpha \in \mathbb{R}$ tem-se:

N1) $\sup_{x\in X} |f| \geq\ 0$ *e* $\sup_{x\in X} |f| = 0\ \Leftrightarrow f = 0$,

N2) $\sup_{x\in X} |\alpha f| = |\alpha| \sup_{x\in X} |f|$,

N3) $\sup_{x\in X} |f + g| \leq \sup_{x\in X} |f| + \sup_{x\in X} |g|$ (*desigualdade triangular*).

A demonstração do item $N3$) é análoga a do exemplo 3.5 (página 31).

Exemplo 4.17. Seja $I = [a, b] \subset \mathbb{R}$. No espaço $\mathcal{C}(I, \mathbb{R})$ de funções contínuas pode-se definir outras normas da seguinte forma: a cada função $f \in \mathcal{C}(I, \mathbb{R})$, associa ou ao número $\|f\|_1 = \displaystyle\int_a^b |f|$, que é igual a área da figura 4.1, ou ao número $\|f\|_2 = \left(\displaystyle\int_a^b f^2\right)^{\frac{1}{2}}$, $\forall\, f \in \mathcal{C}(I, \mathbb{R})$.

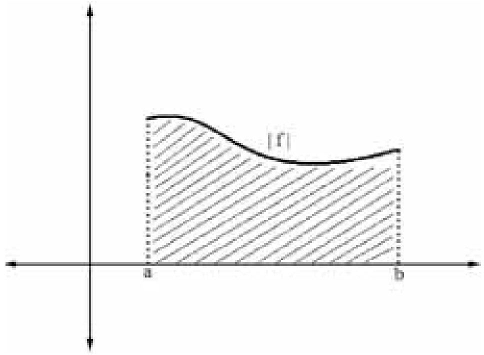

Figura 4.1: *Área da função abaixo da curva*

Observação 4.5. Para definir um importante exemplo de espaço normado, a saber o espaço $\mathcal{L}^p(\Omega)$, é preciso introduzir algumas notações, definições e resultados referentes a medida exterior, funções mensuráveis e integral de Lebesgue de uma função. Estes resultados estão descritos no apêndice A.

As proposições seguintes serão utilizadas na construção do espaço de funções reais p-integráveis no sentido *Lebesgue.*

Proposição 4.4. ***Desigualdade de Hölder****: Se f e g são funções mensuráveis e $|f|^p$ e $|g|^q$ são integráveis com $1 < p, q < \infty$ e $\dfrac{1}{p} + \dfrac{1}{q} = 1$ (diz-se que p e q são expoentes conjugados), então a função fg é integrável e, para cada conjunto mensurável $\Omega \subset \mathbb{R}$, tem-se:*

$$\int_\Omega |fg| \leq \left(\int_\Omega |f|^p\right)^{\frac{1}{p}} \left(\int_\Omega |g|^q\right)^{\frac{1}{q}}. \tag{4.7}$$

Demonstração. Se $f = 0$ *q.t.p.* e $g = 0$ *q.t.p.* então (4.7) está satisfeita. Suponha que $f \neq 0$ e $g \neq 0$, assim a *Desigualdade Hölder*, como na proposição 3.1 do capítulo anterior, segue da desigualdade

$$\rho^\alpha \nu^\beta \leq \alpha\rho + \beta\nu, \tag{4.8}$$

onde $\alpha > 0$, $\beta > 0$, $\alpha + \beta = 1$, $\rho \geq 0$ e $\nu \geq 0$. Para cada $x \in \Omega$ considere

$$\rho = \frac{|f(x)|^p}{\displaystyle\int_\Omega |f|^p},\ \nu = \frac{|g(x)|^q}{\displaystyle\int_\Omega |g|^q},\ \alpha = \frac{1}{p} \text{ e } \beta = \frac{1}{q}$$

Substituindo na desigualdade (4.8) segue que

$$\frac{|f(x)g(x)|}{\left(\displaystyle\int_\Omega |f|^p\right)^{\frac{1}{p}} \left(\displaystyle\int_\Omega |g|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p}\frac{|f(x)|^p}{\displaystyle\int_\Omega |f|^p} + \frac{1}{q}\frac{|g(x)|^q}{\displaystyle\int_\Omega |g|^q}.$$

Integrando sobre Ω, obtem-se

$$\frac{\displaystyle\int_\Omega |f(x)g(x)|}{\left(\displaystyle\int_\Omega |f|^p\right)^{\frac{1}{p}}\left(\displaystyle\int_\Omega |g|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p}\frac{\displaystyle\int_\Omega |f(x)|^p}{\displaystyle\int_\Omega |f|^p} + \frac{1}{q}\frac{\displaystyle\int_\Omega |g(x)|^q}{\displaystyle\int_\Omega |g|^q}$$

ou seja,

$$\frac{\displaystyle\int_\Omega |f(x)g(x)|}{\left(\displaystyle\int_\Omega |f|^p\right)^{\frac{1}{p}}\left(\displaystyle\int_\Omega |g|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

Portanto,

$$\int_\Omega |fg| \leq \left(\int_\Omega |f|^p\right)^{\frac{1}{p}}\left(\int_\Omega |g|^q\right)^{\frac{1}{q}}.$$

□

Proposição 4.5. ***Desilgualdade de Minkowski**: Se f e g são funções mensuráveis tais que para cada $1 \leq p < \infty$, as funções $|f|^p$ e $|g|^p$ são integráveis, então $|f+g|^p$ é integrável e, para cada conjunto mensurável $\Omega \subset \mathbb{R}$, verifica-se a seguinte propriedade:*

$$\left(\int_\Omega |f+g|^p\right)^{\frac{1}{p}} \leq \left(\int_\Omega |f|^p\right)^{\frac{1}{p}} + \left(\int_\Omega |g|^p\right)^{\frac{1}{p}}. \tag{4.9}$$

Demonstração. Para $p = 1$, imediato. Suponha $1 < p < \infty$ então, pela proposição A.4 do capítulo anterior, segue

$$\begin{aligned} \int_\Omega |f+g|^p = & \int_\Omega |f+g||f+g|^{p-1} \\ \leq & \int_\Omega (|f|+|g|)|f+g|^{p-1} \\ \leq & \int_\Omega |f||f+g|^{p-1} + \int_\Omega |g||f+g|^{p-1} \end{aligned}$$

Aplicando a *Desigualdade de Hölder* a cada uma das integrais a direita da desigualdade acima e como $\dfrac{1}{p} + \dfrac{1}{q} = 1$, obtem-se

$$\begin{aligned}\int_\Omega |f+g|^p \leq & \left(\int_\Omega |f|^p\right)^{\frac{1}{p}}\left(\int_\Omega |f+g|^{(p-1)q}\right)^{\frac{1}{q}}\\ + & \left(\int_\Omega |g|^p\right)^{\frac{1}{p}}\left(\int_\Omega |f+g|^{(p-1)q}\right)^{\frac{1}{q}}\\ = & \left(\int_\Omega |f+g|^p\right)^{\frac{1}{q}}\left[\left(\int_\Omega |f|^p\right)^{\frac{1}{p}}+\left(\int_\Omega |g|^p\right)^{\frac{1}{p}}\right].\end{aligned}$$

Suponha $\left(\int_\Omega |f+g|^p\right)^{\frac{1}{q}} \neq 0$ e como $\dfrac{1}{p} = 1 - \dfrac{1}{q}$ segue

$$\left(\int_\Omega |f+g|^p\right)^{\frac{1}{p}} \leq \left(\int_\Omega |f|^p\right)^{\frac{1}{p}} + \left(\int_\Omega |g|^p\right)^{\frac{1}{p}}.$$

□

Exemplo 4.18. Espaço $\mathcal{L}^p(\Omega)$.

Seja $1 \leq p < \infty$ e $\Omega \subset \mathbb{R}$ um conjunto mensurável de medida finita. Define-se $\mathcal{L}^p(\Omega)$ como o espaço das funções reais p-integráveis no sentido *Lesbesgue*, ou seja,

$$\mathcal{L}^p(\Omega) : \left\{ f : \Omega \to \mathbb{R}\,\text{mensuráveis};\ \int_\Omega |f|^p < \infty \right\}$$

com a norma

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

Note que $\mathcal{L}^p(\Omega)$ é um espaço vetorial com as operações usuais, de soma e multiplicação por escalar, de funções. Tais operações estão bem definidas. Com efeito, se $f,\, g \in \mathcal{L}^p(\Omega)$ então $\int_\Omega |f|^p < \infty$ e $\int_\Omega |g|^p < \infty$. Por desigualdade triangular segue que para cada $x \in \Omega$

$$|f(x)+g(x)| \leq |f(x)| + |g(x)| \leq 2\ \max\{|f(x)|, |g(x)|\},$$

assim

$$|f(x)+g(x)| \leq 2\max\{|f(x)|, |g(x)|\}.$$

Suponha que o $\max\{|f(x)|, |g(x)|\} = |f(x)|$ (se o $\max\{|f(x)|, |g(x)|\} = |g(x)|$, análogo) e como $|g(x)| > 0$ segue

$$|f(x)+g(x)| \leq 2|f(x)| \leq 2(|f(x)| + |g(x)|),$$

logo

$$|f(x)+g(x)|^p \leq 2^p(|f(x)|^p + |g(x)|^p).$$

Integrando sobre Ω em ambos os lados da desigualdade acima, obtem-se

$$\int_\Omega |f+g|^p \leq 2^p \left(\int_\Omega |f|^p + \int_\Omega |g|^p \right) < \infty.$$

Portanto, $f+g \in \mathcal{L}^p(\Omega)$. Agora, $\alpha f \in \mathcal{L}^p(\Omega)$, pois para todo $x \in \Omega$

$$|\alpha f(x)| = |\alpha||f(x)|,$$

logo

$$|\alpha f(x)|^p = |\alpha|^p |f(x)|^p,$$

ou seja,

$$\int_\Omega |\alpha f|^p = |\alpha|^p \int_\Omega |f|^p < \infty.$$

É fácil ver que $\|.\|_p$ é uma norma pois, $\|f\|_p = 0 \Leftrightarrow f = 0$ $q.t.p.$. Com efeito, se $f = 0$ $q.t.p.$, então $\|f\|_p = 0$. Por outro lado, se $f \neq 0$ $q.t.p.$ então existe $\Omega' \subset \Omega$, com $m(\Omega') < \infty$ tal que $|f|^p > 0$. Logo,

$$\int_\Omega |f|^p \geq \int_{\Omega'} |f|^p > 0.$$

Portanto, $\left(\int_\Omega |f|^p \right)^{\frac{1}{p}} > 0$ para $f \neq 0$ $q.t.p.$. Além disso, $\|\alpha f\|_p = |\alpha| \|f\|_p$, pois

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

A desigualdade triangular segue da Desigualdade de Minkowski, proposição 4.5, ou seja,

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Portanto, $\mathcal{L}^p(\Omega)$ é um espaço normado.

Exemplo 4.19. O espaço ℓ^p, definido no exemplo 3.8 (página 37), com a norma $\|x\|_p = \left(\sum_{i=1}^{\infty} |\eta_i|^p \right)^{\frac{1}{p}}$ onde $x = (\eta_i)_{i \in \mathbb{N}} \in \ell^p$, é um espaço normado. As condições $N1$) e $N2$) da definição de norma são facilmente verificadas e a condição $N3$) resulta da Desigualdade de Minkowski, proposição 3.2 (página 36).

Exemplo 4.20. O espaço de sequência ℓ^∞, definido no exemplo 3.6 (página 33), com a norma definida por $\|x\| = \sup_{i \in N} |\eta_i|$ onde $x = (\eta_i)_{i \in \mathbb{N}} \in \ell^\infty$, é um espaço normado. A demonstração é semelhandte a do exemplo 3.5 (página 31), do capítulo anterior.

4.3 Espaços de Banach

Um espaço normado que é completo com a métrica induzida pela norma é chamado de espaço de Banach. Todo espaço normado de dimensão finita é Banach (será demonstrado mais adiante), porém se o espaço for de dimensão infinita é preciso mostrar que toda sequência de Cauchy converge para um ponto que pertence ao espaço.

Definição 4.13. *Seja E um espaço normado. Diz-se que E é um* ***espaço de Banach*** *se ele é completo, ou seja, se toda sequência de Cauchy em E é convergente em E.*

Exemplo 4.21. O espaço $\mathbb{R}^n$ com a norma definida por

$$\|x\| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}$$

onde $x = (x_1, x_2, ..., x_n)$, é um espaço de Banach.

De fato, seja $(x_p)_{p\in\mathbb{N}}$ uma sequência de Cauchy em $\mathbb{R}^n$, com $x_p = (\xi_1^{(p)}, \xi_2^{(p)}, ..., \xi_n^{(p)}) = (\xi_j^{(p)})$, $1 \leq j \leq n$. Como (x_p) é de Cauchy, para todo $\epsilon > 0$ existe n_0 tal que $p, q > n_0$ tem-se

$$\|x_p - x_q\| = \left(\sum_{j=1}^{n} |\xi_j^{(p)} - \xi_j^{(q)}|^2\right)^{\frac{1}{2}} < \epsilon. \tag{4.10}$$

Agora, para cada $1 \leq j \leq n$ segue que

$$|\xi_j^{(p)} - \xi_j^{(q)}| \leq \|x_p - x_q\|.$$

Assim, dado $\epsilon > 0$ existe n_0 tal que $p, q > n_0$ então

$$|\xi_j^{(p)} - \xi_j^{(q)}| < \epsilon.$$

Isto mostra que para cada j fixo, $1 \leq j \leq n$, a sequência $(\xi_j^{(1)}, \xi_j^{(2)}, \xi_j^{(3)}, ...) = (\xi_j^{(p)})$ é uma sequência de Cauchy de números reais, logo convergente pois $\mathbb{R}$ é completo, ou seja, $\xi_j^{(p)} \to \xi_j$ quando $p \to \infty$. Usando estes n limites considere $x = (\xi_1, \xi_2, ..., \xi_n)$ (note que $x \in \mathbb{R}^n$) e fazendo $q \to \infty$ na equação (4.10), tem-se

$$\|x_p - x\| \leq \epsilon.$$

Portanto, $x_p \to x$ em $\mathbb{R}^n$, ou seja, x é limite de $(x_p)_{p\in\mathbb{N}}$.

O espaço $\mathbb{C}^n$ também é de Banach, a demosntração é análoga ao do $\mathbb{R}^n$.

Lema 4.3. *Seja $\{x_1, x_2, ..., x_n\}$ um conjunto de vetores linearmente independentes em um espaço normado E. Então existe um número $c > 0$ tal que para quaisquer escalares $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{K}$, tem-se*

$$\|\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n\| \geq c(|\alpha_1| + |\alpha_2| + ... + |\alpha_n|). \tag{4.11}$$

Demonstração. Seja $s = |\alpha_1| + \ldots + |\alpha_n| = \sum_{j=1}^{n} |\alpha_j|$. Note que se $s = 0$, então todos os α_j são nulos e (4.11) é satisfeita para qualquer c. Suponha agora $s > 0$. Dividindo (4.11) por s segue que

$$\|\beta_1 x_1 + \ldots + \beta_n x_n\| \geq c \tag{4.12}$$

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

Então, provar (4.11) equivale a provar que existe $c > 0$ tal que (4.12) seja satisfeita para quaisquer escalares $\beta_1, \ldots, \beta_n \in \mathbb{K}$ com $\sum_{j=1}^{n} |\beta_j| = 1$.

Suponha que (4.12) seja falsa. Ou seja, para todo $c > 0$ existem escalares $\beta_1^c, \beta_2^c, \ldots, \beta_n^c$ com $\sum_{j=1}^{n} \beta_j^{(c)} = 1$ tais que $\|y_c\| < c$ onde $y_c = \beta_1^c x_1 + \beta_2^c x_2 + \ldots + \beta_n^c x_n$. Conforme c varia obtem-se uma sequência $(y_m)_{m \in \mathbb{N}}$ $(c = \dfrac{1}{m}, m \in \mathbb{N})$ satisfazendo

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

ou seja,

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

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 $|\beta_j^{(m)}| \leq 1$. Assim, para cada j fixo a sequência $(\beta_j^{(m)}) = (\beta_j^{(1)}, \beta_j^{(2)}, \ldots)$ é limitada. Pelo teorema 2.4, de Bolzano-Weierstrass, cada uma das sequências

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

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

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

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

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

ou seja,

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

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 $\{x_1, x_2, ..., x_n\}$ é um conjunto linearmente independente segue que $y \neq 0$.

Por outro lado, pela continuidade da norma (lema 4.2), 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 subsequência de (y_m) segue que $y_n^{(m)} \to 0$ quando $m \to \infty$. Agora, pela definição 4.11 de norma, $\|y\| = 0 \Leftrightarrow y = 0$. Contradição, pois $y \neq 0$. □

Como aplicação deste lema segue o resultado abaixo.

Teorema 4.5. *Todo espaço normado de dimensão finita é Banach.*

Demonstração. Considere $(x_m)_{m \in \mathbb{N}}$ uma sequência de Cauchy qualquer em um espaço normado de dimensão finita E. Basta mostrar que $(x_m)_{m \in \mathbb{N}}$ converge para um elemento $x \in E$, ou seja, $x_m \to x$ em E.

Sejam $\dim E = n$ e $\{e_1, e_2, ..., e_n\}$ uma base para E. Então, para cada $x_m \in E$ existem únicos $\alpha_1^{(n)}$,...,$\alpha_n^{(m)} \in \mathbb{K}$ tais que

$$x_m = \alpha_1^{(m)} e_1 + \alpha_2^{(m)} e_2 + ... + \alpha_n^{(m)} e_n.$$

Como $(x_m)_{m \in \mathbb{N}}$ é uma sequência de Cauchy então para todo $\epsilon > 0$ existe n_0 tal que $m, q > n_0$ tem-se $\|x_m - x_q\| < \epsilon$, ou ainda,

$$\left\| \sum_{j=1}^{n} (\alpha_j^{(m)} - \alpha_j^{(q)}) e_j \right\| = \|x_m - x_n\| < \epsilon.$$

Como dim $E = n$, pelo lema 4.3, existe uma constante $c > 0$ tal que

$$c\sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j^{(q)}| \leq \left\|\sum_{j=1}^{n}(\alpha_j^{(m)} - \alpha_j^{(q)})e_j\right\| = \|x_m - x_q\| < \epsilon,$$

onde $m, q > n_0$. Dividindo a desigualdade acima por $c > 0$ obtem-se

$$|\alpha_j^{(m)} - \alpha_j^{(q)}| \leq \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j^{(q)}| < \frac{\epsilon}{c}.$$

Isto mostra que cada uma das n sequências

$$(\alpha_j^{(m)}) = (\alpha_j^{(1)},\, \alpha_j^{(2)},\, \alpha_j^{(3)}, ...),\ 1 \leq j \leq n,$$

é de Cauchy em $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$), por isso convergem. Seja α_j o limite de $(\alpha_j^{(m)})$, $1 \leq j \leq n$. Com estes n limites $\alpha_1, \alpha_2, ..., \alpha_n$ defina

$$x = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n.$$

Note que $x \in E$, além disso,

$$\|x_m - x\| = \left\|\sum_{j=1}^{n}(\alpha_j^{(m)} - \alpha_j)e_j\right\| \leq \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j|\|e_j\|.$$

À direita, $\alpha_j^{(m)} \to \alpha_j$. Por isso $\|x_m - x\| \to 0$, ou seja, $x_m \to x$. Portanto, x_m é convergente em E, logo E é Banach. $\square$

Os exemplos seguintes são espaços de Banach de dimensão infinita.

Exemplo 4.22. O espaço normado ℓ^p com a norma $\|x\|_p = \left(\sum_{j=1}^{\infty} |\xi_j|^p\right)^{\frac{1}{p}}$ é Banach, em que p é um número fixo e $1 \leq p < \infty$.

Com efeito, seja $(x_m)_{m\in\mathbb{N}}$ qualquer sequência de Cauchy no espaço ℓ^p, onde $x_m = (\xi_j^{(m)})_{j\in\mathbb{N}}$. Então para todo $\epsilon > 0$ existe n_0 tal que $m, n > n_0$

$$\|x_m - x_n\|_p = \left(\sum_{j=1}^{\infty} |\xi_j^{(m)} - \xi_j^{(n)}|^p\right)^{\frac{1}{p}} < \epsilon. \tag{4.13}$$

Para cada $j \in \mathbb{N}$, $|\xi_j^{(m)} - \xi_j^{(n)}| \leq \|x_m - x_n\|_p$ e assim, para todo $\epsilon > 0$ existe n_0 (como acima) tal que $m, n > n_0$

$$|\xi_j^{(m)} - \xi_j^{(n)}| < \epsilon.$$

Assim, para cada j fixo, a sequência $(\xi_j^{(1)}, \xi_j^{(2)}, \xi_j^{(3)}, ...)$ é uma sequência de Cauchy em $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$), logo cada sequência $\xi_j^{(m)}$ desta nova sequência converge para um elemento ξ_j quando $m \to \infty$. Usando este fato, considere $x = (\xi_1, \xi_2, \xi_3, ...)$. Resta mostrar que $x \in \ell^p$ e $x_m \to x$.

Da equação (4.13) tem-se para todo $m, n > n_0$,

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

com $k \in \mathbb{N}$. Fazendo $n \to \infty$, para $m > n_0$ segue que

$$\sum_{j=1}^{k} |\xi_j^{(m)} - \xi_j|^p \leq \epsilon^p, \ k \in \mathbb{N}.$$

Fazendo $k \to \infty$, tem-se para todo $m > n_0$

$$\sum_{j=1}^{\infty} |\xi_j^{(m)} - \xi_j|^p \leq \epsilon^p.$$

Isto mostra que $x_m - x = (\xi_j^{(m)} - \xi_j) \in \ell^p$. Segue da Desigualdade de Minkowski, proposição 3.2, que

$$\|x\|_p = \|x - x_m + x_m\|_p \leq \|x - x_m\|_p + \|x\|_p < \infty.$$

Logo $x \in \ell^p$ e $x_m \to x$.

Exemplo 4.23. O espaço normado ℓ^∞ com a norma $\|x\| = \sup_{j \in \mathbb{N}} |x_j|$ é Banach.

De fato, seja $(x_m)_{m \in \mathbb{N}}$ qualquer sequência de Cauchy no espaço ℓ^∞. Assim, para todo $\epsilon > 0$ existe n_0 tal que $m, n > n_0$ tem-se

$$\|x_m - x_n\| = \sup_{j \in \mathbb{N}} |\xi_j^{(m)} - \xi_j^{(m)}| < \epsilon.$$

Agora, para qualquer j fixo e $m, n > n_0$

$$|\xi_j^{(m)} - \xi_j^{(n)}| < \epsilon. \tag{4.14}$$

Assim, para todo j fixo, a sequência $(\xi_j^{(1)}, \xi_j^{(2)}, \xi_j^{(3)}, ...)$ é uma sequência de Cauchy em $\mathbb{K}$, logo convergente, ou seja, para cada j fixo $\xi_j^{(m)} \to \xi_j$ com $m \to \infty$. Usando este fato para todo $j \in \mathbb{N}$, considere a sequência $x = (\xi_1, \xi_2, \xi_3, ...)$. Resta mostrar que $x \in \ell^\infty$ e que $x_m \to x$.

Fazendo $n \to \infty$ em (4.14), segue que

$$|\xi_j^{(m)} - \xi_j| \leq \epsilon. \tag{4.15}$$

para cada j fixo e $m > n_0$. Como $x_m = (\xi_j^{(m)}) \in \ell^\infty$, existe um número real k_m tal que $|\xi_j^{(m)}| \leq k_m$ para todo j. Com isso, para todo j e $m > n_0$ segue que

$$|\xi_j| = |\xi_j - \xi_j^{(m)} + \xi_j^{(m)}| \leq |\xi_j - \xi_j^{(m)}| + |\xi_j^{(m)}| \leq \epsilon + k_m.$$

Note que o lado direito desta desigualdade não depende de j, assim a sequência de números $x = (\xi_j)_{j\in\mathbb{N}}$ é limitada, ou seja, $x \in \ell^\infty$. Ainda, de (4.15) obtem-se

$$\|x_m - x_n\| = \sup_{j\in\mathbb{N}} |\xi_j^{(m)} - \xi_j| \leq \epsilon$$

com $m > n_0$. Isto mostra que $x_m \to x \in \ell^\infty$.

Exemplo 4.24. O espaço das funções contínuas $\mathcal{C}([a,b],\mathbb{R})$ com a norma definida por $\|f\| = \sup\limits_{x\in[a,b]} |f(x)|$, é Banach.

Com efeito, seja $(f_m)_{m\in\mathbb{N}}$ qualquer sequência de Cauchy em $\mathcal{C}([a,b],\mathbb{R})$. Então, dado $\epsilon > 0$ existe n_0, independente do ponto $x \in [a,b]$, tal que $m,n > n_0$ tem-se

$$\|f_m - f_n\| = \sup_{x\in[a,b]} |f_m(x) - f_n(x)| < \epsilon. \tag{4.16}$$

Para cada $x = x_0 \in [a,b]$ fixo e de (4.16), segue

$$|f_m(x_0) - f_n(x_0)| \leq \|f_m - f_n\| < \epsilon.$$

com $m,n > n_0$.

Isto mostra que a sequência $(f_1(x_0), f_2(x_0), f_3(x_0), ...) = (f_m(x_0))$ é uma sequência de números reais, logo convergente. Assim, a sequência $(f_m(x_0))$ converge para o número real $f(x_0)$ quando $m \to \infty$, ou seja,

$$\begin{array}{rcl} f : [a,b] & \to & \mathbb{R} \\ x & \mapsto & f(x) := \lim\limits_{n\to\infty} f_n(x). \end{array}$$

Desta forma é possível associar a cada $x \in [a,b]$ um único número real $f(x)$. Resta mostrar que $f \in \mathcal{C}([a,b],\mathbb{R})$ e que $f_m \to f$.

Fazendo $n \to \infty$ (4.16) e mantendo m fixo, segue que

$$\sup_{x\in[a,b]} |f_m(x) - f(x)| \leq \epsilon, \, \forall\, x \in [a,b].$$

Assim, para cada $x \in [a,b]$, com $m > n_0$ tem-se

$$|f_m(x) - f(x)| \leq \epsilon.$$

Sendo f_m contínua em $[a,b]$, pelo teorema 2.7 (página 27), (f_m) converge uniformemente para f em $[a,b]$. Agora, como f_m é contínua em $[a,b]$ e a convergência é uniforme, pelo teorema 2.6 (página 26) a função limite f é contínua em $[a,b]$, logo $f \in \mathcal{C}([a,b],\mathbb{R})$ e também, $f_m \to f$.

Dependendo a norma considerada em um espaço normado E, este pode não ser de Banach, como mostra o exemplo abaixo.

Exemplo 4.25. Considere o espaço de todas as funções reais e contínuas $\mathcal{C}([0,1],\mathbb{R})$ com a norma $\|f\| = \displaystyle\int_0^1 |f(x)|\,dx$. Com esta norma, o espaço $\mathcal{C}([0,1],\mathbb{R})$ **não é de Banach**.

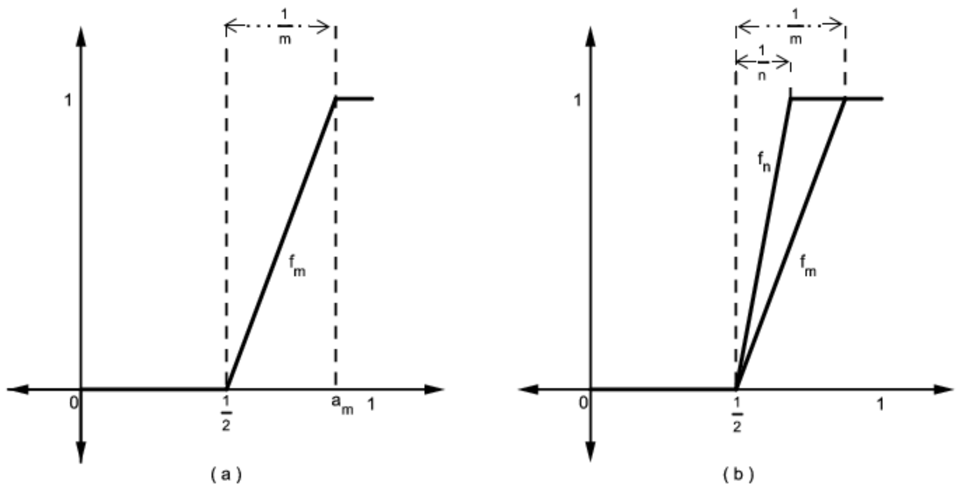

Figura 4.2: *Sequência* $(f_m)_{m\in\mathbb{N}}$

Com efeito, a função f_m, na figura 4.2-(a), é dada por

$$f_m(x) = \begin{cases} 0, & \text{se } 0 \le x \le \dfrac{1}{2}, \\ m\left(x - \frac{1}{2}\right), & \text{se } \dfrac{1}{2} < x \le a_m, \\ 1, & \text{se } a_m < x \le 1. \end{cases}$$

onde $a_m = \dfrac{1}{2} + \dfrac{1}{m}$. Note que quando m varia, f_m forma uma sequência de Cauchy que não é convergente em $\mathcal{C}([0,1],\mathbb{R})$.

A sequência $(f_m)_{m\in\mathbb{N}}$ é de Cauchy, pois

$$\|f_m - f_n\| = \int_0^1 |f_m(x) - f_n(x)|\,dx = \frac{1}{2}\left(\frac{1}{m} - \frac{1}{n}\right) \longrightarrow 0,$$

quando $m, n \to \infty$.

Observe que a integral acima representa a área do triângulo da figura 4.2-(b).

Ainda, $f_m \to f$, onde f é dada por

$$f(x) = \begin{cases} 0, & \text{se } 0 \le x \le \dfrac{1}{2}, \\ 1, & \text{se } \dfrac{1}{2} < x \le 1, \end{cases}$$

pois,

$$\|f_m - f\| = \int_0^1 |f_m(x) - f(x)|\, dx = \frac{1}{2m} \longrightarrow 0, \text{ quando } m \to \infty,$$

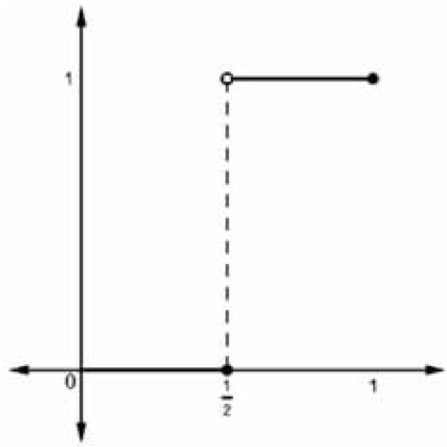

Figura 4.3: *Gráfico da função f*

Como o limite é único e $f \notin \mathcal{C}([0,1],\mathbb{R})$ então $(f_m)_{m\in\mathbb{N}}$ é uma sequência de Cauchy em $\mathcal{C}([0,1],\mathbb{R})$ que não converge em $\mathcal{C}([0,1],\mathbb{R})$.

Exemplo 4.26. O espaço normado $\mathcal{L}^p(\Omega)$, onde $1 \leq p < \infty$ e $\Omega \subset \mathbb{R}$ um conjunto mensurável de medida finita, com a norma $\|f\|_p = \left(\int_\Omega |f|^p\right)^{\frac{1}{p}}$ é Banach.

Seja $(f_n)_{n\in\mathbb{N}}$ uma sequência de Cauchy em $\mathcal{L}^p(\Omega)$, então para todo $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $m, n > n_0$ tem-se $\|f_n - f_m\|_p < \epsilon$.

Pelo teorema 3.4 (página 49), basta mostrar que existe uma subsequência convergente para um ponto de $\mathcal{L}^p(\Omega)$. Seja $k \in \mathbb{N}$. Dado $\epsilon = \frac{1}{2^k}$, considere f_{n_k} tal que

$$\|f_{n_{k+1}} - f_{n_k}\|_p \leq \frac{1}{2^k}, \forall\, k \geq 1.$$

Agora, seja

$$g_n(x) = \sum_{k=1}^{n} |f_{n_{k+1}}(x) - f_{n_k}(x)|, \forall\, n \in \mathbb{N}.$$

Note que

$$g_n(x) \leq \sum_{k=1}^{n} \frac{1}{2^k} \overset{n\to\infty}{\longrightarrow} \sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{1-\frac{1}{2}} - 1 = 1,$$

então $g_n(x)$ converge, pelo teorema 2.8 (página 28). Assim, existe g tal que $g_n \to g$ *q.t.p.*, ou seja, dado $\epsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que $k > k_0$, $\|g_k - g\|_p < \epsilon$.

Agora,

i. g_n é monótona crescente, não-negativa e mensurável.

ii. $g_n \to g$ *q.t.p.* então, pelo teorema A.1 (página 204), $\displaystyle\int |g|^p = \lim_{n\to\infty} \int |g_n|^p < \infty$.

Logo, $\displaystyle\int_\Omega |g|^p < \infty$ e $g(x) < \infty$ *q.t.p.* em Ω, portanto, $g \in \mathcal{L}^p(\Omega)$.

Agora, utilizando g_n e seu limite g segue que a sequência (f_{n_k}) é de Cauchy. Com efeito, sejam k, l tais que $k > l = \max\{k_0, 2\}$ então

$$\begin{aligned}
|f_{n_k}(x) - f_{n_l}(x)| &= |f_{n_k}(x) - f_{n_{k-1}}(x) + f_{n_{k-1}}(x) - f_{n_{k-2}} + \ldots + f_{n_{l+1}}(x) - f_{n_l}(x)| \\
&\leq |f_{n_k}(x) - f_{n_{k-1}}(x)| + |f_{n_{k-1}}(x) - f_{n_{k-2}}(x)| + \ldots + |f_{n_{l+1}}(x) - f_{n_l}(x)| \\
&= g_{k-1}(x) - g_{l-1}(x) \\
&\leq g(x) - g_{l-1}(x) < \epsilon
\end{aligned}$$

pois $g_{l-1} \to g$. Como $k > l$ segue que $|f_{n_k}(x) - f_{n_l}(x)| < \epsilon$ sempre que $k, l > k_0$.

Defina para cada $x \in \Omega$, $f(x) = \lim_{k\to\infty} f_{n_k}(x)$. Note que $f(x)$ está bem definida, uma vez que $(f_{n_k}(x))$ é uma sequência de Cauchy em $\mathbb{R}$, portanto convergente. Resta mostrar que $f \in \mathcal{L}^p(\Omega)$. Fazendo $k \to \infty$ na desigualdade

$$|f_{n_k}(x) - f_{n_l}(x)| \leq g(x) - g_{l-1}(x),$$

obtem-se

$$|f(x) - f_{n_l}(x)| \leq g(x) - g_{l-1}(x) \in \mathcal{L}^p(\Omega) \Rightarrow |f(x) - f_{n_l}(x)| \in \mathcal{L}^p(\Omega),$$

ou seja, $f(x) - f_{n_l}(x) \in \mathcal{L}^p(\Omega)$ e como $\mathcal{L}^p(\Omega)$ é um espaço vetorial segue que $f(x) \in \mathcal{L}^p(\Omega)$ pois $f(x) = h(x) + f_{n_l}(x)$. Portanto, $\mathcal{L}^p(\Omega)$ é Banach.

Observação 4.6. Pela proposição 3.11 (página 47) em espaços métricos de dimensão finita todo conjunto fechado e limitado é compacto. Para espaços normados de dimensão infinita esta afirmação não é, em geral, verdadeira. Considere o conjunto $S \subset \ell^1$ definido por

$$S = \{(1, 0, 0, \ldots, 0, \ldots), (0, 1, 0, \ldots, 0, \ldots), \ldots\}.$$

Cada elemento $x = (\xi_1, \xi_2, \ldots) \in S$, tem a propriedade $\displaystyle\sum_{j=1}^{\infty} |\xi_j| < \infty$. Note que S é fechado e limitado e, para todo $x \in S$, $\|x\| = 1$, logo $x \in B[0, 1]$ mas S não é compacto.

Teorema 4.6. *Um subespaço de um espaço de Banach é um espaço de Banach se, e somente se, é fechado.*

Demonstração. ($\Rightarrow$) Sejam E um espaço de Banach e $F \subset E$ subespaço de E. Por hipótese F é Banach, logo toda sequência (x_m) de Cauchy em F é convergente em F. Portanto, F é fechado.

($\Leftarrow$) Seja $F \subset E$ um subespaço fechado de um espaço de Banach E e considere (x_m) uma sequência de Cauchy em F. Como $F \subset E$ e a norma em F é a restrição da norma de E segue que (x_m) é uma sequência de Cauchy em E. Como E é de Banach, então (x_m) converge para um elemento $x \in E$ e, por F ser fechado, segue que $x \in F$. Portanto, $x_m \to x$ em F, logo F é de Banach. $\square$

Teorema 4.7. *Todo subespaço de dimensão finita de um espaço normado E é fechado em E.*

A demonstração segue diretamente do teorema 4.5 (página 74) e do teorema 4.6 (página 80).

Lema 4.4. ***(Lema de Riesz)*** *Seja E um espaço normado de qualquer dimensão. Sejam Y e Z subespaços de E e suponha que Y é fechado e é um subespaço próprio de Z. Então, para todo número real θ no intervalo $(0,1)$ existe um $z \in Z$ tal que para todo $y \in Y$ tem-se*

$$\|z - y\| \geq \theta,$$

onde $\|z\| = 1$.

Demonstração. Seja $z_0 \in Z \setminus Y$ e seja $a = d(z_0, Y) = \inf\limits_{y \in Y} \|z_0 - y\|$ a sua distância de Y, como na figura 4.4.

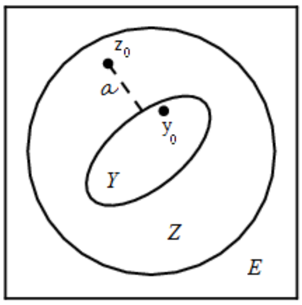

Figura 4.4: *Ideia geométrica para a demonstração do lema*

Note que $a > 0$, já que Y é fechado. Seja $0 < \theta < 1$, pela definição de ínfimo, para todo $\epsilon > 0$, existe $y_0 \in Y$ tal que $d(z_0, y_0) < a + \epsilon$ o que implica

$$a \leq \|z_0 - y_0\| < a + \epsilon.$$

Tomando $\epsilon = a\left(\dfrac{1}{\theta} - 1\right)$, tem-se

$$a \leq \|z_0 - y_0\| < \frac{a}{\theta}. \tag{4.17}$$

Considere $z = c(z_0 - y_0)$, onde $c = \dfrac{1}{\|z_0 - y_0\|}$, claramente $\|z\| = 1$. Resta mostrar que para todo $y \in Y$, $\|z - y\| \geq \theta$. De fato,

$$\begin{aligned} \|z-y\| &= \|c(z_0-y_0)-y\| \\ &= c\left\|z_0-y_0-\frac{1}{c}y\right\| \qquad (4.18) \\ &= c\|z_0-y_1\| \end{aligned}$$

onde $y_1 = y_0 + \dfrac{1}{c}y$. Note que $y_1 \in Y$, uma vez que é uma combinação linear de elementos de Y.

Agora, $\|z_0 - y_1\| \geq a$, pela forma como a foi definido. Como $c > 0$ segue que $c\|z_0 - y_1\| \geq c\,a$. Por (4.18) e (4.17) segue que

$$\|z-y\| = c\|z_0-y_1\| \geq c\,a = \frac{1}{\|z_0-y_0\|}a \geq \frac{a}{a/\theta} = \theta.$$

Portanto, $\|z-y\| \geq \theta$. □

Como aplicação do Lema de Riesz, tem-se o seguinte resultado.

Teorema 4.8. *Se um espaço normado E tem a propriedade de que a bola fechada unitária $B = \{x;\ \|x\| \leq 1\}$ é compacta, então E tem dimensão finita.*

Demonstração. Seja B um conjunto compacto e suponha que $\dim E = \infty$. Basta provar que esta afirmação conduz a uma contradição. Considere qualquer vetor $x_1 \in E$ tal que $\|x_1\| = 1$. Este x_1 gera um subespaço próprio F_1 de E de dimensão um e, pelo teorema 4.7, é fechado. Pelo lema 4.4, existe $x_2 \in E - F_1$, com $\|x_2\| = 1$, tal que

$$\|x_2 - x_1\| \geq \theta = \frac{1}{2}.$$

Os vetores x_1 e x_2 geram um subespaço próprio F_2 de E de dimensão dois e, pelo teorema 4.7, é fechado. Novamente, pelo lema 4.4, existe $x_3 \in E - F_2$, com $\|x_3\| = 1$, tal que para todo $x \in F_2$ tem-se

$$\|x_3 - x\| \geq \frac{1}{2}.$$

Em particular,

$$\|x_3 - x_1\| \geq \frac{1}{2},$$

$$\|x_3 - x_2\| \geq \frac{1}{2}.$$

Procedendo por indução, obtem-se a sequência $(x_n)_{n\in\mathbb{N}}$ de elementos $x_n \in B$ tal que

$$\|x_m - x_n\| \geq \frac{1}{2},\ m \neq n.$$

Dessa forma, $(x_n)_{n\in\mathbb{N}}$ não possui subsequência convergente, o que contradiz a compacidade de B. Portanto, dim $E < \infty$. □

Assim, se o espaço normado E for de dimensão infinita, todo subconjunto compacto tem interior vazio. Uma importante aplicação deste teorema é o fato de que toda aplicação contínua num conjunto compacto possui imagem também compacta, como mostra o teorema abaixo.

Teorema 4.9. *O conjunto imagem de um subconjunto compacto M de um espaço normado E por uma aplicação contínua é um conjunto compacto.*

Demonstração. Sejam $f : M \subset E \to F$ uma aplicação contínua, onde M é um conjunto compacto, E e F são espaços normados em que a norma em F é a restrição da norma de E. Basta mostrar que toda sequência (y_n) em $f(M) \subset F$ possui uma subsequência que converge para algum ponto em $f(M)$, conforme definição 3.21.

Como $y_n \in f(M)$ tem-se que $y_n = f(x_n)$, com $x_n \in M$ e $n \in \mathbb{N}$. Como M é compacto, a sequência $(x_n)_{n\in\mathbb{N}}$ possui uma subsequência (x_{n_k}) convergente para um elemento $x \in M$. Da continuidade de f segue que $f(x_{n_k}) = y_{n_k}$ é uma subsequência de $(y_n)_{n\in\mathbb{N}}$ em $f(M)$. Ainda, como f é contínua em x, tem-se que $f(x) = y \in f(M)$, ou seja, $y_{n_k} \to y$ em $f(M)$. Portanto, $f(M)$ é compacto. □

A partir deste teorema, tem-se o seguinte resultado.

Corolário 4.2. *Uma aplicação contínua em um subconjunto compacto M de um espaço nomado E em $\mathbb{R}$ assume máximo e mínimo em M.*

Demonstração. Seja $f : M \subset E \to \mathbb{R}$. Pelo teorema 4.9, $f(M)$ é compacto e pelo lema 3.2 (página 47) é fechado e limitado. Sejam $\alpha = \inf f(M)$ e $\beta = \sup f(M)$, tem-se $\alpha, \beta \in f(M)$, ou seja, existem x_1 e x_2 em M tais que $f(x_1) = \alpha$ e $f(x_2) = \beta$, nos quais f assume mínimo e máximo, respectivamente. □

4.4 Transformações lineares

No estudo de espaços vetoriais, em particular, espaços normados, uma aplicação é transformação linear quando preserva as operações de adição vetorial e multiplicação de vetor por escalar, como segue na definição abaixo.

Definição 4.14. *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$. Uma **transformação linear** é uma aplicação $T : D(T) \to W$ tal que para todo $u, v \in D(T)$ e escalar $\alpha \in \mathbb{K}$, tem-se*

1) $T(u + v) = T(u) + T(v)$,

2) $T(\alpha v) = \alpha T(v)$.

O domínio $D(T)$ de T é um espaço vetorial e a imagem $Im(T)$ de T é um subespaço vetorial de W.

Observação 4.7. Note, da definição acima, que $T : D(T) \to W$ é uma transformação linear se, e somente se, $T(\alpha u + v) = \alpha T(u) + T(v)$, para todo $\alpha \in \mathbb{K}$ e $u, v \in D(T)$.

Definição 4.15. *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$ e seja $T : D(T) \to W$ uma transformação linear. O **espaço nulo** ou **núcleo** de T, denotado por $\mathcal{N}(T)$, é o conjunto de todos os vetores $v \in D(T)$ tais que $T(v) = 0$. Ou seja,*

$$\mathcal{N}(T) = \{v \in D(T);\, T(v) = 0\}.$$

Note que $\mathcal{N}(T)$ é um subespaço vetorial de V. Com efeito, para quaisquer vetores v_1 e v_2 em $\mathcal{N}(T)$ e $\alpha \in \mathbb{K}$, se $T(v_1) = T(v_2) = 0$ então,

$$T(\alpha v_1 + v_2) = \alpha T(v_1) + T(v_2) = \alpha 0 + 0 = 0,$$

logo, $v_1 + \alpha v_2 \in \mathcal{N}(T)$. Ainda, da propriedade 2) da definicão 4.14, segue que

$$T(0) = T(0.0) = 0T(0) = 0.$$

Ou seja, toda transformação linear de $D(T)$ em W, leva o elemento neutro de V no elemento neutro de W.

Segue alguns exemplos de transformações lineares.

Exemplo 4.27. A aplicação $T : V \to V$ definida por $T(v) = v$ para todo $v \in V$ é uma transformação linear. De fato, sejam $u, v \in V$ e $\alpha \in \mathbb{K}$ então,

$$T(\alpha u + v) = \alpha u + v = \alpha T(u) + T(v).$$

Diz-se que T é a transformação identidade.

Exemplo 4.28. A aplicação $T : V \to W$ definida por $T(v) = 0$ para todo $v \in V$ é chamada transformação nula. Com efeito, para todo $u, v \in V$ e $\alpha \in \mathbb{K}$ tem-se

$$T(\alpha u + v) = 0 = 0 + 0 = \alpha T(u) + T(v).$$

Exemplo 4.29. Seja $T : \mathcal{C}(\mathbb{R}, \mathbb{R}) \to \mathcal{C}(\mathbb{R}, \mathbb{R})$ definida por $T(f)$ e

$$(Tf)(x) = \int_0^x f(t)\, dt.$$

Então T é uma transformação linear. Com efeito, para todo $x \in \mathbb{R}$ sejam $f, g \in$

$\mathcal{C}(\mathbb{R}, \mathbb{R})$ e $\alpha \in \mathbb{R}$ então,

$$\begin{aligned}(T(\alpha f + g))(x) &= \int_0^x (\alpha f + g)(t)\, dt \\ &= \int_0^x [\alpha f(t) + g(t)]\, dt \\ &= \int_0^x \alpha f(t)\, dt + \int_0^x g(t)\, dt \\ &= \alpha \int_0^x f(t)\, dt + \int_0^x g(t)\, dt \\ &= \alpha (Tf)(x) + (Tg)(x),\ \forall\, x \in \mathbb{R}.\end{aligned}$$

Portanto, $T(\alpha f + g) = \alpha T(f) + T(g)$.

4.4.1 Transformações lineares contínuas e limitadas

Nesta seção será enfatizado o fato de que toda transformação linear limitada em um espaço normado é também contínua (vale a recíproca) e, através disto poderá ser observado que algumas propriedades surgem naturalmente.

Definição 4.16. *Sejam E e F espaços normados e $T : D(T) \to F$ uma transformação linear, onde $D(T) \subset E$. A transformação T é chamada* ***limitada*** *se existe um número real $c > 0$ tal que para todo $x \in D(T)$ tem-se*

$$\|T(x)\|_F \leq c\|x\|_E, \tag{4.19}$$

onde $\|.\|_F$ é uma norma em F e $\|.\|_E$ é a norma em E.

Observação 4.8. De (4.19) se $x \neq 0$ então existe $c > 0$ tal que

$$\frac{\|T(x)\|_F}{\|x\|_E} \leq c. \tag{4.20}$$

Considerando o supremo de $\left\{\dfrac{\|T(x)\|}{\|x\|},\ \forall\, x \in D(T) - \{0\}\right\}$ pode-se definir a norma de T no espaço das transformações por

$$\|T\| = \sup_{\substack{x \in D(T) \\ x \neq 0}} \frac{\|T(x)\|}{\|x\|}. \tag{4.21}$$

Ainda, com $c = \|T\|$ tem-se de (4.19) que

$$\|T(x)\|_F \leq \|T\|\|x\|_E.$$

Lema 4.5. *Seja T uma transformação linear limitada. Então*

i. Uma fórmula alternativa para a norma de T é

$$\|T\| = \sup_{\substack{x\in D(T)\\ \|x\|_E=1}} \|T(x)\|_F.$$

ii. $\|T\| = \sup\limits_{\substack{x\in D(T)\\ x\neq 0}} \dfrac{\|T(x)\|_F}{\|x\|_E}$ *define uma norma.*

Demonstração. i. Para todo $x \in D(T)$ com $x \neq 0$ e da linearidade de T, de (4.21) segue que

$$\|T\| = \sup_{\substack{x\in D(T)\\ x\neq 0}} \frac{1}{\|x\|_E}\|T(x)\|_F = \sup_{\substack{x\in D(T)\\ x\neq 0}} \left\|T\left(\frac{x}{\|x\|_E}\right)\right\|_F = \sup_{\substack{x\in D(T)\\ \|y\|_E=1}} \|T(y)\|_F.$$

onde $y = \dfrac{x}{\|x\|_E}$.

ii. Note que a norma definida em *ii*. satisfaz as três propriedades de norma. Para toda transformação linear limitada T e T_1 e $\alpha \in \mathbb{K}$ onde $\|T\| = \sup\limits_{\substack{x\in D(T)\\ \|x\|_E=1}} \|T(x)\|_F$ segue que

N1) $\|T\| \geq 0$, $\forall\, x \in D(T)$ pela própria definição de norma e ainda,

$$\|T\| = 0 \Leftrightarrow \sup_{\substack{x\in D(T)\\ \|x\|_E=1}} \|T(x)\|_F = 0 \Leftrightarrow T(x) = 0,\ \forall\, x \in D(T), \Leftrightarrow T = 0,$$

ou seja, T é a transformação nula.

N2) $\|\alpha T\| = \sup\limits_{\substack{x\in D(T)\\ \|x\|_E=1}} \|\alpha T(x)\|_F = \sup\limits_{\substack{x\in D(T)\\ \|x\|_E=1}} |\alpha|\|T(x)\|_F = |\alpha| \sup\limits_{\substack{x\in D(T)\\ \|x\|_E=1}} \|T(x)\|_F = |\alpha|\|T\|.$

N3) Desigualdade triangular,

$$\begin{aligned}
\|(T+T_1)\| &= \sup_{\substack{x\in D(T)\\ \|x\|_E=1}} \|(T+T_1)(x)\|_F \\
&= \sup_{\substack{x\in D(T)\\ \|x\|_E=1}} \|T(x)+T_1(x)\|_F \\
&\leq \sup_{\substack{x\in D(T)\\ \|x\|_E=1}} \|T(x)\|_F + \sup_{\substack{x\in D(T)\\ \|x\|_E=1}} \|T_1(x)\|_F \\
&= \|T\| + \|T_1\|.
\end{aligned}$$

□

Segue alguns exemplos de transformações lineares limitadas.

Exemplo 4.30. A transformação identidade $T : V \to V$ definida por $T(v) = v$, para todo $v \in V$, é uma transformação linear limitada cuja norma é $\|T\| = 1$. De fato, para todo $v \in V$,

$$\|T(v)\| = \|v\| = 1\|v\| = \|T\|\|v\|.$$

Exemplo 4.31. O transformação nula $T : V \to W$ definida por $T(v) = 0$, para todo $v \in V$, é uma transformação linear limitada com a norma $\|T\| = 0$. De fato, para todo $v \in V$,

$$\|T(v)\| = \|0\| = 0 = 0\|v\| = \|T\|\|v\|.$$

Exemplo 4.32. Considere a transformação $T : \mathcal{C}([0,1],\mathbb{R}) \to \mathcal{C}([0,1],\mathbb{R})$ definida por $T(f)$ e

$$(Tf)(t) = \int_0^1 k(t,\tau) f(\tau)\, d\tau$$

onde a função $k : [0,1] \times [0,1]$, chamada núcleo de T, é assumida como sendo contínua sobre o quadrado fechado $G = [0,1] \times [0,1]$ no plano $t\tau$. Esta transformação T é linear e limitada. De fato, para toda $f, g \in \mathcal{C}([0,1],\mathbb{R})$ e $\alpha \in \mathbb{R}$ tem-se

$$\begin{aligned}
(T(\alpha f + g))(t) &= \int_0^1 k(t,\tau)(\alpha f + g)(\tau)\, d\tau \\
&= \int_0^1 \left(k(t,\tau)\alpha f(\tau) + k(t,\tau)g(\tau)\right)\, d\tau \\
&= \alpha \int_0^1 k(t,\tau) f(\tau)\, d\tau + \int_0^1 k(t,\tau) g(\tau)\, d\tau \\
&= \alpha (Tf)(t) + (Tg)(t).
\end{aligned}$$

Portanto, $T(\alpha f + g) = \alpha T(f) + T(g)$.

Note que a continuidade de k sobre o quadrado fechado G implica que k é limitado, ou seja, existe $k_0 \in \mathbb{R}$ tal que

$$|k(t,\tau)| \leq k_0,\ \forall\, (t,\tau) \in G.$$

Além disso,

$$|f(t)| \leq \sup_{t \in [0,1]} |f(t)| = \|f\|,$$

para todo $t \in [0,1]$.

Assim,

$$\begin{aligned}
\|T(f)\| = \sup_{t\in[0,1]} \left| \int_0^1 k(t,\tau)f(\tau)\,d\tau \right| &\leq \sup_{t\in[0,1]} \int_0^1 |k(t,\tau)|\,|f(\tau)|\,d\tau \\
&\leq \int_0^1 k_0|f(\tau)|\,d\tau \\
&= k_0 \int_0^1 |f(\tau)|\,d\tau \\
&\leq k_0 \int_0^1 \sup_{s\in[0,1]} |f(s)|\,d\tau \\
&= k_0\|f\| \int_0^1 d(\tau) = k_0\|f\|.
\end{aligned}$$

Portanto, $\|T(f)\| \leq k_0\|f\|$ onde $c = k_0$.

Exemplo 4.33. Seja E o espaço normado de todos os polinômios no intervalo $[0,1]$ com a norma dada por $\|p\| = \max\limits_{0\leq t\leq 1} |p(t)|$. Seja T uma transformação diferenciável em E definida por $T(p)$ e

$$(Tp)(t) = p'(t),$$

ou seja, T representa a primeira derivada do polinômio p na variável t. **Esta transformação é linear mas não é limitada**. Com efeito, sejam $p, q \in E$ e $\alpha \in \mathbb{R}$ então

$$\begin{aligned}
(T(\alpha p + q))(t) &= (\alpha p + q)'(t) \\
&= (\alpha p)'(t) + (q)'(t) \\
&= \alpha(p)'(t) + (q)'(t) \\
&= \alpha(Tp)(t) + (Tq)(t),\ \forall\, t \in [0,1].
\end{aligned}$$

Portanto, $T(\alpha p + q) = \alpha T(p) + T(q)$.

Para ver que a transformação T não é limitada, considere o polinônio $p_n(t) = t^n$, $n \in \mathbb{N}$. Como $\|p\| = \max\limits_{0\leq t\leq 1} |p(t)|$ então $\|p_n\| = 1$ e assim,

$$\|(Tp_n)(t)\| = \|(p_n)'(t)\| = \|nt^{n-1}\| = |n|\|t^{n-1}\| = n\|t^{n-1}\| = n1 = n$$

ou seja,

$$\frac{\|T(p_n)\|}{\|p_n\|} = n$$

e como $n \in \mathbb{N}$, fixando $c > 0$ sempre existe $n \in \mathbb{N}$ tal que $n > c$, portanto a transformação T não é limitada.

Note que o espaço de todos os polinômios tem dimensão infinita. Para espaços de dimensão finita tem-se o seguinte teorema.

Teorema 4.10. *Se o espaço normado E é de dimensão finita, então toda transformação linear em E é limitada.*

Demonstração. Considere $\dim E = n$ e $\{e_1, \ldots, e_n\}$ uma base para E. Assim, para todo $x \in E$ existem únicos escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que $x = \sum_{j=1}^{n} \alpha_j e_j$. Seja T uma transformação linear qualquer em E. Como T é linear então

$$\begin{aligned} \|T(x)\| = \left\| T\left(\sum_{j=1}^{n} \alpha_j e_j \right) \right\| &= \left\| \left(\sum_{j=1}^{n} \alpha_j T(e_j) \right) \right\| \\ &\leq \sum_{j=1}^{n} |\alpha_j| \|T(e_j)\| \\ &\leq \max_{1 \leq j \leq n} \|T(e_j)\| \sum_{j=1}^{n} |\alpha_j|. \end{aligned} \tag{4.22}$$

Agora, pelo lema 4.3, como

$$\left\| \sum_{j=1}^{n} \alpha_j e_j \right\| \geq c \left(\sum_{j=1}^{n} |\alpha_j| \right)$$

segue que

$$\sum_{j=1}^{n} |\alpha_j| \leq \frac{1}{c} \left\| \sum_{j=1}^{n} \alpha_j e_j \right\| = \frac{1}{c} \|x\|. \tag{4.23}$$

Substituindo (4.23) em (4.22) obtem-se

$$|T(x)\| \leq \max_{1 \leq j \leq n} \|T(e_j)\| \frac{1}{c} \|x\| = \frac{1}{c} \max_{1 \leq j \leq n} \|T(e_j)\| \|x\| = \kappa \|x\|,$$

onde $\kappa = \frac{1}{c} \max_{1 \leq j \leq n} \|T(e_j)\|$. Portanto,

$$\|T(x)\| \leq \kappa \|x\|,$$

ou seja, T é limitada. □

Note que quando se fala de continuidade em uma aplicação, não necessariamente linear, diz-se que tal aplicação é contínua se ela é contínua em todo ponto do domínio. Agora, se tal aplicação é uma transformação linear, basta mostrar que ela é contínua em um único ponto para concluir daí que é contínua em todo ponto do seu domínio, como mostra o teorema seguinte.

Teorema 4.11. *Sejam E e F espaços normados. Considere $T : D(T) \to F$ uma transformação linear, onde $D(T) \subset E$. Então:*

i. T é contínua se, e somente se, T é limitada.

ii. Se T é contínua em um único ponto, então T é contínua.

Demonstração. i. Seja T uma transformação linear contínua em um ponto $x_0 \in D(T)$, então dado $\epsilon_0 > 0$ fixo existe $\delta_0 > 0$ tal que para todo $x \in D(T)$ tem-se

$$\|x - x_0\| < \delta_0 \Rightarrow \|T(x) - T(x_0)\| < \epsilon_0.$$

Considere qualquer $y \neq 0 \in D(T)$. Escrevendo $x = x_0 + \dfrac{\delta_0}{\|y\|} y$ então $x - x_0 = \dfrac{\delta_0}{\|y\|} y$ onde $\|x - x_0\| = \delta_0$. Assim, da continuidade e da linearidade de T segue que

$$\|T(x) - T(x_0)\| = \|T(x - x_0)\| = \left\| T\left(\frac{\delta_0}{\|y\|} y\right) \right\| = \frac{\delta_0}{\|y\|} \|T(y)\| < \epsilon_0.$$

Logo, $\|T(y)\| \leq \dfrac{\epsilon_0}{\delta_0} \|y\|$. Considerando $c = \dfrac{\epsilon_0}{\delta_0}$ tem-se $\|T(y)\| \leq c\|y\|$. Portanto, T é limitada.

Reciprocamente, seja T uma transformação linear limitada. Se $T = 0$ segue o resultado. Suponha então $T \neq 0$, onde $\|T\| \neq 0$. Sendo T limitada, seja $c = \|T\|$ e considere qualquer $x_0 \in D(T)$. Assim, considerando $\delta = \dfrac{\epsilon}{\|T\|}$ e sendo T linear, então para todo $x \in D(T)$ se $\|x - x_0\| < \delta$ tem-se

$$\|T(x) - T(x_0)\| = \|T(x - x_0)\| \leq \|T\| \|x - x_0\| < \|T\| \frac{\epsilon}{\|T\|} = \epsilon.$$

Ou seja, $\|x - x_0\| < \delta \Rightarrow \|T(x) - T(x_0)\| < \epsilon$ e portanto, T é contínua.

ii. Se T é contínua em um ponto do seu domínio, então pela prova do item *i.* segue que T é limitada, que por sua vez implica na continuidade de T por *i.*. □

O corolário abaixo se refere à continuidade e ao espaço nulo da transformação linear limitada T.

Corolário 4.3. *Sejam T uma transformação linear limitada, $x_n, x \in D(T)$ e (x_n) uma sequência em $D(T)$. Então:*

i. Se $x_n \to x$ então $T(x_n) \to T(x)$.

ii. O espaço nulo $\mathcal{N}(T)$ é fechado.

Demonstração. i. Como T é limitada, pelo teorema 4.11, T é contínua. Usando o teorema 3.3 (página 47), segue o resultado.

ii. Para todo $x \in \overline{\mathcal{N}(T)}$ existe uma sequência (x_n) em $\mathcal{N}(T)$ tal que $x_n \to x$. Pelo item (i) deste corolário, $T(x_n) \to T(x)$. Como $(x_n) \in \mathcal{N}(T)$ então $T(x_n) = 0$ para todo $n \in \mathbb{N}$, pela unicidade do limite $T(x) = 0$, logo $x \in \mathcal{N}(T)$. Portanto, $\overline{\mathcal{N}(T)} = \mathcal{N}(T)$ e, pela proposição 3.6 (página 42), $\mathcal{N}(T)$ é fechado. □

Proposição 4.6. *Sejam E e F espaços normados. As seguintes afirmações a respeito de uma transformação linear $T : E \to F$ são equivalentes:*

i. T é contínua;

ii. T é contínua no ponto $0 \in E$;

iii. Existe $c > 0$ tal que $\|T(x)\| \leq c\|x\|$ para todo $x \in E$;

iv. Existe $c > 0$ tal que $\|T(x) - T(y)\| \leq c\|x - y\|$ para todo $x, y \in E$.

Demonstração. *i.* $\Rightarrow$ *ii.* Imediato.

ii. $\Rightarrow$ *iii.* Se T é contínua no ponto $0 \in E$, pelo item *ii* do teorema 4.11 (página 89), T é contínua e por sua vez é limitada, pelo item *i* do teorema 4.11.

iii. $\Rightarrow$ *iv.* Note que sendo T linear, a hipótese *iii.* implica que existe $c > 0$ tal que

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq c\|x - y\|, \forall\, x, y \in E.$$

iv. $\Rightarrow$ *i.* Para qualquer $\epsilon > 0$ existe $\delta = \dfrac{\epsilon}{c} > 0$ tal que

$$\|x - y\| < \delta \Rightarrow \|T(x) - T(y)\| \leq c\|x - y\| < c\delta = \epsilon, \forall\, x, y \in E.$$

Logo, T é contínua. □

Obtem-se como resultados da proposição anterior os seguintes corolários.

Corolário 4.4. *Seja $T : E \to F$ uma bijeção linear. Para que T seja um homeomorfismo, é necessário e suficiente que existam $\alpha > 0$ e $\beta > 0$ tais que*

$$\alpha\|x\| \leq \|T(x)\| \leq \beta\|x\|, \forall\, x \in E.$$

Corolário 4.5. *Sejam $\|.\|_0$ e $\|.\|_1$ normas sobre um espaço vetorial E. As seguintes propriedades são equivalentes:*

i. $\|.\|_0$ e $\|.\|_1$ são normas equivalentes.

ii. A aplicação identidade $Id : (E, \|.\|_0) \to (E, \|.\|_1)$ é contínua e a inversa também é contínua.

Segue um exemplo de normas definidas sobre o mesmo espaço, mas que não são equivalentes.

Exemplo 4.34. Considere $\|f\|_0 = \sup\{|f_k|; x \in [0, 1]\}$ e $\|f\|_2 = \left(\int_0^1 f_k^2\right)^{\frac{1}{2}}$ normas em $\mathcal{C}([0, 1], \mathbb{R})$. Então, $\|f\|_0$ e $\|f\|_2$ **não são equivalentes**.

Com efeito, considere a sequência (f_k), definida da seguinte forma

$$f_k(x) = \begin{cases} kx, & \text{se } 0 \leq x \leq \frac{1}{k} \\ 2 - kx, & \text{se } \frac{1}{k} \leq x \leq \frac{2}{k} \\ 0, & \text{se } \frac{2}{k} \leq x \leq 1 \end{cases}$$

conforme figura 4.5.

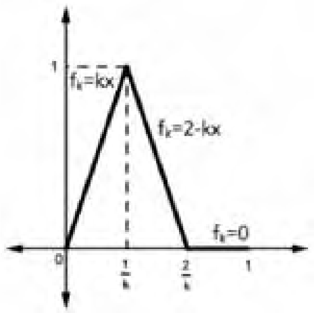

Figura 4.5: *Gráfico de f_k*

Note que na norma $\|f\|_2$, $f_k \to 0$ quando $k \to \infty$, pois

$$\begin{aligned} \|f_k - 0\|_2 = \left(\int_0^1 f_k^2\right)^{\frac{1}{2}} &= \left(\int_0^{1/k} (kx)^2\, dx + \int_{1/k}^{2/k} (2-kx)^2\, dx\right)^{\frac{1}{2}} \\ &= \left(\int_0^{1/k} (k^2x^2)\, dx + \frac{1}{k}\int_0^1 u^2\, du\right)^{\frac{1}{2}} \\ &= \left(k^2\left[\left.\frac{x^3}{3}\right|_0^{1/k}\right] + \frac{1}{k}\left[\left.\frac{u^3}{3}\right|_0^1\right]\right)^{\frac{1}{2}} \\ &= \left(\frac{k^2}{3k^3} + \frac{1}{3k}\right)^{\frac{1}{2}} = \left(\frac{2}{3k}\right)^{\frac{1}{2}}. \end{aligned}$$

Por outro lado, (f_k) não converge para a função nula na norma $\|\|_0$, pois

$$\|f_k - 0\|_0 = \sup\{|f_k|; x \in [0,1]\} = 1.$$

Portanto, a aplicação identidade $Id : (C[0,1], \|.\|_0) \to (C[0,1], \|.\|_2)$ não é contínua no ponto $f = 0$ de $\mathcal{C}([0,1], \mathbb{R})$, implicando, pelo corolário 4.5, que $\|f\|_0$ e $\|f\|_2$ não são equivalentes.

Definição 4.17. *Sejam T_1 e T_2 transformações lineares. Diz-se que T_1 e T_2 são **transformações iguais**, se $D(T_1) = D(T_2)$ e se $T_1(x) = T_2(x)$ para todo $x \in D(T_1) =$*

$D(T_2)$.

*Seja T uma transformação linear. A **restrição** de $T : D(T) \to W$ a um subconjunto $B \subset D(T)$ é denotado por $T\mid_B$ e esta transformação é definida por $T\mid_B : B \to W$ tal que*

$$T\mid_B (x) = T(x), \, \forall \, x \in B.$$

*A **extensão** da transformação T a um conjunto $M \supset D(T)$ é uma transformação $\tilde{T} : M \to W$ tal que*

$$\tilde{T}\mid_{D(T)} = T$$

isto é, $\tilde{T}(x) = T(x)$ para todo $x \in D(T)$. Assim, T é a restrição de $\tilde{T}$ em $D(T)$.

O teorema a seguir se refere a extensão de uma transformação linear limitada.

Teorema 4.12. *Sejam E um espaço normado e F um espaço de Banach. Considere $T : D(T) \to F$ uma transformação linear limitada, onde $D(T) \subset E$. Então T possui uma extensão*

$$\tilde{T} : \overline{D(T)} \to F$$

onde $\tilde{T}$ é uma transformação linear limitada de norma $\|\tilde{T}\| = \|T\|$.

Demonstração. Considere qualquer $x \in \overline{D(T)}$. Se $x \in D(T)$ então existe uma sequência $(x_n) \in D(T)$ tal que $x_n \to x$, logo existe $T(x)$. Suponha que $x \notin D(T)$, então dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $m, n > n_0$ tem-se $\|x_m - x_n\| < \dfrac{\epsilon}{\|T\|}$. Como T é linear e limitada, tomando $m, n > n_0$

$$\|T(x_m) - T(x_n)\| = \|T(x_m - x_n)\| = \|T\|\|x_m - x_n\| \leq \|T\|\frac{\epsilon}{\|T\|} = \epsilon.$$

Portanto, a sequência $(T(x_n))$ é de Cauchy em F e como F é completo, existe $y \in F$ tal que $T(x_n) \to y$ em F.

Defina $\tilde{T}(x) = y$. Note que esta definição não depende da escolha da sequência em $D(T)$ que converge para x. Suponha que $x_n, z_n \in D(T)$ tais que $x_n \to x$ e $z_n \to x$. Considere a sequência (v_m) tal que $(v_m) = (x_1, z_1, x_2, z_2, \ldots)$. Como cada um dos termos desta sequência converge para x, então a sequência $v_m \to x$. Assim, pelo item i. do corolário 4.3 (página 90), a sequência $(T(v_m))$ é convergente e possui o mesmo limite que as subsequências $(T(x_n))$ e $(T(z_n))$, ou seja, $T(v_m) \to y$.

Isto mostra que a transformação $\tilde{T}$ é unicamente definida para cada $x \in \overline{D(T)}$. Como T é linear então $\tilde{T}$ é linear, pois

$$\begin{aligned}
\tilde{T}(\alpha x + z) = & \ \tilde{T}\left(\lim_{n\to\infty} \alpha x_n + \lim_{n\to\infty} z\right) \\
= & \ \lim_{n\to\infty} T(\alpha x_n + z) \\
= & \ \lim_{n\to\infty} T(\alpha x_n) + \lim_{n\to\infty} T(z) \\
= & \ \alpha\tilde{T}(x) + \tilde{T}(z)
\end{aligned}$$

e como $\tilde{T}(x) = T(x)$ para todo $x \in D(T)$, então $\tilde{T}$ é uma extensão de T.

Agora, como T é limitada então $\|T(x_n)\| \leq \|T\|\|x_n\|$. Além disso $x_n \to x$ e $T(x_n) \to y = \tilde{T}(x)$, quando $n \to \infty$, e sendo a norma uma aplicação contínua, segue que

$$\|\tilde{T}(x)\| \leq \|T\|\|x\|.$$

Ou seja, $\tilde{T}$ é limitada e $\|\tilde{T}\| \leq \|T\|$. Por outro lado, $\|\tilde{T}\| \geq \|T\|$ pois sendo a norma definida por um supremo, não pode diminuir em uma extensão. Portanto, $\|\tilde{T}\| = \|T\|$. □

4.5 Funcionais lineares e espaço dual

Um funcional linear é uma transformação linear cuja imagem está no corpo dos escalares. Se o espaço considerado for um espaço vetorial V, então o espaço dual de V, que será denotado por V^*, é um espaço vetorial constituído de todos os funcionais lineares definidos em V. Agora, se E for um espaço vetorial normado então o espaço dual de E, denotado por E', é um espaço normado formado de todos os funcionais lineares limitados.

Definição 4.18. *Seja V um espaço vetorial sobre o corpo $\mathbb{K}$. Um* ***funcional linear*** *f é uma transformação linear com domínio $D(f)$ em um espaço vetorial V e imagem no corpo dos escalares $\mathbb{K}$, ou seja, $f : D(f) \subset V \to \mathbb{K}$, onde $\mathbb{K} = \mathbb{R}$ se V é o espaço vetorial real e $\mathbb{K} = \mathbb{C}$ se V é o espaço vetorial complexo.*

Exemplo 4.35. Considere o espaço vetorial $V = \mathbb{R}^n$ e seja $v_0 \neq 0$ um vetor fixo de V. A aplicação $f : V \to \mathbb{R}$ tal que $f(v) = \langle v, v_0 \rangle$, para todo $v \in V$, onde $\langle , \rangle$ representa o produto escalar de v por v_0, é um funcional linear. De fato, sejam $v_1, v_2 \in V$ e $\alpha \in \mathbb{R}$ então

$$\begin{aligned} f(\alpha v_1 + v_2) &= \langle \alpha v_1 + v_2, v_0 \rangle \\ &= \langle \alpha v_1, v_0 \rangle + \langle v_2, v_0 \rangle \\ &= \alpha \langle v_1, v_0 \rangle + \langle v_2, v_0 \rangle \\ &= \alpha f(v_1) + f(v_2). \end{aligned}$$

Definição 4.19. *Um* ***funcional linear limitado*** *f é uma transformação linear limitada com domínio $D(f)$ em um espaço normado E e cuja imagem está no corpo dos escalares $\mathbb{K}$, ou seja, $f : D(f) \subset E \to \mathbb{K}$. Assim, existe um número real $c > 0$, tal que para todo $x \in D(f)$,*

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

Além disso, a norma de f é

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

ou

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

Observação 4.9. Note que com $c = \|f\|$ na definição acima, então, $|f(x)| \leq \|f\|\|x\|$.

Teorema 4.13. *Sejam E um espaço normado. Um funcional linear $f : D(f) \to \mathbb{K}$, onde $D(f) \subset E$, é contínuo se, e somente se, f é limitado.*

A demonstração é análoga à do teorema 4.11 (página 89) para o caso em que $F = \mathbb{K}$.

Exemplo 4.36. Seja $f \in \mathcal{C}([a,b],\mathbb{R})$ definida por

$$f(\varphi) = \int_a^b \varphi(t)\,dt$$

onde $\varphi \in \mathcal{C}([a,b],\mathbb{R})$. Note que f é um funcional linear, pois para toda $\varphi,\psi \in C([a,b],\mathbb{R})$ e $\alpha \in \mathbb{R}$ tem-se

$$\begin{aligned} f(\alpha\varphi+\psi) &= \int_a^b (\alpha\varphi+\psi)(t)\,dt \\ &= \int_a^b \alpha\varphi(t)+\psi(t)\,dt \\ &= \alpha\int_a^b \varphi(t)\,dt + \int_a^b \psi(t)\,dt \\ &= \alpha f(\varphi)+f(\psi). \end{aligned}$$

Além disso, f é um funcional limitado e têm norma $\|f\| = b-a$. Com efeito, a norma considerada no espaço normado $\mathcal{C}([a,b],\mathbb{R})$ é dada por $\|\varphi\| = \sup_{t\in[a,b]} |\varphi(t)|$, conforme exemplo 4.24 (página 77). Assim,

$$|f(\varphi)| = \left|\int_a^b \varphi(t)\,dt\right| \leq (b-a)\sup_{t\in[a,b]} |\varphi(t)| = (b-a)\|\varphi\|,$$

ou seja,

$$|f(\varphi)| \leq (b-a)\|\varphi\|. \tag{4.24}$$

Escrevendo $c = b-a$, segue que $|f(\varphi)| \leq c\|\varphi\|$, ou seja, f é limitado.

Agora, tomando o supremo sobre toda $\varphi \in \mathcal{C}([a,b],\mathbb{R})$ de norma $\|\varphi\| = 1$ em (4.24), ou seja,

$$\sup_{\substack{\varphi\in\mathcal{C}\\ \|\varphi\|=1}} |f(\varphi)| \leq (b-a)\sup_{\substack{\varphi\in\mathcal{C}\\ \|\varphi\|=1}} \|\varphi\|$$

obtem-se

$$\|f\| \leq (b-a). \tag{4.25}$$

Escolhendo, em particular, $\varphi = \varphi_0 = 1$, note que $\|\varphi_0\| = 1$. Da observação 4.9 (página 95), segue que $|f(\varphi_0)| \leq \|f\|\|\varphi_0\|$. Assim,

$$\|f\| \geq \frac{|f(\varphi_0)|}{\|\varphi_0\|} = |f(\varphi_0)| = \int_a^b dt = b-a,$$

ou seja,

$$\|f\| \geq b-a. \tag{4.26}$$

Portanto, por (4.25) e (4.26), $\|f\| = b-a$.

Exemplo 4.37. Seja f um funcional linear sobre o espaço ℓ^2, definido no exemplo 3.8 (página 37). Seja $a = (\alpha_i) = (\alpha_1, \alpha_2, \ldots) \in \ell^2$ um elemento fixado. O funcional definido por

$$f(x) = \sum_{i=1}^{\infty} \alpha_i \eta_i, \text{ com } x = (\eta_i),$$

é linear, pois

$$f(\lambda x + y) = \sum_{i=1}^{\infty} (\lambda \alpha_i \eta_i + \alpha_i \mu_i) = \lambda \sum_{i=1}^{\infty} \alpha_i \eta_i + \sum_{i=1}^{\infty} \alpha_i \mu_i = \lambda f(x) + f(y),$$

onde $\lambda \in \mathbb{K}$ e $y = (\mu_i) = (\mu_1, \mu_2, \ldots) \in \ell^2$. Além disso f é limitado. Com efeito, note que fazendo $p = 2$ na proposição 3.1 (página 34), Desigualdade de Hölder, obtem-se a Desigualdade de Cauchy-Schwarz. Assim,

$$|f(x)| = \left| \sum_{i=1}^{\infty} \alpha_i \eta_i \right| \leq \sum_{i=1}^{\infty} |\alpha_i \eta_i| \leq \sqrt{\sum_{i=1}^{\infty} |\alpha_i|^2} \sqrt{\sum_{i=1}^{\infty} |\eta_i|^2} = \|a\|\|x\|.$$

Como a é fixo, tomando $c = \|a\|$ obtem-se $|f(x)| \leq c\|x\|$, ou seja, f é limitado.

Agora, dado um espaço vetorial V pode-se a partir deste definir outros espaços vetoriais, como por exemplo, o espaço vetorial formado por todos os funcionais lineares definidos em V.

Definição 4.20. *Seja V um espaço vetorial. Considere o conjunto V^* formado por todos os funcionais lineares definidos sobre V, ou seja,*

$$V^* = \{f : V \to \mathbb{K};\ f \text{ é linear}\}.$$

Este espaço vetorial definido por V^* é chamado ***espaço dual*** [2] *de V. As operações em V^* são definidas da seguinte forma:*

[2] Também chamado espaço dual *algébrico*. Note que esta definição não utiliza norma.

i. $(f+g)(x) = f(x) + g(x)$, $\forall\, f, g \in V^*$, $x \in V$.

ii. $(\alpha f)(x) = \alpha f(x)$, $\forall\, f \in V^*$, $x \in V$, $\alpha \in \mathbb{K}$.

Considere, inicialmente, espaços de dimensão finita. Neste caso é possível expressar uma base para V^* a partir de uma base de V, como será visto no teorema 4.14. Observe que Se V um espaço vetorial de dimensão finita e $\{e_1, e_2, \ldots, e_n\}$ uma base para V, então para todo $x \in V$ existem únicos escalares $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{K}$ tais que $x = \sum_{j=1}^{n} \xi_j e_j$. Como visto na definição anterior, os funcionais em V constituem o espaço dual V^* de V. Assim, para cada $f \in V^*$ e $x = \sum_{j=1}^{n} \xi_j e_j \in V$, tem-se

$$f(x) = f\left(\sum_{j=1}^{n} \xi_j e_j\right) = \sum_{j=1}^{n} \xi_j f(e_j) = \sum_{j=1}^{n} \xi_j \alpha_j \tag{4.27}$$

onde $\alpha_j = f(e_j), j \in \mathbb{N}$. f é unicamente determinado pelos valores α_j nos n vetores da base de V. Por outro lado, toda n-upla de escalares $\alpha_1, \ldots, \alpha_n$ determina um funcional linear em V, como em (4.27).

Definição 4.21. *Sejam V um espaço vetorial de dimensão finita e $B = \{e_1, e_2, \ldots, e_n\}$ uma base de V. Então o conjunto $B^* = \{f_1, f_2, \ldots, f_n\}$ é chamado a **base dual** de B, onde*

$$f_i(e_j) = \delta_{ij} = \begin{cases} 1, & se\ i = j \\ 0, & se\ i \neq j. \end{cases} \tag{4.28}$$

Isto é justificado pelo seguinte teorema.

Teorema 4.14. *Sejam V um espaço vetorial n-dimensional e $B = \{e_1, \ldots, e_n\}$ uma base de V. Então, $B^* = \{f_1, \ldots, f_n\}$, dado por (4.28), é uma base para o dual V^* de V, e $\dim V^* = \dim V = n$.*

Demonstração. Para mostrar que B^* é uma base para V^*, duas condições precisam ser verificadas:

i. B^* é um conjunto linearmente independente.

ii. $[B^*] = V^*$, ou seja, que B^* gera V^*.

i. Considere a seguinte combinação linear

$$\beta_1 f_1 + \ldots + \beta_n f_n = \sum_{i=1}^{n} \beta_i f_i = 0.$$

Assim, para todo $x \in V$, da igualdade anterior segue que

$$(\beta_1 f_1 + \ldots + \beta_n f_n)(x) = \left(\sum_{i=1}^{n} \beta_i f_i\right)(x) = \sum_{i=1}^{n} \beta_i f_i(x) = 0.$$

Em particular, se $x = e_j$, $j \in \{1, \ldots, n\}$, então

$$\sum_{i=1}^{n} \beta_i f_i(e_j) = \sum_{i=1}^{n} \beta_i \delta_{ij} = \sum_{i=1}^{n} \beta_i = \beta_j = 0,$$

ou seja, $\beta_1 = \beta_2 = \ldots = \beta_n = 0$.

ii. Como foi mencionado anteriormente, cada $f \in V^*$ pode ser unicamente representado como uma combinação linear dos elementos de B^*, ou seja, para cada $f \in V^*$ e $x = \sum_{j=1}^{n} \xi_j e_j \in V$, tem-se

$$f(x) = \sum_{j=1}^{n} \xi_j \alpha_j \tag{4.29}$$

com $\alpha_j = f(e_j)$.

Por outro lado, usando a definição 4.21,

$$f_j(x) = f_j\left(\sum_{k=1}^{n} \xi_k e_k\right) = \sum_{k=1}^{n} \xi_k f_j(e_k) = \sum_{k=1}^{n} \xi_k = \xi_j.$$

Portanto, substituindo $\xi_j = f_j(x)$ na equação (4.29) obtem-se

$$f(x) = \sum_{j=1}^{n} \alpha_j f_j(x).$$

Como x é qualquer, a única representação do funcional linear arbitrário f em V em termos dos funcionais $f_1, \ldots, f_n$ é

$$f = \alpha_1 f_1 + \ldots + \alpha_n f_n,$$

ou seja, o funcional f pode ser escrito em termos dos elementos do conjunto B^*, logo B^* gera V^*. Assim, $\dim V^* = n = \dim V$. □

Exemplo 4.38. Seja $B = \{(1, 1, 0), (0, 1, 0), (0, 0, 2)\}$ uma base do $\mathbb{R}^3$, de forma que $x_1 = (1, 1, 0)$, $x_2 = (0, 1, 0)$ e $x_3 = (0, 0, 2)$. Para encontrar uma base $B^* = \{f_1, f_2, f_3\}$ dual de B, note que

$$\begin{array}{ccc} f_1 : \mathbb{R}^3 \to \mathbb{R} & f_2 : \mathbb{R}^3 \to \mathbb{R} & f_3 : \mathbb{R}^3 \to \mathbb{R} \\ f_1(x_1) = \delta_{11} = 1 & f_2(x_1) = \delta_{21} = 0 & f_3(x_1) = \delta_{31} = 0 \\ f_1(x_2) = \delta_{12} = 0 & f_2(x_2) = \delta_{22} = 1 & f_3(x_2) = \delta_{32} = 0 \\ f_1(x_3) = \delta_{13} = 0 & f_2(x_3) = \delta_{23} = 0 & f_3(x_3) = \delta_{33} = 1 \end{array}$$

Agora, qualquer elemento de $\mathbb{R}^3$ pode ser escrito como combinação linear dos vetores da base de $\mathbb{R}^3$. Logo,

$$(x, y, z) = \alpha(1, 1, 0) + \beta(0, 1, 0) + \gamma(0, 0, 2)$$

ou seja,

$$\begin{aligned}
&\alpha = x \\
&\alpha + \beta = y \Rightarrow \beta = y - x \\
&2\gamma = z \Rightarrow \gamma = \frac{z}{2}
\end{aligned}$$

Assim,

$$\begin{aligned}
f_1(x, y, z) = &\ \alpha f_1(x_1) + \beta f_1(x_2) + \gamma f_1(x_3) \\
= &\ x.1 + (y - x).0 + \left(\frac{z}{2}\right).0 \\
= &\ x
\end{aligned}$$

$$\begin{aligned}
f_2(x, y, z) = &\ \alpha f_2(x_1) + \beta f_2(x_2) + \gamma f_2(x_3) \\
= &\ x.0 + (y - x).1 + \left(\frac{z}{2}\right).0 \\
= &\ y - x
\end{aligned}$$

$$\begin{aligned}
f_3(x, y, z) = &\ \alpha f_3(x_1) + \beta f_3(x_2) + \gamma f_3(x_3) \\
= &\ x.0 + (y - x).0 + \left(\frac{z}{2}\right).1 \\
= &\ \frac{z}{2}
\end{aligned}$$

Portanto, a base dual B^* de B é dada pelos funcionais lineares f_1, f_2, f_3 onde $f_1(x, y, z) = x$, $f_2(x, y, z) = y - x$ e $f_3(x, y, z) = \frac{z}{2}$.

Considerando, agora, quaisquer dois espaços normados E e F, o conjunto $\mathcal{B}(E, F)$, formado de todas as transformações lineares limitadas de E em F, também forma um espaço normado. Como consequência do lema 4.5 (página 86), tem-se o seguinte resultado.

Teorema 4.15. *Sejam E e F espaços normados. O espaço vetorial $\mathcal{B}(E, F)$ é um espaço normado com a norma definida por*

$$\|T\| = \sup_{\substack{x \in E \\ x \neq 0}} \frac{\|T(x)\|}{\|x\|} = \sup_{\substack{x \in E \\ \|x\| = 1}} \|T(x)\|.$$

Demonstração. As operações definidas no conjunto $\mathcal{B}(E, F)$ são as usuais, ou seja, $T_1, T_2 \in \mathcal{B}(E, F)$ e escalar $\alpha \in \mathbb{K}$, tem-se

i. $(T_1 + T_2)(x) = T_1(x) + T_2(x)$, $\forall\, x \in E$.

ii. $(\alpha T)(x) = \alpha T(x)$, $\forall\, x \in E$.

A demonstração de que o espaço vetorial $\mathcal{B}(E, F)$ com a norma do supremo, definida acima, é um espaço normado é análoga à prova do item *ii.* do lema 4.5 (página 86). $\square$

O resultado seguinte afirma que se E é um espaço normado, não necessariamente completo, o espaço normado $\mathcal{B}(E, F)$ é completo se F o for.

Teorema 4.16. *Se E é um espaço normado e F é um espaço de Banach, então $\mathcal{B}(E, F)$ é Banach.*

Demonstração. Seja (T_n) uma sequência de Cauchy em $\mathcal{B}(E, F)$, então dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $m, n > n_0$ tem-se $\|T_n - T_m\| < \epsilon$. Como (T_n) é limitada, para todo $x \in E$ e $m, n > n_0$ obtem-se

$$\|T_n(x) - T_m(x)\| = \|(T_n - T_m)(x)\| \leq \|T_n - T_m\|\|x\| < \epsilon\|x\|. \tag{4.30}$$

Agora, para qualquer x fixo e dado $\tilde{\epsilon} > 0$ pode-se escolher $\epsilon = \epsilon_x$ tal que $\epsilon_x\|x\| < \tilde{\epsilon}$. De (4.30), tem-se

$$\|T_n(x) - T_m(x)\| < \tilde{\epsilon}, \tag{4.31}$$

logo a sequência $(T_n(x))$ é de Cauchy em F, para todo $x \in E$. Como F é de Banach, existe $y = T(x) \in F$ tal que $T_n(x) \to y$. Isto define uma aplicação $T : E \to F$, onde $y = T(x)$. Afirmação: $T = \lim_{n\to\infty} T_n$. Esta aplicação é linear, pois $\forall\, x, z \in E$ e $\alpha \in K$, tem-se

$$\begin{aligned} T(\alpha x + z) &= \lim_{n\to\infty} T_n(\alpha x + z) \\ &= \lim_{n\to\infty} (\alpha T_n(x) + T_n(z)) \\ &= \alpha \lim_{n\to\infty} T_n(x) + \lim_{n\to\infty} T_n(z) \\ &= \alpha T(x) + T(z). \end{aligned}$$

Fazendo $m \to \infty$ em (4.30), obtem-se

$$\|T_n(x) - T(x)\| = \|(T_n - T)(x)\| \leq \epsilon\|x\|, \tag{4.32}$$

para todo $x \in E$ e $n > n_0$. Isto mostra que $(T_n - T) \in \mathcal{B}(E, F)$. Como T_n é limitada, $T = T_n - (T_n - T)$ é limitada, isto é $T \in \mathcal{B}(E, F)$. Além disso, considerando em (4.32) o supremo sobre todo x de norma 1, obtem-se para $n > n_0$

$$\|T_n - T\| \leq \epsilon.$$

Isto é, $T_n \to T$ em $\mathcal{B}(E, F)$, portanto $\mathcal{B}(E, F)$ é Banach. □

Este teorema possibilita a seguinte definição.

Definição 4.22. *Seja E um espaço normado. Então o conjunto de todos os funcionais lineares limitados em E constitui um espaço normado com a norma definida por*

$$\|f\| = \sup_{\substack{x\in E \\ x\neq 0}} \frac{|f(x)|}{\|x\|} = \sup_{\substack{x\in E \\ \|x\|=1}} |f(x)|,$$

e é chamado ***espaço dual*** [3] *de E e denotado por E'.*

Teorema 4.17. *O espaço dual E' de um espaço normado E é um espaço de Banach, sendo E de Banach ou não.*

Demonstração. Note que $E' = \{f : E \to \mathbb{K};\, f \text{ é linear e limitado}\}$. Assim, E' é um subconjunto de $\mathcal{B}(E, F)$, onde $F = \mathbb{K}$, ou seja, $F = \mathbb{R}$ ou $F = \mathbb{C}$. Como $\mathbb{K}$ é completo, aplicando o teorema 4.16, segue que o espaço normado E' é Banach. □

Antes de apresentar alguns exemplos de espaço dual, será apresentado um conceito muito utilizado quando se quer investigar determinados espaços e que, em geral, inclui seus respectivos duais. Neste caso, o conceito de isomorfismo será útil, pois o dual de um espaço normado pode ser "identificado" com outro espaço normado conhecido, ou seja, o isomorfismo preserva certas características que facilitam tal investigação. Para tanto, algumas propriedades precisam ser verificadas. Segue então a definição de isomorfismo entre espaços normados.

Definição 4.23. *Um* ***isomorfismo*** *de um espaço normado E em um espaço normado $\tilde{E}$ é uma transformação linear bijetora $T : E \to \tilde{E}$ que preserva a norma, isto é, para todo $x \in E$, $\|T(x)\| = \|x\|$.*

Portanto T é uma transformação isométrica. Diz-se que E e $\tilde{E}$ são isomorfos e denota-se $E \cong \tilde{E}$.

Exemplo 4.39. O espaço dual de $\mathbb{R}^n$ é $\mathbb{R}^n$. Basta mostrar cada elemento em $(\mathbb{R}^n)'$ pode ser identificado com um elemento em $\mathbb{R}^n$, isto é, $(\mathbb{R}^n)' \cong \mathbb{R}^n$ (pois todo funcional linear é contínuo).

Para todo $x \in \mathbb{R}^n$, considere a norma definida por $\|x\| = \left(\sum_{k=1}^{n} |\xi_k|^2\right)^{\frac{1}{2}}$, em que $x = (\xi_1, \ldots, \xi_n)$. Seja $\{e_1, \ldots, e_n\}$ a base canônica do $\mathbb{R}^n$. Cada $x \in \mathbb{R}^n$ pode ser representado de forma única como combinação linear dos vetores da base, ou seja,

$$x = \sum_{j=1}^{n} \xi_j e_j.$$

Note que como o espaço é de dimensão finita, $(\mathbb{R}^n)' = (\mathbb{R}^n)^*$. Seja $f \in (\mathbb{R}^n)'$, isto é, $f : \mathbb{R}^n \to \mathbb{R}$ linear e limitado. Ainda, para todo funcional f e $x \in \mathbb{R}^n$ tem-se

$$f(x) = \sum_{j=1}^{n} \xi_j \gamma_j, \text{ onde } \gamma_j = f(e_j), \tag{4.33}$$

e os números $\gamma_j = f(e_j)$ são unicamente determinados por f.

[3]Também chamado *dual*, *espaço adjunto* e espaço *conjugado*. Note que esta definição utiliza norma.

Defina a seguinte aplicação:

$$\begin{array}{rcl} T : (\mathbb{R}^n)' & \to & \mathbb{R}^n \\ f & \mapsto & (f(e_1), \dots, f(e_n)) := (\gamma_1^{(f)}, \dots, \gamma_n^{(f)}) \end{array}$$

Note $T(f) \in \mathbb{R}^n$ e $T(f)$ é única, pois o funcional f é unicamente determinado pelos n vetores da base, logo T está bem definida. Além disso T é linear, pois para quaisquer funcionais $f, g \in (\mathbb{R}^n)'$ e $\alpha \in \mathbb{R}$ tem-se

$$\begin{aligned} T(\alpha f + g) = & \; ((\alpha f + g)(e_1), \dots, (\alpha f + g)(e_n)) \\ = & \; (\alpha f(e_1) + g(e_i), \dots, \alpha f(e_n) + g(e_n)) \\ = & \; \alpha(f(e_1), \dots, f(e_n)) + (g(e_1), \dots, g(e_n)) \\ = & \; \alpha T(f) + T(g). \end{aligned}$$

Afirmação: T preserva norma, ou seja, $\|T(f)\| = \|f\|$ ou, de forma equivalente, $\|(\gamma_1^{(f)}, \dots, \gamma_n^{(f)})\| = \|f\|$. De fato, aplicando a Desigualdade de Cauchy-Schwarz em (4.33), resulta que

$$|f(x)| \leq \sum_{j=1}^{n} |\xi_j \gamma_j^{(f)}| \leq \left(\sum_{k=1}^{n} |\xi_j|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} |\gamma_j^{(f)}|^2 \right)^{\frac{1}{2}} = \|x\| \, \|T(f)\|. \tag{4.34}$$

Agora, considerando o supremo sobre todo x de norma 1, segue que

$$\|f\| \leq \|T(f)\|. \tag{4.35}$$

No entanto, para $x = (\gamma_1^{(f)}, \dots, \gamma_n^{(f)})$ e da linearidade de f, obtem-se de (4.34) que

$$\begin{aligned} |f(x)| = & \; |f(\gamma_1^{(f)}, \dots, \gamma_n^{(f)})| \\ = & \; |f(f(e_1), \dots, f(e_n))| \\ = & \; |f(f(e_1)e_1 + \ldots + f(e_n)e_n)| \\ = & \; |f(e_1)f(e_1) + \ldots + f(e_n)f(e_n)| \\ = & \; |f^2(e_1) + \ldots + f^2(e_n)| \\ = & \; f^2(e_1) + \ldots + f^2(e_n) \\ = & \; \sum_{j=1}^{n} |\gamma_j|^2 \\ = & \; \|T(f)\|^2 \end{aligned}$$

Por outro lado, como f é um funcional limitado, tem-se

$$|f(x)| \leq \|f\| \|x\|,$$

onde $\|f\| = \sup\limits_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} |f(x)|$. Logo, substituindo $|f(x)| = \|T(f)\|^2$ na desigualdade acima e como $x = (\gamma_1^{(f)}, \ldots, \gamma_n^{(f)}) = T(f)$, resulta que

$$\|T(f)\|^2 \leq \|f\| \|T(f)\|,$$

ou seja,

$$\|T(f)\| \leq \|f\|. \tag{4.36}$$

Portanto, por (4.35) e (4.36) segue a igualdade desejada,

$$\|f\| = \|T(f)\|.$$

Note que $\|T(f)\| = \|f\| = 0 \Rightarrow f = 0$, ou seja, T é injetora e portanto, $\dim \mathcal{N}(T) = 0$. Usando o *Teorema do Núcleo e da Imagem*, segue que T é sobrejetora. Logo, como T é bijetora e preserva norma, então $T : (\mathbb{R}^n)' \to \mathbb{R}^n$ é um isomorfismo e $(\mathbb{R}^n)' \cong \mathbb{R}^n$.

Os próximos exemplos serão com espaços de dimensão infinita, mais especificamente, espaços de sequências. Neste caso é natural pensar como seria uma base para estes espaços. Uma base para um espaço vetorial V de dimensão finita é um conjunto linearmente independente que gera todo o espaço. Porém, se o espaço for de dimensão infinita, qualquer base vai possuir infinitos elementos e, neste caso, estes infinitos vetores podem formar uma base chamada *base de Schauder*, como a seguir.

Definição 4.24. *Uma* ***base de Schauder*** *de um espaço normado E é um conjunto de vetores $\{e_1, e_2, \ldots\}$ satisfazendo que para todo $x \in E$ existe uma única sequência de escalares (α_n) tal que*

$$\|x - (\alpha_1 e_1 + \ldots + \alpha_n e_n)\| \to 0 \textit{ quando } n \to \infty.$$

A partir da definição acima, seguem os exemplos referentes ao dual do espaço ℓ^1 e do espaço ℓ^p.

Exemplo 4.40. O espaço dual de ℓ^1 é isomorfo ao espaço ℓ^∞, isto é, $(\ell^1)' \cong \ell^\infty$. Como visto na seção anterior, cada elemento do espaço ℓ^1 é uma sequência de números (reais ou complexos) cuja soma em módulo converge, ou seja,

$$\ell^1 = \left\{ x = (\xi_j)_{j \in \mathbb{N}} = (\xi_1, \xi_2, \ldots) \in \ell^1;\ \sum_{j=1}^{\infty} |\xi_j| < \infty \right\},$$

e a norma é definida como segue

$$\|x\| = \sum_{j=1}^{\infty} |\xi_j|.$$

Cada elemento do espaço ℓ^∞ é uma sequência limitada de números (reais ou complexos), isto é,

$$\ell^\infty = \left\{ \; y = (\eta_j)_{j\in\mathbb{N}} = (\eta_1, \eta_2, \ldots) \in \ell^\infty; \; |\eta_j| \leq c_y, \; \right\},$$

onde c_y é um número real que pode depender de y. A norma definida neste espaço é a do supremo,

$$\|y\| = \sup_{j\in\mathbb{N}} |\eta_j|.$$

Uma base de Schauder para ℓ^1 é dada por $e_1 = (1,0,0,\ldots)$, $e_2 = (0,1,0,\ldots)$, $e_3 = (0,0,1,\ldots)$, e assim sucessivamente. Cada $x \in \ell^1$ pode ser representado de forma única como combinação linear dos vetores da base, ou seja,

$$x = \sum_{j=1}^{\infty} \xi_j e_j.$$

Seja $f \in (\ell^1)'$, isto é, f linear e limitado. Ainda, para todo funcional f e $x \in \ell^1$ tem-se

$$f(x) = \sum_{j=1}^{\infty} \xi_j \gamma_j, \text{ onde } \gamma_j = f(e_j), \tag{4.37}$$

e os números $\gamma_j = f(e_j)$ são unicamente determinados por f.

Defina a seguinte aplicação:

$$\begin{array}{rcl} T : (\ell^1)' & \to & \ell^\infty \\ f & \mapsto & (\gamma_j) = (f(e_j)),\ j = 1, 2, \ldots \end{array}$$

Note que T está bem definida, pois como o funcional f é unicamente determinado pelos vetores da base, $T(f)$ é única e além disso, como f é linear e limitado e $\|e_j\| = 1$ então

$$|\gamma_j| = |f(e_j)| \leq \|f\|\|e_j\| = \|f\|,$$

assim, considerando o supremo da desigualdade acima obtem-se

$$\sup_{j\in\mathbb{N}} |\gamma_j| = \sup |f(e_1), \ldots, f(e_n), \ldots| \leq \|f\|. \tag{4.38}$$

ou seja,

$$\|T(f)\| \leq \|f\|. \tag{4.39}$$

Assim, $(\gamma_j) \in \ell^\infty$, isto é, $T(f) \in \ell^\infty$. T é linear, pois para quaiquer $f, g \in (\ell^1)'$ e $\alpha \in \mathbb{K}$ tem-se

$$\begin{aligned} T(\alpha f + g) = &\ ((\alpha f + g)(e_1), \ldots, (\alpha f + g)(e_n), \ldots) \\ = &\ (\alpha f(e_1) + g(e_i), \ldots, \alpha f(e_n) + g(e_n), \ldots) \\ = &\ \alpha(f(e_1), \ldots, f(e_n), \ldots) + (g(e_1), \ldots, g(e_n), \ldots) \\ = &\ \alpha T(f) + T(g). \end{aligned}$$

Por outro lado, para todo $y = (\eta_j) \in \ell^\infty$ pode-se obter um funcional linear limitado

$g \in \ell^1$ tal que $T(g) = y$. De fato, definindo $g \in \ell^1$ por

$$g(x) = \sum_{j=1}^{\infty} \xi_j \eta_j,$$

onde $x = (\xi_j) \in \ell^1$, então g é linear, pois

$$\begin{aligned} g(\alpha x + z) &= \sum_{j=1}^{\infty}(\alpha \xi_j + \mu_j)\eta_j \\ &= \sum_{j=1}^{\infty}(\alpha \xi_j \eta_j + \mu_j \eta_j) \\ &= \alpha \sum_{j=1}^{\infty} \xi_j \eta_j + \sum_{j=1}^{\infty} \mu_j \eta_j \\ &= \alpha g(x) + g(z), \end{aligned}$$

onde $z = (\mu_j) \in \ell^1$ e $\alpha \in \mathbb{K}$, e sua limitação segue de

$$|g(x)| \leq \sum_{j=1}^{\infty} |\xi_j \eta_j| \leq \sup_{j \in \mathbb{N}} |\eta_j| \sum_{j=1}^{\infty} |\xi_j| = \sup_{j \in \mathbb{N}} |\eta_j| \|x\| = \|y\| \, \|x\|.$$

Portanto, $g \in (\ell^1)'$, ou seja, T é sobrejetora. Agora, resta mostrar que T preserva norma. De (4.37), obtem-se

$$|f(x)| = \left| \sum_{j=1}^{\infty} \xi_j \gamma_j \right| \leq \sum_{j=1}^{\infty} |\xi_j||\gamma_j| \leq \sup_{j \in \mathbb{N}} |\gamma_j| \sum_{j=1}^{\infty} |\xi_j| = \|T(f)\| \, \|x\|.$$

Considerando o supremo sobre todo x de norma 1, a desigualdade acima resulta em

$$\|f\| \leq \|T(f)\|. \tag{4.40}$$

Logo, por (4.39) e (4.40) segue que

$$\|T(f)\| = \|f\|.$$

Agora, $\|T(f)\| = \|f\| = 0 \Rightarrow f = 0$, ou seja, T é injetora. Portanto, como T é bijetora e preserva norma, segue que $T : (\ell^1)' \to \ell^\infty$ é um isomorfismo e $(\ell^1)' \cong \ell^\infty$.

Exemplo 4.41. O espaço dual de ℓ^p é isomorfo ℓ^q, onde $1 < p < \infty$ e q é o conjugado de p, isto é, $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Todo elemento de ℓ^p é uma sequência de números (reais ou complexos) cuja soma em módulo converge e a norma é definida como segue

$$x = (\xi_j) = (\xi_1, \xi_2, \ldots) \in \ell^p \Rightarrow \sum_{j=1}^{\infty} |\xi_j|^p < \infty \quad \Leftrightarrow \quad \|x\|_p = \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{\frac{1}{p}}.$$

Como no exemplo anterior, uma *base Schauder* para o ℓ^p é dada por $e_1 = (1, 0, 0, \ldots)$, $e_2 = (0, 1, 0, \ldots)$, $e_3 = (0, 0, 1, \ldots)$, e assim por diante. Cada $x \in \ell^p$ pode ser representado de forma única como combinação linear dos vetores da base, ou seja,

$$x = \sum_{j=1}^{\infty} \xi_j e_j.$$

Seja $f \in (\ell^p)'$, onde $(\ell^p)'$ é o dual do espaço ℓ^p. Como f é linear e limitado, para todo $x \in \ell^p$ tem-se

$$f(x) = \sum_{j=1}^{\infty} \xi_j \gamma_j, \text{ onde } \gamma_j = f(e_j), \tag{4.41}$$

e os números $\gamma_j = f(e_j)$ são unicamente determinados por f.

Seja q o expoente conjugado de p e para cada $n \in \mathbb{N}$, considere $x_n = (\xi_j^{(n)})$ com

$$\xi_j^{(n)} = \begin{cases} \dfrac{|\gamma_j|^q}{\gamma_j}, & \text{se } j \leq n \text{ e } \gamma_j \neq 0, \\ 0, & \text{se } j > n \text{ ou } \gamma_j = 0. \end{cases} \tag{4.42}$$

Substituindo em (4.41), obtem-se

$$f(x_n) = \sum_{j=1}^{\infty} \xi_j^{(n)} \gamma_j = \sum_{j=1}^{n} \frac{|\gamma_j|^q}{\gamma_j} \gamma_j = \sum_{j=1}^{n} |\gamma_j|^q. \tag{4.43}$$

Considerando (4.42) e que $(q-1)p = q$, como f é limitado tem-se

$$\begin{aligned} |f(x_n)| \leq \|f\| \|x_n\| &= \|f\| \left(\sum_{j=1}^{n} |\xi_j^{(n)}|^p \right)^{\frac{1}{p}} \\ &= \|f\| \left(\sum_{j=1}^{n} \left| \frac{\gamma_j^q}{\gamma_j} \right|^p \right)^{\frac{1}{p}} \\ &= \|f\| \left(\sum_{j=1}^{n} |\gamma_j|^{(q-1)p} \right)^{\frac{1}{p}} \\ &= \|f\| \left(\sum_{j=1}^{n} |\gamma_j|^q \right)^{\frac{1}{p}} \end{aligned}$$

ou seja,

$$|f(x_n)| \leq \|f\| \left(\sum_{j=1}^{n} |\gamma_j|^q \right)^{\frac{1}{p}}. \tag{4.44}$$

Substituindo (4.43) em (4.44) resulta que

$$\sum_{j=1}^{n} |\gamma_j|^q \leq \|f\| \left(\sum_{j=1}^{n} |\gamma_j|^q \right)^{\frac{1}{p}}.$$

Como $\gamma_j \neq 0$ e $\dfrac{1}{q} = 1 - \dfrac{1}{p}$, então

$$\|f\| \geq \frac{\displaystyle\sum_{j=1}^{n} |\gamma_j|^q}{\left(\displaystyle\sum_{j=1}^{n} |\gamma_j|^q \right)^{\frac{1}{p}}} = \left(\sum_{j=1}^{n} |\gamma_j|^q \right)^{1-\frac{1}{p}} = \left(\sum_{j=1}^{n} |\gamma_j|^q \right)^{\frac{1}{q}}.$$

Fazendo $n \to \infty$ na desigualdade acima, obtem-se

$$\|f\| \geq \left(\sum_{j=1}^{\infty} |\gamma_j|^q \right)^{\frac{1}{q}}, \tag{4.45}$$

ou seja, $(\gamma_j) \in \ell^q$.

Defina a seguinte aplicação:

$$\begin{array}{rcl} T : (\ell^p)' & \to & \ell^q \\ f & \mapsto & (\gamma_j) = (f(e_j)),\ j = 1, 2, \ldots \end{array}$$

Note que T está bem definida, pois $T(f)$ é unica e $T(f) \in \ell^q$. Além disso, T é linear (como no exemplo anterior). Para ver que T é sobrejetora, é preciso obter para todo $y = (\eta_j) \in \ell^q$ um funcional linear limitado $g \in \ell^p$ tal que $T(g) = y$. De fato, definindo $g \in \ell^q$ por

$$g(x) = \sum_{j=1}^{\infty} \xi_j \eta_j,$$

onde $x = (\xi_j) \in \ell^1$, então g é linear (como no exemplo anterior). Além disso g é limitado, pois utilizando a Desigualdade de Hölder (página 34) segue que

$$|g(x)| \leq \sum_{j=1}^{\infty} |\xi_j \eta_j| \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^{\infty} |\eta_j|^q \right)^{\frac{1}{q}} = \|x\| \, \|y\|.$$

Considerando o supremo sobre todo x de norma 1, obtem-se $\|g\| \leq \|y\|$, ou seja, $g \in (\ell^p)'$. Resta mostrar que T preserva norma. De fato, novamente aplicando a Desigualdade de Hölder em (4.41), resulta que

$$|f(x)| \leq \sum_{j=1}^{\infty} |\xi_j \gamma_j| \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^{\infty} |\gamma_j|^q \right)^{\frac{1}{q}} = \|x\| \, \|T(f)\|.$$

Considerando o supremo sobre todo $x \in \ell^p$ de norma 1, obtem-se

$$\|f\| \leq \|T(f)\|. \tag{4.46}$$

Logo, por (4.45) e (4.46) tem-se

$$\|f\| = \|T(f)\|.$$

Agora, note que $\|T(f)\| = \|f\| = 0 \Rightarrow f = 0$, ou seja, T é injetora. Portanto, como T é bijetora e preserva norma, seque que $T : (\ell^p)' \to \ell^q$ é um isomorfismo e assim, $(\ell^p)' \cong \ell^q$.

A construção do espaço dual das funções contínuas será considerado no capítulo 6, pois requer em especial um teorema chamado *Teorema de Hahn-Banach*. O mesmo será abordado posteriormente.

5 Espaços de Hilbert

O espaço $\mathcal{H}$ de Hilbert é um espaço vetorial munido de produto interno cuja métrica determinada por esse produto interno o torne um espaço completo. Com a norma definida a partir do produto interno conclui-se que todo espaço de Hilbert é, em particular, um espaço de Banach.

Em um espaço vetorial V de dimensão finita com produto interno é possível definir o conceito de ortogonalidade entre vetores e partir disto obter algumas propriedades como, por exemplo, escrever o espaço V como soma direta de um subespaço de V com seu complemento ortogonal. A ideia é generalizar estes resultados para espaços de dimensão infinita, que neste caso, são os espaços de Hilbert. Conjuntos ortonormais estão fortemente presentes neste contexto.

Quando se fala em "base" para um espaço de Hilbert, algumas situações precisam ser consideradas, pois o conceito de base não se apresenta da mesma forma que nos espaços vetoriais (a menos que o espaço seja de dimensão finita).

Um teorema muito conhecido é o Teorema da Representação de Riesz que se refere à representação de funcionais lineares contínuos no espaço de Hilbert, o qual afirma que o espaço dual de um espaço $\mathcal{H}$ de Hilbert é isometricamente isomorfo a $\mathcal{H}$.

5.1 Definição e exemplos

Além da definição e exemplos de espaços de Hilbert, serão abordados: identidades de polarização, continuidade do produto interno, completamento, entre outros resultados.

Definição 5.1. *Seja E um espaço vetorial sobre um corpo $\mathbb{K}$. Um* ***produto interno*** *sobre E é uma aplicação $\langle .,. \rangle : E \times E \to \mathbb{K}$ que associa a cada par $(x, y) \in E \times E$ um escalar $\langle x, y\rangle \in \mathbb{K}$, de forma que para quaisquer vetores $x, y, z \in E$ e escalar $\alpha \in \mathbb{K}$ são verificadas as seguintes propriedades:*

PI1) Distributiva: $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$;

PI2) Homogeneidade: $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$;

PI3) Hermitiana: $\langle x, y\rangle = \overline{\langle y, x\rangle}$;

PI4) Positiva definida: $\langle x,x\rangle \geq 0$ *e* $\langle x,x\rangle = 0 \Leftrightarrow x = 0$.

Observação 5.1. Das 3 primeiras propriedades citadas acima, segue que

(a) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x,z\rangle + \beta\langle y,z\rangle$;

(b) $\langle x, \alpha y\rangle = \overline{\alpha}\langle x,y\rangle$;

(c) $\langle x, \alpha y + \beta z\rangle = \overline{\alpha}\langle x,y\rangle + \overline{\beta}\langle x,z\rangle$.

Exemplo 5.1. Sejam V e W espaços vetoriais sobre $\mathbb{K}$. Seja $T : V \to W$ uma transformação linear injetora. Considerando que a aplicação $f : W \times W \to \mathbb{K}$ define um produto interno sobre W, então a aplicação $g : V \times V \to \mathbb{K}$ tal que, para todo $x, y \in V$, $g(x,y) = f(T(x),T(y))$ define um produto interno sobre V.

De fato, g está bem definida e para todo $x, y, z \in V$ e $\alpha \in \mathbb{K}$ tem-se:

PI1)

$$\begin{aligned} g(x+y,z) = &\ f(T(x+y),T(z)) \\ = &\ f(T(x)+T(y),T(z)) \\ = &\ f(T(x),T(z)) + f(T(y),T(z)) \\ = &\ g(x,z) + g(y,z). \end{aligned}$$

PI2)

$$\begin{aligned} g(\alpha x,y) = &\ f(T(\alpha x),T(y)) \\ = &\ f(\alpha T(x),T(y)) \\ = &\ \alpha f(T(x),T(y)) \\ = &\ \alpha g(x,y). \end{aligned}$$

PI3)

$$\begin{aligned} g(x,y) = &\ f(T(x),T(y)) \\ = &\ \overline{f(T(y),T(x))} \\ = &\ \overline{g(y,x)}. \end{aligned}$$

PI4)

$$g(0,0) = f(T(0),T(0)) \overset{T}{=} f(0,0) \overset{f}{=} 0.$$

Como T é injetora, então para $x \in V$, $x \neq 0$ implica que $T(x) \neq 0$ e por f ser um produto interno, segue que

$$g(x,x) = f(T(x),T(x)) > 0.$$

Note que $g(x,y) = \langle x,y\rangle = \langle T(x),T(y)\rangle = f(T(x),T(y))$, para todo $x,y \in V$.

Definição 5.2. *Seja E um espaço com produto interno. Define-se a função $N : E \to \mathbb{R}$ por $N(x) = \sqrt{\langle x,x\rangle}$, para todo $x \in E$.*

Observação 5.2. Observe que $N(x) \geq 0$ para todo $x \in E$ e $N(x) = 0$ se, e somente se, $x = 0$. E além disso, $N(\alpha x) = |\alpha| N(x)$, para todo $\alpha \in \mathbb{K}$ e $x \in E$. Com efeito,

$$N(\alpha x) = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha\overline{\alpha}\langle x, x\rangle} = \sqrt{|\alpha|^2\langle x, x\rangle} = |\alpha|\sqrt{\langle x, x\rangle} = |\alpha| N(x).$$

Em vez de $N(x)$ será utilizada a notação $N(x) = \|x\|$.

Segue a Desigualdade de Cauchy-Schwarz para espaço com produto interno.

Proposição 5.1. *Seja E um espaço com produto interno. Então, para todo $x, y \in E$ tem-se*

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

Demonstração. Note que se $x = 0$ ou $y = 0$ segue o resultado. Suponha então $x \neq 0$ e $y \neq 0$. Assim, para todo $\alpha \in \mathbb{K}$ tem-se

$$\begin{aligned} 0 \leq \|x - \alpha y\|^2 = & \ \langle x - \alpha y, x - \alpha y\rangle \\ = & \ \langle x, x\rangle - \overline{\alpha}\langle x, y\rangle - \alpha\langle y, x\rangle + \alpha\overline{\alpha}\langle y, y\rangle \\ = & \ \langle x, x\rangle - \overline{\alpha}\langle x, y\rangle - \alpha[\langle y, x\rangle - \overline{\alpha}\langle y, y\rangle]. \end{aligned}$$

Como $y \neq 0$, considerando $\overline{\alpha} = \dfrac{\langle y, x\rangle}{\langle y, y\rangle}$ e sabendo que $\langle y, x\rangle = \overline{\langle x, y\rangle}$, a desigualdade acima resulta em

$$\begin{aligned} 0 \leq & \ \langle x, x\rangle - \frac{\langle y, x\rangle}{\langle y, y\rangle}\langle x, y\rangle - \alpha\left[\langle y, x\rangle - \frac{\langle y, x\rangle}{\langle y, y\rangle}\langle y, y\rangle\right] \\ = & \ \langle x, x\rangle - \frac{\langle y, x\rangle}{\langle y, y\rangle}\langle x, y\rangle \\ = & \ \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2}, \end{aligned}$$

ou seja,

$$0 \leq \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2}.$$

Multiplicando ambos os lados da desigualdade por $\|y\|^2$ tem-se

$$0 \leq \|x\|^2\|y\|^2 - |\langle x, y\rangle|^2,$$

isto é,

$$|\langle x, y\rangle|^2 \leq \|x\|^2\|y\|^2.$$

Extraindo a raiz quadrada em ambos os lados segue o resultado,

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

□

Corolário 5.1. *Um produto interno sobre E define uma norma em E, ou seja, para todo $x \in V$ define-se*

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

Demonstração. Basta verificar a desigualdade triangular, que é a condição $N3)$ de norma. Note que

$$\begin{aligned} \|x+y\|^2 = & \ \langle x+y, x+y\rangle \\ = & \ \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ = & \ \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2. \end{aligned}$$

Pela Desigualdade de Cauchy-Schwarz

$$|\langle x,y\rangle| = |\langle y,x\rangle| \leq \|x\|\|y\|,$$

logo,

$$\begin{aligned} \|x+y\|^2 \leq & \ \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ \leq & \ \|x\|^2 + 2\|x\|\|y\|| + \|y\|^2 \\ = & \ (\,\|x\| + \|y\|\,)^2. \end{aligned}$$

Extraindo a raiz quadrada em ambos os lados da desigualdade, resulta que

$$\|x+y\| \leq \|x\| + \|y\|.$$

□

Uma propriedade importante em um espaço com produto interno é o fato de que a norma definida neste espaço satisfaz a lei do paralelogramo, como mostra a proposição abaixo. Se a norma não satisfizer esta propriedade então diz-se que ela não provém de um produto interno.

Proposição 5.2. ***(Lei do Paralelogramo)*** *Sejam x e y vetores em um espaço com produto interno. Então*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Demonstração. Seja E um espaço vetorial com produto interno. Para todo $x, y \in V$ tem-se

$$\begin{aligned} \|x+y\|^2 + \|x-y\|^2 = & \ \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\ = & \ \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ + & \ \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle \\ = & \ 2\langle x,x\rangle + 2\langle y,y\rangle \\ = & \ 2(\|x\|^2 + \|y\|^2). \end{aligned}$$

□

Definição 5.3. *Seja E um espaço com produto interno. Dois vetores $x, y \in V$ são* ***ortogonais*** *se $\langle x, y\rangle = 0$ e denota-se por $x \perp y$. Dados quaisquer subconjuntos $A, B \subset V$, diz-se que $x \perp A$ se $x \perp a$ para todo $a \in A$, e $A \perp B$ se $a \perp b$ para todo $a \in A$ e todo $b \in B$.*

Exemplo 5.2. Seja E um espaço vetorial com produto interno sobre $\mathbb{R}$. Se para todo $x, y \in V$ ocorrer que $\|x+y\|=\|x-y\|$, então $x \perp y$. Com efeito,

$$\begin{aligned} \|x+y\| = \|x-y\| &\Rightarrow \langle x+y, x+y\rangle = \langle x-y, x-y\rangle \\ &\Rightarrow \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 = \|x\|^2 - 2\langle x, y\rangle + \|y\|^2 \\ &\Rightarrow \langle x, y\rangle = 0, \end{aligned}$$

ou seja, x e y são ortogonais.

Se em um espaço com produto interno dois vetores são ortogonais, então vale o **Teorema de Pitágoras**, como abaixo.

Teorema 5.1. *Seja E um espaço com produto interno. Considere $x, y \in E$ tais que $x \perp y$, então*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

Demonstração. Como $x \perp y$ então $\langle x, y\rangle = 0 = \langle y, x\rangle$. Logo,

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2.$$

□

Os espaços à serem considerados neste capítulo são definidos como segue.

Definição 5.4. *Um* ***espaço de Hilbert*** *é um espaço com produto interno que é completo (relativamente a norma proveniente do produto interno).*

Exemplo 5.3. O espaço $\mathbb{C}^n$ é um espaço de Hilbert com produto interno definido por

$$\langle x, y\rangle = \xi_1\overline{\eta_1} + \ldots + \xi_n\overline{\eta_n},$$

onde $x = (\xi_1, \ldots, \xi_n)$ e $y = (\eta_1, \ldots, \eta_n)$ são elementos de $\mathbb{C}^n$. A norma proveniente deste produto interno é definida por

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{\xi_1\overline{\xi_1} + \ldots + \xi_n\overline{\xi_n}} = \sqrt{|\xi_1|^2 + \ldots + |\xi_n|^2} = \left(\sum_{j=1}^{n} |\xi_j|^2\right)^{\frac{1}{2}},$$

e a métrica é dada por

$$d(x, y) = \|x-y\| = \sqrt{\langle x-y, x-y\rangle} = \left(\sum_{j=1}^{n} |\xi_j - \eta_j|^2\right)^{\frac{1}{2}}.$$

Observação 5.3. A partir da definição acima, note que todo espaço de Hilbert é um espaço de Banach cuja norma provém do produto interno, ou seja, o espaço de Hilbert é um espaço de Banach com a norma associada ao produto interno. Este fato será muito útil, pois os resultados apresentados para o espaço de Banach, no capítulo anterior, seguem de modo análogo para espaços de Hilbert, o que simplificará algumas demonstrações. Porém note que nem todo espaço de Banach é um espaço se Hilbert, como será exemplificado a seguir.

Exemplo 5.4. Foi demonstrado no capítulo anterior, que o espaço ℓ^p, $1 \leq p \leq \infty$, é Banach. Porém, para $p \neq 2$ este espaço não é um espaço com produto interno, logo não é de Hilbert. Note que considerando $p = 2$, a norma definida neste espaço coincide com a norma do espaço $\mathbb{C}^n$ ou $\mathbb{R}^n$, conforme as sequências forem complexas ou reais, ou seja, para todo $x, y \in \ell^2$, tem-se

$$\|x\| = \sqrt{\langle x, x\rangle} = \left(\sum_{j=1}^{\infty} |\xi_j|^2\right)^{\frac{1}{2}},$$

onde $x = (\xi_1, \xi_2, \ldots)$, $y = (\eta_1, \eta_2, \ldots)$ e o produto interno é definido por

$$\langle x, y\rangle = \sum_{j=1}^{\infty} \xi_j \overline{\eta_j}. \tag{5.1}$$

Note que este produto interno está bem definido. Com efeito, como $x, y \in \ell^2$ então

$$\sum_{j=1}^{\infty} |\xi_j|^2 < \infty \quad \text{e} \quad \sum_{j=1}^{\infty} |\eta_j|^2 < \infty,$$

logo, pela Desigualdade de Hölder (página 34)

$$\sum_{j=1}^{\infty} |\xi_j \overline{\eta_j}| \leq \left(\sum_{j=1}^{\infty} |\xi_j|^2\right)^{\frac{1}{2}} \left(\sum_{j=1}^{\infty} |\eta_j|^2\right)^{\frac{1}{2}} < \infty.$$

Não é difícil verificar que (5.1) satisfaz as propriedades de produto interno. Portanto, ℓ^2 é um espaço de Hilbert cujo produto interno é definido por (5.1).

Agora, para $p \neq 2$ a norma não pode ser obtida a partir do produto interno. Isto porque a norma não satisfaz a lei do paralelogramo (proposição 5.2). De fato, basta considerar as sequências $x = (1, 1, 0, 0, \ldots) \in \ell^p$ e $y = (1, -1, 0, 0 \ldots) \in \ell^p$. Note que

$$\|x\| = 2^{\frac{1}{p}} = \|y\|,$$

e além disso,

$$\|x + y\| = 2 = \|x - y\|,$$

ou seja,

$$\|x+y\|^2 + \|x-y\|^2 \neq 2(\|x\|^2 + \|y\|^2).$$

Portanto, o espaço ℓ^p, com $p \neq 2$ não é um espaço de Hilbert, pois a norma proveniente do produto interno não satisfaz a lei do paralelogramo.

Exemplo 5.5. No espaço $\mathcal{C}([a,b],\mathbb{R})$ a norma definida por

$$\|f\| = \sup_{t\in[a,b]} |f(t)|$$

não pode ser obtida a partir de um produto interno, uma vez que não satisfaz a lei do paralelogramo. Com efeito, considere $f(t)=1$ e $g(t) = \dfrac{t-a}{b-a}$. Note que

$$\|f\| = 1 = \|g\|.$$

Além disso,

$$f(t)+g(t) = 1 + \frac{t-a}{b-a} \quad \text{e} \quad f(t)-g(t) = 1 - \frac{t-a}{b-a}.$$

Assim,

$$\|f+g\| = \sup_{t\in[a,b]} |f(t)+g(t)| = 2 \quad \text{e} \quad \|f-g\| = \sup_{t\in[a,b]} |f(t)-g(t)| = 1.$$

Logo,

$$\|f+g\|^2 + \|f-g\|^2 = 5 \neq 4 = 2(\|f\|^2 + \|g\|^2).$$

Portanto, como a norma não satisfaz a lei do paralelogramo, ela não provém de um produto interno, implicando que este espaço não é de Hilbert.

No próximo exemplo a norma provém do produto interno, mas o espaço não é de Hilbert.

Exemplo 5.6. Considere o espaço vetorial $V = \mathcal{C}([a,b],\mathbb{R})$ munido do produto interno

$$\langle f, g\rangle = \int_a^b f(t)g(t)\,dt, \;\; f,g \in V. \tag{5.2}$$

Note que este produto interno está bem definido e satisfaz para toda $f,g,h \in V$ e $\alpha \in \mathbb{R}$:

PI1)

$$\begin{aligned}
\langle f+g, h\rangle = \int_a^b [f(t)+g(t)]h(t)\,dt &= \int_a^b [f(t)h(t)+g(t)h(t)]\,dt \\
&= \int_a^b f(t)h(t)\,dt + \int_a^b g(t)h(t)\,dt \\
&= \langle f,h\rangle + \langle g,h\rangle.
\end{aligned}$$

PI2)

$$\langle \alpha f, g\rangle = \int_a^b \alpha f(t)g(t)\,dt = \alpha \int_a^b f(t)g(t)\,dt = \alpha\langle f, g\rangle.$$

PI3)

$$\langle f, g\rangle = \int_a^b f(t)g(t)\,dt = \int_a^b g(t)f(t)\,dt = \langle g, f\rangle.$$

PI4)

$$\langle f, f\rangle = \int_a^b f(t)f(t)\,dt = \int_a^b [f(t)]^2\,dt.$$

e, neste caso, $\langle f, f\rangle = 0 \Leftrightarrow f = 0$, pois se $\langle f, f\rangle = 0$ então $f = 0$. Por outro lado, se $f \neq 0$ então existe $x_0 \in [a, b]$ tal que $f(x_0) \neq 0$. Como f é contínua em $[a, b]$, deve existir uma vizinhança de x_0, $V_{x_0} \subset [a, b]$, tal que $f(x_0) \neq 0$ para todo $x \in [a, b]$ (Teorema da Conservação do Sinal). Assim,

$$\int_a^b [f(t)]^2\,dt \geq \int_{V_{x_0}} [f(t)]^2\,dt > 0,$$

e, portanto, $\langle f, f\rangle > 0$ se $f \neq 0$.

Considere agora a norma proveniente do produto interno

$$\|f\| = \langle f, f\rangle^{\frac{1}{2}} = \left(\int_a^b [f(t)]^2\,dt\right)^{\frac{1}{2}}. \tag{5.3}$$

Agora, embora a norma provenha do produto interno, o espaço não é completo. Com efeito, considere $[a, b] = [-1, 1]$ e

$$f_m(t) = \begin{cases} -1, & \text{se } -1 \leq t \leq -\dfrac{1}{m}, \\ mt, & \text{se } -\dfrac{1}{m} \leq t \leq \dfrac{1}{m}, \\ 1, & \text{se } \quad \dfrac{1}{m} \leq t \leq 1. \end{cases}$$

Note que quando m varia, a sequência (f_m) é de Cauchy em $\mathcal{C}([-1, 1], \mathbb{R})$. De fato, observe que

$$\|f_m - f_n\|^2 = \int_{-1}^1 [f_m(t) - f_n(t)]^2\,dt = \int_{-\frac{1}{m}}^{\frac{1}{m}} (m-n)^2 t^2\,dt = \frac{2}{3}\frac{(m-n)^2}{m^3},$$

onde $m, n \in \mathbb{N}$. Supondo $m < n$ (análogo para $m > n$), se $n - m \leq m$ então

$$\frac{2}{3}\frac{(m-n)^2}{m^3} \leq \frac{2}{3}\frac{m^2}{m^3} = \frac{2}{3}\frac{1}{m} \to 0, \text{ quando } m \to \infty.$$

Caso $n - m > m$ então $n > 2m$, por exemplo, $n = 2m + k$, $k \in \mathbb{N}$. Assim, quando $m \to \infty$ tem-se $n \to \infty$ e, neste caso,

$$\frac{2}{3}\frac{(m-n)^2}{m^3} \leq \frac{2}{3}\frac{(-m-k)^2}{m^3} = \frac{2}{3}\frac{(m+k)^2}{m^3} = \frac{2}{3}\left(\frac{m+k}{m}\right)^2\frac{1}{m} = \frac{2}{3}\left(1+\frac{k}{m}\right)^2\frac{1}{m} \to 0,$$

quando $m \to \infty$. Assim, em ambos os casos tem-se

$$\|f_m - f_n\| = \left(\int_{-1}^{1}[f_m(t) - f_n(t)]^2\,dt\right)^{\frac{1}{2}} \to 0, \text{ quando } m, n \to \infty.$$

Contudo, (f_m) não converge para uma função em $\mathcal{C}([-1,1],\mathbb{R})$. Denote

$$\psi(t) = \begin{cases} -1, & \text{se } t < 0, \\ 1, & \text{se } t \geq 0, \end{cases}$$

que não é uma função contínua. Dada qualquer função f em $\mathcal{C}([-1,1],\mathbb{R})$, segue a partir da desigualdade triangular que

$$\begin{aligned} \|f(t) - \psi(t)\| = {} & \|f(t) - f_m(t) + f_m(t) - \psi(t)\| \\ \leq {} & \|f(t) - f_m(t)\| + \|f_m(t) - \psi(t)\|, \end{aligned}$$

ou seja,

$$\|f(t) - \psi(t)\| \leq \|f(t) - f_m(t)\| + \|f_m(t) - \psi(t)\|.$$

Logo,

$$\left(\int_{-1}^{1}[f(t) - \psi(t)]^2\,dt\right)^{\frac{1}{2}} \leq \left(\int_{-1}^{1}[f(t) - f_m(t)]^2\,dt\right)^{\frac{1}{2}} + \left(\int_{-1}^{1}[f_m(t) - \psi(t)]^2\,dt\right)^{\frac{1}{2}}.$$

Como f é uma função contínua, a integral a esquerda da desigualdade é diferente de zero. Além disso, é fácil ver que

$$\lim_{m\to\infty}\int_{-1}^{1}[f_m(t) - \psi(t)]^2\,dt = 0.$$

Portanto,

$$\int_{-1}^{1}[f(t) - f_m(t)]^2\,dt$$

não converge para zero quando $m \to \infty$. Resulta disto que o espaço $\mathcal{C}([a,b],\mathbb{R})$, com a norma em (5.3), não é completo.

É conhecido que um produto interno corresponde a uma norma, pelo corolário 5.1 (página 112). Agora, é possível a partir de uma norma descobrir qual o produto interno associado a esta norma, desde que a norma satisfaça a lei do paralelogramo (proposição

5.2). Neste caso, a proposição abaixo mostra como deve ser definido o produto interno a partir da norma.

Proposição 5.3. *(**Identidade de Polarização**) Em todo espaço com produto interno tem-se*

i. Se o espaço for real:

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2). \tag{5.4}$$

ii. Se o espaço for complexo:

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2), \tag{5.5}$$

onde

$$Re\,\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) \text{ parte real}$$

e

$$Im\,\langle x, y\rangle = \frac{1}{4}(\|x+iy\|^2 - \|x-iy\|^2) \text{ parte imaginária.}$$

Na referência [7] (em apêndice) é demonstrado que sendo E um espaço normado cuja norma satisfaz a lei do paralelogramo, então a sua norma provém de um produto interno, onde o mesmo é definido como na proposição acima.

O lema a seguir se refere à continuidade do produto interno que, assim como a Desigualdade de Cauchy-Schwarz, é frequentemente utilizada.

Lema 5.1. *Sejam E um espaço com produto interno e (x_n) e (y_n) duas sequências em E. Se $x_n \to x \in E$ e $y_n \to y \in E$, então $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Demonstração. Note que

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| = |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle|.$$

Utilizando a desigualdade triangular, segue que

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| \leq |\langle x_n, y_n\rangle - \langle x_n, y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle|.$$

Aplicando a Desigualdade de Cauchy-Schwarz resulta que

$$0 \leq |\langle x_n, y_n\rangle - \langle x, y\rangle| \leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\|.$$

Como $x_n \to x$ e $y_n \to y$, então $(x_n - x) \to 0$ e $(y_n - y) \to 0$ quando $n \to \infty$, logo pela desigualdade acima e pelo Teorema do Confronto, tem-se

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| \to 0$$

quando $n \to \infty$, ou seja, $\langle x_n, y_n \rangle \to \langle x, y \rangle$. Portanto, pelo teorema 3.3, o produto interno é contínuo. $\square$

Uma importante aplicação deste lema é o fato de que todo espaço com produto interno pode ser completado. O completamento é um espaço de Hilbert e este é único exceto por isomorfismo. Segue então a definição de isomorfismo para espaço com produto interno.

Definição 5.5. *Um* ***isomorfismo*** *T de um espaço com produto interno E em um espaço com produto interno $\tilde{E}$ sobre o mesmo corpo é uma transformação linear bijetora $T : E \to \tilde{E}$ que preserva produto interno, ou seja, para todo $x, y \in E$ tem-se*

$$\langle T(x), T(y) \rangle = \langle x, y \rangle.$$

Diz-se que E e $\tilde{E}$ são isomorfos e denota-se por $E \cong \tilde{E}$.

Observação 5.4. Note que T é uma isometria de E em $\tilde{E}$, pois a distância de E a $\tilde{E}$ é determinada pela norma definida pelo produto interno entre E e $\tilde{E}$.

Teorema 5.2. *(**Completamento**) Para todo espaço E com produto interno existe um espaço de Hilbert $\mathcal{H}$ e um isomorfismo T de E em um subespaço denso $W \subset H$. O espaço $\mathcal{H}$ é único exceto por isomorfismo.*

Na referência [9] é demonstrado que todo espaço com produto interno possui um completamento $\mathcal{H}$, nas condições citadas acima.

Definição 5.6. *Um* ***subespaço*** *F de um espaço com produto interno E é definido como sendo um subespaço vetorial de E considerando o produto interno de E restrito a $F \times F$.*

Semelhantemente, um ***subespaço*** *F de um espaço $\mathcal{H}$ de Hilbert é definido como sendo um subespaço de $\mathcal{H}$, considerado como um espaço com produto interno.*

Teorema 5.3. *Seja F um subespaço de um espaço $\mathcal{H}$ de Hilbert. Então:*

i. F é completo se, e somente se, F é fechado.

ii. Se F é de dimensão finita, então F é completo.

Demonstração. Considerando que todo espaço de Hilbert é um espaço de Banach, cuja norma provém do produto interno, então a demonstração do item *i.* segue do teorema 4.6 (página 80) e a do item *ii.* segue do teorema 4.5 (página 74). $\square$

5.2 Algumas propriedades referentes ao complemento ortogonal

Considere os subconjuntos X e Y de $\mathbb{R}^2$ definidos por $X = \{(t, 1/t),\, t \in \mathbb{R}^+\}$ e $Y = \{(t, 0),\, t \in \mathbb{R}^+\}$, como na figura 5.1.

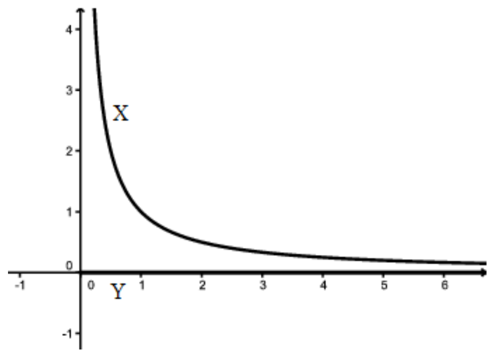

Figura 5.1: *Distância de X à Y*

Note que $X \cap Y = \emptyset$ e no entanto a distância entre os conjuntos X e Y é zero, isto é, $d(X, Y) = 0$. Agora, considere um subespaço W de um espaço vetorial E com produto interno e um vetor $v_0 \in E$ de forma que $v_0 \notin W$. Neste caso $d(v_0, W) > 0$ e além disso é possível determinar um vetor $w_0 \in W$ tal que

$$d(v_0, w) = \inf\{d(v_0, w),\, w \in W\} = d(v_0, w_0).$$

Para determinar $w_0 \in W$, que é o vetor que minimiza a distância entre $v_0 \in E$ e W, precisa-se da definição de complemento ortogonal. Mas, inicialmente, fazem-se necessárias as seguintes definições.

Definição 5.7. *Sejam V um espaço vetorial sobre o corpo $\mathbb{K}$ e $x, y \in V$. O* ***segmento*** *que une x e y em V é definido como sendo o conjunto de todos $z \in V$ da forma*

$$z = \alpha x + (1 - \alpha) y,$$

onde $\alpha \in \mathbb{R}$ e $0 \leq \alpha \leq 1$.

Definição 5.8. *Seja V um espaço vetorial sobre o corpo $\mathbb{K}$. O subconjunto $W \subset V$ é dito ser* ***convexo*** *se para todo $x, y \in W$ o segmento que une x e y está contido em W.*

Considere E um espaço com produto interno e F subconjunto de E. O teorema seguinte se refere à melhor aproximação de $x \in E$ por vetores de F, na forma como segue.

Teorema 5.4. *Seja E um espaço com produto interno e considere $F \neq \emptyset$ um subconjunto convexo que é completo (na norma induzida pelo produto interno). Então para cada $x \in E$ existe um único $y_0 \in F$ tal que*

$$d(x,F) = \inf_{y \in F} \|x - y\| = \|x - y_0\|. \tag{5.6}$$

Demonstração. Note que é preciso mostrar a existência e a unicidade. Para simplificar a notação será utilizado $\delta = d(x,F)$.

i. Existência: Seja $\delta = \inf_{y \in F} \|x - y\|$. Pela definição de ínfimo, dado $\epsilon = \dfrac{1}{n}$ existe uma sequência $(y_n) \in F$ tal que

$$\delta \leq \|x - y_n\| < \delta + \frac{1}{n}.$$

Fazendo $n \to \infty$, segue que

$$\|x - y_n\| = \delta_n \to \delta. \tag{5.7}$$

Esta sequência (y_n) é de Cauchy. Com efeito, escrevendo $y_n - x = v_n$, tem-se $\|v_n\| = \delta_n$ e ainda

$$\begin{aligned} \|v_n + v_m\| = \|(y_n - x) + (y_m - x)\| = \|y_n + y_m - 2x\| &= \left\| 2\left(\frac{y_n + y_m}{2} - x\right)\right\| \\ &= 2\left\|\frac{y_n + y_m}{2} - x\right\| \\ &\geq 2\delta, \end{aligned} \tag{5.8}$$

pois $(y_n + y_m)/2$, já que F é convexo. Além disso, $y_n - x = v_n$ implica $y_n - y_m = v_n - v_m$. Assim, como $\|v_n\| = \delta_n$, de (5.8) e pela lei do paralelogramo (proposição 5.2) resulta que

$$\|y_n - y_m\|^2 = \|v_n - v_m\|^2 = -\|v_n + v_m\|^2 + 2(\|v_n\|^2 + \|v_m\|^2) \leq -(2\delta)^2 + 2(\delta_n^2 + \delta_m^2),$$

ou seja,

$$\|y_n - y_m\|^2 \leq -(2\delta)^2 + 2(\delta_n^2 + \delta_m^2) \to 0$$

quando $n, m \to \infty$, por (5.7). Isto implica que (y_n) é uma sequência de Cauchy. Como F é completo, existe $y_0 \in F$ tal que $y_n \to y_0$. Pela definição de δ, tem-se que $\|x - y_0\| \geq \delta$. Usando este fato, segue que

$$\delta \leq \|x - y_0\| = \|x - y_n + y_n - y_0\| \leq \|x - y_n\| + \|y_n - y_0\| = \delta_n + \|y_n - y_0\|.$$

Como $\delta_n \to \delta$ e $\|y_n - y_0\| \to 0$ quando $n \to \infty$, a desigualdade acima resulta que

$$\delta \leq \|x - y_0\| \leq \delta,$$

ou seja,

$$\delta = \|x - y_0\|.$$

Portanto, $\delta = \inf_{y \in F} \|x - y\| = \|x - y_0\|$.

ii. Unicidade: Para provar a unicidade suponha que existam $y, y_0 \in F$ tais que

$$\|x - y_0\| = \delta \quad \text{e} \quad \|x - y\| = \delta.$$

Pela lei do paralelogramo (página 112),

$$\begin{aligned}
\|y - y_0\|^2 = \|y - x + x - y_0\|^2 = &\ \|(y - x) - (y_0 - x)\|^2 \\
= &\ 2\|y - x\|^2 + 2\|y_0 - x\|^2 - \|(y - x) + (y_0 - x)\|^2 \\
= &\ 2\delta^2 + 2\delta^2 - \left\|2\left(\frac{y + y_0}{2} - x\right)\right\|^2 \\
= &\ 4\delta^2 - 2^2\left\|\frac{y + y_0}{2} - x\right\|^2 \\
= &\ 4\delta^2 - 4\left\|\frac{y + y_0}{2} - x\right\|^2.
\end{aligned}$$

Como F é convexo, $(y + y_0)/2 \in F$, e

$$\left\|\frac{y + y_0}{2} - x\right\| \geq \delta.$$

Então,

$$\|y - y_0\|^2 = 4\delta^2 - 4\left\|\frac{y + y_0}{2} - x\right\|^2 \leq 4\delta^2 - 4\delta^2 = 0,$$

ou seja,

$$\|y - y_0\| = 0,$$

e portanto, $y = y_0$. Isto mostra que existe único y_0 satisfazendo (5.6). □

Lema 5.2. *Com as hipóteses do teorema anterior, seja E um espaço com produto interno e considere $Y \neq \emptyset$ um subespaço completo (na norma induzida pelo produto interno). Então $z = x - y_0$ é ortogonal à Y.*

Demonstração. Suponha que $z \perp Y$ seja falso. Neste caso, existe $y_1 \in Y$ tal que

$$\langle z, y_1 \rangle = \beta \neq 0.$$

Note que $y_1 \neq 0$, pois caso contrário $\langle z, y_1 \rangle = 0$. Além disso para qualquer escalar

$\alpha \in \mathbb{K}$ tem-se

$$\begin{aligned}\|z-\alpha y_1\|^2 = &\ \langle z-\alpha y_1, z-\alpha y_1\rangle\\ = &\ \langle z,z\rangle - \overline{\alpha}\langle z,y_1\rangle - \alpha\langle y_1,z\rangle + \alpha\overline{\alpha}\langle y_1,y_1\rangle\\ = &\ \langle z,z\rangle - \overline{\alpha}\langle z,y_1\rangle - \alpha[\langle y_1,z\rangle - \overline{\alpha}\langle y_1,y_1\rangle]\\ = &\ \langle z,z\rangle - \overline{\alpha}\beta - \alpha[\overline{\beta} - \overline{\alpha}\langle y_1,y_1\rangle].\end{aligned}$$

Como $y_1 \neq 0$, escrevendo $\overline{\alpha} = \dfrac{\overline{\beta}}{\langle y_1,y_1\rangle}$, resulta que

$$\|z-\alpha y_1\|^2 = \langle z,z\rangle - \frac{|\beta|^2}{\langle y_1,y_1\rangle}. \tag{5.9}$$

Pelo teorema 5.4 (página 121), $\delta = \|x-y_0\| = \|z\|$. Assim, de (5.9) segue que

$$\|z-\alpha y_1\|^2 = \|z\|^2 - \frac{|\beta|^2}{\|y_1\|^2} = \delta^2 - \frac{|\beta|^2}{\|y_1\|^2} < \delta^2,$$

ou seja,

$$\|z-\alpha y_1\|^2 < \delta^2 \Rightarrow \|z-\alpha y_1\| < \delta. \tag{5.10}$$

Mas isto não é possível, pois

$$z-\alpha y_1 = x-y_2 \quad \text{onde} \quad y_2 = y+\alpha y_1 \in Y,$$

e como $\delta = \inf\limits_{y\in F}\|x-y\| = \|x-y_0\| = \|z\|$ segue que $\|z-\alpha y_1\| \geq \delta$. Portanto, (5.10) é uma contradição, isto implica que $\langle z,y_1\rangle = 0$ para todo $y_1 \in Y$, ou seja, $z \perp Y$. □

Definição 5.9. *Sejam E um espaço como produto interno e Y subespaço de E. O* ***complemento ortogonal*** *de Y é definido por*

$$Y^{\perp} = \{x \in E;\ x \perp y,\ \forall y \in Y\} = \{x \in E;\ x \perp Y\}.$$

Em um espaço vetorial V de dimensão finita, dado um subespaço qualquer W de V sempre é possível escrever $V = W \oplus W^{\perp}$ e quando se trata de um espaço $\mathcal{H}$ de Hilbert, o subconjunto de maior interesse é o complemento ortogonal de um subespaço fechado $Y \subset \mathcal{H}$. Isto fornece um resultado muito importante chamado Teorema da Projeção, como segue.

Teorema 5.5. *Sejam $\mathcal{H}$ um espaço de Hilbert e $Y \subset \mathcal{H}$ um subespaço fechado qualquer. Então*

$$\mathcal{H} = Y \oplus Z \quad \textit{onde} \quad Z = Y^{\perp}.$$

Demonstração. i. Existência: Como $\mathcal{H}$ é completo e Y é fechado segue que Y é completo, conforme teorema 5.3 (página 119). Como Y é subespaço então ele é convexo.

Pelo teorema 5.4 (página 121) e pelo lema 5.2 (página 122) implica que para cada $x \in \mathcal{H}$ existe um $y_0 \in Y$ tal que

$$x = y_0 + z \quad \text{onde} \quad z \in Z = Y^\perp. \tag{5.11}$$

o que mostra que $\mathcal{H} = Y + Z$.

ii. Unicidade: Para provar a unicidade, basta mostrar que para cada $x \in \mathcal{H}$ a decomposição em (5.11) é única. Assim, suponha que existam $y_0, y_1 \in Y$ e $z, z_1 \in Z$ tais que

$$x = y_0 + z \quad \text{e} \quad x = y_1 + z_1.$$

Logo,

$$y_0 + z = y_1 + z_1,$$

ou seja,

$$y_0 - y_1 = z_1 - z.$$

Agora, $y_0 - y_1 \in Y$ enquanto $z_1 - z \in Y^\perp = Z$. Logo,

$$y_0 - y_1 = z_1 - z \in Y \cap Y^\perp.$$

Por outro lado, se existe $\kappa \in \mathcal{H}$ tal que $\kappa \in Y \cap Y^\perp$, resulta que $\langle \kappa, \kappa \rangle = 0$ e, assim, $\kappa = 0$. Logo,

$$y_0 - y_1 = 0 \quad \text{e} \quad z_1 - z = 0,$$

isto é,

$$y_0 = y_1 \quad \text{e} \quad z = z_1.$$

Portanto, por *i.* e *ii.* segue que $\mathcal{H} = Y \oplus Z$. □

Definição 5.10. *Sejam $\mathcal{H}$ um espaço de Hilbert e $Y \subset \mathcal{H}$ um subespaço fechado qualquer. Na equação* (5.11) *o vetor $y_0 \in Y$ é chamado de* ***projeção ortogonal*** *de x em Y. Esta projeção define uma aplicação $P : \mathcal{H} \to Y$ tal que $y_0 = P(x)$ para todo $x \in \mathcal{H}$. Analogamente o vetor $z \in Y^\perp$ é a projeção ortogonal de x em $Y^\perp$ e define uma aplicação $Q : \mathcal{H} \to Y^\perp$ tal que $z = Q(x)$ para todo $x \in \mathcal{H}$.*

Teorema 5.6. *Com as mesmas hipóteses do teorema 5.5 tem-se*

(a) As projeções P e Q são lineares contínuas e satisfazem $P^2 = P$ e $Q^2 = Q$.

(b) O complemento ortogonal $Y^\perp$ de Y é o espaço nulo $\mathcal{N}(P)$ da projeção ortogonal $P : \mathcal{H} \to Y$.

Demonstração. (*a*) A demonstração será feita para a projeção P, pois para Q segue de modo análogo. Pelo teorema 5.5, para cada $x \in \mathcal{H}$ exitem únicos $y \in Y$ e $z \in Y^\perp$ tais que

$$x = y + z. \tag{5.12}$$

Pela definição anterior $P(x) = y$ e $Q(x) = z$, logo a decomposição (5.12) pode ser reescrita na forma

$$x = P(x) + Q(x). \tag{5.13}$$

Para ver que a projeção P define uma transformação linear, considere as seguintes decomposições

$$x = y + z \quad \text{e} \quad w = u + v,$$

onde $x, w \in \mathcal{H}$, $y, u \in Y$ e $z, v \in Y^{\perp}$. Pela definição de P tem-se $P(x) = y$ e $P(w) = u$, então para todo $\alpha \in \mathbb{K}$ tem-se

$$\begin{aligned} P(\alpha x + w) = & \ P(\alpha(y+z) + (u+v)) \\ = & \ P(\alpha y + \alpha z + u + v) \\ = & \ P((\alpha y + u) + (\alpha z + v)) \\ = & \ \alpha y + u \\ = & \ \alpha P(x) + P(w). \end{aligned}$$

Agora, para mostrar que P é contínua, basta mostrar que P é limitada, conforme o teorema 4.11 (página 89). Com efeito, como $\langle y, z\rangle = 0$ para todo $y \in Y$ e $z \in Y^{\perp}$, de (5.13) note que

$$\begin{aligned} \|x\|^2 = \langle x, x\rangle = & \ \langle P(x) + Q(x), P(x) + Q(x)\rangle \\ = & \ \langle P(x), P(x)\rangle + \langle P(x), Q(x)\rangle + \langle Q(x), P(x)\rangle + \langle Q(x), Q(x)\rangle \\ = & \ \langle P(x), P(x)\rangle + \langle Q(x), Q(x)\rangle \\ = & \ \|P(x)\|^2 + \|Q(x)\|^2 \\ \geq & \ \|P(x)\|^2, \end{aligned}$$

ou seja,

$$\|P(x)\|^2 \leq \|x\|^2,$$

ou ainda,

$$\|P(x)\| \leq \|x\|, \ \forall\, x \in \mathcal{H}.$$

Considerando o supremo sobre todo $x \in \mathcal{H}$ de norma 1, segue da desigualdade acima que

$$\|P(x)\| \leq 1, \ \forall\, x \in \mathcal{H}, \ \|x\| = 1.$$

Logo, P é uma transformação linear limitada e, portanto, contínua.

Resta mostrar que $P^2 = P$. De fato, como $P(x) = y \in Y$ então

$$P(P(x)) = P(y) = P(y + 0) = y = P(x).$$

(b) Deve-se provar que $Y^{\perp} = \mathcal{N}(P) = \{x \in \mathcal{H};\ P(x) = y = 0\}$, ou seja, $\mathcal{N}(P) \subset Y^{\perp}$ e que $Y^{\perp} \subset \mathcal{N}(P)$.

i. $\mathcal{N}(P) \subset Y^{\perp}$. Com efeito, seja $x_0 \in \mathcal{N}(P)$, então $P(x_0) = 0$. Como $\mathcal{H} = Y \oplus Y^{\perp}$

segue que $x_0 = y_0 + z_0$. Pela definição de P, $0 = P(x_0) = y_0$, ou seja, $y_0 = 0$ o que resulta $x_0 = z_0 \in Y^{\perp}$.

ii. $Y^{\perp} \subset \mathcal{N}(P)$. De fato, se $z_0 \in Y^{\perp}$ então

$$P(z_0) = P(0 + z_0) = 0,$$

ou seja, $z_0 \in \mathcal{N}(P)$. □

Observação 5.5. (1) Sabe-se que o complemento ortogonal de um conjunto $N \neq \emptyset$, denotado por $N^{\perp}$, num espaço com produto interno E é o conjunto

$$N^{\perp} = \{x \in E;\, x \perp N\} = \{x \in E;\, \langle x, n\rangle = 0,\, \forall n \in N\}.$$

Mesmo que N não seja subespaço de E, $N^{\perp}$ é um subespaço vetorial de E, pois dados $x, y \in N^{\perp}$, para quaiquer $n \in N$ e escalares $\alpha, \beta \in \mathbb{K}$ tem-se

$$\langle \alpha x + \beta y, n\rangle = \alpha\langle x, n\rangle + \beta\langle y, n\rangle = 0 + 0 = 0,$$

ou seja, $\alpha x + \beta y \in N^{\perp}$. Além disso $N^{\perp}$ é um conjunto fechado. Basta mostrar que qualquer sequência de elementos de $N^{\perp}$ converge para um elemento de $N^{\perp}$. Com efeito, seja $x_m \in N^{\perp}$ uma sequência convergente, isto é, $x_m \to x \in E$. Como $x_m \in N^{\perp}$ então para todo $n \in N$ tem-se

$$\langle x, n\rangle = \langle x - x_m + x_m, n\rangle = \langle x - x_m, n\rangle + \langle x_m, n\rangle.$$

Como $\langle x_m, n\rangle = 0$ para todo $m \in \mathbb{N}$, então

$$\langle x, n\rangle = \langle x - x_m, n\rangle.$$

Agora, pela Desigualdade de Cauchy-Schwarz,

$$0 \leq |\langle x, n\rangle| = |\langle x - x_m, n\rangle| \leq \|x - x_m\|\|n\|.$$

Fazendo $m \to \infty$, $\|x - x_m\| \to 0$, logo,

$$0 \leq |\langle x, n\rangle| \leq 0,$$

ou seja,

$$\langle x, n\rangle = 0, \text{ para todo } n \in N,$$

o que implica que $x \in N^{\perp}$, resultando que $N^{\perp}$ é fechado.

(2) Outro fato a ser observado é que, em geral,

$$N \subset (N^{\perp})^{\perp} = N^{\perp\perp},$$

pois,

$$x \in N \quad \Rightarrow \quad x \perp N^{\perp} \quad \Rightarrow \quad x \in N^{\perp\perp}.$$

Porém, quando o subespaço for fechado em um espaço com produto interno que é completo, segue que $N = N^{\perp\perp}$, como mostra o lema seguinte.

Lema 5.3. *Se Y é um subespaço fechado de um espaço $\mathcal{H}$ de Hilbert, então $Y = Y^{\perp\perp}$.*

Demonstração. Como $Y \subset Y^{\perp\perp}$, basta mostrar que $Y \supset Y^{\perp\perp}$. Seja $x \in Y^{\perp\perp}$. Pelo teorema 5.5, para cada $x \in \mathcal{H}$ existe $y_0 \in Y \subset Y^{\perp\perp}$ tal que

$$x = y_0 + z \quad \text{onde} \quad z \in Y^{\perp}. \tag{5.14}$$

Como $Y^{\perp\perp}$ é um espaço vetorial e $x \in Y^{\perp\perp}$, por hipótese, então

$$z = x - y_0 \in Y^{\perp\perp},$$

e pelo lema 5.2,

$$z \perp Y^{\perp}.$$

Agora, como $z \in Y^{\perp}$ e $z \perp Y^{\perp}$, resulta que $z \perp z$, ou seja, $\langle z, z\rangle = 0$, o que implica $z = 0$. De (5.14), segue que $x = y_0$, isto é, $x \in Y$. Portanto, como $Y \subset Y^{\perp\perp}$ e $Y \supset Y^{\perp\perp}$, então

$$Y = Y^{\perp\perp}.$$

□

Lema 5.4. *Sejam $\mathcal{H}$ um espaço de Hilbert e $N \neq \emptyset$ um subconjunto qualquer de $\mathcal{H}$. O subespaço vetorial gerado por N é denso em $\mathcal{H}$ se, e somente se, $N^{\perp} = \{0\}$.*

Demonstração. i. Suponha que o subespaço gerado por N é denso em $\mathcal{H}$, ou seja, $\overline{[N]} = \mathcal{H}$ e seja $x \in N^{\perp}$. É preciso mostrar que $x = 0$. De fato, se $x \in N^{\perp}$ então $x \in \mathcal{H} = \overline{[N]}$. Logo, existe uma sequência (x_n) em $[N]$ tal que $x_n \to x$. Agora, como $x \in N^{\perp}$ e $N^{\perp} \perp [N]$, segue que $\langle x_n, x\rangle = 0$. Pela continuidade do produto interno (lema 5.1 página 118), resulta que $\langle x_n, x\rangle \to \langle x, x\rangle$ quando $n \to \infty$. Pela unicidade do limite, seque que $\langle x, x\rangle = \|x\|^2 = 0$, ou seja, $x = 0$. Portanto, $N^{\perp} = \{0\}$.

ii. Reciprocamente, suponha que $N^{\perp} = \{0\}$. Se $x \perp [N]$, então $x \perp N$, de modo que $x \in N^{\perp}$ e $x = 0$. Consequentemente, $[N]^{\perp} = \{0\}$. No teorema 5.5, fazendo $Y = \overline{[N]}$ tem-se

$$\mathcal{H} = \overline{[N]} \oplus \overline{[N]}^{\perp}.$$

Como $\overline{[N]}^{\perp} = 0$, segue que $\mathcal{H} = \overline{[N]}$, ou seja, o subespaço gerado por N é denso em $\mathcal{H}$. □

5.3 Conjuntos ortonormais

Nesta seção destacam-se: Desigualdade de Bessel e a Relação de Parseval, processo de ortonormalização de Gram-Schmidt e conjunto ortonormal total, o qual define uma "base" para o espaço de Hilbert. Inicialmente segue a definição de conjunto ortonormal.

Definição 5.11. *Seja E um espaço com produto interno. Um* ***conjunto ortogonal*** *$S \subset E$ é um conjunto cujos elementos, distintos dois a dois, são ortogonais. Um* ***conjunto ortonormal*** *$S \subset E$ é um conjunto ortogonal cujos elementos tem norma 1, ou seja, para todo $x, y \in S$ tem-se*

$$\langle x, y\rangle = \begin{cases} 0, & \text{se } x \neq y, \\ 1, & \text{se } x = y. \end{cases}$$

Se o conjunto ortogonal ou ortonormal S for enumerável, é possível organizá-lo em uma sequência (x_n), denominada sequência ortogonal ou sequência ortonormal, respectivamente.

Geralmente, uma família (x_α), $\alpha \in I$, é chamada ortogonal se para todo $\alpha, \beta \in I$, com $\alpha \neq \beta$, tem-se $x_\alpha \perp x_\beta$. A família (x_α) é chamada ortonormal se é ortogonal e todo elemento x_α tem norma 1, ou seja, para todo $\alpha, \beta \in I$ tem-se

$$\langle x_\alpha, x_\beta\rangle = \delta_{\alpha\beta} = \begin{cases} 0, & \text{se } \alpha \neq \beta, \\ 1, & \text{se } \alpha = \beta. \end{cases}$$

Observação 5.6. Foi mencionado anteriormente que se dois vetores são ortogonais, em um espaço com produto interno, então vale o Teorema de Pitágoras. O mesmo acontece para conjuntos ortogonais finitos, ou seja, se $S = \{x_1, \ldots, x_n\}$ é um conjunto ortogonal, então

$$\|x_1 + \ldots + x_n\|^2 = \|x_1\|^2 + \ldots + \|x_n\|^2.$$

De fato, pela definição anterior, $\langle x_i, x_j\rangle = 0$ se $i \neq j$. Logo,

$$\left\|\sum_{i=1}^n x_i\right\|^2 = \left\langle \sum_{i=1}^n x_i, \sum_{j=1}^n x_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \langle x_i, x_j\rangle = \sum_{i=1}^n \langle x_i, x_i\rangle = \sum_{i=1}^n \|x_i\|^2.$$

Se o conjunto é ortonormal tem-se o seguinte resultado.

Lema 5.5. *Um conjunto ortonormal é linearmente independente.*

Demonstração. Seja $S = \{e_1, \ldots, e_n\}$ um conjunto ortonormal e considere a seguinte combinação linear

$$\alpha_1 e_1 + \ldots + \alpha_n e_n = 0.$$

Como $\langle e_i, e_j \rangle = 0$ se $i \neq j$, então

$$0 = \langle 0, e_j \rangle = \left\langle \sum_{i=1}^{n} \alpha_i e_i, e_j \right\rangle = \sum_{i=1}^{n} \alpha_i \langle e_i, e_j \rangle = \alpha_j \langle e_j, e_j \rangle = \alpha_j,$$

ou seja, $\alpha_j = 0$. Este lema permanece válido se S for um conjunto infinito, pela própria definição de independência linear. A demonstração segue de forma análoga. $\square$

Exemplo 5.7. Considere o espaço $\mathbb{R}^n$ com o produto interno canônico, ou seja, se $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ em $\mathbb{R}^n$, então

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

O conjunto de vetores $e_1 = (1, 0, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, ..., $e_n = (0, 0, \ldots, 1)$, formam um conjunto ortonormal em $\mathbb{R}^n$. Note que $\langle e_i, e_j \rangle = 0$ se $i \neq j$ e além disso, para todo $j \in \{1, 2, \ldots, n\}$, tem-se $\|e_j\| = 1$.

Exemplo 5.8. No espaço ℓ^2 a sequência (e_n), onde $e_n = (\delta_{nj})$, é uma sequência ortonormal, basta observar que $e_1 = (1, 0, 0, 0, \ldots)$, $e_2 = (0, 1, 0, 0, \ldots)$, $e_3 = (0, 0, 1, 0, \ldots)$, e assim sucessivamente.

Exemplo 5.9. Seja $E = \mathcal{C}([0, 2\pi], \mathbb{R})$ o espaço das funções contínuas com produto interno definido por

$$\langle f, g \rangle = \int_0^{2\pi} f(t) g(t)\, dt.$$

(A) Considere (x_n) uma sequência em E definida por

$$x_n(t) = \cos(nt), \;\; n \in \{0, 1, 2, \ldots\}.$$

A partir da sequência (x_n) será construída uma sequência ortonormal denotada por (e_n). Considerando o produto interno definido neste espaço, tem-se

$$\langle x_n, x_m \rangle = \int_0^{2\pi} x_n(t) x_m(t)\, dt = \int_0^{2\pi} \cos(nt)\cos(mt)\, dt.$$

Note que se $m = n = 0$, então

$$\int_0^{2\pi} \cos(nt)\cos(mt)\, dt = \int_0^{2\pi} dt = 2\pi.$$

Se $m = n = 1, 2, \ldots$, então

$$\int_0^{2\pi} \cos(nt)\cos(mt)\, dt = \int_0^{2\pi} \cos^2(nt)\, dt.$$

Utilizando identidade trigonométrica e técnica de integração, segue que

$$\begin{aligned}
\int_0^{2\pi} \cos^2(nt)\,dt &= \frac{1}{n}\int_0^{2\pi n} \cos^2(u)\,du \\
&= \frac{1}{n}\int_0^{2\pi n} \frac{1+\cos(2u)}{2}\,du \\
&= \frac{1}{2n}\left[u\Big|_0^{2\pi n}\right] + \frac{1}{4n}\left[\operatorname{sen}(v)\Big|_0^{4\pi n}\right] \\
&= \pi,
\end{aligned}$$

ou seja,

$$\int_0^{2\pi} \cos(nt)\cos(mt)\,dt = \pi, \text{ com } n = m.$$

Agora, se $m \neq n$, utilizando identidade trigonométrica resulta que

$$\begin{aligned}
\int_0^{2\pi} \cos(nt)\cos(mt)\,dt &= \frac{1}{2}\int_0^{2\pi} [\cos(nt-mt)+\cos(nt+mt)]\,dt \\
&= \frac{1}{2}\int_0^{2\pi} \cos((n-m)t)\,dt + \frac{1}{2}\int_0^{2\pi} \cos((n+m)t)\,dt.
\end{aligned}$$

Fazendo $n - m = p$ e $n + m = q$ e utilizando técnica de integração, segue que

$$\begin{aligned}
\int_0^{2\pi} \cos(nt)\cos(mt)\,dt &= \frac{1}{2}\int_0^{2\pi} \cos(pt)\,dt + \frac{1}{2}\int_0^{2\pi} \cos(qt)\,dt \\
&= \frac{1}{2p}\int_0^{2\pi p} \cos(u)\,du + \frac{1}{2q}\int_0^{2\pi q} \cos(v)\,dv \\
&= \frac{1}{2p}\left[\operatorname{sen}(u)\Big|_0^{2\pi p}\right] + \frac{1}{2q}\left[\operatorname{sen}(v)\Big|_0^{2\pi q}\right] \\
&= 0,
\end{aligned}$$

ou seja,

$$\int_0^{2\pi} \cos(nt)\cos(mt)\,dt = 0, \text{ com } n \neq m.$$

Resumindo,

$$\langle x_n, x_m\rangle = \int_0^{2\pi} \cos(nt)\cos(mt)\,dt = \begin{cases} 0, & \text{se } m \neq n, \\ \pi, & \text{se } m = n = 1, 2, \ldots, \\ 2\pi, & \text{se } m = n = 0. \end{cases}$$

Assim, uma sequência ortonormal é definida por (e_n), em que se $n = 0$ então

$$e_0(t) = \frac{x_0(t)}{\|x_0\|} = \frac{\cos(0t)}{\sqrt{\langle x_0, x_0 \rangle}} = \frac{1}{\sqrt{2\pi}}$$

e se $n = 1, 2, \ldots$, tem-se

$$e_n(t) = \frac{x_n(t)}{\|x_n\|} = \frac{\cos(nt)}{\sqrt{\langle x_n, x_n \rangle}} = \frac{\cos(nt)}{\sqrt{\pi}}.$$

Note que

$$\langle e_n, e_m \rangle = \begin{cases} 0, & \text{se } n \neq m \\ 1, & \text{se } n = m \end{cases} \quad \text{e ainda} \quad \|e_n\| = 1, \; n = 0, 1, 2, \ldots.$$

(B) Considere no mesmo espaço E a seguinte sequência

$$y_n(t) = \operatorname{sen}(nt), \; n \in \{1, 2, \ldots\}.$$

Como em (A), a partir da sequência (y_n) definida acima, será construída uma sequência $(\tilde{e}_n)$ ortonormal. Assim,

$$\langle y_n, y_m \rangle = \int_0^{2\pi} y_n(t) y_m(t)\, dt = \int_0^{2\pi} \operatorname{sen}(nt) \operatorname{sen}(mt)\, dt.$$

Se $m = n = 1, 2, \ldots$, então

$$\begin{aligned}
\int_0^{2\pi} \operatorname{sen}(nt)\operatorname{sen}(mt)\, dt &= \int_0^{2\pi} \operatorname{sen}^2(nt)\, dt \\
&= \frac{1}{n}\int_0^{2\pi n} \operatorname{sen}^2(u)\, du \\
&= \frac{1}{n}\int_0^{2\pi n} \frac{1 - \cos(2u)}{2}\, du \\
&= \frac{1}{2n}\left[u \Big|_0^{2\pi n} \right] - \frac{1}{4n}\left[\operatorname{sen}(v) \Big|_0^{4\pi n} \right] \\
&= \pi.
\end{aligned}$$

Agora, se $m \neq n$ então

$$\begin{aligned}\int_0^{2\pi} \operatorname{sen}(nt)\operatorname{sen}(mt)\,dt &= \frac{1}{2}\int_0^{2\pi}[\cos(nt-mt)-\cos(nt+mt)]\,dt\\ &= \frac{1}{2}\int_0^{2\pi}\cos((n-m)t)\,dt - \frac{1}{2}\int_0^{2\pi}\cos((n+m)t)\,dt.\end{aligned}$$

Fazendo $n-m=p$ e $n+m=q$ obtem-se

$$\begin{aligned}\int_0^{2\pi} \operatorname{sen}(nt)\operatorname{sen}(mt)\,dt &= \frac{1}{2}\int_0^{2\pi}\cos(pt)\,dt - \frac{1}{2}\int_0^{2\pi}\cos(qt)\,dt\\ &= \frac{1}{2p}\int_0^{2\pi p}\cos(u)\,du - \frac{1}{2q}\int_0^{2\pi q}\cos(v)\,dv\\ &= \frac{1}{2p}\left[\operatorname{sen}(u)\Big|_0^{2\pi p}\right] - \frac{1}{2q}\left[\operatorname{sen}(v)\Big|_0^{2\pi q}\right]\\ &= 0.\end{aligned}$$

Ou seja,

$$\langle y_n, y_m\rangle = \int_0^{2\pi}\operatorname{sen}(nt)\operatorname{sen}(mt)\,dt = \begin{cases} 0, & \text{se } m \neq n,\\ \pi, & \text{se } m = n = 1, 2, \ldots. \end{cases}$$

Então a sequência ortonormal $(\tilde{e}_n)$ é definida por

$$\tilde{e}_n(t) = \frac{y_n(t)}{\|y_n\|} = \frac{\operatorname{sen}(nt)}{\sqrt{\langle y_n, y_n\rangle}} = \frac{\operatorname{sen}(nt)}{\sqrt{\pi}}.$$

Observe também que

$$\langle \tilde{e}_n, \tilde{e}_m\rangle = \begin{cases} 0, & \text{se } n \neq m\\ 1, & \text{se } n = m \end{cases} \quad \text{e ainda} \quad \|\tilde{e}_n\| = 1,\ n = 1, 2, \ldots.$$

Note ainda que $x_n \perp y_m$ para todo n e m. De fato,

$$\begin{aligned}\langle x_n, y_m\rangle = \int_0^{2\pi}\cos(nt)\operatorname{sen}(mt)\,dt &= \frac{1}{2}\int_0^{2\pi}[\operatorname{sen}(nt-mt)+\operatorname{sen}(nt+mt)]\,dt\\ &= \frac{1}{2}\int_0^{2\pi}\operatorname{sen}((n-m)t)\,dt + \frac{1}{2}\int_0^{2\pi}\operatorname{sen}((n+m)t)\,dt.\end{aligned}$$

Fazendo $m-n=p$ e $n+m=q$ resulta que

$$
\begin{aligned}
\langle x_n, y_m\rangle = \int_0^{2\pi} \cos(nt)\,\text{sen}(mt)\,dt = &\ \frac{1}{2}\int_0^{2\pi}\text{sen}(pt)\,dt + \frac{1}{2}\int_0^{2\pi}\text{sen}(qt)\,dt \\
= &\ \frac{1}{2p}\int_0^{2\pi p}\text{sen}(u)\,du + \frac{1}{2q}\int_0^{2\pi q}\text{sen}(v)\,dv \\
= &\ -\frac{1}{2p}\left[\cos(u)\Big|_0^{2\pi p}\right] - \frac{1}{2q}\left[\cos(v)\Big|_0^{2\pi q}\right] \\
= &\ -\frac{1}{2p}\left[\cos(2\pi p) - \cos(0)\right] - \frac{1}{2p}\left[\cos(2\pi q) - \cos(0)\right] \\
= &\ -\frac{1}{2p}[1-1] - \frac{1}{2q}[1-1] \\
= &\ 0,
\end{aligned}
$$

ou seja,

$$\langle x_n, y_m\rangle = 0.$$

Outras sequências ortonormais aparecerão mais adiante.

Observação 5.7. Sejam V um espaço vetorial n-dimensional e B uma base ortonormal de V, então para qualquer $v \in V$ tem-se

$$v = \langle v_1, e_1\rangle e_1 + \ldots + \langle v_n, e_n\rangle e_n,$$

ou seja, $v = \displaystyle\sum_{j=1}^{n}\langle v_j, e_j\rangle e_j$. Isto motiva o seguinte resultado.

Teorema 5.7. ***(Desigualdade de Bessel)*** *Seja $(e_k)_{k\in\mathbb{N}}$ uma sequência ortonormal em um espaço com produto interno E. Então para todo $x \in E$*

$$\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 \leq \|x\|^2.$$

Demonstração. Seja $(e_1, e_2, e_3, \ldots)$ é uma sequência ortonormal em E. Se para um elemento $v \in E$ ocorrer que $v \in Y_n = [e_1, e_2, \ldots, e_n]$, onde n é fixo, então pela definição de subespaço gerado, existem escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que

$$v = \sum_{k=1}^{n} \alpha_k e_k. \tag{5.15}$$

Como $\langle e_k, e_j\rangle = 0$ se $k \neq j$, considerando o produto interno de v por um vetor fixo

e_j segue que

$$\langle v, e_j \rangle = \left\langle \sum_{k=1}^{n} \alpha_k e_k, e_j \right\rangle = \sum_{k=1}^{n} \alpha_k \langle e_k, e_j \rangle = \alpha_j,$$

ou seja, (5.15) pode ser reescrita como

$$v = \sum_{k=1}^{n} \langle v, e_k \rangle e_k. \tag{5.16}$$

Agora, considere um elemento $x \in E$ arbitrário, não necessariamente em Y_n, e seja $y \in Y_n$ definido por

$$y = \sum_{k=1}^{n} \langle x, e_k \rangle e_k, \tag{5.17}$$

onde n é fixo. Defina $z \in E$ tal que $x = y + z$, ou seja, $z = x - y$, então $z \perp y$. Com efeito, por (5.17) segue que

$$\|y\|^2 = \left\langle \sum_{k=1}^{n} \langle x, e_k \rangle e_k, \sum_{k=1}^{n} \langle x, e_k \rangle e_k \right\rangle = \sum_{k=1}^{n} \langle x, e_k \rangle \overline{\langle x, e_k \rangle} \langle e_k, e_k \rangle = \sum_{k=1}^{n} |\langle x, e_k \rangle|^2. \tag{5.18}$$

Logo,

$$\begin{aligned} \langle z, y \rangle = \langle x - y, y \rangle &= \langle x, y \rangle - \langle y, y \rangle \\ &= \left\langle x, \sum_{k=1}^{n} \langle x, e_k \rangle e_k \right\rangle - \langle y, y \rangle \\ &= \sum_{k=1}^{n} \overline{\langle x, e_k \rangle} \langle x, e_k \rangle - \|y\|^2 \\ &= \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 - \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \\ &= 0. \end{aligned}$$

Agora, como $z = x - y$ e $z \perp y$ então pelo teorema 5.1 de Pitágoras (página 113),

$$\|z + y\|^2 = \|y\|^2 + \|z\|^2 \Rightarrow \|x\|^2 = \|y\|^2 + \|z\|^2,$$

ou seja,

$$\|z\|^2 = \|x\|^2 - \|y\|^2.$$

De (5.18) segue que

$$0 \leq \|z\|^2 = \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k \rangle|^2,$$

isto é,

$$\sum_{k=1}^{n} |\langle x, e_k\rangle|^2 \leq \|x\|^2. \tag{5.19}$$

Note que esta soma possui termos não negativos, de modo que eles formam uma sequência monótona crescente e limitada (pelo termo $\|x\|^2$) e, portanto, convergente. Além disso esta é uma sequência da soma parcial de uma série infinita, que é convergente. Logo, (5.19) implica que

$$\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 \leq \|x\|^2. \tag{5.20}$$

□

Observação 5.8. No teorema acima, se o espaço E tiver dimensão finita, então todo conjunto ortonormal em E deve ser finito, pois pelo lema 5.5 (página 128) este conjunto é linearmente independente. Logo, (5.20) deve ter soma finita.

Se em (5.20) ocorrer a igualdade, ou seja,

$$\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 = \|x\|^2 \tag{5.21}$$

esta relação passa a ser chamada de **Relação de Parseval**, cuja soma será finita se o espaço for de dimensão finita, ou infinita, caso contrário.

Definição 5.12. *Seja (e_k) uma sequência ortonormal em um espaço E com produto interno. Para todo $x \in E$, os produtos internos $\langle x, e_k\rangle$ (em (5.20)) são chamados* ***coeficientes de Fourier*** *de x em relação à sequência ortonormal (e_k).*

Observação 5.9. Seja V um espaço vetorial munido de produto interno. Note que se $\{e_1, \ldots, e_n\}$ é um conjunto ortonormal em V e $v \in V$ então

$$w = v - \sum_{j=1}^{n} \langle v, e_j\rangle e_j \tag{5.22}$$

é ortogonal a cada e_k, $1 \leq k \leq n$. Com efeito,

$$\begin{aligned} \langle w, e_k\rangle = \langle v - \sum_{j=1}^{n} \langle v, e_j\rangle e_j, e_k\rangle &= \langle v, e_k\rangle - \sum_{j=1}^{n} \langle v, e_j\rangle\langle e_j, e_k\rangle \\ &= \langle v, e_k\rangle - \langle v, e_k\rangle\langle e_k, e_k\rangle \\ &= \langle v, e_k\rangle - \langle v, e_k\rangle = 0. \end{aligned}$$

Portanto, como o vetor em (5.22) é ortogonal a cada e_k, segue que ele é ortogonal ao subespaço gerado pelo conjunto $\{e_1, \ldots, e_n\}$.

É muito útil trabalhar com sequência ortonormal, devido às propriedades que a mesma possui. Porém, caso a sequência não seja ortonormal, dada uma sequência $(x_1, \ldots, x_n)$ linearmente independente em um espaço E com produto interno é possível, a partir desta sequência, construir uma sequência ortonormal $(e_1, \ldots, e_n)$ tal que para todo $n \in \mathbb{N}$, estas sequências gerem o mesmo subespaço de E, ou seja, $[x_1, \ldots, x_n] = [e_1, \ldots, e_n]$ para todo $n \in \mathbb{N}$. Este processo de construção chama-se *processo de ortonormalização de Gram-Schmidt*, como mostra a teorema abaixo.

Teorema 5.8. *Seja $(x_j)_{j\in\mathbb{N}}$ uma sequência de vetores linearmente independentes em um espaço E com produto interno. Então existe uma sequência ortonormal $(e_j)_{j\in\mathbb{N}}$ em E tal que*

$$[x_1, \ldots, x_j] = [e_1, \ldots, e_j], \ \forall\, j \in \mathbb{N}.$$

Demonstração. A construção dos elementos e_j é feita utilizando um processo de recorrência. Note que a sequência (e_j) precisa ser ortogonal e seus elementos ter norma 1, então o processo ocorrerá em duas etapas, como segue.

1) Inicialmente o primeiro passo é construir um conjunto ortogonal $\{v_1, \ldots, v_n\}$ a partir do conjunto $\{x_1, \ldots, x_n\}$ que contém elementos da sequência $(x_j)_{j\in\mathbb{N}}$. Para isto, defina $v_1 = x_1$. O vetor v_2 é obtido considerando $v_2 = x_2 - \alpha x_1$, onde o escalar α é escolhido de forma que $\langle v_2, x_1\rangle = 0$, ou seja, $\langle x_2 - \alpha x_1, x_1\rangle = 0$, resultando que

$$0 = \langle x_2, x_1\rangle - \alpha\langle x_1, x_1\rangle \Rightarrow \alpha = \frac{\langle x_2, x_1\rangle}{\|x_1\|^2},$$

ou seja,

$$v_2 = x_2 - \frac{\langle x_2, x_1\rangle}{\|x_1\|^2} x_1 = x_2 - \frac{\langle x_2, v_1\rangle}{\|v_1\|^2} v_1,$$

ilustrado na figura abaixo,

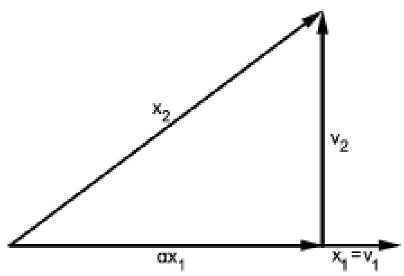

Figura 5.2: *Ideia geométrica do processo de Gram-Schmidt*

Para obter o vetor v_3, considere $v_3 = x_3 - \beta_1 v_1 - \beta_2 v_2$ onde os escalares β_1 e β_2 são escolhidos de forma que $\langle v_3, v_1\rangle = \langle v_3, v_2\rangle = 0$, ou seja,

$$\langle x_3 - \beta_1 v_1 - \beta_2 v_2, v_1\rangle = 0 \quad \text{e} \quad \langle x_3 - \beta_1 v_1 - \beta_2 v_2, v_2\rangle = 0$$

o que resulta

$$0 = \langle x_3, v_1\rangle - \beta_1\|v_1\|^2 - \beta_2\langle v_2, v_1\rangle = \langle x_3, v_1\rangle - \beta_1\|v_1\|^2 \Rightarrow \beta_1 = \frac{\langle x_3, v_1\rangle}{\|v_1\|^2}$$

e

$$0 = \langle x_3, v_2\rangle - \beta_1\langle v_1, v_2\rangle - \beta_2\|v_2\|^2 = \langle x_3, v_2\rangle - \beta_2\|v_2\|^2 \Rightarrow \beta_2 = \frac{\langle x_3, v_2\rangle}{\|v_2\|^2},$$

isto é,

$$v_3 = x_3 - \frac{\langle x_3, v_1\rangle}{\|v_1\|^2}v_1 - \frac{\langle x_3, v_2\rangle}{\|v_2\|^2}v_2.$$

Procedendo desta forma, na n-ésima etapa de construção deve-se ter

$$v_n = x_n - \sum_{j=1}^{n-1}\frac{\langle x_n, v_j\rangle}{\|v_j\|^2}v_j.$$

Logo, o conjunto $\{v_1, \ldots, v_n\}$ é ortogonal. Note que o vetor v_n é não nulo para todo $n \in \mathbb{N}$, pois caso contrário x_n seria um combinação linear, contradizendo o fato de que o conjunto $\{x_1, \ldots, x_n\}$ é linearmente independente. Portanto, definindo $e_j = v_j/\|v_j\|$, $1 \leq j \leq n$, segue que o conjunto $\{e_1, \ldots, e_n\}$ é ortonormal e, além disso, tem-se $[e_1, \ldots, e_n] = [x_1, \ldots, x_n]$.

2) Supondo o processo construído para $e_1, \ldots, e_n$, considere o vetor x_{n+1} e por construção define-se

$$e_{n+1} = \frac{v_{n+1}}{\|v_{n+1}\|}, \tag{5.23}$$

onde

$$v_{n+1} = x_{n+1} - \sum_{j=1}^{n}\langle x_{n+1}, e_j\rangle e_j. \tag{5.24}$$

Resta provar que

i. $\{e_1, \ldots, e_n, e_{n+1}\}$ é ortonormal,

ii. $[e_1, \ldots, e_n, e_{n+1}] = [x_1, \ldots, x_n, x_{n+1}]$.

Com efeito, o item *i.* segue da observação 5.9, pois o vetor e_{n+1} é ortogonal ao conjunto gerado por $\{e_1, \ldots, e_n\}$.

Agora, em *ii.* tem-se por hipótese que $[e_1, \ldots, e_n] = [x_1, \ldots, x_n]$. Como o vetor e_{n+1} é uma combinação linear dos elementos do conjunto $\{x_1, \ldots, x_{n+1}\}$ segue que $[e_1, \ldots, e_{n+1}] \subset [x_1, \ldots, x_{n+1}]$. Por outro lado, substituindo (5.24) em (5.23) resulta que

$$x_{n+1} = \|v_{n+1}\|e_{n+1} + \sum_{j=1}^{n}\langle x_{n+1}, e_j\rangle e_j,$$

ou seja, o vetor x_{n+1} também é uma combinação linear dos elementos do conjunto $\{e_1, \ldots, e_{n+1}\}$, de forma que $[x_1, \ldots, x_{n+1}] \subset [e_1, \ldots, e_{n+1}]$ concluindo, portanto, que

$[e_1, \dots, e_{n+1}] = [x_1, \dots, x_{n+1}]$, completando a demonstração do teorema. $\square$

Exemplo 5.10. Considere o espaço do polinômios $\mathcal{P}_2(\mathbb{R})$ com o produto interno definido por

$$\langle p(t), q(t) \rangle = \int_0^1 p(t)q(t)\, dt$$

e seja $B = \{1, t, t^2\}$ a base canônica deste espaço. Utilizando o processo de ortonormalização de Gram-Schmidt será obtida uma base ortonormal $B' = \{e_1, e_2, e_3\}$ a partir da base B.

Com efeito, como $B = \{1, t, t^2\}$, considere $x_1 = 1$, $x_2 = t$ e $x_3 = t^2$. Inicialmente, seja $v_1 = x_1 = 1$. Assim,

$$v_2 = t - \frac{\langle t, 1 \rangle}{\|1\|^2} 1$$

onde

$$\langle t, 1 \rangle = \int_0^1 t\, dt = \frac{t^2}{2}\bigg|_0^1 = \frac{1}{2} \quad \text{e} \quad \|1\|^2 = \langle 1, 1 \rangle = \int_0^1 1\, dt = t\Big|_0^1 = 1,$$

logo,

$$v_2 = t - \frac{1/2}{1} = t - \frac{1}{2}.$$

O vetor v_3 é definido por

$$v_3 = t^2 - \frac{\langle t^2, 1 \rangle}{\|1\|^2} 1 - \frac{\langle t^2, v_2 \rangle}{\|v_2\|^2} v_2,$$

onde

$$\langle t^2, 1 \rangle = \int_0^1 t^2\, dt = \frac{t^3}{3}\bigg|_0^1 = \frac{1}{3},$$

$$\langle t^2, v_2 \rangle = \int_0^1 t^2 \left(t - \frac{1}{2}\right) dt = \int_0^1 \left(t^3 - \frac{t^2}{2}\right) dt = \left[\frac{t^4}{4} - \frac{t^3}{6}\right]\bigg|_0^1 = \frac{1}{4} - \frac{1}{6} = \frac{1}{12},$$

$$\|v_2\|^2 = \langle v_2, v_2 \rangle = \int_0^1 \left(t - \frac{1}{2}\right)^2 dt = \int_0^1 \left(t^2 - t + \frac{1}{4}\right) dt = \left[\frac{t^3}{3} - \frac{t^2}{2} + \frac{t}{4}\right]\bigg|_0^1 = \frac{1}{12},$$

assim,

$$v_3 = t^2 - \frac{1/3}{1} - \frac{1/12}{1/12}\left(t - \frac{1}{2}\right) = t^2 - t + \frac{1}{6}.$$

Como $v_1 = 1$, $v_2 = t - \dfrac{1}{2}$ e $v_3 = t^2 - t + \dfrac{1}{6}$ segue que $e_1 = \dfrac{v_1}{\|v_1\|}$, $e_2 = \dfrac{v_2}{\|v_2\|}$ e $e_3 = \dfrac{v_3}{\|v_3\|}$, ou seja, $e_1 = 1$,

$$e_2 = \frac{t - 1/2}{\sqrt{1/12}} = \frac{t}{\sqrt{1/12}} - \frac{1/2}{\sqrt{1/12}} = \frac{t}{1/2\sqrt{3}} - \frac{1/2}{1/2\sqrt{3}} = 2\sqrt{3}t - \sqrt{3} = \sqrt{3}(2t - 1).$$

Para obter e_3, note inicialmente que

$$\begin{aligned}\|v_3\|^2 = \int_0^1 \left(t^2 - t + \frac{1}{6}\right)^2 dt &= \int_0^1 \left(t^4 - 2t^3 + \frac{4}{3}t^2 - \frac{1}{3}t + \frac{1}{36}\right) dt \\ &= \left[\frac{t^5}{5} - \frac{t^4}{2} + \frac{4t^3}{9} - \frac{t^2}{6} + \frac{t}{36}\right]\Bigg|_0^1 \\ &= \frac{1}{5} - \frac{1}{2} + \frac{4}{9} - \frac{1}{6} + \frac{1}{36} = \frac{1}{180},\end{aligned}$$

assim,

$$e_3 = \frac{t^2 - t + 1/6}{\sqrt{1/180}} = \frac{t^2 - t + 1/6}{1/6\sqrt{5}} = 6\sqrt{5}t^2 - 6\sqrt{5}t + \sqrt{5} = \sqrt{5}(6t^2 - 6t + 1).$$

Portanto, $B' = \{1, \sqrt{3}(2t-1), \sqrt{5}(6t^2 - 6t + 1)\}$ é a base ortonormal. Não é difícil verificar que os vetores de B' são ortogonais e possuem norma 1.

Observação 5.10. Dada qualquer sequência ortonormal (e_k) em um espaço $\mathcal{H}$ de Hilbert, pode-se associar com (e_k) a sequência (s_n) da soma parcial

$$s_n = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n,\ n \in \mathbb{N},$$

onde $\alpha_1, \alpha_2, \ldots$, são escalares. Se (s_n) é convergente, por exemplo, para um elemento $s \in \mathcal{H}$, ou seja, $s_n \to s$ quando $n \to \infty$, então a série infinita é chamada convergente e s é chamada a soma da série

$$s = \sum_{k=1}^{\infty} \alpha_k e_k = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 + \ldots,$$

em outras palavras, $\|s_n - s\| \to 0$ quando $n \to \infty$.

Isto motiva o seguinte teorema, chamado *Teorema da Convergência.*

Teorema 5.9. *Seja (e_k) uma sequência ortonormal em um espaço $\mathcal{H}$ de Hilbert e considere a seguinte série*

$$\sum_{k=1}^{\infty} \alpha_k e_k. \tag{5.25}$$

Então:

i. A série (5.25) *converge (na norma em $\mathcal{H}$) se, e somente se, a seguinte série converge:*

$$\sum_{k=1}^{\infty} |\alpha_k|^2. \tag{5.26}$$

ii. Se (5.25) *converge, então os coeficientes α_k são os coeficientes de Fourier $\langle x, e_k\rangle$, onde x denota a soma de* (5.25) *e, neste caso, a série* (5.25) *pode ser escrita* *como*

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k. \tag{5.27}$$

iii. Para qualquer $x \in \mathcal{H}$, a série (5.25) com $\alpha_k = \langle x, e_k \rangle$ converge (na norma em $\mathcal{H}$).

Demonstração. i. Considere as somas parciais da seguinte forma:

$$s_m = \alpha_1 e_1 + \ldots + \alpha_m e_m \quad \text{e} \quad s_n = \alpha_1 e_1 + \ldots + \alpha_m e_m + \alpha_{m+1} e_{m+1} \ldots + \alpha_n e_n,\ n > m,$$

e ainda,

$$\sigma_m = |\alpha_1|^2 + \ldots + |\alpha_m|^2 \quad \text{e} \quad \sigma_n = |\alpha_1|^2 + \ldots + |\alpha_m|^2 + |\alpha_{m+1}|^2 + \ldots + |\alpha_n|^2,\ n > m.$$

Como (e_k) é ortonormal, obtem-se

$$\begin{aligned}
\|s_n - s_m\|^2 = {} & \|\alpha_{m+1} e_{m+1} \ldots + \alpha_n e_n\|^2 \\
= {} & \langle \alpha_{m+1} e_{m+1} \ldots + \alpha_n e_n \,,\, \alpha_{m+1} e_{m+1} \ldots + \alpha_n e_n \rangle \\
= {} & \alpha_{m+1} \overline{\alpha_{m+1}} \langle e_{m+1}, e_{m+1} \rangle + \ldots + \alpha_n \overline{\alpha_n} \langle e_n, e_n \rangle \\
= {} & |\alpha_{m+1}|^2 + \ldots + |\alpha_n|^2 \\
= {} & \sigma_n - \sigma_m.
\end{aligned}$$

Assim, (s_n) é uma sequência de Cauchy em $\mathcal{H}$ se, e somente se, (σ_n) é uma sequência de Cauchy em $\mathbb{R}$. Como $\mathcal{H}$ e $\mathbb{R}$ são completos, segue o resultado.

ii. Sejam $x = \sum_{k=1}^{\infty} \alpha_k e_k$ em $\mathcal{H}$ e $s_n = \alpha_1 e_1 + \ldots + \alpha_n e_n$ uma soma parcial. Como (e_k) ortonormal, considere o produto interno

$$\langle s_n, e_j \rangle = \left\langle \sum_{i=1}^{n} \alpha_i e_i, e_j \right\rangle = \sum_{i=1}^{n} \alpha_i \langle e_i, e_j \rangle = \alpha_j \langle e_j, e_j \rangle = \alpha_j,$$

para todo $j = 1, \ldots, k$, onde $k \leq n$ fixo. Por hipótese, $s_n \to x$. Pela continuidade do produto interno (lema 5.1), tem-se

$$\alpha_j = \langle s_n, e_j \rangle \to \langle x, e_j \rangle,$$

com $j \leq k$. Note que fazendo $n \to \infty$, pode-se tomar k ($k \leq n$) tão grande quanto se queira, de modo que $\alpha_j = \langle x, e_j \rangle$, $j = 1, 2, \ldots$, o que conclui a demonstração.

iii. Pela Desigualdade de Bessel (teorema 5.7), a série

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2$$

é convergente. Usando este fato e o item i, segue o resultado. $\square$

Até o momento, quando necessário, os resultados apresentados utilizaram de sequências enumeráveis e ortonormais. Isto de fato é suficiente, pelo seguinte resultado chamado *Lema de Coeficientes de Fourier.*

Lema 5.6. *Qualquer x em um espaço E munido de produto interno pode ter, no máximo, uma quantidade enumerável de coeficientes de Fourier $\langle x, e_\alpha \rangle$ não nulos em relação a uma família ortonormal (e_α), $\alpha \in I$, em E.*

Demonstração. Ainda que uma família ortonormal $(e_\alpha)_{\alpha \in I}$ em um espaço E munido de produto interno seja não enumerável (uma vez que o conjunto I de índices é não enumerável), pode-se obter os coeficientes de Fourier $\langle x, e_\alpha \rangle$ de x em E. Utilizando a Desigualdade de Bessel (página 133), pode-se concluir que, para cada $m = 1, 2, \ldots,$ fixo, o número de coeficientes de Fourier tal que $|\langle x, e_\alpha \rangle| \geq 1/m$ é finito. Com efeito, para cada m fixo, tem-se

$$s_m = \sum_{j=1}^{m} |\langle x, e_j \rangle|^2 \leq \|x\|^2,$$

ou seja, $|s_m| \leq \|x\|^2$ para todo m. Assim, como (s_m) é monótona crescente e limitada segue que a série

$$\sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2,$$

é convergente, e além disso

$$|\langle x, e_j \rangle| \leq \|x\|^2$$

para todo $j = 1, 2, \ldots$. Agora, para cada $m \geq 1$ considere o conjunto

$$J_m = \{\alpha \in I;\ |\langle x, e_\alpha \rangle| \geq 1/m\}.$$

Note que J_m é finito, pois caso contrário J_m teria um subconjunto enumerável $(\alpha_j)_{j \in \mathbb{N}}$ tal que $|\langle x, e_{\alpha_j} \rangle| \geq 1/m$ e então

$$s_{\alpha_j} = |\langle x, e_{\alpha_1} \rangle| + \ldots + |\langle x, e_{\alpha_j} \rangle|,$$

resultando que

$$s_{\alpha_j} \geq j \frac{1}{m}. \tag{5.28}$$

Fazendo $j \to \infty$ em (5.28) segue que $s_{\alpha_j} \to \infty$, o que contradiz a Desigualdade de Bessel. Logo, J_m é finito. Portanto

$$A = \{\alpha \in I;\ |\langle x, e_\alpha \rangle| \neq 0\} = \bigcup_{m=1}^{\infty} J_m.$$

Como A é uma união enumerável de conjuntos finitos segue que A é enumerável. □

Proposição 5.4. *Nas condições do lema anterior, a soma*

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$$

não depende da ordem em que os vetores $(e_\alpha)_{\alpha \in I}$ *estão dispostos na sequência.*

Demonstração. Seja (ξ_m) uma reorganização da sequência (e_n). Por definição, isto significa que existe uma aplicação $f : \mathbb{N} \to \mathbb{N}$ bijetora, com $n \mapsto m(n)$, de modo que os termos correspondentes das duas sequências sejam iguais, ou seja, $\xi_{m(n)} = e_n$.

Defina $\alpha_n = \langle x, e_n \rangle$, $\beta_m = \langle x, \xi_m \rangle$ e ainda

$$x_1 = \sum_{n=1}^{\infty} \alpha_n e_n \quad \text{e} \quad x_2 = \sum_{m=1}^{\infty} \beta_m \xi_m.$$

Pelo item *ii*. do teorema 5.9,

$$\alpha_n = \langle x, e_n \rangle = \langle x_1, e_n \rangle \quad \text{e} \quad \beta_m = \langle x, \xi_m \rangle = \langle x_2, \xi_m \rangle.$$

Usando este fato e como $\xi_{m(n)} = e_n$, obtem-se

$$\langle x_1 - x_2, e_n \rangle = \langle x_1, e_n \rangle - \langle x_2, e_n \rangle = \langle x_1, e_n \rangle - \langle x_2, \xi_{m(n)} \rangle = \langle x, e_n \rangle - \langle x, \xi_{m(n)} \rangle = 0.$$

Analogamente,

$$\langle x_1 - x_2, \xi_m \rangle = \langle x_1, \xi_m \rangle - \langle x_2, \xi_m \rangle = \langle x_1, e_n \rangle - \langle x_2, \xi_m \rangle = \langle x, e_n \rangle - \langle x, \xi_m \rangle = 0.$$

Isto implica que

$$\begin{aligned} \|x_1 - x_2\|^2 &= \left\langle x_1 - x_2, \sum_{n=1}^{\infty} \alpha_n e_n - \sum_{m=1}^{\infty} \beta_m \xi_m \right\rangle \\ &= \sum_{n=1}^{\infty} \overline{\alpha_n} \langle x_1 - x_2, e_n \rangle - \sum_{m=1}^{\infty} \overline{\beta_m} \langle x_1 - x_2, \xi_m \rangle \\ &= 0, \end{aligned}$$

ou seja, $\|x_1 - x_2\| = 0$ implicando, por definição de norma, que $x_1 = x_2$. Como (ξ_m) é uma reorganização arbitrária de (e_n), isto completa a demonstração.

□

Definição 5.13. *Um* ***conjunto total*** *(ou fundamental) em um espaço E com produto interno é um subconjunto $N \subset E$ cujo subespaço gerado é denso em E. Em outros termos, o subconjunto N é total em E se, e somente se, $\overline{[N]} = E$. Consequentemente, um conjunto ortonormal (ou sequência ou família) no espaço E que é total em E é* *chamado um* ***conjunto ortonormal total*** *(ou sequência ou família, repectivamente) em E.*

Definição 5.14. *Uma* ***base ortonormal*** *em um espaço $\mathcal{H}$ de Hilbert é um conjunto ortonormal total.*

Teorema 5.10. *Em todo espaço $\mathcal{H} \neq \{0\}$ de Hilbert existe uma base ortonormal.*

Demonstração. Seja N o conjunto de todos os subconjuntos ortonormais de $\mathcal{H}$. Como $\mathcal{H} \neq \{0\}$, existe um elemento $x \neq 0$ e um subconjunto ortonormal de $\mathcal{H}$ é $\{y\}$ onde $y = x/\|x\|$, logo $N \neq \emptyset$. A inclusão de conjuntos define uma relação de ordem parcial sobre N (exemplo 4.11 página 60). Assim, toda cadeia $C \subset N$ possui um limitante superior, a saber, é a união dos elementos de C. Pelo Lema de Zorn (lema 4.1, página 60), o conjunto N possui um elemento maximal F.

Resta mostrar que F é total em $\mathcal{H}$. Com efeito, suponha que F não é total. Pelo teorema 5.11, existe um elemento não nulo $z \in \mathcal{H}$ tal que $z \perp F$. Assim, como $z \perp F$ e $e = z/\|z\|$ é ortonormal, então $F_1 = F \cup \{e\}$ onde F é um subespaço próprio de F_1. Absurdo, pois isto contradiz o fato de F ser maximal. Portanto, como F é um conjunto ortonormal total, pela definição 5.14, F é uma base ortonormal para $\mathcal{H}$. $\square$

Observação 5.11. *Todas as bases ortonormais em um dado espaço $\mathcal{H}$ de Hilbert possuem a mesma cardinalidade.* Este conceito de cardinalidade está diretamente relacionado com a dimensão do espaço de Hilbert, a qual pode-se observar que:

1) Se o espaço de Hilbert for de dimensão finita, então a dimensão de Hilbert é a dimensão no sentido algébrico, ou seja, a dimensão é o número de elementos da base. Logo, a demonstração segue como no corolário 4.1.

2) Para um espaço de Hilbert *separável* de dimensão infinita, o resultado segue do teorema 5.13, que será apresentado mais adiante.

3) Para um espaço de Hilbert mais geral deve-se mostrar que quaisquer duas bases do espaço possuem sempre a mesma cardinalidade. A demonstração faz uso de ferramentas mais avançadas da Teoria dos Conjuntos e, portanto será omitida.

O teorema abaixo mostra que um conjunto ortonormal total não pode ser estendido por adição de novos elementos.

Teorema 5.11. *Seja N um subconjunto de um espaço E com produto interno. Então:*

i. Se N é total em E, então não existe um elemento não nulo $x \in E$ que é ortogonal a todo elemento de N, em outras palavras, se $x \perp N$ então $x = 0$.

ii. Se E é completo, ou seja, se $E = \mathcal{H}$ é um espaço de Hilbert, a condição em i. também é suficiente para a totalidade de N em $E = \mathcal{H}$.

Demonstração. *i.* Pelo teorema 5.2 (página 119), existe um espaço $\mathcal{H}$ de Hilbert que é o completamento de E. Então E, que é um subespaço de $\mathcal{H}$, é denso em $\mathcal{H}$. Como N é total em E, pela definição 5.13 o subespaço gerado por N é denso em E, ou seja, $\overline{[N]} = E$. Assim, por E ser denso em $\mathcal{H}$, segue que $\overline{[N]} = \mathcal{H}$. Pelo lema 5.4 (página 127), se $\overline{[N]} = \mathcal{H}$ então o complemento ortogonal de N em $\mathcal{H}$ é o conjunto nulo, ou seja, $N^{\perp} = \{0\}$, portanto, se $x \in E$ e $x \perp N$ segue que $x = 0$.

ii. Se $E = \mathcal{H}$ é um espaço de Hilbert satisfazendo $N^{\perp} = \{0\}$, então pelo lema 5.4, segue que $\overline{[N]} = \mathcal{H}$, isto é, o conjunto gerado por N é denso em $\mathcal{H}$. Logo, pela definição 5.13, resulta que N é total em E. □

Teorema 5.12. *Um conjunto ortonormal N em um espaço $\mathcal{H}$ de Hilbert é total em $\mathcal{H}$ se, e somente se, todo $x \in \mathcal{H}$ satisfaz a Relação de Parseval (equação* (5.21)*, página 135) (soma sobre todos os coeficientes de Fourier não nulos de x em relação à N).*

Demonstração. ($\Rightarrow$) Suponha que N seja total em $\mathcal{H}$. Considere qualquer $x \in \mathcal{H}$ e seus coeficientes de Fourier não nulos dispostos em uma sequência $\langle x, e_1\rangle$, $\langle x, e_2\rangle$, ... (conforme lema 5.6, página 141), ou escrito em alguma ordem definitiva, se houver uma quantidade finita de termos. Defina $y \in \mathcal{H}$ por

$$y = \sum_k \langle x, e_k\rangle e_k \tag{5.29}$$

(no caso em que a série é infinita, a convergência segue do teorema 5.9). O vetor $x - y$ é ortogonal à N. Com efeito, para cada e_j em (5.29), usando a ortonormalidade segue que

$$\begin{aligned}
\langle x - y, e_j\rangle &= \langle x, e_j\rangle - \langle y, e_j\rangle = \langle x, e_j\rangle - \left\langle \sum_k \langle x, e_k\rangle e_k, e_j \right\rangle \\
&= \langle x, e_j\rangle - \sum_k \langle x, e_k\rangle\langle e_k, e_j\rangle \\
&= \langle x, e_j\rangle - \langle x, e_j\rangle \\
&= 0
\end{aligned}$$

Seja $v \in N$ distinto de todo (e_k), então $\langle x, v\rangle = 0$, de modo que

$$\begin{aligned}
\langle x - y, v\rangle &= \langle x, v\rangle - \langle y, v\rangle \\
&= \langle x, v\rangle - \left\langle \sum_k \langle x, e_k\rangle e_k, v \right\rangle \\
&= \langle x, v\rangle - \sum_k \langle x, e_k\rangle\langle e_k, v\rangle \\
&= 0 - 0 = 0.
\end{aligned}$$

Assim, $x - y \perp N$, ou seja, $x - y \in N^{\perp}$. Como N é total em $\mathcal{H}$ então o conjunto gerado por N é denso em $\mathcal{H}$ e, pelo lema 5.4 (página 127), tem-se $N^{\perp} = \{0\}$, resultando que $x - y = 0$, isto é, $x = y$. Logo,

$$x = \sum_k \langle x, e_k \rangle e_k. \tag{5.30}$$

Portanto, usando a ortonormalidade segue que

$$\begin{aligned}
\|x\|^2 = \langle x, x \rangle &= \left\langle \sum_k \langle x, e_k \rangle e_k, \sum_k \langle x, e_k \rangle e_k \right\rangle \\
&= \sum_k \langle x, e_k \rangle \overline{\langle x, e_k \rangle} \langle e_k, e_k \rangle \\
&= \sum_k \langle x, e_k \rangle \overline{\langle x, e_k \rangle} \\
&= \sum_k |\langle x, e_k \rangle|^2,
\end{aligned}$$

ou seja, todo $x \in \mathcal{H}$ satisfaz a Relação de Parseval.

($\Leftarrow$) Suponha que N não é total. Pelo teorema 5.11, existe um elemento não nulo $x \in \mathcal{H}$ tal que $x \perp N$ em $\mathcal{H}$. Como $x \perp N$, em (5.21) (na Desigualdade de Bessel) deve-se ter que $\langle x, e_k \rangle = 0$ para todo $k \in \mathbb{N}$. Assim, na Relação de Parseval,

$$\sum_k |\langle x, e_k \rangle| = 0.$$

Por outro lado, como $x \neq 0$ segue que $\|x\|^2 \neq 0$, logo

$$\sum_k |\langle x, e_k \rangle| \neq \|x\|^2,$$

contradizendo a Relação de Parseval. Portanto, se

$$\sum_k |\langle x, e_k \rangle|^2 = \|x\|^2$$

para todo $x \in \mathcal{H}$, então N é total em $\mathcal{H}$. □

Teorema 5.13. *Seja $\mathcal{H}$ um espaço de Hilbert. Então:*

i. Se $\mathcal{H}$ é separável então todo conjunto ortonormal em $\mathcal{H}$ é enumerável.

ii. Se $\mathcal{H}$ contém uma sequência ortonormal que é total em $\mathcal{H}$, então $\mathcal{H}$ é separável.

Demonstração. i. Sejam $\mathcal{H}$ um espaço separável, D qualquer conjunto denso em $\mathcal{H}$ e N algum conjunto ortonormal. Como N é ortonormal, para quaisquer dois elementos

$x, y \in N$ tem-se

$$\begin{aligned} d(x,y) = \|x-y\| &= \sqrt{\langle x-y, x-y\rangle} \\ &= \sqrt{\langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle} \\ &= \sqrt{1+1} = \sqrt{2}. \end{aligned}$$

Considere as vizinhanças B_x de x e B_y de y de raio $\sqrt{2}/3$, note que $B_x \cap B_y = \emptyset$. Como D é denso em $\mathcal{H}$, existe $a \in D$ em B_x e $b \in D$ em B_y com $a \neq b$ e $B_x \cap B_y = \emptyset$.

Assim, se N fosse não enumerável, para cada $x \in N$ haveriam inúmeras vizinhanças disjuntas duas a duas, de modo que D seria não enumerável. Agora, como D é qualquer conjunto denso, segue que $\mathcal{H}$ deveria conter um conjunto denso que é não enumerável, contrariando a hipótese de que $\mathcal{H}$ é separável. Portanto, N é enumerável.

ii. Sejam (e_k) uma sequência ortonormal total em $\mathcal{H}$ e C o conjunto de todas as combinações lineares

$$\gamma_1^{(n)} e_1 + \ldots + \gamma_n^{(n)} e_n$$

$n = 1, 2, \ldots$, onde $\gamma_k^{(n)} = a_k^{(n)} + i b_k^{(n)}$ e $a_k^{(n)}$ e $b_k^{(n)}$ são racionais (se $\mathcal{H}$ for real então $b_k^{(n)} = 0$). Note que C é enumerável. Basta provar que C é denso em $\mathcal{H}$, ou seja, mostrar que para cada $x \in \mathcal{H}$ e $\epsilon > 0$ existe um $v \in C$ tal que $\|x - v\| < \epsilon$. Com efeito, como a sequência (e_k) é total em $\mathcal{H}$, existe um n tal que $Y_n = [e_1, \ldots, e_n]$ contém um ponto cuja distância de x é menor que $\epsilon/2$.

Considerando y a projeção ortogonal de x sobre o espaço Y_n tem-se

$$\|x - y\| < \frac{\epsilon}{2}, \tag{5.31}$$

e como

$$y = \sum_{k=1}^{n} \langle x, e_k\rangle e_k \tag{5.32}$$

segue que

$$\left\| x - \sum_{k=1}^{n} \langle x, e_k\rangle e_k \right\| < \frac{\epsilon}{2}.$$

Como o conjunto dos racionais é denso em $\mathbb{R}$, para cada $\langle x, e_k\rangle$ existe um $\gamma_k^{(n)}$ tal que

$$\left\| \sum_{k=1}^{n} \langle x, e_k\rangle e_k - \sum_{k=1}^{n} \gamma_k^{(n)} e_k \right\| = \left\| \sum_{k=1}^{n} [\langle x, e_k\rangle - \sum_{k=1}^{n} \gamma_k^{(n)}] e_k \right\| < \frac{\epsilon}{2}.$$

Defina $v \in C$ por

$$v = \sum_{k=1}^{n} \gamma_k^{(n)} e_k.$$

Assim,

$$\begin{aligned}
\|x - v\| &= \left\| x - \sum_{k=1}^{n} \gamma_k^{(n)} e_k \right\| \\
&= \left\| x - \sum_{k=1}^{n} \langle x, e_k\rangle e_k + \sum_{k=1}^{n} \langle x, e_k\rangle e_k - \sum_{k=1}^{n} \gamma_k^{(n)} e_k \right\| \\
&\leq \left\| x - \sum_{k=1}^{n} \langle x, e_k\rangle e_k \right\| + \left\| \sum_{k=1}^{n} \langle x, e_k\rangle e_k - \sum_{k=1}^{n} \gamma_k^{(n)} e_k \right\| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Portanto, C é denso em $\mathcal{H}$ e como C é enumerável, segue que $\mathcal{H}$ é separável. □

Definição 5.15. *A dimensão de um espaço de Hilbert, chamada* ***dimensão hilbertiana****, é a cardinalidade de uma base ortonormal desse espaço.*

Teorema 5.14. *Sejam $\mathcal{H}$ e $\tilde{\mathcal{H}}$ dois espaços de Hilbert, ambos reais ou ambos complexos. $\mathcal{H}$ e $\tilde{\mathcal{H}}$ são isomorfos se, e somente se, possuem a mesma dimensão hilbertiana.*

Demonstração. ($\Rightarrow$) Sejam $\mathcal{H}$ e $\tilde{\mathcal{H}}$ dois espaços de Hilbert e $T : \mathcal{H} \to \tilde{\mathcal{H}}$ um isomorfismo entre estes espaços, ou seja, $\mathcal{H} \cong \tilde{\mathcal{H}}$. Como T é um isomorfismo, pela definição 5.5, para todo $x, y \in \mathcal{H}$ tem-se

$$\langle T(x), T(y)\rangle = \langle x, y\rangle,$$

isto é, todo elemento ortonormal em $\mathcal{H}$ tem imagem ortonormal em T. Uma vez que T é bijetora, pode-se concluir que T aplica cada conjunto ortonormal total de $\mathcal{H}$ em um conjunto ortonormal total de $\tilde{\mathcal{H}}$, em outras palavras, T "leva" base ortonormal em base ortonormal. Portanto, $\mathcal{H}$ e $\tilde{\mathcal{H}}$ possuem a mesma dimensão hilbertiana.

($\Leftarrow$) Suponha que $\mathcal{H}$ e $\tilde{\mathcal{H}}$ possuem a mesma dimensão hilbertiana. Note que se $\mathcal{H} = \{0\}$ e $\tilde{\mathcal{H}} = \{0\}$, segue o resultado. Suponha que $\mathcal{H} \neq \{0\}$, então $\tilde{\mathcal{H}} \neq \{0\}$ e, além disso, qualquer base ortonormal B em $\mathcal{H}$ e $\tilde{B}$ em $\tilde{\mathcal{H}}$ tem a mesma cardinalidade (conforme definição 5.15). Assim, pode-se escrever $B = (e_k)$ e $\tilde{B} = (\tilde{e}_k)$ com o mesmo conjunto de índices $\{k\}$.

Para mostrar que $\mathcal{H}$ e $\tilde{\mathcal{H}}$ são isomorfos, será construído um isomorfismo $T : \mathcal{H} \to \tilde{\mathcal{H}}$. Para todo $x \in \mathcal{H}$ tem-se

$$x = \sum_{k} \langle x, e_k\rangle e_k \tag{5.33}$$

onde o lado direito é uma soma finita ou uma série infinita (conforme lema 5.6, página 141) e pela Desigualdade de Bessel (teorema 5.7, página 133)

$$\sum_{k} |\langle x, e_k\rangle|^2 < \infty.$$

Defina

$$\tilde{x} = T(x) = \sum_k \langle x, e_k \rangle \tilde{e}_k, \tag{5.34}$$

de modo que $\tilde{x} \in \tilde{\mathcal{H}}$ e a convergência segue do teorema 5.9 (página 139). Como o produto interno é linear com respeito a primeira componente, segue que a transformação T é linear. De fato, sejam $x, y \in \mathcal{H}$ e $\alpha \in \mathbb{K}$ então

$$T(\alpha x + y) = \sum_k \langle \alpha x + y, e_k \rangle \tilde{e}_k = \alpha \sum_k \langle x, e_k \rangle \tilde{e}_k + \sum_k \langle y, e_k \rangle \tilde{e}_k = \alpha T(x) + T(y).$$

Além disso, T é uma isometria. Com efeito, de (5.34) e (5.33) obtem-se

$$\|\tilde{x}\|^2 = \|T(x)\|^2 = \left\| \sum_k \langle x, e_k \rangle \tilde{e}_k \right\| = \sum_k |\langle x, e_k \rangle|^2 = \|x\|^2$$

ou seja, T preserva norma e, portanto, T é injetora.

Finalmente, resta mostrar que T é sobrejetora. Para qualquer $\tilde{x} \in \tilde{\mathcal{H}}$,

$$\tilde{x} = \sum_k \alpha_k \tilde{e}_k$$

tem-se $\displaystyle\sum_k |\alpha_k|^2 < \infty$, pela Desigualdade de Bessel. Assim,

$$\sum_k \alpha_k e_k$$

é uma soma finita ou uma série infinita que converge para um elemento $x \in \mathcal{H}$ pelo teorema 5.9 (página 139), e pelo mesmo teorema $\alpha_k = \langle x, e_k \rangle$. Tem-se então $\tilde{x} = T(x)$ por (5.34). Como $\tilde{x} \in \tilde{\mathcal{H}}$ foi arbitrário, segue que T é sobrejetora. Portanto, como T é uma isometria e é bijetora, conclui-se que T é um isomorfismo entre $\mathcal{H}$ e $\tilde{\mathcal{H}}$. □

5.4 Funcionais lineares em espaços de Hilbert

Um resultado muito importante na teoria dos funcionais lineares em um espaço de Hilbert é o Teorema da Representação de Riesz, o qual fornece uma associação entre o espaço de Hilbert e seu respectivo dual.

Teorema 5.15. ***(Teorema da Representação de Riesz)*** *Todo funcional linear limitado f em um espaço $\mathcal{H}$ de Hilbert pode ser representado em termos do produto interno, ou seja,*

$$f(x) = \langle x, z_f \rangle, \ \forall \ x \in \mathcal{H},$$

onde z_f (que depende de f) é unicamente determinado por f e tem norma

$$\|f\| = \|z_f\|.$$

Demonstração. Deve-se provar que

i. f tem a representação $f(x) = \langle x, z_f \rangle$, para todo $x \in \mathcal{H}$,

ii. z_f é unicamente determinado por f,

iii. $\|f\| = \|z_f\|$.

i. Note que se $f = 0$, basta considerar $z_f = 0$. Dessa forma, $f(x) = \langle x, z_f \rangle = 0$ para todo $x \in \mathcal{H}$. Sejam $f \neq 0$ e $\mathcal{N}(f)$ o núcleo de f, ou seja, $\mathcal{N}(f) = \{x \in \mathcal{H};\ f(x) = 0\}$. Como f é contínuo, pois é limitado, segue do corolário 4.3 (página 90) que $\mathcal{N}(f)$ é um subconjunto fechado de $\mathcal{H}$. Pelo teorema 5.5 (página 123), vem que

$$\mathcal{H} = \mathcal{N}(f) \oplus \mathcal{N}(f)^{\perp}.$$

Note que $\mathcal{N}(f)^{\perp} \neq \{0\}$, pois como $f \neq 0$ segue que $\mathcal{H} \neq \mathcal{N}(f)$. Seja $\xi \in \mathcal{N}(f)^{\perp}$ com $\|\xi\| = 1$. Observe que para todo $x \in \mathcal{H}$ o vetor $(f(x)\xi - f(\xi)x) \in \mathcal{N}(f)$, pois

$$f(f(x)\xi - f(\xi)x) = f(x)f(\xi) - f(\xi)f(x) = 0.$$

Assim, como $\xi \in \mathcal{N}(f)^{\perp}$, então

$$\langle f(x)\xi - f(\xi)x, \xi \rangle = 0,\ \forall\, x \in \mathcal{H},$$

ou seja,

$$f(x)\langle \xi, \xi \rangle - f(\xi)\langle x, \xi \rangle = 0,$$

resultando que

$$f(x) = \frac{f(\xi)}{\langle \xi, \xi \rangle}\langle x, \xi \rangle = \frac{f(\xi)}{\|\xi\|^2}\langle x, \xi \rangle = f(\xi)\langle x, \xi \rangle = \langle x, \overline{f(\xi)}\xi \rangle.$$

Considerando $z_f = \overline{f(\xi)}\xi$ tem-se $f(x) = \langle x, z_f \rangle$ para todo $x \in \mathcal{H}$.

ii. Para mostrar a unicidade de z_f, suponha que para todo $x \in \mathcal{H}$ existem z_1 e z_2 tais que

$$\langle x, z_1 \rangle = f(x) = \langle x, z_2 \rangle.$$

Logo, $\langle x, z_1 \rangle - \langle x, z_2 \rangle = \langle x, z_1 - z_2 \rangle = 0$ para todo x. Escolhendo, em particular, $x = z_1 - z_2$ deve-se ter que

$$0 = \langle x, z_1 - z_2 \rangle = \langle z_1 - z_2, z_1 - z_2 \rangle = \|z_1 - z_2\|^2.$$

Por definição de norma $z_1 - z_2 = 0$, ou seja $z_1 = z_2$, o que implica a unicidade o vetor z_f.

iii. A igualdade é obvia se $f = 0$, pois $z_f = 0$. Suponha $f \neq 0$, então $z_f \neq 0$. Como f é um funcional linear limitado, pela observação 4.9 (página 95), segue que

$$|f(x)| \leq \|f\|\|x\|, \tag{5.35}$$

e pelo item *i.* tem-se $f(x) = \langle x, z_f \rangle$. Considerando $x = z_f$ e utilizando (5.35) resulta que

$$f(z_f) = \langle z_f, z_f \rangle = \|z_f\|^2 \leq \|f\|\|z_f\|,$$

ou seja,

$$\|z_f\| \leq \|f\|. \tag{5.36}$$

Agora, pela Desigualdade de Cauchy-Schwarz

$$|f(x)| = |\langle x, z_f \rangle| \leq \|x\|\|z_f\|. \tag{5.37}$$

Considerando o supremo sobre todo $x \in \mathcal{H}$ de norma 1, de (5.37) segue que

$$\|f\| = \sup_{\substack{x\in\mathcal{H} \\ \|x\|=1}} |f(x)| = \sup_{\substack{x\in\mathcal{H} \\ \|x\|=1}} |\langle x, z_f \rangle| \leq \|z_f\|,$$

isto é,

$$\|f\| \leq \|z_f\|. \tag{5.38}$$

De (5.36) e (5.38) obtem-se

$$\|f\| = \|z_f\|,$$

concluindo a demonstração do teorema. □

A ideia da unicidade (item *ii.*) no teorema acima parte do seguinte lema.

Lema 5.7. *Seja E um espaço munido de produto interno. Se $\langle v_1, w \rangle = \langle v_2, w \rangle$ para todo $w \in E$, então $v_1 = v_2$. Em particular, $\langle v_1, w \rangle = 0$ para todo $w \in E$ então $v_1 = 0$.*

Demonstração. Por hipótese, $\langle v_1, w \rangle = \langle v_2, w \rangle$ para todo $w \in E$, logo

$$\langle v_1 - v_2, w \rangle = \langle v_1, w \rangle - \langle v_2, w \rangle = 0.$$

Para $w = v_1 - v_2$ obtem-se

$$\langle v_1 - v_2, v_1 - v_2 \rangle = \|v_1 - v_2\|^2 = 0 \Rightarrow v_1 - v_2 = 0 \Rightarrow v_1 = v_2.$$

Em particular, $\langle v_1, w \rangle = 0$ com $w = v_1$ resulta que $\|v_1\|^2 = 0$, ou seja, $v_1 = 0$. □

Observação 5.12. No Teorema da Representação de Riesz a hipótese do espaço com produto interno ser completo não pode ser retirada. Considere, por exemplo, o subespaço S de $\ell^2(\mathbb{N})$ constituído por elementos com apenas um número finito de entradas não nulas. Este subespaço não é completo. Para ver isto basta considerar sequências em S da seguinte forma:

$$\begin{array}{l}(1, 1/2, 1/3, 1/4, 0, 0, 0, 0, \ldots)\\(1, 1/2, 1/3, 1/4, 1, 5, 0, 0, 0, \ldots)\\(1, 1/2, 1/3, 1/4, 1/5, 1/6, 0, 0, \ldots).\end{array}$$

Note que estas sequências convergem para

$$(1, 1/2, 1/3, 1/4, 1/5, 1/6, 1/7, 1/8, \ldots)$$

que não é elemento de S, logo S não é completo.

Agora, considere $f : S \to \mathbb{K}$ definida por

$$f(x) = \sum_{j=1}^{\infty} \frac{x_j}{j},$$

onde $x = (x_1, x_2, \ldots, x_k, 0, 0, 0, \ldots) \in S$. Como f é um funcional linear limitado, então $f \in S^*$, mas não existe $z_f \in S$ de forma que $f(x) = \langle x, z_f \rangle$ para todo $x \in S$. Com efeito, suponha que existe $z_f \in S$, com $z_f = (z_1, z_2, \ldots, z_{k-1}, 0, 0, 0, \ldots)$, $k \in \mathbb{N}$, satisfazendo $f(x) = \langle x, z_f \rangle$ para todo $x \in S$. Considere $x \in S$ tal que $x = (\delta_{nk}) = (0, 0, \ldots, 0, 1, 0, 0, \ldots)$, ou seja, a k-ésima coordenada de x é igual a 1 e as demais coordenadas nulas. Note que

$$f(x) = f((\delta_{nk})) = \frac{1}{k} \neq 0.$$

Por outro lado,

$$\begin{aligned} f(x) = \langle x, z_f \rangle &= \langle (\delta_{nk}), z_f \rangle \\ &= \langle (0, 0, \ldots, 0, 1, 0, 0, \ldots), (z_1, z_2, \ldots, z_{k-1}, 0, 0, 0, \ldots) \rangle \\ &= 0.x_1 + 0.x_2 + \ldots + 0.x_{k-1} + 1.0 + 0.0 + 0.0 + \ldots \\ &= 0, \end{aligned}$$

o que é um absurdo.

Observação 5.13. O espaço dual $\mathcal{H}'$ de um espaço de Hilbert $\mathcal{H}$ é o conjunto de todos os funcionais lineares limitados e, consequentemente, contínuos. A aplicação $T : \mathcal{H}' \to \mathcal{H}$ que associa a cada funcional linear $f \in \mathcal{H}'$ o elemento $z_f \in \mathcal{H}$, define um isomorfismo entre $\mathcal{H}'$ e $\mathcal{H}$.

Com efeito, pelo Teorema da Representação de Riesz todo funcional linear limitado f em $\mathcal{H}$ é representado por único elemento $z_f \in \mathcal{H}$ tal que $f(x) = \langle x, z_f \rangle$, para todo

$x \in \mathcal{H}$, e tem-se $\|f\| = \|z_f\|$. Por outro lado, cada elemento $z_f \in \mathcal{H}$ é unicamente identicado com um funcional linear limitado $f \in \mathcal{H}$ e $\|f\| = \|z_f\|$. Portanto, como T é bijetora e é uma isometria, a aplicação

$$\begin{aligned} T : \mathcal{H}' &\to \mathcal{H} \\ f &\mapsto z_f \end{aligned}$$

define um isomorfismo entre $\mathcal{H}'$ e $\mathcal{H}$.

Exemplo 5.11. Foi demonstrado no capítulo anterior que o dual do espaço ℓ^p é o espaço ℓ^q com $1 < p, q < \infty$ e $1/p + 1/q = 1$. Neste capítulo foi visto que somente para $p = 2$ o espaço ℓ^p é Hilbert. Logo, pode-se concluir que o dual do espaço ℓ^2 é o próprio espaço ℓ^2, ou seja, a aplicação

$$\begin{aligned} T : (\ell^2)' &\to \ell^2 \\ f &\mapsto z_f, \end{aligned}$$

onde

$$f(x) = \langle x, z_f \rangle = \sum_{j=1}^{\infty} \xi_j \overline{z_j},$$

para todo $x = (\xi_1, \xi_2, \ldots), z_f = (z_1, z_2, \ldots) \in \ell^2$ define um isomorfismo entre $(\ell^2)'$ e ℓ^2. Note que se ℓ^2 for um espaço complexo, então a aplicação é o conjugado linear, pois $\alpha f \mapsto \overline{\alpha} z_f$.

6 Aplicações

Como a Análise Funcional é uma área com várias aplicações, neste capítulo serão apresentadas algumas delas, a saber: o espaço dual do espaço $\mathcal{C}([a,b],\mathbb{R})$, o Teorema do Ponto fixo de Banach e alguns resultados envolvendo aproximação no espaço de funções.

6.1 O espaço dual de $\mathcal{C}([a,b],\mathbb{R})$

O espaço dual de $\mathcal{C}([a,b],\mathbb{R})$ pode ser identificado com um subespaço do espaço das funções de variação limitada em $[a,b]$. Esta identificação é feita utilizando um Teorema de Riesz que afirma que qualquer funcional linear limitado em $\mathcal{C}([a,b],\mathbb{R})$ é uma integral de Riemann-Stieltjes com respeito a uma função de variação limitada. Este estudo terá início com a definição e a existência da integral de Riemann-Stieltjes, como segue.

6.1.1 A integral de Riemann-Stieltjes: definição e existência da integral

Inicialmente, será feita uma abordagem sobre a integral de Riemann e posteriormente, a integral de Riemann-Stieltjes.

Definição 6.1. *Seja $[a,b]$ um intervalo dado. Uma* ***partição*** *do intervalo $[a,b]$ é um subconjunto finito de pontos $P = \{x_0, x_1, \ldots, x_n\} \subset [a,b]$ tal que $a, b \in P$ e ainda*

$$a = x_0 \leq x_1 \leq \ldots \leq x_n = b.$$

Escreve-se $\Delta x_i = x_i - x_{i-1}$ onde $i = 1, 2, \ldots, n$.

Agora suponha $f : [a,b] \to \mathbb{R}$ uma função real limitada e $P = \{x_0, x_1, \ldots, x_n\}$ uma partição de $[a,b]$. Para $i = 1, 2, \ldots, n$, sejam m_i e M_i o ínfimo e o supremo, repectivamente, dos valores de f no intervalo $[x_{i-1}, x_i]$, ou seja,

$$m_i = \inf f(x) \ (x_{i-1} \leq x \leq x_i),$$
$$M_i = \sup f(x) \ (x_{i-1} \leq x \leq x_i).$$

A partir disto define-se a ***soma inferior*** $s(P,f)$ *e a* ***soma superior*** $S(P,f)$ *da função* f *fazendo*

$$s(P,f) = m_1\Delta x_1 + \ldots + m_n\Delta x_n = \sum_{i=1}^{n} m_i\Delta x_i,$$

$$S(P,f) = M_1\Delta x_1 + \ldots + M_n\Delta x_n = \sum_{i=1}^{n} M_i\Delta x_i.$$

e finalmente,

$$\underline{\int_a^b} f(x)\,dx = \sup s(P,f) \tag{6.1}$$

$$\overline{\int_a^b} f(x)\,dx = \inf S(P,f) \tag{6.2}$$

em que o ínfimo e o supremo são relativos a todas as partições P *de* $[a,b]$*. Os membros a esquerda em* (6.1) *e* (6.2) *são chamados, repectivamente,* ***Integral de Riemann*** *inferior e superior aplicada ao intervalo* $[a,b]$*.*

Se (6.1) *e* (6.2) *coincidem, diz-se que* f *é* ***Riemann-integrável*** *em* $[a,b]$*, ou ainda, que* f *é R-integrável e o valor comum de* (6.1) *e* (6.2) *é denotado por*

$$\int_a^b f(x)\,dx, \tag{6.3}$$

que é a integral de Riemann de f *aplicada em* $[a,b]$*.*

Sendo m *o ínfimo e* M *o supremo de* f *em* $[a,b]$*, então*

$$m \leq f(x) \leq M \ (a \leq x \leq b),$$

logo, para toda partição P *do intervalo* $[a,b]$

$$m(b-a) \leq s(P,f) \leq S(P,f) \leq M(b-a)$$

de modo que os números $s(P,f)$ *e* $S(P,f)$ *contituem um conjunto limitado. Isto mostra que as integrais inferior e superior estão definidas para toda função* f *limitada.*

Observação 6.1. Uma propriedade útil das integrais de Riemann é que se f é integrável no intervalo $[a,b]$, então $|f|$ também é integrável em $[a,b]$ e satisfaz

$$\left|\int_a^b f\right| \leq \int_a^b |f|.$$

Inicialmente será introduzida a integral de Riemann-Stieltjes com respeito às funções monótonas crescentes. Posteriormente serão consideradas funções de variação limitada.

Definição 6.2. *Seja α uma função monótona crescente em $[a,b]$. Como os números $\alpha(a)$ e $\alpha(b)$ são finitos, segue-se que α é limitada em $[a,b]$. Para cada partição P de $[a,b]$, escreva*

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

Claramente, $\Delta\alpha_i \geq 0$. Para qualquer função real f limitada em $[a,b]$ considere

$$s(P,f,\alpha) = m_1\Delta x_1 + \ldots + m_n\Delta x_n = \sum_{i=1}^{n} m_i\Delta\alpha_i,$$

$$S(P,f,\alpha) = M_1\Delta x_1 + \ldots + M_n\Delta x_n = \sum_{i=1}^{n} M_i\Delta\alpha_i,$$

onde m_i e M_i tem o mesmo significado da definição anterior. Assim, por definição

$$\underline{\int_a^b} f(x)\,d\alpha = \sup s(P,f,\alpha), \tag{6.4}$$

$$\overline{\int_a^b} f(x)\,d\alpha = \inf S(P,f,\alpha), \tag{6.5}$$

sendo, novamente, o ínfimo e o supremo relativos a toda partição P de $[a,b]$. Se os membros a esquerda de (6.4) e (6.5) são iguais, o seu valor comum é denotado por

$$\int_a^b f(x)\,d\alpha(x), \tag{6.6}$$

ou simplesmente,

$$\int_a^b f\,d\alpha. \tag{6.7}$$

*Esta é a **integral de Riemann-Stieltjes**, ou simplesmente **Integral de Stieltjes**, de f com respeito a aplicação α em $[a,b]$. Se (6.6) existe, ou seja, se (6.4) e (6.5) são iguais, diz-se que f é integrável em relação a α no sentido de Riemann, e escreve-se que f é RS-integrável.*

Observação 6.2. No caso em que $\alpha(x) = x$, a integral de Riemann torna-se um caso particular da integral de Riemann-Stieltjes. Exceto quando explícito, no caso geral, α não precisa sequer ser contínua.

Note que a integral de Riemann-Stieltjes depende de f, α, a e b, portanto, quando for conveniente, será utilizada a notação de (6.7), uma vez que a letra utilizada para representar a "variável de integração" é muito subjetiva.

A partir de agora será investigada a existência da integral (6.7). Para tanto, considere f real e limitada e α monótona crescente em $[a,b]$. Quando não houver possibilidade de equívoco, será utilizado $\int$ em vez de $\int_a^b$.

Definição 6.3. *Diz-se que P^* é um **refinamento** de P se $P \subset P^*$, ou seja, se todo* *ponto de P é um ponto de P^*. Dadas duas partições P_1 e P_2, diz-se que P^* é um* ***refinamento comum*** *se $P^* = P_1 \cup P_2$.*

Teorema 6.1. *Se P^* é um refinamento de P, então*

$$s(P, f, \alpha) \leq s(P^*, f, \alpha) \tag{6.8}$$

e

$$S(P^*, f, \alpha) \leq S(P, f, \alpha). \tag{6.9}$$

Demonstração. Para demonstrar (6.8), suponha inicialmente que a partição P^* resulte de P pelo acréscimo de um único ponto x^*, ou seja, $P^* = P \cup \{x^*\}$, de forma que $x_{i-1} < x^* < x_i$, em que x_{i-1} e x_i são dois pontos consecutivos de P. Sejam

$$\begin{aligned} w_1 &= \inf f(x) \quad (x_{i-1} \leq x \leq x^*), \\ w_2 &= \inf f(x) \quad (x^* \leq x \leq x_i). \end{aligned}$$

Como

$$m_i = \inf f(x) \, (x_{i-1} \leq x \leq x_i),$$

segue que

$$w_1 \geq m_i \quad \text{e} \quad w_2 \geq m_i,$$

e além disso,

$$x_i - x_{i-1} = (x_i - x^*) + (x^* - x_{i-1}).$$

Sabendo que

$$\begin{aligned} s(P^*, f, \alpha) = &\ m_1[\alpha(x_1) - \alpha(x_0)] + m_2[\alpha(x_2) - \alpha(x_1)] + \ldots + \\ + &\ w_1[\alpha(x^*) - \alpha(x_{i-1})] + w_2[\alpha(x_i) - \alpha(x^*)] + \ldots \end{aligned}$$

$$\begin{aligned} s(P, f, \alpha) = &\ m_1[\alpha(x_1) - \alpha(x_0)] + m_2[\alpha(x_2) - \alpha(x_1)] + \ldots + \\ + &\ m_i[\alpha(x_i) - \alpha(x_{i-1})] + \ldots \end{aligned}$$

segue que

$$\begin{aligned} s(P^*, f, \alpha) - s(P, f, \alpha) = &\ w_1[\alpha(x^*) - \alpha(x_{i-1})] + w_2[\alpha(x_i) - \alpha(x^*)] \\ - &\ m_i[\alpha(x_i) - \alpha(x_{i-1})] \\ = &\ (w_1 - m_i)[\alpha(x^*) - \alpha(x_{i-1})] + (w_2 - m_i)[\alpha(x_i) - \alpha(x^*)] \geq 0 \end{aligned}$$

ou seja,

$$s(P^*, f, \alpha) \geq s(P, f, \alpha).$$

Se P^* contém k pontos a mais que P, basta repetir o procedimento acima para k pontos que obtem-se o resultado. A demonstração de (6.9) segue de forma análoga. □

Teorema 6.2. $\underline{\int_a^b} f\,d\alpha \leq \overline{\int_a^b} f\,d\alpha.$

Demonstração. Seja P^* um refinamento comum de duas partições P_1 e P_2. Pelo teorema 6.1,

$$s(P_1,f,\alpha) \leq s(P^*,f,\alpha) \leq S(P^*,f,\alpha) \leq S(P_2,f,\alpha),$$

ou seja,

$$s(P_1,f,\alpha) \leq S(P_2,f,\alpha). \tag{6.10}$$

Mantendo a partição P_2 fixa e considerando o supremo sobre toda P_1, resulta de (6.10) que

$$\sup s(P_1,f,\alpha) = \underline{\int} f\,d\alpha \leq S(P_2,f,\alpha).$$

Da mesma forma, mantendo a partição P_1 fixa e considerando o ínfimo sobre toda P_1 tem-se

$$s(P_1,f,\alpha) \leq \inf S(P_2,f,\alpha) = \overline{\int} f\,d\alpha.$$

Portanto, segue o resultado. □

Teorema 6.3. *A função f é RS-integrável se, e somente se, para todo $\epsilon > 0$, existe uma partição P tal que*

$$S(P,f,\alpha) - s(P,f,\alpha) < \epsilon.$$

Demonstração. ($\Rightarrow$) Suponha f RS-integrável. Dado $\epsilon > 0$, existem partições P_1 e P_2 tais que

$$S(P_2,f,\alpha) - \int f\,d\alpha < \frac{\epsilon}{2}, \tag{6.11}$$

$$\int f\,d\alpha - s(P_1,f,\alpha) < \frac{\epsilon}{2}. \tag{6.12}$$

Seja P um refinamento comum de P_1 e P_2. O teorema 6.1, juntamente com (6.11) e (6.12), mostra que

$$S(P,f,\alpha) \leq S(P_2,f,\alpha) < \int f\,d\alpha + \frac{\epsilon}{2} < s(P_1,f,\alpha) + \epsilon \leq s(P,f,\alpha) + \epsilon,$$

ou seja,

$$S(P,f,\alpha) < s(P,f,\alpha) + \epsilon.$$

Portanto,

$$S(P,f,\alpha) - s(P,f,\alpha) < \epsilon.$$

($\Leftarrow$) Para toda partição P tem-se

$$s(P,f,\alpha) \leq \underline{\int} f\,d\alpha \leq \overline{\int} f\,d\alpha \leq S(P,f,\alpha),$$

e como, por hipótese, dado $\epsilon > 0$ existe uma partição P de $[a, b]$ tal que

$$S(P, f, \alpha) - s(P, f, \alpha) < \epsilon,$$

então pela observação 2.1 (página 19), tem-se

$$\sup s(P, f, \alpha) = \inf S(P, f, \alpha).$$

Portanto, a função f é RS-integrável. $\square$

Teorema 6.4. *Se f é contínua em $[a, b]$, então f é RS-integrável em relação a α em $[a, b]$. Além disso, a cada $\epsilon > 0$ corresponde $\delta > 0$ tal que*

$$\left| \sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \int_a^b f\, d\alpha \right| < \epsilon \tag{6.13}$$

qualquer que seja a partição $P = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$ com $\max \Delta x_i < \delta$, e a escolha dos pontos $t_i \in [x_{i-1}, x_i]$.

Demonstração. Dado $\epsilon > 0$, seja $\eta > 0$ tal que

$$[\alpha(b) - \alpha(a)]\eta < \epsilon.$$

Como f é contínua no intervalo compacto $[a, b]$, então f é uniformemente contínua em $[a, b]$. Assim, existe $\delta > 0$ tal que $x, t \in [a, b]$

$$|x - t| < \delta \Rightarrow |f(x) - f(t)| < \eta. \tag{6.14}$$

Se P é qualquer partição de $[a, b]$ tal que $\max \Delta x_i < \delta$ $(1 \leq i \leq n)$, então de (6.14) segue que

$$M_i - m_i \leq \eta \qquad (i = 1, \ldots, n).$$

Logo,

$$\begin{aligned} S(P, f, \alpha) - s(P, f, \alpha) = & \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i \leq \sum_{i=1}^{n} \eta\Delta\alpha_i \\ = & \ \eta \sum_{i=1}^{n} \Delta\alpha_i = \eta[\alpha(b) - \alpha(a)] < \epsilon, \end{aligned}$$

ou seja,

$$S(P, f, \alpha) - s(P, f, \alpha) < \epsilon.$$

Portanto, pelo teorema 6.3, segue que f é RS-integrável. Agora, para provar (6.13), note que sendo f RS-integrável, suponha que $S(P, f, \alpha) - s(P, f, \alpha) < \epsilon$ seja válida para $P = \{x_0, x_1, \ldots, x_n\}$ e $t_i \in [x_{i-1}, x_i]$. Assim, $f(t_i) \in [m_i, M_i]$. Logo,

$$\begin{aligned} m_i \leq f(t_i) \leq M_i \Rightarrow & \sum_{i=1}^{n} m_i \Delta\alpha_i \leq \sum_{i=1}^{n} f(t_i)\Delta\alpha_i \leq \sum_{i=1}^{n} M_i \Delta\alpha_i \\ \Rightarrow & \ s(P,f,\alpha) \leq \sum_{i=1}^{n} f(t_i)\Delta\alpha_i \leq S(P,f,\alpha). \end{aligned}$$

Por outro lado, $s(P,f,\alpha) \leq \int_a^b f\,d\alpha \leq S(P,f,\alpha)$. Portanto,

$$\left| \sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \int_a^b f\,d\alpha \right| < \epsilon.$$

□

Ainda supondo que α é monótona crescente em $[a,b]$, segue o teorema.

Teorema 6.5. *Se f é monótona em $[a,b]$ e se α é contínua em $[a,b]$, então f é RS-integrável.*

Demonstração. Dado $\epsilon > 0$, qualquer que seja o inteiro positivo n, considere uma partição P tal que

$$\Delta\alpha_i = \frac{\alpha(b) - \alpha(a)}{n} \qquad (i = 1, \ldots, n),$$

o que é possível pois, por hipótese, α é contínua.

Supondo que f é monótona crescente (a demosntração para o outro caso é análoga), sejam

$$M_i = f(x_i), \quad m_i = f(x_{i-1}) \quad (i = 1, \ldots, n).$$

Assim,

$$\begin{aligned} S(P,f,\alpha) - s(P,f,\alpha) = & \ \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \\ = & \ \frac{\alpha(b) - \alpha(a)}{n} [f(b) - f(a)] < \epsilon \end{aligned}$$

desde que n seja suficientemente grande. Logo,

$$S(P,f,\alpha) - s(P,f,\alpha) < \epsilon.$$

Portanto, pelo teorema 6.3, segue que f é RS-integrável. □

A seguir serão enunciadas algumas propriedades da integral de Riemann-Stieltjes. As demonstrações serão omitidas, mas podem ser encontradas na referência [15]. Demais propriedades e resultados fundamentais sobre a integral de Riemann-Stieltjes podem ser obtidos na mesma referência.

Teorema 6.6. *a) Se f e g são RS-integráveis em $[a,b]$, então $f+g$ é RS-integrável e tem-se*

$$\int_a^b (f+g)\,d\alpha = \int_a^b f\,d\alpha + \int_a^b g\,d\alpha.$$

b) Se f é RS-integrável em $[a,b]$ e c é uma constante qualquer, então cf é RS-integrável e

$$\int_a^b cf\,d\alpha = c\int_a^b f\,d\alpha.$$

c) Se f e g são RS-integráveis em $[a,b]$ e se $f(x) \le g(x)$ em $[a,b]$, então

$$\int_a^b f\,d\alpha \le \int_a^b g\,d\alpha.$$

d) Se f é RS-integrável em $[a,b]$ e se $a < c < b$, então f é RS-integrável em $[a,c]$ e $[c,b]$ e tem-se

$$\int_a^c f\,d\alpha + \int_c^d f\,d\alpha = \int_a^b f\,d\alpha.$$

e) Se f é RS-integrável em $[a,b]$ e se $|f(x)| \le M$ em $[a,b]$, então

$$\left|\int_a^b f\,d\alpha\right| \le M[\alpha(b) - \alpha(a)].$$

f) Se f é RS-integrável em relação a α_1 em $[a,b]$ e f é RS-integrável em relação a α_2 em $[a,b]$, então f é RS-integrável em relação a $\alpha_1+\alpha_2$ em $[a,b]$ e tem-se

$$\int_a^b f\,d(\alpha_1+\alpha_2) = \int_a^b f\,d\alpha_1 + \int_a^b f\,d\alpha_2.$$

g) Se f é RS-integrável em $[a,b]$ e c uma constante positiva, então f é RS-integrável em relação a $c\alpha$ em $[a,b]$ e tem-se

$$\int_a^c f\,d(c\alpha) = c\int_a^b f\,d\alpha.$$

Segue um exemplo de integral de Riemann-Stieltjes.

Exemplo 6.1. Sejam

$$\alpha(x) = \begin{cases} 0, & \text{se } x < 1/2 \\ 2, & \text{se } x \ge 1/2 \end{cases} \quad \text{e} \quad f(x) = x^2.$$

Será demonstrado que f é RS-integrável e que

$$\int_0^1 f\,d\alpha = \frac{1}{2}.$$

A ideia é utilizar o teorema 6.3. Para tanto, dado $\epsilon > 0$ considere uma partição P do intervalo $[0,1]$ onde

$$0 = x_0 \leq x_1 \leq x_2 \leq \ldots \leq x_k \leq \ldots \leq x_{2k} = 1 \;\text{ e }\; \Delta x_k = \frac{1}{2k}.$$

É preciso determinar k. Note que

$$\begin{aligned}
&x_0 = 0\\
&x_1 = x_0 + \frac{1}{2k} = \frac{1}{2k}\\
&x_2 = x_1 + \frac{1}{2k} = \frac{1}{2k} + \frac{1}{2k} = 2\left(\frac{1}{2k}\right)\\
&x_3 = x_2 + \frac{1}{2k} = 2\left(\frac{1}{2k}\right) + \frac{1}{2k} = 3\left(\frac{1}{2k}\right)\\
&\vdots\\
&x_{k-1} = (k-1)\frac{1}{2k} = \frac{1}{2} - \frac{1}{2k}\\
&x_k = k\left(\frac{1}{2k}\right) = \frac{1}{2}.
\end{aligned}$$

Lembrando que $M_i = \sup\{f(x);\, x \in [x_{i-1}, x_i]\}$ e $m_i = \inf\{f(x);\, x \in [x_{i-1}, x_i]\}$, e sendo $f(x)$ crescente no intervalo $[0,1]$ então

$$\begin{aligned}
S(P,f,\alpha) &= \left(\frac{1}{2}\right)^2 [2-0] = \frac{1}{2},\\
s(P,f,\alpha) &= \left(\frac{1}{2} - \frac{1}{2k}\right)^2 [2-0] = 2\left[\frac{1}{4} - \frac{1}{2k} + \frac{1}{4k^2}\right] = \frac{1}{2} - \frac{1}{k} + \frac{1}{2k^2},
\end{aligned}$$

assim,

$$S(P,f,\alpha) - s(P,f,\alpha) = \frac{1}{k} - \frac{1}{2k^2}.$$

Sendo $\epsilon \leq \frac{1}{2}$, basta considerar k como sendo o maior inteiro positivo tal que

$$k > \frac{1 + \sqrt{1 - 2\epsilon}}{2\epsilon}, \text{ pois } \left(\frac{1}{k} - \frac{1}{2k^2}\right) < \epsilon,$$

logo

$$S(P,f,\alpha) - s(P,f,\alpha) < \epsilon.$$

No caso em que $\epsilon > \frac{1}{2}$, todo $k > 0$ satisfaz. Portanto, dado $\epsilon > 0$ existe uma partição P tal que $S(P,f,\alpha) - s(P,f,\alpha) < \epsilon$ determinado por k. Assim, existe $\int_a^b f\, d\alpha$, pois

$$\underline{\int_0^1} f\, d\alpha = \overline{\int_0^1} f\, d\alpha = \frac{1}{2} \Rightarrow \int_0^1 f\, d\alpha = \frac{1}{2}.$$

Observação 6.3. Até o momento, a integração da função f em $[a,b]$ foi referente a funções monótonas crescentes α. Agora, toda a teoria de integração apresentada anteriormente pode ser ampliada substituindo a classe das funções monótonas pela classe de todas as funções de *variação limitada*, como segue na definição.

Definição 6.4. *Seja $P = \{t_0, t_1, \ldots, t_n\}$ uma partição de $[a, b]$. Uma função ω definida em $[a, b]$ é chamada de **variação limitada** em $[a, b]$ se a **variação total** $Var(\omega, a, b)$ de ω em $[a, b]$ for finita, onde*

$$Var(\omega, a, b) = \sup \left\{ \sum_{i=1}^{n} |\omega(t_i) - \omega(t_{i-1})| < \infty;\ P \text{ partição de } [a, b] \right\}, \tag{6.15}$$

em que o supremo é relativo a todas as partições de $[a, b]$. Quando o intervalo é evidente, escreve-se $Var(\omega)$.

Funções de variação limitada podem, por exemplo, ser obtidas a partir do seguinte resultado.

Lema 6.1. *Se f é uma função Riemann-integrável no intervalo $[a, b]$, então a função F definida por*

$$F(x) = \int_a^x f(t)\, dt$$

é um função de variação limitada em $[a, b]$.

Demonstração. Para mostrar que F é de variação limitada, seja

$$a = x_0 < x_1 < \ldots < x_k = b$$

uma partição P de $[a, b]$. Então

$$\sum_{i=1}^{k} |F(x_i) - F(x_{i-1})| = \sum_{i=1}^{k} \left| \int_{x_{i-1}}^{x_i} f(t)\, dt \right| \leq \sum_{i=1}^{k} \int_{x_{i-1}}^{x_i} |f(t)|\, dt = \int_a^b |f(t)|\, dt.$$

Portanto, considerando o supremo sobre todas as possíveis partições do intervalo $[a, b]$, tem-se

$$\mathrm{Var}(F) \leq \int_a^b |f(t)|\, dt < \infty.$$

□

Observação 6.4. O conjunto das funções de variação limitada em $[a, b]$ forma um espaço vetorial, pois a classe de funções de variação limitada é fechada em relação às operações de adição e multiplicação por escalar (ver referência [15]). A norma neste espaço é definida por

$$\|\omega\| = |\omega(a)| + \mathrm{Var}(\omega). \tag{6.16}$$

Para mostrar que (6.16) define uma norma, considere $P = \{t_0, t_1, \ldots, t_n\}$ qualquer partição de $[a, b]$ e a

$$\mathrm{Var}(\omega) = \sup \sum_{j=1}^{n} |\omega(t_j) - \omega(t_{j-1})|,$$

em que o supremo é relativo a todas as partições de $[a, b]$.

Note que se $\omega = 0$, então

$$|\omega(a)| + \mathrm{Var}(\omega) = 0 + 0 = 0 \Rightarrow \|\omega\| = 0.$$

Por outro lado, se $\|\omega\| = 0$, então

$$0 \geq -|\omega(a)| = \mathrm{Var}(\omega) = \sup \sum_{j=1}^{n} |\omega(t_j) - \omega(t_{j-1})|,$$

implicando que $\mathrm{Var}(\omega) = 0$. Logo, deve-se ter $\omega = 0$. Ainda, $\|\omega\| \geq 0$ pela própria definição de norma.

Também vale,

$$\begin{aligned} \|\alpha\omega\| = |\alpha\omega| + \mathrm{Var}(\alpha\omega) = & \ |\alpha||\omega(a)| + |\alpha|\mathrm{Var}(\omega) \\ = & \ |\alpha|(|\omega(a)| + \mathrm{Var}(\omega)) \\ = & \ |\alpha|\|\omega\|. \end{aligned}$$

Resta mostrar a desigualdade triangular. Para tanto, sejam

$$\Delta\omega_j = \omega(t_j) - \omega(t_{j-1}) \text{ e } \Delta\tilde{\omega}_j = \tilde{\omega}(t_j) - \tilde{\omega}(t_{j-1}).$$

Então, para qualquer partição P de $[a,b]$, tem-se

$$\sum_{j=1}^{n} |\Delta\omega_j - \Delta\tilde{\omega}_j| \leq \sum_{j=1}^{n} |\Delta\omega_j| + \sum_{j=1}^{n} |\Delta\tilde{\omega}_j|.$$

Logo,

$$\sup \sum_{j=1}^{n} |\Delta\omega_j - \Delta\tilde{\omega}_j| \leq \sup \sum_{j=1}^{n} |\Delta\omega_j| + \sup \sum_{j=1}^{n} |\Delta\tilde{\omega}_j|,$$

ou seja,

$$\mathrm{Var}(\omega + \tilde{\omega}) \leq \mathrm{Var}(\omega) + \mathrm{Var}(\tilde{\omega}).$$

Além disso, por desigualdade triangular, tem-se

$$|\omega(a) + \tilde{\omega}(a)| \leq |\omega(a)| + |\tilde{\omega}(a)|.$$

Assim,

$$\begin{aligned} \|\omega + \tilde{\omega}\| = |(\omega + \tilde{\omega})(a)| + \mathrm{Var}(\omega + \tilde{\omega}) = & \ |\omega(a) + \tilde{\omega}(a)| + \mathrm{Var}(\omega + \tilde{\omega}) \\ \leq & \ |\omega(a)| + |\tilde{\omega}(a)| + \mathrm{Var}(\omega) + \mathrm{Var}(\tilde{\omega}) \\ = & \ [|\omega(a)| + \mathrm{Var}(\omega)] + [|\tilde{\omega}(a)| + \mathrm{Var}(\tilde{\omega})] \\ = & \ \|\omega\| + \|\tilde{\omega}\|, \end{aligned}$$

portanto,

$$\|\omega + \tilde{\omega}\| \leq \|\omega\| + \|\tilde{\omega}\|.$$

O espaço das funções de variação limitada, com a norma definida em (6.16), é denotado por $\mathrm{BV}[a,b]$, onde BV está associado ao termo “bounded variation”, que quer dizer variação limitada.

Teorema 6.7. *Se α é monótona crescente em $[a,b]$, então α é de variação limitada em $[a,b]$ e $V(\alpha) = \alpha(b) - \alpha(a)$.*

Demonstração. Seja P qualquer partição do intervalo $[a,b]$, onde

$$a = x_0 < x_1 < \ldots < x_k = b.$$

Assim,

$$\begin{aligned}
\sum_{k=1}^{n} |\alpha(x_i) - \alpha(x_{i-1})| &= \sum_{k=1}^{n} [\alpha(x_i) - \alpha(x_{i-1})] \\
&= \alpha(x_1) - \alpha(x_0) + \alpha(x_2) - \alpha(x_1) + \ldots + \alpha(x_n) - \alpha(x_{n-1}) \\
&= \alpha(x_n) - \alpha(x_0) \\
&= \alpha(b) - \alpha(a).
\end{aligned}$$

Portanto, considerando o supremo sobre todas as possíveis partições do intervalo $[a,b]$, segue que

$$\mathrm{Var}(\alpha) = \alpha(b) - \alpha(a).$$

$\square$

Observação 6.5. O resultado permance válido se α for monótona decrescente e, neste caso, $\mathrm{Var}(\alpha) = \alpha(a) - \alpha(b)$.

Teorema 6.8. *A função f é de variação limitada em $[a,b]$ se, e somente se, é a diferença de duas funções monótonas crescentes.*

Demonstração. ($\Rightarrow$) Seja

$$\beta(t) = \mathrm{Var}(f,a,t), \ \ a \leq t \leq b. \tag{6.17}$$

Note que β é montótona crescente, pois a variação total de qualquer função de variação limitada sobre qualquer intervalo é não negativa.

Considere a função $\gamma(t) = \beta(t) - f(t)$. Resta mostrar que γ é monótona crescente. Com efeito, se $t_1 \leq t_2$, então

$$\gamma(t_2) - \gamma(t_1) = [\beta(t_2) - \beta(t_1)] - [f(t_2) - f(t_1)]. \tag{6.18}$$

Agora, $|f(t_2) - f(t_1)| \leq \beta(t_2) - \beta(t_1)$, pela própria definição de β. Logo, em (6.18) deve-se ter

$$\gamma(t_2) - \gamma(t_1) \geq 0 \ \Rightarrow \ \gamma(t_2) \geq \gamma(t_1),$$

resultando que γ é monótona crescente.

Portanto, de (6.17) segue que $f = \beta - \gamma$ é a diferença de duas funções monótonas crescentes.

($\Leftarrow$) Se $f = \beta - \gamma$, onde β e γ são monótonas crescentes, então pelo teorema 6.7 segue que β e γ são de variação limitada em $[a,b]$. Logo, como f é a diferença de duas funções de variação limitada em $[a,b]$, segue que f é de variação limitada em $[a,b]$ (ver observação (6.4)).

$\square$

6.1.2 O Teorema de Hahn-Banach e o dual de $\mathcal{C}([a,b],\mathbb{R})$

Nesta seção o Teorema de Hahn-Banach, muito importante em análise funcional, será utilizado para obter uma "formula" geral para representação de funcionais lineares limitados em $\mathcal{C}([a,b],\mathbb{R})$ e tal representação será em termos da integral de Riemann-Stieltjes.

O teorema abaixo, onde o espaço vetorial V é real, se refere a extensão de funcionais lineares de um subespaço $S \subset V$ em que o funcional f a ser estendido é majorado em S por um funcional sublinear p (definição 6.5) definido em V, de forma que a extensão $\tilde{f}$ em V ainda é linear e majorado por p. Antes de enunciar o teorema, segue a definição de funcional sublinear.

Definição 6.5. *Um* ***funcional sublinear*** *sobre um espaço vetorial* V *é uma aplicação* $p : V \to \mathbb{R}$ *tal que*

i. $p(x+y) \leq p(x) + p(y),\ \forall\ x, y \in V,$

ii. $p(\lambda x) = \lambda p(x),\ \forall\ x \in V\ e\ \lambda \geq 0 \in \mathbb{R}.$

Teorema 6.9. ***(Teorema de Hahn-Banach: extensão de funcionais lineares)*** *Seja* V *um espaço vetorial real e* p *um funcional sublinear em* V*. Seja* f *um funcional linear definido em um subespaço* S *de* V *satisfazendo*

$$f(x) \leq p(x),\ \forall\ x \in S. \tag{6.19}$$

Então f *tem uma extensão linear* $\tilde{f}$ *satisfazendo*

$$\tilde{f}(x) \leq p(x),\ \forall\ x \in V, \tag{6.20}$$

isto é, $\tilde{f}$ *é um funcional linear em* V*, satisfazendo* (6.20) *para todo* $x \in V$ *e* $\tilde{f}(x) = f(x)$ *para todo* $x \in S$*.*

Demonstração. Seja $\mathcal{F}$ o conjunto de todas as extensões g de f, com $D(g) \subset V$, satisfazendo $g(x) \leq p(x)$ para todo $x \in D(g)$. Note que $\mathcal{F} \neq \emptyset$. Com efeito, como $S \subset D(g)$, considere o funcional linear $h_0 : \overline{S} \to \mathbb{R}$ como segue

$$h_0(v) = \begin{cases} f(v), & \text{se } v \in S \\ \lim\limits_{n\to\infty} f(v_n), & \text{se } v \in \overline{S} \setminus S. \end{cases}$$

Pelo teorema 4.12 (página 93) (com $F = \mathbb{R}$), h_0 é uma extensão para f. Resta mostrar que $h_0(v) \leq p(v)$ para todo $v \in \overline{S}$. Como $f(x) \leq p(x)$ para todo $x \in S$, se $v \in S$, pela definição de h_0 tem-se

$$h_0(v) = f(v) \leq p(v) \Rightarrow h_0(v) \leq p(v).$$

Por outro lado, se $v \in \overline{S} \setminus S$, então

$$h_0(v) = \lim_{n \to \infty} f(v_n) \leq \lim_{n \to \infty} p(v_n) = p(v),$$

ou seja,

$$h_0(v) \leq p(v).$$

Portanto, $h_0 \in \mathcal{F}$. Assim, pode-se definir em $\mathcal{F}$ uma relação de ordem parcial por

$$g \preceq h \text{ (significa que } h \text{ é uma extensão de } g),$$

isto é, por definição, $D(g) \subset D(h)$ e $g(x) = h(x)$ para todo $x \in D(g)$. Agora, para qualquer conjunto totalmente ordenado $C \subset \mathcal{F}$, considere um funcional linear $\hat{g}$ tal que

$$\hat{g}(x) = g(x) \qquad \text{se } x \in D(g),$$

onde $g \in C$, cujo domínio de $\hat{g}$,

$$D(\hat{g}) = \bigcup_{g \in C} D(g),$$

é um espaço vetorial, pois C é um conjunto totalmente ordenado. Note que se $x \in D(g_1) \cap D(g_2)$, com $g_1, g_2 \in C$, então $g_1(x) = g_2(x)$, pois C é totalmente ordenado. Assim, $g_1 \preceq g_2$ ou $g_2 \preceq g_1$. Note que $g \preceq \hat{g}$, para todo $g \in C$. Logo, $\hat{g}$ é um limitante superior de C. Como $C \subset \mathcal{F}$, pelo Lema de Zorn (página 60), $\mathcal{F}$ tem um elemento maximal $\tilde{f}$. Portanto, pela definição de $\mathcal{F}$, $\tilde{f}$ é uma extensão de f satisfazendo

$$\tilde{f}(x) \leq p(x), \ \forall \ x \in D(\tilde{f}).$$

Para completar a demonstração, observe inicialmente que demonstrar que $\tilde{f}$ é definida sobre todo V, satisfazendo (6.20), é o mesmo que mostrar que $D(\tilde{f}) = V$. Suponha que isto seja falso, ou seja, $D(\tilde{f}) \neq V$. Neste caso, sejam $z \in V \setminus D(\tilde{f})$ e Y_0 o subespaço gerado por $D(\tilde{f})$ e z, ou seja, se $x \in Y_0$ então existem $y \in D(\tilde{f})$ e $\lambda \in \mathbb{R}$ tais que

$$x = y + \lambda z. \tag{6.21}$$

Esta representação é única. Com efeito, sejam $y, \tilde{y} \in D(\tilde{f})$ tais que

$$y + \lambda z = \tilde{y} + \beta z \Rightarrow (y - \tilde{y}) = (\beta - \lambda) z.$$

Note que $(y-\tilde{y}) \in D(\tilde{f})$, mas $z \notin D(\tilde{f})$. Logo, deve-se ter $y-\tilde{y}=0$ e $\beta-\lambda=0$, ou seja, $y=\tilde{y}$ e $\beta=\lambda$. Isto mostra que para cada $x \in Y$, a representação em (6.21) é única.

Considere agora a seguinte função

$$g_0(x) = \tilde{f}(y) + \lambda c, \tag{6.22}$$

onde c é alguma constante real. g_0 é um funcional linear, pois se $x_1, x_2 \in Y_0$, tais que $x_1 = y_1 + \lambda z$ e $x_2 = y_2 + \beta z$, então

$$\begin{aligned} g_0(x_1+\alpha x_2) = & \ \tilde{f}(y_1+\alpha y_2) + (\lambda+\alpha\beta)c \\ = & \ \tilde{f}(y_1) + \tilde{f}(\alpha y_2) + \lambda c + \alpha\beta c \\ = & \ (\tilde{f}(y_1) + \lambda c) + \alpha[\tilde{f}(y_2) + \beta c] \\ = & \ g_0(x_1) + \alpha g_0(x_2). \end{aligned}$$

Considerando $\lambda = 0$ em (6.22), tem-se $g_0(x) = \tilde{f}(x)$. Assim, g_0 é uma extensão própria de $\tilde{f}$, ou seja, uma extensão tal que $D(\tilde{f})$ é um subconjunto próprio de $D(g_0)$, ou ainda, $D(\tilde{f}) \subsetneq D(g_0)$. Neste caso, a ideia é provar que $g_0 \in \mathcal{F}$, ou seja, que $g_0(x) \leq p(x)$ para todo $x \in D(g_0)$, o que contradiz a maximalidade de $\tilde{f}$ e, assim, a hipótese de que $D(\tilde{f}) \neq V$ não é verdadeira.

Para isto, considere qualquer $y_1, y_2 \in D(\tilde{f})$. Como $\tilde{f}(x) \leq p(x)$ para todo $x \in D(\tilde{f})$ e sendo p um funcional sublinear, então

$$\begin{aligned} \tilde{f}(y_1) - \tilde{f}(y_2) = \tilde{f}(y_1 - y_2) \leq & \ p(y_1 - y_2) \\ = & \ p(y_1 + z - z - y_2) \\ \leq & \ p(y_1 + z) + p(-z - y_2), \end{aligned}$$

ou seja,

$$\tilde{f}(y_1) - \tilde{f}(y_2) \leq p(y_1 + y_0) + p(-y_0 - y_2),$$

logo,

$$-p(-z-y_2) - \tilde{f}(y_2) \leq p(y_1 + z) - \tilde{f}(y_1), \tag{6.23}$$

onde z é fixo. Observando a desigualdade acima, note que y_1 não aparece a esquerda e y_2 não aparece a direita da desigualdade. Neste caso, considerando o supremo sobre todo $y_2 \in D(\tilde{f})$ a esquerda e o ínfimo sobre $y_1 \in D(\tilde{f})$ a direita, a desigualdade permanece verdadeira, ou seja,

$$\sup_{y_2 \in D(\tilde{f})} [-p(-z-y_2) - \tilde{f}(y_2)] \leq \inf_{y_1 \in D(\tilde{f})} [p(y_1+z) - \tilde{f}(y_1)].$$

Assim, para algum c (que pode ser escolhido de forma conveniente) satisfazendo

$$\sup_{y_2 \in D(\tilde{f})} [-p(-z-y_2) - \tilde{f}(y_2)] \leq c \leq \inf_{y_1 \in D(\tilde{f})} [p(y_1+z) - \tilde{f}(y_1)],$$

tem-se de (6.23)

$$-p(-z-y_2) - \tilde{f}(y_2) \leq c \leq p(y_1+z) - \tilde{f}(y_1).$$

Provar que $g_0(x) \leq p(x)$ para todo $x \in D(g_0)$, onde $x = y + \lambda z$, equivale a mostrar que

$$\tilde{f}(y) + \lambda c \leq p(y + \lambda z), \tag{6.24}$$

para todo $y \in D(\tilde{f})$.

Note que se $\lambda = 0$, então (6.24) está satisfeita. Suponha $\lambda \neq 0$. Se $\lambda < 0$, considerando $y_2 = y/\lambda$, e como

$$-p(-z-y_2) - \tilde{f}(y_2) \leq c,$$

segue que

$$-p\left(-z - \frac{y}{\lambda}\right) - \tilde{f}\left(\frac{y}{\lambda}\right) \leq c.$$

Multiplicando esta desigualdade por $-\lambda > 0$, tem-se

$$\lambda p\left(-z - \frac{y}{\lambda}\right) + \lambda \tilde{f}\left(\frac{y}{\lambda}\right) \leq -\lambda c.$$

Da linearidade de $\tilde{f}$ e sendo p sublinear, pode-se escrever

$$p\left(-\lambda z - y\right) + \tilde{f}(y) \leq -\lambda c$$

ou seja,

$$\tilde{f}(y) + \lambda c \leq -p\left(-\lambda z - y\right) = p(\lambda z + y),$$

logo,

$$g_0(x) \leq p(x).$$

Se $\lambda > 0$, considerando $y_1 = y/\lambda$, e como

$$c \leq p(y_1 + z) - \tilde{f}(y_1),$$

segue que

$$c \leq p\left(\frac{y}{\lambda} + z\right) - \tilde{f}\left(\frac{y}{\lambda}\right).$$

Multiplicando esta desigualdade por $\lambda > 0$, tem-se

$$\lambda c \leq \lambda p\left(\frac{y}{\lambda} + z\right) - \lambda \tilde{f}\left(\frac{y}{\lambda}\right),$$

e como $\tilde{f}$ é linear e p sublinear, segue que

$$\lambda c \leq p(y + \lambda z) - \tilde{f}(y),$$

ou seja,

$$\tilde{f}(y) + \lambda c \leq p(y + \lambda z),$$

mas isto significa que

$$g_0(x) \leq p(x).$$

Logo, a hipótese de que $D(\tilde{f}) \neq V$ é falsa, pois isto contraria a maximalidade de $\tilde{f}$. Portanto, $D(\tilde{f}) = V$, ou seja, $\tilde{f}$ é definida sobre todo o espaço vetorial V. $\square$

O próximo teorema é uma generalização para o caso em que V é um espaço vetorial complexo (ver demonstração em [9]).

Teorema 6.10. ***(Teorema de Hahn-Banach generalizado)*** *Seja V um espaço vetorial complexo e $p : V \to \mathbb{R}$ um funcional sublinear tal que*

i. $p(x + y) \leq p(x) + p(y), \ \forall \ x, y \in V,$

ii. $p(\lambda x) = |\lambda| p(x), \ \forall \ x \in V$ *e* $\lambda \in \mathbb{K}.$

Seja f um funcional linear definido em um subespaço S de V satisfazendo

$$f(x) \leq p(x), \ \forall \ x \in S. \tag{6.25}$$

Então f tem uma extensão linear $\tilde{f}$ atendendo

$$\tilde{f}(x) \leq p(x), \ \forall \ x \in V. \tag{6.26}$$

Segue o Teorema de Hahn-Banach para espaços normados cujos funcionais lineares são limitados e, consequentemente, contínuos.

Teorema 6.11. ***(Teorema de Hahn-Banach para espaços normados)*** *Seja f um funcional linear limitado em um subespaço F de um espaço normado E. Então existe um funcional linear limitado $\tilde{f}$ que é extensão de f em E e tem a mesma norma,*

$$\|\tilde{f}\|_E = \|f\|_F, \tag{6.27}$$

onde

$$\|\tilde{f}\|_E = \sup_{\substack{x \in E \\ \|x\|=1}} |\tilde{f}(x)| \quad e \quad \|f\|_F = \sup_{\substack{x \in F \\ \|x\|=1}} |f(x)|.$$

($\|f\|_F = 0$ no caso trivial $F = \{0\}$)

Demonstração. Se $F = \{0\}$, então $f = 0$ e a extensão é $\tilde{f} = 0$, logo $\|\tilde{f}\|_E = \|f\|_F$. Suponha $F \neq \{0\}$. A ideia é utilizar o teorema 6.10. Inicialmente é preciso encontrar um funcional sublinear p. Como $f \in F$ é um funcional linear limitado, para todo $x \in F$ tem-se

$$|f(x)| \leq \|f\|_F \|x\|.$$

De (6.25) no teorema 6.10,

$$f(x) \leq p(x), \ \forall \ x \in F.$$

Considere o funcional sublinear p como sendo

$$p(x) = \|f\|_F \|x\|.$$

Note que p pode definir-se em todo E e, neste conjunto, utilizando desigualdade triangular tem-se

$$\begin{aligned} p(x+y) = \|f\|_F \|x+y\| \leq & \ \|f\|_F (\|x\| + \|y\|) \\ = & \ \|f\|_F \|x\| + \|f\|_F \|y\| \\ = & \ p(x) + p(y), \end{aligned}$$

para todo $x, y \in E$. Além disso,

$$p(\lambda x) = \|f\|_F \|\lambda x\| = |\lambda| \|f\|_F \|x\| = |\lambda| p(x),$$

para todo $x \in E$ e $\lambda \in \mathbb{K}$.

Então, pelo teorema 6.10 existe um funcional linear $\tilde{f}$ satisfazendo

$$|\tilde{f}(x)| \leq p(x) = \|f\|_F \|x\|.$$

Considerando o supremo sobre todo $x \in E$ de norma 1 e utilizando a desigualdade acima,

$$\|\tilde{f}\|_E = \sup_{\substack{x \in E \\ \|x\|=1}} |\tilde{f}(x)| \leq \|f\|_F,$$

ou seja,

$$\|\tilde{f}\|_E \leq \|f\|_F. \tag{6.28}$$

Por outro lado, como em uma extensão a norma não diminui (e neste caso é a norma do supremo), tem-se também

$$\|\tilde{f}\|_E \geq \|f\|_F. \tag{6.29}$$

Portanto, de (6.28) e (6.29) segue que

$$\|\tilde{f}\|_E = \|f\|_F.$$

$\square$

Observação 6.6. Quando o espaço normado E for de Hilbert, o teorema anterior torna-se simples. Com efeito, se F é um subespaço fechado de um espaço $E = \mathcal{H}$ de Hilbert, então, pelo Teorema da Representação de Riesz (página 148), o funcional $f : F \subset \mathcal{H} \to \mathbb{K}$ tem a representação

$$f(x) = \langle x, z \rangle, \text{ para todo } z \in F,$$

e $\|z\| = \|f\|$. Como o produto interno está definido em todo o espaço $\mathcal{H}$, isto fornece uma extensão $\tilde{f}$ de F em $\mathcal{H}$ com $\|\tilde{f}\| = \|z\| = \|f\|$. Note que no exemplo 5.6 (página 115), o espaço $\mathcal{C}([a,b],\mathbb{R})$ não é completo, então não é possível aplicar o Teorema da Representação de Riesz. Neste caso, o teorema abaixo mostra que a integral de Stieltjes pode ser usada para representar um funcional linear contínuo em $\mathcal{C}([a,b],\mathbb{R})$.

Teorema 6.12. ***(Teorema de Riesz)*** *Todo funcional linear limitado f em $\mathcal{C}([a,b],\mathbb{R})$ pode ser representado pela integral de Riemann-Stieltjes*

$$f(\phi) = \int_a^b \phi(t)\, d\omega(t), \tag{6.30}$$

onde $\phi \in \mathcal{C}([a,b],\mathbb{R})$, ω é uma função de variação limitada em $[a,b]$ e a variação de ω é igual a $\|f\|$, ou seja, $Var(\omega) = \|f\|$.

Demonstração. O espaço $\mathcal{C}([a,b],\mathbb{R})$ pode ser considerado como um subespaço do espaço $\mathcal{B}[a,b]$ das funções limitadas, cuja norma é definida por

$$\|\phi\| = \sup_{a \le t \le b} |\phi(t)|.$$

Seja f um funcional linear limitado em $\mathcal{C}([a,b],\mathbb{R})$. Pelo teorema 6.11 de Hahn-Banach, f tem uma extensão $\tilde{f}$ do espaço $\mathcal{C}([a,b],\mathbb{R})$ para o espaço normado $\mathcal{B}[a,b]$. Além disso, pelo mesmo teorema, o funcional linear $\tilde{f}$ é limitado e tem a mesma norma de f, ou seja, $\|\tilde{f}\| = \|f\|$.

Para definir a função ω, necessária em (6.30), considere a função $\phi_s \in \mathcal{B}[a,b]$, tal que $\phi_a(\xi) = 0$ para todo $\xi \in [a,b]$, e para $a < \xi \le b$ define-se

$$\phi_s(\xi) = \begin{cases} 1, & \text{se } a \le \xi \le s \\ 0, & \text{se } s < \xi \le b. \end{cases}$$

Utilizando ϕ_s e o funcional $\tilde{f}$, defina ω em $[a,b]$ por

$$\omega(t) = \begin{cases} 0, & \text{se } t = a \\ \tilde{f}(\phi_t), & \text{se } a < t \le b. \end{cases} \tag{6.31}$$

ω é de variação limitada em $[a,b]$, de fato, dada uma partição P qualquer de $[a,b]$,

$$a = t_0 < t_1 < \ldots < t_n = b,$$

seja $\varepsilon_j = \mathrm{sgn}[\omega(t_j) - \omega(t_{j-1})]$, onde

$$\mathrm{sgn}(t) = \begin{cases} -1, & \text{se } t < 0, \\ \ \ 0, & \text{se } t = 0, \\ \ \ 1, & \text{se } t > 0, \end{cases}$$

representa o sinal de t, com $t \in \mathbb{R}$.

Assim, como $\tilde{f}$ é um funcional linear limitado e $\|\tilde{f}\| = \|f\|$, segue da definição de ω que

$$\begin{aligned} \sum_{j=1}^{n} |\omega(t_j) - \omega(t_{j-1})| &= \sum_{j=1}^{n} \varepsilon_j [\omega(t_j) - \omega(t_{j-1})] \\ &= \sum_{j=1}^{n} \varepsilon_j [\tilde{f}(\phi_{t_j}) - \tilde{f}(\phi_{t_{j-1}})] \\ &= \tilde{f}\left(\sum_{j=1}^{n} \varepsilon_j [\phi_{t_j} - \phi_{t_{j-1}}] \right) \\ &\leq \|\tilde{f}\| \left\| \sum_{j=1}^{n} \varepsilon_j [\phi_{t_j} - \phi_{t_{j-1}}] \right\| \\ &= \|f\| \left\| \sum_{j=1}^{n} \varepsilon_j [\phi_{t_j} - \phi_{t_{j-1}}] \right\| \\ &= \|f\|, \end{aligned}$$

pois $\left\| \sum_{j=1}^{n} \varepsilon_j [\phi_{t_j} - \phi_{t_{j-1}}] \right\| = 1$ para todo $t \in [a, b]$. Considerando o supremo sobre todas as partições P de $[a, b]$ tem-se

$$\mathrm{Var}(\omega) \leq \|f\|, \tag{6.32}$$

ou seja, ω é de variação limitada em $[a, b]$.

Sejam $\phi \in \mathcal{C}([a, b], \mathbb{R})$ e considere P_n partição arbitrária de $[a, b]$,

$$P_n : a = t_0 < t_1 < \ldots < t_n = b.$$

Considere ϕ_n (ϕ_n depende de P_n) da seguinte forma

$$\phi_n = \sum_{j=1}^{n} \phi(t_{j-1})[\phi_{t_j} - \phi_{t_{j-1}})]. \tag{6.33}$$

Então $\phi_n \in \mathcal{B}[a,b]$. Pela definição de ω,

$$\begin{aligned} \tilde{f}(\phi_n) = & \sum_{j=1}^{n} \phi(t_{j-1})[\tilde{f}(\phi_{t_j}) - \tilde{f}(\phi_{t_{j-1}})] \\ = & \sum_{j=1}^{n} \phi(t_{j-1})[\omega(t_j) - \omega(t_{j-1})]. \end{aligned} \tag{6.34}$$

Agora, escolhendo qualquer sequência (P_n) de partições de $[a,b]$ com

$$\eta(P_n) = \max\{\Delta t_j\} \to 0,$$

segue que

$$\sum_{j=1}^{n} \phi(t_{j-1})[\omega(t_j) - \omega(t_{j-1})] \longrightarrow \int_a^b \phi(t)\, d\omega(t),$$

quando $n \to \infty$. Basta mostrar que $\tilde{f}(\phi_n) \to \tilde{f}(\phi)$, pois $\phi \in \mathcal{C}([a,b],\mathbb{R})$ e, neste espaço, $\tilde{f} = f$. De fato, pela definição de ϕ_t, considerando $t = a$, segue de (6.33) que

$$\phi_n(a) = \phi(a).1 + 0 \Rightarrow \phi_n(a) - \phi(a) = 0.$$

Além disso, se $t_{j-1} \leq t \leq t_j$, ainda por (6.33) tem-se

$$\phi_n(t) = \phi(t_{j-1}).1 + 0 \Rightarrow \phi_n(t) = \phi(t_{j-1}),$$

ou seja,

$$|\phi_n(t) - \phi(t)| = |\phi(t_{j-1}) - \phi(t)|,$$

para $t_{j-1} \leq t \leq t_j$. Agora, se $\eta(P_n) = \max\{\Delta t_j\} \to 0$, então $\|\phi_n - \phi\| \to 0$ pois ϕ é contínua em $[a,b]$. Como $[a,b]$ é compacto, segue que ϕ é uniformemente contínua e sendo $\tilde{f}$ contínua, segue que $\tilde{f}(\phi_n) \to \tilde{f}(\phi) = f(\phi)$. Portanto, obtem-se (6.30).

Finalmente, como

$$|f(\phi)| = \left| \int_a^b \phi(t)\, d\omega(t) \right| \leq \sup_{t\in[a,b]} |\phi(t)| \mathrm{Var}(\omega) = \|\phi\| \mathrm{Var}(\omega),$$

considerando o supremo sobre toda $\phi \in \mathcal{C}([a,b],\mathbb{R})$ de norma 1, obtem-se

$$\|f\| \leq \mathrm{Var}(\omega). \tag{6.35}$$

Portanto, de (6.32) e de (6.35), segue que

$$\mathrm{Var}(\omega) = \|f\|.$$

□

6.2 Teorema do Ponto Fixo de Banach

O Teorema do Ponto Fixo de Banach ou *Teorema da Contração* é um resultado sobre espaços métricos completos que garante a existência e a unicidade de um ponto que é mantido fixo por uma aplicação. Tem aplicações em várias áreas da Matemática pura e aplicada, por exemplo, é utilizado para demonstrar a existência de soluções em equações diferenciais.

Definição 6.6. *Sejam X um conjunto não vazio e $T : X \to X$ uma aplicação. Um ponto $x \in X$ é chamado* ***ponto fixo*** *de T se $T(x) = x$.*

Exemplo 6.2. Toda função contínua $f : [0,1] \to [0,1]$ possui um ponto fixo. Com efeito, considere a aplicação $g(x) = f(x) - x$, que é contínua em $[0,1]$. Note que $0 \leq f(x) \leq 1$ para todo $x \in [0,1]$. Logo,

$$g(0) = f(0) \geq 0 \quad \text{e} \quad g(1) = f(1) - 1 \leq 0.$$

Assim, pelo Teorema do Valor Intermediário (ver referência [10]), existe $x \in [0,1]$ tal que $g(x) = 0$, ou seja, $f(x) = x$. Portanto, x é ponto fixo de f.

Este resultado permance válido para $f : [a,b] \subset \mathbb{R} \to \mathbb{R}$ contínua, pois sendo $[a,b]$ um conjunto compacto, a imagem de $[a,b]$ em $\mathbb{R}$ é um conjunto compacto, e a demonstração segue de forma análoga.

No caso do Teorema do Ponto Fixo de Banach, uma condição necessária é que a aplicação envolvida seja uma contração, cuja definição segue abaixo.

Definição 6.7. *Sejam $M = (M,d)$ e $N = (N,\tilde{d})$ dois espaços métricos. Diz-se que uma aplicação $T : M \to N$ é uma* ***contração*** *quando existe uma constante real k, com $0 < k < 1$, tal que para todo $x, y \in M$,*

$$\tilde{d}(T(x), T(y)) \leq kd(x,y).$$

Observação 6.7. Toda contração entre espaços métricos é uniformemente contínua. Com efeito, se $T : M \to N$ é uma contração, então existe uma constante real positiva $k < 1$ tal que para todo $x, y \in M$ tem-se

$$\tilde{d}(T(x), T(y)) \leq kd(x,y).$$

Assim, dado $\epsilon > 0$ basta considerar $\delta = \epsilon/k$, tal que se $d(x,y) < \delta$ para quaisquer $x, y \in M$, então

$$\tilde{d}(T(x), T(y)) \leq kd(x,y) < k\delta = k\frac{\epsilon}{k} = \epsilon,$$

e, portanto, $\tilde{d}(T(x), T(y)) < \epsilon$.

Exemplo 6.3. Se a aplicação $f : I \to \mathbb{R}$ é diferenciável no intervalo aberto I e $|f'(x)| \leq k < 1$ para todo $x \in I$, então f é uma contração. De fato, sejam $x, y \in I$ com $x < y$ (análogo para $x > y$). Como f é contínua em $[x, y]$ e diferenciável em (x, y), pelo Teorema do Valor Médio (ver referência [10]), existe $z \in (x, y)$ tal que

$$f'(z) = \frac{f(y) - f(x)}{y - x} \Rightarrow f(y) - f(x) = f'(z)(y - x).$$

Portanto,

$$|f(x) - f(y)| = |f'(z)||(x - y)| \leq k|x - y|.$$

Teorema 6.13. ***(Teorema do Ponto Fixo de Banach)*** *Sejam $M \neq \emptyset$ um espaço métrico completo e $T : M \to M$ uma contração em M. Então T tem exatamente um ponto fixo.*

Demonstração. Deve-se provar que:

a) Para qualquer $x_0 \in M$, a “sequência iterativa” definida por

$$x_1 = T(x_0) \quad \text{e} \quad x_n = T^n(x_0),$$

converge para o ponto fixo $x \in M$.

b) x obtido em a) é único.

a) Seja $x_0 \in M$ qualquer e defina a sequência $(x_n)_{n\in\mathbb{N}}$ como segue

$$x_0, \quad x_1 = T(x_0), \quad x_2 = T(x_1) = T^2(x_0), \quad \ldots, \quad x_n = T^n(x_0), \quad \ldots, \tag{6.36}$$

de forma que T^n é a n-ésima composição de T consigo mesma no ponto x_0. A “sequência iterativa” (x_n) é uma sequência de Cauchy em M. De fato, sendo T uma contração, de (6.36) tem-se

$$\begin{aligned} d(x_{m+1}, x_m) = d(T(x_m), T(x_{m-1})) \leq\ & kd(x_m, x_{m-1}) \\ =\ & kd(T(x_{m-1}), T(x_{m-2})) \\ \leq\ & k[kd(x_{m-1}, x_{m-2})] = k^2 d(x_{m-1}, x_{m-2}) \\ =\ & k^2 d(T(x_{m-2}), T(x_{m-3})) \\ \leq\ & k^2[kd(x_{m-2}, x_{m-3})] = k^3 d(x_{m-2}, x_{m-3}) \\ \vdots\ & \\ \leq\ & k^m d(x_1, x_0), \ m \in \mathbb{N}. \end{aligned}$$

Agora, sejam $m, n \in \mathbb{N}$ com $m < n$. Utilizando desigualdade triangular segue que

$$\begin{aligned} d(x_m, x_n) \leq\ & d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \ldots + d(x_{n-1}, x_n) \\ \leq\ & k^m d(x_0, x_1) + k^{m+1} d(x_0, x_1) + \ldots + k^{n-1} d(x_0, x_1) \\ =\ & [k^m + k^{m+1} + \ldots + k^{n-1}] d(x_0, x_1), \end{aligned}$$

ou seja,

$$d(x_m, x_n) \leq [k^m + k^{m+1} + \ldots + k^{n-1}]d(x_0, x_1). \tag{6.37}$$

Multiplicando (6.37) por k $(0 < k < 1)$, obtem-se

$$kd(x_m, x_n) \leq [k^{m+1} + k^{m+2} \ldots + k^n]d(x_0, x_1). \tag{6.38}$$

Subtraindo (6.38) de (6.37) tem-se

$$d(x_m, x_n) - kd(x_m, x_n) \leq [k^m - k^n]d(x_0, x_1),$$

ou seja,

$$(1 - k)d(x_m, x_n) \leq k^m[1 - k^{n-m}]d(x_0, x_1),$$

e segue disto que

$$d(x_m, x_n) \leq k^m \frac{(1 - k^{n-m})}{1 - k} d(x_0, x_1).$$

Como $0 < k < 1$, note que $1 - k^{n-m} < 1$, logo

$$d(x_m, x_n) \leq \frac{k^m}{k - 1} d(x_0, x_1), \ \ m < n. \tag{6.39}$$

Note que k^m pode ser tão pequeno quanto se queira considerando m suficientemente grande, pois $0 < k < 1$. Além disso, $d(x_0, x_1)$ é uma quantidade fixa. Logo a sequência (x_m) é de Cauchy em M e, sendo M completo, existe $x \in M$ tal que $x_m \to x$ em M. Agora, este limite x é o ponto fixo da aplicação T. Com efeito, como T é uma contração, utilizando desigualdade triangular e (6.39) tem-se

$$\begin{aligned} d(x, T(x)) \leq d(x, x_m) + d(x_m, T(x)) = & \ \ d(x, x_m) + d(T(x_{m-1}), T(x)) \\ \leq & \ \ d(x, x_m) + kd(x_{m-1}, x), \end{aligned}$$

ou seja,

$$0 \leq d(x, T(x)) \leq d(x, x_m) + kd(x_{m-1}, x),$$

e neste caso, fazendo $m \to \infty$, segue que $d(x, T(x)) = 0$, pois $x_m \to x$. Portanto, $T(x) = x$, ou seja, x é ponto fixo de T.

b) Unicidade do ponto fixo: sejam $x, \tilde{x} \in M$ com $x \neq \tilde{x}$ e $T : M \to M$ uma contração. Se x e $\tilde{x}$ são pontos fixos de T, então $T(x) = x$ e $T(\tilde{x}) = \tilde{x}$. Logo, sendo T uma contração, existe uma constante real k $(0 < k < 1)$ tal que

$$d(x, \tilde{x}) = d(T(x), T(\tilde{x})) \leq kd(x, \tilde{x}),$$

o que é um absurdo já que $0 < k < 1$. Portanto, deve-se ter $d(x, \tilde{x}) = 0$, resultando em $x = \tilde{x}$. □

Uma das mais relevantes aplicações deste teorema se refere à solução de *Problemas de Valor Inicial* (PVI) em equações diferenciais ordinárias. O teorema abaixo, chamado Teorema de Picard, fornece condições suficientes para a existência e unicidade de tais problemas, como será demonstrado a seguir.

Teorema 6.14. ***(Teorema de Picard: Existência e unicidade de soluções)*** *Seja f um função contínua no retângulo fechado*

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

e, assim, limitada em R. Seja $c > 0$ tal que, para todo $(t,x) \in R$,

$$|f(t,x)| \le c.$$

Suponha que f satisfaz a ***condição de Lipschitz*** *em R com respeito ao segundo argumento, isto é, existe uma constante $\eta \ge 0$ (constante de Lipschitz) tal que para todos $(t,x),(t,y) \in R$ valha*

$$|f(t,x) - f(t,y)| \le \eta |x-y|. \tag{6.40}$$

Então, ao menos no intervalo fechado $J = [t_0 - \beta, t_0 + \beta]$, onde

$$\beta < \min\left\{a,\ \frac{b}{c},\ \frac{1}{\eta}\right\}, \tag{6.41}$$

o problema de valor inicial descrito pelas relações

$$x'(t) = f(t,x(t)) \text{ com } x(t_0) = x_0, \tag{6.42}$$

apresenta uma única solução.

Demonstração. Seja $\mathcal{C}(J,\mathbb{R})$ o espaço das funções contínuas definidas no intervalo J, em que $J = [t_0 - \beta, t_0 + \beta]$, com a métrica d definida por

$$d(x,y) = \sup_{t \in J} |x(t) - y(t)|,\ x, y \in \mathcal{C}(J,\mathbb{R}).$$

Note que $\mathcal{C}(J,\mathbb{R})$ é um espaço métrico completo (conforme exemplo 4.24, página 77). Seja $\tilde{\mathcal{C}}$ um subespaço de $\mathcal{C}(J,\mathbb{R})$ consistindo de todas as funções $x \in \mathcal{C}(J,\mathbb{R})$ satisfazendo

$$|x(t) - x_0| \le c\beta,\ \forall\, t \in J. \tag{6.43}$$

Logo, $\tilde{\mathcal{C}} \subset \mathcal{C}(J,\mathbb{R})$ é fechado. Para ver isto, seja (x_n) uma sequência convergente em $\tilde{\mathcal{C}}$, ou seja, existe $\tilde{x} \in \mathcal{C}(J,\mathbb{R})$ tal que $x_n \to \tilde{x}$, isto é, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que

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

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

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

Como $c\beta$ não depende de t, considerando o supremo sobre todo $t \in J$ segue que

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

Resta mostrar que $\tilde{x} \in \tilde{\mathcal{C}}$. Por desigualdade triangular

$$d(\tilde{x}, x_0) \leq d(\tilde{x}, x_n) + d(x_n, x_0) \leq \epsilon + c\beta, \ n > n_0.$$

Fazendo $\epsilon \to 0$, obtem-se

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

e segue disto que $\tilde{x} \in \tilde{\mathcal{C}}$. Como $\tilde{\mathcal{C}}$ é um subespaço fechado em $\mathcal{C}(J, \mathbb{R})$ que é completo, pelo teorema 4.6 (página 80), segue que $\tilde{\mathcal{C}}$ é completo.

Note que resolver o problema de valor inicial

$$x'(t) = f(t, x(t)) \text{ onde } x(t_0) = x_0,$$

é equivalente a resolver a equação integral

$$x(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau))\, d\tau, \ t \in J,$$

cuja demonstração segue do Teorema Fundamental do Cálculo.

Defina $T : \tilde{\mathcal{C}} \to \tilde{\mathcal{C}}$ por

$$(T(x))(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau))\, d\tau. \tag{6.44}$$

T está definida para todo $x \in \tilde{\mathcal{C}}$, pois $c\beta < b$ (por (6.41)). Assim se $x \in \tilde{\mathcal{C}}$, então $\tau \in J$ e $(\tau, x(\tau)) \in R$, e a integral em (6.44) está bem definida, pois f é contínua em R. Segue disto que T pode ser aplicada a toda função $x \in \tilde{\mathcal{C}}$. Como $|f(t, x)| \leq c$ segue que

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

satisfazendo (6.43), ou seja, T é uma aplicação de $\tilde{\mathcal{C}}$ em si mesmo.

Além disso, T é uma contração em $\tilde{\mathcal{C}}$. De fato, de (6.40) segue que

$$
\begin{aligned}
|(T(x))(t) - (T(y))(t)| = & \left| \int_{t_0}^{t} [f(\tau, x(\tau)) - f(\tau, y(\tau))] \, d\tau \right| \\
\leq & \int_{t_0}^{t} |f(\tau, x(\tau)) - f(\tau, y(\tau))| \, d\tau \\
\leq & \ \eta \int_{t_0}^{t} |x(\tau) - y(\tau)| \, d\tau \\
\leq & \ \eta |t - t_0| \sup_{\tau \in J} |x(\tau) - y(\tau)| \\
\leq & \ \eta\beta \, d(x, y).
\end{aligned}
$$

Como o lado direito da desigualdade não depende de t, considerando o supremo sobre $t \in J$ à esquerda, resulta que

$$d(T(x), T(y)) \leq kd(x, y), \text{ onde } k = \eta\beta.$$

De (6.41) tem-se $k = \eta\beta < 1$ e, portanto, T é uma contração em $\tilde{\mathcal{C}}$. Nestas condições, pelo Teorema do Ponto Fixo de Banach 6.13, T tem um único ponto fixo $x \in \tilde{\mathcal{C}}$, ou seja, uma função x contínua em J satisfazendo $x(t) = (T(x))(t)$ para todo $t \in J$. Logo, de (6.44), tem-se

$$x(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau)) \, d\tau. \tag{6.45}$$

Como $(\tau, x(\tau)) \in R$, onde f é contínua, segue que (6.45) é diferenciável, logo satisfaz (6.42). Por outro lado, cada solução em (6.42) tem que satisfazer (6.45). $\square$

Observação 6.8. Note que, utilizando o Teorema do Ponto fixo de Banach, é possível obter um processo iterativo chamado **Método de Picard** para encontrar a solução de um problema de valor inicial, em que a solução é o limite da sequência $(x_0, x_1, x_2, \ldots)$ obtida pela iteração

$$x_{n+1}(t) = x_0 + \int_{t_0}^{t} f(\tau, x_n(\tau)) \, d\tau,$$

com $n = 0, 1, 2, \ldots$.

Por exemplo, considere o seguinte PVI:

$$\begin{cases} x'(t) = & x \\ x(0) = & 1 \end{cases}$$

Note que $f(t, x(t)) = x$, $t_0 = 0$ e $x_0 = 1$. Utilizando o Método de Picard obtem-se

$$\begin{aligned}
x_1(t) = &\ 1 + \int_0^t f(\tau, x_0(\tau))\, d\tau = 1 + \int_0^t 1\, d\tau = 1 + t \\
x_2(t) = &\ 1 + \int_0^t f(\tau, x_1(\tau))\, d\tau = 1 + \int_0^t (1+\tau)\, d\tau = 1 + t + \frac{t^2}{2!} \\
x_3(t) = &\ 1 + \int_0^t f(\tau, x_2(\tau))\, d\tau = 1 + \int_0^t \left(1 + \tau + \frac{\tau^2}{2!}\right) d\tau = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} \\
\vdots & \\
x_{n+1}(t) = &\ 1 + \int_0^t f(\tau, x_n(\tau))\, d\tau = 1 + \int_0^t \left(1 + \tau + \frac{\tau^2}{2!} + \ldots + \frac{\tau^n}{n!}\right) d\tau \\
= &\ 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \ldots + \frac{t^{n+1}}{(n+1)!}.
\end{aligned}$$

Como $e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \ldots + \frac{t^{n+1}}{(n+1)!} + \ldots$, segue que as iteradas de Picard $x_n(t)$ convergem para a solução $x(t) = e^t$.

6.3 Convergência em $\mathcal{C}([a,b], \mathbb{R})$

Nesta seção pretende-se discutir a convergência de uma sequência de melhores aproximações em $\mathcal{C}([a,b], \mathbb{R})$ com relação a norma

$$\|f\| = \langle f, f\rangle^{\frac{1}{2}} = \left(\int_a^b [f(t)]^2\, dt\right)^{\frac{1}{2}},$$

que pelo exemplo 5.6 (página 115) não é completo.

A convergência com respeito a esta norma é normalmente chamada de *convergência na média.*

Definição 6.8. *Seja (f_n) uma sequência de funções em $\mathcal{C}([a,b], \mathbb{R})$. Então (f_n) é dita* ***convergir na média*** *para uma função $f \in \mathcal{C}([a,b], \mathbb{R})$ se*

$$d(f_n, f) = \|f_n - f\|_2 = \left(\int_a^b [f_n(t) - f(t)]^2\, dt\right)^{\frac{1}{2}} \longrightarrow 0 \text{ quando } n \to \infty.$$

Lema 6.2. *Se (f_n) é uma sequência de funções contínuas que converge uniformemente, então (f_n) também converge na média.*

Demonstração. Se (f_n) converge uniformemente em $\mathcal{C}([a,b], \mathbb{R})$ para uma função contínua $f \in \mathcal{C}([a,b], \mathbb{R})$, então dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ (n_0 depende apenas de ϵ) tal que

$$|f_n(t) - f(t)| < \frac{\epsilon}{\sqrt{b-a}},$$

para todo $n > n_0$ e todo $t \in [a,b]$.

Segue disto que

$$\begin{aligned}
\|f_n - f\|_2 = \left(\int_a^b [f_n(t) - f(t)]^2\, dt\right)^{\frac{1}{2}} < \left(\int_a^b \frac{\epsilon^2}{b-a}\, dt\right)^{\frac{1}{2}} &= \frac{\epsilon}{\sqrt{b-a}}\left(\int_a^b dt\right)^{\frac{1}{2}} \\
&= \frac{\epsilon}{\sqrt{b-a}}\left(t\Big|_a^b\right)^{\frac{1}{2}} \\
&= \frac{\epsilon}{\sqrt{b-a}}\sqrt{b-a} \\
&= \epsilon,
\end{aligned}$$

ou seja,

$$d(f_n, f) = \|f_n - f\|_2 \longrightarrow 0 \text{ quando } n \to \infty.$$

□

O teorema abaixo afirma que toda função real contínua em um intervalo compacto pode ser aproximada uniformemente por polinômios. A sua demonstração fará uso do chamado *polinômio de Bernstein.*

Definição 6.9. *Dada uma função $f : [0,1] \to \mathbb{R}$ o n-ésimo* ***polinômio de Bernstein*** *associado a f é definido por*

$$\mathrm{B}_n^f(x) = \sum_{k=1}^{n} \begin{pmatrix} n \\ k \end{pmatrix} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right), \tag{6.46}$$

onde o coeficiente binomial é dado por

$$\begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{k!(n-k)!},\ n \in \mathbb{N},\ 0 \leq k \leq n.$$

Teorema 6.15. ***(Teorema da Aproximação de Weierstrass)*** *Seja $f \in \mathcal{C}([a,b],\mathbb{R})$, então dado $\epsilon > 0$ existe $n \in \mathbb{N}$ e um polinômio $p \in P_n$,*

$$P_n = \left\{ p \in \mathcal{C}([a,b],\mathbb{R});\ p(x) = \sum_{k=1}^{n} c_k x^k \right\},$$

tal que $\|f - p\| < \epsilon$, ou seja, p converge uniformemente para f em $[a,b]$.

Demonstração. Suponha, sem perca de generalidade, que $[a,b] = [0,1]$[1]. A ideia é demosntrar que dado $\epsilon > 0$ fixo, existe $n \in \mathbb{N}$ tal que $|\mathrm{B}_n^f(x) - f(x)| < \epsilon$ para todo $x \in [0,1]$.

[1]ver referência [16].

Um caso particular do **Teorema do Binômio de Newton** é dado por

$$\sum_{k=1}^{n} \binom{n}{k} x^k(1-x)^{n-k} = [x+(1-x)]^n = 1^n = 1. \tag{6.47}$$

Derivando (6.47) com respeito a x, obtem-se

$$\sum_{k=1}^{n} \binom{n}{k} (kx^{k-1}(1-x)^{n-k} + x^k[(n-k)(1-x)^{n-k-1}(-1)]) = 0 \Rightarrow$$

$$\sum_{k=1}^{n} \binom{n}{k} (kx^{k-1}(1-x)^{n-k} + x^k[(k-n)(1-x)^{n-k-1}]) = 0 \Rightarrow$$

$$\sum_{k=1}^{n} \binom{n}{k} x^{k-1}(1-x)^{n-k-1}[k(1-x) + x(k-n)] = 0 \Rightarrow$$

$$\sum_{k=1}^{n} \binom{n}{k} x^{k-1}(1-x)^{n-k-1}[k - kx + kx - nx] = 0 \Rightarrow$$

$$\sum_{k=1}^{n} \binom{n}{k} x^{k-1}(1-x)^{n-k-1}[k - nx] = 0,$$

e multiplicando esta última igualdade por $x(1-x)$, segue que

$$\sum_{k=1}^{n} \binom{n}{k} x^{k}(1-x)^{n-k}[k - nx] = 0.$$

Derivando novamente em relação a x e aplicando (6.47), obtem-se

$$\sum_{k=1}^{n} \binom{n}{k} (x^{k-1}(1-x)^{n-k-1}[k-nx].[k-nx] + x^k(1-x)^{n-k}(-n)) = 0 \Rightarrow$$

$$\sum_{k=1}^{n} \binom{n}{k} (x^{k-1}(1-x)^{n-k-1}[k-nx]^2 - nx^k(1-x)^{n-k}) = 0 \Rightarrow$$

$$\sum_{k=1}^{n} \binom{n}{k} x^{k-1}(1-x)^{n-k-1}[k-nx]^2 = n\sum_{k=1}^{n} \binom{n}{k} x^k(1-x)^{n-k} = n,$$

ou seja,

$$\sum_{k=1}^{n} \binom{n}{k} x^{k-1}(1-x)^{n-k-1}[k-nx]^2 = n,$$

e multiplicando por $\dfrac{x(1-x)}{n^2}$, resulta que

$$\sum_{k=1}^{n}\binom{n}{k}x^k(1-x)^{n-k}\left[\frac{k}{n}-x\right]^2=\frac{x(1-x)}{n}. \tag{6.48}$$

Utilizando (6.47) e (6.46), note que

$$\begin{aligned} f(x)-\mathrm{B}_n^f(x) = {} & f(x)\sum_{k=1}^{n}\binom{n}{k}x^k(1-x)^{n-k}-\sum_{k=1}^{n}\binom{n}{k}x^k(1-x)^{n-k}f\left(\frac{k}{n}\right) \\ = {} & \sum_{k=1}^{n}\binom{n}{k}x^k(1-x)^{n-k}\left[f(x)-f\left(\frac{k}{n}\right)\right], \end{aligned}$$

assim,

$$|f(x)-\mathrm{B}_n^f(x)|\leq\sum_{k=1}^{n}\binom{n}{k}x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right|. \tag{6.49}$$

Como f é contínua no conjunto compacto $[0,1]$, pelo teorema 2.7, segue que f é uniformemente contínua em $[0,1]$, logo dado $\epsilon>0$ existe $\delta>0$ (não dependendo de x) tal que

$$\left|x-\frac{k}{n}\right|<\delta\Rightarrow\left|f(x)-f\left(\frac{k}{n}\right)\right|<\frac{\epsilon}{2}, \tag{6.50}$$

para todo $x,k/n\in[0,1]$.

Agora, para todo $x\in[0,1]$ considere os conjuntos N' e N'' como seguem

$$N'=\left\{k\in\mathbb{N},\ 0\leq k\leq n;\ \left|x-\frac{k}{n}\right|<\delta\right\},$$

$$N''=\left\{k\in\mathbb{N},\ 0\leq k\leq n;\ \left|x-\frac{k}{n}\right|\geq\delta\right\}.$$

Separe a soma à direita de (6.49) em duas partes fazendo

$$\begin{aligned} |f(x)-\mathrm{B}_n^f(x)|\leq {} & \sum_{k=1}^{n}\binom{n}{k}x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right| \\ = {} & \sum_{k\in N'}\binom{n}{k}x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right| \\ + {} & \sum_{k\in N''}\binom{n}{k}x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right|. \end{aligned} \tag{6.51}$$

Assim, utilizando (6.50) e (6.47) para a soma em N', segue que

$$\begin{aligned}\sum_{k\in N'}\binom{n}{k}x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right| &< \frac{\epsilon}{2}\sum_{k\in N'}\binom{n}{k}x^k(1-x)^{n-k}\\ &\leq \frac{\epsilon}{2}\sum_{k=1}^{n}\binom{n}{k}x^k(1-x)^{n-k}\\ &= \frac{\epsilon}{2}.\end{aligned} \tag{6.52}$$

Como f é limitada, pois $[0,1]$ é compacto, existe um número real $M>0$ tal que $|f(x)|\leq M$ para todo $x\in[0,1]$. Logo, para a soma em N'' tem-se

$$\begin{aligned}\sum_{k\in N''}\binom{n}{k}x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right| &\leq 2M\sum_{k\in N''}\binom{n}{k}x^k(1-x)^{n-k}\\ &= 2M\left(\frac{\delta^2}{\delta^2}\right)\sum_{k\in N''}\binom{n}{k}x^k(1-x)^{n-k}\\ &= \frac{2M}{\delta^2}\sum_{k\in N''}\binom{n}{k}x^k(1-x)^{n-k}\,\delta^2,\end{aligned}$$

e como $\delta\leq\left|x-\dfrac{k}{n}\right|$, então $\delta^2\leq\left[x-\dfrac{k}{n}\right]^2$. Segue disto e de (6.48) que

$$\begin{aligned}\sum_{k\in N''}\binom{n}{k}x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right| &\leq \frac{2M}{\delta^2}\sum_{k\in N''}\binom{n}{k}x^k(1-x)^{n-k}\left[x-\frac{k}{n}\right]^2\\ &= \frac{2M}{\delta^2}\frac{x(1-x)}{n}.\end{aligned}$$

Note que quando $x=1/2$, $x(1-x)$ atinge seu valor máximo $1/4$. Logo,

$$\sum_{k\in N''}\binom{n}{k}x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right|\leq\frac{2M}{\delta^2}\frac{1}{4n}=\frac{M}{2\delta^2 n}.$$

Considerando $n\in\mathbb{N}$ tal que $n>\dfrac{M}{\delta^2\epsilon}$, então

$$\sum_{k\in N''}\binom{n}{k}x^k(1-x)^{n-k}\left|f(x)-f\left(\frac{k}{n}\right)\right|\leq\frac{M}{2\delta^2 n}<\frac{M}{2\delta^2\dfrac{M}{\delta^2\epsilon}}=\frac{M\epsilon}{2M}=\frac{\epsilon}{2}. \tag{6.53}$$

Portanto, utilizando (6.52) e (6.53) em (6.51) deve-se ter

$$|f(x) - \mathrm{B}_n^f(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} < \epsilon,$$

para todo $x \in [0,1]$, ou seja, a sequência (B_n^f) converge uniformemente para a função f em $[0,1]$. □

Teorema 6.16. ***(Teorema da Convergência)*** *Sejam $f \in \mathcal{C}([a,b],\mathbb{R})$ e $(p_j)_{j\in\mathbb{N}}$ uma sequência formada pelos elementos de P_j, onde cada $p_j \in P_j$ é a melhor aproximação de f com relação a norma $\|\,.\,\|_2$. Então p_j converge para f na média.*

Demonstração. Pelo Teorema da Aproximação de Weierstrass 6.15, existe uma sequência $(q_n)_{n\in\mathbb{N}}$ de polinômios $q_n \in P_n$, convergindo uniformemente para f. Pelo lema 6.2, a convergência uniforme implica a convergência desta sequência na média, ou seja,

$$\|f - q_n\|_2 \to 0 \text{ quando } n \to \infty.$$

Sendo $(p_j)_{j\in\mathbb{N}}$ uma sequência de melhores aproximações de f, deve-se ter

$$\|f - p_j\|_2 \leq \|f - q_n\|_2,$$

e, portanto,

$$\|f - p_j\|_2 \to 0 \text{ quando } j \to \infty.$$

□

6.4 Aproximação de funções seccionalmente contínuas

Definição 6.10. *Diz-se que uma função f definida em $[a,b]$ é* ***seccionalmente contínua*** *se existir uma partição $P = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$ tal que f é contínua em cada subintervalo (x_{i-1}, x_i), possuindo limites laterais finitos nos pontos da partição.*

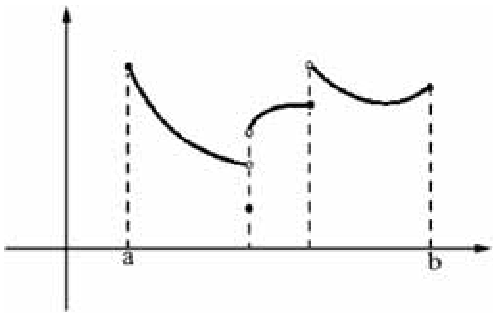

Figura 6.1: *Função seccionalmente contínua*

Denote por $\mathcal{C}_{-1}([a,b],\mathbb{R})$ o espaço de todas as funções seccionalmente contínuas no intervalo $[a,b]$. Sejam $f, g \in \mathcal{C}_{-1}([a,b],\mathbb{R})$ e a partição

$$a \leq \xi_0 < \xi_1 < \ldots < \xi_{m-1} < \xi_m \leq b,$$

em que $\xi_1, \ldots, \xi_{m-1}$ são os pontos onde a função produto $f.g$ apresenta saltos, ou seja, são os pontos de descontinuidade. Defina o seguinte produto interno

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,dx = \sum_{j=0}^{m-1} \int_{\xi_j}^{\xi_{j+1}} f(x)g(x)\,dx, \tag{6.54}$$

onde a norma é dada por

$$\|f\|_2 = \langle f, f \rangle^{\frac{1}{2}} = \left(\sum_{j=0}^{m-1} \int_{\xi_j}^{\xi_{j+1}} [f(x)]^2\,dx \right)^{\frac{1}{2}}$$

Note que $V = \mathcal{C}_{-1}([a,b],\mathbb{R})$ define um espaço vetorial com o produto interno dado por (6.54). Seja W qualquer subespaço vetorial de V de dimensão finita, neste caso sabe-se que dada uma função $f \in V$, existe uma única melhor aproximação $\tilde{f}$ de f em W. Isto motiva o seguinte resultado.

Teorema 6.17. *Seja $f \in \mathcal{C}_{-1}([a,b],\mathbb{R})$. Então a sequência $(p_k)_{k\in\mathbb{N}}$ de melhores aproximações $p_k \in P_k$ converge para f na média.*

Demonstração. A ideia é aproximar a função f por uma função contínua h, considerando suas melhores aproximações.

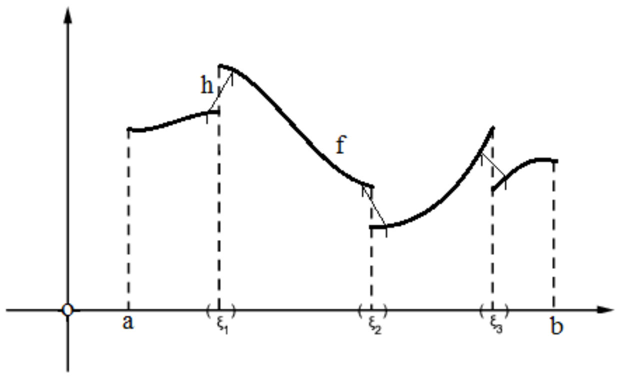

Figura 6.2: *Ideia geométrica para a demonstração do teorema*

Dada $f \in \mathcal{C}_{-1}([a,b],\mathbb{R})$ com saltos em $\xi_1, \ldots, \xi_{m-1}$, construa a função contínua h como segue

$$h(x) = \begin{cases} f(\xi_j - \delta) + \dfrac{f(\xi_j + \delta) - f(\xi_j - \delta)}{2\delta}[x - (\xi_j - \delta)], & \text{se } x \in [\xi_j - \delta, \xi_j + \delta], \\ & 1 \leq j \leq m - 1, \\ \\ f(x), & \text{caso contrário,} \end{cases}$$

onde $\delta \leq \dfrac{1}{2} \min\limits_{0 \leq j \leq m-1} (\xi_{j+1} - \xi_j)$.

Seja $(q_k)_{k \in \mathbb{N}}$ a sequência de melhores aproximações $q_k \in P_k$ de h. Neste caso, como $q_k \to h$ quando $k \to \infty$, existe $k_0 \in \mathbb{N}$ tal que

$$\|h - q_k\|_2 < \frac{\epsilon}{2} \text{ para todo } k > k_0. \tag{6.55}$$

Além disso, por desigualdade triangular

$$\|f - q_k\|_2 = \|f - h + h - q_k\|_2 \leq \|f - h\|_2 + \|h - q_k\|_2. \tag{6.56}$$

Agora,

$$\|f - h\|_2^2 = \sum_{j=0}^{m-1} \int_{\xi_j}^{\xi_{j+1}} [f(x) - h(x)]^2 \, dx = \sum_{j=1}^{m-1} \int_{\xi_j - \delta}^{\xi_j + \delta} [f(x) - h(x)]^2 \, dx. \tag{6.57}$$

Seja $M = \max\limits_{x \in [a,b]} |f(x)|$. Pela definição de h e por desigualdade triangular tem-se

$$|h(x) - f(x)| \leq |h(x)| + |f(x)| \leq 2M, \; \forall \, x \in [a, b],$$

independente da escolha do δ. Substituindo esta desigualdade em (6.57), obtem-se

$$\begin{aligned} \|f - h\|_2^2 \leq \sum_{j=1}^{m-1} \int_{\xi_j - \delta}^{\xi_j + \delta} [2M]^2 \, dx = & \; 4M^2 \left[\sum_{j=1}^{m-1} \int_{\xi_j - \delta}^{\xi_j + \delta} dx \right] \\ = & \; 4M^2 \underbrace{\left[\int_{\xi_1 - \delta}^{\xi_1 + \delta} dx + \int_{\xi_2 - \delta}^{\xi_2 + \delta} dx + \ldots + \int_{\xi_{m-1} - \delta}^{\xi_{m-1} + \delta} dx \right]}_{(m-1) \text{ vezes}} \\ = & \; 4M^2 \underbrace{[(\xi_1 + \delta) - (\xi_1 - \delta) + \ldots + (\xi_{m-1} + \delta) - (\xi_{m-1} - \delta)]}_{(m-1) \text{ vezes}} \\ = & \; 4M^2 \underbrace{[2\delta + \ldots + 2\delta]}_{(m-1) \text{ vezes}} = 4M^2(m-1)2\delta, \end{aligned}$$

ou seja,

$$\|f-h\|_2^2 \leq 8M^2(m-1)\delta.$$

Dado $\epsilon > 0$ considere $\delta < \dfrac{\epsilon^2}{32M^2(m-1)}$, logo na desigualdade acima tem-se

$$\|f-h\|_2 \leq (8M^2(m-1)\delta)^{\frac{1}{2}} < \left(8M^2(m-1)\frac{\epsilon^2}{32M^2(m-1)}\right)^{\frac{1}{2}} = \left(\frac{\epsilon^2}{4}\right)^{\frac{1}{2}} = \frac{\epsilon}{2},$$

isto é,

$$\|f-h\|_2 < \frac{\epsilon}{2}. \tag{6.58}$$

Substituindo (6.55) e (6.58) em (6.56), obtem-se

$$\|f-q_k\|_2 \leq \|f-h\|_2 + \|h-q_k\|_2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

para todo $k > k_0$.

Uma vez que a melhor aproximação $p_k \in P_k$ está mais próxima de f do que a melhor aproximação $q_k \in P_k$ de h, deve-se ter

$$\|f-p_k\|_2 \leq \|f-q_k\|_2 < \epsilon,$$

e, portanto, p_k converge para f na média. □

6.5 Aproximação de funções contínuas por escalonadas

Nesta seção pretende-se mostrar um resultado que afirma ser possível aproximar uma função contínua por uma que seja escalonada, o que em algumas situações pode ser vantajoso, por exemplo, na realização de cálculos.

Definição 6.11. *Uma função $g_e : [a,b] \to \mathbb{R}^n$ é chamada **função escalonada** se existe uma decomposição D de $[a,b]$,*

$$a = t_0 \leq t_1 \leq \ldots \leq t_m \leq b,$$

tal que a cada intervalo aberto (t_{k-1}, t_k), $k = 1, 2, \ldots, m$, g_e toma um valor constante c_k.

Exemplo 6.4. A função $g_e : [-2,2] \to \mathbb{R}$ que a cada $x \in [-2,2]$ associa o inteiro n tal que $n \leq x < n+1$, é uma função escalonada como mostra a figura 6.3

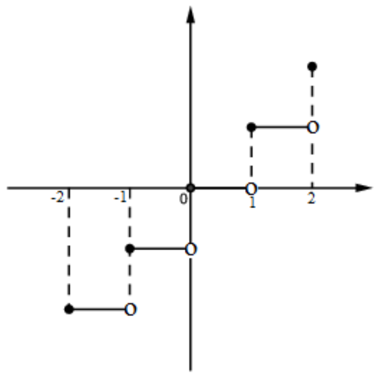

Figura 6.3: *Função escalonada*

Proposição 6.1. *O espaço das funções escalonadas $g_e : [a, b] \to \mathbb{R}^n$ é denso no espaço das funções contínuas.*

Demonstração. Note que é preciso demonstrar que dada uma função contínua f, existe uma função escalonada g_e convergindo uniformemente para f para todo $x \in [a, b]$.

Seja $f : [a, b] \to \mathbb{R}^n$ uma função uniformemente contínua. Assim, dado $\epsilon > 0$ existe $\delta > 0$ tal que

$$\|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \epsilon.$$

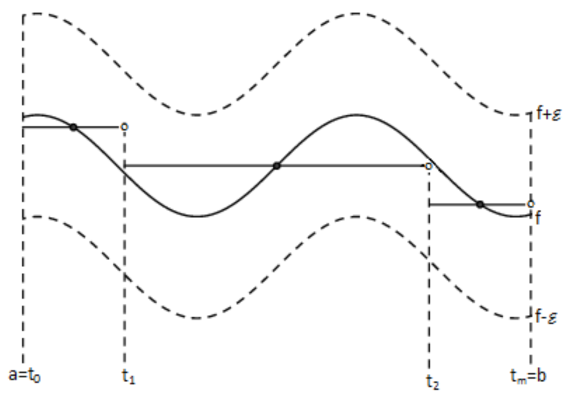

Figura 6.4: *Ideia geométrica para a demonstração da proposição*

Considere uma decomposição D do intervalo $[a, b]$ de forma que $\max\{\Delta t_k\} < \delta$, $k = 1, 2, \ldots, m$. Defina $g_e : [a, b] \to \mathbb{R}^n$ por meio de suas restrições a cada intervalo semiaberto $[t_{k-1}, t_k)$ fazendo

$$g_e(x) = \begin{cases} f(c_k), & \text{para algum } c_k \in [t_{k-1}, t_k),\ \forall\, x \in [t_{k-1}, t_k), \\ f(b), & \text{se } x = b. \end{cases}$$

Note que, pela definição da g_e, se $x \neq b$ em $[a, b]$, existe um único k tal que se

$x \in [t_{k-1}, t_k)$ então $\|x - c_k\| < \delta$. Logo, pela continuidade da f,

$$\|f(x) - g_e(x)\| = \|f(x) - f(c_k)\| < \epsilon,$$

ou seja,

$$\|x - c_k\| < \delta \Rightarrow \|f(x) - f(c_k)\| < \epsilon,$$

concluindo a demonstração. $\square$

7 Comentários finais

Durante o desenvolvimento deste trabalho foi possível observar que a Álgebra Linear é um pré-requisito indispensável quando o assunto é Análise Funcional. Na disciplina de Álgebra Linear apresentada aos cursos de graduação são estudados os espaços vetoriais de dimensão finita. No presente trabalho, alguns resultados conhecidos destes espaços puderam ser provados para espaços de dimensão infinita, utilizando novas ferramentas como, por exemplo, o Lema de Zorn.

Além disso, é possível observar que vários resultados envolveram propriedades topológicas, principalmente no que diz respeito à norma e convergência. Em particular, o estudo de convergência envolvendo sequência de Cauchy foi de fundamental importância.

Finalmente, foi possível constatar que a Análise Funcional é um ramo da Matemática que possui muitas aplicações e é de fundamental importância na obtenção de novos resultados da Análise como, por exemplo, a demonstração do Teorema da Existência e Unicidade de Soluções para equações diferenciais ordinárias (EDO).

Enfim, neste trabalho foi possível apresentar algumas ferramentas da Análise Funcional relacionando as mesmas com várias áreas da Matemática, apresentando um texto que envolve resultados fundamentais e aplicações indispensáveis no estudo da mesma.

Referências

[1] BARRETO, A. C. *Tópicos de Análise*. Rio de Janeiro: IMPA, 1971.

[2] BIEZUNER, R. J. *Análise Funcional* - Notas, Minas Gerais, 2009, 124 pag.

[3] CARVALHO, A. N. *Análise I* - Notas, São Carlos, 2005, 171 pag.

[4] COELHO, F. U. *Um Curso de Álgebra Linear*. 2. ed. São Paulo: Editora da Universidade de São Paulo, 2010.

[5] HOFFMAN, K.; KUNZE, R. *Linear Algebra*. Second edition. New Jersey: Prentice-Hall, Inc., 1971.

[6] HÄMMERLIN, G.; HOFFMANN, K.-H. *Numerical Mathematics*. Translate by: Larry Schumaker. New York: Springer-Verlag, 1991.

[7] HÖNIG, C. S. *Análise Funcional e Aplicações*. São Paulo: IME - USP, 1970.

[8] KOLMOGOROV, A. N.; FOMIN, S. V. *Introductory Real Analysis*. Translate and edited: Richard A. Silverman. New York: Dover Publications, Inc., 1975.

[9] KREYSZIG, E. *Introductory Functional Analysis with Applications*. New York: Wiley, 1989.

[10] LIMA, E. L. *Análise Real - Volume 1*. 10. ed. Rio de Janeiro: IMPA, 2009.

[11] LIMA, E. L. *Curso de Análise - Volume 1*. 11. ed. Rio de Janeiro: IMPA, 2004.

[12] LIMA, E. L. *Espaços Métricos*. 4. ed. Rio de Janeiro: IMPA, 2011.

[13] OLIVEIRA, C. R. *Introdução à Análise Funcional*. 2. ed. Rio de Janeiro: IMPA, 2009.

[14] ROYDEN, H. L. *Real Analysis*. Second edition. New York: Macmilan Publishing CO., 1968.

[15] RUDIN, W. *Princípios de Análise Matemática*. Tradução: Eliana Rocha Henrique de Brito. Rio de Janeiro: Ao Livro Técnico S.A., 1971.

[16] SIMMONS, G. F. *Introduction to Topology and Modern Analysis*. New York: McGraw-Hill, 1963.

A Funções Lebesgue mensuráveis

Os conceitos aqui introduzidos tem como referência a bibliografia [14]. Alguns resultados serão demonstrados, os que não forem podem ser encontrados na mesma referência. Para definir a integral de Lebesgue é preciso definir a teoria de funções Lebesgue mensuráveis. Serão descritos os resultados necessários para introduzir a integral de Lebesgue.

A.1 Os números reais estendidos

Será introduzida a noção de reta estendida e algumas operações definidas neste conjunto.

Definição A.1. *Um sistema de* ***números reais estendidos*** *consiste da reta real* $\mathbb{R}$ *acrescentando os elementos* $+\infty$ *e* $-\infty$*. Notação:* $\mathbb{R}_e = \mathbb{R} \cup \{-\infty, +\infty\}$.

Preserva-se a relação de ordem ($<$) *para o conjunto dos números reais estendidos, definindo para cada número real* x, $-\infty < x < \infty$.

Quanto as operações aritméticas é comum fazer algumas convenções, como segue na definição.

Definição A.2. *Para todo número real* x, $+\infty$ *e* $-\infty$ *em* $\mathbb{R}_e$ *define-se:*

a) $x + \infty = +\infty$;

b) $x - \infty = -\infty$;

c) Se $x > 0$*, então* $x\,.\,(+\infty) = +\infty, \quad x\,.\,(-\infty) = -\infty$;

d) Se $x < 0$*, então* $x\,.\,(+\infty) = -\infty, \quad x\,.\,(-\infty) = +\infty$;

e) $\dfrac{x}{+\infty} = \dfrac{x}{-\infty} = 0$;

f) $\infty + \infty = \infty$;

g) $-\infty - \infty = -\infty$;

h) $\infty\,.\,(\pm\infty) = \pm\infty$;

i) $-\infty \,.\, (\pm\infty) = \mp\infty$;

j) *A operação* $\infty - \infty$ *não é definida, e por convenção* $0 \,.\, (\pm\infty) = 0$.

Uma utilização de números reais estendidos é que todo conjunto tem supremo e ínfimo, como mostra a seguinte definição.

Definição A.3. *Seja X um conjunto cujos elementos pertencem ao conjunto dos números reais estendidos. Então*

*a) Se X **não é limitado superiormente**, ou seja, se para cada real y existe $x \in X$ tal que $y < x$, diz-se que o* $\sup X = +\infty$.

*b) Se X **não é limitado inferiormente**, ou seja, se para cada real y existe $x \in X$ tal que $x < y$, diz-se que o* $\inf X = -\infty$.

*c) Se X é um **conjunto vazio** definem-se o supremo e o ínfimo de X (quando este é um subconjunto do reais) como sendo:*

$$i)\ \sup \emptyset = -\infty, \qquad\qquad ii)\ \inf \emptyset = +\infty.$$

Definição A.4. *Uma função cujos valores estão no conjunto de números reais estendidos é chamada uma **função a valores reais estendidos**.*

A.2 Medida exterior

Considere o conjunto $\mathbb{R}$ dos números reais e seus subconjuntos da forma (a,b), $[a,b]$, $(a,b]$ e $[a,b)$, ou seja, intervalos da reta. Sabe-se que o comprimento de $I = (a,b)$, por exemplo, é dado por $l(I) = l(a,b) = b - a$, $b > a$, que é a medida de I. Para subconjuntos mais gerais tem-se a seguinte definição.

Definição A.5. *Dado um conjunto A de números reais considere uma coleção enumerável de intervalos abertos $\{I_n\}$ que cobre A, isto é, $A \subset \bigcup_{n=1}^{\infty} I_n$, e para esta coleção considere a soma dos comprimentos dos intervalos da coleção. Define-se a **medida exterior** de A como sendo*

$$m_e(A) = \inf \left\{ \sum_{n=1}^{\infty} l(I_n);\ (I_n)\ \textit{cobertura aberta de } A \right\}.$$

Segue da definição acima que $m_e(\emptyset) = 0$ e que se $A \subset B$ então $m_e(A) \leq m_e(B)$.

Exemplo A.1. A medida exterior de um intervalo é seu comprimento. A demonstração pode ser encontrada na referência [14].

Definição A.6. *Um conjunto E é dito ser* ***mensurável*** *se para cada conjunto A tem-se $m_e(A) = m_e(A \cap E) + m_e(A \cap E^c)$, em que E^c é o complementar de E. Neste caso m_e é a medida exterior de Lebesgue e diz-se que E é* ***Lebesgue mensurável****.*

Exemplo A.2. O conjunto $\emptyset$ e o conjunto $\mathbb{R}$ de números reais são mensuráveis.

Observação A.1. Sejam os conjuntos B e B' definidos por

$$B = (A \cap E_1) \cup (A \cap E_2) \quad \text{e} \quad B' = (A \cap E_1) \cup (A \cap E_2 \cap E_1^c).$$

Desta forma, $B = B'$. Com efeito, inicialmente note que $B \subset B'$, pois

$$\text{se } x \in B \Rightarrow \begin{cases} x \in (A \cap E_1) \Rightarrow & x \in A \text{ e } x \in E_1, \\ & \text{ou} \\ x \in (A \cap E_2) \Rightarrow & x \in A \text{ e } x \in E_2. \end{cases}$$

Se $x \in (A \cap E_1)$, então $x \in B'$. Por outro lado, se $x \in (A \cap E_2)$, e neste caso $x \in A$ e $x \in E_2$, restam duas possibilidades para x

$$\begin{cases} x \in E_1 \Rightarrow & x \in (A \cap E_1) \Rightarrow x \in B', \\ \quad\text{ou} & \\ x \in E_1^c \Rightarrow & x \in (A \cap E_2 \cap E_1^c) \Rightarrow x \in B'. \end{cases}$$

Logo, se $x \in B$ então $x \in B'$ e, portanto, $B \subset B'$. Para ver que $B' \subset B$, note que

$$\text{se } x \in B' \Rightarrow \begin{cases} x \in (A \cap E_1) \Rightarrow & x \in B, \\ & \text{ou} \\ x \in (A \cap E_2 \cap E_1^c) \Rightarrow & x \in (A \cap E_2 \cap E_1^c) \subset (A \cap E_2) \Rightarrow x \in B. \end{cases}$$

Portanto, como $B' \subset B$ e $B \subset B'$, segue que $B = B'$.

Esta observação será utilizada na demonstração do lema abaixo.

Lema A.1. *Se E_1 e E_2 são mensuráveis, então $E_1 \cup E_2$ é mensurável.*

Demonstração. Seja A um conjunto qualquer. Como E_2 é mensurável, pela definição A.6, tem-se

$$m_e(A \cap E_1^c) = m_e(A \cap E_1^c \cap E_2) + m_e(A \cap E_1^c \cap E_2^c). \tag{A.1}$$

Como $A \cap (E_1 \cup E_2) = (A \cap E_1) \cup (A \cap E_2) = (A \cap E_1) \cup (A \cap E_2 \cap E_1^c)$ segue que

$$m_e(A \cap [E_1 \cup E_2]) \leq m_e(A \cap E_1) + m_e(A \cap E_2 \cap E_1^c).$$

Assim, somando $m_e(A \cap E_1^c \cap E_2^c)$ em ambos os lados, a desigualdade anterior torna-se

$$m_e(A \cap [E_1 \cup E_2]) + m_e(A \cap E_1^c \cap E_2^c) \leq m_e(A \cap E_1) + m_e(A \cap E_2 \cap E_1^c) + m_e(A \cap E_1^c \cap E_2^c).$$

Utilizando a equação (A.1) na desigualdade acima obtem-se

$$m_e(A \cap [E_1 \cup E_2]) + m_e(A \cap E_1^c \cap E_2^c) \leq m_e(A \cap E_1) + m_e(A \cap E_1^c) = m_e(A)$$

pois, por hipótese, E_1 é mensurável.

Logo,

$$m_e(A) \geq m_e(A \cap [E_1 \cup E_2]) + m_e(A \cap E_1^c \cap E_2^c)$$

e como $(E_1 \cup E_2)^c = E_1^c \cap E_2^c$ segue, da definição A.6, que $E_1 \cup E_2$ é mensurável. □

Definição A.7. *Diz-se que um conjunto $A \subset \mathbb{R}$ tem* ***medida nula*** *em $\mathbb{R}$ e escreve-se $m(A) = 0$ se, para todo $\epsilon > 0$, existe uma coleção de intervalos $\{I_n\}$, $n \in \mathbb{N}$, tal que*

i. $A \subset \displaystyle\bigcup_{n=1}^{\infty} I_n$;

ii. $\displaystyle\sum_{n=1}^{\infty} m_e(I_n) < \epsilon$.

Caso a cobertura $\{I_n\}$ seja finita, diz-se que A possui conteúdo nulo.

Exemplo A.3. Todo conjunto enumerável tem medida nula. Com efeito, considere $A = \{a_1, a_2, ..., a_n, ...\}$, $n \in \mathbb{N}$, um conjunto enumerável. Dado $\epsilon > 0$, considere para cada $a_n \in A$ o intervalo aberto I_n de centro a_n e de comprimento $\dfrac{\epsilon}{2^{n+1}}$. Assim, $A \subset \displaystyle\bigcup_{n=1}^{\infty} I_n$ e

$$\sum_{n=1}^{\infty} m(I_n) = \sum_{n=1}^{\infty} \frac{\epsilon}{2^{n+1}} = \frac{\epsilon}{2} \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n = \frac{\epsilon}{2}\left(\frac{1}{1-\frac{1}{2}} - 1\right) = \frac{\epsilon}{2} < \epsilon.$$

Portanto, $m(A) = 0$.

Lema A.2. *Se $m_e(E) = 0$, então E é mensurável.*

Demonstração. Seja A um conjunto qualquer. Note que $A \cap E \subset E$. Da definição A.5, tem-se

$$m_e(A \cap E) \leq m_e(E).$$

Como por hipótese $m_e(E) = 0$ e do fato de que medida é um número positivo, pois é a soma dos comprimentos de intervalos, segue que

$$m_e(A \cap E) = 0.$$

Observe também que $A \cap E^c \subset A$ e, pelo mesmo argumento que o anterior,

$$m_e(A \cap E^c) \leq m_e(A),$$

ou seja,

$$m_e(A) \geq m_e(A \cap E^c) + m_e(A \cap E).$$

Portanto, pela definição A.6, tem-se que E é mensurável. □

Observação A.2. Um conjunto de medida nula não precisa, necessariamente, ser enumerável. Um exemplo clássico é o *Conjunto de Cantor*[1]. O Conjunto de Cantor $\mathbf{C}$ é um subconjunto fechado do intervalo $[0, 1]$, obtido como complementar de uma reunião de intervalos abertos. $\mathbf{C}$ é obtido de $[0, 1]$ removendo o intervalo $(\frac{1}{3}, \frac{2}{3})$, depois retira-se o intervalo $(\frac{1}{9}, \frac{2}{9})$ e $(\frac{7}{9}, \frac{8}{9})$, e assim por diante. O conjunto $\mathbf{C}$ dos pontos não retirados é o conjunto de Cantor, ou seja, $\mathbf{C} = [0, 1] - \bigcup_{n=1}^{\infty} I_n$, veja a figura A.1.

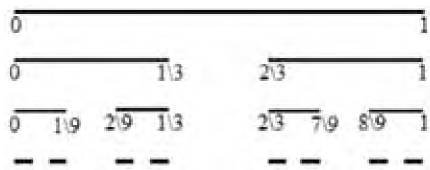

Figura A.1: *Construção do conjunto de Cantor*

Para ver que $m(\mathbf{C}) = 0$, observe que remove-se de $[0, 1]$ um intervalo de comprimento $\frac{1}{3}$ depois retira-se 2 intervalos de comprimento $\frac{1}{3^2}$ e em seguida 4 intervalos de comprimento $\frac{1}{3^3}$ e assim sucessivamente, assim na n-ésima etapa de sua construção restam apenas intervalos de comprimento $\frac{1}{3^n}$. Portanto, dado qualquer intervalo $J \subset [0, 1]$ de comprimento $c > 0$, considerando n tal que $\frac{1}{3^n} < c$, o intervalo J estará mutilado após a n-ésima etapa da construção de $\mathbf{C}$. Assim, $\mathbf{C}$ não contém intervalos, logo $m(\mathbf{C}) = 0$. Ou ainda,

$$m(\mathbf{C}) = 1 - \sum_{n=0}^{\infty} \frac{2^n}{3^{n+1}} = 1 - \frac{1}{3} \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n = 1 - \frac{1}{3} \left(\frac{1}{1 - \frac{2}{3}}\right) = 0.$$

Exemplo A.4. Um exemplo de conjunto não mensurável é um *conjunto de Vitali*[2]. A este conjunto não é possível atribuir um comprimento, nem nulo, nem finito, nem infinito. A demosntração será omitida, pois requer outros conceitos que não serão abordados neste trabalho.

Outro exemplo de conjunto não mensurável pode ser obtido na referência [14].

Observação A.3. Sendo $\mathcal{M}$ uma coleção de conjuntos mensuráveis, então o complemento de um conjunto mensurável é mensurável e a união (e a interseção) de uma coleção enumerável de conjuntos mensuráveis é mensurável. (Ver teorema em [14])

[1]O conjunto de Cantor é um subconjunto do intervalo [0,1] definido pelo matemático George Ferdinand Ludwig Philipp Cantor (1845-1918). Este conjunto é não enumerável, ver demonstração em [10], página 57.

[2]Giuseppe Vitali (1875-1932). Um conjunto de Vitali é um subconjunto dos números reais cuja existência é consequência do axioma da escolha.

Como nem todos os conjuntos são mensuráveis, é muito importante saber que os conjuntos que surgem naturalmente em certas construções são mensuráveis. Por exemplo, os listados na proposição abaixo.

Proposição A.1. *Seja f uma função a valores reais estendidos cujo domínio é mensurável. Então as seguintes afirmações são equivalentes:*

i. Para cada número real α, o conjunto $\{x;\, f(x) > \alpha\}$ é mensurável.

ii. Para cada número real α, o conjunto $\{x;\, f(x) \geq \alpha\}$ é mensurável.

iii. Para cada número real α, o conjunto $\{x;\, f(x) < \alpha\}$ é mensurável.

iv. Para cada número real α, o conjunto $\{x;\, f(x) \leq \alpha\}$ é mensurável.

Estas afirmações implicam

v. Para cada número real α o conjunto $\{x;\, f(x) = \alpha\}$ é mensurável.

Demonstração. $(i) \Rightarrow (iv)$ Seja D o domínio mensurável da f. Como $\{x;\, f(x) > \alpha\}$ é o complementar de $\{x;\, f(x) \leq \alpha\}$, segue que

$$\{x;\, f(x) \leq \alpha\} = D - \{x;\, f(x) > \alpha\}.$$

Por hipótese D e $\{x;\, f(x) > \alpha\}$ são mensuráveis, e como a diferença de dois conjuntos mensuráveis é mensurável, segue disto que $\{x;\, f(x) \leq \alpha\}$ é mensurável. Portanto, $(i) \Rightarrow (iv)$. Analogamemente, $(iv) \Rightarrow (i)$, $(ii) \Rightarrow (iii)$ e $(iii) \Rightarrow (ii)$.

Agora, como

$$\{x;\, f(x) \geq \alpha\} = \bigcap_{n=1}^{\infty} \left\{x;\, f(x) > \alpha - \frac{1}{n}\right\},$$

e a interseção de uma sequência de conjuntos mensuráveis é mensurável (observação A.3), segue que $(i) \Rightarrow (ii)$. Analogamente, $(ii) \Rightarrow (i)$, pois

$$\{x;\, f(x) > \alpha\} = \bigcup_{n=1}^{\infty} \left\{x;\, f(x) \geq \alpha + \frac{1}{n}\right\}$$

e a união de uma sequência de conjuntos mensuráveis é mensurável. Isto mostra que as quatro primeiras propriedades são equivalentes. Se α é um número real,

$$\{x;\, f(x) = \alpha\} = \{x;\, f(x) \geq \alpha\} \cap \{x;\, f(x) \leq \alpha\}$$

e tem-se que (ii) e (iv) implicam (v) para α real. Como

$$\{x;\, f(x) = \infty\} = \{x;\, f(x) \geq n\},$$

$(ii) \Rightarrow (v)$ para $\alpha = \infty$. Da mesma forma, $(iv) \Rightarrow (v)$ para $\alpha = -\infty$ e tem-se que (ii) e (iv) implicam (v). $\square$

Definição A.8. *Uma função a valores reais estendidos f é chamada* ***Lesbesgue mensurável*** *se seu domínio é mensurável e se satisfaz uma das quatro primeiras afirmações da proposição A.1.*

A proposição seguinte diz que certas operações realizadas com funções mensuráveis tornam-se funções mensuráveis.

Proposição A.2. *Sejam c uma constante e f e g duas funções reais mensuráveis definidas na interseção dos domínios de f e g. Então as funções $f+c$, cf, $f+g$, $g-f$ e fg também são mensuráveis.*

Demonstração. $f+g$ é Lebesgue mensurável. Com efeito, como $D(f+g) = D(f)$, que é mensurável por hipótese, basta provar a condição (iii) da proposição A.1. Note que

$$\{x;\, f(x) + c < \alpha\} = \{x;\, f(x) < \alpha - c\}$$

o que implica $f + c$ ser mensurável. Analogamente, cf é mensurável em $D(f)$.

Dado $x \in D(f)$ se $f(x) + g(x) < \alpha$ então $f(x) < \alpha - g(x)$, logo existe $r_x \in \mathbb{Q} \cap (f(x), \alpha - g(x))$. Como

$$\{x;\, f(x) + g(x) < \alpha\} = \bigcup_{r_x} (\{x;\, f(x) < r\} \cap \{x;\, g(x) < \alpha - r\})$$

e cada conjunto $\{x;\, f(x) < r\}$ e $\{x;\, g(x) < \alpha - r\}$ é enumerável, então esta união enumerável é mensurável.

Como $-g = (-1)g$ e g é uma função mensurável por hipótese, então $f+(-g) = f-g$ é mensurável.

A função f^2 é mensurável. Com efeito, pela proposição A.1 é preciso mostrar que $\{x;\, f^2(x) > \alpha\}$ é mensurável.

Se $\alpha < 0$, note que

$$\{x;\, f^2(x) > \alpha\} = D(f),$$

logo mensurável.

Se $\alpha > 0$, observe que

$$\{x;\, f^2(x) > \alpha\} = \{x;\, f(x) > \sqrt{\alpha}\} \cup \{x;\, f(x) < -\sqrt{\alpha}\},$$

e como cada conjunto $\{x;\, f(x) > \sqrt{\alpha}\}$ e $\{x;\, f(x) < -\sqrt{\alpha}\}$ é mensurável segue que $\{x;\, f^2(x) > \alpha\}$ é mensurável. Finalmente,

$$fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2]$$

é mensurável, pois é soma de funções mensuráveis. $\square$

Definição A.9. *Se uma determinada propriedade é válida para um conjunto, exceto para um subconjunto deste de medida nula, diz-se que esta propriedade se mantém em* ***quase toda parte*** *e usa-se a abreviação q.t.p..*

Proposição A.3. *Se f é uma função mensurável e $f = g$ q.t.p., então g é mensurável.*

Demonstração. Seja $E = \{x;\, f(x) \neq g(x)\}$ onde, por hipótese, $m(E) = 0$. Note que o conjunto $\{x;\, f(x) > \alpha\}$ é mensurável, pelo item (i) da proposição A.1. Observe também que os conjuntos $\{x \in E;\, g(x) > \alpha\}$ e $\{x \in E;\, g(x) \leq \alpha\}$ são mensuráveis, pois são subconjuntos de E e $m(E) = 0$. Como,

$$\{x;\, g(x) > \alpha\} = (\{x;\, f(x) > \alpha\} \cup \{x \in E;\, g(x) > \alpha\}) - \{x \in E;\, g(x) \leq \alpha\},$$

segue que $\{x;\, g(x) > \alpha\}$ é mensurável para todo α e, portanto, g é mensurável. $\square$

A.3 Integral de Lebesgue de uma função

Nesta seção será introduzida a integral de Lebesgue de uma função f mensurável. Para isto é necessário definir primeiramente a integral de funções simples, como abaixo.

Definição A.10. *Seja $E \subset \mathbb{R}$ um conjunto qualquer. Define-se a* ***função característica*** *$\mathcal{X}_E$ do conjunto E como*

$$\mathcal{X}_E(x) = \begin{cases} 1, & se\ x \in E, \\ 0, & se\ x \notin E,\ \forall\, x \in \mathbb{R}. \end{cases}$$

A função $\mathcal{X}_E$ é mensurável se, e somente se, E é mensurável.

Definição A.11. *A combinação linear*

$$\varphi(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{E_i}(x)$$

é chamada ***função simples*** *se os conjuntos E_i são mensuráveis.*

Observação A.4. A representação de φ não é única. Porém, note que φ é uma função simples se, e somente se, ela é mensurável e assume apenas um número finito de valores. Se φ é uma função simples e $\{a_1, a_2, ..., a_n\}$ o conjunto de valores não nulos de φ, então

$$\varphi = \sum_{i=1}^{n} a_i \mathcal{X}_{A_i},$$

onde $A_i = \{x;\, \varphi(x) = a_i\}$. Esta representação para φ é chamada a **representação canônica** e caracteriza-se pelo fato de que os A_i são disjuntos e os a_i são distintos e não nulos.

Se a função φ é nula fora de um conjunto de medida finita, define-se a integral de φ por

$$\int \varphi(x)\,dx = \sum_{i=1}^{n} a_i m(A_i),$$

onde φ tem a representação canônica. Se E é qualquer conjunto mensurável, define-se

$$\int_E \varphi(x)\,dx = \int \varphi(x)\mathcal{X}_E(x)\,dx.$$

Lema A.3. *Seja* $\varphi = \sum_{i=1}^{n} a_i \mathcal{X}_{E_i}$, *com* $E_i \cap E_j = \emptyset$ *para* $i \neq j$. *Supõe que cada* E_i *é um conjunto mensurável de medida finita. Então*

$$\int \varphi = \sum_{i=1}^{n} a_i m(E_i).$$

Demonstração. O conjunto $A_a = \{x;\, \varphi(x) = a\} = \bigcup_{a_i = a} E_i$. Portanto,

$$am(A_a) = \sum_{a_i = a} a_i m(E_i)$$

pela aditividade de m, e assim

$$\int \varphi(x)\,dx = \sum am(A_a) = \sum_{a_i = a} a_i m(E_i).$$

□

Definição A.12. *Se* f *é uma função mensurável limitada definida em um conjunto mensurável* E *com* $m(E)$ *finita, define-se a* ***integral de Lebesgue*** *de* f *sobre* E *por*

$$\int_E f(x)\,dx = \inf_{\psi \geq f} \int_E \psi(x)\,dx$$

para todas as funções simples $\psi \geq f$.

Observação A.5. Pode-se também definir a *integral de Lebesgue* de f sobre E como sendo

$$\int_E f(x)\,dx = \sup_{\varphi \leq f} \int_E \varphi(x)\,dx$$

para todas as funções simples $\varphi \leq f$, pois

$$\inf_{f \leq \psi} \int_E \psi(x)\,dx = \sup_{f \geq \varphi} \int_E \varphi(x)\,dx$$

para todas as funções simples φ e ψ.

Proposição A.4. *Se f e g são funções mensuráveis limitadas definidas no conjunto E de medida finita, então:*

i. $\int_E (\alpha f + \beta g) = \alpha \int_E f + \beta \int_E g.$

ii. Se $f = g$ *q.t.p., então* $\int_E f = \int_E g.$

iii. Se $f \leq g$ *q.t.p., então* $\int_E f \leq \int_E g$*. Consequentemente,* $\left| \int_E f \right| \leq \int_E |f|.$

iv. Se $A \leq f(x) \leq B$*, então* $Am(E) \leq \int_E f \leq Bm(E).$

v. Se A e B são conjuntos mensuráveis de medida finita, então $\int_{A \cup B} f = \int_A f + \int_B f.$

O teorema a seguir é um resultado importante, pois permite, sob algumas hipóteses, permutar o limite com a integral. Além disso, foi utilizado para mostrar que o espaço $\mathcal{L}^p$ com a norma definida pela integral de Lebesgue é completo (este conceito foi apresentado no capítulo 4).

Teorema A.1. ***(Convergência Monótona)*** *Seja (f_n) uma sequência crescente de funções mensuráveis não-negativas e seja $f = \lim_{n\to\infty} f_n$. Então*

$$\int f = \lim_{n\to\infty} \int f_n.$$